ZACHODNIOPOMORSKI UNIWERSYTET TECHNOLOGICZNY W SZCZECINIE

WYDZIAŁ INFORMATYKI

mgr Kamil Bortko

Sekwencyjna metoda nieinwazyjnej redukcji habituacji w systemach interaktywnych

Rozprawa doktorska

Promotor: dr hab. inż. Jarosław Jankowski, prof. ZUT

Szczecin 2023

Na podstawie art. 187 ust. 3 Ustawy z dnia 20 lipca 2018 roku - Prawo o szkolnictwie wyższym i nauce (Dz. U. 2018 poz. 1668 z późn. zm.) przedkładam rozprawę doktorską w formie zbioru powiązanych tematycznie artykułów opublikowanych w czasopismach naukowych i recenzowanych materiałach konferencyjnych, które stanowią oryginalne rozwiązanie problemu naukowego. Tytuł prezentowanej rozprawy to: „***Sekwencyjna metoda nieinwazyjnej redukcji habituacji w systemach interaktywnych***”. W skład rozprawy wchodzi cykl 12 publikacji naukowych z lat 2018-2023.

W dalszej części znajduje się syntetyczny opis uzyskanych wyników w postaci streszczenia rozprawy doktorskiej, a w szczególności omówienie:

— problemu badawczego,

— głównego celu rozprawy,

— cyklu publikacji,

— podstawowych pojęć,

— opisu badań,

— dorobku akademickiego kandydata do stopnia doktora.

Następnie zostały zamieszczone pełne teksty opublikowanych artykułów naukowych wchodzących w skład rozprawy, jako załączniki numerowane od A1 do A12:

A1. **Bortko, K.**, Bartków, P., Pazura, P., Jankowski, J. (2019). Increasing User Engagement and Virtual Goods Life Span Through Products Diversity and Intensity of Content Updates. In Intelligent Information and Database Systems: 11th Asian Conference, ACIIDS 2019, Yogyakarta, Indonesia, April 8–11, 2019, Proceedings, Part II 11 (pp. 519-530). Springer International Publishing.

Liczba punktów ministerialnych: 20
Udział w artykule: 30%
Wkład: Opracowanie koncepcji i założeń, przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.

A2. **Bortko, K.**, Pazura, P., Hamari, J., Bartków, P., Jankowski, J. (2019). From the Hands of an Early Adopter's Avatar to Virtual Junkyards: Analysis of Virtual Goods' Lifetime Survival. Applied Sciences, 9(7), 1268.

Liczba punktów ministerialnych: 70
Liczba cytowań: 2
Udział w artykule: 30%
Wkład: Opracowanie koncepcji i założeń, przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.

A3. **Bortko, K.**, Jankowski, J., Pazura, P. (2020). From perceptual to algorithmic evaluation of recommending interfaces survival in visual space. Procedia Computer Science, 176, 2736-2745.

Liczba punktów ministerialnych: 70
Liczba cytowań: 1
Udział w artykule: 50%

Wkład: Przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.

A4. **Bortko, K.**, Jankowski, J., Bartków, P., Pazura, P., Śmiałkowska, B. (2020). Attracting user attention to visual elements within website with the use of Fitts's law and flickering effect. Procedia Computer Science, 176, 2756-2763.

Liczba punktów ministerialnych: 70
Liczba cytowań: 6
Udział w artykule: 50%
Wkład: Opracowanie koncepcji i założeń, przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.

A5. **Bortko, K.**, Bartków, P., Jankowski, J., Kuras, D., Sulikowski, P. (2019). Multi-criteria evaluation of recommending interfaces towards habituation reduction and limited negative impact on user experience. Procedia Computer Science, 159, 2240-2248.

Liczba punktów ministerialnych: 70
Liczba cytowań: 22
Udział w artykule: 30%
Wkład: Przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.

A6. **Bortko, K.**, Fornalczyk, K., Jankowski, J. (2021). Towards Effective Peripherial Chatbot Communication with Adjustable Intensity of Content Changes. 29th International Conference on Information Systems Development (ISBN: 9781908358981, Strony od - do: [1-6]).

Liczba punktów ministerialnych: 140
Udział w artykule: 50%
Wkład: Opracowanie koncepcji i założeń, przeprowadzenie badan i opracowanie wyników, wizualizacja, opracowanie tekstu.

A7. Fornalczyk, K., **Bortko, K.**, Jankowski, J. (2021). Improving User Attention to Chatbots through a Controlled Intensity of Changes within the Interface. Procedia Computer Science, 192, 5112-5121.

Liczba punktów ministerialnych: 70
Udział w artykule: 40%
Wkład: Przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.

A8. **Bortko, K.**, Fornalczyk K., Jankowski, J., Sulikowski P., Dziedziak K. (2023). Impact of changes in chatbot's facial expressions on user attention and reaction time. Plos one.

Publikacja zaakceptowana, w druku
Liczba punktów ministerialnych: 100
Udział w artykule: 50%
Wkład: Opracowanie koncepcji i założeń, przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.

A9. Bartków, P., **Bortko, K.**, Jankowski, J., Pazura, P. (2023). Modeling the impact of the habituation effect on information spreading processes with repeated contacts under an SI model. Plos one, 18(4), e0280266.

Liczba punktów ministerialnych: 100
Udział w artykule: 20%
Wkład: Analiza statystyczna i opracowanie wyników, planowanie eksperymentu, dobór modeli habituacji obliczeniowej.

A10. Jankowski, J., Bartków, P., Pazura, P., **Bortko, K.** (2020). Evaluation of the costs of delayed campaigns for limiting the spread of negative content, panic and rumours in complex networks. In Computational Science–ICCS 2020: 20th International Conference, Amsterdam, The Netherlands, June 3–5, 2020, Proceedings, Part IV 20 (pp. 291-304). Springer International Publishing.

Liczba punktów ministerialnych: 140
Liczba cytowań: 3
Udział w artykule: 25%
Wkład: Opracowanie i analiza wyników.

A11. Fornalczyk, K., **Bortko, K.**, Disterheft A., Jankowski, J. (2023). Modeling the Impact of Video Dynamics on User Engagement and Eye Tracking Patterns. Procedia Computer Science.

Publikacja zaakceptowana, w druku
Liczba punktów ministerialnych: 70
Udział w artykule: 35%
Wkład: Opracowanie koncepcji i założeń, przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.

A12. **Bortko, K.**, Bartków P., Jankowski, J. (2023). Modeling the Impact of Habituation and Breaks in Exploitation Process on Multi-Armed Bandits Performance. Procedia Computer Science.

Publikacja zaakceptowana, w druku
Liczba punktów ministerialnych: 70
Udział w artykule: 70%
Wkład: Opracowanie koncepcji i założeń, przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.

Spis treści

1. Problem badawczy

Rozwój technologii interaktywnych powiązany jest ze wzrostem złożoności procesów komunikacji człowieka z systemami komputerowymi. Interakcja jest często odpowiedzią na sygnały i bodźce generowane przez system informatyczny w formie głosowej, wizualnej czy multimodalnej [1]. Interfejsy użytkownika oparte na głosie, systemy automatycznego rozpoznawania mowy i inteligentni asystenci stwarzają nowe możliwości w zakresie interakcji człowiek-komputer w takich obszarach jak: komunikacja i asystowanie w samochodach [2], automatyzacja domowa [3], obsługa klienta i chatboty [4], zastosowania medyczne [5] [6] [7], a także sektor edukacyjny [8].

Twórcy systemów dysponują szeroką paletą możliwości i technologii, które mogą być wykorzystywane na poziomie interfejsu użytkownika dla różnych celów. Przykładowo komunikacja może być rozpatrywana na poziomie interfejsu użytkownika i dostarczane wiadomości zapewniają możliwość efektywnego działania systemu użytkowego, dostępu do wymaganych funkcji [9]. Mogą to być również komunikaty ostrzegawcze [10], elementy przekazu marketingowego, treści prezentowane w systemach handlu elektronicznego, systemach rozrywkowych [11], grach czy przekazie wideo w mediach strumieniowych [12].

Dynamiczne treści są nasycone bodźcami przyciągającymi uwagę i wpływającymi na doświadczenie użytkownika (ang. user experience) w sposób pozytywny lub negatywny. Nasycenie bodźcami może prowadzić do efektu przeciążenia informacyjnego i negatywnie wpływać na koncentrację, pamięć i ogólną jakość interakcji użytkownika z systemem i kontakt wieloma elementami marketingowymi dziennie [13] [14].

W naturalny sposób działają tu mechanizmy habituacji, która jest definiowana jako "zmniejszenie reakcji na powtarzający się bodziec w wyniku eksponowania się na niego"[15]. Habituacja jest naturalnym procesem poznawczym, w którym użytkownik staje się mniej reaktywny na dany bodziec w miarę jego powtarzalności. Jest to podstawowa forma uczenia, która pozwala organizmowi dostosować się do rutynowych sytuacji [16].

Warto zaznaczyć, że poza wymiarem psychologicznym czy marketingowym habituacja rozpatrywana jest również z punktu widzenia obliczeniowego [17] [18], modelowania, znajduje zastosowanie w uczeniu maszynowym [19]. Powstały modele komputerowe, które umozliwiaja symulacyjne odzwierciedlenie procesu. Obszar ten daje mozliwość wdrażania nowych rozwiązań w obszarach informatyki pozwiązanych z klasyfikacją ACM Digital Library Computing Classification System w obszarach *Human-centered computing · Human computer interaction (HCI) · HCI design and evaluation methods · Laboratory experiments.*

W kontekście interfejsów użytkownika, habituacja odnosi się do zmniejszenia uwagi i reaktywności użytkownika wobec powtarzających się elementów interfejsu. Kiedy dany bodziec lub interakcja staje się powszechna lub przewidywalna, użytkownik może stać się mniej zainteresowany lub mniej skłonny do reakcji [15]. Wynika z ograniczonych możliwości poznawczych człowieka i zdolności do przetwarzania informacji. Mózg człowieka posiada ograniczone zasoby i nie jest w stanie prztwarzać wszystkich dostępnych bodźców i informacji [20].

W przypadku systemów interaktywnych, habituacja może prowadzić do spadku jakości i skuteczności interakcji, co negatywnie wpływa na doświadczenie użytkownika oraz wydajność systemu. Z punktu widzenia efektywności systemu informatycznego obserwowany jest stopniowy spadek skuteczności w wyniku powtarzających się interakcji powiązanych z generowanymi bodźcami. Nadmiarowa informacja jest eliminowana. W literaturze rozpatrywane sa zjawiska powiązane z habituacją, mające bezpośredni związek z efektywnością systemu, na przykład ślepota banerowa [21], ograniczenie efektywności komunikatów ostrzegawczych [22] i ograniczenie wpływu komunikatów motywacyjnych [23].

W praktyce projektowania stosowane są różne podejścia do redukcji habituacji, które mają na celu zapobieganie spadkowi skuteczności komunikacji. Jednak wiele z tych podejść ma charakter inwazyjny, przykładowo bazuje na zwiększonym kontraście, efektach intensywnego migotania czy przesłaniania treści redakcyjnych, co przekłada się na wzrost intensywności przekazu oraz obniżenie doświadczeń pozytywnych użytkownika.

Natomiast w wymiarze badawczym w publikacjach podejmowane są próby opracowania nieinwazyjnych metod redukcji habituacji, które nie wymagają znaczących zmian w systemach interaktywnych ani zwiększania intensywności bodźców. Przykładowo opierają się one na rozwiązaniach takich jak:

1. Komunikaty polimorficzne: Wykorzystują różnorodność form i stylów komunikacji, aby zapobiec habituacji. Komunikaty polimorficzne są dynamicznie generowane i prezentowane w różnych wariantach, co utrzymuje zainteresowanie użytkownika [24].

2. Personalizacja adaptacyjna: Dostosowanie interakcji do indywidualnych preferencji i potrzeb użytkownika. Wykorzystuje się techniki uczenia maszynowego i analizy danych, aby system mógł się dynamicznie dostosowywać do zachowań i potrzeb użytkownika [25].

3. Zastosowanie wzorców interakcji: Wykorzystuje się różnorodność wzorców i sekwencji interakcji, aby uniknąć powtarzalności i monotonii. Systemy interaktywne mogą być zaprojektowane tak, aby oferowały różne ścieżki i opcje interakcji dla użytkownika [26].

Warto zauważyć, że komunikacja i interakcja jest realizowana często w sposób sekwencyjny. Sekwencyjność w komunikacji odnosi się do posiadania logicznej kolejności i ciągłości w przekazywaniu informacji. Oznacza to, że wiadomości są przekazywane w sposób zgodny z określonym porządkiem lub układem. Problemy te występują w obszarach powiązanych z m.in. interaktywnymi systemami rozrywkowymi, mediami społecznościowymi, systemami informacyjnymi, mediami stumieniowymi i systemami dialogowymi. Problemy powiązane z komunikacją sekwencyjną mogą prowadzić do dezorientacji odbiorcy, utraty istotnych informacji, błędnego zrozumienia lub nieprawidłowego interpretowania przekazywanych wiadomości.

Dotychczasowe rozwiązania w ograniczonym stopniu uwzględniają sekwencyjność wyświetlanych komunikatów oraz ekspozycji obiektów. Istnieje wiele problemów, które mogą wpływać na efektywność komunikacji realizowanej sekwencyjnie, takie jak:

1. Przerwy lub zakłócenia w przekazywaniu informacji, na przykład zbyt długi czas między wiadomościami lub utrata części komunikatu [27].

2. Błędy w przekazywaniu informacji, na przykład pomijanie lub zamiana kolejności elementów [28].

3. Brak jasności w przekazywaniu informacji, co prowadzi do niejasności w zrozumieniu kolejności lub ciągłości przekazu [29].

Badania w ramach rozprawy doktorskiej miały na celu opracowanie sekwencyjnej metody nieinwazyjnej redukcji habituacji. Komunikaty powinny być dostarczane do odbiorcy w sposób efektywny, powinny być zauważane natomiast nie powinny odbiorcy irytować. Powinny sygnalizować lub nienachalnie zwracać naszą uwagę. Użytkownik szuka przede wszystkim treści informacyjnych i tych związanych z bieżącymi potrzebami, a podświadomie pomija treści, które są powtarzane. Eliminowane są również informacje niezwiązane z wykonywanym zadaniem, na przykład komunikaty systemowe, komunikaty związane z bezpieczeństwem, komunikaty marketingowe itd. Obszary występowania problemów to również rekomendacje w handlu elektronicznym, wizualny i tekstowy przekaz marketingowy, gry wideo i platformy rozrywkowe, interfejsy użytkownika, interfejsy czasu rzeczywistego, produkty wirtualne, komunikaty w systemach antywirusowych, komunikaty związane z aktualizacjami i inne.

W trakcie badań reprezentowanych w poszczególnych publikacjach w przedstawionym cyklu pozyskano wiedzę, którą nastepnie zagregowano i wykorzystano do osiągnięcia głównego celu rozprawy. Autor przeanalizował istniejące prace naukowe związane z redukcją habituacji w systemach interaktywnych, aby zidentyfikować kluczowe podejścia i techniki. Następnie wiedza z tych publikacji została syntezowana i wykorzystana do opracowania sekwencyjnej metody nieinwazyjnej redukcji habituacji w systemach interaktywnych, którą uogólniono w publikacji [A12]. Zaproponowano w niej system uczący, ponadto, omówiono techniki analizy danych, które zostały zastosowane w celu oceny skuteczności metody redukcji habituacji.

W realizacji planowanych badań napotkano problemy i trudności wynikających z trwającej przez blisko 2 lata pandemii COVID-19. Dodatkowe badania realizowane w okresie kiedy nie było możliwe prowadzenie badań z udziałem uzytkowników w laboratorium oparto na środowiskach eksperymentalnych wykorzystujących podejście obliczeniowe i symulacje agentowe. Opracowano wirtualne środowisko zorientowane na sieci społeczne, które umożliwiło prowadzenie eksperymentów z redukcją habituacji i ocenę skuteczności sekwencyjnej metody nieinwazyjnej redukcji. Wykorzystanie podejścia agentowego pozwoliło na przeprowadzenie symulacji bez udziału użytkowników. Dzięki temu możliwe było zgromadzenie danych i wyciągnięcie wniosków dotyczących sekwencyjnej metody nieinwazyjnej redukcji habituacji w systemach interaktywnych.

Wyniki rozprawy mają istotne implikacje dla projektantów i twórców systemów interaktywnych, którzy starają się zapewnić pozytywne doświadczenia użytkownika, a jednocześnie osiągać zakładane cele. Stosowanie sekwencyjnej metody nieinwazyjnej redukcji habituacji może przyczynić się do zwiększenia skuteczności i zadowolenia użytkowników systemów interaktywnych.

2. Główny cel rozprawy

Celem rozprawy jest opracowanie metody sekwencyjnej, która będzie stanowiła podstawę systemu uczącego się, zorientowanego na modyfikację interfejsu systemu interaktywnego w czasie rzeczywistym w celu redukcji efektu habituacji.

Teza rozprawy:

Analiza charakterystyki wizualnej obiektów wyświetlanych sekwencyjnie w procesie korzystania z systemu interaktywnego oraz reakcji odbiorcy umożliwia wyznaczanie poziomu oddziaływania wizualnego dla stanów przyszłych, które zapewnią pozyskanie uwagi użytkownika.

3. Cykl publikacji wchodzących w skład rozprawy

Jako osiągnięcie naukowe w dyscyplinie Informatyka techniczna i telekomunikacja wskazuję cykl dwunastu powiązanych tematycznie publikacji pt. Sekwencyjna metoda nieinwazyjnej redukcji habituacji w systemach interaktywnych. Powiązania pomiędzy publikacjami zostały przedstawione na rysunku 3.1.

W skład cyklu publikacji wchodzą następujące prace:[1]

A1. **Bortko, K.**, Bartków, P., Pazura, P., Jankowski, J. (2019). Increasing User Engagement and Virtual Goods Life Span Through Products Diversity and Intensity of Content Updates. In Intelligent Information and Database Systems: 11th Asian Conference, ACIIDS 2019, Yogyakarta, Indonesia, April 8–11, 2019, Proceedings, Part II 11 (pp. 519-530). Springer International Publishing.

Liczba punktów ministerialnych: 20
Udział w artykule: 30%
Wkład: Opracowanie koncepcji i założeń, przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.

A2. **Bortko, K.**, Pazura, P., Hamari, J., Bartków, P., Jankowski, J. (2019). From the Hands of an Early Adopter's Avatar to Virtual Junkyards: Analysis of Virtual Goods' Lifetime Survival. Applied Sciences, 9(7), 1268.

Liczba punktów ministerialnych: 70
Liczba cytowań: 2
Udział w artykule: 30%
Wkład: Opracowanie koncepcji i założeń, przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.

A3. **Bortko, K.**, Jankowski, J., Pazura, P. (2020). From perceptual to algorithmic evaluation of recommending interfaces survival in visual space. Procedia Computer Science, 176, 2736-2745.

Liczba punktów ministerialnych: 70
Liczba cytowań: 1
Udział w artykule: 50%
Wkład: Przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.

[1] Informacje o cytowaniach pochodzą z Google Scholar, stan na dzień 28.06.2023 r.

A4. **Bortko, K.**, Jankowski, J., Bartków, P., Pazura, P., Śmiałkowska, B. (2020). Attracting user attention to visual elements within website with the use of Fitts's law and flickering effect. Procedia Computer Science, 176, 2756-2763.

Liczba punktów ministerialnych: 70
Liczba cytowań: 6
Udział w artykule: 50%
Wkład: Opracowanie koncepcji i założeń, przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.

A5. **Bortko, K.**, Bartków, P., Jankowski, J., Kuras, D., Sulikowski, P. (2019). Multi-criteria evaluation of recommending interfaces towards habituation reduction and limited negative impact on user experience. Procedia Computer Science, 159, 2240-2248.

Liczba punktów ministerialnych: 70
Liczba cytowań: 22
Udział w artykule: 30%
Wkład: Przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.

A6. **Bortko, K.**, Fornalczyk, K., Jankowski, J. (2021). Towards Effective Peripherial Chatbot Communication with Adjustable Intensity of Content Changes. 29th International Conference on Information Systems Development (ISBN: 9781908358981, Strony od - do: [1-6]).

Liczba punktów ministerialnych: 140
Udział w artykule: 50%
Wkład: Opracowanie koncepcji i założeń, przeprowadzenie badan i opracowanie wyników, wizualizacja, opracowanie tekstu.

A7. Fornalczyk, K., **Bortko, K.**, Jankowski, J. (2021). Improving User Attention to Chatbots through a Controlled Intensity of Changes within the Interface. Procedia Computer Science, 192, 5112-5121.

Liczba punktów ministerialnych: 70
Udział w artykule: 40%
Wkład: Przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.

A8. **Bortko, K.**, Fornalczyk K., Jankowski, J., Sulikowski P., Dziedziak K. (2023). Impact of changes in chatbot's facial expressions on user attention and reaction time. Plos one.

Publikacja zaakceptowana, w druku
Liczba punktów ministerialnych: 100
Udział w artykule: 50%
Wkład: Opracowanie koncepcji i założeń, przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.

A9. Bartków, P., **Bortko, K.**, Jankowski, J., Pazura, P. (2023). Modeling the impact of the habituation effect on information spreading processes with repeated contacts under an SI model. Plos one, 18(4), e0280266.

Liczba punktów ministerialnych: 100
Udział w artykule: 20%
Wkład: Analiza statystyczna i opracowanie wyników, planowanie eksperymentu, dobór modeli habituacji obliczeniowej.

A10. Jankowski, J., Bartków, P., Pazura, P., **Bortko, K.** (2020). Evaluation of the costs of delayed campaigns for limiting the spread of negative content, panic and rumours in complex networks. In Computational Science–ICCS 2020: 20th International Conference, Amsterdam, The Netherlands, June 3–5, 2020, Proceedings, Part IV 20 (pp. 291-304). Springer International Publishing.

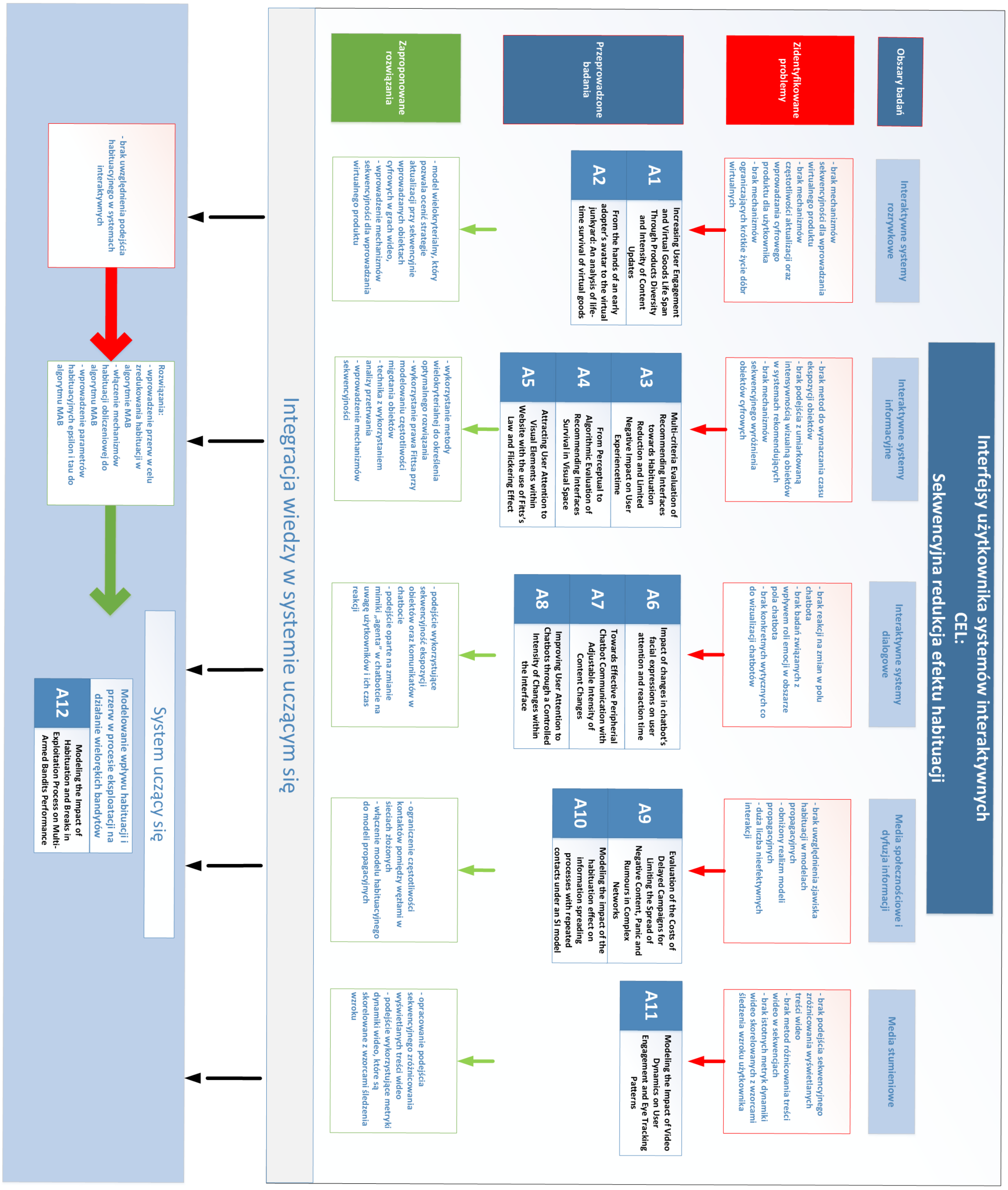
Liczba punktów ministerialnych: 140
Liczba cytowań: 3
Udział w artykule: 25%
Wkład: Opracowanie i analiza wyników.

A11. Fornalczyk, K., **Bortko, K.**, Disterheft A., Jankowski, J. (2023). Modeling the Impact of Video Dynamics on User Engagement and Eye Tracking Patterns. Procedia Computer Science.

Publikacja zaakceptowana, w druku
Liczba punktów ministerialnych: 70
Udział w artykule: 35%
Wkład: Opracowanie koncepcji i założeń, przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.

A12. **Bortko, K.**, Bartków P., Jankowski, J. (2023). Modeling the Impact of Habituation and Breaks in Exploitation Process on Multi-Armed Bandits Performance. Procedia Computer Science.

Publikacja zaakceptowana, w druku
Liczba punktów ministerialnych: 70
Udział w artykule: 70%
Wkład: Opracowanie koncepcji i założeń, przeprowadzenie badań i opracowanie wyników, wizualizacja, opracowanie tekstu.



Rysunek 3.1. Wizualizacja powiązań pomiędzy poszczególnymi publikacjami A1–A12. W wersji drukowanej wizualizacja została dołączona w formacie A3.

4. Podstawowe pojęcia

W nawiązaniu do tematu rozprawy dotkorskiej omówiono podstawowe pojęcia takie jak: habituacja, sekwencyjność, metoda nieinwazyjna, algorytm MAB, metoda, systemy interaktywne, bodźce w systemach interaktywnych oraz interfejs. Pozwolą one szerzej i dokładniej zrozumieć cel oraz tezę rozprawy doktorskiej.

Pierwszym z pojęć jest **habituacja**. Jest ona naturalną reakcją człowieka na powtarzające się bodźce, które nie dostarczają nowych informacji lub nie wymagają intensywnej uwagi. Może wystąpić w różnych kontekstach interakcyjnych, takich jak interfejsy użytkownika, aplikacje mobilne, strony internetowe czy systemy głosowe [16].

Habituacja w odniesieniu do systemów interaktywnych odnosi się do procesu, w którym użytkownicy stają się bardziej przyzwyczajeni do określonych cech interakcji lub funkcji w systemie [30]. Jest to rodzaj adaptacji, w którym użytkownik stopniowo przestaje reagować na pewne bodźce lub elementy interfejsu, ponieważ stają się one dla niego znane, rutynowe lub mało istotne [31]. W przypadku systemów interaktywnych, habituacja może mieć zarówno pozytywne, jak i negatywne konsekwencje. z jednej strony, użytkownicy mogą się przyzwyczaić do wygodnych i intuicyjnych funkcji, co może poprawić ich efektywność i wydajność w korzystaniu z systemu. Na przykład, jeśli użytkownik regularnie korzysta z aplikacji mobilnej o prostym i spójnym interfejsie, może szybko opanować nawigację i przeprowadzać operacje bez dużego wysiłku po pewnym czasie [32]. Z drugiej strony, habituacja może prowadzić do niezauważenia nowych lub istotnych funkcji, które mogą być wprowadzane w systemie. Użytkownicy mogą przestać zauważać ważne powiadomienia lub opcje, ponieważ stają się one dla nich częścią tła. Jest to wyzwanie dla projektantów systemów interaktywnych, którzy muszą uwzględnić równowagę między przyzwyczajeniami użytkowników a wprowadzaniem innowacji i ulepszeń [31].

Aby zidentyfikować habituacją w systemach interaktywnych, ważne jest monitorowanie i analiza zachowań użytkowników. Projektanci mogą używać danych analitycznych, badań użytkowników i informacji zwrotnych, aby zrozumieć, jak użytkownicy korzystają z systemu i jakie działania stają się rutynowe. Na podstawie tych informacji można wprowadzać zmiany w interfejsie użytkownika, aby zwiększyć uwagę użytkowników na istotne elementy i uniknąć badania habituacji [16].

Wnioski płynące z **habituacji w systemach interaktywnych** mogą również wpływać na strategię rozwoju i wprowadzania nowych funkcji. Projektanci mogą zaplanować regularne aktualizacje systemu, które wprowadzają zmiany i nowości, aby pobudzić użytkowników do ponownej uwagi i uniknąć stagnacji czy zaniedbania innowacji.

W rozprawie zwrócono uwagę na sekwencyjność komunikacji. W informatyce, termin **"sekwencyjność"** odnosi się do sposobu wykonywania instrukcji lub operacji w określonej kolejności. Oznacza to, że program lub proces przetwarza dane lub wykonuje operacje krok po kroku, jeden za drugim, bez skoków ani równoczesnych działań [33]. Przyjęta sekwencyjność zakłada działanie etapowe i stopniowe.

Wyrażenie **"metoda jest nieinwazyjna"** odnosi się do techniki, procedury lub badania, które nie utrudniają procesów. W kontekście dziedziny informatyki, termin nieinwazyjny odnosi się do metody lub techniki, która nie wymaga ingerencji ani zmian w istniejącym systemie lub strukturze. Jest to podejście, które ma na celu minimalizację zakłóceń lub zmian w działającym środowisku [34]. W skrócie, stwierdzenie, że metoda jest **"nieinwazyjna"**, sugeruje, że jest to bezpieczna, pozbawiona negatywnego oddziaływania na użytkownika.

Algorytm MAB (Multi-Armed Bandit) odnosi się do problemu wyboru optymalnej decyzji w warunkach niepewności, gdzie agent ma do dyspozycji wiele "ramion"(ang. arms), z których każde generuje nagrody zgodnie z nieznanym rozkładem. W ramach eksploracji i eksploatacji, algorytm MAB podejmuje kolejne decyzje w celu maksymalizacji ogólnego zysku [35] [36]. System uczący się oparty na MAB charakteryzuje się zdolnością do poprawiania swojego zachowania na podstawie zebranych doświadczeń. W przypadku algorytmu MAB, agent ma za zadanie nauczyć się, które z ramion generują największą nagrodę. Początkowo, agent eksploruje różne ramiona, aby zdobyć informacje na temat ich potencjalnej wydajności. Na podstawie zebranych danych, algorytm stopniowo zdobywa wiedzę i adaptuje swoje działania, skupiając się na wybieraniu najlepszych ramion w celu maksymalizacji nagrody.

Algorytm MAB spełnia podstawowe cechy systemu uczącego się. Posiada zdolność do uczenia się na podstawie doświadczeń, dostosowywania swojego zachowania w oparciu o zebrane informacje oraz optymalizacji wyników w celu osiągnięcia lepszych rezultatów [35] [37].

Algorytm MAB używany w wielu systemach interaktywnych do podejmowania decyzji w oparciu o informacje zwrotne. Algorytm MAB jest wykorzystywany w sytuacjach, w których system musi wybierać spośród kilku możliwych akcji, nie mając pełnej wiedzy o ich skutkach [35]. System interaktywny może korzystać z algorytmu MAB, aby zoptymalizować wybór akcji na podstawie sukcesów i porażek wcześniejszych interakcji. Przykładem zastosowania algorytmu MAB w systemie interaktywnym może być platforma reklamowa, która musi wybierać spośród różnych reklam do wyświetlenia użytkownikom. Algorytm MAB pozwala na eksplorację różnych reklam i adaptacyjne dostosowywanie wyboru w oparciu o otrzymane informacje zwrotne [36].

Algorytm MAB może być wykorzystywany jako składnik systemu interaktywnego do podejmowania optymalnych decyzji w warunkach niepewności i interakcji z użytkownikami [37].

Zaproponowana **"metoda"** [38] jest traktowana jako sposób postępowania, systematyczne rozwiązanie problemu, zbiór reguł, procedur lub technik stosowanych w określonym celu. Metoda jest planowanym i uporządkowanym sposobem osiągnięcia określonego rezultatu lub rozwiązania danego zadania.

Systemy interaktywne to szeroka kategoria systemów, które umożliwiają interakcję pomiędzy użytkownikiem a technologią w celu realizacji określonych zadań, dostarczania informacji lub otrzymywania odpowiedzi na pytania [39][40]. Istnieje wiele różnych typów systemów interaktywnych, z których niektóre to:

1. **Interaktywne systemy informacyjne**: Są to systemy, które dostarczają użytkownikowi informacji na podstawie jego preferencji i zapytań. Przykłady to wyszukiwarki internetowe, systemy rekomendacyjne czy personalizowane portale informacyjne [41].

2. **Interaktywne systemy dialogowe**: Są to systemy, które prowadzą dialog z użytkownikiem, odpowiadając na pytania, udzielając informacji lub realizując zadania. Przykłady to asystenci głosowi, chatboty czy interaktywne systemy rezerwacyjne [42].

3. **Interaktywne systemy rozrywkowe**: Są to systemy, które zapewniają użytkownikom rozrywkę, zabawę i interakcję z wirtualnymi światami. Przykłady to gry komputerowe, wirtualna rzeczywistość (VR) czy interaktywne wystawy [43].

4. **Wideo interaktywne w reklamach**: Wideo reklamowe stają się coraz bardziej interaktywne, umożliwiając użytkownikom bezpośrednie zaangażowanie się w treść reklamy. Przykłady to reklamy online, które pozwalają użytkownikom na kliknięcie w interaktywne elementy wideo, aby uzyskać więcej informacji, dokonać zakupu lub podjąć interaktywne działania [44].

5. **Media społecznościowe**: Obejmują platformy internetowe, które umożliwiają użytkownikom tworzenie profili, nawiązywanie kontaktów, udostępnianie treści i interakcję z innymi użytkownikami. Sieci społecznościowe stanowią wirtualne przestrzenie, w których użytkownicy mogą dzielić się informacjami, komunikować się, tworzyć grupy tematyczne i uczestniczyć w różnego rodzaju interakcjach społecznych [45].

Bodźce w systemach interaktywnych, które mogą wpływać na sekwencyjną metodę nieinwazyjnej redukcji habituacji, mogą być różnorodnei zależą od kontekstu aplikacji [46] [47]. Istotna jest tutaj:

1. **Częstotliwość prezentacji bodźca**: Sekwencyjna metoda nieinwazyjnej redukcji habituacji opiera się na stopniowym zmniejszaniu intensywności bodźca w celu zmniejszenia habituacji. Częstotliwość prezentacji bodźca może być jednym z czynników wpływających na skuteczność metody. Może to obejmować interwały czasowe między kolejnymi prezentacjami bodźca.

2. **Intensywność bodźca**: W sekwencyjnej metodzie nieinwazyjnej redukcji habituacji, intensywność bodźca jest stopniowo zmniejszana. Stopień redukcji intensywności bodźca oraz tempo zmniejszania mogą wpływać na efektywność metody.

3. **Zróżnicowanie bodźców**: Wprowadzenie zróżnicowanych bodźców może pomóc w uniknięciu habituacji. Systemy interaktywne mogą oferować różnorodne bodźce, takie jak zmienione kolory, kształty, dźwięki czy tekstury, aby utrzymać zainteresowanie użytkownika i minimalizować efekt habituacji [48].

Interfejs użytkownika to pojęcie powszechnie stosowane w dziedzinie informatyki i technologii [49] punkt kontaktu lub powiązanie między dwoma lub więcej systemami, urządzeniami lub komponentami, które umożliwiają im komunikację i wymianę informacji. W kontekście oprogramowania, interfejs oznacza zbiór reguł, protokołów i narzędzi, które umożliwiają interakcję użytkownika z programem lub systemem. Może to obejmować graficzny interfejs użytkownika (GUI), w którym elementy takie jak przyciski, menu i pola tekstowe są używane do manipulacji programem za pomocą myszy i klawiatury. Innym rodzajem interfejsu jest wiersz poleceń, gdzie użytkownik wprowadza komendy tekstowe, aby wykonać określone zadania.

5. Wyniki badań

5.1. Interaktywne systemy rozrywkowe [A1] i [A2][1]

A1. "Increasing User Engagement and Virtual Goods Life Span Through Products Diversity and Intensity of Content Updates"

Sektor wirtualnych produktów jest kluczowym modelem biznesowym dla platform społecznościowych, gier i wirtualnych światów. Użytkownicy tych systemów oczekują regularnych aktualizacji treści i funkcji, aby zapobiec habituacji i utrzymać zainteresowanie. Przedstawione badanie bada wpływ różnorodności i liczby wirtualnych produktów na zaangażowanie użytkowników. Proponowany jest również model decyzyjny, który wspomaga ocenę strategii aktualizacji i dystrybucji wirtualnych produktów. Rosnące znaczenie środowisk cyfrowych wymaga nowych narzędzi i metod analizy, ponieważ zjawiska znane z rynków offline mają również zastosowanie w systemach elektronicznych. Badanie pokazuje, że strategie aktualizacji treści mają wpływ na zaangażowanie użytkowników i żywotność wirtualnych produktów. Wprowadzenie dużej ilości różnorodnych produktów w ramach jednej aktualizacji wpływa na zainteresowanie użytkowników i propagację produktów w systemie. Badane czynniki obejmują również liczbę elementów w aktualizacji i dynamikę użytkowania produktów. Proponowany model wielokryterialny pozwala ocenić strategie aktualizacji, przy uwzględnieniu kosztów implementacji i dynamiki użytkowania.

Bodziec, który został tutaj wykorzystany do nieinwazyjnej redukcji habituacji to zróżnicowanie oraz intensywność kontentu przez różnorodność sekwencyjną.

A2. "From the hands of an early adopter's avatar to the virtual junkyard: An analysis of life-time survival of virtual goods"

Badanie koncentruje się na pomiarze i przewidywaniu czasu życia wirtualnych dóbr oraz wpływie cech wczesnych użytkowników na przetrwanie tych dóbr. Wykorzystuje się uczenie maszynowe i drzewa decyzyjne do modelowania predykcyjnego. Komunikacja i aktywność społeczna są kluczowymi czynnikami propagowania wirtualnych dóbr i są pożądanymi cechami wczesnych użytkowników. Potrzebne są nowe modele analityczne i strategie, aby lepiej wykorzystać wirtualne produkty w systemach online. Habitacja klientów, czas życia i techniki ulepszania produktów są istotne w przypadku rynków elektronicznych. Wyniki badania pokazują, że cechy wczesnych użytkowników mają wpływ na zaangażowanie i przetrwanie wirtualnych dóbr w środowisku elektronicznym. Monitorowanie wzorców użytkowania i cech użytkowników umożliwia identyfikację produktów o krótkim czasie przetrwania i podejmowanie działań, takich jak zachęty i techniki zapraszania dodatkowych użytkowników, aby zmniejszyć efekt habituacji i zwiększyć czas użytkowania produktu.

[1] Sekcja powstała w oparciu o publikacje **A1** i **A2**

Przedstawiona koncepcja zakłada redukcje habituacji poprzez zróżnicowanie oraz intensywność kontentu przez różnorodność sekwencyjną.

5.2. Interaktywne systemy informacyjne [A3], [A4] i [A5] [2]

A3. "From Perceptual to Algorithmic Evaluation of Recommending Interfaces Survival in Visual Space"

Projektowanie systemów rekomendacyjnych koncentruje się na dobraniu odpowiednich produktów i usług dla klientów. Jednak samo zastosowanie skutecznych algorytmów nie zapewni sukcesu platformy e-commerce, jeśli interfejs rekomendacji nie przyciąga uwagi użytkowników. Ślepotę banerową, czyli adaptację użytkowników na bodźce wizualne, można podświadomie eliminować, a intensywność elementów graficznych w interfejsie ma na to wpływ. Badanie przedstawione w artykule analizuje przetrwanie obiektów graficznych i ich związki z intensywnością wizualną, zamiast tradycyjnych badań percepcyjnych. Wyniki wskazują, że wzrost intensywności nie zawsze prowadzi do większej uwagi użytkowników. Analiza przetrwania pozwala identyfikować różnice między projektami i użytkownikami otrzymującymi te same bodźce. Dla interfejsów o wyższej intensywności, analiza przetrwania pozwala określić proporcje użytkowników bardziej skoncentrowanych na interfejsie. Wyższa intensywność nie gwarantuje dłuższego skupienia uwagi użytkownika. Wyniki pokazują, że analiza przetrwania w kontekście długości przetrwania wpływa na odbiór informacji. Większa liczba wyróżnionych elementów nie zawsze przekłada się na lepsze wyniki. Warto więc dokładnie dobrać liczbę wyróżnionych elementów, a nie skupiać się na zwiększaniu intensywności i liczby. Zaproponowana technika analizy przetrwania pozwala ocenić skuteczność różnych wariantów interfejsów rekomendacyjnych.

Redukcja habituacji następuje poprzez zróżnicowanie oraz częstotliwość prezentacji kontentu przez różnorodność sekwencyjną.

A4. "Attracting User Attention to Visual Elements within Website with the use of Fitts's Law and Flickering Effect"

Projektowanie aplikacji wymaga skutecznego zarządzania treściami, interakcjami i przestrzenią. w przypadku systemów internetowych, przepływ interakcji często opiera się na wyróżniających się elementach nawigacyjnych i reklamowych. Skalowanie rozmiaru obiektów zgodnie z prawem Fittsa może wpływać na zachowanie użytkownika, poprzez kontrolowanie odległości od centralnej części interfejsu. w badaniu skupiono się na analizie efektów migotania obiektów o różnej częstotliwości, kierunku i odległości od centralnych elementów strony internetowej, oraz jego wpływu na uwagę użytkownika i redukcję habituacji. Celem badania było sprawdzenie, czy wyższa częstotliwość migotania może zrekompensować większą odległość obiektów i zwiększyć skupienie użytkownika, podobnie jak skalowanie rozmiarów obiektów w prawie Fittsa. Wyniki wykazały, że migotanie zwiększa uwagę na obiekty znajdujące się dalej, jednak nie ma wpływu na najbliższe obiekty. Zaobserwowano poprawę o dwadzieścia procent całkowitego czasu przyciągania uwagi obiektów odległych, jednak większą skuteczność stwierdzono dla najbliższych obiektów, niezależnie od częstotliwości migotania.

[2] Sekcja powstała w oparciu o publikacje **A3**, **A4** i **A5**

Przedstawiona koncepcja zakłada redukcje habituacji poprzez częstotliwość migotania obiektów przy wykorzystaniu prawa Fittsa.

A5. "Multi-criteria Evaluation of Recommending Interfaces towards Habituation Reduction and Limited Negative Impact on User Experience"

Rosnące natężenie treści marketingowych wizualnych może prowadzić do ograniczonej uwagi użytkowników i spadku efektywności, ze względu na zjawisko habituacji, znane jako ślepotę banerową. Aby przeciwdziałać temu problemowi, marketerzy stosują różne techniki wizualne, takie jak animacje i intensywność, jednak mogą one prowadzić do negatywnego postrzegania marki. Podobne wyzwanie dotyczy systemów rekomendacyjnych, gdzie badanie opisane w tekście analizuje wpływ natężenia wizualnego w interfejsie rekomendacyjnym na uwagę użytkownika. Badanie wykorzystuje śledzenie wzroku i analizę wielokryterialną, aby ocenić doświadczenie użytkownika oraz skupienie uwagi na elementach wizualnych i ich natężenie. Celem badania jest znalezienie rozwiązań, które są zarówno skuteczne, jak i mają ograniczony negatywny wpływ na użytkowników. Analiza pokazuje, że wysokie natężenie treści marketingowych zazwyczaj prowadzi do gorszego doświadczenia użytkownika, więc ważne jest poszukiwanie rozwiązań, które przezwyciężają habituację. Wyniki badania wskazują, że umiarkowana intensywność wizualna rekomendacji jest preferowana w porównaniu do braku wyróżnienia lub silnego wyróżnienia. z perspektywy projektowania interfejsu rekomendacyjnego i dostarczania rekomendowanych treści, umiarkowane natężenie wizualne jest bardziej korzystne, zgodnie z wnioskami badania opartymi na analizie wielokryterialnej.

Mechanizm redukcji habituacji bazuje na intensywność wizualna obiektów przy wykorzystaniu braku lub silnego ich wyróżnienia.

5.3. Interaktywne systemy dialogowe [A6], [A7] i [A8] [3]

A6. "Towards Effective Peripherial Chatbot Communication with Adjustable Intensity of Content Changes"

Projektowanie interfejsów użytkownika wymaga skutecznego zarządzania wszystkimi elementami, które na nich się znajdują. Jednym z problemów jest przekazywanie wiadomości użytkownikom na stronach internetowych, które są dynamiczne i często się zmieniają. Przykładem są chatboty, które pojawiają się w obszarach peryferyjnych i są często pomijane przez użytkowników podczas przeglądania strony. Przedstawione badanie skupia się na wpływie intensywności zmian w obszarze chatbota na reakcję użytkowników. Badanie przeprowadzono, zmieniając treści chatbota w obszarze peryferyjnym przy różnych czasach reakcji. Wyniki wykazały, że intensywność zmian w obszarze chatbota głównie wpływa na reakcje użytkowników w ciągu 400 ms. Dla dłuższych czasów reakcji, różnica między znaczącymi a niewielkimi zmianami była mniejsza. Oznacza to, że dynamika chatbota może być dostosowywana do zachowań użytkowników w celu zwiększenia świadomości i efektywności przekazywania wiadomości. Głównym celem badania było sprawdzenie wpływu intensywności zmian w chatbocie na świadomość nowej zawartości w obszarze peryferyjnym, bez zakłócania głównego zadania użytkownika. Zastosowano różne poziomy intensywności zmian w chatbocie, oparte na sekwencji par treści oraz różnicach między

[3] Sekcja powstała w oparciu o publikacje **A6**, **A7** i **A8**

nimi. Wyniki jednoznacznie pokazały, że intensywność zmian w obszarze chatbota miała istotny wpływ na reakcje użytkowników, którzy spędzali na stronie około 400 ms. Dla dłuższych czasów reakcji, od około 500 ms do końca, różnica między dużymi a małymi zmianami była mniejsza. To sugeruje, że dynamika zmian w chatbocie może być dostosowywana do indywidualnych zachowań użytkowników, aby zwiększyć ich świadomość zmian i efektywność przekazywania konkretnej wiadomości.

Mechanizm redukcji habituacji bazuje na sekwencyji intensywności wizualnej obiektów.

A7. "Improving User Attention to Chatbots through a Controlled Intensity of Changes within the Interface"

W kontekście rozwoju technologii wspierających chatboty, niewiele badań skupiło się na aspektach wizualnych interfejsu. Zmiany w obszarze chatbota mogą mieć różny charakter, takie jak zwiększanie lub zmniejszanie liczby wiadomości. Ze względu na zjawisko habituacji, czyli ślepoty na zmiany, wiadomości dostarczane przez chatboty mogą nie być zauważane przez użytkowników. Głównym celem przedstawionego badania było zbadanie wpływu zmian wewnątrz chatbota na zachowanie użytkowników oraz zdolność skierowania ich uwagi na ten obszar. Badanie to wykazało, że intensywność i rodzaj zmian wewnątrz chatbota mogą poprawić wydajność i skuteczność dostarczania wiadomości do użytkowników. Badanie skoncentrowało się na tym, jak zmiany treści były reprezentowane w obszarze chatbota. Wyniki oparte na dekompozycji wiadomości chatbota na osobne składniki wykazały, że intensywne zmiany w treści wizualnej były bardziej skuteczne niż stopniowe aktualizacje. Zarówno zwiększające się, jak i zmniejszające się zmiany w obszarze widzenia chatbota miały równie duży wpływ, podobnie jak stopniowe zmiany w obszarze widzenia. Pojawienie się nowych wiadomości w obszarze chatbota przyciągało o około 60% więcej użytkowników niż ich zniknięcie. Badanie potwierdziło również intuicyjne założenie, że im większe zmiany zachodzą w obszarze chatbota, tym większa jest zdolność chatbota do przyciągania uwagi użytkownika. Stopniowe zmiany w przestrzeni wizualnej były trudniejsze do wykrycia. z tego powodu deweloperzy chatbota powinni rozważyć wyświetlanie dużych bloków treści naraz, zamiast stopniowego wprowadzania wiadomości o mniejszym polu widzenia.

Redukcja habituacji następuje poprzez stopniowe sekwencyjne zmiany obiektów w przestrzeni wizualnej.

A8. "Impact of changes in chatbot's facial expressions on user attention and reaction time"

Komunikacja online z wykorzystaniem chatbotów wymaga skutecznych algorytmów, przetwarzania języka i atrakcyjnej reprezentacji wizualnej. Aby zwiększyć zaangażowanie użytkowników i sprawić, że komunikacja jest bardziej naturalna, kluczowe są habituacja i sekwencyjność. Chatboty muszą konkurować z innymi elementami na stronach internetowych i w aplikacjach, więc przyciągnięcie uwagi użytkownika jest wyzwaniem. Badanie dotyczyło wpływu emocjonalnych wyrazów chatbota na czas reakcji użytkownika. Odkryto, że użytkownicy bardziej reagują na pozytywne emocje, które są zauważane szybciej niż negatywne. Analiza zmian emocji wyrażanych przez chatbota, przedstawianych sekwencyjnie, wykazała, że szczęście, zaskoczenie i strach skutkowały szybszym czasem reakcji. Średni czas reakcji na wizualne zmiany w sekwencji zmniejszał się z każdą kolejną zmianą.

To badanie sugeruje, że przy tworzeniu chatbota warto przemyśleć, które emocje są odpowiednie dla konkretnych komunikatów, aby osiągnąć szybką reakcję użytkownika.

Redukcja habituacji następuje poprzez stopniowe sekwencyjne zmiany obiektów w przestrzeni wizualnej.

5.4. Media społecznościowe i dyfuzja informacji [A9] i [A10] [4]

A9. "Modeling the impact of the habituation effect on information spreading processes with repeated contacts under an SI model"

Informacje, którym jesteśmy narażeni, pochodzą z różnych źródeł, niezależnie od naszych preferencji. Ze względu na ograniczoną zdolność przetwarzania informacji, część komunikatów może być absorbowana, podczas gdy inne są ignorowane. Powtarzające się bodźce prowadzą do habituacji, co oznacza, że reakcje stają się niższe ze względu na proces uczenia się. Ten efekt jest badany przede wszystkim w kontekście bodźców wizualnych, ale ma również znaczenie w rozprzestrzenianiu informacji. Modele rozprzestrzeniania informacji często zakładają powtarzający się kontakt, ale nie uwzględniają habituacji, która obniża reaktywność w sieci. w badaniu analizuje się wpływ habituacji na rozprzestrzenianie informacji za pomocą modelu podatny-zarażony (SI), który jest często wykorzystywany w innych modelach. Wykazano, że spadek habituacji znacząco wpływa na procesy propagacji informacji, pogarszając ich wyniki. Nawet niewielkie zmiany w poziomie habituacji mają istotny wpływ na te procesy propagacyjne.

Mechanizm redukcji habituacji bazuje na zróżnicowaniu oraz częstotliwości propagacji informacji.

A10. "Evaluation of the Costs of Delayed Campaigns for Limiting the Spread of Negative Content, Panic and Rumours in Complex Networks"

Wzrost wydajności procesów rozprzestrzeniania informacji i maksymalizacja wpływu są istotne z perspektywy marketingu i działań w sieciach społecznościowych. Istnieje także potrzeba tłumienia procesów rozprzestrzeniania w celu ograniczenia rozpowszechniania dezinformacji, informacji szkodliwych dla zdrowia publicznego lub redukcji roli konkurentów na rynku. Działania tłumiące mogą polegać na rozpowszechnianiu konkurencyjnych treści, a ich skuteczność zależy od czasu i intensywności kampanii. Przedstawione badanie wskazuje, że opóźnienie w uruchomieniu działań tłumiących może być zrekompensowane poprzez odpowiednie dostosowanie parametrów, co nadal może przynieść pożądane rezultaty.

Redukcja habituacji następuje poprzez zróżnicowanie oraz częstotliwość propagacji informacji.

[4] Sekcja powstała w oparciu o publikacje **A9** i **A10**

5.5. Media strumieniowe [A11] [5]

A11. "Modeling the Impact of Video Dynamics on User Engagement and Eye Tracking Patterns"

Badania dotyczące potencjału treści wideo w przyciąganiu uwagi użytkowników oraz poprawy ich skuteczności bez zwiększania obciążenia poznawczego i unikania reklam są istotne dla prowadzenia skutecznych kampanii marketingowych. w publikacji zbadano, jak dynamika wideo, mierzona za pomocą dedykowanych metryk, jest powiązana z wzorcami śledzenia wzroku i czy może przewidywać zaangażowanie użytkownika, mierzone liczbą fiksacji i czasem spędzonym na oglądaniu treści wideo. Wyniki wskazują, że wybrane metryki mogą służyć jako predyktory wzorców ruchu oczu i że dynamika wideo ma rzeczywisty wpływ na zaangażowanie użytkowników w oglądane treści wideo. Ogólnie rzecz biorąc, wyniki sugerują, że im większa dynamika wideo, tym większe zaangażowanie użytkowników, wyrażające się w dłuższym czasie oglądania, większej liczbie odwiedzin obszaru wyświetlania wideo i większej liczbie fiksacji. Jednak należy unikać nadmiernego nasycenia treścią wideo dynamiką, ponieważ może to prowadzić do przeciwnego efektu i zmniejszonego zaangażowania użytkowników. Skuteczne treści wideo marketingowe wymagają integracji kilku elementów, takich jak apel emocjonalny, dynamika i techniki przyciągania uwagi, aby zwiększyć zaangażowanie użytkowników. Niemniej jednak, mimo dobrze zaprojektowanych treści, intensywne korzystanie z treści wideo w celach marketingowych może prowadzić do zachowań unikania, pomijania lub blokowania reklam wideo. Użytkownicy mają coraz krótszą tolerancję dla długich treści wideo. w takiej sytuacji warto eksplorować metody, które pozwolą na projektowanie efektywnych treści bez uciekania się do technik, które pogarszają doświadczenie użytkownika. Badanie sugeruje, że twórcy treści wideo mogą próbować kontrolować zaangażowanie użytkowników poprzez stosowanie scen o różnym stopniu dynamizmu. W przyszłości istotne będzie identyfikowanie cech wideo nie tylko na podstawie różnic kolorystycznych, ale także na podstawie cech obiektów obecnych w kadrze, aby możliwe było ocenianie różnic na podstawie elementów sceny, a nie tylko intensywności wizualnej.

Mechanizm redukcji habituacji bazuje na zróżnicowaniu, częstotliwości oraz intensywności kontentu wideo.

5.6. Modelowanie wpływu habituacji i przerw w procesie eksploatacji na działanie wielorękich bandytów [A12][6]

A12. "Modeling the Impact of Habituation and Breaks in Exploitation Process on Multi-Armed Bandits Performance"

Habituacja jest powszechnym zjawiskiem w procesie uczenia, gdzie reakcja na powtarzający się bodziec maleje wraz z upływem czasu. w algorytmach Multi-Armed Bandits (MAB), które są wykorzystywane do optymalizacji dostarczania treści wizualnych, habituacja może prowadzić do suboptymalnej wydajności. Na przykład, jeśli agent staje się zhabituowany do podwójnej ramy, może kontynuować wybieranie tej ramy, nawet

[5]Sekcja powstała w oparciu o publikacje **A11**
[6]Sekcja powstała w oparciu o publikacje **A12**

jeśli są dostępne lepsze opcje. Habituację można modelować jako formę "zapominania", gdzie agent stopniowo traci pewność w swoje oszacowania prawdopodobieństw nagrody dla każdej ramy w miarę upływu czasu. Proponowane podejście pozwala na aktualizację oszacowań zgodnie z modelem habituacji, jednocześnie wykorzystując ramę o najwyższym oszacowanym prawdopodobieństwie nagrody. Wyniki pokazały, że jest to bardzo dobre rozwiązanie wprowadzenia przerw w modelu habituacji Multi-Armed Bandit, który został zaimplementowany i przetestowany w naszym badaniu. w artykule opisano zjawisko habituacji w kontekście uczenia oraz wpływ tego zjawiska na algorytmy Multi-Armed Bandits (MAB), które są stosowane do optymalizacji dostarczania treści wizualnych. Wskazano, że habituacja może prowadzić do wybierania suboptymalnych opcji przez agenta, nawet gdy są dostępne lepsze możliwości. Zaproponowano podejście, które uwzględnia model habituacji i pozwala na aktualizację oszacowań nagród dla poszczególnych rąk w zależności od tego zjawiska. Jednocześnie wykorzystuje się informację o najwyższym oszacowanym prawdopodobieństwie nagrody dla ramy. Przeprowadzone badania potwierdziły, że wprowadzenie przerw w modelu habituacji MAB jest skutecznym rozwiązaniem. Badanie to zostało zaimplementowane i przetestowane, co pozwoliło uzyskać obiecujące wyniki. Ogólnie rzecz biorąc, habituacja jest złożonym zjawiskiem, które może wpływać na wydajność algorytmów Multi-Armed Bandits (MAB). Jednak poprzez zastosowanie adaptacyjnych strategii eksploracji, algorytmów "restless bandit"lub algorytmów bandyty kontekstowego, badacze mogą zmniejszyć szansę wystąpienia habituacji i poprawić wydajność algorytmów MAB. Ważne jest zauważenie, że jest to ogólny przegląd i konieczne są dalsze badania i studia w celu opracowania bardziej szczegółowego i kompletnego artykułu. Podsumowując, habituacja jest istotnym aspektem algorytmów Multi-Armed Bandit (MAB), ponieważ pomaga agentowi w równoważeniu eksploracji i eksploatacji. Poprzez modelowanie habituacji jako formy "zapominania", agent może kontynuować aktualizację swoich oszacowań prawdopodobieństw nagrody dla każdej ramy, jednocześnie wykorzystując ramę z najwyższym oszacowanym prawdopodobieństwem nagrody. Ogólnie rzecz biorąc, habituacja jest istotnym czynnikiem do rozważenia podczas projektowania i wdrażania algorytmów MAB w celu osiągnięcia optymalnych rezultatów. Oprócz zmniejszenia wpływu habituacji, przerwy w ekspozycji na bodźce w problemie MAB mają kilka innych korzyści. Po pierwsze, umożliwiają użytkownikowi dostęp do innych ram algorytmu, co może prowadzić do odkrycia bardziej korzystnych decyzji. Po drugie, przerwy zapobiegają szybkiemu wyczerpaniu najlepszej ramy i uniemożliwiają utknięcie algorytmu w suboptymalnych rozwiązaniach. Ostatecznie, zastosowanie przerw w ekspozycji na bodźce może poprawić stabilność i efektywność algorytmu MAB, zmniejszając wpływ losowości na wyniki i zwiększając precyzję decyzji. Ogólnie rzecz biorąc, przerwy w habituacji są ważnym narzędziem w problemie MAB i mogą pomóc osiągnąć lepsze wyniki i lepsze doświadczenia użytkowników.

Artykuł ten uogólnia dotychczasowe badania i mechanizmy redukcji habituacji w systemie uczącym się.

6. Podsumowanie

Sekwencyjna metoda nieinwazyjnej redukcji habituacji odnosi się do strategii mających na celu zmniejszenie wpływu habituacji, czyli utraty zainteresowania użytkowników, na systemy interaktywne, w tym systemy rekomendujące, media strumieniowe, platformy społecznościowe czy platformy komunikacyjne.

Badania zostały posadowione w dyscyplinie *Informatyka techniczna i telekomunikacja*, a w szczególności zgodnie z klasyfikacją ACM Digital Library Computing Classification System w obszarach *Human-centered computing · Human computer interaction (HCI) · HCI design and evaluation methods · Laboratory experiments.*

W interaktywnych systemach rozrywkowych, sekwencyjne podejście może być stosowane do utrzymania zaangażowania użytkowników poprzez dostosowywanie interfejsu użytkownika, stosowanie gamifikacji, organizowanie interaktywnych działań społecznych i promowanie różnorodnych form komunikacji. Dodatkowo, wykorzystanie technik personalizacji treści czy modelowanie preferencji użytkownika, pozwala dostarczyć rekomendacje dokładnie dopasowane do indywidualnych potrzeb i zainteresowań **(A1, A2)**.

W interaktywnych systemach informacyjnych, zwłaszcza rekomendujących, sekwencyjna redukcja habituacji znajduje odzwierciedlenie w strategiach, które mają na celu zapobieganie habituacji użytkowników poprzez różnorodność i zmienność prezentowanych rekomendacji. System rekomendacyjny dąży do prezentowania różnorodnych rekomendacji zamiast skupiać się tylko na najpopularniejszych lub najbardziej podobnych elementach. Może to obejmować rekomendacje z różnych kategorii, różne style lub warianty podobnych elementów. Celem jest też dostarczanie zróżnicowanych i interesujących rekomendacji, które utrzymują zainteresowanie użytkowników. Poprzez zastosowanie różnych algorytmów, uwzględnianie kontekstu, eksplorowanie nowych obszarów i prezentowanie różnorodnych rekomendacji. System rekomendacyjny może przeciwdziałać habituacji i zapewnić bardziej satysfakcjonujące doświadczenie użytkownikom **(A3, A4, A5)**.

W interaktywnych systemach dialogowych, w szczególności dotyczących komunikatów oraz obiektów wizualnych w obszarze chatbota, wykorzystuje się podejścia, które mają na celu zmniejszenie efektu habituacji i utrzymanie zainteresowania użytkownika. W przypadku komunikatów tekstowych, chatbot może korzystać z różnych wariantów treści, aby uniknąć monotonii. Może to obejmować zmienne frazy, struktury zdania, długość komunikatów, a także różne sposoby przedstawiania informacji. Rotacja treści zapewnia większą różnorodność w komunikatach i minimalizuje przewidywalność. W obszarze chatbota zaproponowana metoda polega na dostarczaniu różnych obiektów wizualnych w odpowiedzi na zapytania użytkownika. Może to obejmować pokazywanie różnych obrazów, zmienne układy graficzne lub dynamiczne prezentowanie danych w atrakcyjny sposób. Celem sekwencyjnej względem redukcji habituacji jest zapewnienie różnorodności, zaskoczenia i personalizacji w interakcjach z chatbotem. Poprzez stosowanie tych technik, chatbot staje się bardziej angażujący i trzyma użytkownika zainteresowanego, co przekłada się na lepsze doświadczenie użytkownika z systemem dialogowym **(A6, A7, A8)**.

W przypadku mediów społecznościowych i dyfuzji informacji, wykorzystuje się podjęście, które umożliwia skuteczną kontrolę i zarządzanie procesem adaptacji i utraty zainteresowania użytkowników, co przekłada się na długotrwałe i trwałe oddziaływanie treści mediów społecznościowych. Dzięki temu zastosowaniu możliwe jest utrzymanie uwagi użytkowników oraz pobudzanie ich zaangażowania, co stanowi kluczowy czynnik sukcesu w kontekście dyfuzji informacji w mediach społecznościowych.**(A9, A10)**.

W przypadku mediów stumieniowych, wprowadzono dynamiczne dostosowanie rekomendacji na podstawie bieżącego kontekstu użytkownika, dostosowanie dynamiki scen wideo, wykorzystanie technik personalizacji treści oraz zróżnicowanie prezentowanych propozycji, aby uniknąć monotonii i znużenia użytkownika **(A11)**.

Biorąc pod uwagę publikacje, które wchodza w skład rozprawy tj. od A1 do A11 pozyskano niezbędną wiedzę, która wykorzystano w publikacji [A12]. Na ich podstawie opracowano sekwencyjną metodę nieinwazyjnej redukcji habituacji. Zaproponowano w nim podejście wykorzystujące algorytm MAB, który spełnia podstawowe cechy systemu uczącego się. W algorytmie MAB redukcja habituacji wiąże się z wprowadzeniem przerw w eksploracji i eksploatacji, aby uniknąć przewlekłej eksploatacji wybranych opcji i zachęcić do eksploracji innych, potencjalnie lepszych wariantów **(A12)**.

Przeprowadzone badania wzbogacają wiedzę na temat skutecznych strategii utrzymania zaangażowania użytkowników. Proponowane podejście może przyczynić się do tworzenia systemów interaktywnych, które lepiej odpowiadają zmieniającym się oczekiwaniom użytkowników. W każdym z rozpatrywanych obszarów metoda dąży do zapewnienia dynamiki, personalizacji i różnorodności doświadczeń użytkownika. Poprzez unikanie monotonii i znużenia, systemy interaktywne mogą lepiej odpowiadać na zmienne oczekiwania i preferencje użytkowników, co przekłada się na wyższe zaangażowanie i satysfakcję.

Podsumowując, sekwencyjna metoda nieinwazyjnej redukcji habituacji ma na celu zapewnienie efektywnych komunikatów dla których nie sa stosowane techniki inwazyjne użytkownikom w systemach interaktywnych, takich jak systemy rekomendujące, platformy społecznościowe oraz wideo strumieniowe. Poprzez zastosowanie odpowiednich strategii, takich jak personalizacja, zróżnicowanie, sekwencyjność i interaktywność, można zmniejszyć wpływ habituacji i utrzymać zaangażowanie użytkowników na wysokim poziomie. Wiedza z cyklu publikacji wchodzących w skład rozprawy doktorskiej została syntezowana i wykorzystana do opracowania sekwencyjnej metody nieinwazyjnej redukcji habituacji w systemach interaktywnych, którą ugoólniono w publikacji [A12]. Zaproponowano w niej system uczący się i w nim włączono mechanizmy habituacji obliczeniowej, wprowadzono parametry habituacyjne, oraz wprowadzono przerwyw komunkacji. w ramach realizowanych badań cel rozprawy został osiągnięty.

Osiągnięcia uzyskane w rozprawie, które stanowią wkład do *Informatyki technicznej i telekomunikacji* jednocześnie potwierdzają realizację celu rozprawy oraz postawioną tezę obejmują:

— włączenie mechanizmów habituacji obliczeniowej do algorytmu MAB;

— wprowadzenie parametrów habituacyjnych do algorytmu MAB;

— wprowadzenie przerw w celu zredukowania habituacji w algorytmie MAB;

- opracowanie podejścia wykorzystujacego sekwencyjność ekspozycji obiektów oraz komunikatów w chatbocie z wykorzystaniem zmian w poszczególnych etapach sekwencji;
- opracowane podejście analityczne z wykorzystaniem analizy przetrwania, która pozwala ocenić efektywność prezentowanych sekwencyjnie rekomendacji;
- zbadanie i zaproponowanie analizy zróżnicowania sekwencji scen w wideo;
- zbadanie i opracowanie podejścia o umiarkowanej intensywności wizualnej w systemach rekomendujących przy preferowanym braku wyróżnienia lub silnego wyróżnienia;
- zaproponowano model wielokryterialny, który pozwala ocenić strategie aktualizacji przy sekwencyjnie wprowadzanych obiektach cyfrowych w grach wideo;
- zaproponowano podejście wykorzystujące wyższą częstotliwość poziomów migotania o możliwym najniższym oddziaływaniu na użytkownika przy niskiej intensywności wizualnej;
- zintegrowanie mechanizmów habituacji w modelu rozprzestrzeniania informacji w sieciach złożonych;
- opracowanie, zaimplementowanie oraz zrealizowanie szeregu eksperymentów z wykorzystaniem eyetrackera;

Przeprowadzone badania pozwoliły na potwierdzenie postawionej w rozprawie tezy. Analiza charakterystyki wizualnej obiektów wyświetlanych sekwencyjnie w procesie korzystania z systemu interaktywnego oraz reakcji odbiorcy umożliwia wyznaczanie poziomu oddziaływania wizualnego dla stanów przyszłych, które zapewnią pozyskanie uwagi użytkownika.

W trakcie prac nad rozprawą doktorską zidentyfikowano potencjalne obszary dalszych badań. Przykładowo dla dalszych efektów w zakresie stosowania przerw w komunikacji kluczowe może być dynamicznie dostosowywanie przerw w zależności od reakcji użytkowników i ewolucji ich preferencji. Algorytm MAB może analizować dane zwrotne i dostosowywać ich częstotliwość oraz długość przerw, tak, aby utrzymać zaangażowanie użytkowników na wystarczającym poziomie. Na przykład, jeśli użytkownicy wykazują większe zainteresowanie po krótkiej przerwie, system może dostosować czas trwania przerw w celu uniknięcia habituacji i utrzymać jego zaangażowania.

7. Dorobek akademicki

Niniejszy rozdział prezentuje dorobek akademicki mgra Kamila Bortko, kandydata do stopnia naukowego doktora. W ramach prezentowanego dorobku wyodrębniono osiągnięcia naukowe, dydaktyczne, organizacyjne i zawodowe.

7.1. Dorobek naukowy

7.1.1. Profile internetowe[1]

Wskaźniki z profilu internetowego Google Scholar zebrane zostały w tabeli 7.1.

Tabela 7.1. Profile internetowe (stan na dzień 28.06.2023).

Profil	Liczba artykułów	Liczba cytowań	h-index
Google Scholar	17	47	4

7.1.2. Wykaz pozostałych prac naukowych[2]

Poniżej przedstawiono wykaz pozostałych prac naukowych opublikowanych oraz w druku:

1. Lewandowska A., Rejer I., **Bortko K.**, Jankowski J. (2022). Eye-Tracker study of influence of affective disruptive content on user's visual attention and emotional state. Sensors, 22(2), 547.
2. Pazura P., **Bortko K.**, Jankowski J., Michalski R. (2020). A Dynamic Vote-Rank Based Approach for Effective Sequential Initialization of Information Spreading Processes Within Complex Networks. In Computational Science–ICCS 2020: 20th International Conference, Amsterdam, The Netherlands, June 3–5, 2020, Proceedings, Part I 20 (pp. 638-651). Springer International Publishing.
3. Karczmarczyk A., **Bortko K.**, Bartków P., Pazura P., Jankowski J. (2018, August). Influencing information spreading processes in complex networks with probability spraying. In 2018 IEEE/ACM international conference on advances in social networks analysis and mining (ASONAM) (pp. 1038-1046). IEEE.
4. Pazura P., Jankowski J., **Bortko K.**, Bartków P. (2019, August). Increasing the diffusional characteristics of networks through optimal topology changes within sub-graphs.

[1]Stan na dzień 28.06.2023 r.
[2]Stan na dzień 28.06.2023 r.

In Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (pp. 994-1000).

5. Cypryjański J., Grzesiuk A., **Bortko K.**(2020). What Really Helps Us Make a Choice? An Experimental Evaluation of AHP. In Experimental and Quantitative Methods in Contemporary Economics: Computational Methods in Experimental Economics (CMEE) 2018 Conference (pp. 353-363). Springer International Publishing.

7.1.3. Pozostałe

— Udział w grantach badawczych w charakterze wykonawcy/stypendysty:

• Wspomaganie procesów rozprzestrzeniania treści marketingowych w mediach społecznościowych (OPUS, 2016/21/B/HS4/01562)

• Nieinwazyjne metody redukcji zjawiska habituacji w marketingu elektronicznym (OPUS, 2017/27/B/HS4/01216)

— Komercyjny projekt badawczy „Badanie ścieżek wzrokowych użytkowników systemu rezerwacji biletów lotniczych przy pomocy eye trackingu":

• Celem głównym badania była analiza efektywności systemu rezerwacji biletów lotniczych po wprowadzeniu poprawek na podstawie dwóch poprzednich testów użyteczności oraz wykryciu ewentualnych nowych problemów użyteczności. Dostarczone przez zleceniodawcę prototypy symulowały finalny produkt, a ewentualne poprawki mogły zostać wdrożone jeszcze przed ostatecznymi pracami deweloperskimi.

• Badanie miało w zamiarze zweryfikowanie wzorców zachowań z wykorzystaniem trzech scenariuszy obsługi systemu rezerwacji biletów lotniczych. Kolejnym celem badania była identyfikacja kluczowych zachowań w odniesieniu do elementów stron i etapów interakcji z punktu widzenia

— Pełniłem funkcję tutora w projekcie dofinansowanym ze środków UE "Szkoła Orłów ZUT"w roku akademickim 22/23.

7.2. Dorobek dydaktyczny

7.2.1. Kursy

W ramach dotychczasowej pracy dydaktycznej prowadziłem zajęcia na poziomie S1 zgodnie z wykazem:

1. Bazy danych
   – cykl 15 laboratoriów
2. Narzędzia inżynierskie
   – cykl 4 laboratoriów
3. Wprowadzenie do informatyki
   – cykl 3 laboratoriów

4. Projektowanie zorientowane na użytkownika
   – cykl 15 laboratoriów
5. Projektowanie zorientowane na człowieka
   – cykl 15 wykładów i 15 laboratoriów
6. Pracownia dyplomowa
   – cykl 8 seminariów
7. Inżynierski projekt zespołowy
   – cykl 15 spotkań projektowych
8. Zarządzanie informacją 2
   – cykl 3 laboratoriów
9. Duże zbiory danych
   – cykl 4 laboratoriów

7.2.2. Prace dyplomowe

Podczas dotychczasowej pracy dydaktycznej byłem promotorem następujących prac inżynierskich:

1. Opracowanie aplikacji do automatycznego testowania stron internetowych szkół wyższych w Szczecinie.
2. Model interfejsu użytkownika do badania reakcji użytkownika poprzez ocenę, wyszukiwanie oraz zamawianie online potraw w restauracjach.
3. Model interfejsu użytkownika do badania reakcji na wyświetlane treści multimedialne przy wykorzystaniu prawa Fittsa.
4. Modelowanie wpływu zmienności treści w botach internetowych na poziom uwagi użytkowników.
5. Model aplikacji do oceny oraz tworzenia rankingów filmów i seriali przy wykorzystaniu danych z zewnętrznych serwisów filmowych.
6. Badanie wpływu elementów wizualnych na zainteresowanie użytkowników treścią botów internetowych.
7. Model systemu doboru poziomu intensywności oddziaływania wizualnego w systemach internetowych przy użyciu okulografu.
8. Analiza oddziaływania treści reklamowych na stronach internetowych na odbiorcę.
9. Model interfejsu użytkownika do badania reakcji na prezentowane treści multimedialne.

7.3. Dorobek organizacyjny

— Poprowadzenie prelekcji na targach pracy R@BBIT 2022.

— Jestem opiekunem oraz brałem udział w przygotowaniu laboratorium do badań systemów informatycznych i gier na WI sala 001 z udziałem eyetrackingu.

7.4. Dorobek zawodowy

7.4.1. Historia zatrudnienia

2016-2017 Bertelsmann Media Sp. z o.o. z siedzibą w Warszawie Oddział Arvato Services Polska w Plewiskach
Analityk specjalista ds. raportowania i analiz danych

2017-2018 STU Ergo Hestia
Analityk zespołu specjalistów ds. oceny ryzyka

2018-obecnie Wydział Informatyki ZUT w Szczecinie
Pracownik naukowo-dydaktyczny na stanowisku aystent

Bibliografia

[1] J. Smith and A. Johnson, “The impact of interactive technologies on human-computer communication,” *Journal of Interactive Systems*, vol. 45, no. 2, pp. 78–92, 2022.

[2] R. Brown and M. Davis, “Advancements in human-computer interaction for automotive communication and assistance,” in *Proceedings of the International Conference on Human-Computer Interaction*, pp. 125–136, 2021.

[3] S. Wilson and L. Thompson, “Home automation and interactive interfaces: Enhancing user experience,” *Journal of Interactive Technology*, vol. 33, no. 4, pp. 187–200, 2020.

[4] C. Johnson and E. Williams, “Chatbots and customer service: Improving interactions in the digital age,” *International Journal of Customer Experience*, vol. 12, no. 3, pp. 112–125, 2019.

[5] M. Adams *et al.*, “Applications of interactive technologies in the medical field,” *Journal of Medical Informatics*, vol. 26, no. 1, pp. 45–58, 2018.

[6] A. Lopez *et al.*, “Interactive voice-based systems for medical diagnosis: A review,” *International Journal of Healthcare Technology*, vol. 9, no. 2, pp. 73–88, 2017.

[7] R. Patel *et al.*, “Multimodal interfaces for medical education and training,” *Journal of Medical Education*, vol. 15, no. 4, pp. 201–215, 2019.

[8] K. Smith and B. Wilson, “Enhancing business and educational interactions through interactive technologies,” *International Journal of Business Communication*, vol. 55, no. 3, pp. 178–192, 2023.

[9] R. Johnson and S. Brown, “Role of user interface communication in system functionality,” *International Journal of Human-Computer Interaction*, vol. 37, no. 3, pp. 145–158, 2022.

[10] J. Smith *et al.*, “Warning messages in user interfaces: Effectiveness and design guidelines,” in *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 231–242, 2021.

[11] M. Davis *et al.*, “The role of marketing content in user interaction: A case study of entertainment systems,” *Journal of Interactive Marketing*, vol. 28, no. 2, pp. 89–102, 2020.

[12] A. Wilson *et al.*, “Impact of stimuli-rich content on user experience,” *International Journal of Human-Computer Studies*, vol. 76, no. 4, pp. 278–291, 2019.

[13] L. Thompson and M. Adams, “Information overload in user interfaces: Effects on user attention and interaction quality,” in *Proceedings of the International Conference on Human-Computer Interaction*, pp. 112–125, 2018.

[14] R. Patel *et al.*, “The influence of daily exposure to marketing elements on user interaction,” *Journal of Marketing Research*, vol. 45, no. 1, pp. 67–82, 2017.

[15] P. M. Groves and R. F. Thompson, “Habituation: a dual-process theory,” *Psychological Review*, vol. 77, no. 5, pp. 419–450, 1970.

[16] C. H. Rankin, T. Abrams, R. J. Barry, S. Bhatnagar, D. F. Clayton, J. Colombo, and J. L. Lissemore, "Habituation revisited: an updated and revised description of the behavioral characteristics of habituation," *Neurobiology of Learning and Memory*, vol. 92, no. 2, pp. 135–138, 2009.

[17] A. Wilson *et al.*, "Computational approaches to habituation in marketing," *Journal of Marketing Analytics*, vol. 15, no. 2, pp. 78–92, 2022.

[18] R. Johnson and S. Brown, "Computational modeling of habituation in interactive systems," in *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 231–242, 2021.

[19] M. Davis *et al.*, "Machine learning models for simulating habituation processes," *Journal of Artificial Intelligence Research*, vol. 28, no. 4, pp. 187–200, 2020.

[20] C. D. Wickens, "Multiple resources and mental workload," *Human Factors*, vol. 50, no. 3, pp. 449–455, 2008.

[21] J. Smith *et al.*, "Banner blindness and habituation: The impact on user interaction in interactive systems," *Journal of Interactive Advertising*, vol. 38, no. 2, pp. 89–102, 2022.

[22] R. Johnson and S. Brown, "Effects of habituation on the effectiveness of warning messages in user interfaces," in *Proceedings of the International Conference on Human-Computer Interaction*, pp. 125–136, 2021.

[23] M. Davis *et al.*, "The influence of habituation on the impact of motivational messages in interactive systems," *Journal of Applied Psychology*, vol. 45, no. 3, pp. 178–192, 2020.

[24] Y. Li and Q. Zhu, "Polymorphic messages: A non-intrusive approach to mitigate habituation in human-computer interaction," 2015.

[25] A. Felfernig, G. Friedrich, and M. Mandl, "Personalized user interfaces: Non-intrusive adaptation to habituation," 2008.

[26] S. Oviatt and P. Cohen, *Beyond Human-Computer Interaction: New Modes of Work, Collaboration, and Technology.* Morgan Kaufmann, 2000.

[27] I. Szkudlarek-Mikhoels and D. Taff, "Interpersonal dynamics in synchronous computer-mediated communication: Exploring sequential patterns and their impact on perceptions of communication quality," *Journal of Language and Social Psychology*, vol. 37, no. 6, pp. 682–703, 2018.

[28] M. Dainton and E. D. Zelley, *Applying Communication Theory for Professional Life: A Practical Introduction.* 2019.

[29] K. Miller, *Communication Theories: Perspectives, Processes, and Contexts.* 2020.

[30] A. Amran, Z. F. Zaaba, and M. K. Mahinderjit Singh, "Habituation effects in computer security warning," *Information security journal: A global perspective*, vol. 27, no. 4, pp. 192–204, 2018.

[31] O. Déniz, J. Lorenzo, and M. Hernández, "A simple habituation mechanism for perceptual user interface," *Inteligencia Artificial. Revista Iberoamericana de Inteligencia Artificial*, vol. 8, no. 23, p. 0, 2004.

[32] K. T. Hanson and E. P. Riley, "Beyond neutrality: The human–primate interface during the habituation process," *International Journal of Primatology*, vol. 39, pp. 852–877, 2018.

[33] J. Johnson, *Designing with the mind in mind: simple guide to understanding user interface design guidelines.* Morgan Kaufmann, 2020.

[34] Y. Li and L. Zhang, *Non-Invasive Methods in Computer Science: Advancements and Applications.* 2020.

[35] Y. Wang and L. Li, *Multi-Armed Bandit Algorithms for Machine Learning Applications.* 2019.

[36] T. Lattimore and C. Szepesvári, *Bandit algorithms.* Cambridge University Press, 2020.

[37] S. Agrawal and N. Goyal, "Analysis of thompson sampling for the multi-armed bandit problem," 2012.

[38] "Pwn - polski słownik języka polskiego." Dostęp online, 2023.

[39] W. M. Newman and M. G. Lamming, *Interactive system design.* Addison-Wesley Reading, 1995.

[40] M. Johnson, *Interactive Systems: Principles and Design.* AMC Publishing, 2022.

[41] A. M. Rashid, I. Albert, D. Cosley, S. K. Lam, S. M. McNee, J. A. Konstan, and J. Riedl, "Getting to know you: learning new user preferences in recommender systems," in *Proceedings of the 7th international conference on Intelligent user interfaces*, pp. 127–134, 2002.

[42] H. Chen, X. Liu, D. Yin, and J. Tang, "A survey on dialogue systems: Recent advances and new frontiers," *Acm Sigkdd Explorations Newsletter*, vol. 19, no. 2, pp. 25–35, 2017.

[43] L. Nacke, A. Drachen, and D. Johnson, *Entertainment Computing and Serious Games: Game-based Frameworks for Prevention and Intervention.* Routledge, 2019.

[44] J. Kim and J.-H. Lee, "Exploring the effectiveness of interactive advertising on smart tv: Focus on consumers' attitude and engagement," *Journal of Interactive Advertising*, vol. 16, no. 1, pp. 36–48, 2016.

[45] D. Boyd and N. Ellison, "Social network sites: Definition, history, and scholarship," *Journal of Computer-Mediated Communication*, vol. 13, no. 1, pp. 210–230, 2007.

[46] D. Benyon, P. Turner, and S. Turner, *Designing interactive systems: People, activities, contexts, technologies.* Pearson Education, 2005.

[47] A. Dix, "Human–computer interaction, foundations and new paradigms," *Journal of Visual Languages & Computing*, vol. 42, pp. 122–134, 2017.

[48] J. Triesch and D. H. Ballard, "Efficient and adaptive sensory processing," *Network: Computation in Neural Systems*, vol. 8, no. 4, pp. 503–522, 1997.

[49] J. Preece, Y. Rogers, and H. Sharp, *Interaction Design: Beyond Human-Computer Interaction.* Wiley, 5th ed., 2019.

A1.

Bortko, K., Bartków, P., Pazura, P., Jankowski, J. (2019). Increasing User Engagement and Virtual Goods Life Span Through Products Diversity and Intensity of Content Updates. In Intelligent Information and Database Systems: 11th Asian Conference, ACIIDS 2019, Yogyakarta, Indonesia, April 8–11, 2019, Proceedings, Part II 11 (pp. 519-530). Springer International Publishing.

Increasing User Engagement and Virtual Goods Life Span Through Products Diversity and Intensity of Content Updates

Kamil Bortko, Piotr Bartków, Patryk Pazura, and Jarosław Jankowski( )

Faculty of Computer Science and Information Technology,
West Pomeranian University of Technology in Szczecin,
ul. Żołnierska 49, 71-210 Szczecin, Poland
{kbortko,pbartkow,ppazura,jjankowski}@wi.zut.edu.pl

Abstract. The virtual goods sector has become one of the main business models for social platforms, games and virtual worlds. While online systems are under continuous development, their users require frequent updates of virtual goods, new digital content and functionality. System developers face dilemmas concerning the frequency of updates or content drops to decrease the habituation effect and increase the life span of digital products. The presented research shows how the diversity and number of virtual products can increase user engagement and interest in new products. Apart from the empirical study, a multi-criteria decision support model is proposed for the evaluation of strategies used in system updates and virtual goods distribution.

1 Introduction

Nowadays, virtual worlds live in symbiosis with the real world. They can be implemented within games, social platforms or dedicated systems with complex social and economic phenomena observed [3]. They enable the possibility of correlating real-time communication with respect to economic activity between users. Together with their evolution, new business models were developed based on the distribution of virtual goods [4] such as avatars [7]. The sale of virtual goods has become an important business model for online games and virtual worlds [1,2]. Virtual goods usually refer to virtual items such as avatar clothing, weapons, pets, coins, characters and tokens [10,12]. While there is substantial research on the motivation to purchase virtual goods [11], the research related to identifying the role of content characteristics for product life span and factors affecting user engagement is limited. Online systems are continuously being updated with new game content [20] and then content spreads within the system [16]. Adjusting the frequency, volume and quality of this new content requires analysis and planning. The presented research shows how strategies of content updates affect user engagement and product life span. The empirical study is

N. T. Nguyen et al. (Eds.): ACIIDS 2019, LNAI 11432, pp. 519–530, 2019.
https://doi.org/10.1007/978-3-030-14802-7_45

followed by a multi-criteria evaluation of these strategies. The rest of the paper is organized as follows: Sect. 2 includes the literature review and Sect. 3 presents the conceptual framework and assumptions for the proposed approach. In the next section, the empirical results are presented, followed by a multi-criteria strategy evaluation presented in Sect. 5 and summary in Sect. 6.

2 Related Work

Virtual goods are the basic source of income for many internet ventures and platforms [10]. There has been a substantial increase in research on the purchases of virtual goods over the last decade [11]. The literature focuses primarily on the relationship between game design and the business model of selling content within games and virtual worlds [11] and their role in creating a positive user experience [2]. Players motivations to acquire various virtual goods are analyzed, as well as developing advantages over the competition. An important aspect of this is related to expressing yourself through, for example, a special outfit, theme or avatar. The motivation shown by users in participating in activities and using the services they engage in is related to their attitudes towards virtual goods [12].

Virtual goods are usually closely related to the specific game in which customers buy them. Therefore, game developer should foresee the continuity and benefits of using and purchasing virtual goods. However, before taking advantage of this, the intention of further use must be assessed [9]. An important aspect of this is the subjective assessment of the player based on the psychological assessment of motivation and exploring the mechanisms of the decision-making process. This is measured by the level of user involvement when buying virtual goods.

From the perspective of MMO (massively multiplayer online) games, we can see a relationship between revenue and the motivation of players [24]. Motivation itself can result from various aspects such as maintaining distance from other players, acquiring new things or even identifying with a virtual character. It was discovered that revenue is strongly correlated with the user's motivations. Therefore, we must analyze how these aspects, such as usefulness and attitude, prompt the user to purchase virtual goods.

Often, because it costs nothing to get the game itself, free online games are strongly focused on the sale of virtual items. The publishers, however, encourage the purchase of virtual goods to improve the character's abilities and collect unique items such as armor or clothing. They allow the user to emphasize their individuality. However, users can still use the website's services free of charge. This enables the publishers to build a growing base of new customers [13].

There is currently a strong emphasis on customer engagement in the usage and distribution of digital content. This is achieved with the help of various mechanisms of content propagation. People are more willing to promote and distribute different types of information [8]. People who trust the content themselves are more likely to recommend it to others whilst also being more susceptible to receiving this type of information [22].

Developers face dilemmas related to optimal frequency of virtual goods updates, their volumes and intensity with focus on continuous development [20]. A low frequency of updates can result in user churn, while frequent development of new content increases operational costs. From another point of view, users may have a limited ability for digital content consumption when content is updated frequently. This may be considered as an unwise budget allocation when content production is significantly higher than demand. Another problem is taking into account limited availability of users within the system [17].

The life span of online gaming products is usually shorter than that of traditional products, and users are constantly expecting new content and system updates [18,19]. Another problem is the habituation effect resulting from the short life span of virtual goods and limited time in which the product can attract web users. This opens up new directions for research since so far it has mainly only been studied for traditional markets [23].

3 Conceptual Framework

To gather knowledge about virtual products usage and behavior the system is monitored towards engagement, content life span and gathered knowledge is used during platform development. Generalization of presented problem assumes n content updates $U_1, U_2, ..., U_n$ with the use of m content categories $C_1, C_2, ..., C_m$ (Fig. 1). Within each update U_i and category C_j new elements $E_{i,j,k}$ are introduced, where k represents the number of element from category j assigned to content update i, where $i = 1, 2, ..., n$, $j = 1, 2, ..., m$ and $k = 1, 2, ..., N(U_i, C_j)$.

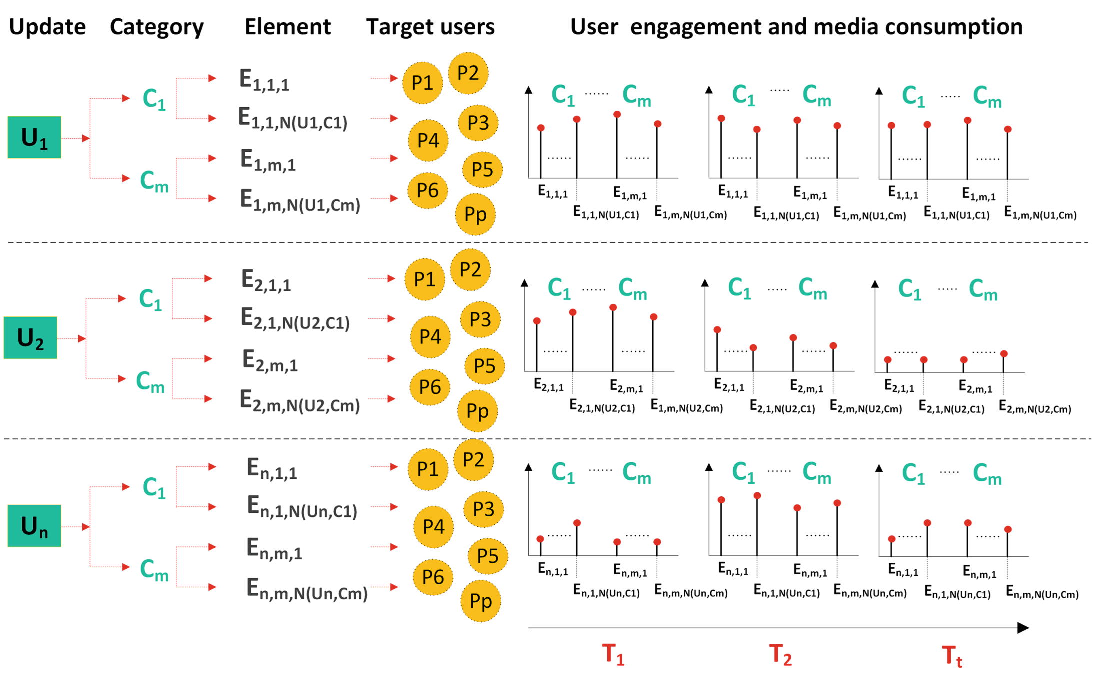


Fig. 1. Evaluation of digital content distribution strategies based on product diversification and scale of updates

Value $N(U_i, C_j)$ represents the number of elements within category C_j used during update U_i. Content is delivered to target platform's users $P = P_1, P_2, ..., P_p$. In the following time periods T_1, T_2,..., T_t user activity and consumption of each content unit is monitored. Regarding used strategies and content elements assigned to each update U_i different effect can be observed. For example update U_1 resulted continuous usage of delivered virtual products with long life span. For U_2 high intensity of usage is observed only at the beginning period and then users loose interest in new product. Update U_n resulted growing usage till saturation point and then dropping user interest observed. Different strategies for new content introduction can be considered. The complexity of the problem leads to the need for decision support tools and analytical methods for better planning and the optimization of system development strategies. It is assumed that several factors are monitored, such as total product usage or consumption dynamics, over monitored time periods. If updates are based on new elements, content production costs should be considered alongside the number of elements in each update and their diversity. A multi-criteria evaluation of results can deliver guidelines for future development planning. This leads to multi-criteria problems with preferences assigned to evaluation measures. Various methods can be used for strategies evaluation and results ranking. The PROMETHEE method was selected to create the presented research [5,6,21]. It delivers the ability of building of an outranking between different strategies. It uses be a set of solutions, in our case possible strategies A, each $a \in A$, $f_j(a)$ represents the evaluation of a solution a, to a given criterion f_j. The preference function $P_j(a, b)$ represents the degree of preference of solution a over solution b for a given evaluation criterion f_j. A multi-criteria preference index $\pi(a, b)$ of a over b is used. It takes all the criteria into account with the expression: $\pi(a, b) = \sum_{j=1}^{k} w_j P_j(a, b)$. In the decision process positive $\phi + (a)$ and negative outranking $\phi - (a)$ is used for strategies evaluation and final ranking creation. In the next sections PROMETHEE method is used for the evaluation of content development strategies used in empirical research based on multiplayer platform.

4 Empirical Results

Empirical study is based on behavior within virtual world and usage of newly introduced avatars [14,15]. Content updates where focused on avatars delivered to users with four types of elements E1, E2, E3, E4. Total twenty one updates U_1–U_{21} were taken into account with data of user activity gathered after new content was introduced. Result is shown in Table 1 each content update with showed number of times each element of avatar was used. The table shows the division into various types of data, i.e. the number of types of elements (their diversity), the number of elements, the sum of total changes and statistics for individual types of elements: E1, E2, E3 and E4. Also results divided into four weeks are presented as well as in aggregated form for all periods (Figs. 2 and 3). Comprehensive analysis of the dynamics of the division for specific weeks is presented as a reversed geometric function graph. In first week 59%

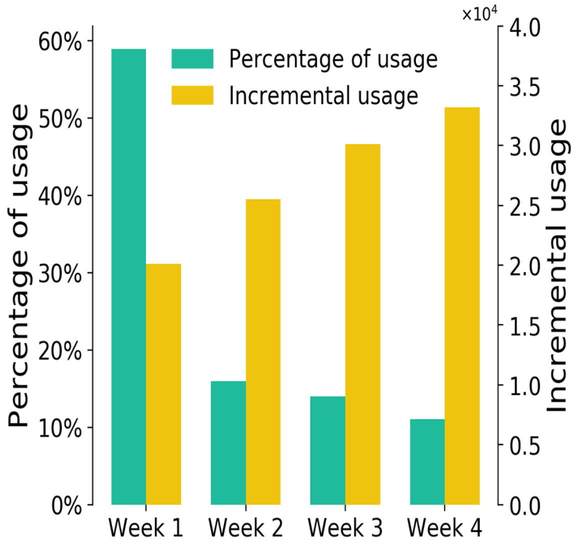


Fig. 2. Percentage of usage and incremental usage

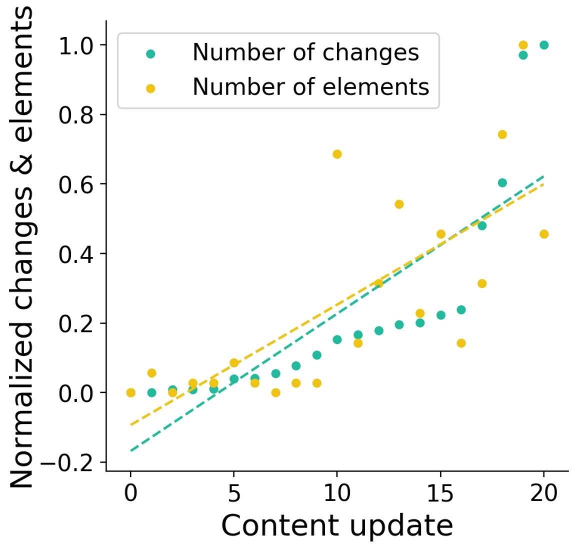


Fig. 3. Number of changes and number of elements

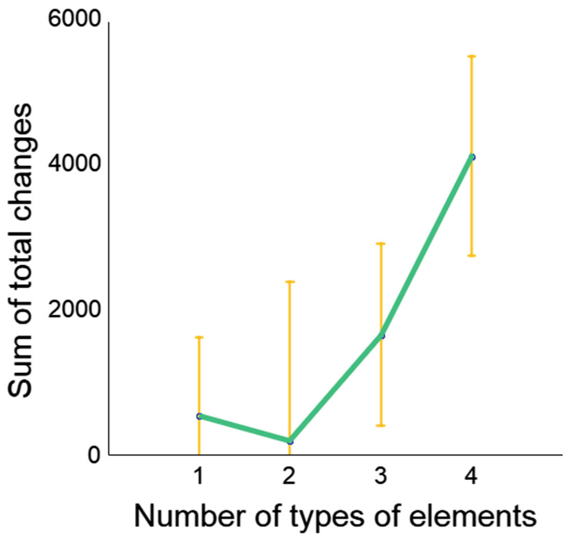


Fig. 4. Relation between total usage and product diversity represented by the number of introduced types of elements

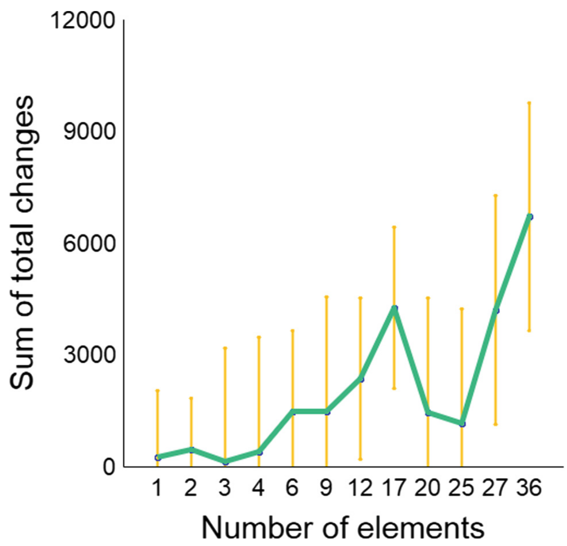


Fig. 5. Relation between total usage and the number of elements used for content update

of all usages take place. This value in the following weeks decreases accordingly: second week to 16%, third week to 14% and last week to 11%. The values of changes in individual weeks have changed as follows: Week I equal to 20060, Week II equal 25484, Week III equal to 30083 and 33797 in Week IV. The first week had the largest share in the number of changes within analyzed four weeks. In the next step ANOVA analysis was performed. We analysed number of types of elements in relation to the sum of changes within 28 days (Fig. 4). Each content updates could have a maximum of 4 types of elements. The factor F is greater than one (6.9586), it is close to the dependent variable, i.e. the number

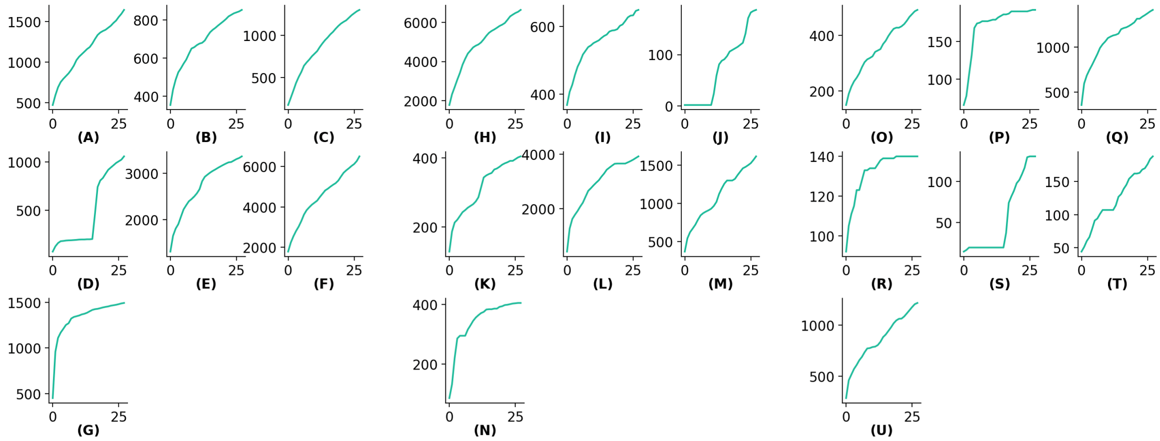


Fig. 6. Incremental usage after each content update U1–U21

Table 1. Content updates with the division into various types of data, i.e. the number of types of elements, the number of elements, the sum of total changes and statistics for individual types of elements: E1, E2, E3 and E4

Content update	Types of elements	E1	E2	E3	E4	Number of elements	Sum of changes	Week I	Week II	Week III	Week IV
U1	3	0	2	2	2	6	1750	861	300	238	246
U2	1	0	0	0	2	2	871	592	100	100	60
U3	1	0	12	0	0	12	1340	639	266	235	163
U4	4	5	6	6	8	25	1176	190	11	675	180
U5	3	7	3	0	2	12	3394	2312	611	257	184
U6	4	15	8	8	5	36	6721	3620	1026	847	998
U7	4	2	5	1	1	9	1499	1270	115	64	43
U8	4	7	3	3	4	17	6914	4142	997	751	759
U9	1	0	0	0	2	2	665	517	49	36	47
U10	2	1	0	1	0	2	198	2	79	34	72
U11	3	1	2	0	1	4	407	249	94	37	24
U12	4	8	6	9	4	27	4224	2215	953	483	264
U13	3	10	2	0	5	17	1649	844	280	246	245
U14	1	0	0	0	2	2	411	295	80	19	11
U15	1	0	1	0	0	1	507	285	65	77	64
U16	1	1	0	0	0	1	192	175	7	8	2
U17	3	10	6	4	0	20	1463	923	214	138	138
U18	1	0	0	0	3	3	140	128	10	2	0
U19	1	0	0	1	0	1	133	19	0	79	33
U20	2	1	1	0	0	2	201	94	20	48	26
U21	3	2	3	1	0	6	1259	688	147	225	155
Sum:		70	60	36	41	207	35114	20060	5424	4599	3714

of types of elements (6.95). A significance level $p <= 0.0029$ confirms the dependence of number of types of elements on the sum of changes. Another goal was to examine whether a number of changes depends on the number of elements used during update. Our dependent variable was the sum of changes within 28 days (Fig. 5). The sum of the items ranged from 1 to 36. The factor F is greater than 1 (3.16890), it is close to the dependent variable, i.e. the number of elements.

Thanks to this we can determine if the test is statistically significant. A significance level $p <= 0.047058$ obtained confirms the dependence of sum of changes within 28 days on the number of used elements. The graph shows the spread of normalized data taking into account the number of elements and the sum of changes for four weeks. We can see a clear coverage of the number of items versus the sum of changes within 28 days. Both trend lines running through the graph almost overlap. Incremental usage after each content update is presented in Fig. 6.

5 Multicriteria Strategy Evaluation

Empirical study delivered data from content updates within the real system. Various approaches were used with different number of elements, their diversity and different number of elements in each category. Analysis showed that results represented by user engagement where dependent on a number of used elements and their diversity. Used strategies can be evaluated from the perspective of costs, number of elements and results represented by user engagement depending on preferences of decision maker.

In first scenario analysis was performed from the perspective of four criteria: content production cost, number of types of elements used, number of elements and the total usage in analyzed period. Three variants were analyzed with different weights assigned to criteria. For the first variant (Cost variant I) the same weights to all the criteria were assigned. Results for this variant shows the ranking of the strategies with best results achieved by eight content update (U8).

Analyzing the result from PROMETHEE we can conclude that U8 is preferred to all the other actions in the PROMETHEE ranking. They have cost 47, number of types of elements 4, number of elements 17 and total of changes 6914. U12 is on top of U6 but they are very close to each other. U6 and U12 is incomparable with the U8, because they have a worse score on $\Phi+$, but still above 0,4. U16, U19 and U15 have the worst position in the ranking with $\Phi+$ below 0.3 (Table 2).

In diagram A in Fig. 7, we see that U2, U9, U14 and U19 are clearly the cheapest options as it project completely to the left side. U15 and U18 the second best choices with respect to cost. They are very close to each other. U6 and U12 are very close to each other and are the most expensive options. This information is of course highly dependent on the localization on the GAIA plane. For lower level one can expect more distortions with respect to actual evaluations. U4 is the best one in number of types of elements. U8 is the best option on sum of changes and number of elements and not so bad on number of type of elements but it is weak on cost. If our determining criterion will be sum of changes, with weight 70% (Cost variant III), we should chose U8 with cost equal to 47, the number of types of elements equal to 4, number of elements 17, total number of changes equal to 6914. If our determining criterion will be cost with weight 70% (Cost variant II), we should chose U2 with low cost equal to 2, one type of elements, two elements used and total 871 changes. In second variant

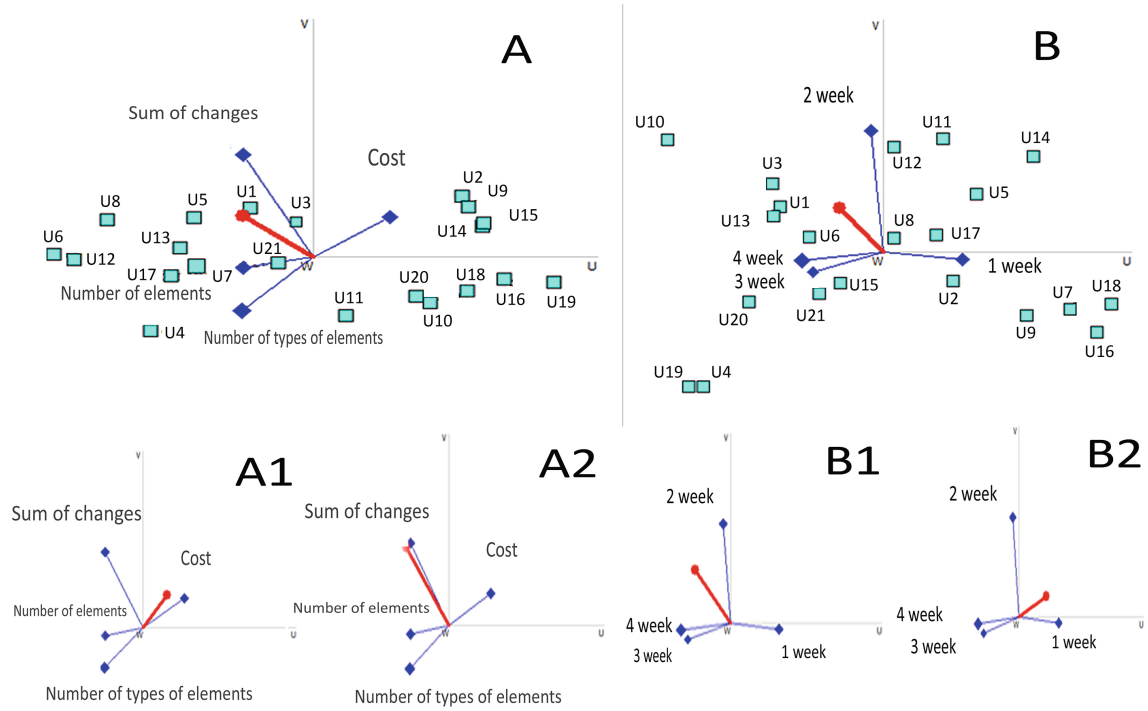


Fig. 7. GAIA analysis diagrams from scenario I cost - A, A1, A2 and scenario II dynamics B, B1, B2.

criteria have different weight: cost = 70%, number of types of elements = 10%, number of elements = 10% and the total of changes = 10%. Results for Cost variant II show the ranking of the actions according to $\Phi+$ with content update number 2 (U2) with the best result, followed by others. Analyzing data from a different side the ranking according to $\Phi-$: U2 is still on top, but it is followed by U9 and U14. We can conclude that U2 is preferred to all the other actions in the PROMETHEE ranking. They have cost equal to 2, number of types of elements equal to 1, number of elements equal to 2 and total of changes at the level of 871. U9 is on top of U14 but they are very close to each other. U9 and U14 is incomparable with the U2, because they has a worse score on $\Phi+$, but above 0.64 updates U6 and U17 have the worst position in the ranking with $\Phi+$ below 0.3.

In third variant criteria have different weights: cost = 10%, number of types of elements = 10%, number of elements = 10%, total number of changes = 70%. Results tab for cost variant III shows the ranking of the actions according to $\Phi+$: content update number 8 (U8) with best result, followed by other. Results shows the ranking according to $\Phi-$: U8 is still on top, but it is followed by U6 and U12. We can conclude that U8 is preferred to all the other actions in the PROMETHEE ranking.

U8 is on top of U6 but they are very close to each other. U6 and U12 is incomparable with the U8, because they has a worse score on $\Phi+$, but above 0.4. U18, U16 and U19 have the worst position in the ranking with $\Phi+$ below 0.3.

Table 2. Scenario I: cost variants for three variants: cost variant I, cost variant II, cost variant III with MCDA analysis

Scenario I	Cost variant I				Cost variant II				Cost variant III			
Rank	Update	Φ	$\Phi+$	$\Phi-$	Update	Φ	$\Phi+$	$\Phi-$	Update	Φ	$\Phi+$	$\Phi-$
1	U8	0,46	0,70	0,24	U2	0,47	0,66	0,19	U8	0,79	0,88	0,09
2	U6	0,43	0,69	0,26	U9	0,46	0,65	0,19	U6	0,71	0,85	0,14
3	U12	0,40	0,68	0,28	U14	0,44	0,64	0,20	U12	0,64	0,81	0,17
4	U7	0,30	0,63	0,33	U19	0,34	0,60	0,26	U5	0,51	0,74	0,23
5	U4	0,23	0,59	0,36	U18	0,21	0,57	0,36	U1	0,45	0,71	0,26
6	U1	0,23	0,58	0,35	U15	0,20	0,56	0,36	U13	0,37	0,67	0,30
7	U5	0,23	0,58	0,35	U1	0,09	0,53	0,44	U7	0,36	0,67	0,31
8	U13	0,18	0,55	0,38	U10	0,08	0,53	0,45	U17	0,23	0,60	0,38
9	U17	0,11	0,53	0,41	U16	0,05	0,50	0,46	U4	0,09	0,54	0,45
10	U21	0,08	0,50	0,43	U11	0,04	0,51	0,47	U21	0,09	0,53	0,44
11	U11	−0,06	0,44	0,50	U20	0,02	0,50	0,48	U3	0,08	0,52	0,44
12	U3	−0,10	0,40	0,50	U7	0,00	0,49	0,49	U2	−0,10	0,42	0,52
13	U2	−0,10	0,36	0,46	U21	−0,03	0,47	0,50	U9	−0,17	0,38	0,55
14	U9	−0,13	0,35	0,48	U8	−0,12	0,43	0,55	U14	−0,31	0,31	0,62
15	U14	−0,18	0,33	0,50	U5	−0,15	0,41	0,56	U15	−0,31	0,32	0,63
16	U10	−0,26	0,34	0,60	U3	−0,22	0,37	0,59	U11	−0,33	0,33	0,65
17	U20	−0,26	0,34	0,60	U13	−0,29	0,34	0,63	U20	−0,47	0,26	0,72
18	U18	−0,30	0,30	0,60	U4	−0,33	0,33	0,66	U10	−0,53	0,23	0,75
19	U15	−0,33	0,28	0,60	U12	−0,38	0,30	0,68	U18	−0,66	0,15	0,81
20	U19	−0,43	0,21	0,64	U6	−0,43	0,28	0,71	U16	−0,68	0,14	0,82
21	U16	−0,49	0,20	0,69	U17	−0,44	0,27	0,71	U19	−0,77	0,09	0,86

In second scenario user engagement and dynamics in analyzed time periods was taken into account. It represents situations when decision maker is interested in specific results, for example high usage in short time after content update, or longer product life span. Four criterias were used at this stage of analysis: dynamics of the first week as percentage of total usage in all periods, dynamics in the second week, dynamics in the third week and dynamics of the fourth last analysed week.

Results for used scenarios in the linear variant shows the ranking of the actions according to $\Phi+$: content update number ten (U10) is on top, followed by other. Analyzing data from a different side the ranking according to $\Phi-$: U10 is still on top, but it is followed by U13 and exequo U1 and U3.

We can conclude that U10 is preferred to all other update in PROMETHEE ranking with $\Phi+$ equal to 0.6875, and $\Phi-$ equal to 0.2875. Overall score Φ is 0.4. It is almost a double advantage over the previous one (U20). They have dynamics of the first week in percent 1%, dynamics of the second week as a percentage 42%, dynamics of the third week as a percentage 18%, dynamics of the fourth last week as a percentage 39%. U1 and U3 they are very close to each other. They have the same scores in Φ (0.1875), $\Phi+$ (0.55) and $\Phi-$ (0.3625). U7, U9, U16 and U18 have the worst position in the ranking with $\Phi+$ values within limits −0.2375 to −0.4 (Table 3).

Table 3. Scenario II: dynamics for three variants: linear, reversed geometric and Gaussian with MCDA analysis

Scenario II	Linear				Reversed geometric				Gaussian			
Rank	Update	Φ	$\Phi+$	$\Phi-$	Update	Φ	$\Phi+$	$\Phi-$	Update	Φ	$\Phi+$	$\Phi-$
1	U10	0,40	0,69	0,29	U14	0,41	0,69	0,28	U10	0,64	0,80	0,16
2	U13	0,20	0,55	0,35	U18	0,35	0,65	0,30	U3	0,45	0,69	0,24
3	U1	0,19	0,55	0,36	U16	0,33	0,64	0,31	U13	0,28	0,58	0,31
4	U3	0,19	0,55	0,36	U5	0,31	0,62	0,31	U12	0,27	0,58	0,31
5	U6	0,16	0,51	0,35	U7	0,28	0,63	0,34	U1	0,26	0,58	0,33
6	U21	0,09	0,48	0,39	U9	0,25	0,61	0,36	U20	0,18	0,57	0,39
7	U20	0,06	0,51	0,45	U2	0,22	0,57	0,35	U11	0,16	0,54	0,38
8	U8	0,06	0,48	0,41	U11	0,21	0,57	0,37	U21	0,14	0,51	0,37
9	U12	0,06	0,46	0,40	U17	0,18	0,58	0,40	U6	0,13	0,49	0,37
10	U11	0,05	0,48	0,43	U8	0,11	0,52	0,41	U15	0,07	0,49	0,43
11	U15	0,05	0,48	0,43	U12	0,05	0,46	0,41	U4	0,02	0,50	0,48
12	U17	0,00	0,48	0,48	U6	−0,06	0,41	0,47	U8	0,01	0,45	0,44
13	U4	0,00	0,49	0,49	U15	−0,06	0,42	0,48	U5	−0,02	0,45	0,47
14	U5	0,00	0,45	0,45	U21	−0,08	0,39	0,47	U19	−0,02	0,48	0,50
15	U19	−0,01	0,49	0,50	U1	−0,16	0,39	0,55	U14	−0,06	0,43	0,49
16	U14	−0,06	0,44	0,50	U13	−0,16	0,38	0,54	U17	−0,06	0,43	0,49
17	U2	−0,06	0,40	0,46	U3	−0,23	0,36	0,58	U2	−0,12	0,37	0,49
18	U9	−0,24	0,34	0,58	U10	−0,34	0,33	0,67	U9	−0,47	0,23	0,70
19	U7	−0,34	0,30	0,64	U20	−0,38	0,28	0,67	U7	−0,56	0,18	0,74
20	U16	−0,40	0,28	0,68	U4	−0,57	0,21	0,78	U16	−0,64	0,16	0,80
21	U18	−0,40	0,28	0,68	U19	−0,64	0,18	0,82	U18	−0,64	0,16	0,80

In diagram B in Fig. 7 in two out of three variants (Linear and Gaussian) except reversed geometric variant we could see that U10 was the best options. In Week I we could see the highest growth rate for U16 and U18 (91%). The lowest growth rate have, our best option U10 (1%). Next options with lowest rate is U19 and U4 with 15% and 18%. Then next have dynamics over 45% and even higher. Average of dynamics in Week I is 57.71%. The case is different in the case of the dynamics of the Week II U10 has rate of dynamics 42%. It was the best rate of all options. In Week II U10 has rate 18% with (average for all is equal to 16.43%). Then the last period had dynamics of 40%, which was again the highest value. In linear and gaussian options, U10 is the most desirable option.

In reversed geometric variant the best is U14. In Week I value is equal to 73%, in Week II is equal to 20% in Week II 5% and Week IV 3%. Φ is equal to 0.4075, $\Phi+$ equal to 0.69 and $\Phi-$ equal to 0.2825. These are the best values, but slightly relative to the next three options: U18 (Week I 91%, Week II 7%, Week III 1%, Week IV 0%) and U16 (Week I 91%, Week II 4%, Week III 4%, Week IV 1%). Both options showed a similar relationship each week. In revered geometric variant U14 is the best option.

Performed analysis shows how evaluation of results from content updates is dependent on preferences of decision maker and strategical goals. Different strategies can be considered as successfully when main target is high coverage with.

6 Conclusions

The increased importance of digital environments and the role of virtual goods in online business models is creating the need for new analytical tools and methods. Phenomena typical to offline markets are also often observed within electronic systems and are related to product life cycles, consumer habituation and strategies of new products development. The presented research shows how content update strategies can affect user engagement and the life span of virtual products. High diversification of products within a single content drop influences user interest and the products' propagation within the system. Other factors analyzed include the number of elements within a single content drop and the dynamics of product usage after introduction. The proposed conceptual framework based on multi-criterial model makes an evaluation of the used strategies possible. Two main approaches were discussed based on implementation costs and usage dynamics. Future work will focus on a more detailed analysis of propagation within social networks and use behavior prediction based on earlier behaviors.

Acknowledgments. The work was supported by the National Science Centre of Poland, the decision no. 2017/27/B/HS4/01216.

References

1. Alha, K., Koskinen, E., Paavilainen, J., Hamari, J.: Critical acclaim and commercial success in mobile free-to-play games. In: DiGRA/FDG (2016)
2. Alha, K., Koskinen, E., Paavilainen, J., Hamari, J., Kinnunen, J.: Free-to-play games: professionals' perspectives. In: Proceedings of Nordic DiGRA 2014 (2014)
3. Bainbridge, W.S.: The scientific research potential of virtual worlds. Science **317**(5837), 472–476 (2007)
4. Bakshy, E., Karrer, B., Adamic, L.A.: Social influence and the diffusion of user-created content. In: Proceedings of the 10th ACM Conference on Electronic Commerce, pp. 325–334. ACM (2009)
5. Brans, J.P., Mareschal, B.: The PROMCALC & GAIA decision support system for multicriteria decision aid. Decis. Support Syst. **12**(4–5), 297–310 (1994)
6. Brans, J.P., Vincke, P., Mareschal, B.: How to select and how to rank projects: the Promethee method. Eur. J. Oper. Res. **24**(2), 228–238 (1986)
7. Castronova, E.: Virtual worlds: a first-hand account of market and society on the Cyberian frontier (2001)
8. Chiu, H.C., Hsieh, Y.C., Kao, Y.H., Lee, M.: The determinants of email receivers' disseminating behaviors on the internet. J. Advertising Res. **47**(4), 524–534 (2007)
9. Hamari, J.: Why do people buy virtual goods? Attitude toward virtual good purchases versus game enjoyment. Int. J. Inf. Manage. **35**(3), 299–308 (2015)

10. Hamari, J., Järvinen, A.: Building customer relationship through game mechanics in social games. In: Business, Technological, and Social Dimensions of Computer Games: Multidisciplinary Developments, pp. 348–365. IGI Global (2011)
11. Hamari, J., Keronen, L.: Why do people buy virtual goods? A literature review. In: 2016 49th Hawaii International Conference on System Sciences (HICSS), pp. 1358–1367. IEEE (2016)
12. Hamari, J., Lehdonvirta, V.: Game design as marketing: how game mechanics create demand for virtual goods (2010)
13. Hanner, N., Zarnekow, R.: Purchasing behavior in free to play games: concepts and empirical validation. In: 2015 48th Hawaii International Conference on System Sciences (HICSS), pp. 3326–3335. IEEE (2015)
14. Jankowski, J., Bródka, P., Hamari, J.: A picture is worth a thousand words: an empirical study on the influence of content visibility on diffusion processes within a virtual world. Behav. Inf. Technol. **35**(11), 926–945 (2016)
15. Jankowski, J., Michalski, R., Bródka, P.: A multilayer network dataset of interaction and influence spreading in a virtual world. Sci. Data **4**, 170144 (2017)
16. Jankowski, J., Michalski, R., Kazienko, P.: The multidimensional study of viral campaigns as branching processes. In: Aberer, K., Flache, A., Jager, W., Liu, L., Tang, J., Guéret, C. (eds.) SocInfo 2012. LNCS, vol. 7710, pp. 462–474. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-35386-4_34
17. Jankowski, J., Michalski, R., Kazienko, P.: Compensatory seeding in networks with varying avaliability of nodes. In: 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 1242–1249. IEEE (2013)
18. Kaplan, A.M., Haenlein, M.: Consumer use and business potential of virtual worlds: the case of second life. Int. J. Media Manage. **11**(3–4), 93–101 (2009)
19. Kwong, J.A.: Getting the goods on virtual items: a fresh look at transactions in multi-user online environments. Wm. Mitchell L. Rev. **37**, 1805 (2010)
20. Lu, H.-P., Wang, S.-m.: The role of internet addiction in online game loyalty: an exploratory study. Internet Res **18**(5), 499–519 (2008)
21. Mareschal, B., Brans, J.P., Vincke, P., et al.: Promethee: a new family of outranking methods in multicriteria analysis. ULB-Universite Libre de Bruxelles, Technical report (1984)
22. McPherson, M., Smith-Lovin, L., Cook, J.M.: Birds of a feather: homophily in social networks. Annu. Rev. Sociol. **27**(1), 415–444 (2001)
23. Wathieu, L.: Consumer habituation. Manage. Sci. **50**(5), 587–596 (2004)
24. Wohn, D.Y.: Spending real money: purchasing patterns of virtual goods in an online social game. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 3359–3368. ACM (2014)

A2.

Bortko, K., Pazura, P., Hamari, J., Bartków, P., Jankowski, J. (2019). From the Hands of an Early Adopter's Avatar to Virtual Junkyards: Analysis of Virtual Goods' Lifetime Survival. Applied Sciences, 9(7), 1268.

applied sciences

MDPI

Article

From the Hands of an Early Adopter's Avatar to Virtual Junkyards: Analysis of Virtual Goods' Lifetime Survival

Kamil Bortko [1], Patryk Pazura [1], Juho Hamari [2,3], Piotr Bartków [1] and Jarosław Jankowski [1,*]

1 Faculty of Computer Science and Information Technology, West Pomeranian University of Technology, 71-550 Szczecin, Poland; kbortko@wi.zut.edu.pl (K.B.); patryk.pazura@gmail.com (P.P.); pbartkow@wi.zut.edu.pl (P.B.)
2 Gamification Group, Faculty of Information Technology and Communication Sciences, Tampere University, 33100 Tampere, Finland; juho.hamari@tut.fi
3 Gamification Group, Faculty of Humanities, University of Turku, 20500 Turku, Finland
* Correspondence: jjankowski@wi.zut.edu.pl

Received: 21 December 2018; Accepted: 18 March 2019; Published: 27 March 2019

check for updates

Abstract: One of the major questions in the study of economics, logistics, and business forecasting is the measurement and prediction of value creation, distribution, and lifetime in the form of goods. In "real" economies, a perfect model for the circulation of goods is impossible. However, virtual realities and economies pose a new frontier for the broad study of economics, since every good and transaction can be accurately tracked. Therefore, models that predict goods' circulation can be tested and confirmed before their introduction to "real life" and other scenarios. The present study is focused on the characteristics of early-stage adopters for virtual goods, and how they predict the lifespan of the goods. We employ machine learning and decision trees as the basis of our prediction models. Results provide evidence that the prediction of the lifespan of virtual objects is possible based just on data from early holders of those objects. Overall, communication and social activity are the main drivers for the effective propagation of virtual goods, and they are the most expected characteristics of early adopters.

Keywords: virtual goods; product life span; innovation diffusion; virtual world; social networks

---

1. Introduction

Virtual worlds and games have been postulated to provide unprecedented possibilities for research in general [1,2], but especially for the study of economics [3] due to their ability to systematically track every event in that reality, but also due to the possibility of creating controllable environments while having people exhibit natural behaviors.

Perhaps one of the most prominent veins of study related to virtual economies has been the study of consumer behavior related to adopting and purchasing virtual goods in virtual worlds and games [4–7]. This has especially been the case since games and virtual world operators have been the forerunners in implementing the so-called freemium or free-to-play business model ([8–10]), where playing or using the virtual environment is free of charge, but the operator generates revenue through different manifold marketing strategies combining classical sales tactics imbued with platform design that further encourages virtual-goods purchases [11–13].

Virtual goods mostly take up the forms of in-game items related to the theme of the game, such as avatar clothing, gear, vehicles, pets, emoticons, and other customization options [5,14], as well as different types of items related to the recent proliferation of "gamblification", where acquiring virtual goods is increasingly based on gambling-like mechanics, effectively blurring the line between gaming and gambling [15].

The largest vein of research in this continuum has been the investigation into why people purchase virtual goods [4,5] in primary or secondary markets within the virtual world. Popularly, this question was initially motivated by the sheer anecdotal amazement of why people would spend considerable amount of real money on products that "do not exist" [11,16]. However, since the initial combination of hype and disillusionment, virtual and game economies have entered into the realm of everyday consumer-facing services. Studying the question of why people purchase and trade virtual goods has primarily focused on latent psychological factors such as motivations, attitudes, experiences, and belief, and how they predict virtual-goods transactions as well as the internal design of the environment (see, for example, Reference [4] for a review of the area). However, the limitation within this sphere of research is that it can only provide a glimpse of the reasons why users purchase virtual goods as a singular event since it is focused on the consumer rather than the object of consumption and trade—the virtual good itself. Only few studies [17] have taken the initiative in an attempt to map the longer lifespan of virtual goods from their inception to circulation and to their ultimate end, destroyed from the virtual world, forgotten in a user's virtual bag, or existing in an account of a user who has stopped visiting the virtual world.

Additionally, one of the major hurdles in governing and maintaining virtual economies, in addition to increasing consumer demand for virtual goods [11], has been the balancing act between "sources" and "sinks" [18] of virtual goods within a virtual economy. There is no practical or technical reason why any virtual good could not exist in complete abundance within the virtual economy. However, this would create problems both in relation to the meaningfulness of acting within the virtual world due to extreme inflation, which would also effectively void any need for users to purchase or trade virtual goods. Therefore, the lifetime management of virtual goods is of vital importance for any virtual-economy operator (see References [6,11,18]). Some of the methods in the game-operator palette have been, for example, contrived durability and planned obsolescence of virtual goods (see, for example, Reference [19]).

Game developers are confronted with issues identified with the ideal recurrence of virtual-product updates, their volumes, and intensity, with an emphasis on ceaseless development [20]. Reduced recurrence of updates can result in user churn, while the consistent improvement of new content increases operational expenses. From another perspective, users may have a constrained capacity for digital content used when content is updated as often as possible. This might be regarded as unwise budget allocation when content production is fundamentally higher than demand. The life expectancy of web-based gaming items is generally shorter than that of traditional items, and users always expect system updates and new content [21,22]. Another issue is the habituation impact resulting from the short life expectancy of virtual products, and the limited time in which the item can attract online users. This opens up new research directions since, so far, it has principally been researched for traditional markets [23].

To address this research problem, the present study is focused on the characteristics of early-stage adopters of virtual goods and how they predict the lifespan of the goods. Rogers [24] treats 2.5% of users as innovators, 13.5% of users as early adopters, 34% as an early majority, and 34% and 16% as the late majority and laggards, respectively. This research shows how characteristics of early-stage adopters affect user engagement and product lifespan. The main contributions include the identification of the role of early adopters of virtual goods for product lifespan, and building a predictive model for product life with the use of data.

The empirical study is followed by analysis based on survival prediction models and identification of the role of the characteristics of early-stage adopters for product lifespan. Decision trees showed the ability to predict product lifespan with the use of product-adopter characteristics. The rest of the paper is organized as follows. The Methodology section contains the conceptual framework, dataset description, and methodological background. The Results section includes descriptive statistics and results from the lifespan models based on user characteristics. This is followed by results from product classification in terms of their lifespan and user characteristics with an accuracy higher than 80%. The study is concluded in the final section.

2. Methodology

2.1. Research Questions and Study Design

The presented study assumes the ability of virtual-product survival prediction with user attributes, especially those interested in the product at different stages of the product lifecycle. This research is based on the conceptual framework presented in Figure 1. A set of virtual products, Pi, was introduced to the audience of a social platform. Behaviors related to user engagement and products usage were collected. The node position within the example social network is represented by node size. Small circles were used for low degree nodes with one connection through medium sizes up to biggest ones for nodes with four connections. In general, user characteristics can represent various attributes related to network centrality and activity within the system like communication frequency and intensity of platform usage. They create parameters space with m distinguished variables assigned to each user in the form of vector $V = [V_1, V_2, ..., V_m]$. Users adopted to each product can be divided into five adoption groups with 2.5% of users interested in product distinguished as innovators, next 13.5% classified as early adopters, 34% as early majority, 34% of late majority, and users adopting to product at the end (laggards) as 16% of all adopters.

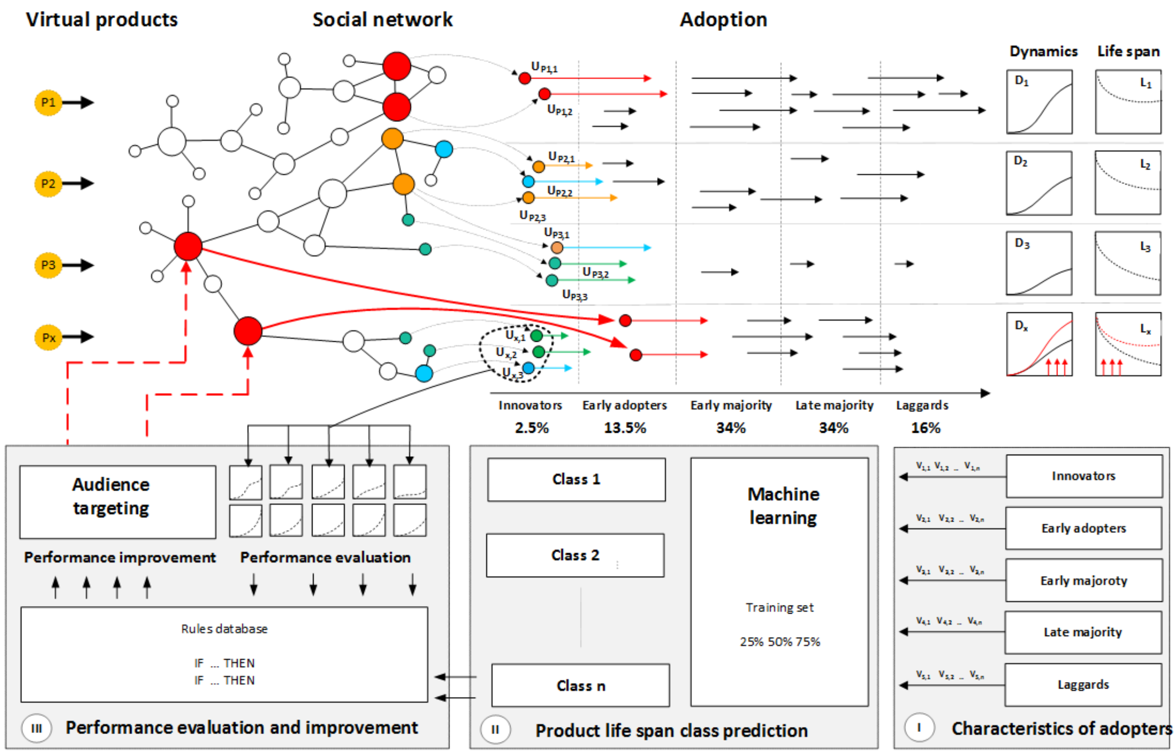


Figure 1. (**I**) Analytical system integration with the platform with the ability to detect the characteristics of users engaged in a new product, and the stages when adoption takes place; (**II**) product classification according to survival time and audience characteristics; (**III**) monitoring the performance of new products, predicting their usage, and additional audience targeting.

One of the research questions is whether innovator and early-adopter characteristics can affect product lifespan. It would be possible to identify the characteristics of initial users and, in cases of low-performance prediction, the interest of users with central network positions could be increased by sample delivery, trial accounts, or other incentives. As a result, followers and late adopters would be influenced and motivated for new-product usage.

Three exemplary possible scenarios are presented. In Scenario 1, Product P1 is introduced, and two users, namely, $UP_{1,1}$ and $UP_{1,2}$, with high positions represented by red nodes adopt the product as innovators. They are followed by users adopted at several stages of the product lifecycle, and it is considered a successful product launch, boosted by top-user engagement. The campaign is characterized by high dynamics D_1 and long product lifespan L_1. Scenario 2 for Product P2 assumes three innovators, $UP_{2,1}$, $UP_{2,2}$, and $UP_{2,3}$, characterized by medium metrics within a social network, and this is represented by orange and blue nodes. They build the interest of other users in the launched product, and overall campaign evaluation results in medium dynamics D_2 and product lifespan L_2. Scenario 3 assigned to Product P3 is based on the interest of innovators with the lowest network positions. It results in dynamics D_3 and lifespan L_3. In an analytical system, historical product data are used to analyze the influence of user characteristics, especially innovators and early adopters, on the product's lifespan and engagement among other users. This is based on three stages of data processing. In Stage (I), the characteristics of adopters from all groups are measured. In Stage (II), classification is performed to build class descriptors of users who are characteristic for a product with different survival time. Results are used to build a knowledge base and rules set for further use within the system and future product evaluation. In the next stage, new product Px is launched and introduced to the system. Innovators and early adopters were monitored, and prediction of the product lifespan was performed. If the product that is assigned to the class with possible low lifespan, actions to improve performance can be implemented by the selection of users with high network positions to build interest in the new product, denoted as red arrows. The main goal is to increase the dynamics of product consumption D_x and its lifespan L_x. In practice, it can be performed by product samples, trial accounts, or various other forms of incentives.

2.2. Dataset Description and Participants

The experimental study is based on data from the virtual world and the use of avatars within the platform [25,26]. The introduced dataset covers information from 195 items included in the form of user avatars. Items are utilized in the virtual-world platform providing various forms of entertainment and chat functions. Graphical symbols represent users who all have the chance to participate in the life of the online network, with 850,000 accounts initiated. Clients interact in the space of public graphical rooms that are related to various themes. They can configure and supply their private rooms and also utilize web-based games and unique entertainment alternatives.

The fundamental functions of the service are related to chatting, meeting new people, communication, and creating social relations. Other features include clothes and virtual products, styles, avatars, and a decorative element. New-product information can be distributed through private messages, sent through the use of an internal communication system. The analytical module concentrated on new items and this enhanced monitoring of content distribution and collecting information related to data-dissemination procedures. Clients accessed various amounts of functions that are commonly available, and also paid for premium services, which provide more potential outcomes. Virtual products appear in the form of products equivalent to real goods, special effects for avatars, or avatars themselves. Account extensions used within the system had different characteristics and purposes. For example, animations, flashing elements, and active objects handled by avatars were used.

While innovation-diffusion theory emphasizes the role of innovation characteristics, it was important to take into account objects with similar characteristics to minimize the impact of individual product features and the level of innovation. This led to analysis of comparable static-avatar elements with similar characteristics without special effects usually attracting more attention than static objects.

2.3. Survival Analysis Methods for Measuring Product Lifespan

The presented study uses survival analysis to analyze the expected time duration when interest in new products exists, which represents the product lifespan. In the field of survival analysis length of time taken is referred to event time [27]—product usage time in our case. Survival analysis was originally developed in the medical field, as a means of analyzing the time between medical intervention and death. Over the past few decades, the field was expanded to include other events as well as events that occur multiple times for a given individual [28].

Survival analysis has wide applications in the field of marketing, including customer-relationship management (CRM), marketing-campaign management, and trigger-event management [29]. If we denote the time taken for an event to occur as T, we can construct a frequency histogram and model a series of events as a function of time. The probability distribution function for T can be denoted by $f(t)$. The cumulative distribution function can be denoted by $F(t)$. This provides the following equation:

$$F(t) = p(T \leq t)$$

Using the above approach, we can represent survival as a function of time $S(t)$ such that: for $t = 0$, $S(t) = 1$ for the specific time that a failure occurs, the value of $S(t)$ is zero [30]. In some cases, the time to failure will not be observable and only partial observation will be possible. In this case, we consider a specific 'censoring time' c. The survival function is then denoted as:

$$S(t) = P(T > t) = 1 - F(t)$$

Instantaneous hazard or conditional failure rate is the instantaneous rate at which a randomly selected individual—who is known to be alive at time $(t-1)$ and will die at time t [31]. Mathematically, instantaneous hazard is equal to the number of failures between time t and time $t + \Delta(t)$, divided by the size of the population at risk (at time t), divided by $\Delta(t)$. This gives us the proportion of the present population at time t that fail, per unit of time, represented by the equation:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t < T \leq t + \Delta(t) | T > t)}{\Delta(t)} = \frac{f(t)}{S(t)}$$

Widely used, the Kaplan–Meier method is used to estimate time-related events [27]. Most commonly, it is used in biostatistics to analyze death as outcome. However, in more recent years, the technique has seen adoption in the fields of social sciences and industrial statistics. For example, in economics, we might measure how long people tend to remain unemployed after being let go by an employer; in engineering, we might measure how long a certain mechanical component tends to last before mechanical failure takes place. The survival function is theoretically a smooth curve, but it can be estimated using the Kaplan–Meier (KM) curve. Plotting the Kaplan–Meier estimate entails a series of horizontal steps of declining magnitude that, for a sufficiently large sample approach, estimate the true survival function for the given population. When applying this approach, survival-function value between successive sampled observations is presumed constant [32]. An important advantage of the Kaplan–Meier curve is its ability to take into account censored data loss within the sample before the final outcome is observed. In cases where no truncation or censoring occurs, the Kaplan–Meier curve is equivalent to empirical distribution [33,34].

As mentioned, survival analysis has wide applications for marketing, including CRM, marketing-campaign management, and trigger-event management [29]. Depending on the business setting, e.g., contractual versus noncontractual, different techniques can be applied [29]. For example, a goal might be to analyze the performance of a marketing campaign (while in progress), and how different customer features affect its performance. In this case, recurrent survival analysis techniques are used and the hazard function models the tendency of customers to buy a given product [35,36]

Survival analysis also has wide applications in the field of customer-behavior analysis. Among other things, it has been used to make predictions regarding customer retention in the banking [37] and insurance industries [38], credit scoring (with macroeconomic variables) [39], credit-granting decisions [40], and risk predictions of small-business loans [41].

Aside from customer behavior, survival analysis has been used to make predictions regarding the survival of online companies [42], as well as the duration of open-source projects [43]. Similarly, product survival in given markets was analyzed with network effects based on product compatibility [44].

The advent of digital marketing has provided additional streams of rich behavior data and subsequently new fertile ground for the application of survival analysis. With these data, survival analysis can be used to make predictions regarding the survival of music albums and distribution [45], the survival of mobile applications [46], as well as e-commerce recommendations to users [47].

For social platforms, survival analysis has been applied to triadic relationships within a social network [48], as well as participation in online entertainment communities with the use of entertainment and community-based mechanisms [49]. Player activity in online games provides valuable data for analysis, with a focus on game hours, subscription cancellations [50], and the adjustment of game parameters. In this context, a primary goal is to achieve the optimal user experience in terms of game speed and design [51].

Another area that is being explored is churn prediction in mobile games using survival ensembles [52] and player-motivation theories [53]. While game-time survival analysis can be used as a predictor of user engagement, it can also provide knowledge regarding factors that affect gameplay duration [54]. Similarly, it can provide insight in how player activity and popularity affects retention within games [55]. It can also be used to uncover predictors of game-session length, such as character level or age within the game [56]. The ability to quantify user satisfaction provides greater ability to target user needs [57].

2.4. Classification Methods Used for Product-Lifespan Prediction

Decision-making involves several approaches, including decision-tree classifiers [58]. Making a decision based on the structure of a decision tree allows complex decisions to be broken into a few small ones to deeply understand a problem. Decision trees are pervasive in a variety of real-world applications, including and not limited to medicinal research [59], biology, credit risk assessment, financial-market modeling, electrical engineering, quality control, biology, chemistry and so on. The evolution of web applications and social media resulted new areas of decision support and data analytics focused on user interaction and online behaviors. Decision trees are used for e-commerce, social media, online games, player segmentation, and other areas. Among other areas, applications include decision-tree usage for the future adoption of e-commerce-service predictions [60]. In social media, decision trees are used, for example, to predict the distance between users with Twitter activity data [61] and Twitter message classification with the use of the Classification and Regression Tree (CART) algorithm [62]. This wide area of applications includes online games with a focus on player-segmentation strategies based on self-recognition and game behaviors in the online game world to improve player satisfaction [63]. Integrated data-mining techniques such as association rule discovery, decision trees, and self-organizing map neural networks within the Kano model are used for customer-preference analysis in massively multiplayer online role-playing games [64].

Predicting aspects of playing behavior with the use of supervised learning algorithms is trained on large-scale player-behavior data. Decision-tree learning induces well-performing and informative solutions [65]. Rule databases can be used in a form of rule reasoner in online games for the detection of cheating activities [66], while a case-based reasoning approach can be applied for the purpose of training our system to learn and predict player strategies [67]. Educational games can be improved with decision trees used for the identification of factors affecting user behavior and knowledge acquisition within educational online games [68]. In other applications, decision trees are used for Internet game addiction in adolescents [69] and game-traffic analysis at the transport layer [70].

Clusterization techniques are used for player-behavior segmentation in computer games with the use of K-means and simplex volume maximization clustering [71], and user segmentation is used for retention management in online social games [72]. Integrated data-mining and experiential-marketing techniques can be used to segment online-game customers [73].

Owing to their structure, trees are easy to interpret, and hence result in better insights to problems. Nodes in decision-tree ramify from root nodes, and each node represents a condition related to a single input variable (feature), each branch represents a condition outcome, and each leaf node represents the class label. In this study, we applied CART [74], which is a binary tree. The method is to generate binary-tree-utilized binary-recursive partitioning that divides the dataset into two subsets, as per the minimization of a heterogeneity criterion computed on the resulting subsets. Each division made is based on a single variable, and some variables may not be used at all, while others may be used several times. Each subset is then further split based on independent rules.

Let's take into account decision tree T, with one of its leaves t. T is a mapping that assigns a leaf t to each sample $(X_i^1, \ldots, X_i^p)$, where i is an index for the samples. T can be viewed as a mapping to assign a value $\widehat{Y}_i = T(X_i^1, \ldots, X_i^p)$ to each sample. Let $p(j|t)$ be the proportion of a class j in a leaf t. The Gini index and entropy are the two most popular heterogeneity criteria. The entropy index is:

$$E_t = \sum_j p(j|t) logp(j|t)$$

with, by convention, $xlogx = 0$ when $x = 0$. The Gini Index is an impurity-based criterion that measures divergence between the probability distributions of the target attribute's values [75]. The Gini index is defined as:

$$D_t = \sum_{i \neq j} p(i|t)p(j|t) = 1 - \sum_{i \neq j} p(i|t)^2$$

For the purpose of our research, we followed the formal definitions proposed by Maimon and Rokach [76], with bag algebra in the background [77]. Following the definitions, the training set in typical supervised learning consists of labeled examples in order to form a description that can be used to predict previously unseen examples. Many data descriptions were created, and the most frequently used is the bag instance of a certain bag schema. The bag schema is denoted as $R(A \cup y)$ and provides the description of the attributes and their domains. A indicates the set of input attributes containing n attributes: $A = \{a_1, \ldots, a_i, \ldots, a_n\}$ and y represents the class variable or the target attribute. Attributes appear in one of two forms, nominal or numeric. If attribute a_i is nominal, we denote it by $dom(a_i) = \{v_{i,1}, \ldots v_{i,2}, \ldots, v_{i,|dom(a_i)|}\}$ where $dom(a_i)$ stands for its finite cardinality.

The domain of the target attribute appears in a similar way, $dom(y) = \{c_1, \ldots, c_{|dom(a_i)|}\}$. All possible examples that make up the set are called instance space: $X = dom(a_1) \times dom(a_2) \times \ldots \times dom(x_n)$. The Cartesian product of all input-attribute domains define the instance space.

The Cartesian product of all input-attribute domains and target-attribute domain defines the universal instance space, i.e., $U = X \times dom(y)$. Training consists of a set of tuples. Each tuple is described by a vector of attribute values. The training set is denoted as $S(R) = (\langle x_1, y_1 \rangle, \ldots, \langle x_n, y_n \rangle)$ where $x_q \in X$ and $y_q \in dom(y)$. The algorithm needs these data to learn how to match the input variables with the dependent variables—briefly, how to fit into the algorithm.

The test dataset was used to verify how our algorithm learns from the training data by checking its classification accuracy. We achieved this through matching classified observation with a real-observation class.

3. Results

3.1. Descriptive Statistics

Statistical analysis was based on 195 elements divided into four types of virtual elements, E1, E2, E3, and E4, used within the system representing avatar head, body, legs, and shoes. The data contain the anonymized behavioral patterns of 8139 unique users. The analyzed products were introduced to system users within 21 content updates (CUs).

In order to perform statistical analysis, we used two groups of separate variables related to user activities. Variable abbreviations and their explanation can be found in Table 1.

Table 1. Abbreviations of variables with their short description used in the article.

Short		Explanation of the Variables
Activity Factors	CA	communication activity
	SD	social dynamics
	CP	communication popularity
	SP	social position
	AA	adoption activity
Experience Factors	FR_in	all messages sent by the user until they are changed to unique users
	FR_out	all messages received by the user until changed from unique users
	MSG_in	all messages sent by the user until the change
	MSG_out	all messages received by the user until the change
	FR_total	total amount of FR_in and FR_out
	MSG_total	total amount of MSG_in and MSG_out
	U_log	number of logins before the change
Adoption Group	AG1	innovators
	AG2	innovators + early adopters
	AG3	innovators + early adopters+early majority
	AG4	innovators + early adopters+early majority+late majority
	AG5	innovators + early adopters+early majority+late majority+laggards

The first group includes five variables treated as Activity Factors with the symbols CA–AA. These are, respectively: CA, communication activity represented by an average number of messages received by users adopting the product divided by the number of logins; SD, social dynamics, represented by an average of a number of friends of the product adopter divided by the number of logins; CP, communication popularity, represented by an average number of outgoing messages divided by incoming messages; SP, social position, represented by the average number of received messages divided by the number of incoming messages; and AA, adoption activity, represented by averaging the number of new avatar-element usages divided by the number of logins.

The second group of variables represents Experience Factors related to user activity since account creation, such as MSG_in, the average number of all messages received by the user until the avatar changes; MSG_out, the average number of all messages sent by the user until the change; MSG_total, the average number of total messages sent and received by the user; FR_in—the number of unique friends contacting the user until the avatar change; FR_out—the number of friends contacted before the avatar usage, and FR_total, the average total number of friends.

For each product, users were assigned to Adoption Groups in five classes: innovators, early adopters, early majority, late majority, and laggards, according to time of adoption.

For the purpose of determining the role of used variables, user-related factors were used for the statistical models of survival analysis. We took into account the User Activity and User Experience factors. Initial analysis showed that, for most products, survival time was shorter than one month, and only few of them reached nearly three months. To cover usage periods with more detail, five time periods were taken into account during analysis: one week, two weeks, one month, two months, and three months. One week as the shortest period makes it possible to analyze behavior each day of the week after product launch. Analyzing the statistical significance of predictors that influenced the lifetime dependent variable, we can see that mean CA and AA showed statistical significance of $p < 0.05$ for all periods. The CP variable, on the other hand, is one that has no effect and is not relevant in any given period. Separately analyzing each period, we can see that the periods of one month, two months, and three months showed the significance of the CA, SD, SP and AA variables. Wald's statistics with results presented in Table 2 showed the highest value with CA in the periods of two weeks month and two months. In the three-month period, Wald pointed to the significance of AA. The influence of predictors on the dependent variable over seven days showed significance in CA, SD, and AA. However, in the 14 day period, only two predictors, CA and AA, showed statistical significance, which affected the product's life expectancy. In the next step, Kaplan–Meier (Figures 2–6) survival probability charts for one month with division for user parameters, and the three user groups were analyzed. The diagrams show the emergence of a growing number of increasingly shorter episodes that, at the border, seek the real function of survival. Figure 6 shows a survival model without division into classes as a general model for divisional and nondivisional variables.

The next stages show statistical regression models with division into aggregated groups of adopters, i.e., AG1–AG5. Regarding the explanation of these classes, we can refer to Table 3. Regression analysis was divided into two groups of variables, and product life is a dependent variable. The first group of variables (predictors) include average variable values from CA to AA. The second group include experience-related variables, i.e., MSG_in, MSG_out, FR_in, and FR_out. In Table 3, we can see the statistical-significance parameter (p) and the strength-of-significance factor (f).

Table 2. Survival analysis with five user variables divided into five periods with Wald statistics and statistical significance showed.

Variables	7 Days		14 Days		1 Month		2 Months		3 Months	
	Wald	p	Wald	p	Wald	p	Wald	p	Wald	p
CA	24.90	<0.01	28.90	<0.01	20.16	<0.01	23.73	<0.01	21.61	<0.01
SD	4.99	0.03	2.82	0.09	6.09	<0.01	4.11	0.04	4.49	0.03
CP	0.51	0.47	0.08	0.78	2.00	0.16	1.27	0.26	1.50	0.22
SP	2.30	0.13	1.25	0.26	7.12	0.01	9.71	<0.01	12.37	<0.01
AA	6.09	<0.01	6.18	<0.01	20.06	<0.01	22.45	<0.01	26.58	<0.01

Table 3. Results of regression analysis showing how Activity Factors and Experience Factors are affecting user assignment to adoption group.

Variables	AG1		AG2		AG3		AG4		AG5	
	f	p	f	p	f	p	f	p	f	p
CA	5.07	0.03	4.19	0.04	36.33	<0.01	29.48	<0.01	16.95	<0.01
SD	0.62	0.43	0.10	0.75	7.97	0.01	4.27	0.04	0.39	0.53
CP	0.21	0.65	9.10	<0.01	5.74	0.02	5.36	0.02	8.25	<0.01
SP	0.52	0.47	13.98	<0.01	9.62	<0.01	9.74	<0.01	13.47	<0.01
AA	9.78	<0.01	9.97	<0.01	0.01	0.94	0.23	0.63	0.05	0.82
MSG_in	0.77	0.38	2.99	0.09	14.24	<0.01	7.18	0.01	4.44	0.04
MSG_out	<0.01	0.94	6.07	0.01	18.38	<0.01	9.51	<0.01	7.67	0.01
FR_in	1.22	0.27	6.62	0.01	13.51	<0.01	21.22	<0.01	12.83	<0.01
FR_out	4.49	0.04	15.67	<0.01	30.57	<0.01	36.94	<0.01	29.11	<0.01

The first group of predictors for AG1 showed significance for CA and AA. Average predictor AA was characterized by the strongest impact. For AG2, the case was definitely different. Four of the five predictors, i.e., CA and CP to AA, were significant. The only predictor that did not have statistical significance was average predictor SD. The impact forces of the predictors, especially in SP, were characterized by a strong accent. For AG3 and AG4, regression analysis showed similar significance to AG2 also for four predictors, but in these cases, lack of predictor significance in relation to dependent variables was shown by the mean of AA variables. In both cases, the CA variable strongly affected G5 results, where things were quite different. Statistical significance was only demonstrated in three cases: CA, CP, and SP.

The second group of predictors that affect the dependent variable also showed variability. In AG1, one of the four predictors was statistically significant, namely, FR_out (0.03). The situation looked completely different for AG2. Here, we can clearly see the strength of joining two classes. Significance statistics showed a positive result for up to three predictors, i.e., MSG_out, FR_in, and FR_out. In the case of AG3, AG4, and AG5, statistical significance was shown by 100% of predictors from FR_out, being the one that acts the strongest on the dependent variable.

The next part of analysis was based on an intergroup comparison of user characteristics between products with different survival time. In order to compare individual lifecycles with the Activity and Experience factors, we used the Mann–Whitney U Test. Analysis was presented in four perspectives: analysis of individual user classes, analysis of aggregated user classes based on activity-factor analysis of individual user classes, and analysis of aggregated user classes based on Experience Factors. Periods that we compared with each other are visible in Table A1. By starting division-variable predictor analysis for innovators, we can see the lack of significance of parameters at the first comparison period. In the next two, we can see that predictor CA was significant, which indicates that the periods significantly differed from predictor CA. In the last pair of compared periods, predictors CA, SD, and SP showed the largest differences.

Statistics for innovators show us a tendency for the comparative period to be smaller, in this case, two to three months, so more predictors influenced the differences. Analyzing the four other user classes, we see the opposite relationship. Starting with early adopters, where the differences could be seen in the four predictors in the first two pairs, in the next two the number of differences decreased. In the cases of early-majority, late-majority, and laggard users, significance statistics that point to differences are slowly blurred, as in the case of laggards, where in the last group of period comparisons we see the lack of significance of the given predictor data, which indicates low differences. Based on the aggregated user classes, we can see that the first combination of innovators and early adopters positively affects predictor significance, and this indicates large differences for most of the analyzed pairs (from three to four strongly affecting differences). We can see that the shorter the comparison period is, the smaller the differences are, such as two months versus three months. By analyzing the statistics of nondivisional variables that also include four period pairs, we see that statistics for the innovators themselves did not show any significance. We can see statistical significance at subsequent classes. Analyzing the remaining classes together, we see that differences in individual periods clearly increase. So, for early adopters, when analyzing the last two pairs of periods, statistical significance was less than 0.05, which indicates an increase in differences. By analyzing the last group, laggards, we could see that, in each group of periods, differences are clear and quite significant in each of the period pairs being compared. The same applies to aggregated users. Here, we compare the first period and, only in the case of AG2, in 14 days versus three months, we see the lack of slight differences between predictors. Other groups indicate strong differences, as we can see in Tables A2 and A3.

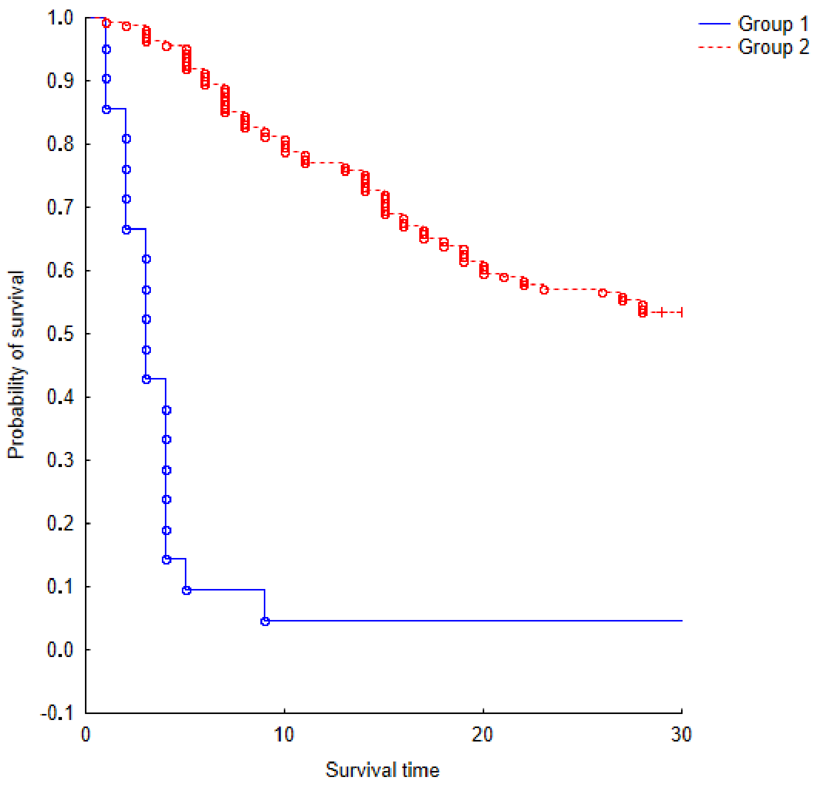


Figure 2. The Kaplan-Meier survival model for two groups of Experience Factors over a period of one month.

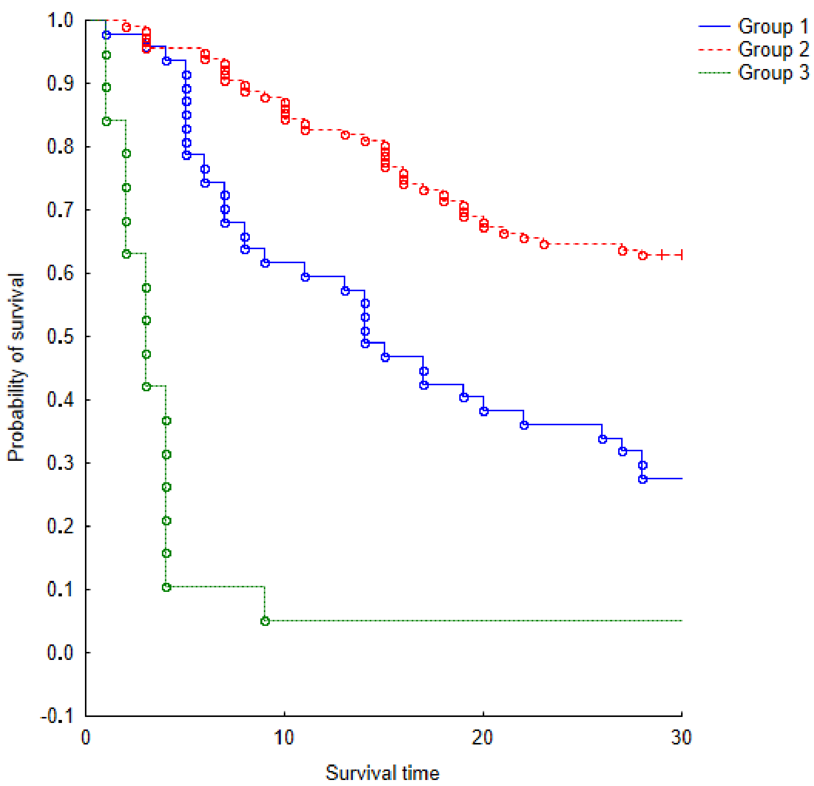


Figure 3. The Kaplan-Meier survival model for three groups of Experience Factors over a period of one month.

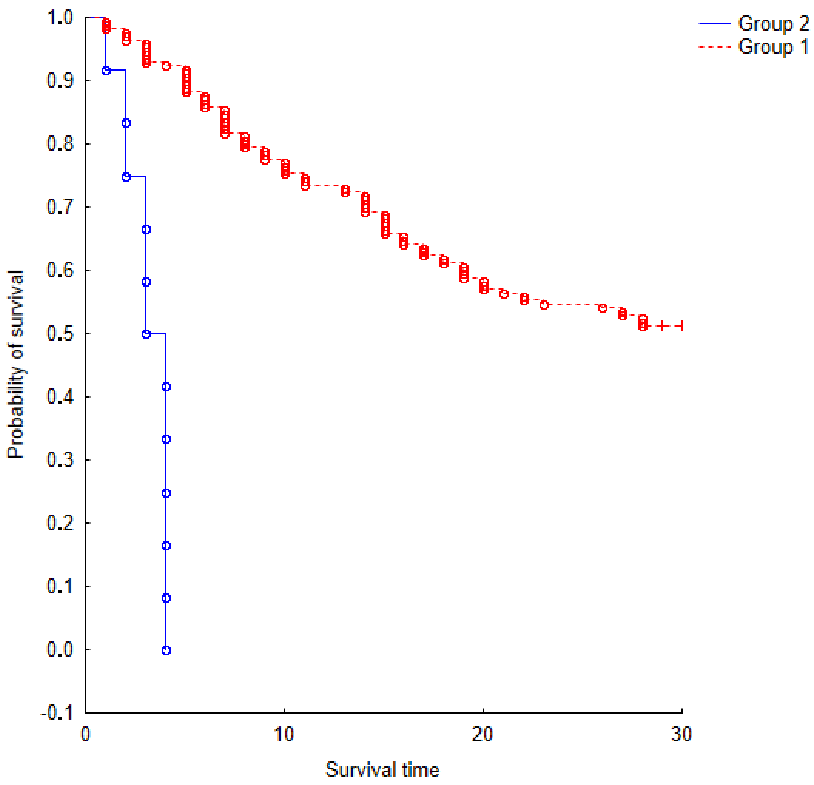


Figure 4. The Kaplan-Meier survival model for two groups of Activity Factors over a period of one month.

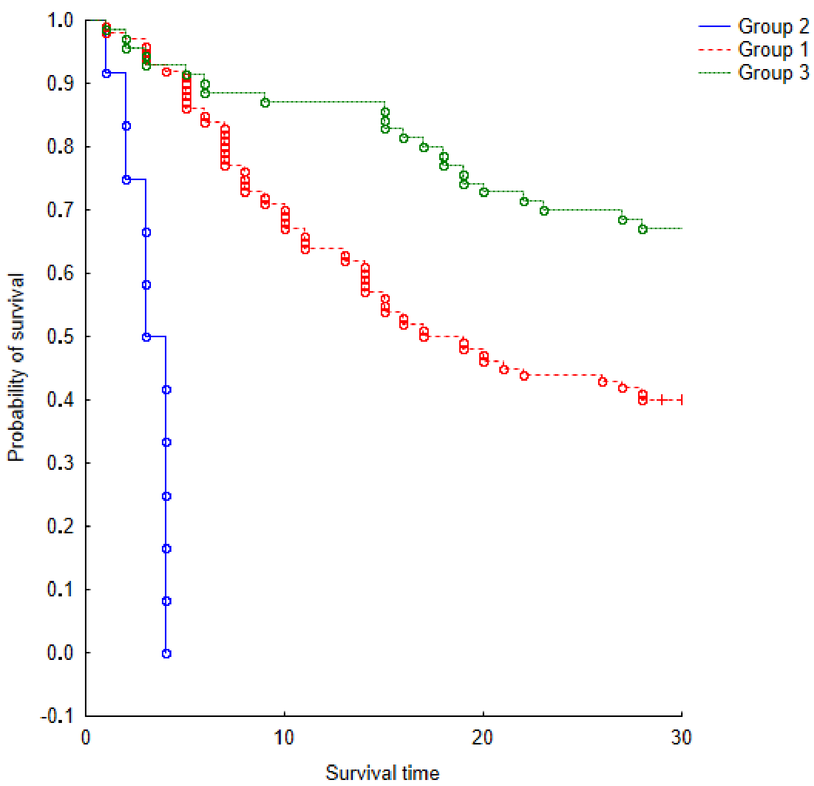


Figure 5. The Kaplan-Meier survival model for three groups of activity of experience over a period of one month.

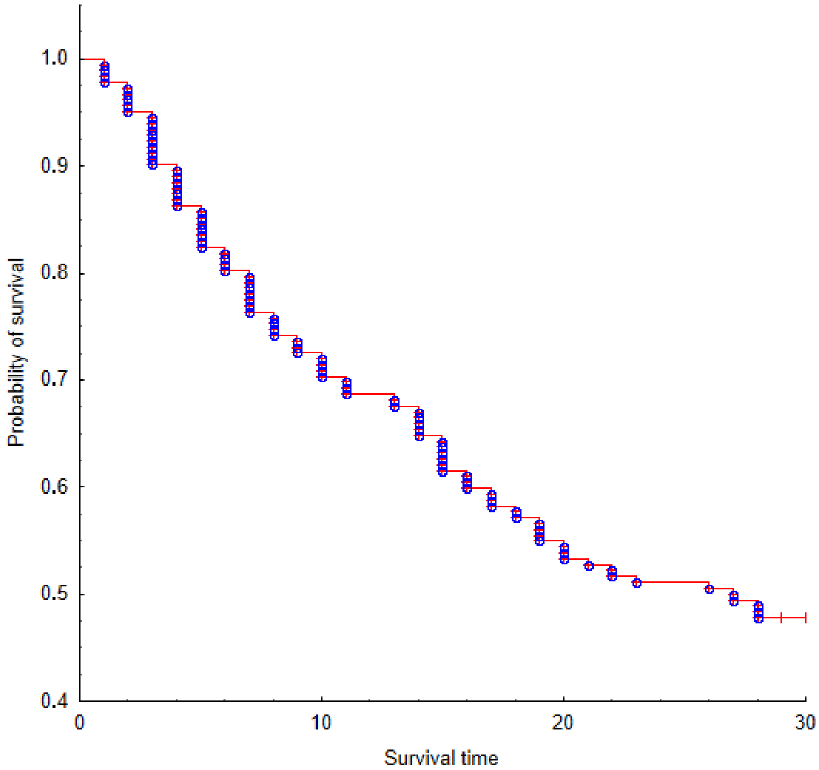


Figure 6. The Kaplan-Meier survival model for one group of divisional and non-divisional variables over a period of one month.

3.2. Survival-Time Prediction with Early-Adopter Characteristics

For survival-time, a prediction dataset containing the usage statistics of 195 newly introduced products was used. Usage statistics for each product are defined by product and user identifiers, and a timestamp representing time when the product (in this case, the avatar) was used by a specific user. For each product, only the first usage per user was taken into account. For each product, data were collected from a newly added product starting from product launch until last product usage. For each analysis, two sets of variables were used based on the User Activity and User Experience factors presented earlier in Table 1.

In Figure 7, we see a high increase in the CP variable for products with seven-day survival with simultaneous small CA values. In other periods, we see density with slight deviations, as in the case of the three-month period, where we see growth in the CA variable; in a 14-day survival period, an increase in the SD variable was observed. As in the previous chart, Figure 8 shows a clear division into survival-period groups. Within seven days, an increase in the AA ratio with a simultaneous drop in SD was visible, which may indicate a drop in interest from users with low SD. In the remaining periods, we can see in Figure 9 a clear decrease in the SD index with a simultaneous increase in CA; this showed that the more users communicate with others, the less likely it is for the product to be accepted.

In the case of this chart, we can clearly see that the fewer users log in, fewer messages are sent to others, and fewer sent to the circle of potential friends. There is a clear decline from that period to the next. In the case of the last graph, Figure 10, we can see density against the FR_out indicator at initial values oscillating at 150–250. Here, however, we also see a decline from period to period. In the initial period, the MSG_in indicator is small, but increases with survival time. However, the last period (three months) oscillates near the first period, which indicates a lower number of messages sent by the users adopting products in that group. Results from all Experience Factors and Activity Factors are presented in Figures A1 and A2 within Appendix A.

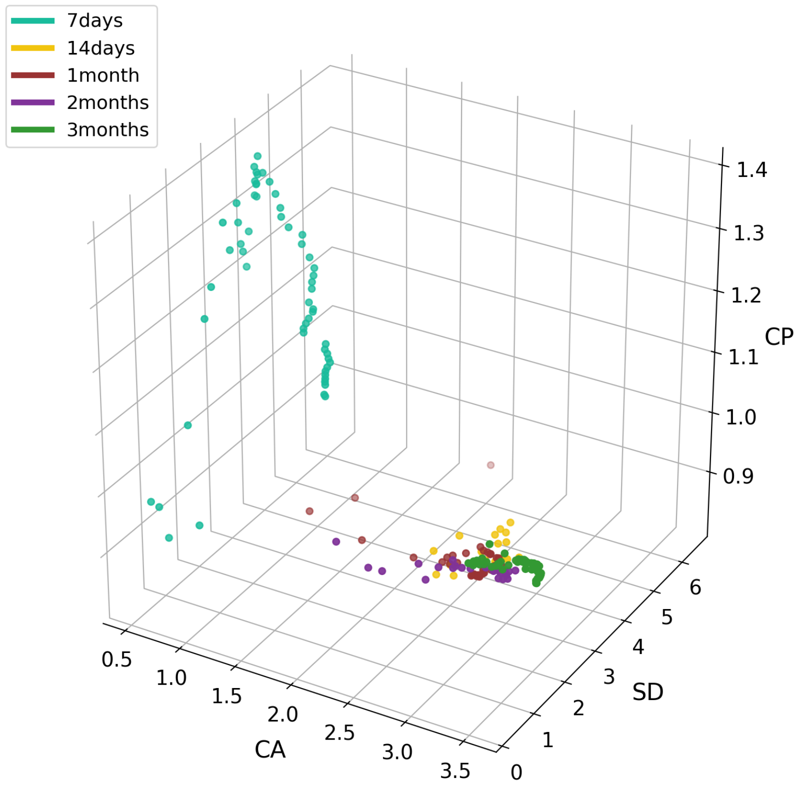


Figure 7. Dependence of objects in classes from CA, SD and CP variables.

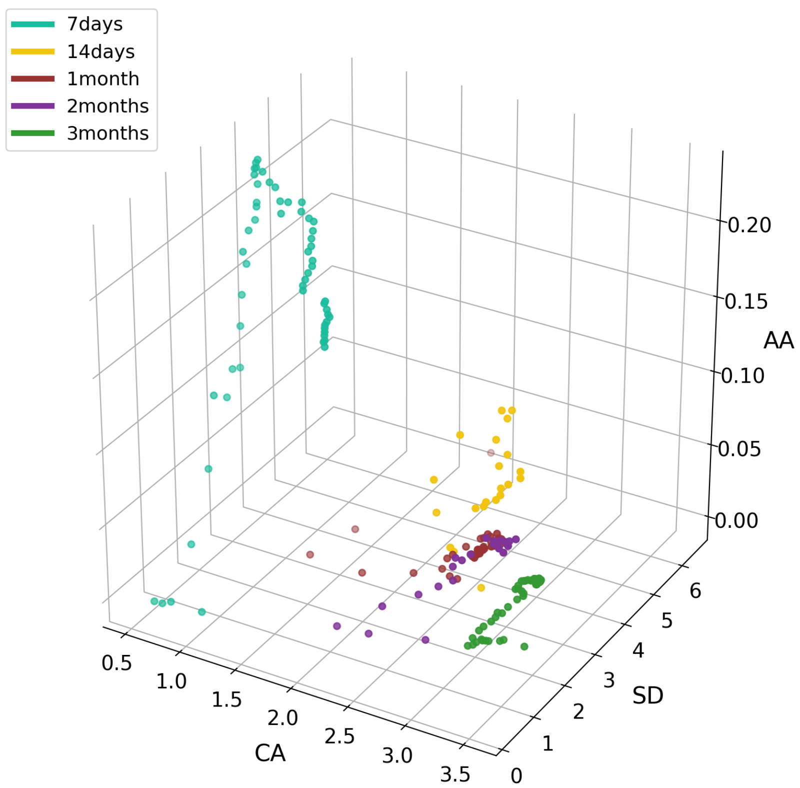


Figure 8. Dependence of objects in classes from CA, SD and AA variables.

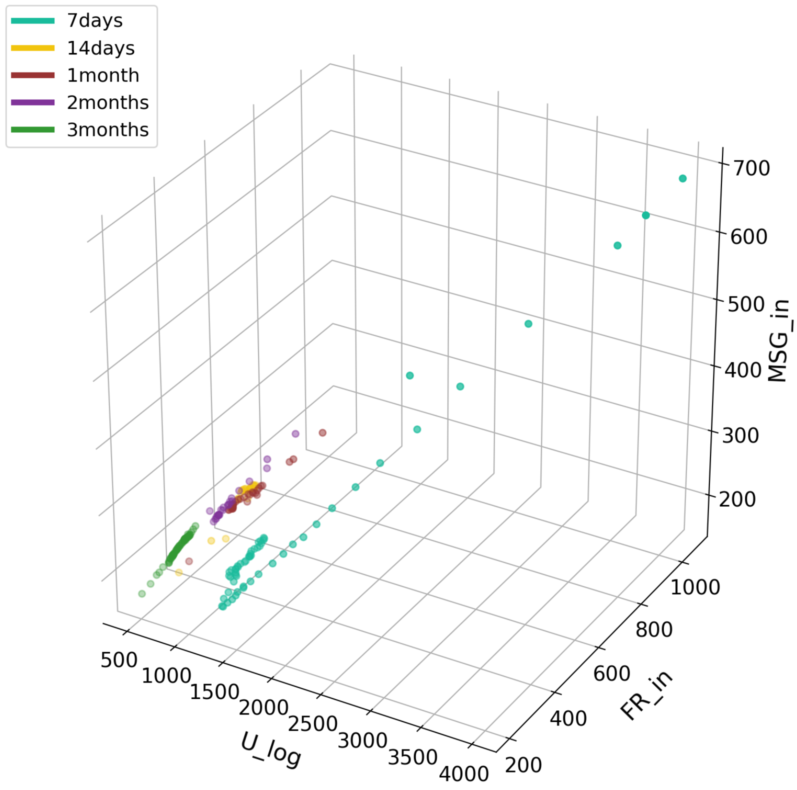


Figure 9. Dependence of objects in classes from U_log, FR_in and MSG_in variables.

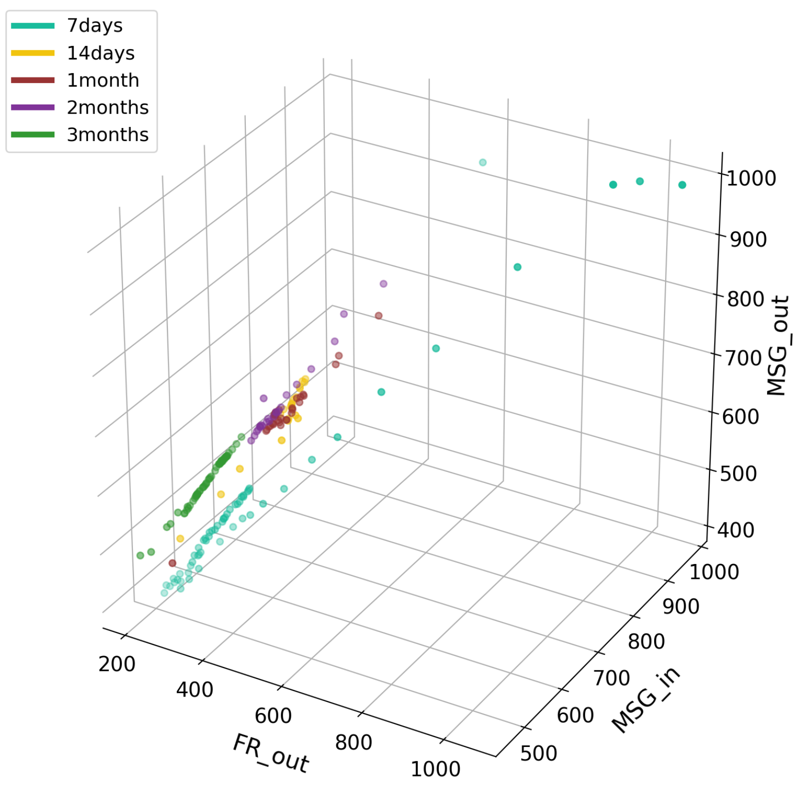


Figure 10. Dependence of objects in classes from FR_out, MSG_in and MSG_out variables.

Another stage investigates how the number of analyzed adopted users from 1% to 100% and their characteristics affect classification accuracy for the prediction of product lifetime and survival-class assignment (one week, two weeks, one month, two months, three months). The selection of observations to the training dataset was randomly performed; therefore, to stabilize the results, we repeated and averaged classification one hundred times for each dataset measure to obtain accurate information.

The experiment was carried out in three training-dataset sizes: 25%, 50%, and 75%. Classification and the decision-tree model were implemented with the help of the scikit-learn machine-learning library for the programming language Python. Classification was performed and, in the first stage, user-activity factors were used. Results are presented in Figure 11. They show high classification accuracy achieved for the training set based on 50% and 75% of the analyzed products. Accuracy at a level higher than 90% is achieved with less than 20% of product-usage statistics with activity factors taken into account. The training set based on 25% of the products delivered low accuracy, with a percentage of adopters lower than 60%, but it reached 90% when 70% of data were used for each product. Higher fluctuation of results was observed with a low number of analyzed adopted users.

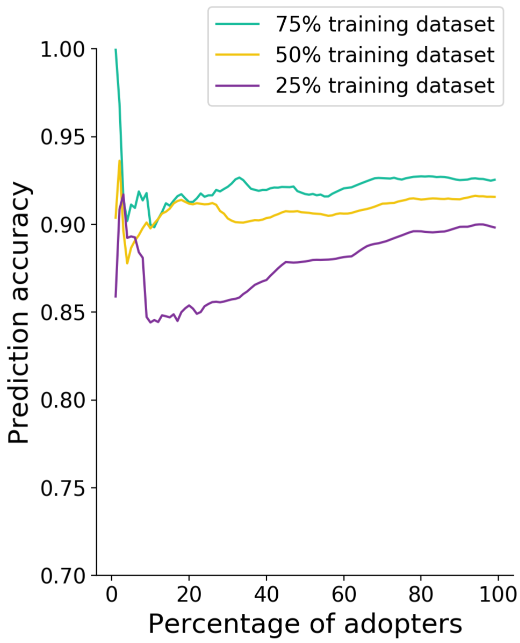


Figure 11. Accuracy of classification results with the use of Activity Factors for 25%, 50%, 75% training set and 10%, 20%, . . . , 90% of adopters used.

Detailed numerical results are presented in Table 4. It shows that analysis of characteristics of even only 10% of product adopters makes it possible to predict product assignment to a class with low or longer survival time.

Table 4. Accuracy of classification results with the use of Activity Factors for 25%, 50%, 75% training set and 10%, 20%, . . . , 90% of adopters used.

Traning Set	Number of Adopters Used for Classification								
	10%	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**
75%	0.900	0.915	0.920	0.920	0.917	0.921	0.926	0.927	0.925
50%	0.898	0.912	0.903	0.904	0.907	0.906	0.912	0.914	0.914
25%	0.844	0.854	0.857	0.868	0.879	0.881	0.890	0.896	0.899

Apart from social activity, factor classification was performed with the use of incremental data about user activity within the system. Results are presented in Figure 12. It shows that, for 50 and 70 training, the initial accuracy of the used low-fraction data at 1–3% of used data is very high due to innovator characteristics.

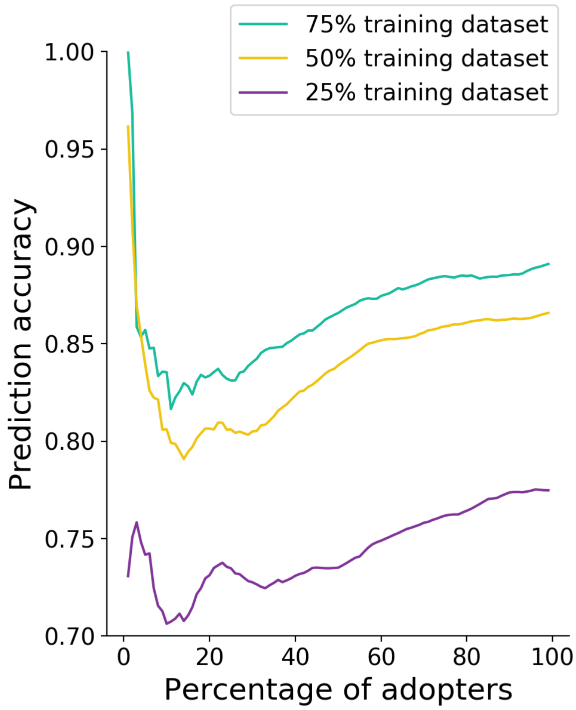


Figure 12. Accuracy of classification results with the use of Experience Factors for 25%, 50%, 75% training set and 10%, 20%, . . . , 90% of adopters used.

Additionally used data dropped accuracy to 80%. Subsequently, it grew with data acquisition. For the training set with the size of 75% of used products, the lowest accuracy was achieved after using data from 15% of adopters, and 65% accuracy for 50% of products with 15% of the adopter sample used. For all training-set sizes, accuracy continuously grew together with the increased number of adopters.

Detailed numerical results for classification based on incremental usage statistics represented by Experience Factors for each user are presented in Table 5.

Table 6 shows classification-accuracy statistics with identified user groups as innovators, then innovators together with adopters, and extended by early majority, late majority, and laggards. For Activity Factors, this shows that even using data from only innovators (2.5% of first adopters) creates the ability to assign a product to one of five adoption classes. Innovators used together with adopters delivered results above 19% for training sets with 50% and 75% size. Classification based on the 25% training dataset delivered accuracy above 18%. Further connection of the adopter group slightly improved classification, but from a practical-application perspective, it delays the time during which product survival abilities are predicted and additional adopter targeting is performed. The worst results were obtained for Experience Factors, but they were still above 80% accuracy for the training sets with 50% and 75%.

Table 5. Accuracy of classification results with the use of Experience Factors for 25%, 50%, 75% training set and 10%, 20%, . . . , 90% of adopters used.

Training Set	Number of Adopters Used for Classification								
	10%	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**
75%	0.835	0.834	0.841	0.853	0.866	0.875	0.882	0.885	0.885
50%	0.806	0.807	0.805	0.823	0.839	0.852	0.856	0.861	0.863
25%	0.706	0.731	0.728	0.731	0.735	0.749	0.758	0.764	0.773

Table 6. Accuracy of classification results with the use of Activity and Experience Factors for combined adoption group.

Group ID	Rogers Title	Size of Training Dataset					
		Activity Factors			**Experience Factors**		
		75%	**50%**	**25%**	**75%**	**50%**	**25%**
G1	innovators	0.961	0.912	0.895	0.943	0.914	0.747
G2	+ early adopters	0.919	0.901	0.869	0.855	0.831	0.725
G3	+ early majority	0.919	0.905	0.865	0.849	0.821	0.729
G4	+ late majority	0.921	0.907	0.874	0.861	0.834	0.739
G5	+ laggards	0.922	0.908	0.878	0.865	0.839	0.744

4. Discussion and Conclusions

For expanded virtual product usage within online systems, new analytical models and strategies are required. Common phenomena to offline markets are regularly seen in electronic systems and are identified with lifespan, customer habituation, and new-product improvement techniques. This research indicates how the attributes of early adopters to new items can influence user engagement and the survival of virtual goods within dynamic electronic environments. Achieved results, from product classification based on decision trees, showed that it is possible to predict product lifespan with the use of adopter characteristics. Adopter communication activity, represented by Activity Factors, positively affected product survival time. This shows that adopters with high experience factors are the main influencers in the system, and their behavior is adopted by other users.

Monitoring of product-usage patterns and adopter characteristics makes it possible to identify products with possible low survival time, and invite additional adopters with the use of incentives and other techniques. Gathered knowledge can be used to reduce the habituation effect and increase product-usage time due to social influence and follower behavior.

Results from the conducted study lead to the following main conclusions:

- characteristics of early adopters related to social activity positively influence product lifespan and the engagement of other users within the system;
- product lifespan can be estimated with the use of initial-audience and early-adopter characteristics;
- the combination of innovators and adopters positively affects the statistical significance of the dependent variable that represents survival time;
- initial-user characteristics can be used to classify products in terms of future usage for the detection of low-potential products, for performance improvement and targeting additional adopters with the desired specifics.

Future work will concentrate on a progressive point-by-point evaluation of distribution in the use of social networks and behavior prediction, which is dependent on interpersonal organizations, and the use of conduct forecast, which is dependent on earlier behaviors.

Author Contributions: conceptualization, K.B., P.P., J.H., P.B. and J.J.; methodology, K.B., P.P., J.H., P.B. and J.J.; validation, K.B., P.P., P.B.; investigation, K.B., P.P. and P.B.; resources, P.B. and P.P.; data curation, P.B. and P.P.; writing–original draft preparation, K.B., P.P., J.H., P.B. and J.J.; writing–review and editing, K.B., P.P., J.H., P.B. and J.J.; visualization, K.B., P.P.; supervision, J.J.; project administration, J.J.; funding acquisition, J.J.

Funding: The work was supported by the National Science Centre of Poland, the decision no. 2017/27/B/HS4/01216.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Supporting information

Table A1. Intergroup comparisons using the life expectancy of a product as a dependent variable and as divisional predictors divided into two groups.

Variables		Activity Factors							
		7 Days vs. 3 m		14 Days vs. 3 m		1 m vs. 3 m		2 m vs. 3 m	
		p	z	p	z	p	z	p	z
innovator	CA	0.21	−1.25	−2.00	0.05	<0.01	−3.17	<0.01	−2.88
	SD	0.28	−1.07	−0.41	0.69	0.27	−1.10	0.04	−2.10
	CP	0.64	−0.47	0.69	0.49	0.32	0.99	0.74	−0.33
	SP	0.11	−1.62	−0.42	0.67	0.08	−1.75	0.02	−2.26
	AA	0.85	−0.19	1.29	0.20	0.93	−0.09	0.36	0.91
early adopter	CA	<0.01	−5.90	−5.57	0.00	<0.01	−2.84	0.01	−2.64
	SD	<0.01	−3.17	−4.22	<0.01	<0.01	−4.14	<0.01	−3.99
	CP	0.52	−0.64	−1.64	0.10	0.44	0.77	0.57	0.57
	SP	<0.01	−4.50	−4.46	<0.01	<0.01	−2.85	0.06	−1.88
	AA	<0.01	3.70	2.39	0.02	0.15	1.44	0.55	0.60
early majority	CA	<0.01	−8.13	−5.61	<0.01	<0.01	−2.96	0.50	−0.67
	SD	<0.01	−5.14	−5.48	<0.01	<0.01	−4.48	0.21	−1.25
	CP	0.15	1.44	0.92	0.36	0.56	-0.58	0.64	−0.46
	SP	0.03	−2.16	−3.39	<0.01	0.01	−2.46	0.19	−1.31
	AA	0.66	0.45	0.69	0.49	0.64	0.47	0.12	1.54
late majority	CA	<0.01	−7.00	−1.37	0.17	0.05	−1.94	<0.01	−3.17
	SD	<0.01	−3.17	−2.41	0.02	<0.01	−2.87	0.40	−0.84
	CP	0.04	2.04	1.65	0.10	0.68	0.41	0.79	−0.27
	SP	0.06	−1.86	−1.20	0.23	0.65	−0.46	0.03	−2.18
	AA	0.71	0.37	2.30	0.02	0.76	0.31	0.36	0.92
laggards	CA	<0.01	−5.32	−0.66	0.51	0.03	−2.21	0.82	−0.23
	SD	0.02	−2.43	−0.37	0.71	0.56	−0.58	0.80	0.25
	CP	0.89	0.14	−0.34	0.74	0.97	0.04	0.76	0.30
	SP	0.50	−0.68	−0.60	0.55	0.04	−2.10	0.46	−0.74
	AA	0.35	0.93	2.11	0.03	0.45	0.76	0.22	1.24
All together	CA	<0.01	−8.20	−3.46	<0.01	0.01	−2.46	0.14	−1.47
	SD	0.33	−0.98	−2.32	0.02	<0.01	−2.99	0.66	−0.44
	CP	0.07	1.84	0.12	0.91	0.69	−0.40	0.57	−0.57
	SP	<0.01	−5.87	−3.25	<0.01	0.01	−2.47	0.03	−2.19
	AA	<0.01	4.24	0.87	0.38	0.16	1.40	0.02	2.39
G1	CA	0.21	−1.25	−2.00	0.05	<0.01	−3.17	<0.01	−2.88
	SD	0.28	−1.07	−0.41	0.69	0.27	−1.10	0.04	−2.10
	CP	0.64	−0.47	0.69	0.49	0.32	0.99	0.74	−0.33
	SP	0.11	−1.62	−0.42	0.67	0.08	−1.75	0.02	−2.26
	AA	0.85	−0.19	1.29	0.20	0.93	−0.09	0.36	0.91

Table A1. *Cont.*

Variables		Activity Factors							
		7 Days vs. 3 m		14 Days vs. 3 m		1 m vs. 3 m		2 m vs. 3 m	
		p	z	p	z	p	z	p	z
G2	CA	<0.01	−5.72	−5.07	<0.01	<0.01	−2.84	<0.01	−2.87
	SD	<0.01	−2.92	−3.59	<0.01	<0.01	-3.43	<0.01	−3.80
	CP	0.73	−0.34	−1.23	0.22	0.73	0.34	0.60	0.53
	SP	<0.01	−4.02	−4.36	<0.01	0.01	−2.68	0.01	−2.47
	AA	<0.01	3.12	2.95	<0.01	0.17	1.37	0.54	0.62
G3	CA	<0.01	−8.20	−5.63	<0.01	<0.01	−3.25	0.03	−2.11
	SD	0.01	−2.53	−5.23	<0.01	<0.01	−4.16	0.03	−2.19
	CP	0.70	−0.38	−0.08	0.93	0.48	−0.71	0.64	−0.47
	SP	<0.01	−6.16	−4.13	<0.01	0.01	−2.67	0.01	−2.75
	AA	<0.01	3.14	−0.11	0.92	0.86	−0.17	0.91	0.12
G4	CA	<0.01	−8.26	−3.96	<0.01	0.01	−2.53	<0.01	−3.10
	SD	0.12	−1.56	−3.29	<0.01	<0.01	−3.37	0.62	−0.49
	CP	0.41	0.82	0.53	0.60	0.15	−1.46	0.30	−1.04
	SP	<0.01	−5.80	−3.55	<0.01	0.01	−2.45	<0.01	−3.04
	AA	<0.01	2.98	0.35	0.73	0.95	−0.06	0.56	0.58
G5	CA	<0.01	−8.20	−3.46	<0.01	0.01	−2.46	0.14	−1.47
	SD	0.33	−0.98	−2.32	0.02	<0.01	−2.99	0.66	−0.44
	CP	0.07	1.84	0.12	0.91	0.69	−0.40	0.57	−0.57
	SP	<0.01	−5.87	−3.25	<0.01	0.01	−2.47	0.03	−2.19
	AA	<0.01	4.24	0.87	0.38	0.16	1.40	0.02	2.39

Table A2. Intergroup comparisons using the life-length of a product as a dependent variable and as non-variable predictors.

Variables		Experience Factors							
		7 Days vs. 3 Months		14 Days vs. 3 Months		1 Month vs. 3 Months		2 Months vs. 3 Months	
		p	z	p	z	p	z	p	z
G1	MSG_in	0.99	−0.01	0.16	0.88	0.38	−0.87	0.93	−0.09
	MSG_out	0.13	1.53	−0.48	0.63	0.55	−0.6	0.81	−0.24
	MSG_total	0.83	0.21	−0.13	0.90	0.47	−0.72	0.87	−0.17
	FR_in	0.46	−0.73	−1.17	0.24	0.32	−0.99	0.24	1.17
	FR_out	0.21	−1.25	−1.61	0.11	0.91	−0.11	0.10	1.67
	FR_total	0.43	−0.79	−1.34	0.18	0.74	−0.33	0.10	1.67
G2	MSG_in	0.25	−1.15	0.25	0.80	0.01	2.49	0.04	2.02
	MSG_out	0.18	−1.33	0.67	0.50	0.01	2.69	0.02	2.37
	MSG_total	0.41	−0.83	0.42	0.68	0.01	2.71	0.03	2.20
	FR_in	0.01	−2.6	−0.25	0.80	0.03	2.15	0.01	2.47
	FR_out	0.06	−1.9	1.33	0.18	<0.01	2.86	<0.01	3.85
	FR_total	0.03	−2.12	0.82	0.41	0.01	2.64	<0.01	3.44
G3	MSG_in	0.39	0.86	3.92	<0.01	<0.01	3.61	<0.01	3.12
	MSG_out	0.69	−0.4	4.09	<0.01	<0.01	3.91	<0.01	3.27
	MSG_total	0.96	0.05	4.05	<0.01	<0.01	3.79	<0.01	3.22
	FR_in	0.47	−0.72	4.39	<0.01	<0.01	3.82	<0.01	3.34
	FR_out	0.92	0.10	4.90	<0.01	<0.01	4.76	<0.01	4.24
	FR_total	0.84	−0.20	4.84	<0.01	<0.01	4.57	<0.01	3.91
G4	MSG_in	0.82	0.23	3.90	<0.01	0.01	2.66	<0.01	3.15
	MSG_out	0.96	0.05	4.08	<0.01	<0.01	2.98	<0.01	3.24
	MSG_total	0.95	0.06	3.98	<0.01	<0.01	2.86	<0.01	3.24
	FR_in	0.29	−1.07	3.93	<0.01	<0.01	2.99	<0.01	3.62
	FR_out	0.80	0.25	4.57	<0.01	<0.01	4.18	<0.01	4.20
	FR_total	0.90	−0.12	4.38	<0.01	<0.01	3.83	<0.01	4.01
G5	MSG_in	0.18	1.33	3.57	<0.01	<0.01	4.44	<0.01	3.36
	MSG_out	0.4	0.85	3.72	<0.01	<0.01	4.47	<0.01	3.43
	MSG_total	0.34	0.96	3.65	<0.01	<0.01	4.48	<0.01	3.43
	FR_in	0.63	0.48	3.87	<0.01	<0.01	4.54	<0.01	3.55
	FR_out	0.16	1.40	4.31	<0.01	<0.01	4.98	<0.01	3.97
	FR_total	0.27	1.10	4.18	<0.01	<0.01	4.84	<0.01	3.86

Table A3. Intergroup comparisons using the life-length of a product as a dependent variable and as non-variable predictors.

Variables		Experience Factors							
		7 Days vs. 3 Months		14 Days vs. 3 Months		1 Month vs. 3 Months		2 Months vs. 3 Months	
		p	z	p	z	p	z	p	z
innovator	MSG_in	0.99	−0.01	0.16	0.88	0.38	−0.87	0.93	−0.09
	MSG_out	0.13	1.53	−0.48	0.63	0.55	−0.60	0.81	−0.24
	MSG_total	0.83	0.21	−0.13	0.90	0.47	−0.72	0.87	−0.17
	FR_in	0.46	−0.73	−1.17	0.24	0.32	−0.99	0.24	1.17
	FR_out	0.21	−1.25	−1.61	0.11	0.91	−0.11	0.10	1.67
	FR_total	0.43	−0.79	−1.34	0.18	0.74	−0.33	0.10	1.67
early adopter	MSG_in	0.21	−1.27	0.92	0.36	<0.01	4.50	<0.01	3.14
	MSG_out	0.16	−1.40	1.36	0.17	<0.01	4.64	<0.01	3.17
	MSG_total	0.33	−0.97	1.10	0.27	<0.01	4.69	<0.01	3.14
	FR_in	0.02	−2.42	0.29	0.77	<0.01	3.18	0.02	2.34
	FR_out	0.17	−1.37	2.10	0.04	<0.01	3.83	<0.01	3.42
	FR_total	0.12	−1.56	1.37	0.17	<0.01	3.71	<0.01	3.10
early majority	MSG_in	<0.01	3.96	5.34	<0.01	<0.01	4.66	<0.01	3.31
	MSG_out	<0.01	2.93	5.37	<0.01	<0.01	4.73	<0.01	3.31
	MSG_total	<0.01	3.66	5.31	<0.01	<0.01	4.75	<0.01	3.38
	FR_in	0.12	1.55	5.35	<0.01	<0.01	4.47	<0.01	3.19
	FR_out	0.38	0.88	5.48	<0.01	<0.01	4.89	<0.01	3.18
	FR_total	0.18	1.35	5.43	<0.01	<0.01	4.77	<0.01	3.14
late majority	MSG_in	0.06	1.87	4.26	<0.01	<0.01	3.18	<0.01	3.26
	MSG_out	0.01	2.81	4.32	<0.01	<0.01	3.30	<0.01	3.39
	MSG_total	0.02	2.37	4.30	<0.01	<0.01	3.29	<0.01	3.30
	FR_in	0.81	0.24	3.80	<0.01	0.01	2.59	<0.01	3.07
	FR_out	0.37	0.90	4.00	<0.01	<0.01	3.15	<0.01	3.75
	FR_total	0.53	0.63	3.97	<0.01	<0.01	3.02	<0.01	3.43
laggards	MSG_in	0.03	2.19	3.15	<0.01	<0.01	3.16	<0.01	3.32
	MSG_out	0.01	2.64	3.36	<0.01	<0.01	3.29	<0.01	3.37
	MSG_total	0.02	2.36	3.31	<0.01	<0.01	3.29	<0.01	3.41
	FR_in	0.04	2.08	2.69	0.01	<0.01	2.99	<0.01	3.09
	FR_out	0.04	2.01	2.30	0.02	<0.01	2.89	<0.01	2.87
	FR_total	0.04	2.05	2.39	0.02	<0.01	3.04	<0.01	2.91
All together	MSG_in	0.18	1.33	3.57	<0.01	<0.01	4.44	<0.01	3.36
	MSG_out	0.40	0.85	3.72	<0.01	<0.01	4.47	<0.01	3.43
	MSG_total	0.34	0.96	3.65	<0.01	<0.01	4.48	<0.01	3.43
	FR_in	0.63	0.48	3.87	<0.01	<0.01	4.54	<0.01	3.55
	FR_out	0.16	1.40	4.31	<0.01	<0.01	4.98	<0.01	3.97
	FR_total	0.27	1.10	4.18	<0.01	<0.01	4.84	<0.01	3.86

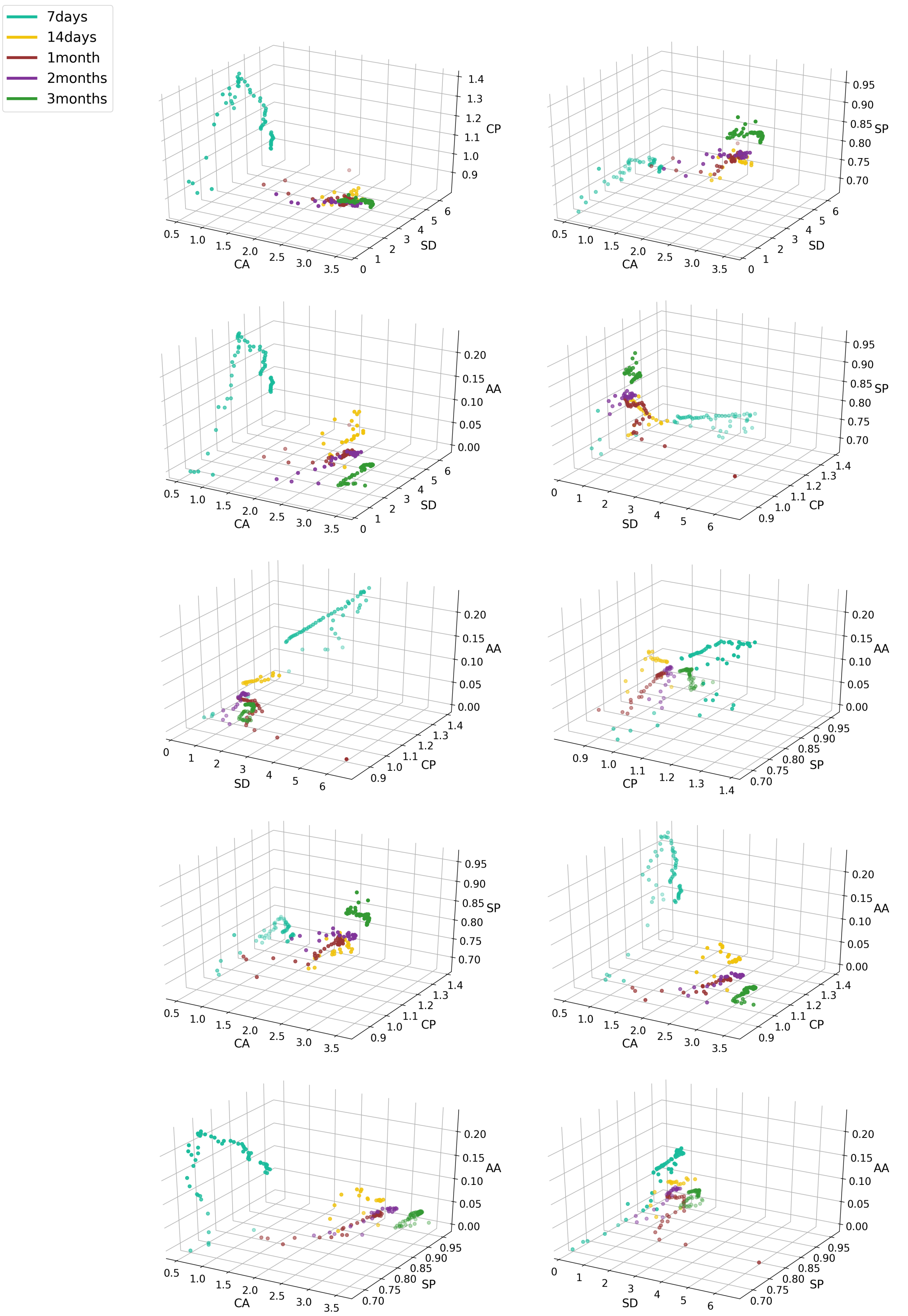


Figure A1. Dependence of objects in classes from all Activity Factors.

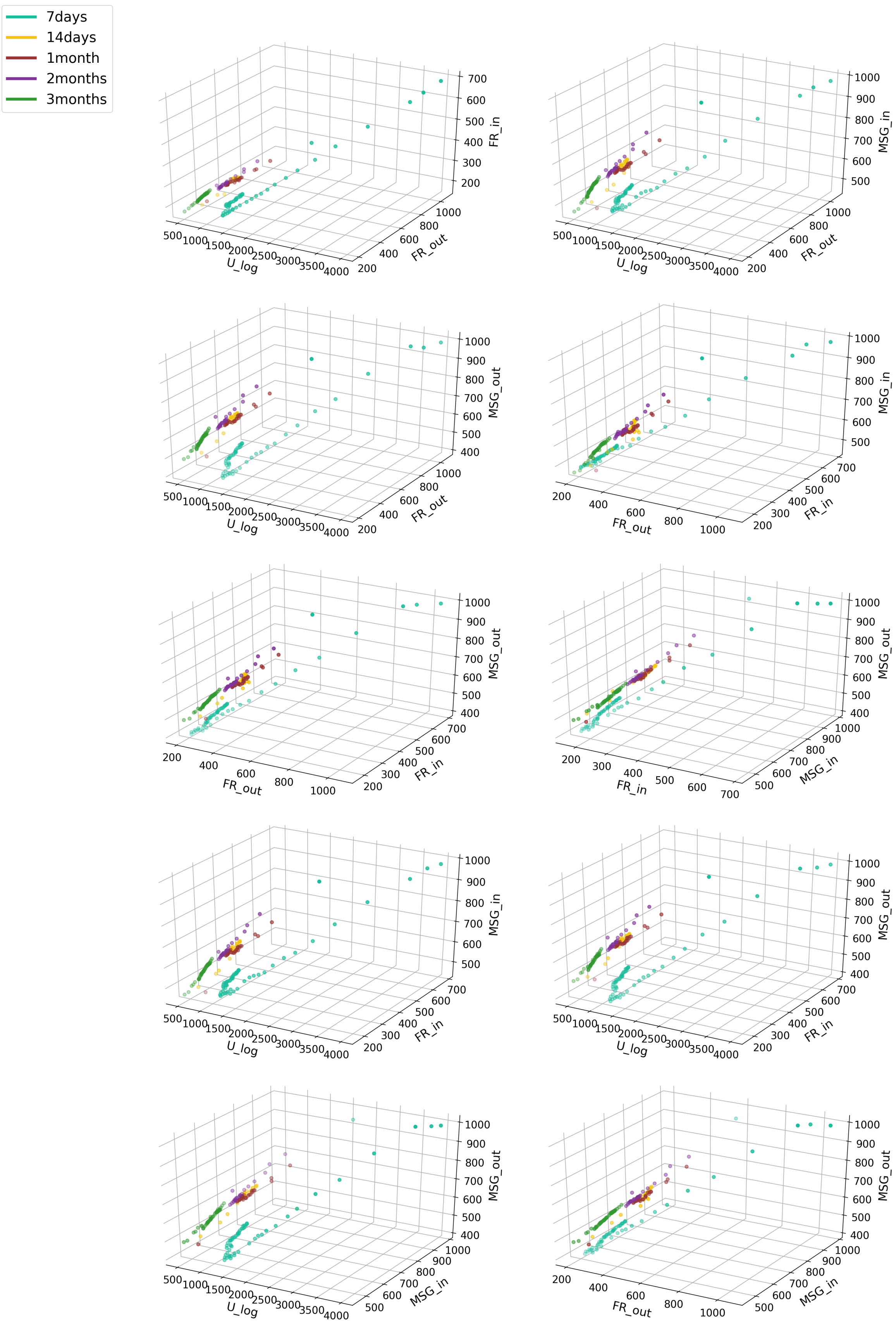


Figure A2. Dependence of objects in classes from all Experience Factors.

References

1. Bainbridge, W.S. The scientific research potential of virtual worlds. *Science* **2007**, *317*, 472–476. [CrossRef]
2. Lazer, D.; Pentland, A.; Adamic, L.; Aral, S.; Barabási, A.L.; Brewer, D.; Christakis, N.; Contractor, N.; Fowler, J.; Gutmann, M.; et al. Computational social science. *Science* **2009**, *323*, 721–723. [CrossRef]
3. Horton, J.J.; Rand, D.G.; Zeckhauser, R.J. The online laboratory: Conducting experiments in a real labor market. *Exp. Econ.* **2011**, *14*, 399–425. [CrossRef]
4. Hamari, J.; Keronen, L. Why do people buy virtual goods: A meta-analysis. *Comput. Hum. Behav.* **2017**, *71*, 59–69. [CrossRef]
5. Lehdonvirta, V. Virtual item sales as a revenue model: Identifying attributes that drive purchase decisions. *Electron. Commer. Res.* **2009**, *9*, 97–113. [CrossRef]
6. Lehdonvirta, V.; Castronova, E. *Virtual Economies: Design and Analysis*; MIT Press: Cambridge, MA, USA, 2014.
7. Lin, H.; Sun, C.T. Cash trade in free-to-play online games. *Games Cult.* **2011**, *6*, 270–287. [CrossRef]
8. Alha, K.; Koskinen, E.; Paavilainen, J.; Hamari, J.; Kinnunen, J. Free-to-play games: Professionals' perspectives. In Proceedings of the 2014 International DiGRA Nordic Conference, Upsalla, Sweden, 29–30 May 2014.
9. Alha, K.; Koskinen, E.; Paavilainen, J.; Hamari, J. Critical Acclaim and Commercial Success in Mobile Free-to-Play Games. In Proceedings of the First International Joint Conference of DiGRA and FDG, Dundee, UK, 1–6 August 2016.
10. Hamari, J.; Hanner, N.; Koivisto, J. Service quality explains why people use freemium services but not if they go premium: An empirical study in free-to-play games. *Int. J. Inf. Manag.* **2017**, *37*, 1449–1459. [CrossRef]
11. Hamari, J.; Lehdonvirta, V. Game design as marketing: How game mechanics create demand for virtual goods. *Int. J. Bus. Sci. Appl. Manag.* **2010**, *5*, 14–29.
12. Hamari, J.; Järvinen, A. Building customer relationship through game mechanics in social games. In *Business, Technological, and Social Dimensions of Computer Games: Multidisciplinary Developments*; IGI Global: Hershey, PA, USA, 2011; pp. 348–365.
13. Heimo, O.I.; Harviainen, J.T.; Kimppa, K.K.; Mäkilä, T. Virtual to virtuous money: A virtue ethics perspective on video game business logic. *J. Bus. Ethics* **2018**, *153*, 95–103. [CrossRef]
14. Fairfield, J.A. Virtual property. *B.U.L. Rev.* **2005**, *85*, 1047.
15. Macey, J.; Hamari, J. eSports, skins and loot boxes: Participants, practices and problematic behaviour associated with emergent forms of gambling. *New Media Soc.* **2019**, *21*, 20–41. [CrossRef]
16. Taylor, T.L. Living digitally: Embodiment in virtual worlds. In *The Social Life of Avatars*; Springer: London, UK, 2002; pp. 40–62.
17. Jankowski, J.; Kolomvatsos, K.; Kazienko, P.; Watróbski, J. Fuzzy Modeling of User Behaviors and Virtual Goods Purchases in Social Networking Platforms. *J. UCS* **2016**, *22*, 416–437.
18. Lehdonvirta, V. Virtual economics: Applying economics to the study of game worlds. In Proceedings of the 2005 Conference on Future Play (Future Play 2005), Lansing, MI, USA, 13–15 October 2005.
19. Orbach, B.Y. The durapolist puzzle: Monopoly power in durable-goods markets. *Yale J. Regul.* **2004**, *21*, 67.
20. Lu, H.P.; Wang, S.m. The role of Internet addiction in online game loyalty: An exploratory study. *Internet Res.* **2008**, *18*, 499–519. [CrossRef]
21. Kwong, J.A. Getting the goods on virtual items: A fresh look at transactions in multi-user online environments. *William Mitchell Law Rev.* **2010**, *37*, 1805.
22. Kaplan, A.M.; Haenlein, M. Consumer use and business potential of virtual worlds: The case of second life. *Int. J. Media Manag.* **2009**, *11*, 93–101. [CrossRef]
23. Wathieu, L. Consumer habituation. *Manag. Sci.* **2004**, *50*, 587–596. [CrossRef]
24. Rogers, E.M. *Diffusion of Innovations*, 4th ed.; Free Press: New York, NY, USA, 1962.
25. Jankowski, J.; Bródka, P.; Hamari, J. A picture is worth a thousand words: An empirical study on the influence of content visibility on diffusion processes within a virtual world. *Behav. Inf. Technol.* **2016**, *35*, 926–945. [CrossRef]
26. Jankowski, J.; Michalski, R.; Bródka, P. A multilayer network dataset of interaction and influence spreading in a virtual world. *Sci. Data* **2017**, *4*, 170144. [CrossRef] [PubMed]

27. Cleves, M.; Gould, W.; Gould, W.W.; Gutierrez, R.; Marchenko, Y. *An Introduction to Survival Analysis Using Stata*; Stata Press: Lakeway Drive, TX, USA, 2008.
28. Drye, T.; Wetherill, G.; Pinnock, A. When are customers in the market? Applying survival analysis to marketing challenges. *J. Target. Meas. Anal. Mark.* **2001**, *10*, 179–188. [CrossRef]
29. Fader, P.S.; Hardie, B.G. Probability models for customer-base analysis. *J. Interact. Mark.* **2009**, *23*, 61–69. [CrossRef]
30. Miller, R.G., Jr. *Survival Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 2011; Volume 66.
31. Larivière, B.; Van den Poel, D. Investigating the role of product features in preventing customer churn, by using survival analysis and choice modeling: The case of financial services. *Expert Syst. Appl.* **2004**, *27*, 277–285. [CrossRef]
32. Giot, P.; Schwienbacher, A. IPOs, trade sales and liquidations: Modelling venture capital exits using survival analysis. *J. Bank. Financ.* **2007**, *31*, 679–702. [CrossRef]
33. Greene, W.H. *Econometric Analysis (International Edition)*; Pearson Education: London, UK, 2000.
34. Kettunen, J. Education and unemployment duration. *Econ. Educ. Rev.* **1997**, *16*, 163–170. [CrossRef]
35. Schmittlein, D.C.; Morrison, D.G.; Colombo, R. Counting your customers: Who-are they and what will they do next? *Manag. Sci.* **1987**, *33*, 1–24. [CrossRef]
36. Fader, P.S.; Hardie, B.G.; Lee, K.L. "Counting your customers" the easy way: An alternative to the Pareto/NBD model. *Mark. Sci.* **2005**, *24*, 275–284. [CrossRef]
37. Mavri, M.; Ioannou, G. Customer switching behaviour in Greek banking services using survival analysis. *Manag. Financ.* **2008**, *34*, 186–197. [CrossRef]
38. Harrison, T.; Ansell, J. Customer retention in the insurance industry: Using survival analysis to predict cross-selling opportunities. *J. Financ. Serv. Mark.* **2002**, *6*, 229–239. [CrossRef]
39. Bellotti, T.; Crook, J. Credit scoring with macroeconomic variables using survival analysis. *J. Oper. Res. Soc.* **2009**, *60*, 1699–1707. [CrossRef]
40. Narain, B. Survival analysis and the credit-granting decision. In *Readings in Credit Scoring: Foundations, Developments, and Aims*; Oxford University Press: Oxford, UK, 2004; p. 235.
41. Glennon, D.; Nigro, P. Measuring the default risk of small business loans: A survival analysis approach. *J. Money Credit Bank.* **2005**, *37*, 923–947. [CrossRef]
42. Kauffman, R.J.; Wang, B. The success and failure of dotcoms: A multi-method survival analysis. In Proceedings of the 6th INFORMS Conference on Information Systems and Technology (CIST), Miami, FL, USA, 3–4 November 2001.
43. Samoladas, I.; Angelis, L.; Stamelos, I. Survival analysis on the duration of open source projects. *Inf. Softw. Technol.* **2010**, *52*, 902–922. [CrossRef]
44. Wang, Q.; Chen, Y.; Xie, J. Survival in markets with network effects: Product compatibility and order-of-entry effects. *J. Mark.* **2010**, *74*, 1–14. [CrossRef]
45. Bhattacharjee, S.; Gopal, R.D.; Lertwachara, K.; Marsden, J.R.; Telang, R. The effect of digital sharing technologies on music markets: A survival analysis of albums on ranking charts. *Manag. Sci.* **2007**, *53*, 1359–1374. [CrossRef]
46. Jung, E.Y.; Baek, C.; Lee, J.D. Product survival analysis for the App Store. *Mark. Lett.* **2012**, *23*, 929–941. [CrossRef]
47. Wang, J.; Zhang, Y. Opportunity model for e-commerce recommendation: Right product; right time. In Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval, Dublin, Ireland, 28 July–1 August 2013; pp. 303–312.
48. Kossinets, G.; Watts, D.J. Empirical analysis of an evolving social network. *Science* **2006**, *311*, 88–90. [CrossRef] [PubMed]
49. Deng, Y.; Hou, J.; Ma, X.; Cai, S. A dual model of entertainment-based and community-based mechanisms to explore continued participation in online entertainment communities. *Cyberpsychol. Behav. Soc. Netw.* **2013**, *16*, 378–384. [CrossRef] [PubMed]
50. Tarng, P.Y.; Chen, K.T.; Huang, P. An analysis of WoW players' game hours. In Proceedings of the 7th ACM SIGCOMM Workshop on Network and System Support for Games, Worcester, MA, USA, 21–22 October 2008; pp. 47–52.

51. Isaksen, A.; Gopstein, D.; Nealen, A. Exploring Game Space Using Survival Analysis. In Proceedings of the 10th International Conference on the Foundations of Digital Games (FDG 2015), Pacific Grove, CA, USA, 22–25 June 2015; ISBN 978-0-9913982-4-9.
52. Periáñez, Á.; Saas, A.; Guitart, A.; Magne, C. Churn prediction in mobile social games: Towards a complete assessment using survival ensembles. In Proceedings of the 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), Montreal, QC, Canada, 17–19 October 2016; pp. 564–573.
53. Borbora, Z.; Srivastava, J.; Hsu, K.W.; Williams, D. Churn prediction in mmorpgs using player motivation theories and an ensemble approach. In Proceedings of the 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing, Boston, MA, USA, 9–11 October 2011; pp. 157–164.
54. Viljanen, M.; Airola, A.; Heikkonen, J.; Pahikkala, T. Playtime measurement with survival analysis. *IEEE Trans. Games* **2018**, *10*, 128–138. [CrossRef]
55. Gu, L.; Jia, A.L. Player Activity and Popularity in Online Social Games and their Implications for Player Retention. In Proceedings of the 2018 16th Annual Workshop on Network and Systems Support for Games (NetGames), Amsterdam, The Netherlands, 12–15 June 2018; pp. 1–6.
56. Zhuang, X.; Bharambe, A.; Pang, J.; Seshan, S. *Player Dynamics in Massively Multiplayer Online Games*; Tech. Rep. CMU-CS-07-158; School of Computer Science, Carnegie Mellon University: Pittsburgh, PA, USA, 2007.
57. Huang, T.Y.; Chen, K.T.; Huang, P.; Lei, C.L. A generalizable methodology for quantifying user satisfaction. *IEICE Trans. Commun.* **2008**, *91*, 1260–1268. [CrossRef]
58. Safavian, S.R.D.L. A survey of decision tree classifier methodology. *IEEE Trans. Syst. Man. Cybern.* **1991**, *21*, 660–674. [CrossRef]
59. Podgorelec, V.; Kokol, P.; Stiglic, B.; Rozman, I. Decision Trees: An Overview and Their Use in Medicine. *J. Med. Syst.* **2002**, *26*, 445–463. [CrossRef]
60. Lee, S.; Lee, S.; Park, Y. A prediction model for success of services in e-commerce using decision tree: E-customer's attitude towards online service. *Expert Syst. Appl.* **2007**, *33*, 572–581. [CrossRef]
61. McGee, J.; Caverlee, J.; Cheng, Z. Location prediction in social media based on tie strength. In Proceedings of the 22nd ACM international conference on Information & Knowledge Management, San Francisco, CA, USA, 27 October– 1 November 2013; pp. 459–468.
62. Nivedha, R.; Sairam, N. A machine learning based classification for social media messages. *Indian J. Sci. Technol.* **2015**, *8*. [CrossRef]
63. Doh, Y.Y. Player Segmentation Strategies Based on the Types of Self-recognition in Online Game World. In *Morphological Analysis of Cultural DNA*; Springer: Singapore, 2017; pp. 139–148.
64. Chen, L.S.; Chang, P.C. Extracting knowledge of customers' preferences in massively multiplayer online role playing games. *Neural Comput. Appl.* **2013**, *23*, 1787–1799. [CrossRef]
65. Mahlmann, T.; Drachen, A.; Togelius, J.; Canossa, A.; Yannakakis, G.N. Predicting player behavior in tomb raider: Underworld. In Proceedings of the 2010 IEEE Symposium on Computational Intelligence and Games (CIG), Dublin, Ireland, 18–21 August 2010; pp. 178–185.
66. Hsieh, J.L.; Sun, C.T. Building a player strategy model by analyzing replays of real-time strategy games. In Proceedings of the 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), Hong Kong, China, 1–8 June 2008; pp. 3106–3111.
67. Kang, A.R.; Woo, J.; Park, J.; Kim, H.K. Online game bot detection based on party-play log analysis. *Comput. Math. Appl.* **2013**, *65*, 1384–1395. [CrossRef]
68. Tsai, F.H.; Kuang-Chao, Y.; Hsiao, H.S. Exploring the factors influencing learning effectiveness in digital game-based learning. *J. Educ. Technol. Soc.* **2012**, *15*, 240.
69. Kim, K.S.; Kim, K.H. A prediction model for internet game addiction in adolescents: Using a decision tree analysis. *J. Korean Acad. Nurs.* **2010**, *40*, 378–388. [CrossRef]
70. Han, Y.T.; Park, G.S. Game traffic classification using statistical characteristics at the transport layer. *ETRI J.* **2010**, *32*, 22–32. [CrossRef]
71. Drachen, A.; Sifa, R.; Bauckhage, C.; Thurau, C. Guns, swords and data: Clustering of player behavior in computer games in the wild. In Proceedings of the 2012 IEEE Conference on Computational Intelligence and Games (CIG), Granada, Spain, 11–14 September 2012; pp. 163–170.
72. Fu, X.; Chen, X.; Shi, Y.T.; Bose, I.; Cai, S. User segmentation for retention management in online social games. *Decis. Support Syst.* **2017**, *101*, 51–68. [CrossRef]

73. Sheu, J.J.; Su, Y.H.; Chu, K.T. Segmenting online game customers—The perspective of experiential marketing. *Expert Syst. Appl.* **2009**, *36*, 8487–8495. [CrossRef]
74. Breiman, L.J.; Friedman, R.O.C.S. *Classification and Regression Trees*; Chapman and Hall/CRC: London, UK, 1984.
75. Bel, L.D.; Allard, J.R.A.H. CART algorithm for spatial data: Application to environmental and ecological data. *Comput. Stat. Data Anal.* **2009**, *53*, 3082–3093. [CrossRef]
76. Rokach, L.O.M. Top-down induction of decision trees classifiers—A survey. *IEEE Trans. Syst. Man. Cybern. Part C Appl. Rev.* **2005**, *35*, 476–487. [CrossRef]
77. Stéphane Grumbach, T.M. Towards Tractable Algebras for Bags. *J. Comput. Syst. Sci.* **1996**, *52*, 570–588. [CrossRef]

A3.

Bortko, K., Jankowski, J., Pazura, P. (2020). From perceptual to algorithmic evaluation of recommending interfaces survival in visual space. Procedia Computer Science, 176, 2736-2745.

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 176 (2020) 2736–2745

Procedia Computer Science

www.elsevier.com/locate/procedia

24th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems

From Perceptual to Algorithmic Evaluation of Recommending Interfaces Survival in Visual Space

Kamil Bortko[a,*], Jarosław Jankowski[a], Patryk Pazura[a]

[a] *Faculty of Computer Science and Information Technology, West Pomeranian University of Technology, Szczecin, Poland*

Abstract

The design of recommending systems is mainly related to algorithms targeting customer needs, products, and service selection. However, even the most effective algorithms will not increase the performance of an e-commerce platform if the recommendation interface is not noticed by the user. Recommendations, like all other visual stimuli, can be subconsciously eliminated by users as a result of habituation identified for online environment as a banner blindness. This article presents the results of an experiment, in which the influence of the intensity of graphic elements within the interface on user attention was examined. The study explores the visual space and survival of graphic objects, analysed with the use of survival analysis instead of typical perceptual studies. Results show that an increase in visual intensity does not always lead to an increase in user attention, and survival analysis makes it possible to identify the differences between design variants and between users targeted with the same stimuli.

Keywords: eyetracking, survival analysis, habituation

1. INTRODUCTION

Recent integrated efforts from researchers and practitioners have been focused towards more efficient recommending algorithms and methods. They evolved from early-stage studies, focused on collaborative filtering [18], towards new solutions with the integrated use of explanations [16], context [17], social networks [7], and other techniques [11]. Apart from algorithmic approaches, the studies were focused on recommending interfaces from the HCI perspective [15] and the visual layer, analysed with the use of eye tracking [2]. They showed the importance of the number of presented recommendations, structure of the interface (in terms of presented images), and descriptions.

Even though the accuracy of recommendations grows, due to banner blindness [1] and the habituation effect [14], similar to other repeated content, recommendations within websites may go unnoticed. Banner blindness has been identified for text and other stimuli. Habituation is the main reason for banner blindness phenomena, and is based on a

* Corresponding author. Tel.: +48-91-449-55-93
E-mail address: kbortko@wi.zut.edu.pl

10.1016/j.procs.2020.09.283

lowered response after repeated stimuli. It is especially observed when advertising clutter, with a high number of marketing elements, is used. These observed phenomena create a need for attracting user attention toward recommending interfaces. Several techniques such as increased visual intensity, animations, flickering effects, verbal communication with varying intensity [9] or gradual approach can be used [8]. To avoid over-exposure and negative influences on the user experience, the goal can be a compromised solution, searching for the lowest possible stimuli which overcomes the habituation effect.

In the presented study, elements with scaled visual intensity were used in a typical recommending interface. To compare the effects of different visual techniques, the study proposes an algorithmic approach based on survival analysis. This was used to show how the visual effects used within recommending interfaces affect user attention with low, medium, high, and very high intensities. The results showed that elements using intrusive techniques improved results only to a small extent, when compared to medium levels. The paper is organized as follows: The assumptions are presented within the Conceptual Framework Section. The plan of the experiments, with empirical results, is presented within the Experimental Results Section, and is followed by the Conclusions.

2. The conceptual framework

2.1. General assumptions

The presented study is focused on an analysis of the effect of different visual intensities in a recommending interface on user attention, represented by fixations registered with an eye tracking device. It is assumed that the visual intensity of the recommending interface is scaled from a low level, through medium, and up to high intensity. During the experimental phase, different intensities are used for different users. The total fixation time for each of the users is represented as the survival time of the recommending interface in visual space. The survival of each variant of recommending interface can be analysed, in terms of probability of survival (taking into account the time intervals). Depending on the survival curve, the performance of each technique can be analysed and used to search for compromise solutions with the main goal of increasing performance without decreasing the user experience.

2.2. Methodological background

The presented research utilises survival analysis to obtain probabilities of the expected time duration of focus on the recommendation interface. Survival analysis uses time intervals finished by events [3] - total fixation time is ended when the attention is removed from the recommending interface, in our case. Survival analysis was initially created for medical areas, as a method for breaking down the time between medicinal intercession and passing. In the course of recent decades, the field was extended to incorporate single and multiple events for a given time periods and taking into account individuals, such as customers [4]. With wide areas of applications in the field of advertising or customer-relationship management, increased usage has been observed in recent years [5]. On the off-chance we need to indicate the time taken for an event to happen at T, a histogram and model of the progression of events, in relation to time, can be developed. The likelihood distribution, as a function for T, can be represented by $f(t)$. The aggregate conveyance distribution is denoted by $F(t)$. This gives the accompanying condition: F(t)=p(T≤ t). Utilizing the above methodology, survival is treated as a component of time $S(t)$ with the end goal: For $t = 0$, $S(t) = 1$ for the particular time that a failure happens; the estimation of $S(t)$ is zero [12]. The opportunity of failure won't be detectable, and only incomplete perception will be conceivable. For this situation, we consider a particular censoring time c. The survival is, then, signified as: S(t)=P(T>t)=1-F(t). The conditional failure rate is the rate at which an arbitrarily-chosen individual, who is known to be alive at time $(t-1)$, will pass on at time t [10]. Formally, the instantaneous hazard is equivalent to the quantity of failure between time t and time $t + \Delta(t)$, divided by the measure of the populace in danger (at time t), divided by $\Delta(t)$. This gives the extent of the present populace at time t that come up short, per unit of time, defined by the condition: $\mathrm{h(t)}=\lim_{\Delta t \to 0} \frac{P(t<T\leq t+\Delta(t)|T>t)}{\Delta(t)} = \frac{f(t)}{S(t)}$.

Generally, the Kaplan-Meier approach is utilised to evaluate time-related events [3]. Plotting the Kaplan-Meier curve involves a progression of level strides of declining greatness that, for an adequately substantial example approach, gauge the genuine survival for the given populace. While applying this methodology, an incentive between

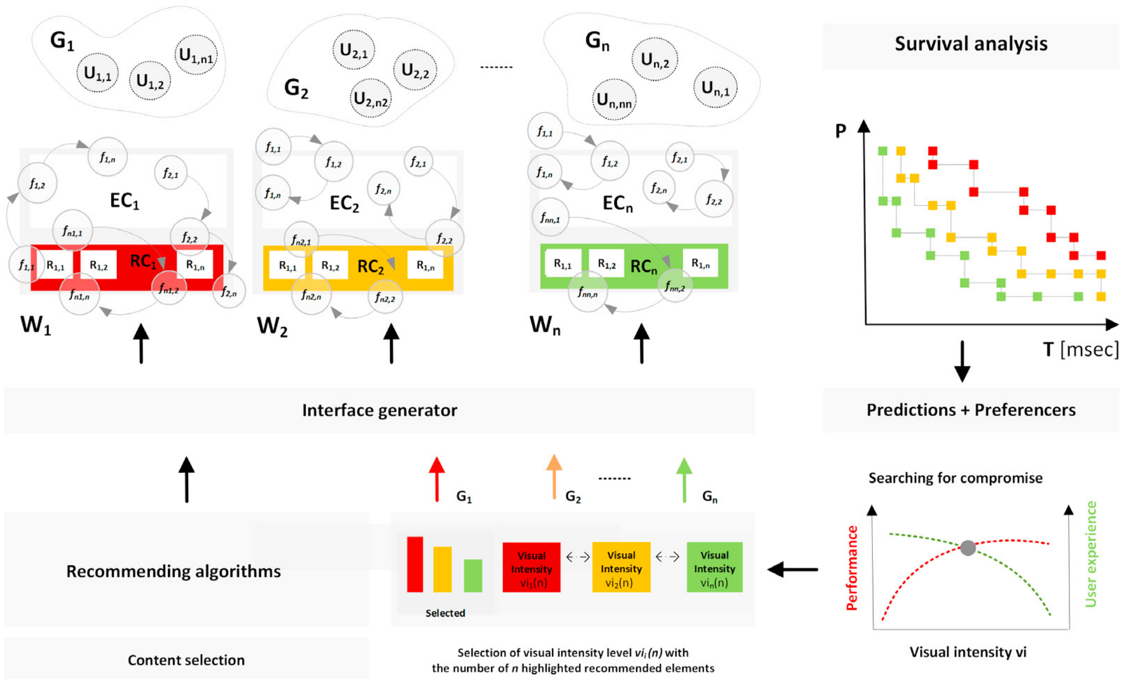


Fig. 1. Model of the system for the evaluation of recommending interfaces with the use of survival analysis and varying visual intensities.

progressive tested perceptions is assumed to be constant [6]. An essential favorable property of the Kaplan-Meier bend is its capacity to consider controlled information inside the example, before the ultimate result is observed.

2.3. System structure

The presented approach assumes that the target users are divided into the groups $G_1, G_2, \ldots, G_n$. Users perform tasks in the website instances W_1, $W_2, \ldots, W_n$ with distinguished editorial content EC_1, $EC_2, \ldots$, EC_n and recommended content RC_1, $RC_2, \ldots, RC_n$. Users within each group are targeted with a recommending interface with different visual intensities, denoted by $vi_1(n_1)$, $vi_2(n_2), \ldots$, $vi_n(n_n)$ and n_i representing the number of highlighted elements. Figure 1 shows an illustrative example, based on three target groups and elements with three levels of visual intensity used; denoted by green for low level intensity, orange for medium intensity, and red for the highest intensity. For each user, an i amount of m_i fixations $f_{i1}, f_{i2}, \ldots, f_{imi}$ is registered with the use of an eye tracker, and they are assigned to areas of interest localized within the editorial content EC and the recommended content RC. For each intensity level, fixations from all users were gathered and processed with the use of survival analysis; it delivers information about the survival probabilities for each time interval. The survival curve can be used to analyse the distance between results from different design variants. A compromise solution is delivered and can be applied for future usage within websites. Visual parameters are assigned to the recommended content obtained from the recommending algorithm.

3. Experimental results

3.1. Plan of the experiment

The experiment consisted of a recommending interface with varying intensity of objects in the area of recommendation. The analysis was based on four levels of visual intensity $vi_i(m)$, where i takes the values from 0 to 3 and n represents the number of highlighted elements. Level 3 meant the maximum intensity of highlighting, through a flickering red frame around the recommended product. Level 2 was highlighted only by a red frame around the recommendation product. The recommended products included a small static bright red element. However, level 0 did not stand out; that is, the products were not highlighted.

Table 1. Intergroup comparison of set-visual intensity of pairs, using the Mann-Whitney U Test.

Juxtaposition of four visual intensity (vi_n)	Time differences		
	U	z	p
vi_0 vs vi_3	802.000	-2.293	0.022
vi_1 vs vi_3	3533.000	-2.791	0.005
vi_2 vs vi_3	4390.500	-0.564	0.573
vi_0 vs vi_2	871.5000	-1.837	0.066
vi_1 vs vi_2	3758.500	-2.20520	0.027
vi_0 vs vi_1	1020.000	-0.863	0.388

Each visual intensity was divided into another four visual variant. Variants that occurred in the experiment are: 1 to 3 (that is, one product highlighted and three not highlighted), 2 to 2, 3 to 1, and 4 to 0 (that is, all products were highlighted). The variants of the highlighted products in heatmaps are shown in Figure 2, Figure 3 and Figure 4. Figure 2 shows visual intensity at the lowest level. Next heatmap in Figure 3 shows visual intensity at the medium level (vi_3). The last, Figure 4 presents visual intensity at the highest level. On the heatmaps it is evident that the level of intensity did not affect users' absorption unambiguously. Higher visual intensity was often much worse compared to the lower visual intensity level. Therefore, our experiment had 19 pages; such that at least one page had one intensity combination with a given sub-level. To detect the impact of the visual characteristics and avoid the impact of content and product type, similar products were recommended. An experiment was based on 24 users (13 men and 11 women) what is typical number of users for attentive behavior analysis [13]. Each respondent watched individual pages with a given variant for 7 seconds, and then another page was shown. In total, 19 pages with different variants were displayed in each sequence. The experiment was carried out in laboratory conditions in daylight. None of the people had visual impairments.

3.2. Overall results

Analysing the time measurements which each user spent on looking at the interface of the page, organised into recompilation, highlighted, and unspecified elements, and first look on the whole interface, one can notice some dependencies. The results of the measurements evidently indicate that vi_3 definitely attracted more user attention. Two of the vi (levels 2 and 3) stand out from the other two. However, at first look, despite appearances, took the longest time for vi_3, amounting to 2451 ms. At vi_2 and vi_1, the time was the lowest—near 2100 ms. Interestingly, vi_0 turned out to be slightly worse, with a time close to 2268 ms. It is almost 200 ms faster than vi_3. The difference in the average time of looking at the recommendation at vi_2 and vi_3 was only 8%. A significant difference was shown by the results of looking at vi_0 and vi_1 , which (up to the highest vi) were 50% and nearly 27%, respectively. We can see that the viewing time was twice as high as at intensity 0.

3.3. Intergroup comparison

In order to compare the individual visual intensity of elements with each other, based on the time spent focusing on the recommendation interface, we used the Mann-Whitney U Test. The analysis is presented in Table 1, comparing the four intensity groups: from 0 (denoting a static level) to 3 (indicating the highest vi). Starting the comparative analysis of the static intensity with the highest visual intensity, we notice the significance of the parameter.

In the case of a dependent variable, when we compare vi_1 and vi_2, we see a similar level of significance. In comparison with vi_1 and vi_3, we see the best level of significance. Based on accurate statistics, we can assume that there are statistically significant differences between the variables. In the case of the variable time, in relation to the remaining pair of groups (i.e., vi_2 and vi_3, vi_0 and vi_2, vi_0 and vi_1), we can see a clear lack of statistical differences, which shows us the significance coefficient (where the p-value >0.05).

3.4. Survival analysis with the use of fixation time

The statistical analysis of the survival was based on four levels of intensity, from 0 to 3, with a breakdown into four variants each. The variable that we will use to analyse the survival is time. In the analysis, we compared four groups of different combinations: Aggregated variants in four vi, and each vi from 1 to 3 with all variants, in relation to level 0, where the variants were static (without any distinction). Each of the graphs shows the probability of survival of a given element, highlighted at different intensities in time.

Fig. 2. Instance of the experimental website with low attention to the recommending interface with it's visual intensity at the lowest level (vi_1).

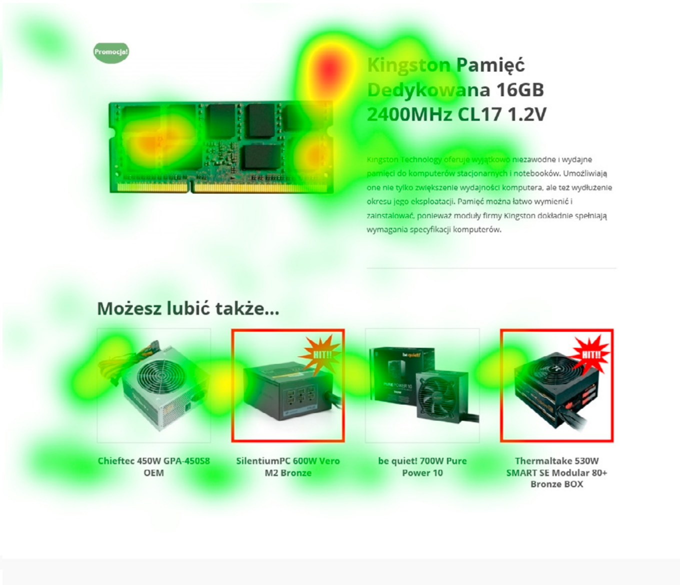


Fig. 3. Instance of the experimental website with medium attention to the recommending interface with it's visual intensity at the medium level (vi_3).

Fig. 4. Instance of the experimental website with high attention to the recommending interface with it's visual intensity at the highest level (vi_2).

Figure 5. shows the probability of survival of the highlighted elements, divided into aggregated intensity groups. The probability in the initial phase of exposure, up to 1 second, spans from 55% (for vi_0) to 80% (for vi_2 and vi_3). The longer the time spent looking at the highlighted elements, the greater the probability of survival we see, for vi_0 and vi_1. The probability of survival of elements for vi_3, at any time, is higher (by a few percentage points) than vi_2; although slightly. In the final phase, we can see a slightly higher probability of survival of the elements of vi_2 (in the phase from 3.5 seconds until the end of the studied period of time).

Figure 6. shows the probability of the survival of the elements highlighted in vi_3, taking into account different variants of the number of elements highlighted at vi_0. The probability in the initial phase of exposure, up to 1 second, stretches from 70% to 85%. $vi_3(3)$ significantly lags behind the rest of the variants (i.e., up to 1 second is about 55%). Subsequently, with an exposure of up to 2.5 seconds, the probability is definitely the highest for $vi_3(1)$, where it is close to 43%. However, for the rest, including for vi_0, it is 10–20%.

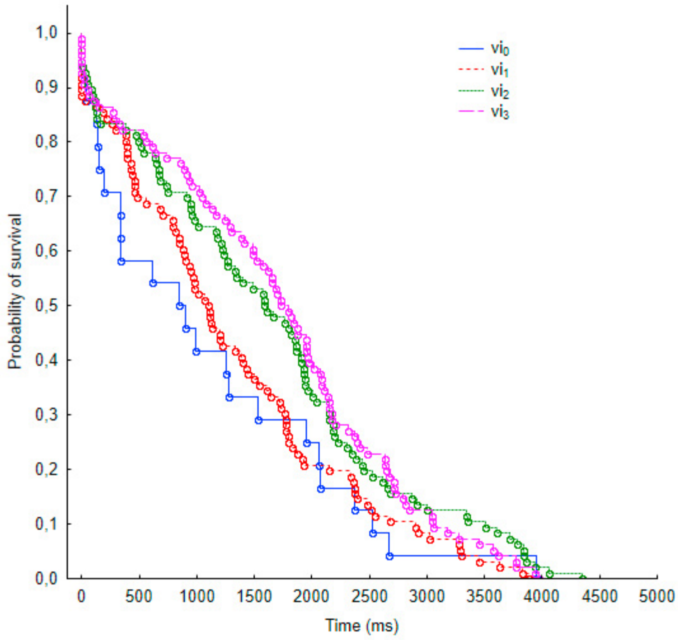


Fig. 5. The Kaplan-Meier survival model for the four groups of *vi*.

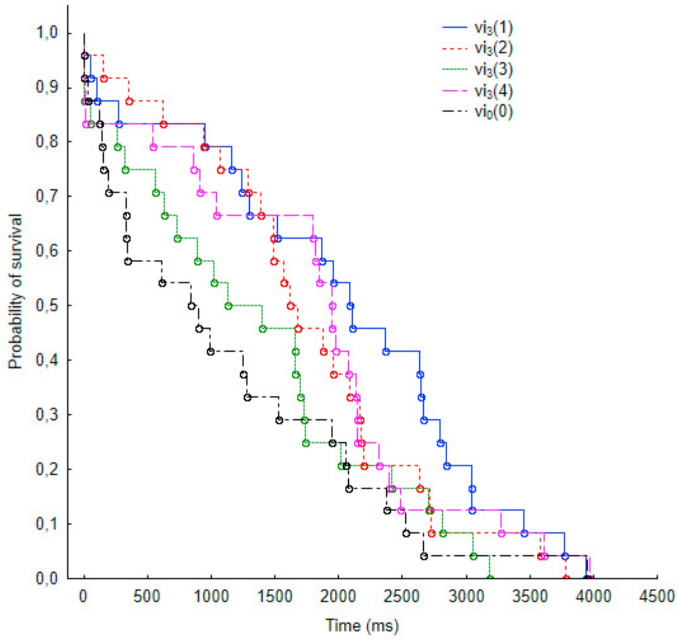


Fig. 6. The Kaplan-Meier survival model for vi_3, with four variants, in comparison with vi_0.

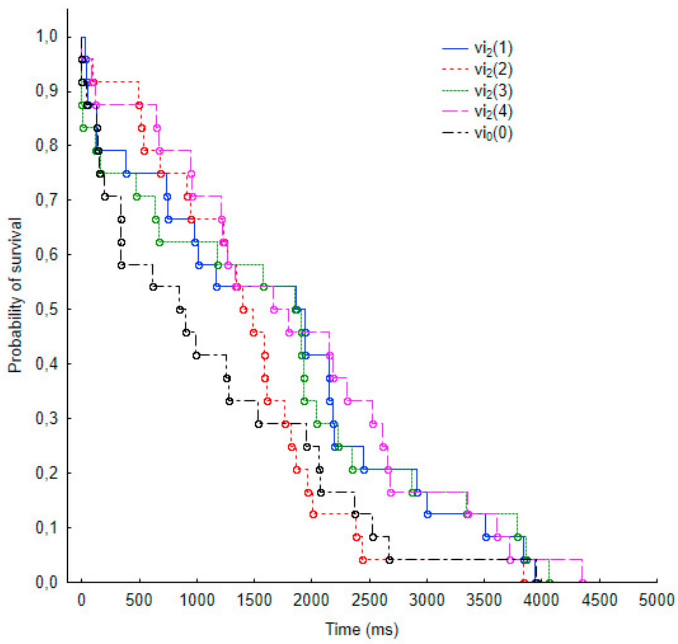


Fig. 7. The Kaplan-Meier survival model for vi_2, with four variants, in comparison with vi_0.

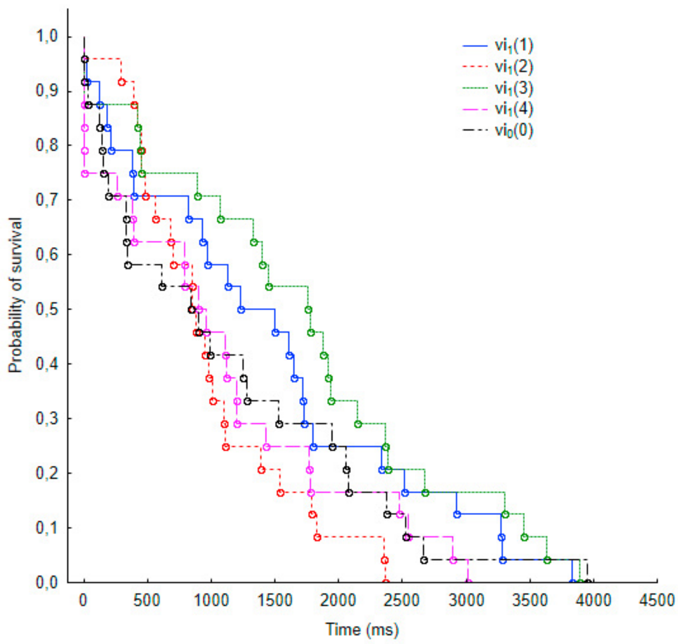


Fig. 8. The Kaplan-Meier survival model for vi_1, with four variants, in comparison with vi_0.

Figure 7. presents the probability of survival of the elements highlighted in vi_2, taking into account different variants of the number of elements highlighted at vi_0. The probability in the initial stage of exposure, up to 1 second, spans from 65% to 80%. The probability of surviving variants at this time does not deviate too much from the public. Next,

with an exposure of up to 2 seconds, the probability is definitely the highest for $vi_2(3)$ and $vi_2(4)$, where it is close to 45%. Here, however, there is a significant decrease in the probability of the $vi_2(2)$, which is close to 10% (i.e., even less than the element in vi_0 with a probability of 25%). In the long-term view of a given element, the probability for each variant oscillates around 15% at 3.5 seconds.

The last figure, Figure 8. shows the probability of survival of the elements distinguished at vi_1. The probability in the initial phase of exposure, up to 1 second, spans from 25% to 75%; which is the largest range of probabilities. The probability of surviving variants, at this time, differs significantly from the previous vi_2. The probability of looking at variations $vi_1(4)$ and $vi_1(2)$, above 1 second, significantly decreases; even from vi 0. The probabilities of looking 1.5 seconds per element for $vi_1(1)$ and $vi_1(3)$ are 50% and 68%, respectively. At more than 2 seconds, the differences in probability are blurred; however, the highest probability of survival of the elements continues to lead the $vi_1(1)$ and $vi_1(3)$.

3.5. Survival analysis with the use of first look time

For the second group of survival analysis, the variable that was used will be the first look, i.e. the time of the first look at a given object starting from displaying the page to the user. First look is also arrival time, time to first sample recorded within the AOI. Here, as in the first variant of survival analysis, we compare 4 groups of combinations. Here, however, contrary to the previous group of survival analyzes, a longer time is not conducive to favorable evaluation. The positive aspect here is the shortest possible time. The faster a given user looked at a given object, the better. It means that the given element had faster absorbing the user's attention.

Figure 9., shows the probability of survival of the elements distinguished in vi_1, taking into account different variants of the number of elements distinguished from vi_0 . The probability in the initial phase of exposure up to 1 second spans from 25 to 41 %, except for $vi_1(2)$, where the probability is close to 55 %. Next, with an exposure of up to 2 seconds, the probability is definitely the highest for the $vi_1(3)$ and $vi_1(2)$, where it is close to 30% and 37% respectively. Here, however, there is a significant decrease in the likelihood of vi_0 . In the longer time of looking at a given element, the probability for each variant oscillates between 9% and 31% for the $vi_1(2)$ at 3.5 seconds. Variant $vi_1(2)$ visibly lags behind the rest of the variants throughout the period maintains a gap of almost 5 percentage points from the rest.

The next, figure 10. shows the probability of survival of elements highlighted in vi_2, taking into account different variants of the number of elements distinguished against vi_0. The probability in the initial phase of exposure to 1 second springs from 83 to 97%. The $vi_2(4)$ and $vi_2(1)$ variants stand out much from the rest of the variants in the initial phase. Subsequent to an exposure of up to 500 milliseconds, the probability definitely drops for each of the variants. A clear division can be seen in around 1 second. Here, we can see that elements in option $vi_2(2)$ have the highest probability of survival within 55%. The remaining variants, along with vi_0 oscillate between 21 and 38%. Surviving in further stages up to 3 seconds evidently evens out. The longest survival time had two variants vi_0 and $vi_2(3)$. However, the longest probability of about 11% was characterized by the $vi_2(4)$.

The next figure 11., however, shows the probability of survival of elements highlighted in vi_3. Probability in the initial phase of exposure up to 1 second spans from 38 to 65%, which is the largest range of probabilities. The probability of variants surviving at this time differs significantly from previous vi_0. The probability of looking at $vi_3(4)$ and $vi_3(3)$ is very similar. Here we see the highest probability of $vi_3(1)$.The probability of looking at about 1.5 seconds for an element for all variants of vi_3 ranges from 39% and 51%. The vi_0 drops drastically to the level of 15% relative to the vi_3 variants. However, up to 3 seconds, the probability begins to compare with the exception of $vi_3(3)$, which holds a distance of several percent. In the final phase, however, we see the variant $vi_3(1)$, whose probability persists the longest. The last figure 12. shows the probability of survival of distinguished elements divided into aggregated intensity groups. The probability in the initial phase of exposure up to 1 second springs from 35% for vi_0, vi_1 and vi_2 up to 55 % vi_3. The longer the time of looking at the highlighted elements, the greater the probability of survival we see for vi_0. In 2 seconds we see that it significantly deviates from the rest and the probability is around 15%. The probability of survival of elements for vi_3 at any point in time is higher by a few percentage points than the rest, although this slightly aligns in 3 seconds. In the final phase we see a slightly higher probability of survival of the elements at vi_3.

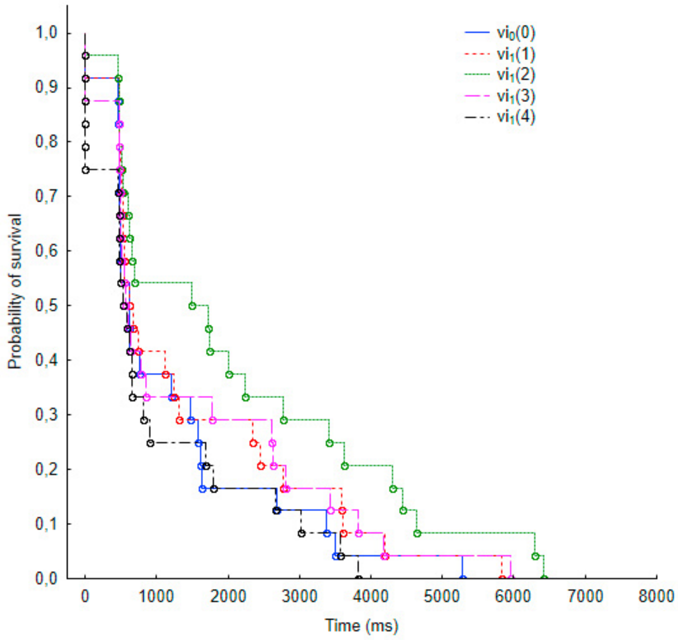


Fig. 9. The Kaplan-Meier survival model for vi_1, with four variants, in comparison with vi_0.

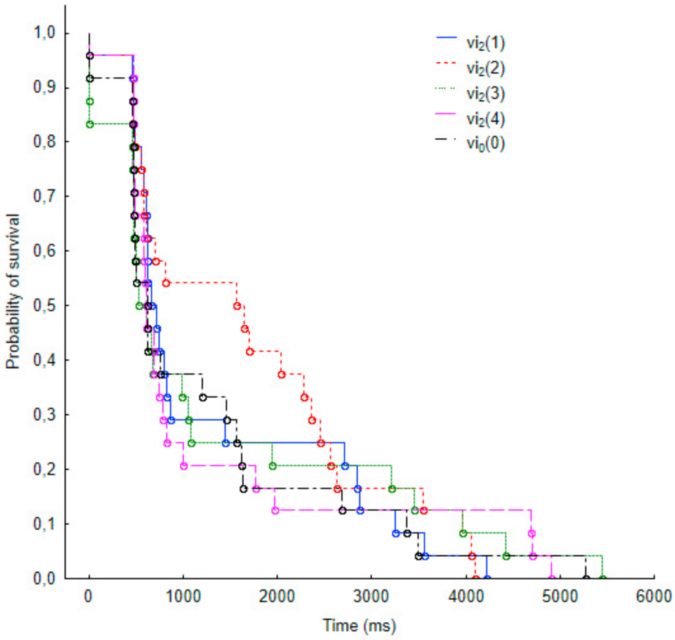


Fig. 10. The Kaplan-Meier survival model for vi_2, with four variants, in comparison with vi_0.

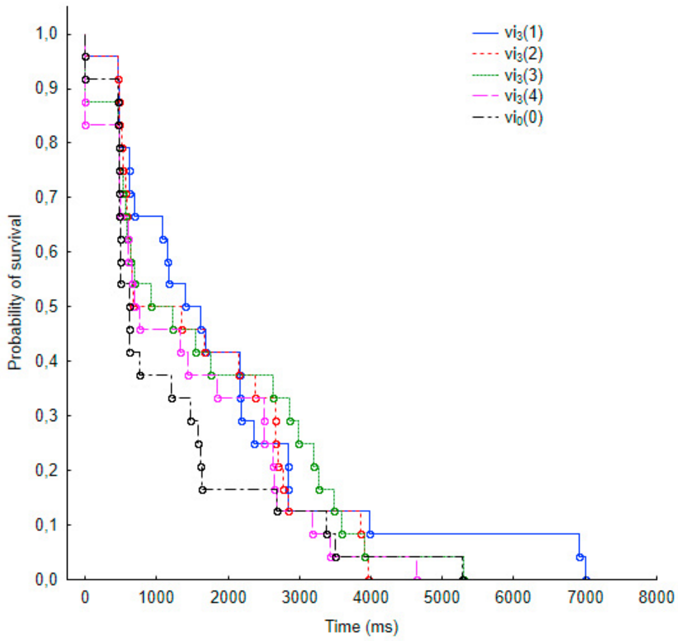


Fig. 11. The Kaplan-Meier survival model for vi_3, with four variants, in comparison with vi_0.

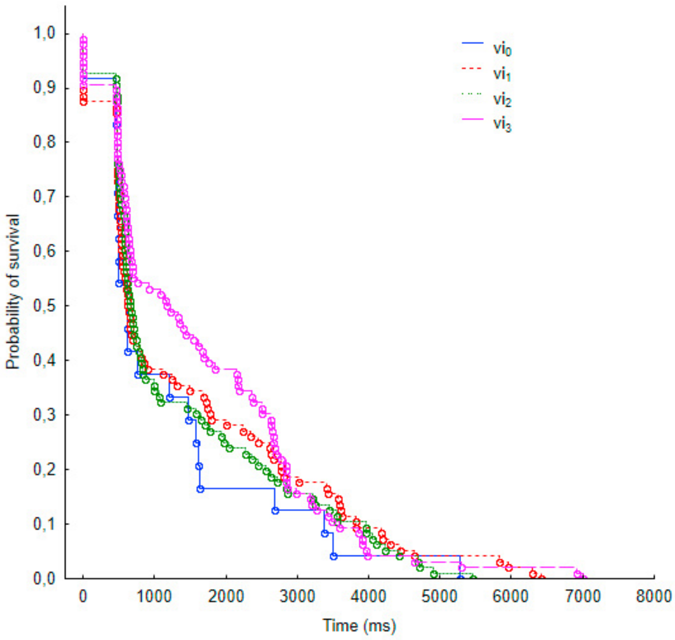


Fig. 12. The Kaplan-Meier survival model for the four groups of vi.

4. Conclusions

Survival analysis makes it possible to identify the number of users, with low and higher attention, focused on the interface within considered time intervals. For interfaces with higher intensity, survival can be analysed by the proportion of users with higher and lower attention among whole analysed groups. The probabilities of high-intensity objects with a larger number of highlighted elements does not guarantee a longer fixation and the user attention.

The results clearly indicate that the intensity level alone is not a sufficient guarantee of a longer fixation of user attention. We do not see any significant differences between vi_2 and vi_3.

On the other hand, analyzing the experience of particular groups or the visual intensity of aggregate length of survival affects the reception of information. Long time means worse result, which means that the user needed more time to look at the element in the visual intensity relationship with the division into visual variant. Similarly to the case of survival analysis focused on the length of time of looking, one can notice that the higher the visual intensity, the result is not necessarily the better result. Visual variant also gives some interesting information such as the fact that the number of highlighted elements does not affect the much better result of fast absorption of the eyesight. The results question here whether a greater number of distinguished elements is definitely necessary and more beneficial in the design of the interface than the distinction of fewer elements.

An increase in the intensity of elements and their number in the recommending interface is not always more beneficial, in relation to intensities at a lower level. From the point of view of the designer of the recommendation interface, it is important to carefully choose the amount of distinguished elements. Therefore, it is not worth focusing on increasing the intensity and increasing the number of highlighted elements, since the differences in perception of users are often very small.

The proposed technique makes it possible to evaluate the effectiveness of individual design variants, in terms of survival in the visual space, and gives new analytic possibilities in the area of effectiveness analysis of recommending interfaces.

The survival analysis used in the study showed the direction of further work. For example other than eye tracking techniques can be used for analysis, such as clicks or mouse movements. The results can be also expanded for a larger sample of users. However, for a pilot study, we limited empirical research to detect key dependencies. The main aspect of the article was to support the analysis of survival as a reliable indicator in this type of research. Another study can be extended towards comparisons with other analytical techniques as well as detailed analysis of customer behavior as a result of attention catching elements with different survival time.

5. Acknowledgments

The work was supported by the National Science Centre of Poland, the decision no. 2017/27/B/HS4/01216.

References

[1] Benway, J.P., 1998. Banner blindness: The irony of attention grabbing on the world wide web, in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, SAGE Publications Sage CA: Los Angeles, CA. pp. 463–467.
[2] Chen, L., Pu, P., 2010. Eye-tracking study of user behavior in recommender interfaces, in: International Conference on User Modeling, Adaptation, and Personalization, Springer. pp. 375–380.
[3] Cleves, M., Gould, W., Gould, W.W., Gutierrez, R., Marchenko, Y., 2008. An introduction to survival analysis using Stata. Stata press.
[4] Drye, T., Wetherill, G., Pinnock, A., 2001. When are customers in the market? applying survival analysis to marketing challenges. Journal of Targeting, Measurement and Analysis for Marketing 10, 179–188.
[5] Fader, P.S., Hardie, B.G., 2009. Probability models for customer-base analysis. Journal of interactive marketing 23, 61–69.
[6] Giot, P., Schwienbacher, A., 2007. Ipos, trade sales and liquidations: Modelling venture capital exits using survival analysis. Journal of Banking & Finance 31, 679–702.
[7] He, J., Chu, W.W., 2010. A social network-based recommender system (snrs), in: Data mining for social network data. Springer, pp. 47–74.
[8] Jankowski, J., Hamari, J., Watróbski, J., 2019. A gradual approach for maximising user conversion without compromising experience with high visual intensity website elements. Internet Research 29, 194–217.
[9] Jankowski, J., Watróbski, J., Ziemba, P., 2015. Modeling the impact of visual components on verbal communication in online advertising, in: Computational Collective Intelligence. Springer, pp. 44–53.
[10] Larivière, B., Van den Poel, D., 2004. Investigating the role of product features in preventing customer churn, by using survival analysis and choice modeling: The case of financial services. Expert Systems with Applications 27, 277–285.
[11] Lu, J., Wu, D., Mao, M., Wang, W., Zhang, G., 2015. Recommender system application developments: a survey. Decision Support Systems 74, 12–32.
[12] Miller Jr, R.G., 2011. Survival analysis. volume 66. John Wiley & Sons.
[13] Ooms, K., De Maeyer, P., Fack, V., 2014. Study of the attentive behavior of novice and expert map users using eye tracking. Cartography and Geographic Information Science 41, 37–54.
[14] Portnoy, F., Marchionini, G., 2010. Modeling the effect of habituation on banner blindness as a function of repetition and search type: Gap analysis for future work, in: CHI'10 Extended Abstracts on Human Factors in Computing Systems, ACM. pp. 4297–4302.

[15] Swearingen, K., Sinha, R., 2001. Beyond algorithms: An hci perspective on recommender systems, in: ACM SIGIR 2001 Workshop on Recommender Systems, Citeseer. pp. 1–11.
[16] Tintarev, N., Masthoff, J., 2007. A survey of explanations in recommender systems, in: 2007 IEEE 23rd international conference on data engineering workshop, IEEE. pp. 801–810.
[17] Verbert, K., Manouselis, N., Ochoa, X., Wolpers, M., Drachsler, H., Bosnic, I., Duval, E., 2012. Context-aware recommender systems for learning: a survey and future challenges. IEEE Transactions on Learning Technologies 5, 318–335.
[18] Yang, X., Guo, Y., Liu, Y., Steck, H., 2014. A survey of collaborative filtering based social recommender systems. Computer Communications 41, 1–10.

A4.

Bortko, K., Jankowski, J., Bartków, P., Pazura, P., Śmiałkowska, B. (2020). Attracting user attention to visual elements within website with the use of Fitts's law and flickering effect. Procedia Computer Science, 176, 2756-2763.

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 176 (2020) 2756–2763

Procedia Computer Science

www.elsevier.com/locate/procedia

24th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems

Attracting User Attention to Visual Elements within Website with the use of Fitts's Law and Flickering Effect

Kamil Bortko[a,*], Jarosław Jankowski[a], Piotr Bartków[a], Patryk Pazura[a], Bożena Śmiałkowska[a]

[a] *Faculty of Computer Science and Information Technology, West Pomeranian University of Technology, Szczecin, Poland*

Abstract

Design of applications requires proper management of content, interaction elements and space. In web systems flow of interactions is often based on attention catching components in a form of elements of navigation or advertising content. Distance from central part of interface is affecting user behavior, what can be overcome by scaling size of the objects within the interface according to Fitts's law. Here we examine the effects of flickering elements with different frequency, direction and distance from central elements of webpage on user attention and habituation reduction. The aim of the study was to examine how to overcome users' resistance to stimuli based on objects flickering with different frequencies located in different directions and distances. We wanted to check if the higher flicker frequency can compensate the larger distance and increase and user focus depending on whether the element is located nearest or farthest, similarly like objects sizes for Fitts's law. Results showed that flicker increases the attention to objects with higher distance but has no effect on nearest objects. I was possible to improve total time attracted by distance objects by twenty percent, however the higher performance was observed to nearest objects, no matter what frequency was used for them.

Keywords: human-computer interaction, Fitts's law, eye-tracking, habituation, web design, online marketing

1. Introduction

Nowadays, the effect of habituation, and dropping user attention to marketing content and elements of websites is a nightmare for many companies and people responsible for marketing performance. Question is how to effectively and non-invasively draw the user's attention to content within websites and applications. Instead of annoying or distracting target users, information that reaches them should signal or unobtrusively draw attention [2]. Users are primarily looking for informational content and those related to current needs and subconsciously skip content that is being repeated. Information not related to the task being performed is also eliminated, e.g. system messages,

* Corresponding author. Tel.: +48-91-449-55-93
E-mail address: kbortko@wi.zut.edu.pl

10.1016/j.procs.2020.09.281

security messages, marketing messages, etc. The main areas of occurrence of problems are related to banner advertising within portals and social platforms, textual advertisements, recommendations in e-commerce and other visual and textual messages. Other areas cover video games and entertainment platforms, security messages in anti-virus systems and messages related to updates [11]. The reason for these phenomena is habituation, which is defined as a cognitive process, consisting in a gradual disappearance of the reaction to a repetitive stimulus, if it does not carry any significant changes (information) [1] [4]. Habituation is associated with the activity of the reticular system. Each new stimulus first produces the stimulating effect of the reticular formation, as the stimulus repeats, the reticular system exerts an inhibitory effect on the transmission of impulses in the associated sensory pathways [14]. The basis of this phenomenon is also the limited ability to process information and allocate attention [6]. Also important are aspects such as: testing the behavior of users during messages appearing in the user interface and analyzing mechanisms of habituation in the interfaces by current and new users.

The goal of the presented study is the analysis of the parameters of visual intensity of objects displayed sequentially in the process of using the interactive system and the recipient's reactions recorded using an eye-tracker. In the experiment, theoretical background based on Fitts's law was used to examine the role of distance from the central part of the interface and role flickering frequency on attracting user attention. Apart from this we study the differences in increasing and decreasing intensity for objects displayed in individual stages of the sequence. The analysis of the impact of graphic object parameters on their durability in visual space with the use survival models was performed. Results showed the moderate effect of flickering for objects located near to the central scene, however performance higher distance was improved. In general it shows that marketers should add used techniques to the location of objects within visual space and intensive technics are not always required to attracts users attention.

2. Review of the Literature

Designers use various techniques to attract users attention to overcome banner blindness effect [2]. They include animations, elements with high intensity. They often take a form of intrusive advertising negatively affecting marketing content perception. The results suggest that animation in banner ads does not necessarily increase the user's attention, but even if the user does not knowingly notice the banner ad, it affects the user's attitude to the brand [7]. Literature deals with this topic very early. To tackle the topic, a literature review with a background of three main currents: habituation, Fitts's law and frequency was performed. The subject of habituation is taken up by many publications. The most commonly cited descriptions of the behavioral characteristics of habituation come from two papers published almost 40 years ago [4]. The results also show that the animation improves the rollback performance of banner content. The subject of the ad, the position of the banner, as well as the inscriptions and colors are better recalled when the banner is animated in contrast to the non-animated banner, while the intensity of the animation does not affect the rollback performance associated with the banner. Importantly, the banner animation does not affect the performance of the actual information retrieval task. What's more, the intensity of the animation does not affect the attitude of the respondents to advertising banners [5]. Banner blindness is strongest for banner ads on the right side of the page and goal-oriented tasks [12]. The results showed that the ads were explicitly visited while reading and that the advertising times were the longest when the above advertisement was static and the second advertisement was animated. The results showed that attention (i.e., the time when the eyes first entered the ad) was related to the time of its appearance. This is especially the case for the ad on the right, which indicates that ads displayed near the text area are paying particular attention [13]. Most advertisers and web designers believe that animation has a "magic" function that attracts attention, improves memory and leads to more favorable attitudes [16]. Some studies show that animation does not help to increase attention, especially when participants perform search tasks [8]. Apart from the visual characteristic of content, other parameters like size and location have effect on user attention and overall performance. Fitts's law is used to study the impact of measuring attention absorption time relative to the distance of an object. In experiments, Fitts' Law is used, supported by eyetracking research [15]. Fitts's law predicts average statistical effectiveness of pointing, but it doesn't tell how a person can cope with a single act of pointing. Finally, practical relevance relates to the importance of these design decisions and events for designers. Fitts law becomes important when improving the commonly used interface for many experienced users. It allows to mathematically predict the time it takes a person to indicate a goal - either by touching it or by pointing it with the cursor, depending on the distance to the goal and the size of the goal [9]. However, some recent findings have shown that larger targets

Table 1. Global ANOVA analysis statistics showing effect of flickering for all objects.

	F	p-values
nearest	0.15	0.85
medium	0.54	0.57
farthest	2.87	0.05

have shortened acquisition times [10]. In summary, thanks to the use of an eye tracking system, this study contributes to important findings regarding the effectiveness of banners, especially with regard to the impact of banner format and animation on the users' attention. Both banner format and animation had a significant impact on attention [7].

3. Goals and plan of experiment

The experiment was based on user sessions with played movie and displayed around objects with different sizes and flickering effects with frequency 7 Hz and 15 Hz. As a reference was used object without flickering effect, denoted as 0 Hz. Presented in this paper study was based on 20 participants taking part in the experiment in same conditions in daylight. Number of participants is similar to other eye tracking based studies focused on visual attention and display blindness[3]. Each of them was tested using the Tobi Pro X3 eye-tracker. Around the film, in 12 places, as shown in the Fig. 1 ads were displayed. Each of them was displayed for 2 seconds with an interval of 1 second. People had previously undergone calibration, which allowed reliable measurement and elimination of divergent results. The objects were displayed in four different sequences. In Seq. 1, the same advertisement was displayed in each individual place, but with a different frequency of flickering. Seq. 2 displayed various advertisements with different flicker frequencies. Both of these sequences proceeded clockwise, i.e. at the beginning the advertisement flickered. In Seq. 3 proceeded analogously to seq. 1, as well as Seq. 4 proceeded adequately to Seq. 2. The difference was only that the ads were displayed in a random way, i.e. each of the ads was displayed in a different place with three different frequencies i.e. 0 Hz, 7 Hz and 15 Hz.

4. Empirical results

4.1. Empirical results

In the next step ANOVA analysis was used to carry out the significance test for the dependent variable, the time of looking at the object we used a qualitative variable as is the flicker frequency in Hz. Results are presented in Tab. 1. We divided the analysis according to the distance of the object from the central part of the page where the film was located. Analyzing the obtained data, we can see that the statistical significance of the qualitative variable in the built statistical model was reached only at the distance "farthest". The significance was p is higher then 0.05. The strength of the coefficient was 2.87, which indicates a strong relationship with the dependent variable, i.e. the time of looking. The other two variants for "nearest" and "medium" distances do not show significance for the variable flicker frequency.

We could see here very low significance. For the "medium" variant it is 0.58, which indicates a poor fit and low significance of this variable to the model. The strength of the coefficient is only 0.54. However, the significance for the "nearest" variant is in this case the worst and amounts to 0.87, i.e. the result is close to 1. The fit force is 0.15, which means that the result is also very poor. Using only one variable, which is the flicker frequency, we see an upward trend in significance in favor of the "farthest" variant, where only this value alone has reasonable results. This result proves that it is better for the user to place graphic elements or objects further away if they want to increase the time of looking at them.

4.2. Descriptive global statistics

Analyzing aggregate results divided into several separate groups and global results in Tab. 2 and Tab. 3, we see some tendencies and relationships. We've grouped the results by several variables. The first of these are the values

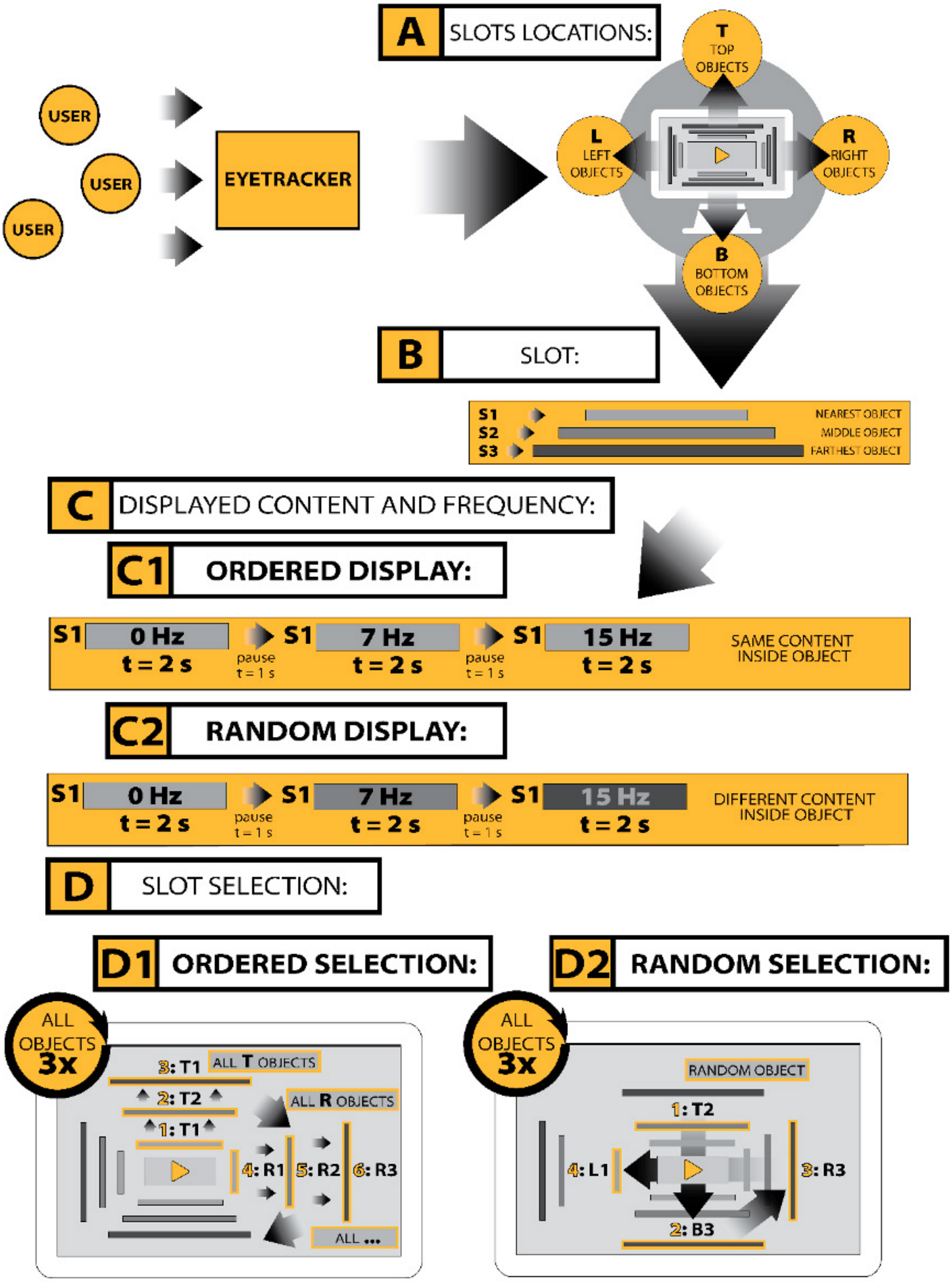


Fig. 1. Conceptual plan of the experiment. Part A shows locations of free slots, used to display content: top, right, bottom and left groups of objects. Part B shows that every group of slots contains three different slots, near central part of the screen, in the middle and far from it. Part C shows two ways of displaying content inside the slot. C1 presents ordered display (same content for every frequency inside selected slot) and C2 presents random displaying (different content for every frequency inside the slot). Part D shows slots selections. D1 represents ordered selection type, and D2 represents random selection.

Table 2. Aggregated looking time for used locations for each flickering frequency.

	Avg. Looking time	Total looking time	Avg looking time			Total looking time		
			0 Hz	7 Hz	15 Hz	0 Hz	7 Hz	15 Hz
Top	81.18	46758	67.23	105.23	71.07	12909	20204	13645
Bottom	76.89	44289	59.02	81.82	89.83	11332	15710	17247
Left	91.93	52953	96.43	90.30	89.07	18515	17337	17101
Right	95.87	55219	82.75	99.82	105.03	15888	19166	20165

Table 3. Performance for used locations for each flickering frequency.

	15 Hz vs 0 Hz	15 Hz vs 7 Hz	7 Hz vs 0 Hz
Top	1.06	0.68	1.57
Bottom	1.52	1.10	1.39
Left	0.92	0.99	0.94
Right	1.27	1.05	1.21

of the average time of looking at objects located in three distances: nearest, medium, farthest with the division into object directions: top, bottom, left and right. The above division into flicker frequencies was isolated: 0 Hz, 7 Hz, 15 Hz and all frequencies combined. Starting from the frequency value together, in close proximity to the central part of the interface, we see a definite advantage of the average time of looking at an object from left to right. The results of these two directions do not differ much and amount to 133 and 137 ms, respectively. The other two, i.e. the top and bottom are 64 and 69 ms, which is almost half the time of looking at the object. With the medium variant, the trend looks barely different. Here, the upper direction result is 118 ms. The other three directions oscillate within 90 ms, which is about 30 ms worse. In the farthest variant, the situation is different. Here, the down direction clearly stands out against the background of the other directions and is close to 72 ms, with an average viewing time of about 60 ms for the remaining three directions. The total average time of looking at three distances is best for left and right directions, are 92 and 96 ms, respectively. Much worse results are top and bottom directions, where their results came out respectively: 82 and 77 ms, or about 10-14 ms shorter average viewing time.

By breaking down the results into groups of distances in relation to the direction for individual frequencies, we see some patterns and relationships. We can see that in Tab. 4 and Tab. 5. Starting from the results for 0 Hz we see that the average time of looking for ” nearest” distances are much higher than. The results for the ”medium” distance are slightly worse, without one case where the result from the ”left” direction is more than twice as bad. The results for “farthest” are the worst results. They oscillate around 20 to 40% of the ”nearest” distance results. The situation is slightly different for the 7 Hz frequency, where the results for the ”near” distance are clearly closer to the ”farthest” distance value with one exception - the ”right” direction (50% better in favor of the ”nearest” distance). The average viewing time for the ”medium” distance differs significantly in the ”top” direction, where it is 140 ms, with values of about 90 ms for the ”nearest” and ” farthest ” distances. At 15 Hz, there is a clear difference between the directions for each distance. For the ”left” and ”right” directions, the average viewing time is nearly three and four times higher. The reverse trend occurs, however, for the ”medium” distance. Here you can see a better result for the ”top” and ”bottom” directions, but this is only a difference of about 20%. The values for the distance ”farthest” are similar in each direction. However, analyzing the distances themselves, it can be seen that the values of the average time of looking for ”near” are much better. By the ”left” and ”right” directions we see a result two and three times better.

4.3. Sequence analysis

Tab. 6 and Tab. 7 presents the results of sequence analysis by distance and frequency. The values for average look time and increases are given there. Analyzing each sequence individually, we can see the division into distances - ”nearest”, ”medium” and ” farthest ” as well as the flicker frequency, i.e. 0 Hz, 7 Hz and 15 Hz, we see some tendencies and relationships. Practically in every sequence and at every frequency, the average viewing time was the largest at the ”nearest” distance. When analyzing globally, we see that increasing the frequency did not significantly affect the user’s focus on a given object. In each variant, the advantage of looking at the user at the object in the

Table 4. Aggregated looking time for used locations for each flickering frequency divided by objects.

	Overall			0 Hz		
	nearest	medium	farthest	nearest	medium	farthest
Top	64.22	118.92	60.40	73.36	94.81	33.53
Bottom	69.12	89.99	71.56	70.00	61.03	46.03
Left	133.13	86.28	56.39	181.50	70.84	36.95
Right	137.42	91.57	58.61	104.63	101.92	41.70
Avg values	100.97	96.69	61.73	107.37	82.15	39.55
	7 Hz			15 Hz		
	nearest	medium	farthest	nearest	medium	farthest
Top	80.53	140.68	95.18	38.77	121.55	51.79
Bottom	70.39	84.42	90.66	66.97	124.53	77.98
Left	87.55	104.67	78.67	130.34	83.33	53.53
Right	142.94	81.63	74.91	164.70	91.16	59.22
Avg values	95.35	102.85	84.85	100.19	105.14	60.63

Table 5. Performance for used locations for each flickering frequency

	15 Hz vs 0 Hz			15 Hz vs 7 Hz			7 Hz vs 0 Hz		
	nearest	medium	farthest	nearest	medium	farthest	nearest	medium	farthest
Top	0.53	1.28	1.54	0.48	0.86	0.54	1.10	1.48	2.84
Bottom	0.96	2.04	1.69	0.95	1.48	0.86	1.01	1.38	1.97
Left	0.72	1.18	1.45	1.49	0.80	0.68	0.48	1.48	2.13
Right	1.57	0.89	1.42	1.15	1.12	0.79	1.37	0.80	1.80

Table 6. Sequence analysis.

flickering in Hz	location	Seq. 1	Seq. 2	Seq. 3	Seq. 4	Avg values
0 Hz	nearest	136.38	130.50	72.89	89.72	107.37
	medium	77.36	71.34	59.19	120.72	82.15
	farthest	50.64	23.05	21.94	62.59	39.55
7 Hz	nearest	87.59	104.81	66.34	122.66	95.35
	medium	112.02	151.19	40.23	107.44	102.72
	farthest	78.12	82.16	71.39	108.02	84.92
w15 Hz	nearest	105.48	115.94	92.09	87.27	100.20
	medium	93.16	98.11	79.09	149.77	105.03
	farthest	51.97	54.34	41.61	94.61	60.63

”nearest” position is visible. The sequences were also divided into two main groups, static, where the same ads were displayed, at different frequencies, in the case of Seq. 1 ordered and Seq. 3 in random positions. The average view time here was 74 ms. The second group consists of Seq. 2 and Seq. 4, i.e. those where there was variability of ads, i.e. different ads with different frequencies appeared in a given object. Similarly to the first group, Seq. 2 proceeded clockwise, while Seq. 4 at random positions. The average look time in this case was 98 ms, which is close to 24 ms more, which gives an approx. 25% increase relative to the sequence from the first group.

5. Conclusions

The presented study has the form of preliminary research with the main purpose to examine the effects of different intensity of visual elements in addition to the size and distance of objects in accordance to Fitts law. The research was intended to show a certain path for future internet system designers placing graphic elements and choosing their intensity.

Table 7. Performance.

flickering in Hz	location	Seq. 1	Seq. 2	Seq. 3	Seq. 4	Avg values
15 Hz vs 0 Hz	nearest	0.77	0.89	1.26	0.97	0.97
	medium	1.20	1.38	1.34	1.24	1.29
	farthest	1.03	2.36	1.90	1.51	1.70
15 Hz vs 7 Hz	nearest	1.20	1.11	1.39	0.71	1.10
	medium	0.83	0.65	1.97	1.39	1.21
	farthest	0.67	0.66	0.58	0.88	0.70
7 Hz vs 0 Hz	nearest	1.54	3.56	3.25	1.73	2.52
	medium	1.20	1.11	1.39	0.71	1.10
	farthest	0.83	0.65	1.97	1.39	1.21

The intensity of flickering graphic elements as well as the direction of their arrangement does not always improve the result of its absorption by the user. The results show that the intensity of the the flickering effect can improve the focus on the subject in the peripheral view with a greater distance from the central scene. However, it is not necessary to apply this type of effect to objects available in central views. It seems more beneficial to place objects closer together without displaying them by flickering. The increase in flicker frequency is also not conducive to much better results of the average time of looking at the object.

In terms of directions, it is obvious that greater focus and average viewing time are best for left and right directions. Analyzing globally, the average viewing time shows that the further the object is at an increased frequency, the more it does not necessarily improve the average viewing time. It has been found that it is not profitable to increase the flicker frequency of objects. The smaller the distance, the more it compensates for the length of the eye attraction. The ”average” distance values slightly exceed the values of the average ”nearest” viewing time, but not in every variant and not in all directions. From this it follows that the place of its occurrence is not so important to the user.

When analyzing the cost aspect, web designers and internet systems designers should pay attention to the appropriate selection of the intensity of graphic elements. More elements and their significant distinction, as can be seen from the study, does not mean an improvement in the user’s eye absorption.

Therefore, it is worth considering whether it is better to place static objects without any distinction by flickering and better focus on distances. It is a kind of flicker frequency compensation as opposed to the direction and distance of placing objects. Using only one variable, which is the flicker frequency, we see an upward trend in favor of the ”farthest” variant, in which only the value itself has reasonable results.

This result clearly proves that it is better for the user to place graphic elements or objects further if they want to extend the time of their viewing. So before designing, it is worth thinking carefully about the layout of graphic elements and choosing their intensity, so that it is not too glaring and at the same time attracts the attention of people in a gentle way.

6. Future work

The next steps of future works can be based on conducting an experiment strictly on the Fitts law, using objects with different sizes, locations and more levels of intensity. This will allow a broader view of the complexity of the habituation problem in combination with the placement of given objects in appropriate distances and directions. Extended study can be based on a larger sample of respondents tested within different conditions such as time of day or type of lighting.

Acknowledgements

The work was supported by the National Science Centre of Poland, the decision no. 2017/27/B/HS4/01216.

References

[1] Basuki, T.A., Cerone, A., Griesmayer, A., Schlatte, R., 2009. Model-checking user behaviour using interacting components. Formal aspects of computing 21, 571.

[2] Benway, J.P., 1998. Banner blindness: The irony of attention grabbing on the world wide web, in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, SAGE Publications Sage CA: Los Angeles, CA. pp. 463–467.

[3] Dalton, N.S., Collins, E., Marshall, P., 2015. Display blindness? looking again at the visibility of situated displays using eye-tracking, in: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, pp. 3889–3898.

[4] Groves, P.M., Thompson, R.F., 1970. Habituation: a dual-process theory. Psychological review 77, 419.

[5] Hamborg, K.C., Bruns, M., Ollermann, F., Kaspar, K., 2012. The effect of banner animation on fixation behavior and recall performance in search tasks. Computers in Human Behavior 28, 576–582.

[6] Lane, S.T., Franklin, J.C., Curran, P.J., 2013. Clarifying the nature of startle habituation using latent curve modeling. International Journal of Psychophysiology 88, 55–63.

[7] Lee, J., Ahn, J.H., 2012. Attention to banner ads and their effectiveness: An eye-tracking approach. International Journal of Electronic Commerce 17, 119–137.

[8] Li, K., Huang, G., Bente, G., 2016. The impacts of banner format and animation speed on banner effectiveness: Evidence from eye movements. Computers in Human Behavior 54, 522–530.

[9] MacKenzie, I.S., 1992. Fitts' law as a research and design tool in human-computer interaction. Human-computer interaction 7, 91–139.

[10] Michalski, R., 2008. The effects of panel location, target size, and gender on efficiency in simple direct manipulation tasks, in: Human Computer Interaction. IntechOpen.

[11] Owens, J.W., Chaparro, B.S., Palmer, E.M., 2011. Text advertising blindness: the new banner blindness? Journal of usability studies 6, 172–197.

[12] Resnick, M., Albert, W., 2014. The impact of advertising location and user task on the emergence of banner ad blindness: An eye-tracking study. International Journal of Human-Computer Interaction 30, 206–219.

[13] Simola, J., Kuisma, J., Öörni, A., Uusitalo, L., Hyönä, J., 2011. The impact of salient advertisements on reading and attention on web pages. Journal of Experimental Psychology: Applied 17, 174.

[14] Soro, A., Brereton, M., Roe, P., 2016. Towards an analysis framework of technology habituation by older users, in: Proceedings of the 2016 ACM Conference on Designing Interactive Systems, pp. 1021–1033.

[15] Vertegaal, R., 2008. A fitts law comparison of eye tracking and manual input in the selection of visual targets, in: Proceedings of the 10th international conference on Multimodal interfaces, pp. 241–248.

[16] Yoo, C.Y., Kim, K., 2005. Processing of animation in online banner advertising: The roles of cognitive and emotional responses. Journal of Interactive marketing 19, 18–34.

A5.

A6 Bortko, K., Bartków, P., Jankowski, J., Kuras, D., Sulikowski, P. (2019). Multi-criteria evaluation of recommending interfaces towards habituation reduction and limited negative impact on user experience. Procedia Computer Science, 159, 2240-2248.

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 159 (2019) 2240–2248

Procedia Computer Science

www.elsevier.com/locate/procedia

23rd International Conference on Knowledge-Based and Intelligent Information & Engineering Systems

Multi-criteria Evaluation of Recommending Interfaces towards Habituation Reduction and Limited Negative Impact on User Experience

Kamil Bortko[a,*], Piotr Bartków[a], Jarosław Jankowski[a], Damian Kuras[a], Piotr Sulikowski[a]

[a] *West Pomeranian University of Technology, Faculty of Computer Science and Information Technology, ul. Zolnierska 49, 71-210 Szczecin, Poland*

Abstract

The growing intensity of visual marketing content may result in limited user attention and dropping performance. The main reason is a habituation effect observed for repeated stimuli causing the banner blindness phenomenon. To overcome that marketers resort to various techniques based on visual effects, such as animations or vividness, with possible side effects like negative brand perception. The same can be observed for recommending systems and the presented study investigates varying visual intensity in a recommending interface with the use of eye tracking. Multi-criteria analysis is performed to evaluate results from the perspective of user experience, attention allocated to visual objects and their intensity with the main goal to recommend the most viable solutions.

Keywords: recommending systems, recommendation interfaces, multi-criteria analysis, eye tracking

1. Introduction

Rapid growth in e-commerce sector creates demand for analytical methods used for increasing performance of trading platforms. Online systems gather data about customer preferences, online behaviors and transactions. They are are used to improve user experience and increase sales volumes. Recommending systems integrated within e-commerce platforms help to better target products according to user needs and reduce efforts to find required products and services. In an online environment it is mainly those systems that substitute a salesperson known from the traditional environment and make up for the lack of personal contact since they allow for presenting items that might be of

* Corresponding author.Tel.: +48-91-449-55-93
E-mail address: kbortko@wi.zut.edu.pl

10.1016/j.procs.2019.09.399

interest to the customer. Recommending systems play an important role in purchase decision making [2]. Early stage solutions in this field based on collaborative filtering algorithms [19] were extended towards explanation interfaces [16] with the use of context [17] and other techniques [10]. Another direction is integration and utilization of data from social media [5].

While algorithms development is required for proper content selection and targeting, other steps of communication with consumers are based on human-computer interaction with recommender interfaces, which can be analyzed with DOM-events-based solutions [14] as well as eye tracking. The HCI perspective [15] and the visual layer of recommending interfaces attracted researchers' attention willing to better integrate recommending interfaces within online platforms [3]. The number of recommendations, the associated images, descriptions and layouts etc. can all be analyzed towards interface optimization.

Recommendation interfaces are elements of marketing strategies and can be treated as a means of marketing content distribution. Unfortunately, even though they are focused on improving user experience, they can suffer from banner blindness phenomenon [1] and, thus, be ignored to some extent by online consumers. Apart from visual content, banner blindness was also identified with regard to texts and other stimuli [11]. Due to habituation effects repeated content may be not noticed and even best recommendations may have no effect [12]. It is especially observed when advertising clutter [4] with a high number of marketing elements is used. It results in unconscious ignorance of marketing objects and is observed especially for task-oriented behaviors of web users [8].

To avoid the blindness marketers often use techniques based on increasing visual intensity [6] of presented objects with the use of animations and flickering effects [9]. At the same time, however, this increases content intrusiveness and may induce irritation and other side effects affecting brand perception [13]. This situation justifies the need for more balanced solutions to overcome habituation effects without negative influence on web users [7].

The presented study concentrates on multi-criteria decision analysis (MCDA) of features of recommending interfaces taking into account their visual intensity, attention represented by fixations measured with eye tracking and time required to attract attention after a website is loaded. Our experiments conducted with varied recommendation presentations showed that highest presentation intensity does not necessarily trigger increased attention. The habituation effect can best be reduced with moderate intensity. Multi-criteria methods can be used to better adjust presentation strategies to preferences, taking into account short- and long-term commercial goals and positive user experience.

This paper is organized as follows. Assumptions for the study and its conceptual framework are presented in Section 2. The structure of our experiment and empirical results are provided in Section 3 and conclusions are presented in Section 4.

2. Conceptual Framework

2.1. General Assumptions

The main goal of the presented study is to evaluate different strategies used to attract user attention to the recommending interface from the perspective of commercial goals and user experience. Commercial goals can be fulfilled by attracting user attention in order to recommend certain products while from the user experience perspective a task-oriented user is more interested in editorial content and their own targets. The methodology of our research assumes the use of a multi-criteria method for strategy evaluation and user behavior analysis with eye tracking.

The experimental group of users can be divided into subgroups $G_1, G_2, \ldots, G_n$. Several design variants of a recommendation interface with recommended content $RC_1, RC_2, \ldots, RC_n$ with different intensities are prepared and presented within experimental websites $W_1, W_2, \ldots, W_n$. Together with recommendations a dedicated section with editorial content $EC_1, EC_2, \ldots, EC_n$ is presented. For each subgroup of users a different visual intensity level of the recommending interface is selected, denoted by vi_1, $vi_2, \ldots, vi_n$.

An example based on three subgroups of users is presented in Fig. 1. Three levels of visual intensity are used and they are denoted by green color for low level intensity, orange color for medium intensity, and the highest intensity is represented by red color. For each user within subgroup m_i fixations are registered, represented by $f_{i,1}, \mathrm{f}_{i,2}, \ldots, f_{i,mi}$. Fixations are aggregated within areas of interest – website sections defined as editorial content EC or recommended content RC. Results for each design variant are used as input for the multi-criteria model. Searching for compromise

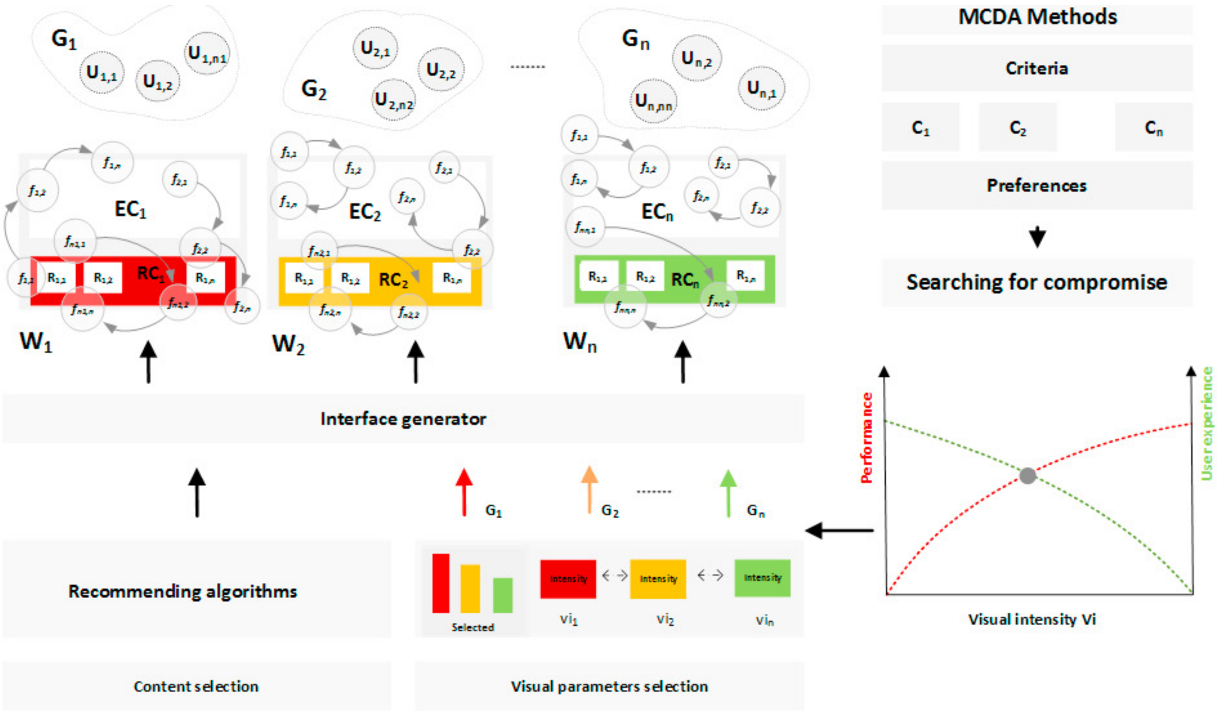


Fig. 1. Model of the system for the evaluation of recommending interfaces with the use of MCDA methods and varying visual intensities.

solutions is performed considering visual intensity, user attention to the recommending interface and to editorial content. Several scenarios will be analyzed using different criteria and goals.

2.2. *Methodological Background*

Multi-criteria analysis of results can provide guidance for future development planning. This leads to multi-criteria problems with preferences assigned to the assessment measures. Various methods can be used to evaluate the strategy and rank the results [18]. The PROMETHEE method was chosen in our research study. It provides an opportunity to effectively combine criteria and find an advantage between different strategies. It uses a set of solutions, in our case the possible strategies are A, each $a \in A$, and $f_j(a)$ represents the solution rating a for the given criterion f_j. The $P_j(a,b)$ preference function represents the degree of a preference over b solution for the given f_j evaluation criterion. A preference index is used for multiple criteria (a,b) from a above b and includes all k criteria: $\pi(a,b) = \sum_{j=1}^{k} w_j P_j(a,b)$, where w_j means the criterion weight. The positive ϕ+(a) and negative outranking ϕ-(a) are used for a meaningful assessment of the chosen strategy and then for creating the final ranking.

3. Experimental Results

3.1. *Structure of the Experiment*

The experiment consisted of a recommending interface with varying intensity of objects in the recommendations area of the website. The analysis was based on four levels of visual intensity vi, from 0–3. Each level was defined as follows: in level 0, the elements in the recommendation interface were static and did not have any distinctions; in level 1, the element had a small red object in the upper right corner; in level 2, the elements had a red frame with a small red object in the right corner; the level 3 component was highlighted by a red frame and a red flickering object. Each intensity level was divided into another four sub-levels taking into account the number of highlighted items. Variants that occurred in the experiment were: 1 to 3 (that is, one product was highlighted and three not highlighted), 2 to 2, 3 to 1, and 4 to 0 (that is, all products were highlighted). Experiment had 19 web pages altogether so that at least one

Table 1. The data that define individual variants. Time values are given in milliseconds.

Variants	Avg time of looking at the interface	Avg time of looking at the area of recommendation	Visual intensity level	First look
VI0	1276 ms	239 ms	0	2268 ms
VI3	1287 ms	350 ms	3	2451 ms
VI2	1345 ms	322 ms	2	2159 ms
VI1	1293 ms	251 ms	1	2131 ms

page had one intensity combination with a given sub-level. An experiment was carried out with 24 participants. Every person watched individual pages with a given variant for 7 seconds. In the next step, another page was shown. All in all, there were 19 subsequent pages with different recommendation variants displayed to each of the participants.

3.2. Results and Visual Analysis

The gaze maps are presented in Fig. 6 - Fig. 9 and represent the most significant figure type popular in eye tracking literature. Looks are the basic unit of measurement – one point of view is equal to one raw sample registered by the eye tracking device. Fig. 6 shows the first look, which focuses on the upper part of the interface. In the first place, a large concentration on the area outside the recommendation interface is observed. Then we can see the attention given to objects in the recommendation interface, as evidenced by large fixation. Fig. 7 shows the first look, which also focuses on the upper interface. This time, however, there is little interest in the recommendation area. The user's attention is definitely focused on the product description. Fig. 8 shows the first look in the vicinity of the recommendation area – the highlighted elements of the interface have focused user attention, which is clearly visible. Fig. 9 shows slight attention of the user to the recommendation area, which we can notice after the first look and the user's path of vision mainly focused on the upper part of the interface.

In short, heatmaps shown in Fig. 10 - Fig. 13 show relative intensity of the values in the table. This means that we have a large number of numbers and each receives a graphical representation. Those areas that have the highest values – in relation to other numbers present – will get a "hot" color, while those with lower values - again, in relation to other current numbers – will receive a "cold" color. Figures with heatmaps depict the views of the users in a given area in our recommendation interface. Fig. 10 shows uniformly apportioned eyesight to each of the objects in the interface – the area of recommendation and the upper part of the interface attracted attention equally. We can see small differences to the disadvantage of the recommending area. Fig. 11 shows a similar relationship that we observe in Fig. 10, but the eye-pull accent is more evenly distributed, and we do not notice great differences in the saturation of either area with the red color. In Fig. 12, it can be clearly seen that there was little focus of sight on the recommending area, which shows yellow color with respect to intense red areas at the top of the interface. The last Fig. 13 shows a strong focus of sight on the recommending area, and we also see an intense red color in the upper interface area – the sight is distributed among all elements with a slight advantage over the upper area.

4. Multi-criteria Analysis of Used Scenarios

The PROMETHEE II method was used in the analysis. We have four scenarios in which we use four variants along with four criteria. The criteria which we have chosen for our analysis are: 1. Average time of looking at the interface, 2. Average time of looking at the area of recommendation, 3. Visual intensity level, 4. Time of the first look at the recommendation interface (First look). Four variants were analyzed with different weights assigned to the criteria. The first variant utilizes a weight of 25% for criteria 1 and 2, and 50% for criterion 4. The second variant has a weight of 25% for each of the four factors. In the third variant a weight of 25% was assigned to criterion 3 and 75% to criterion 4. In the fourth variant we assigned a weight of 100% to criterion 4. The options we will take into account are visual intensity levels from 0 to 3. The data that define individual variants are presented in Table 1, where time values are provided in milliseconds. They are used when creating rankings for individual scenarios.

Table 2. The PROMETHEE II method was used in the analysis. We have four scenarios in which we use four variants along with four criteria. Four variants were analyzed with different weights assigned to the criteria. Criteria: 1. Average time of looking at the interface, 2. Average time of looking at the area of recommendation, 3. Visual intensity level, 4. Time of the first look at the recommendation interface.

Ranking	Variant	ϕ	$\phi+$	$\phi-$	Ranking	Variant	ϕ	$\phi+$	$\phi-$
1-25%, 2-25%, 3- 0%, 4- 50%					1- 0%, 2- 0%, 3- 25%, 4- 75%				
1	VI2	0.5	0.75	0.25	1	VI1	0.8333	0.9167	0.0833
1	VI1	0.5	0.75	0.25	2	VI2	0.1667	0.5833	0.4167
3	VI3	-0.3333	0.3333	0.6667	3	VI0	0	0.5	0.5
4	VI0	-0.6667	0.1667	0.8333	4	VI3	-1	0	1
1- 25%, 2- 25%, 3- 25%, 4- 25%					1- 0%, 2- 0%, 3- 0%, 4- 100%				
1	VI2	0.3333	0.6667	0.3333	1	VI1	1	1	0
2	VI1	0.3333	0.6667	0.3333	2	VI2	0.3333	0.6667	0.3333
3	VI3	-0,3333	0,3333	0.6667	3	VI0	-0.3333	0.3333	0.6667
4	VI0	-0.3333	0.3333	0.6667	4	VI3	-1	0	1

4.1. Scenario I emphasizing first look in relation to first two criteria

In the first scenario, we take into account only three criteria: Average time of looking at the interface, Average time of looking at the area of recommendation and Time of the first look at the recommendation interface. According to our preferences, the first two should reach maximum value while for the third criterion we want it to be minimal. By analyzing Table 2 and Fig. 2 to Fig. 5 we can see that at the head of the ranking there are several dependencies. Analyzing the PROMETHEE results, we can conclude that variants VI1 and VI2 are preferred for activities in the ranking. They have the highest Average time of looking at the interface. As for the Average time of looking at the area of recommendation, they are ranked 3rd and 2nd, respectively. Time of the first look at the recommendation interface for those two variants is also the lowest, i.e. 2131 ms and 2159 ms. Their ϕ score is 0.5 for both variants. The VI3 and VI0 variants have much worse results. The ϕ score for VI3 is -0.3333, which is almost 3 times worse than VI1 and VI2. Variant VI0 is significantly lower than that and has the shortest Average time of looking at the interface. Time of the first look at the recommendation interface also belongs to some of the worst, therefore this variant ranks at the end of the stake with ϕ score of -0.6667. If we would like the first look to be as quick as possible as our priority and give the other two time-related criteria as equal yet less important than the priority criterion, we should consider choosing visual intensity levels between 1 and 2.

4.2. Scenario II with balanced criteria weights

In the second scenario, we take into account all four criteria with evenly spread weights. According to our preferences, the first two should reach the maximum value while for the third and fourth criterion we want them to be minimal. Analyzing the PROMETHEE result, we can conclude that variants VI1 and VI2 are equally preferred for activities in the ranking, similarly as in scenario 2. They have the highest Average time of looking at the interface. Their ϕ score is 0.3333 for both variants. Variants VI3 and VI0 have much produce results – they are on the opposite pole with their ϕ score of -0.3333, almost 2 times worse than in the case of VI1 and VI2. The visual intensity level criterion, however, at VI0 obviously shows the best result, because it has value 0 while the criterion is being kept to a minimum. If we want to ensure that all criteria are equally important to us, we should consider the choice between visual intensity levels between 1 and 2, as in scenario 1.

4.3. Scenario III emphasizing first look in relation to intensity level

The following scenario takes into account two criteria: Visual intensity level and Time of the first look at the recommendation interface, with weights of 25% and 75%, respectively. According to our preferences, we want the third and fourth criterion to be minimal. Analyzing the PROMETHEE result, we can conclude that variant VI1 is definitely preferred for activities in the ranking. The ϕ score is 0.8333. Another variant, VI2 has a ϕ score of 0.1677, which is almost six times worse. Variants VI3 and VI0 again have much worse results. If we want to make sure that visual intensity level and in particular time of the first look at the recommendation interface are important to us, we should definitely choose VI1. The remaining variants remain unattractive.

4.4. Scenario IV with total priority of the first look criterion

In the last scenario, we take into account only one criterion: Time of the first look at the recommendation interface with a weight of 100 %. This criterion is to be minimized. Analyzing the PROMETHEE results, we can conclude that variant VI1 is definitely preferred for activities in the ranking, as in scenario 3. The ϕ score is 1. The next variant, VI2 has a ϕ score of 0.3333, which is almost 3 times worse. Variants VI0 and VI3 invariably have much worse results with ϕ scores of -0.3333 and -1, respectively. If we want to treat the time of the first look at the recommendation interface as the only important to us, we should definitely choose visual intensity level 1. The remaining variants are not worth recommending.

5. Conclusions

Efforts of web content designers are often focused on attracting visual attention of web users with the use of various techniques to overcome habituation effects. Since high marketing content intensity usually leads to worse user experience, it is important to search for effective solutions with limited negative impact on webs users. The presented study showed the impact of different levels of visual intensity of the recommending component on user behavior. The use of eye tracking and fixations measurement allowed for detailed analysis of influence of recommendation presentation modifications on user behavior. Multi-criteria analysis with the use of PROMETHEE II allowed to analyze several scenarios based on four variants related to the levels of visual intensity of elements in the recommendation interface and four criteria, one of which is the crucial time of the first look of the user at the recommendation interface, from the point of view of our analysis. Analyzing all of the scenarios considered, there are several important conclusions. The first and foremost of them is to indicate that according to MCDA results it is advisable to select either VI2 (in scenarios 1 and 2) or VI1 variant of visual intensity (in scenarios 3 and 4). Highest visual intensity of the recommendation (VI3) or complete lack of highlight (VI0) were not selected. All in all, VI1 and VI2 variants turned out to be much more beneficial, and consequently, moderate visual intensity instead of no highlight or severe highlight seems to be the rational choice and, from the perspective of our study, is therefore worth considering while designing a recommendation interface and delivering recommended content.

Acknowledgements

The work was supported by the National Science Centre of Poland (Decision No. 2017/27/B/HS4/01216).

References

[1] Benway, J.P., 1998. Banner blindness: The irony of attention grabbing on the world wide web, in: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, SAGE Publications Sage CA: Los Angeles, CA. pp. 463–467.

[2] Castagnos, S., Jones, N., Pu, P., 2009. Recommenders' influence on buyers' decision process, in: Proceedings of the Third ACM Conference on Recommender Systems, ACM, New York, NY, USA. pp. 361–364.

[3] Chen, L., Pu, P., 2010. Eye-tracking study of user behavior in recommender interfaces, in: International Conference on User Modeling, Adaptation, and Personalization, Springer. pp. 375–380.

[4] Ha, L., McCann, K., 2008. An integrated model of advertising clutter in offline and online media. International Journal of Advertising 27, 569–592.

[5] He, J., Chu, W.W., 2010. A social network-based recommender system (snrs), in: Data mining for social network data. Springer, pp. 47–74.

[6] Jankowski, J., Hamari, J., Watróbski, J., 2019a. A gradual approach for maximising user conversion without compromising experience with high visual intensity website elements. Internet Research 29, 194–217.

[7] Jankowski, J., Hamari, J., Watróbski, J., 2019b. A gradual approach for maximising user conversion without compromising experience with high visual intensity website elements. Internet Research 29, 194–217.

[8] Jin, X., Zhou, Y., Mobasher, B., 2005. Task-oriented web user modeling for recommendation, in: International Conference on User Modeling, Springer. pp. 109–118.

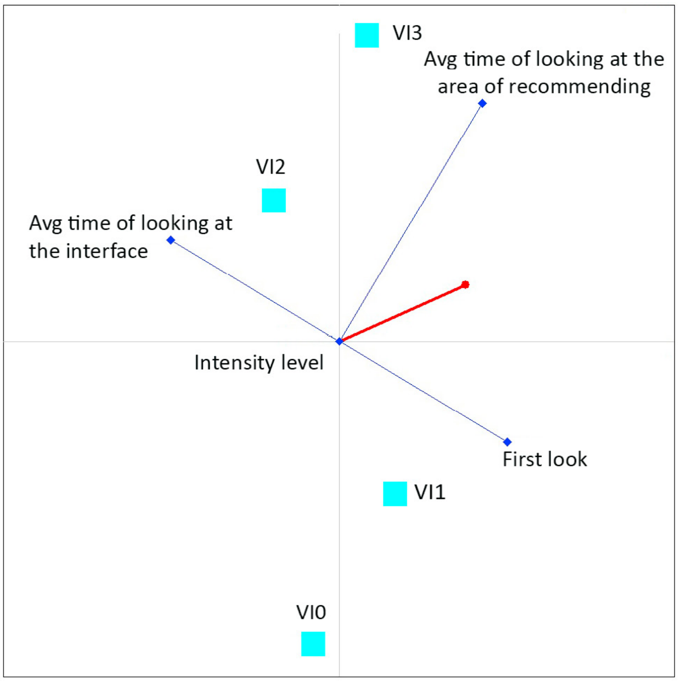


Fig. 2. Scenario 1 of multi-criteria analysis, with first and second criteria with weights of 25% and Time of the first look with weight 50%.

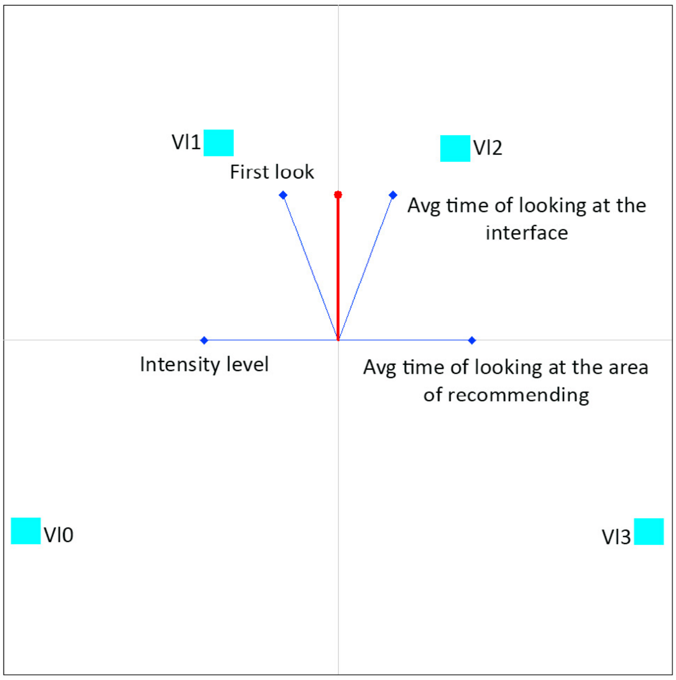


Fig. 3. Scenario 2 of multi-criteria analysis, with all criteria with weights of 25%.

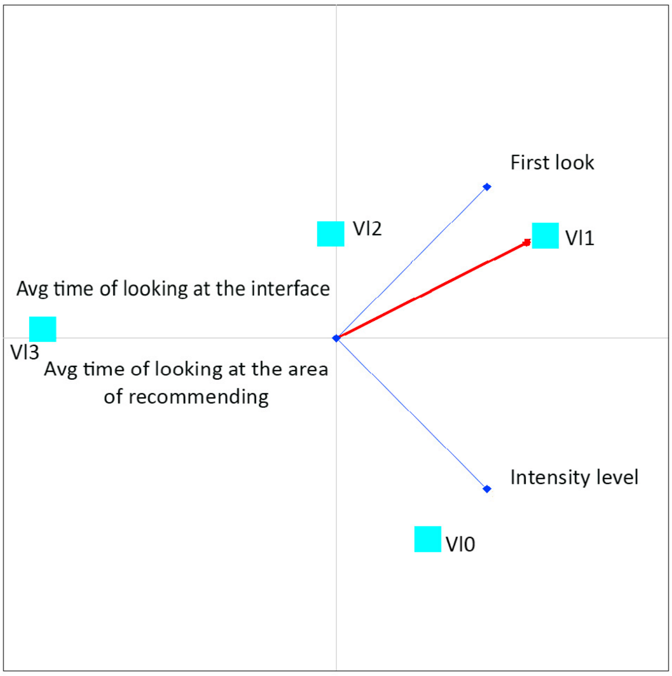


Fig. 4. Scenario 3 of multi-criteria analysis, with Intensity level criterion with weight of 25% and Time of the first look with weight of 75%.

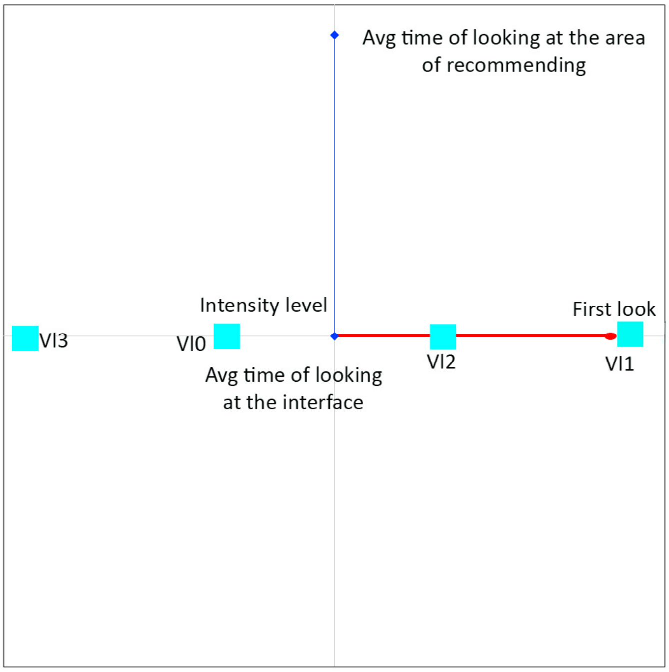


Fig. 5. Scenario 4 of multi-criteria analysis, with Time of the first look with weight of 100%.

[9] Lee, J., Ahn, J.H., Park, B., 2015. The effect of repetition in internet banner ads and the moderating role of animation. Computers in human behavior 46, 202–209.

[10] Lu, J., Wu, D., Mao, M., Wang, W., Zhang, G., 2015. Recommender system application developments: a survey. Decision Support Systems 74, 12–32.

[11] Owens, J.W., Chaparro, B.S., Palmer, E.M., 2011. Text advertising blindness: the new banner blindness? Journal of usability studies 6, 172–197.

Fig. 6. In the first place there is visible a large concentration on the area outside the recommendation area.

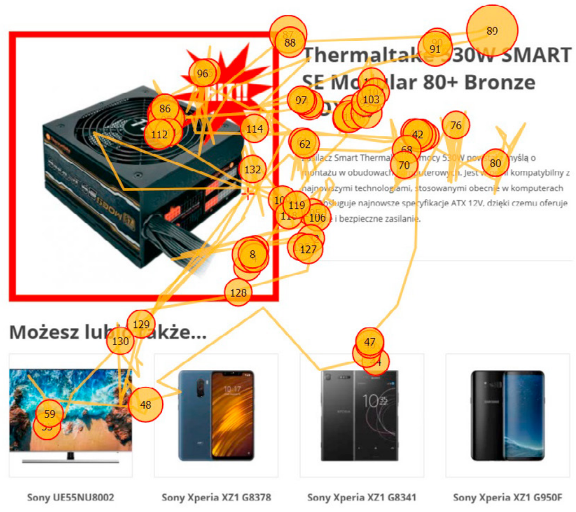


Fig. 7. Shows the first look, which also focuses on the upper interface.

Fig. 8. Clearly shows the first look in the vicinity of the recommending area.

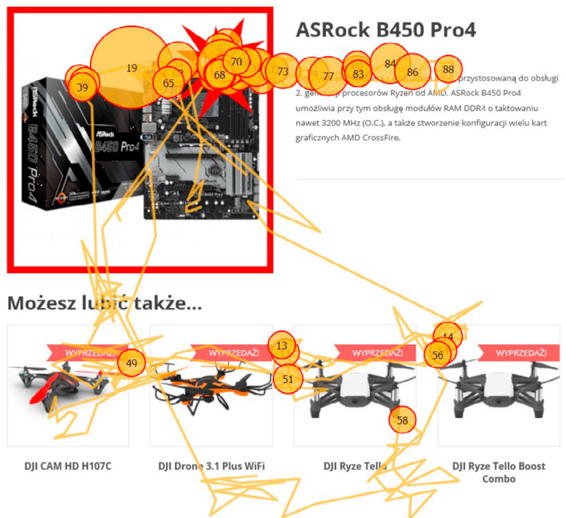


Fig. 9. Shows slight attention of the user to the recommendation area.

[12] Portnoy, F., Marchionini, G., 2010. Modeling the effect of habituation on banner blindness as a function of repetition and search type: Gap analysis for future work, in: CHI'10 Extended Abstracts on Human Factors in Computing Systems, ACM. pp. 4297–4302.

[13] Rohrer, C., Boyd, J., 2004. The rise of intrusive online advertising and the response of user experience research at yahoo!, in: CHI'04 Extended Abstracts on Human Factors in Computing Systems, ACM. pp. 1085–1086.

[14] Sulikowski, P., Zdziebko, T., Turzyński, D., Kańtoch, E., 2018. Human-website interaction monitoring in recommender systems. Procedia Computer Science 126, 1587–1596.

[15] Swearingen, K., Sinha, R., 2001. Beyond algorithms: An hci perspective on recommender systems, in: ACM SIGIR 2001 Workshop on Recommender Systems, Citeseer. pp. 1–11.

[16] Tintarev, N., Masthoff, J., 2007. A survey of explanations in recommender systems, in: 2007 IEEE 23rd international conference on data engineering workshop, IEEE. pp. 801–810.

[17] Verbert, K., Manouselis, N., Ochoa, X., Wolpers, M., Drachsler, H., Bosnic, I., Duval, E., 2012. Context-aware recommender systems for learning: a survey and future challenges. IEEE Transactions on Learning Technologies 5, 318–335.

[18] Watróbski, J., Jankowski, J., Ziemba, P., Karczmarczyk, A., Zioło, M., 2019. Generalised framework for multi-criteria method selection. Omega 86, 107–124.

[19] Yang, X., Guo, Y., Liu, Y., Steck, H., 2014. A survey of collaborative filtering based social recommender systems. Computer Communications 41, 1–10.

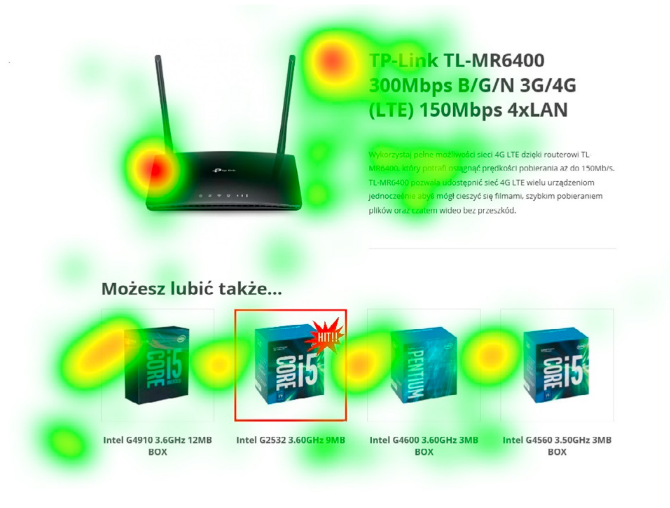


Fig. 10. Heatmap with small differences to the disadvantage of the recommending area.

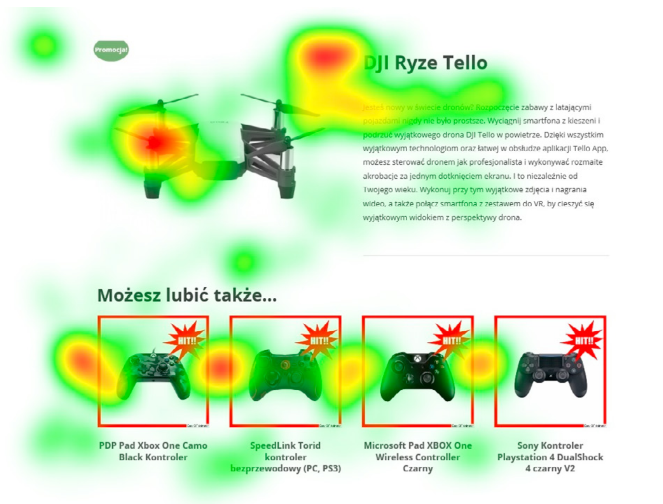


Fig. 11. Heatmap without significant differences between the areas.

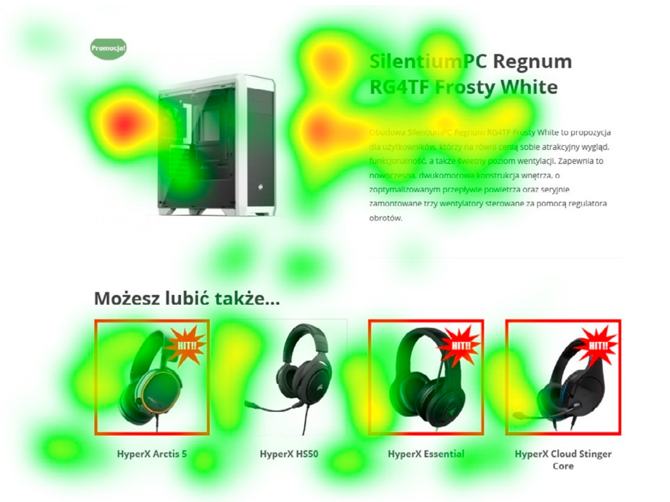


Fig. 12. Heatmap with small focus of sight on therecommending area

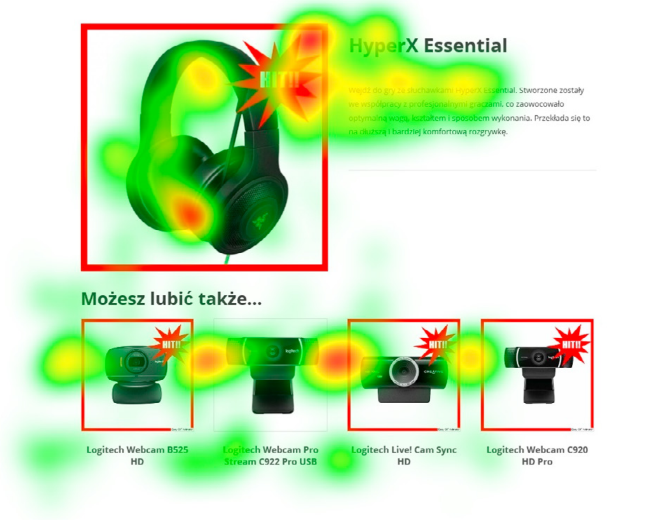


Fig. 13. Heatmap with sight distributed equally with a slight advantage over the upper area.

A6.

Bortko, K., Fornalczyk, K., Jankowski, J. (2021). Towards Effective Peripherial Chatbot Communication with Adjustable Intensity of Content Changes. 29th International Conference on Information Systems Development

Towards Effective Peripherial Chatbot Communication with Adjustable Intensity of Content Changes

Kamil Bortko
West Pomeranian University of Technology in Szczecin
Szczecin, Poland *kbortko@wi.zut.edu.pl*

Kacper Fornalczyk
West Pomeranian University of Technology in Szczecin
Szczecin, Poland *kacper_fornalczyk@wp.pl*

Jarosław Jankowski
West Pomeranian University of Technology in Szczecin
Szczecin, Poland *jjankowski@wi.zut.edu.pl*

Abstract

The art of designing user interfaces requires the efficient management of all the elements with which they cooperate. Problems appear with the effective delivery of messages to users dynamically switching web pages. This happens, for example, with chatbots appearing in peripheral areas. They are difficult to notice and are often missed during website exploration. This study shows the impact of the intensity of content changes in the chatbot area on the user response. The study was performed for changes in the peripheral area for different reaction times. The results showed that the intensity of changes within the chatbot area mainly impacts reactions within 400 ms; for a longer response, the difference between substantial and small changes is lower. Chatbot dynamics can be adjusted to user behavior to increase awareness and the performance of message delivery.

Keywords: chatbot, habituation, visual space, peripheral vision

1. Introduction

Large numbers of chatbots play a significant role in supporting customers and facilitating marketing processes. Chatbots are regularly used by companies looking to improve their customer service, sales, and marketing performance in various sectors. The number of chatbots has continued to grow, especially when Facebook launched its chatbot platforms [17]. The first phase of communication focuses on attracting the user's attention and, importantly, maintaining eye contact. What needs to be considered when designing chatbots and how they work is that excess stimuli can limit a user's ability to process information. This leads to some users ignoring or not noticing chatbots [7]. Excess stimuli can limit a user's ability to process information, leading some users to ignore or not notice chatbots. Phenomena such as habituation leading to a reduction in responses to repeated stimuli were also observed [5], and they are caused by the continuous attempts of chatbots to contact customers. We must deal with the fact that marketing content can quite often be ignored, with phenomena similar to text blindness [14] and banner blindness [6] that appear when users treat chatbot messages as marketing content. Another aspect is related to changing content within the chatbot interface. Some changes may go unnoticed, for example, due to a phenomenon known as change blindness [10]. The main goal of this study was to verify the impact of the intensity of changes within the chatbot area in peripheral vision on the ability to notice them during primary-task execution. This showed to what extent real-time changes in the chatbot area of an interface influence a user's attention toward the chatbot, depending on the possible time frame representing the website visit.

2. Related work

Literature studies treat chatbots as "online systems of human–computer dialogue with natural language" [1]. The first connection that affects the chatbot was assigned to Alan Turing. It is to him that we attribute this concept [9]. Chatbots can be embedded into interactive systems, and a recent study looked at chatbot interfaces [16]. The chatbot research literature focuses both on the visual effects of the conversation itself, and determining the level of anthropomorphism, which is closely related to the topic. Research that deals with the use of eye tracking to test the visual attention of subjects has been studied very little. This attention was paid to the visual effect of the chatbot (e.g., which part of the chatbot attracts the greatest amount of attention of our respondent). One of the main aspects [2] of the research was communication between the chatbot and humans. The main question was whether the chatbot should be more of an agent. The same visual chatbot area, which is the object of study, leaves room for much deeper and more accurate analysis. Many studies showed that the visual aspects of the chatbot significantly showed the importance of the visual appearance of the chatbot to users noticing [8]. Visual-effects research studies the overall design of a chatbot. What is also touched upon is the degree to which the use of thoughtful text [4] is perceived. Designing in terms of usability, in its broad sense, includes interactions with graphical elements of chatbots [3]. There are typical phenomena that can influence, e.g., visual communication. Habituation and poor reaction to repeated [11] content are repetitive and familiar chatbot phenomena. They worsen the performance of the chatbot and are a new area of research. Quite an interesting phenomenon that is correlated with chatbots is the so-called banner blindness. This mainly consists in reducing users' attention to various types of marketing components and filtering out unwanted content [13]. In the literature, we can find many publications that relate to banner blindness. They define it as the inability to perceive specific visual objects, such as advertising banners, in a visual scene. Messages that are displayed in the chatbot visual space are displayed in a very similar way to that of any security alerts on the systems. They are also routinely ignored by users. One of the important factors that influence the ineffectiveness of alerts is habituation. It is defined as a limited response to a realert. A similar phenomenon to habituation that is closely related to chatbots is change blindness. This occurs when users do not notice significant changes in the visual sphere of the interface that occur at the same time as other visualizations do [13]. Research by other authors showed that users perceive visualizations in various user interfaces as a single space that is rich in various types of elements. They often overlook large and significant changes that greatly impact attracting attention to chatbots [15]. However, human attention is limited, leading to the fact that a user's perception of dynamic events in the user's field of view requires a great deal of focus on appearing objects [10]. Noticing a change in the user's field of view requires a great deal of attention based on the high level of variability of elements in the field of view [12].

3. Conceptual framework

On a website, chatbots are usually integrated with other content for attracting user attention, such as product listings and commercial offers. As chatbots are often surrounded by other content and overloaded with marketing messages, their ability to attract user attention can be limited. In typical applications, user behavior is often based on switching between pages and only a limited amount of time is available for message delivery. Under such conditions, messages delivered via chatbots may remain unnoticed. In this study, we assumed that the main goal is to deliver information about chatbot changes in peripheral vision without breaking the primary task. Figure 1 shows an example of two processes with a static chatbot (A) with low potential to absorb user attention, and a dynamic chatbot with an adjustable intensity of changes (B), proposed in this study.

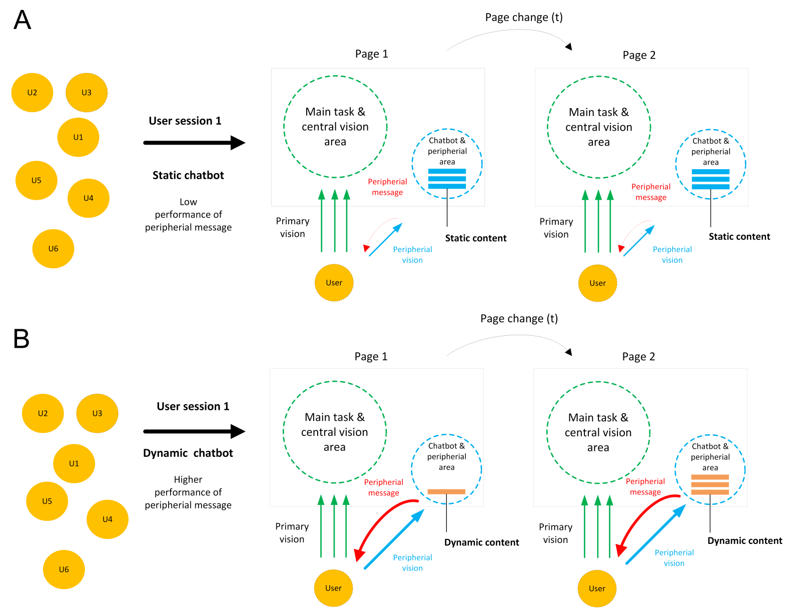


Fig. 1. Exemplary two processes with static chatbot (A) with low potential to absorb user attention and proposed in this study dynamic chatbot with adjustable intensity of changes (B).

In Process A, the user concentrates on the primary task, and messages shown in peripheral area are not delivered. After a short amount of time, the user switches from Page 1 to Page 2, with low ability to notice chatbot messages. In Process B, the user concentrates on the primary task, but chatbot changes are noticed due to the adjustable intensity. They are visible in the peripheral vision, and users, without interrupting their task, are aware of new content before switching to another page. The natural process of communication with a chatbot includes the appearance of new messages and the disappearance of old ones. The intensity of changes within the chatbot, represented by the percentage of content that changes between updates within certain time intervals, may impact user attention. The above assumptions and theoretical background led us to create a decomposed chatbot design with the ability to adjust the number of messages on the basis of content slots located within the interface. The number of messages assigned to each time interval can differ, and the changes in messages between time intervals are the main mechanics of content changes within the visual space that drive user attention.

4. Results

The experiment based on a chatbot with adjustable intensity of changes was attended by 13 people aged in the range of 18–35 years. The test was performed on a 27" monitor. Each person's head was firmly set on a chin rest 54 cm away. This made it possible to obtain a 45-degree angle between the center of the screen and the edge of the chatbot. A Tobii Pro X3 eye tracker with a 120 Hz sampling rate was used to track the user's gaze paths, and the average session lasted 15 minutes.

The main characteristics of the data comprised two categories. Each change of state (SC) was divided into 9 changes in the chatbot areas: 1:1, 1:2, 1:3, 2:1, 2:2, 2:3, 3:1, 3:2, and 3:3. The sequence of changes was based on two states. The first number, Num1, from pair NUm1:Num2 denotes the number of message components in the first state, and the second number, Num2, in the second state.

Figure 2A shows a diagram of user interaction during the experiment. Users looked at the area marked in red. In a given area, a blue dot appeared and randomly pulsed. On the right-hand side of the user's screen, a chatbot was visible in which content appeared sequentially. The user, thanks to peripheral vision, could see objects that were not directly in front of them-in this case, the chatbot with content appearing. If changes were noticed, the user had to press the mouse button. The intensities of changes based on the nine states were tested for each user with three repetitions and a random order.

Table 1. Results of ANOVA analysis for all time intervals and regression analysis of the number of reactions against the SC.

	SC vs time of intervals					**Regression**	
	400 ms	**500 ms**	**600 ms**	**700 ms**	**>700 ms**	**Number of reaction vs SC**	
F	15.573	13.892	14.829	14.481	14.102	SE of estimation	3.541
p	0.002	0.003	0.002	0.009	0.012	**p**	0.001

For analysis, we used variables related to noticed changes in chatbot areas grouped into time intervals based on registered reactions on changes reported by pressing mouse button. They were grouped into five variants: times up to 400, 500, 600, 700, and over 700 ms. Analyzing the impact of the intensity of changes within chatbot areas on the response using statistical ANOVA showed that it exhibited statistical significance $p < 0.05$ in each of the five intervals, ranging from 0.001 to 0.012. In each of these cases, there was a definite clear, strong, and statistically significant impact of SC on the time of intervals. F statistics for each of these groups also looked very similar, with a slight indication of the first interval, or up to 400 ms, in which differences in the continuity of the corrections of the result were slightly smaller. Here, we had results of 15.573, showing the highest impact of intensity changes for the lowest used time interval. Variables using analysis of variance, thanks to which we checked whether certain independent variables (factors (SC)) had an impact on the level of the dependent variable (tested, measured variable), i.e., time intervals in our case.

If we consider regression analysis, we can note several important aspects. To assess the goodness of fit between the models and data, we used the values of the determination coefficients (R^2). For the model, their value was very high, $R^2 = 98.481$. A good fit was also evidenced by the low value of the SE of estimation at the level of 3.541.

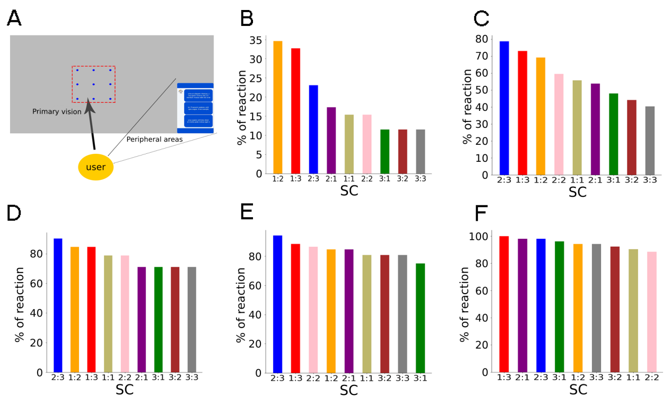


Fig. 2. (A) Experimental screen with diagram of user interaction during the experimental study's and from (B) to (F) bar charts with every time intervals (400 ms, 500 ms, 600 ms, 700 ms, >700 ms)

Figure 2B shows, for time of intervals of 400 ms, the percentage of users reporting noticed changes in the chatbot. The reaction percentage had the highest values with growth SC, i.e., 1:2, 1:3, and 2:3. The worst results were achieved with SC that began with 3 news blocks. The values in this chart ranged from 12% and a maximum of 35%. from Figure 2C–E (500–700 ms) show similar tendencies to those in Figure 2B with slight shifts followed by SC. However, here, the reaction percentage spread ranged between 45% and nearly 90%. The greater the time of intervals was, the greater the reaction percentage. Figure 2F (> 700 ms) shows a reaction percentage in the range from 90% to nearly 98%.

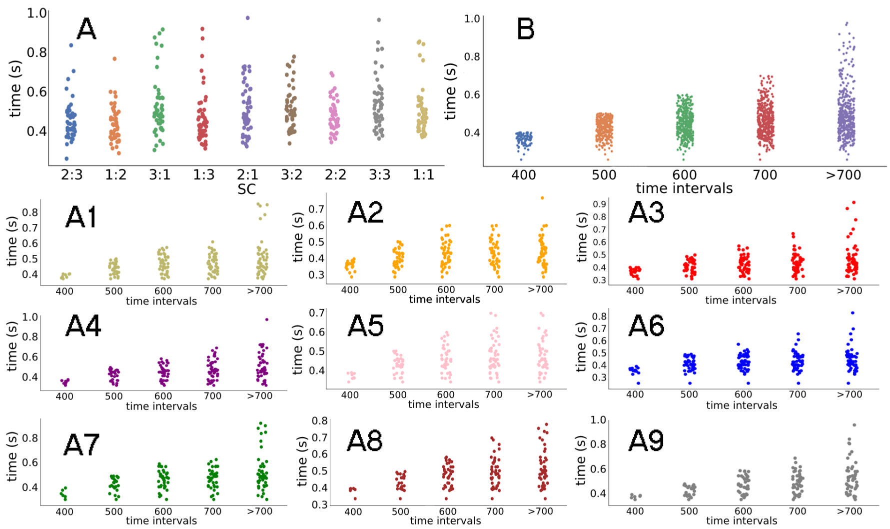


Fig. 3. Scatter plot (A) is number of reactions in particular time periods divided into SC. Scatter plot (B) shows umber of reactions in particular time periods divided into time intervals. Charts form (A1) to (A9) refer in color to Fig.3A.

Figure 3A shows the number of reactions in particular time periods divided into states of change (SC). Clusters of fast reactions up to nearly 500 ms are visible here. The longer the time was, the greater the scatter of the reaction became. Figure 3B shows the reactions of users, broken down into individual time intervals. At the time intervals, the clusters in the upper limits were very scattered. The remaining plots (Figure 3A1–A9) refer in color to Figure 3A, and a breakdown of these reactions against the time intervals is visible here.

5. Conclusions

The main goal of this study was to verify the impact of the intensity of changes in a chatbot on providing awareness about new content in peripheral vision without interrupting a main task. Different intensities of changes within the chatbot were used on the basis of the sequence of pairs of content units and adjusted differences between them. Results clearly showed that the intensity of changes in the chatbot area significantly affected the reactions of users who stayed on the website for nearly 400 ms. For a longer response, starting from up to 500 ms to the end, the difference between large and small changes was smaller. This hints that the dynamics of changes in the chatbot could be adapted to the individual behavior of users in order to increase their awareness of changes in the efficiency of delivering specific messages.

6. Acknowledgments

This work was supported by the National Science Centre of Poland, Decision No. 2017/27/B/ HS4/01216, and by the ZUT Highfliers School (Szkoła Orłów ZUT) project, coordinated by Dr. Piotr Sulikowski, within the framework of the program of the Ministry of Education and Science (grant no. MNiSW/2019/391/DIR/KH, POWR.03.01.00-00-P015/18), cofinanced by the European Social Fund; the amount of financing was PLN 1,704,201.66.

References

1. Cahn, J., 2017. Chatbot: Architecture, design, & development. University of Pennsylvania School of Engineering and Applied Science Department of Computer and Information Science .
2. Chaves, A.P., Gerosa, M.A., 2019. How should my chatbot interact? a survey on human-chatbot interaction design. arXiv preprint arXiv:1904.02743 .
3. Følstad, A., Brandtzæg, P.B., 2017. Chatbots and the new world of hci. interactions 24, 38–42.
4. Følstad, A., Nordheim, C.B., Bjørkli, C.A., 2018. What makes users trust a chatbot for customer service? an exploratory interview study, in: International conference on internet science, Springer. pp. 194–208.
5. Groves, P.M., Thompson, R.F., 1970. Habituation: a dual-process theory. Psychological review 77, 419.
6. Hervet, G., Guérard, K., Tremblay, S., Chtourou, M.S., 2011. Is banner blindness genuine? eye tracking internet text advertising. Applied cognitive psychology 25, 708–716.
7. Itti, L., Koch, C., 2001. Computational modelling of visual attention. Nature reviews neuroscience 2, 194–203.
8. Kuligowska, K., 2015. Commercial chatbot: performance evaluation, usability metrics and quality standards of embodied conversational agents. Professionals Center for Business Research 2.
9. Kumar, P., Sharma, M., Rawat, S., Choudhury, T., 2018. Designing and developing a chatbot using machine learning, in: 2018 International Conference on System Modeling & Advancement in Research Trends (SMART), IEEE. pp. 87–91.
10. Owens, J.W., Chaparro, B.S., Palmer, E.M., 2011. Text advertising blindness: the new banner blindness? Journal of usability studies 6, 172–197.
11. Rankin, C.H., Abrams, T., Barry, R.J., Bhatnagar, S., Clayton, D.F., Colombo, J., Coppola, G., Geyer, M.A., Glanzman, D.L., Marsland, S., et al., 2009. Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. Neurobiology of learning and memory 92, 135–138.
12. Rensink, R.A., 2002. Change detection. Annual review of psychology 53, 245–277.
13. Resnick, M., Albert, W., 2014. The impact of advertising location and user task on the emergence of banner ad blindness: An eye-tracking study. International Journal of Human-Computer Interaction 30, 206–219.
14. Simons, D.J., Rensink, R.A., 2005. Change blindness: Past, present, and future. Trends in cognitive sciences 9, 16–20.
15. Smith, T.J., Henderson, J.M., 2008. Edit blindness: The relationship between attention and global change blindness in dynamic scenes. Journal of Eye Movement Research 2.
16. Yan, M., Castro, P., Cheng, P., Ishakian, V., 2016. Building a chatbot with serverless computing, in: Proceedings of the 1st International Workshop on Mashups of Things and APIs, pp. 1–4.
17. Zarouali, B., Van den Broeck, E., Walrave, M., Poels, K., 2018. Predicting consumer responses to a chatbot on facebook. Cyberpsychology, Behavior, and Social Networking 21, 491–497.

A7.

Fornalczyk, K., Bortko, K., Jankowski, J. (2021). Improving User Attention to Chatbots through a Controlled Intensity of Changes within the Interface. Procedia Computer Science, 192, 5112-5121

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 192 (2021) 5112–5121

Procedia Computer Science

www.elsevier.com/locate/procedia

25th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems

Improving User Attention to Chatbots through a Controlled Intensity of Changes within the Interface

Kacper Fornalczyk[a], Kamil Bortko[a*], Jarosław Jankowski[a]

[a]Faculty of Computer Science and Information Technology, West Pomeranian University of Technology, Szczecin, Poland

Abstract

Designing user interfaces requires the efficient management of elements that interact with the user and attract their attention. In the development of technologies supporting the development of chatbots, few studies have focused on the visual aspects of the interface. Changes in the chatbot area can be of different natures, such as increasing or decreasing the number of messages. Due to habituation (change blindness), messages delivered by chatbots may not be noticed. The main goal of the present study was to examine the impact of changes within the chatbot on user behaviour and the possibility of directing user attention to the chatbot area. This study showed that the intensity and types of changes within the chatbot can improve performance and successful message delivery to users.

Keywords: human-computer interaction, chatbots, habituation, change blindness, eye tracking

1. Introduction

Take, chatbots play an important role in supporting customers and facilitating marketing processes. Thousands of chatbots are already helping businesses improve their customer service, sales, and marketing performance in various sectors. Chatbots are expected to be the one of the most important AI applications for consumers in the next five years [14]. The number of chatbots has increased over time, especially after

* Corresponding author
E-mail address: kbortko@wi.zut.edu.pl

1877-0509
10.1016/j.procs.2021.09.289

Facebook launched its chatbot platforms [17]. The development of chatbots is focused on natural language processing, applications of machine learning [30], and neural networks [26], with the main challenges related to making chatbot–human communications appear closer to human–human communications through more accurate responses [19]. However, while chatbots can participate and grab a user's attention, natural language interfaces to computer systems cannot yet simulate the full range of intelligent human conversation [19]. Another studied aspect is chatbot interfaces. For commercial purposes, it is important to engage users in communication with chatbots and actively use those chatbots for supporting processes within company websites [21]. In the first phase of communication, a visual contact is important, as is attracting user attention [25]. Several problems have also been observed in other areas. Excess stimuli can limit a user's ability to process information, leading some users to ignore chatbots or not notice them. Phenomena like habituation leading to a reduction of responses to repeated stimuli have also been observed [6] and are caused by the continuous attempts of chatbots to contact customers. As marketing content is often subconsciously ignored, phenomena similar to banner blindness [22] and text blindness [12] can be observed when users treat chatbot messages as typical marketing content. Another aspect is related to changing content within the chatbot interface. Changes may be not noticed, for example, due to the change blindness phenomena [24], leading to the user not noticing visual changes, especially when other stimuli demand the user's attention.

While earlier studies focused on other areas of human–computer interactions and the impacts of changes within the visual space, the present research explores how changes within the chatbot area can influence user attention. Chatbot usage can be considered from the perspective of supporting users in task completion while other applications remain focused on the delivery of marketing content. In both cases, the effective delivery of messages is important. Thus, the main goal of this study was to answer the following question: To what extent do real-time changes in the chatbot area of an interface influence a user's attention toward the chatbot? The resulting experiment showed how attention can be increased through intensive changes, rather than gradual effects, as well as the differences in response when different levels of changes are used. The following section presents a literature review and is followed by the conceptual framework. The results are presented in the experimental section and summarized in the conclusions.

2. Review of the Literature

Studies in the literature treat chatbots as "online systems of human–computer dialogue with natural language" [14]. The first chatbot concept is attributed to Alan Turing [14]. Since Turing, chatbot technology has improved with advances in natural language processing and machine learning [30]. A chatbot is a program driven by the rules and intellectual abilities of a human that focuses on conversation strategies; chatbot designers largely use machine learning and AI techniques [5]. Chatbots can enrich an organization with unique knowledge based on consumer feedback [5]. Using replies similar to those of real people, chatbots have the ability to grab a user's attention [16]. Recently, however, the integration of chatbots with messengers has become the focus of attention [18].

While chatbots can be integrated with interactive systems, recent studies have also considered chatbot interfaces [20]. Chatbot research literature focuses on the visual effects of the conversation itself and the level of anthropomorphism. The use of eye tracking to study a user's visual attention has been explored very little, and no particular attention has been given to the appearance of the chatbot itself (e.g., which chatbot elements attract the most attention). Another study [13] presented a literature review on human–chatbot interactions. The main aspects included the communication of the chatbot with a human, whether the chatbot should be more like an agent, etc. The visual area of the chatbot itself, however, leaves room for deeper analysis. The results of many studies on chatbot visualization emphasized the importance of a chatbot's visual appearance [9]. Visual effects studies explore the overall chatbot design, as well as the extent to which the use of well-thought-out text is perceived [10]. In this case, designing for usability involves interactions with the graphical elements of the chatbots [11].

The performance of a chatbot acting within visual space can be influenced by the typical phenomena affecting visual communication. Habituation and lower responsiveness to repeated content [15] are common phenomena that can decrease chatbot performance and represent a new area of study. Another interesting phenomena related to chatbots is banner blindness, which is well-known from digital marketing studies based on a user's decreased attention to marketing components and filtering unwanted content [28]. Literature treats banner blindness as the inability to perceive specific visual objects, e.g., advertising banners, in the visual scene. Chatbot messages can appear within the visual space in a similar way to security warnings and are likewise routinely ignored. The main influencing factor behind the ineffectiveness of warnings is caused by habituation, i.e., a limited response to a repeated alert.

Another phenomenon directly related to chatbots is change blindness, which occurs when users do not perceive even substantial changes in the visual scene that occur simultaneously with other visuals [28]. Earlier studies showed that humans perceive the visual elements in user interfaces as a uniform space that is rich in various types of elements but often fail to notice large and significant changes [23], which can also affect chatbots. Given that human attention is limited, the perception of dynamic events in a scene requires focus [24]. Attention is required to perceive the change, and in the absence of localized motion signals, attention can be shifted based on the high level of variability of elements in the field of view [29].

3. Conceptual framework

On a website, chatbots are usually integrated with other content for attracting user attention, such as product listings and commercial offers. As chatbots are often surrounded by other content and overloaded with marketing messages, their ability to attract user attention can be limited. Under such conditions, messages delivered via chatbots may remain unnoticed. In this study, we assume that the intensity of changes within the chatbot interface can influence user attention during engaging tasks. The main goal is to maintain relatively low visual intensity without the usage of invasive content such as flashing elements, vivid effects, or animations, with the main content of chatbots focused instead on text-based messages. The present conceptual framework uses mechanisms identified in earlier theoretical studies. The natural process of communication with a chatbot includes the appearance of new messages and the disappearance of old ones. While chatbots can be organized in several ways, the main question is as follows: What kind of changes within a chatbot will attract more attention—message elements appearing or disappearing? Cole and Kuhn [1] showed that appearing objects a perceived with higher priority and have a greater potential to capture user attention. An earlier study by Franconeri and Simons [2] proposed a behavioural urgency hypothesis and identified that greater attention was attracted by suddenly appearing objects requiring immediate attention due to representing a potential threat. Therefore, it is expected that appearing elements will initiate a better response, but questions remain regarding the differences between both approaches—for example, whether the disappearance of larger objects and numbers of messages can have a comparable impact to the appearance of a smaller portion of information featuring new content.

Another question relates to the intensity of changes within the chatbot, represented by the percentage of content that changes between updates within certain time intervals. Presumably, large changes with many new message components will deliver better results than small increases in content. Ball, Elzemann, and Busch [3] focused on scene manipulation and measuring changes based on the number of pixels changed. The results showed that response times decreased by increasing the size of the change. In relation to the number of messages updated within chatbots, one important factor is the relationship between the impact of the number of new messages on response time and the proportion of users who notice those changes. Other aspects from earlier perceptual studies with direct applications in chatbots were based on the impact of gradual or disruptive changes. Laloyaux, Devue, and Cleeremans [4] showed that gradual changes within a visual space are more difficult to detect than disruptive modifications, as such changes result in high levels of change blindness. Simons, Franconeri, and Reimer [2] showed that gradual changes can result in change blindness even if those changes are performed in the central part of one's vision area. These early results motivated further experiments using chatbots to measure how the gradual appearance of messages with small changes can differ from showing few messages at once.

The above assumptions and theoretical background led us to create a decomposed chatbot design with the ability to adjust the number of messages based on content slots located within the interface. The structure of the experimental interface is presented in Fig 1. To maintain the assumed changes within the chatbot interface and control them through message exposition, a set of $M_t=(m_1,m_2,\ldots,m_t)$ messages is assigned to each time interval t. The number of messages assigned to each time interval can differ, and the changes in messages between time intervals are the main mechanics of content change within the visual space that drive user attention.

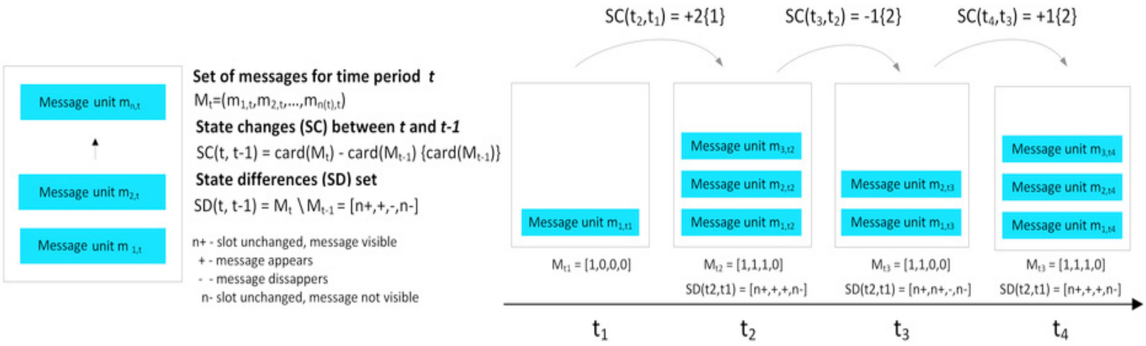


Fig. 1. Chatbot decomposition and incremental and decremental state changes

Each chatbot state St is described by the number of presented messages. The stage changes SC between time t and t-1 are described by the difference in the number of messages shown in time t and t-1. These changes are computed as the difference in the number of messages within set M_t and set M_t-1, with the number of messages in t-1 serving as a reference for added messages or messages in time t additionally denoted as cardinality {card(M_t-1}. The state difference set SD represents the difference between sets M_t and M_t-1 and indicates the type and intensity of changes between t-1 and t. The elements of the set correspond to messages from set $M_t=(m_1,m_2,\ldots,m_n)$, with values that are unchanged and visible denoted as n+ (a message exists at the same position for both t and t-1) and those that are unchanged and not visible denoted as n- (empty message slot for both time intervals), showing a message in a given slot with '+' for appearing and '-' for disappearing. This method makes it possible to describe the chatbot states in each stage and the differences between time intervals in a generalized manner.

Fig. 1 also shows an example process of changes in the number of messages within the chatbot's interface. In the first stage t_1, one slot is used, and a single message unit is shown. In the second stage t_2, another two parts of the message are shown, and potential user would experience changes within the interface from one message to three, with the potential to attract the user's attention. In the next stage t_3, two message units are presented, and the change is based on a reduction in the number of messages between t_3 and t_2. The next step produces one additional message, with the change between t_3 and t_4 denoted as SD(t_4,t_3)=(n+,n+,+,n-) and SC(t_4,t_3)=+1{1}.

While the generalized version of the interface for practical usage is based on using n messages to cover all possible state changes within experimental user sessions, a four-state interface was designed with an empty interface and one, two, and three message units included within the interface. The whole experimental space with all 15 possible change variants is presented in Fig. 2.

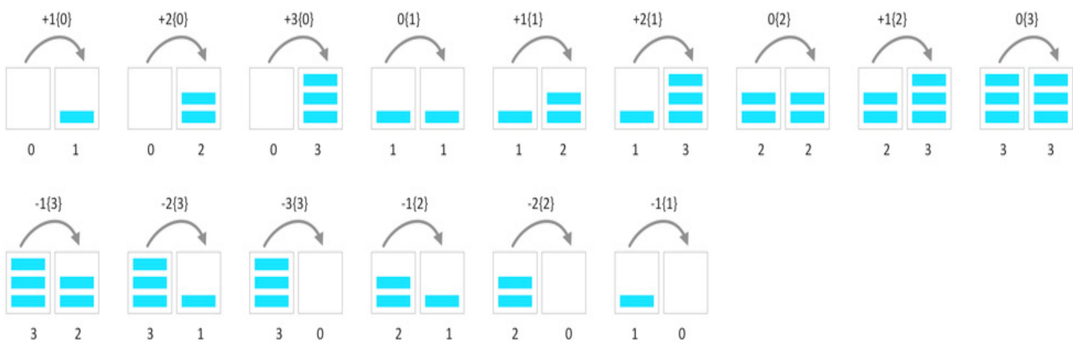


Fig. 2. Complete experimental space for four state object with incremental and decremental changes

The above assumptions consider incremental changes starting from a level with zero, one, or two messages and three combinations featuring unchanged states, with one, two, or three visible elements for both time intervals. This model also covers decreases in the number of messages (three > two > one message(s)), as well as increases and decreases in the number of messages and changes based on one, two, or three messages influencing the intensity of changes within the visual space.

4. Experimental setup and results

The above assumptions were used to design the experimental environment. To engage the user, the experimental chatbot was integrated with a game based on searching for visual elements within complex drawings. Fig. 3 presents the interface layout for the game and integrated chatbot, with typical localization within an application displayed in the bottom-right corner. In this example, the chatbot is providing tips to the user.

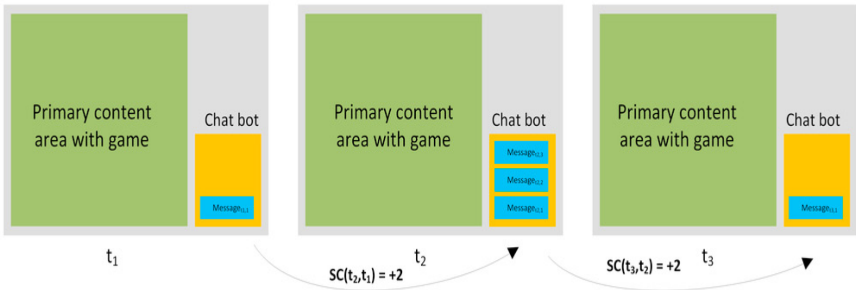


Fig. 3. Experimental setup and integration of the chatbot with a game

The main goal of this research was to study the effect of changes within the chatbot area on user attention and the ability to drive attention toward the chatbot area through visual paths. The chatbot interface was changed using time intervals according to the experimental plan using each pair of changes in a random sequence, with three repetitions per area of change. For each time period $t_1, t_2, \dots, t_n$ messages are loaded. The proposed structure and changes between chatbot updates enabled us to perform an experiment to analyse the effect of unchanged messages between time intervals, with one, two, and three messages for both time intervals and state differences of SD(t,t-1)=(n+, n-, n-), SD(t,t-1)=(n+, n+, n-), and SD(t,t-1)=(n+, n+, n+). Another monitored aspect was the difference between the appearance and disappearance of messages. Messages appeared in five element combinations: SD(t,t-1)=(+, n-, n-), SD(t,t-1)=(+, +, n-), SD(t,t-1)=(+, +, +), SD(t,t-1)=(n+, +, +), and SD(t,t-1)=(n+, +, n-). Messages disappeared in four combinations with state differences of SD(t,t-1)=(n+, n+, -), SD(t,t-1)=(n+, -, -), SD(t,t-1)=(-, -, -), and SD(t,t-1)=(n+, n+, -).

This conceptual framework and experimental environment were used to verify the impact of changes within the chatbot on the user's attention and reactions to the chatbot messages. The experiment was based on 15 users performing tasks that involved searching for items in specific images and then tagging those images. After tagging the images, the chatbot suggested where to look for the relevant elements. A random allocation of changes was used for each user, with three repetitions. A Tobii Pro X3 eye tracker with a 120 Hz sampling rate was used to track the user's gaze paths, and the average session lasted 11 minutes. The eye tracker was calibrated to each user at the beginning of the experiment.

4.1. Descriptive global statistics

The analysed data were grouped and aggregated into three categories based on some common characteristics. We distinguished the obtained measurement results for the time the user needed to switch from the game area to the chatbot as effective message delivery (EF) or ineffective message delivery (INEF). For EF, we included results whose transition times ranged from 100 to 2000 ms, based on a study that excluded reaction times of <100 ms as "too early" and reaction times > 2000 ms as "too late' (i.e., not a direct reaction to the stimuli) (Otaki and Shibata, 2019). We included cases with no reaction to the change in the INFE group.

The main characteristics of the data included three categories: the state difference (SD) between changes in chatbot content, the direction of changes based on appearing or disappearing content (increases or decreases in the number of messages), and the start values ("start" refers to the actual filling of the chatbot at the time of change). The data for measuring the time of transition to the chatbot and the total number of EF and INEF responses were averaged for each state change (SC). Each of these categories was given its own groups. SD included groups 0, 1, 2, and 3 based on the differences in number of messages, followed by the direction groups of none, down, up, and finally starts, which included groups 0, 1, 2, and 3 with the initial number of messages before the occurrence of a change.

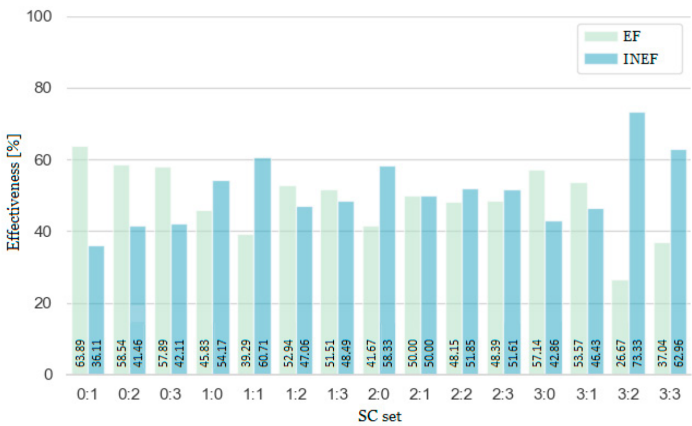


Fig. 4. Effectiveness represented by the percentage of effective (EF) and ineffective (INEF) expositions for each state changes (SC)

The EF and INEF percentages for each state change are presented in Fig. 4. Here, each of the SCs presents a very different EF:INEF ratio. The largest difference in EF can be seen for states 3:2 and 3:3, with a benefit observed for the INEF SCs. For these two SCs, the ratios are around 65-75% the INEF ratios. There is also a significant difference in favour of INEF for SC 1:1 and 2:0 (in both cases, almost 60% are INEF). For the SC group starting with 0 (0:1, 0:2, 0:3), there is a favourable percentage shown for EF. In this case, the percentage of EF is much higher in all cases than the percentages of INEF.

Table 1. Results for changes in parameters: the average and quantitative values for EF and INEF.

Change parameters					Effective delivery		Non-effective delivery	Total EF & INEF
State change SC	Percentage changed	Interface area	State difference SD	Direction	EF	Avg reaction time	INEF	
1 -> 3	66%	100%	2	increase	17	737.9	16	33
0 -> 3	100%	100%	3	increase	22	760	16	38
2 -> 2	0%	66%	0	neutral	13	776.2	14	27
1 -> 2	50%	66%	1	increase	18	786.7	16	34
2 -> 1	50%	33%	1	decrease	14	813.9	14	28
1 -> 0	100%	0%	1	decrease	11	826.2	13	24
2 -> 3	33%	100%	1	increase	15	853.1	16	31
1 -> 1	0%	33%	0	neutral	11	861.2	17	28
0 -> 2	100%	66%	2	increase	24	881	17	41
3 -> 3	0%	100%	0	neutral	10	891.3	17	27
0 -> 1	100%	33%	1	increase	23	916.9	13	36
3 -> 1	66%	33%	2	decrease	15	935.3	13	28
2 -> 0	100%	0%	2	decrease	10	1015.8	14	24

3 -> 0	100%	0%	3	decrease	16	1022.8	12	28
3 -> 2	33%	66%	1	decrease	8	1155.5	22	30

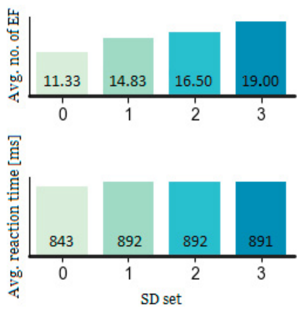


Fig. 5. Mean reaction time and number of EF for state differences SD

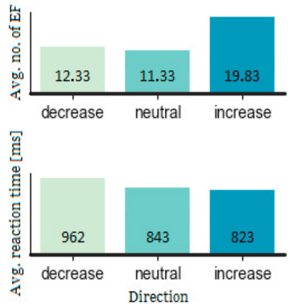


Fig. 6. Mean reaction time and number of EF for each Direction

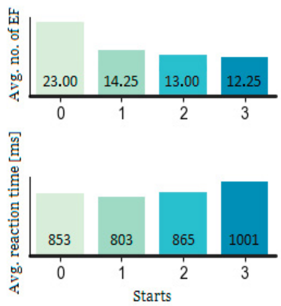


Fig. 7. Mean reaction time and number of EF for each Start

Fig. 5 presents two charts. The image on the top shows the global average number of EF responses by all users for changes of different sizes. Here, the larger the changes are, the better the results become based on the number of EFs by the test subjects. The average number of EFs for three changes was almost twice as high as that for zero changes, where only the content changed, not the number of slots occupied. The graph in the bottom row shows the global average EF response times for the same changes. Here, the averages for the different SCs do not differ significantly. Fig. 6 also consists of two diagrams. In the top row, Fig. 6 shows the global average number of EFs by all users for changes in direction. A significant advantage can be observed in the average number of EFs for directional changes. The downward and neutral changes showed similar results, and the graph on the bottom row shows the global EF response times for the same changes. On average, for downward shifts, users needed about 100 ms more time per reaction compared to neutral shifts and upward shifts. The top row of Fig. 7 shows the global average EF numbers of all users for changes starting with different chatbot fill levels. Here, a significant advantage can be observed in the average number of EFs for changes in which the chatbot was initially empty. There was also a downward trend in the number of EFs depending on the amount of the chatbot filled at the time of change—the more filled the chatbot, the smaller the number of EF responses. On the bottom row, Fig. 7 shows the global EF response times for the same changes. For the changes that took place when the chatbot was completely full, users needed more time on average to react (about 170 ms). The remaining changes were characterized by similar results. Fig. 8 presents two diagrams. The graph in the top row shows the global EF and average EF times of all users for all possible changes. The times are between 730 and 1160 ms. An extended reaction time is observable for the group of SCs with three starts compared to the others. The figure on the bottom row shows the ratio of the global number of EF to INEF responses. An SC of 0:1 provided the best EF, and a change of 3:2 provided the worst EF.

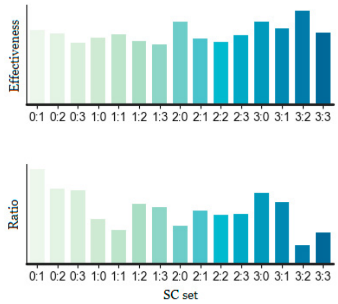


Fig. 8. Ratio and effectiveness for state changes SC set

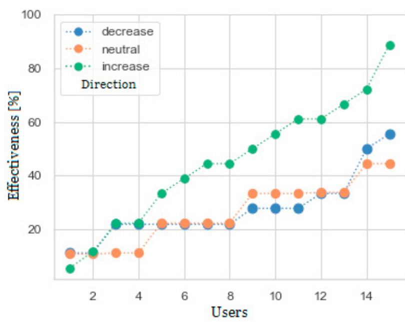


Fig. 9. Effectiveness in % per Users for direction

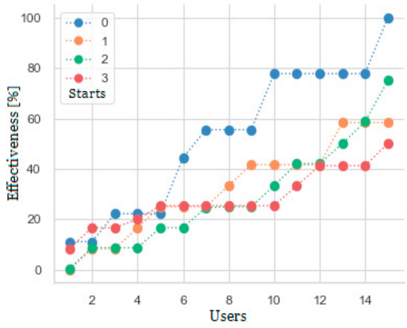


Fig. 10. Effectiveness in % per Users for Starts

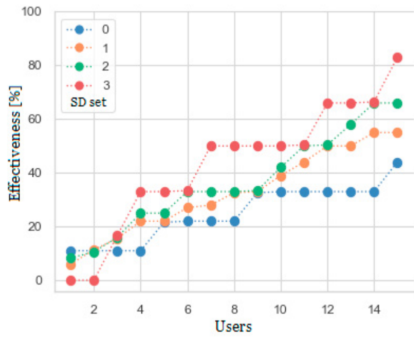


Fig. 11. Effectiveness in % per Users for SD set

Another portion of this analysis is based on individual users. Fig. 9 shows the percentage of EF elements for the direction of the level changes of each user. Users were sorted according to an increasing percentage of EFs. Here, an upward shift increased the number of EFs for nearly every user. Neutral changes and those lower than the percentage of EFs appeared very similar. There were no major differences in these two states. Fig. 10 shows the pairs starting with 0, 1, 2, and 3. The results here are also organised by the percentage of EFs for each user. In five cases, the relationship results were similar. However, in the remaining ten cases, there is a clear and large increase in the percentage of EFs for those starting with 0. In this case, the percentage of EFs clearly exceeds 50%. Lastly, Fig. 11 shows the state changes from level 0 to level 3 for each user, sorted in ascending order. Most users paid greater attention to the changes that took place at level 3. This result is also confirmed by the number of EFs, with a maximum close to 84 EF. Changes at level 1 and level 2 achieved a similar percentage of EFs. Changes at level 0 were the worst in this test. Here, the maximum value of EF is close to 45% of all EFs.

The average transition time for each group of increases and decreases in chatbot exposure number ranged from 0 to 3. Slight differences can be observed in the “differences” groups (1 to 3), where the average time was 892 ms. With the zero SC, this time was around 842 ms, which indicates about a 6% increase in the speed of vision transition compared to that in groups 1 to 3. On the other hand, the average number of EFs tends to increase with an increase in difference. Starting from the difference at a level of 0, with about 11 EF, up to the difference at a level of 1, where the average number of EFs increases by nearly 3,5, we can observe a nearly 25% increase in the number of EFs. Next, level 2 presents an average EF of 16,5, while level 3 has an average of 19 EF. As we can see, subsequent levels tended to increase by 15% to 25% compared to each previous level. The direction of change in the amount of content within the chatbot shows a significant difference in the average eyesight transit time when the direction of change is downward. This time is, on average, 961.5 ms, which is almost 14% longer than that for the results without directional changes and almost 16% longer than the times for upward changes. The EF count here is also significant for upward changes. Upward changes oscillate close to approximately 20 EF in the downward direction and indicate no change when the number ranges from 11 to 12 EF. Based on the data for the average eyesight transit time, the results for pairs starting from 0 to 2 oscillated between 800 and 865 ms. On the other hand, pairs starting with 3 showed the longest time of transition from the user's own patterns to the chatbot, reaching an average of approximately 1000 ms. The highest average number of EFs in this case was achieved for the set change starting with 0, which provided an average of 23. The remaining three variants achieved an average number of EFs from 12 to 14.

Table 2. Average values of reaction time and EF for SD, Direction and Starts.

	S.D.				Direction			Starts			
	0	1	2	3	neutral	decrease	increase	0	1	2	3
Average of reaction time	842.9	892	892.5	891.4	842.9	961.6	822.6	852.6	803	864.7	1001.2
Average of EF	11.3	14.8	16.5	19	11.3	12.3	19.8	23	14.3	13	12.3

To compare the individual steps of the sequence based on the variables in Table 2 (EF with positive and negative details), we used a Mann–Whitney U test. The analysis is presented by comparing the three steps in the sequence with the division of three groups of pairs—from 1, representing the first occurrence in the sequence, to 3, representing the last occurrence in the sequence. Starting with a comparative analysis of step 1 compared to step 2, the significance of the parameter has a p-value of 0.03. For the dependent variable, when we compare 1 and 3, the p-value is 0.01. In comparing levels 2 and 3, there is no observable significance level. However, the p-value is close to 0.05, which is, in principle, almost equal to the adopted limit of significance (p-value < 0.05). Thus, based on accurate statistics, we can assume that there are statistically significant differences between the variables. Between steps 1 and 2 in the direction of changes, step no. 1 has greater importance. The next pair, 1 and 3, also favours step no. 1. Here, the value of step no. 1 is slightly weaker than that in the previous comparison. Nevertheless, this result highlights the significant importance of step no. 1. Between 2 and 3, the trend favours step no. 3. When analysing the two groups of EF and INEF, we used the time of transition as a dependent variable. Statistical significance in these two groups was determined at a p-value of 0.049. This level of statistical significance demonstrates a good fit, indicating a high probability of obtaining the differences that we observed in our study (even greater if the null hypothesis is true).

5. Conclusions

The main goal of this study was to examine the impact of changes in the chatbot area on user attention and the ability to direct attention to the chatbot interface. These impacts reflected the ways in which the content changes were represented in the chatbot area. The results based on decomposition of the chatbot messages into separate components showed that intensive changes in visual content were more effective than gradual updates. Increasing and decreasing changes in the field of view were equally more effective than gradual changes in the field of view. Changes involving the appearance of new messages in the chatbot area were characterized by about 60% more users being directed to the chatbot area compared to changes involving the disappearance of messages. This study also confirmed the intuitive assumption that the greater the changes that occur in the chatbot area, the greater the chatbot's ability to attract user attention. Gradual changes in the visual space were more difficult to detect. Thus, chatbot developers should consider displaying large blocks of content at a time rather than gradually display messages with a smaller field. Interestingly, the lower the percentage of the chatbot that was filled at the time of change, the greater the chatbot's effectiveness in attracting the user's attention. The appearance of new messages when the chatbot was empty (set change 0) was characterized by much more effective user attraction and a shorter reaction time to change compared to other cases. Therefore, when the chatbot intends to display more important information than previously displayed, the chatbot should first reset its area and then present a new message only after some time has passed to maximize the probability that the user will notice this information. In future research, we will explore how to draw the user's attention to messages as quickly as possible and properly fill the information areas in chatbots and other system messages. Future work might also study other influential features relevant to chatbots, such as message characteristics and animated elements.

Acknowledgements

The work was supported by the National Science Centre of Poland, the decision no. 2017/27/B/HS4/01216.

References

[1] Cole, G. G., & Kuhn, G. (2010). Attentional capture by object appearance and disappearance. *Quarterly Journal of Experimental Psychology*, **63(1)**: 147-159.

[2] Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & psychophysics,* **65(7)**: 999-1010.

[3] Ball, F., Elzemann, A., & Busch, N. A. (2014). The scene and the unseen: manipulating photographs for experiments on change blindness and scene memory. *Behavior research methods*, **46(3)**: 689-701.

[4] David, E., Laloyaux, C., Devue, C. and Cleeremans, A., 2006. Change Blindness to Gradual. Changes in Facial Expressions. *Psychologica Belgica,* **46(4)**: 253–268.

[5] Shawar, Bayan Abu, and Eric Steven Atwell. "Using corpora in machine-learning chatbot systems." *International journal of corpus linguistics 10.4 (2005)*: 489-516.

[6] Groves, P. M., & Thompson, R. F. (1970). Habituation: A dual-process theory. *Psychological Review, 77*, pp. 419–450

[7] Thompson, R. F., & Spencer, W. A. (1966). Habituation: A model phenomenon for the study of neuronal substrates of behavior. *Psychological Review*, **73**: 16–43.

[8] Otaki, Makoto, and Katsuyuki Shibata. "The effect of different visual stimuli on reaction times: performance comparison of young and middle-aged people." *Journal of physical therapy science 31.3 (2019):* 250-254.

[9] Kuligowska, K. "Commercial chatbot: performance evaluation, usability metrics and quality standards of embodied conversational agents." *Professionals Center for Business Research 2 (2015)*.

[10] Følstad, Asbjørn, Cecilie Bertinussen Nordheim, and Cato Alexander Bjørkli. "What makes users trust a chatbot for customer service? An exploratory interview study." *International conference on internet science. Springer, Cham, 2018.*

[11] Følstad, Asbjørn, and Petter Bae Brandtzæg. "Chatbots and the new world of HCI." *Interactions 24.4 (2017)*: 38-42.

[12] Owens, Justin W., Barbara S. Chaparro, and Evan M. Palmer. "Text advertising blindness: the new banner blindness?." *Journal of usability studies 6.3 (2011)*: 172-197.

[13] Chaves, Ana Paula, and Marco Aurelio Gerosa. "How should my chatbot interact? A survey on human chatbot interaction design." *arXiv preprint arXiv:1904.02743 (2019).*

[14] Cahn, Jack. "CHATBOT: Architecture, design, & development." *University of Pennsylvania School of Engineering and Applied Science Department of Computer and Information Science (2017)*.

[15] Rankin, Catharine H., et al. "Habituation revisited: an updated and revised description of the behavioral characteristics of habituation." *Neurobiology of learning and memory 92.2 (2009)*: 135-138.

[16] Ghose, Supratip, and Jagat Joyti Barua. "Toward the implementation of a topic specific dialogue based natural language chatbot as an undergraduate advisor." *2013 international conference on informatics, electronics and vision (ICIEV). IEEE, 2013.*

[17] Zarouali, Brahim, et al. "Predicting consumer responses to a chatbot on Facebook." *Cyberpsychology, Behavior, and Social Networking 21.8 (2018)*: 491-497.

[18] Dale, Robert. "The return of the chatbots." *Natural Language Engineering 22.5 (2016)*: 811-817.

[19] Hill, Jennifer, W. Randolph Ford, and Ingrid G. Farreras. "Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations." *Computers in human behavior 49 (2015)*: 245-250.

[20] Yan, Mengting, et al. "Building a chatbot with serverless computing." *Proceedings of the 1st International Workshop on Mashups of Things and APIs. 2016.*

[21] Xu, Anbang, et al. "A new chatbot for customer service on social media." *Proceedings of the 2017 CHI conference on human factors in computing systems. 2017.*

[22] Hervet, Guillaume, et al. "Is banner blindness genuine? Eye tracking internet text advertising." *Applied cognitive psychology 25.5 (2011)*: 708-716.

[23] Smith, Tim J., and John M. Henderson. "Edit Blindness: The relationship between attention and global change blindness in dynamic scenes." *Journal of Eye Movement Research 2.2 (2008)*.

[24] Simons, Daniel J., and Ronald A. Rensink. "Change blindness: Past, present, and future." *Trends in cognitive sciences 9.1 (2005)*: 16-20.

[25] Itti, Laurent, and Christof Koch. "Computational modelling of visual attention." *Nature reviews neuroscience 2.3 (2001)*: 194-203.

[26] Nuruzzaman, Mohammad, and Omar Khadeer Hussain. "A survey on chatbot implementation in customer service industry through deep neural networks." *2018 IEEE 15th International Conference on eBusiness Engineering (ICEBE). IEEE, 2018*.

[27] Cavanaugh, James, and Robert H. Wurtz. "Subcortical modulation of attention counters change blindness." *Journal of Neuroscience 24.50 (2004)*: 11236-11243

[28] Resnick, Marc, and William Albert. "The impact of advertising location and user task on the emergence of banner ad blindness: An eye-tracking study." *International Journal of Human-Computer Interaction 30.3 (2014):*206-219.

[29] Rensink, R. A. (2002). Change detection. *Annual review of psychology*, **53(1)**: 245-277.

[30] Kumar, Praveen, et al. "Designing and developing a chatbot using machine learning." *2018 International Conference on System Modeling & Advancement in Research Trends (SMART). IEEE*, 2018.

[31] Harkous, Hamza, et al. "Pribots: Conversational privacy with chatbots." *Twelfth Symposium on Usable Privacy and Security ({SOUPS}* 2016). 2016.

A8. Bortko, K., Fornalczyk K., Jankowski, J., Sulikowski P., Dziedziak K. (2023). Impact of changes in chatbot's facial expressions on user attention and reaction time. PLoS ONE

PONE-D-22-20948R2: Final Decision Being Processed - [EMID:39be625f311d8f5c]

Odebrane
Powiadomienia

PLOS ONE <em@editorialmanager.com>

śr., 21 cze, 01:01 (7 dni temu)

do mnie

Przetłumacz wiadomość

Wyłącz dla następującego języka: angielski

CC: "Kacper Fornalczyk" kacper_fornalczyk@wp.pl, "Jarosław Jankowski" jjankowski@wi.zut.edu.pl, "Piotr Sulikowski" psulikowski@wi.zut.edu.pl, "Karina Dziedziak" kdziedziak@wi.zut.edu.pl

Impact of changes in chatbot's facial expressions on user attention and reaction time
PONE-D-22-20948R2

Dear Mr. Bortko,

We're pleased to inform you that your manuscript has been judged scientifically suitable for publication and will be formally accepted for publication once it meets all outstanding technical requirements.

Within one week, you'll receive an e-mail detailing the required amendments. When these have been addressed, you'll receive a formal acceptance letter and your manuscript will be scheduled for publication.

An invoice for payment will follow shortly after the formal acceptance. To ensure an efficient process, please log into Editorial Manager at http://www.editorialmanager.com/pone/, click the 'Update My Information' link at the top of the page, and double check that your user information is up-to-date. If you have any billing related questions, please contact our Author Billing department directly at authorbilling@plos.org.

If your institution or institutions have a press office, please notify them about your upcoming paper to help maximize its impact. If they'll be preparing press materials, please inform our press team as soon as possible -- no later than 48 hours after receiving the formal acceptance. Your manuscript will remain under strict press embargo until 2 pm Eastern Time on the date of publication. For more information, please contact onepress@plos.org.

Kind regards,

Impact of changes in chatbot's facial expressions on user attention and reaction time

Kamil Bortko[1], Kacper Fornalczyk[1], Jarosław Jankowski[1], Piotr Sulikowski[1], Karina Dziedziak[1]

1. West Pomeranian University of Technology, Faculty of Computer Science and Information Technology, Żołnierska 49, 71-210 Szczecin, Poland

* kbortko@zut.edu.pl

Abstract

Communication within online platforms supported by chatbots requires algorithms, language processing methods, and an effective visual representation. These are crucial elements for increasing user engagement and making communication more akin to natural conversation. Chatbots compete with other graphic elements within websites or applications, and thus attracting a user's attention is a challenge even before the actual conversation begins. A chatbot may remain unnoticed even with sophisticated techniques at play. Drawing attention to the chatbot area localized within the periphery area can be carried out with the use of various visual characteristics. The presented study analyzed the impact of changes in a chatbot's emotional expressions on user reaction. The aim of this study was to observe, based on user reaction times, whether changes in a chatbot's emotional expressions make it more noticeable. The results showed that users are more sensitive to positive emotions within chatbots, as positive facial expressions were noticed more quickly than negative ones.

Introduction

Chatbots are widely used within online platforms. They are based on human intellectual principles and abilities that focus on conversation. An organization can benefit from a chatbot's unique knowledge based on consumer opinions in addition to its human-computer interaction (HCI) area background, allowing for the redistribution of cognitive tasks between humans and machines [1]. Chatbots integrate many characteristics, making them capable of eliciting different kinds of social responses to varying degrees, including verbal responses [2], gestures, and visual responses [3].

While language processing, machine learning, and artificial intelligence algorithms supporting communication within chatbots have been extensively analyzed [4, 5], only a limited number of studies have focused on their visual parameters [6]. Less focus has been given to the appearance of chatbots, including elements to attract a user's eyesight and attention, but the overall impact of chatbot design has been emphasized [7]. The growing popularity of chatbots has brought about new challenges regarding HCIs such as changing interaction patterns [8], and thus, the role of chatbot usability has also been discussed [9]. The need to further improve their performance is due to raising expectations that chatbots will display social behaviors that are customary for interpersonal communication.

The extent to which a medium, like a chatbot, is designed to resemble and behave like a human, incorporating elements such as a human-like appearance, facial expressions, and gestures, can significantly influence the perceived level of social presence during the interaction. Consequently, users may experience a heightened sense of engaging in a genuine and natural conversation, almost as if they were conversing with an actual human being rather than a machine.

The visualization of chatbots and their social features should target users' expectations to ultimately avoid frustration and dissatisfaction. The effects of an electronic conversation on human behavior and the perceived level of anthropomorphism are a part of the broad issue of human attitudes towards humanoid technologies.

According to the theory proposed by Mori, the more a character resembles a human, the more it is accepted by us and evokes positive feelings. However, when a character becomes too realistic but still has subtle differences, such as unnatural movements or improper proportions, we experience a sense of unease or rejection in our minds. This is the moment when we enter the so-called Uncanny Valley [10] [11].

Mori's theory suggests that the acceptance of artificial figures increases with their level of realism until a certain point, after which there is a sudden decline in acceptance. Only when a character reaches an exceptionally high level of realism, almost indistinguishable from a living human, does acceptance increase again.

The phenomenon of the Uncanny Valley has also been applied in the context of avatars with facial expressions. Irregularities in facial expression movements or inconsistencies with our expectations can make us feel uneasy and result in negative attitudes towards such avatars. Therefore, it is not surprising that the hypothesis regarding the Uncanny Valley has been adopted to explain the poor commercial success of some animated films in the media [12].

In a study conducted by Katsyri et al. (2015), existing research on people's reactions to artificial figures with varying degrees of realism was analyzed to investigate which hypotheses best explain this phenomenon. One hypothesis suggests that the feeling of unease in the Uncanny Valley arises from our social and cultural context. If artificial human-like figures are perceived as strange or inappropriate in our society, they can evoke negative emotions [13]. Study of Kao et al. (2019) finds that avatars with higher anthropomorphism led to higher player experience. Avatars with higher anthropomorphism also led players to identify more highly with their avatars. Independent of avatar type, we find avatar identification significantly promotes player experience. Players playing games doomed by little humanoidness. We will be more successful when the avatar is more a human [14].

Another factor potentially influencing chatbot design is that in natural communication, participants pay attention to emotions expressed through voice and movements such as gestures and facial [15]. The same mechanisms transferred to a virtual environment can be related to emotions expressed within a chatbot visualization. The content displayed in relation to human perception is intended to arouse feelings and the desire to interact with the chatbot [16]. A conversation with a graphic avatar in the chatbot field [17] will have a different impact on a user than a conversation with a human visualization, and may increase the fluency of conversation. Facial expressions improve the effectiveness of avatars in different contexts [18] and are also a way of influencing people's judgment [17]. Naturalness is important, as evidenced by how well a dialogue system can follow a natural course of conversation. A method of evaluating naturalness in conversational dialog systems has been proposed [19] based on a chatbot that summarizes the user's emotional state in a survey. The percentage of the chatbot's facial expressions followed by the user during the conversation with the chatbot and the interaction with the content were analyzed. The main aspects studied were the communication between the chatbot and the human, whether the chatbot should be more of an agent, etc. The chatbot's visual area itself, however, leaves room for a deeper analysis [7]. The results of studies on chatbot visualization have emphasized the importance of the visual appearance of the chatbot [20].

The main goal of the presented study was to verify the impact of emotional changes in the chatbot's face on user attention and reaction time. Earlier studies demonstrated the impact of textual content change intensity within the chatbot on user attention [21]. Expressing emotions can be one of the techniques to attract user attention, and the question is which emotional state can effectively attract user attention during other tasks without the risk of there being negative effects on user perception [22]. Chatbots are usually located within corners of the screen and the peripheral vision area. Our experiment integrates a task within the primary vision area and the presence of the chatbot in the peripheral area. The primary task requires the participant's attention, and changes within the chatbot are carried out in the periphery [23]. In real systems, chatbots (as well as system messages or in-video-game messages) that are aimed at helping users or players, are usually located peripherally. Considering that, they require the users to perform a more accurate peripheral search and identification tasks while performing a central task [24]. This is in line with the attention-utility issue discussed by McCrickard et al. [25] in the context of peripheral design, which refers to the ways in which peripheral cues or messages help users achieve their goals without requiring their full attention. Utility refers to a system's usefulness to its users or customers. McCrickard et al. discussed utility as a value provided by the peripheral system as a whole, did not directly manipulate utility as part of their experiment, and focused on attention-utility trade-off, considering attention as a constrained resource that can be traded for some utility [26]. We would like to consider utility as the meaning of the content of the "individual gaps", which refer to specific needs or requirements of individual users that may not be fulfilled by the existing attention-utility of the system. By addressing these individual gaps, the authors suggest that the attention-utility of the system can be increased and adapted to better meet the diverse needs of users. Utility can be evaluated as an aspect of human-computer interaction for the purpose of identifying aspects of this interaction that can be improved with the help of evaluation methods [27] in different experimental settings. In studying various peripheral or secondary display information representations, many researchers focus only on the information gained without measuring the changes in the primary task performance caused by these display [27].

Similarly, some studies create an unrealistic testing environment given that the distinction between reaction and comprehension is unclear. While user reaction and comprehension tasks are often closely related, the two objectives may imply differences in the notification system information design.

In everyday life, signals such as facial expressions often appear in our peripheral field of vision. Although the processing of facial expressions within the central vision has been widely researched [28], fewer studies have focused on processing objects within the peripheral vision. To date, research has been consistent about the specific and automatic processing of information about positive emotions in the peripheral vision, which can draw attention to emotion-enhanced messages and enable a quick behavioral response [28]. One study showed a decline in recognition and detection performance as eccentricity increases, with happiness and surprise being the best recognized expressions in the peripheral vision. As for detection, however, another well-detected expression is fear, along with happiness and surprise [29].

Based on the presented experiment, we show that positive emotions, such as happiness and surprise in a human chatbot representation, are better detected in terms of reaction time than other expressions. The results show that task constraints shape the perception of expressions in the peripheral vision and provide new evidence that detection and recognition rely on partially separate underlying mechanisms, with the latter being more dependent on the higher spatial content frequency of the facial stimulus.

Experiment setup

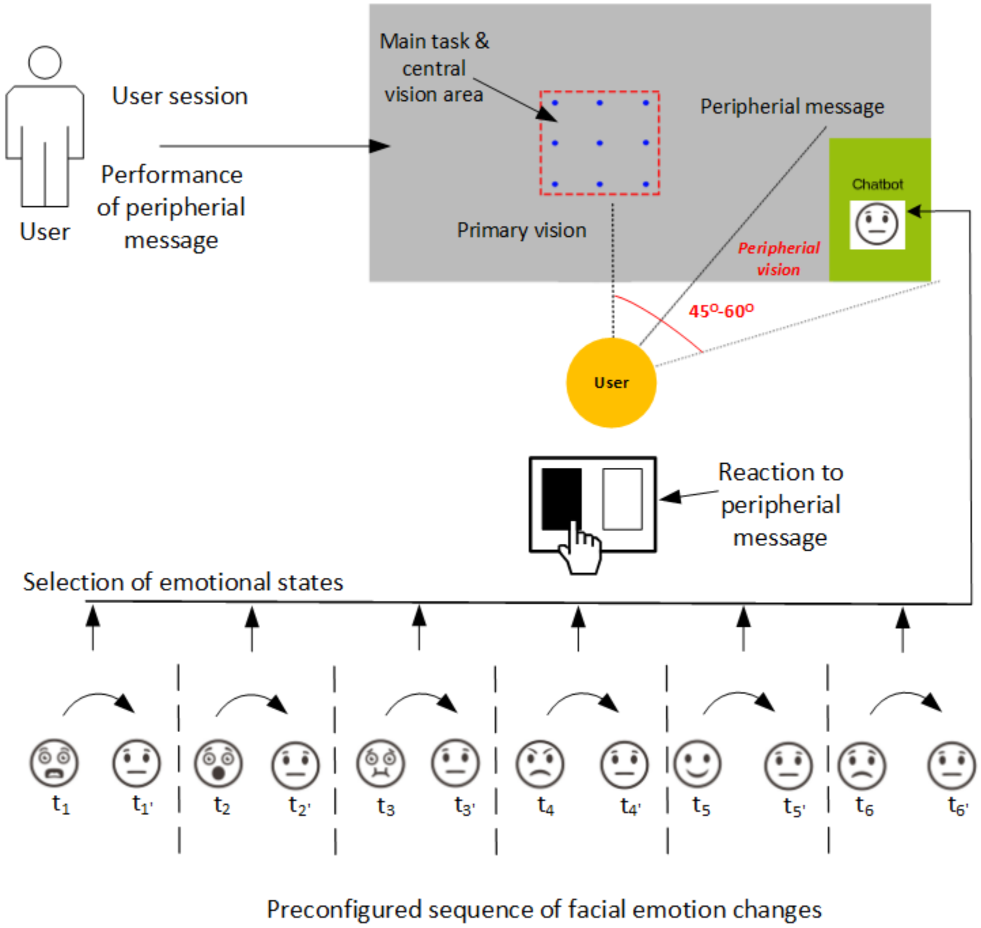


Fig 1. Each study participant's experimental session was based on a primary task within the central vision area and a chatbot component located in the periphery, in the bottom right-hand corner of the screen. A user was performing tasks, and different emotions were expressed within the chatbot area. Each change in facial expressions noticed in the chatbot area was reported by the user's mouse left click. Emotions were always changed from a neutral state, and after each change the neutral state returned.

The main goal was to study the user's reaction to changes in the emotional state of a chatbot located within the area of the peripheral vision. The structure of the experiment is presented in Figure 1. The screen area was organized into two sections: the center of the screen showing the primary task, and the chatbot area in the periphery. The experiment was set up so that the participants could not look directly at the chatbot due to validation and real-time control of the coordinate position of gaze patterns. The experiment was planned to last 15 minutes per user, and each user session was divided into four parts. In the first part, the user's reaction was examined on the basis of Male Face 1. The reaction was pressing the mouse button when a change in the object (in our task, a photo of a face with an expressed emotion) was observed in the chatbot located in the peripheral area. In the second and third parts, Female Face 1 and 2 were used, and in the fourth part, Male Face 2 was used. For each part of the experiment, the individual six emotions were displayed in the following order: happiness, sadness, anger, fear, surprise, and disgust. Each of these emotions was displayed twice. The appearance of each emotion was preceded by the display of a photo with a neutral expression on the face, which was the so-called zero state. The neutral state separated the emergence of each emotion to reset its impact. The change in the photo in the chatbot area happened after a random time of 5–12 s.

The experiment involved 20 people aged 20-25. Of study participants, 50 per cent were female, and 50 per cent were male. The inclusion criteria were defined as 'participants with normal or corrected-to-normal vision and between 18 and 35 years old'. The exclusion criteria were defined as 'participants with a history of eye disease or neurological disorder, or any other condition that could affect the results of the study. The experiment used a 27-inch Dell monitor and the Tobii Pro X3 eye tracker with a sampling frequency of 120 Hz. A special stand with a tripod was prepared for the experiment, which made it possible to keep the head of each participant in a stationary position. The eye-tracking software used was Tobii Pro Lab, version 1.194. The participant sat in front of the monitor at a distance of about 54 cm, which made it possible to create a 45-degree angle between the center of the screen and the edge of the chatbot, which was located in the bottom right-hand corner of the screen. As a result, the chatbot, as assumed, was within the peripheral vision of each of the respondents. The eye tracker was calibrated so that the monitoring software would react at the right moment whenever the participant's eyesight left the designated field and returned to the correct area. This was the role of the 'overseer'. Its main purpose was to keep the user's gaze within the designated frame and not looking away into the chatbot's field.

The task of the user was to stare at the pulsating red dot in the center of the screen. They could not look outside the designated area, otherwise a warning was displayed, and until their gaze returned to the designated area, the further display of photos was suspended. When staring at the pulsating dot, the user would notice a change taking place within the chatbot in the corner of their eye, at which point they had to left-click their mouse. The information about the click time was saved by the software to a text file, which also contained accurate information about the times when individual emotions were displayed. Thus, after determining the difference between those two values, the user's reaction time to a given emotion was calculated and understood as the time between a stimulus and a response. Information was also recorded when the user looked outside the designated area and when they looked at the chatbot. There were cases when the user did not notice the changes taking place in the chatbot area - this was also recorded. The analyzed data were grouped and aggregated into three categories based on common characteristics. We included results whose transition times ranged from 100 to 2000 ms, based on a study that excluded reaction times that were <100 ms as 'too early' and reaction times that were >2000 ms as 'too late' (i.e., not a direct reaction to the stimuli) [30].

Results

Analysis for each emotional expression

For the analysis, we used the variables regarding changes in the chatbot field for individual time intervals, which were grouped into four variants: four repetitions, and groups of emotions (i.e., positive and negative). In Table 1, which reflects the response to changes in the chatbot field using Mann–Whitney U statistical analysis, we can see that the intergroup comparison shows a statistical significance below $p < 0.05$ in a few cases.

The intergroup comparison of the emotion of surprise with the rest of the emotions shows significance in virtually every case, except for the emotion of happiness. This indicates little variation in reaction times.

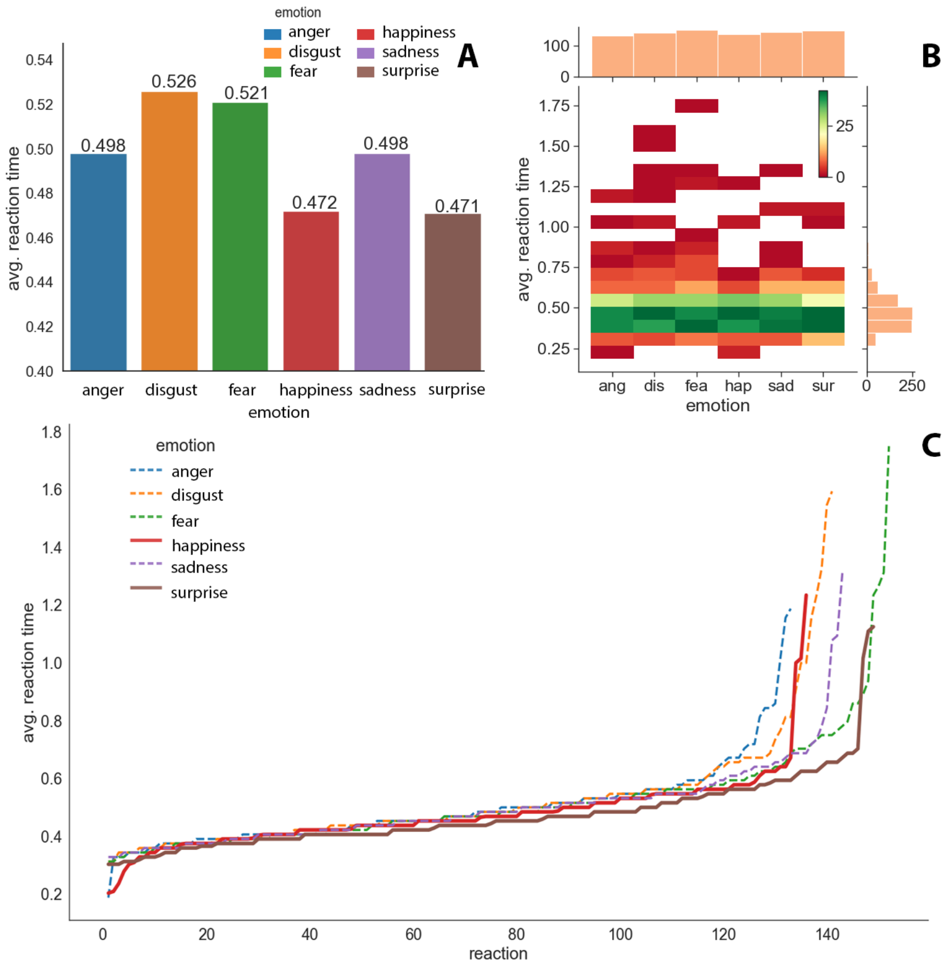


Fig 2. Barplot with average time reaction for each emotion (A) heatmap for each emotion (B) and line graphs (C) showing individual reaction times sorted ascending for each emotion (positive emotions presented as continuous curves and negative emotions as broken curves).

These emotions are both positive and resulted in a similar average response time. Starting the analysis at the base, we can conclude, as shown in Figure 2 (A), that the average user reaction time to the change in emotions was about 0.499 s, with the reaction time to the face expressing disgust (0.526 s) and fear (0.521 s) being the longest. Users reacted the fastest to the faces expressing surprise (0.471 s) and happiness (0.472 s), while faces expressing anger and sadness resulted in reactions at a slightly slower speed (0.498 s).

Figure 2 (B) shows the frequency of reaction times obtained for each of the emotions.

Table 1. Intergroup comparision of emotions

Expression	**anger**	**disgust**	**fear**	**happiness**	**surprise**	**sadness**
anger	x	.705(.37)	.511(.65)	.278(1.08)	.066(1.83)	.891(.13)
disgust		x	.795(.25)	.127(1.52)	.028(2.19)	.751(.31)
fear			x	.067(1.82)	.014(2.44)	.587(.54)
happiness				x	.457(.74)	.217(1.23)
surprise					x	.051(1.95)
sadness						x

A comparison of all emotions. The table shows the significance ($p < 0.05$) or lack thereof of the comparison and its strength, i.e., the value, in parentheses.

The width of reaction time interval is 0.08 s. The legend in the upper right-hand corner indicates which color corresponds to the frequency of the recorded reactions in the given time interval. It can be seen here that the vast majority of reaction times ranged from 0.36 to 0.52 s. In this interval, most of the reactions were to the emotion of surprise. The histogram on the side shows that, above 0.5 s, the number of reactions decreases with the duration of the reaction time. After about 0.7 s, the number of these reactions, regardless of the emotion, is negligible. The top histogram shows the sum of correct reactions obtained by all users. On average, there were 142.5 reactions per emotion.

Figure 2 (C) shows all reaction times for each emotion. The times for each emotion were sorted from best to worst. Solid lines are used to highlight the reactions to the positive emotions of surprise and happiness. Emotions from the negative group are presented with dashed lines. One can see here that, regardless of the emotion, most of the results are similar, but a significant portion of the reactions to positive emotions achieved better results compared with the other emotions. The emotion of surprise stands out here, in particular, which is almost always below the other lines. The line of happiness is higher than that of surprise, but still stands apart from the negative emotions.

Analysis for Aggregated Positive vs. Negative Emotions

As it was mentioned, the emotions were divided into two groups: positive and negative. Positive emotions are a group of positively associated emotions: happiness and surprise. Negative emotions, in this case, are non-positive emotions: anger, fear, disgust, and sadness. Figure 3 shows the results obtained for aggregated positive and negative emotions.

The results of the overall averages indicate that the reactions to positive emotions were much faster than those to negative emotions.

Figure 3 (A) shows all reaction times for each type of aggregated emotion. The times for each type were sorted from best to worst. Solid lines represent the average results for the group. The chart also shows the area defined by the worst and best reaction outcomes from a given group. It is evident here that the group of positive emotions resulted in faster reaction times. From the areas behind the lines, it can be seen that the results for the group of positive emotions were also less varied for each of the emotions than those for the group of negative emotions, but this may be influenced by the fact that there are twice as many emotions in the negative group.

The analysis in Figure 3 (B) of the intergroup comparison showed that the significance group fluctuated at the level of $p < 0.003$. This proves a clear difference between the two groups. The Z factor in this case was 2.913. The average response times for the negative and positive emotion groups were 0.511 s and 0.472 s, respectively. This indicates faster responses to faces that show positive emotions.

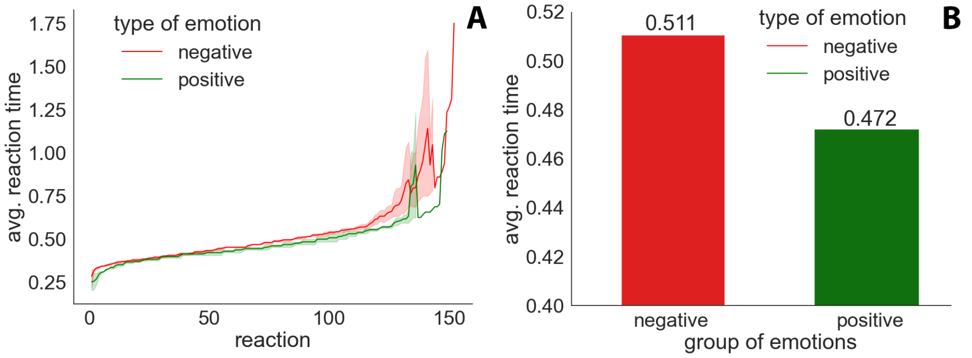


Fig 3. Figure (A) shows individual average reaction times (sorted ascending) for aggregated positive and negative emotions along with the coverage for individual emotions in each aggregate, and Figure (B) shows the average time for each group of emotions.

Analysis of Sequences of Emotions

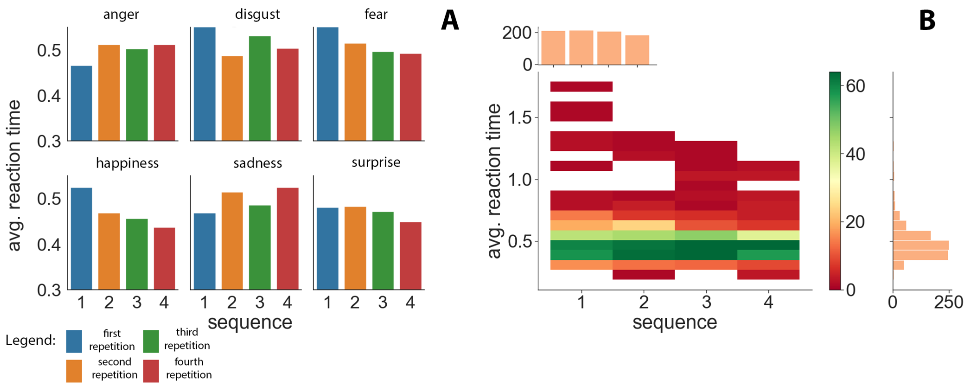


Fig 4. Figure (A) shows the average reaction time for each emotion in relation to the sequence number, and the heatmap (B) shows the frequency of reaction times for each of the sequences for all emotions combined.

Figure 4 (A) shows the average reaction time for each emotion in relation to the sequence number, and the heatmap in Figure4 (B) shows the frequency of reaction times obtained for each of the sequences for all emotions, with the width of the reaction time interval being 0.08 s. The legend on the right indicates what frequency of recorded reactions in a given time interval corresponds to each color.

One can infer interesting conclusions from Figure 4 (A). This figure implies that for positive emotions of happiness and surprise, a decrease in the average reaction time with each repetition was observed; e.g., the reaction time in the case of happiness dropped from being close to 0.51 s to 0.45 s. For the negative emotion of fear, the reaction time decreased with each repetition from 0.55 to 0.48 s. The other negative emotions showed no similar trends. The emotion of anger at the first repetition reached a reaction time of about 0.46 s, after which this time stabilized at a level of nearly 0.5 s. In turn, for the disgust and sadness emotions, highly inconsistent results were observed.

Therefore, with regard to the emotions of happiness, surprise, and fear, users tended to 'learn' in order to elicit a quicker response. The remaining emotions, however, did not show such tendency. Rather, they were highly variable. Globally, by dividing emotions into positive and negative aggregates, we can conclude that the group of positive emotions showed a learning trend. The positive emotions had a stronger influence on the user than the negative emotions, only one of which showed a learning trend.

Figure 4 (B) shows the frequency of reaction times obtained for each of the repetitions for all emotions, with the width of the reaction time interval being 0.08 s. The legend on the right indicates what frequency of recorded reactions in a given time interval corresponds to each color.

Figure 4 (B) depicts the frequency of reaction times observed for all emotions in a sequence. Based on the side histogram, most reaction times oscillated in the range between 0.36 s and 0.52 s. This was most pronounced in Sequence 3. Sequence 1 seemed to be the most diverse in terms of the frequency of reaction times at different intervals. Furthermore, with each successive sequence, an increasing number of reactions were recorded mainly in the aforementioned time interval from 0.36 s to 0.52 s. There was an average of 214 reactions per sequence. It can also be seen that, in the last two sequences, the number of reactions started to decrease. In the successive sequences, 221, 223, 217, and 193 reactions were recorded sequentially.

Discussion and Conclusion

The main goal of this study was to investigate how the user's reaction time was affected by a visual change in emotions expressed through facial expressions of a chatbot represented by real human face photos located within the user's peripheral vision. Fraser W. Smith and Stephanie Rossit showed that different emotions are recognized and detected with varying results, which encouraged our studies [29]. Our results show that faces expressing happiness and surprise were noticed the fastest. Their results convincingly demonstrated that task constraints shape perception of expression in the peripheral vision and provided new evidence that detection and recognition relies on separate underlying mechanisms, with recognition being more stimulus-dependent [29]. Bayle et al. showed that humans have the ability to detect the presence and type of emotion according to the facial expressions of a presented photo in the far periphery [28]. Much research has focused on affect rather than how humanoid objects are perceived within chatbots [11] [31] [32].

Both of these emotions of happiness and surprise can be classified as positive, while the others, which were also present in the study, i.e., anger, disgust, sadness, and fear, are classified as negative or pejorative. By analyzing the emotions in groups, the study showed that changes to emotions in the positive group were noticed 7.6% faster than those in the negative group. As the emotion changes were displayed to users in sequences, the study revealed that the emotions of happiness, surprise, and fear with each successive change led to a decrease in the average response time of users. No tendency was noted for the remaining emotions. Without classifying the emotions, the average response time to visual changes in a given sequence still decreased with each successive change. The average reaction time for the last sequences compared to the first ones was 6.2% faster. The reason for this may be an effect of the user learning/recognizing a given emotion. This study suggests that, when developing a humanoid chatbot model that can express certain emotions, it is worth considering which emotions should be selected for specific messages if a specific and prompt reaction in the user is desired.

We have shown that some expressions are detected faster in the peripheral vision, i.e., happiness and surprise, and some are detected more slowly, i.e., anger, sadness, disgust, and fear, the last two showing the longest reaction times. Our work examined the performance of activities occurring within the user's peripheral vision.

The researched area is represented in a small number of studies. This paper is an extension of the research to date, and is particularly related to Bayle's work on changes in emotional states [28]. This follow-up study was based on adding humanoid features to the chatbot area with the main focus on emotional changes, which proved to be an effective way to catch the attention of the user in another area [29].

An important aspect of chatbots is to increase the naturalness of their assistance and communication with the user. McCrickard et al. [25] discussed utility as a value provided by the peripheral system as a whole and did not directly manipulate utility as part of their experiment, whereas we considered utility as the meaning of the content of the individual gaps; therefore, the concept of utility should be better defined. The emotions that are shown in the chatbot area clearly influence the user's reaction speed. It was found that particular emotions seem to stimulate or slow down the interactions located in the peripheral vision

In terms of guidelines for interface designers, recommender system programmers, and content creators of chatbots based on research, they should consider making the images of people representing the chatbot positive. This will affect reaction speed.

Game designers may also create guidelines for their creations. Messages appearing with adequate avatars that express different emotions in the peripheral vision may affect the speed of the player's reaction. In the excitement of gameplay in MMO (massively multiplayer online) games or even simulators, e.g., piloting a plane or driving a vehicle, every fraction of a second counts. In MMO games there are often various messages with different icons and the gamer needs to perceive peripheral changes to be able to react quickly, e.g., by pressing key combinations or by clicking the appropriate icon, or in some games, e.g., StarCraft, the character's avatar at the bottom of the screen suggests their 'damage', which can affect the decisions made by the user. One way to combine theories about avatars in chatbots and simulators of simulation tasks in relation to the peripheral vision is to use the avatar's peripheral vision as a way to display information about the simulation task.

For example, in a virtual reality driving simulation, the user's avatar could have a heads-up display that shows their speed and other important information, but it could also use the avatar's peripheral vision to display information about the environment, such as traffic signals or other vehicles in the vicinity. This could help the user to more easily stay aware of their surroundings and make more informed decisions while driving. Additionally, the chatbot feature of the avatar could be used to provide verbal cues or warnings, making use of the peripheral vision in a more natural way. By introducing appropriate content expressing individual emotions, we can aim to extend the reaction time or effectively shorten it. However, users should not be overloaded with other stimuli, e.g., smiles and humor, within the chatbot area, because repeated and too invasive representations of emotions often result in habituation [33] [34] [35].

In the presented study, we aimed to analyze the performance of a peripherally located chatbot, represented by human face photos, in terms of the impact of changes in the chatbot's emotional expressions on the user's reaction. The goal in itself was to study the user's reaction to changes in the chatbot field; however, users had no idea that the changing photos in the peripheral vision represented different emotions. Our earlier study was based on the impact of changes in textual-only content within the chatbot, and we view this study as an extension to that research. The results showed that users are more sensitive to positive emotions within chatbots, as they are noticed more quickly than negative facial expressions.

The presented study raises questions for further research related to the impact of techniques used for avatar visualization, e.g., simplified forms vs. realistic avatars, on user reactions. Other research could also focus on another form of emotion expression, e.g., avatar gestures or symbols conveying emotional messages.

Institutional Review Board Statement

The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Ethics Committee of Bioethical Commission of Pomeranian Medical University, Szczecin (09.03.2020); Bioethics Committee Agreement No. KB-0012/24/2020.

Informed Consent Statement

Informed consent was obtained from all subjects involved in the study. Written informed consent was obtained from the patient(s) to publish this paper. The research described for the papers was accepted by Bioethics Committee: agreement No. KB-0012/24/2020.

Acknowledgements

This work was supported by the National Science Centre of Poland, the decision no. 2017/27/B/HS4/01216 and by ZUT Highfliers School /Szkoła Orłów ZUT/ project co-ordinated by Dr. Piotr Sulikowski, within the framework of the program of the Minister of Education and Science /Grant No. MNiSW/2019/391/DIR/KH, POWR.03.01.00-00-P015/18/, co-financed by the European Social Fund, the amount of financing PLN 1,704,201,66.

References

1. Zhang P, Benbasat I, Carey J, Davis F, Galletta DF, Strong D. Human-computer interaction research in the MIS discipline. Communications of the Association for Information Systems. 2002;9(20):334–355.

2. Reeves B, Nass C. How people treat computers, television, and new media like real people and places; 1996.

3. MacDorman KF, Minato T, Shimada M, Itakura S, Cowley S, Ishiguro H. Assessing human likeness by eye contact in an android testbed. In: Proceedings of the XXVII annual meeting of the cognitive science society. Citeseer; 2005. p. 21–23.

4. Cahn J. CHATBOT: Architecture, design, & development. University of Pennsylvania School of Engineering and Applied Science Department of Computer and Information Science. 2017;.

5. Hussain S, Ameri Sianaki O, Ababneh N. A survey on conversational agents/chatbots classification and design techniques. In: Workshops of the International Conference on Advanced Information Networking and Applications. Springer; 2019. p. 946–956.

6. Kuligowska K. Commercial chatbot: performance evaluation, usability metrics and quality standards of embodied conversational agents. Professionals Center for Business Research. 2015;2.

7. Følstad A, Nordheim CB, Bjørkli CA. What makes users trust a chatbot for customer service? An exploratory interview study. In: International conference on internet science. Springer; 2018. p. 194–208.

8. Chaves AP, Gerosa MA. How should my chatbot interact? A survey on human-chatbot interaction design. arXiv preprint arXiv:190402743. 2019;.

9. Følstad A, Brandtzæg PB. Chatbots and the new world of HCI. interactions. 2017;24(4):38–42.

10. Mori M. Bukimi no tani [the uncanny valley]. Energy. 1970;7:33–35.

11. Ciechanowski L, Przegalinska A, Magnuski M, Gloor P. In the shades of the uncanny valley: An experimental study of human–chatbot interaction. Future Generation Computer Systems. 2019;92:539–548.

12. Burleigh TJ, Schoenherr JR. A reappraisal of the uncanny valley: categorical perception or frequency-based sensitization? Frontiers in Psychology. 2015;5:1488.

13. Kätsyri J, Förger K, Mäkäräinen M, Takala T. A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness. Frontiers in psychology. 2015;6:390.

14. Kao D. The effects of anthropomorphic avatars vs. non-anthropomorphic avatars in a jumping game. In: Proceedings of the 14th international conference on the foundations of digital games; 2019. p. 1–5.

15. Hatfield E, Rapson RL, Le YCL. Emotional contagion and empathy. The social neuroscience of empathy. 2011; p. 19.

16. David B, Chalon R, Zhang B, Yin C. Design of a collaborative learning environment integrating emotions and virtual assistants (chatbots). In: 2019 IEEE 23rd International Conference on Computer Supported Cooperative Work in Design (CSCWD). IEEE; 2019. p. 51–56.

17. Angga PA, Fachri WE, Elevanita A, Agushinta RD, et al. Design of chatbot with 3D avatar, voice interface, and facial expression. In: 2015 international conference on science in information technology (ICSITech). IEEE; 2015. p. 326–330.

18. Salem B, Earle N. Designing a non-verbal language for expressive avatars. In: Proceedings of the third international conference on Collaborative virtual environments; 2000. p. 93–101.

19. Carranza KALR, Manalili J, Bugtai NT, Baldovino RG. Expression tracking with OpenCV deep learning for a development of emotionally aware Chatbots. In: 2019 7th international conference on robot intelligence technology and applications (RiTA). IEEE; 2019. p. 160–163.

20. Morrissey K, Kirakowski J. 'Realness' in chatbots: establishing quantifiable criteria. In: International conference on human-computer interaction. Springer; 2013. p. 87–96.

21. Fornalczyk K, Bortko K, Jankowski J. Improving User Attention to Chatbots through a Controlled Intensity of Changes within the Interface. Procedia Computer Science. 2021;192:5112–5121.

22. Sulikowski P, Ryczko K, Bak I, Yoo S, Zdziebko T. Attempts to Attract Eyesight in E-Commerce May Have Negative Effects. Sensors. 2022;22(22):8597.

23. Gluck J, Bunt A, McGrenere J. Matching attentional draw with utility in interruption. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 2007. p. 41–50.

24. Chung S, Sparks JV. Motivated processing of peripheral advertising information in video games. Communication Research. 2016;43(4):518–541.

25. McCrickard DS, Catrambone R, Chewar CM, Stasko JT. Establishing tradeoffs that leverage attention for utility: Empirically evaluating information display in notification systems. International Journal of Human-Computer Studies. 2003;58(5):547–582.

26. McCrickard DS, Chewar CM. Attuning Notification Design to User Goals and Attention Costs. Commun ACM. 2003;46(3):67–72. doi:10.1145/636772.636800.

27. Gray WD, Salzman MC. Damaged merchandise? A review of experiments that compare usability evaluation methods. Human–computer interaction. 1998;13(3):203–261.

28. Bayle DJ, Schoendorff B, Hénaff MA, Krolak-Salmon P. Emotional facial expression detection in the peripheral visual field. PloS one. 2011;6(6):e21584.

29. Smith FW, Rossit S. Identifying and detecting facial expressions of emotion in peripheral vision. PloS one. 2018;13(5):e0197160.

30. Otaki M, Shibata K. The effect of different visual stimuli on reaction times: a performance comparison of young and middle-aged people. Journal of physical therapy science. 2019;31(3):250–254.

31. Atiyah A, Jusoh S, Alghanim F. Evaluation of the Naturalness of Chatbot Applications. In: 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT). IEEE; 2019. p. 359–365.

32. Jenkins MC, Churchill R, Cox S, Smith D. Analysis of user interaction with service oriented chatbot systems. In: International conference on human-computer interaction. Springer; 2007. p. 76–83.

33. Bortko K, Bartków P, Jankowski J, Kuras D, Sulikowski P. Multi-criteria Evaluation of Recommending Interfaces towards Habituation Reduction and Limited Negative Impact on User Experience. Procedia Computer Science. 2019;159:2240–2248.

34. Sulikowski P. Evaluation of Varying Visual Intensity and Position of a Recommendation in a Recommending Interface Towards Reducing Habituation and Improving Sales. In: Chao KM, Jiang L, Hussain OK, Ma SP, Fei X, editors. Advances in E-Business Engineering for Ubiquitous Computing. Cham: Springer International Publishing; 2020. p. 208–218.

35. Sulikowski P, Zdziebko T. Deep Learning-Enhanced Framework for Performance Evaluation of a Recommending Interface with Varied Recommendation Position and Intensity Based on Eye-Tracking Equipment Data Processing. Electronics. 2020;9(2):266.

A9.

Bartków, P., Bortko, K., Jankowski, J., Pazura, P. (2023). Modeling the impact of the habituation effect on information spreading processes with repeated contacts under an SI model. Plos one, 18(4), e0280266.

RESEARCH ARTICLE

Modeling the impact of the habituation effect on information spreading processes with repeated contacts under an SI model

Piotr Bartków*, Kamil Bortko, Jarosław Jankowski, Patryk Pazura

West Pomeranian University of Technology, Faculty of Computer Science and Information Technology, Szczecin, Poland

* pbartkow@zut.edu.pl

Check for updates

OPEN ACCESS

Citation: Bartków P, Bortko K, Jankowski J, Pazura P (2023) Modeling the impact of the habituation effect on information spreading processes with repeated contacts under an SI model. PLoS ONE 18(4): e0280266. https://doi.org/10.1371/journal.pone.0280266

Editor: Hocine Cherifi, Unviersity of Burgundy, FRANCE

Received: May 4, 2022

Accepted: December 23, 2022

Published: April 12, 2023

Data Availability Statement: All used data can be found on: https://github.com/Vladislav-The-Poker/Responsiveness-drop-data.

Funding: This work was supported by the National Science Centre of Poland, the decision no. 2017/27/B/HS4/01216. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

People are exposed to information from different sources whether or not such exposure is desired. Due to a limited ability to process information, only parts of the messages may be absorbed, and other parts may be ignored. Repeated stimuli lead to lower responses due to the learning process and the habituation effect. While this effect has been intensively studied, mainly in relation to visual stimulus, it is also incorporated in information spreading processes. Information spreading models often assume the possibility of repeated contact, but no habituation effect, which lowers the level of responsiveness of nodes in the network, has been implemented. Here, we study the impact of the habituation effect on information spreading with a susceptible–infected (SI) model, which is often the basis for other models. We assume that a decrease in habituation has an impact on propagation processes. Analysis of the results shows that the course of these propagation processes behave differently, significantly worsening their results. These processes are very sensitive, even to small changes in the level of habituation.

1. Introduction

Electronic media are used extensively for marketing operations, often leading to marketing content overload. Users experience advertising clutter with many commercial messages delivered within portals and social media [1]. The perceived intrusiveness of marketing content may negatively impact brand awareness, and overall performance [2]. It often leads to advertising avoidance within social media and the usage of ad blocking software [3]. From the perspective of perception and a limited ability to process information, content is filtered, and only a limited number of messages is absorbed [4]. In the area of visual advertising, banner blindness is resulting in ignored marketing content within visual spaces [5, 6].

The habituation effect is one of the reasons for reduced response [7]. It was initially analyzed from the perspective of biological systems and can be understood as a form of basic learning [8]. While the habituation effect was identified and studied in the 1960s [9], new goals and directions have been identified, including new ways of separating habituation from sensory adaptation or fatigue [10].

Apart from empirical experiments, the need for new predictive models is emphasized [11]. Even at an early stage of research, simulation models were used for modeling synaptic mechanisms [12]. Differential equations from first applications [13] were extended to model inter-stimulus intervals [14] or the impact of long-term memory [15].

Apart from biological systems, habituation was taken into account in artificial systems and robots [16]. The purpose of the model is to represent visual attention for computer programs or robots [17]. Novelty detection algorithms were also inspired by habituation studies [18]. Applications of habituation mechanisms to machine-learning processes were implemented to make the learning process closer to biological systems, because, even in such case, a decrease in the responsiveness of the learning process has been observed over time [19]. Models of habituation were also used for multi-armed bandits algorithms for marketing online content delivery optimization [20]. Recent efforts were related to predictability [21], visual stimulus [22], and modeling emotional habituation [23].

Besides visual communication and display advertising, repeated stimuli are also typical for word-of-mouth actions within real or digital social networks. The social network structure allows for the flow of various kinds of content. This can be any information, idea, visual content, or viral movie. Social network members perceive repeated exposures, and its impact on information spreading was analyzed [24]. Repetitions can deliver a cumulative impact on consumer behavior and increase the probability of purchase [25]. This results in the extension of the influence maximization problem towards repeated contacts. The cumulative influence was also analyzed for threshold models, and pieces of information received by users in each step are accumulated before the final decision takes place [26]. Multiple received signals were used as an extension of single activation models to reach threshold zones [27]. Analogies can be found in epidemiology research, where transmission probability is related to a number of contacts with an infected person [28]. Smieszek et al. focused on contact repetition and proposed an extension of the SIR model [29]. Earlier, the deterministic epidemic model taking into account repeated contacts was proposed, and the repetition impact of spreading effectivity was analyzed [30]. Repetitions of periods of partnership contact are also typical for sexually transmitted diseases [31].

While various models are used to represent behaviors based on repeated messages, the drop in response after repeated messages was not taken into account. Most models assume that repeated massages will not decrease the effectiveness of the spreading process but rather increase. While a cumulative impact can increase performance, from another perspective, repeated communication can be perceived as unwanted, and the probability of purchases can generally decrease [32]. Incentivized viral campaigns generate a high number of repeated contact, and the performance represented by conversion rate can decrease [33]. Intensive informational campaigns focused on changes in social behavior, for example, changes in behavior during a pandemic, are performed [34]. They do not always deliver the expected results, and habituation can be one of the reasons for a decreased response.

In terms of spreading within networks, the habituation effect was modeled earlier under an Independent Cascades Model [35]. The IC model proposed in [36] assumes that repeated contacts are observed only when communication with the same content comes from other different users. A single user has only one chance to activate other user, so repetition between two users never exists.

In the current study, we focused on a situation that is more common for information spreading: when a repeated message can flow between the same nodes. We adopted a susceptible–infected (SI) epidemic model, which has been discussed and used for information spread in many studies [37, 38]. In the basic SI model, a failed attempt to transmit information or a virus does not affect the probability of success of subsequent attempts. The difference in our model is that, as a result of the habituation process, each failed attempt lowers the chances of infection in the next step. It assumes repetitions between the same users, which represents a situation when several attempts to deliver a message are taken between two network members. The habituation effect was integrated within the model, and experiments were performed for different parameters of response decrease.

2. The habituation effect and its integration within the susceptible–infected (SI) model

Habituation is a cognitive process that involves the fading of a response to a stimulus [11]. In practice, it can be assumed that, as a species, we are generally able to adapt to changes that occur. We react to a variety of different external stimuli with the senses. As time goes by, when the stimulus no longer causes the same reaction or simply does not change, the reaction fades. At the beginning of 2020, we faced two major events that may in some way illustrate the fading of reactions and the "taming" of society to the changes that took place. The first is the COVID-19 epidemic, and the second is Russia's attack on Ukraine. It was possible to observe how the amount of information, or, unfortunately, disinformation, on a given topic can increase. Eastern European countries experienced several disinformation attacks when the war began, aimed at weakening the mobilization of the population in providing aid to refugees. This was partially successful because of the strong emotional appeal of this content, fueled by the natural fear of war. Border countries were particularly vulnerable to attacks because of its assistance to refugees. They have experienced two major panic attacks in society, fueled by so-called "trolls" on social media. This is an example of how habituation can work to the advantage of the recipient, weakening their reaction to acutely harmful message. Fear is never a good advisor, and it most easily intensifies the user's reaction. At the beginning of the war, information about sudden increases in fuel prices caused panic and queues. The result was a threat of fuel shortages for special forces vehicles. The second situation concerned a shortage of sugar. The panic caused people to buy up supplies, and the stores ran out of sugar for two weeks. Both situations threatened a direct decline in the quality of life, which resulted in a psychological attack on a large part of the population.

In the middle of these events the idea was born about conducting a study on how habituation impacts the way various content spreads through social networks. In this study we make no distinction between the spread of information or contagion; nor do we specify the exact type of information in question, as the entirety of humanity's developement has occured throught the mechanisms of information spread. It is irrelevant if what is being spread is religion, disease, fashion, lifestyle, ideology, or the elementary skills like the ability to write or basic math. It could even be an emotion, like outrage or fear; in fact we've seen this year how fear of war spreads, and despite the importance and gravity of this subject we may already see a decline of interest and strength of response to the ongoing conflict. That's habituation, and we are interested exclusively in how this process affects the spread of our test outbreak.

This section presents the idea of integration habituation with the spreading model. Fig 1 shows how responsiveness level curve is shaped by the effect of habituation. Imagine having a friend, who is trying to convince you to their point of view at almost every opportunity. During first two meetings an attempt to spread this information occurs, with the subject, namely us, showing no positive response. On the contrary, our level of responsiveness decreases, while our irritation rises. During third meeting there is no spread attempt; it might be a simple conversation about the weather, but whatever the subject, it gives us a welcome break from the annoying content and our level of responsiveness to the subject rebounds. During the next two meetings, however, further attempts occur. This results in our responsiveness falling to almost

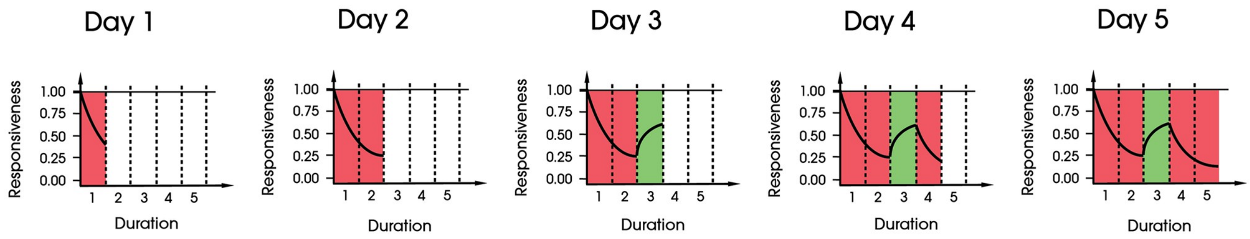


Fig 1. Example illustrating the impact of repeated ineffective/unwanted contacts on node responsiveness.

https://doi.org/10.1371/journal.pone.0280266.g001

zero. The interpretation is that it's not simply that we don't want to listen, it's that at this point we don't want to even meet our troublesome friend.

The SI model generates a very large number of contacts, each one resulting in an attempt to transmit information. This results in a faster coverage process compared to SIR or SIRS models, where nodes can undergo "healing." We found that it will be best suited to test the effectiveness of multiple contacts. This situation can generate a large number of unsuccessful attempts, which cause a decrease in responsiveness with each successive attempt.

In our study, we assume that each node wants to provide the same information. What kind of information it is is not important, because any content under certain conditions can be tiresome. For this reason, in our model, the studied epidemics have fixed parameters, while the edge weights are randomized for each individual contact in the network. As in real life, we have various moods or coincidences from different people under certain conditions, and messages may be more digestible.

The SI model allows infected nodes to contact their neighbors as many times as possible. They are not able to recover from infection like in other, previously mentioned models. For this reason, the situation, as shown on Day 3 in Fig 1, will not have a chance to occur. In the SI model, if an infected node appears in the neighborhood of a node, it will attempt to infect the susceptible node in each successive simulation step until it succeeds. In our assumption, each failed attempt will affect the probability of propagation in each subsequent attempt. Of course, a node can become infected on the first attempt, in which case its level of responsiveness will not change. A song can be "catchy" upon hearing it for the first time, and some may attest to the experience of "love at first sight." Arguably, an interesting assumption would be that the level of responsiveness at which we adopt a given piece of content influences the "fervor" with which we will propagate it further. However, this could already be a different kind of responsiveness, since every stimulus we interact with affects our resistance to a variety of factors.

At the beginning of each simulation, the infection process begins with a group of nodes. In the first step, each infected node gets one chance each to infect susceptible neighbors. Nodes having several active neighbors that will be contacted several times in one turn. A node infected in a given queue is added to the pool of seeders and will be able to infect a neighbor in the next step. Nodes that are not infected after contact have their responsiveness level lowered, which decreases the chance of success in the next attempt. This process will continue until the network is fully covered. In our study, we used a coverage threshold of 80%. We chose this limit because of the significant decrease in process dynamics and the time required for 100% coverage.

Fig 2 shows how the dynamics of network coverage evolves. Graph 2.A shows the average of all runs without the effect of habituation. These processes took 40 simulation steps to reach the threshold, with an average of about 20 attempts per successful infection. Chart 2.B, on the other hand, shows the averaged runs for the same simulations, already showing a decrease in

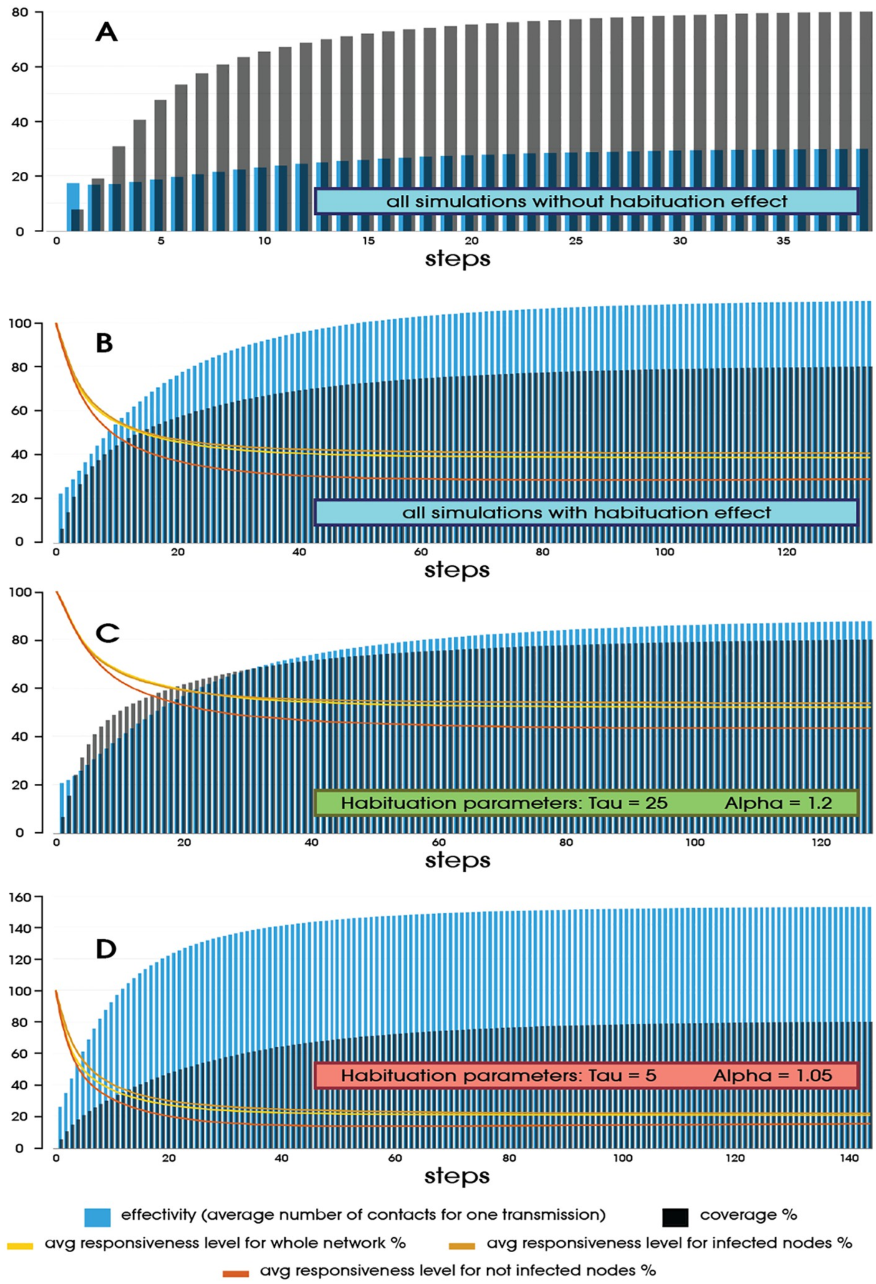


Fig 2. Example illustrating the impact of a drop in responsiveness on process performance.

https://doi.org/10.1371/journal.pone.0280266.g002

responsiveness. It took 100 contacts to infect one node, making the time increase by more than three times. Plot 2.C and 2.D show the results for the most and least favorable habituation parameters, respectively. Overlaid on all graphs having the effect of habituation are curves of decreasing responsiveness for all nodes, infected nodes, and uninfected nodes, which includes those not yet contacted.

In the next section, we describe in more detail the assumptions of the experiment, the parameters used, and the computational method.

Table 1. Main network characteristics for Networks N1–N4, including the number of nodes and edges, the mean degree (DG), network density (ND), global clustering coefficient (CC), mean eigenvector centrality (EV), and modularity (MD).

Networks	Source	Nodes	Edges	DG	ND	CC	EV	MD	Reference
N1	University of California	899	7019	16.62	0.0174	0.07	0.14	0.22	[39]
N2	Political blogs	1224	16715	27.31	0.0223	0.23	0.1	0.43	[40]
N3	Hamsterster friendships	1858	12534	13.49	0.0073	0.09	0.05	0.45	[41]
N4	UoCalifornia messages	1899	13838	14.57	0.0077	0.06	0.08	0.26	[42]

https://doi.org/10.1371/journal.pone.0280266.t001

3. Experiment setup and the mathematical method

Agent-based simulations were performed on four real networks. All were from public repositories. The parameters of the chosen networks are presented in Table 1.

Experiments were simulated within the given network $N(V, E)$ based on the vertex set $V = v_1, v_2, \ldots, v_m$ and edge set $E = e_1, e_2, \ldots, e_n$. Simulations were performed using the proposed Susceptible–Infected model. Each node $u \in V$ has a relationship represented by an edge $(u, v) \in E$. At each step $t+1$ of the simulation, every node $v \in V$ can be infected by his neighbor with a propagation probability $PP(u, v)$ provided that the infecting neighbor was infected in step $t < t + 1$.

In this study, we focused on investigating the impact of habituation on network coverage and infection duration. For the purpose of the experiment, a test space was created, consisting of the following parameters, $R \times N \times PP \times SP \times H \times A \times T$ with a seeding strategy based on single-stage seeding. This provided us with 640 combinations, each of which was performed 10 times, which allowed us to analyze the influence of individual parameters on the course of the infection process. Randomized probabilities were used for each simulation, which was drawn at each attempt to infect a susceptible node on a given edge. Details of the parameters used are given in Table 2.

Each simulation begins with a group of activated nodes $\Phi(s_0)$ in a given graph $G(V, E)$. In each subsequent simulation step s, a set of nodes $\Phi(s-1)$ activated at step $s - 1$ is generated before the contagion process begins. For each node from the set $u \in \Phi(s-1)$, a list of susceptible neighbors is created $\Theta(v, s)$. For each node $v \in \Theta(v, s)$, activation is attempted by node u. Activation occurs when the randomly generated number on the edge between the nodes concerned is lower than the given $PP(u, v)$. Propagation probability is equal for all steps. If the activation attempt is successful, the newly infected node migrates to the set of nodes infected in this step $\Phi(s)$ and will be able to participate in the infection process in the next step $s + 1$ as a spreader.

Due to the integration of the model with the habituation effect, each node is also assigned a responsiveness level $R(v, s)$, which is used to calculate the node-specific propagation

Table 2. Parameters used for diffusion in the simulations.

Symbol	Parameter	Variants	Values
R	Ranking Type	2	Random, Degree
N	Network	4	N1, N2, N3, N4
PP	Propagation probability	2	0.05, 0.1
SP	Seed fraction	2	1%, 5%
H	Habituation	2	Exists, Not exists
A	α	2	1.05, 1.2
T	τ	5	5, 10, 15, 20, 25

https://doi.org/10.1371/journal.pone.0280266.t002

probability at a given simulation stage. Responsiveness decreases due to a failed contagion attempt, which directly affects the probability of propagation according to our model. If $R(v, s)<1.0$, then, for a given PP_s, its new value for a given node is calculated according to the formula $PP_s(u, v, s) = PP(u, v) * R(v, s)$.

The calculation of the responsiveness factor for a given node is performed after each contact. At the beginning of each simulation step, nodes can be in one of two states: 1, active, or 0, inactive. Active nodes may attempt to infect. Each contact in the network can result in one of two possibilities: ineffective activation (unwanted messages) or activation. When a node is activated, the level of responsiveness does not change, and such a node already functions as a spreader.

When an unsuccessful attempt is made, responsiveness is calculated according to Formula (1):

$$y = y_0 - \frac{S}{\alpha}\left(1 - \exp\left(\frac{\alpha \cdot Cnt_{+1}}{\tau}\right)\right) \tag{1}$$

where y_0 represents the initial habituation value. For non-contacted inactive nodes, it is 1.0; for inactive contacted nodes, it is valid for each discreet time point t. S represents stimulus exposition and in this experiment always takes the value of 1 because of the number of actions in the current time step. α is responsible for the recovery rate. τ is a constant influencing habituation process. t is valid for the time that has passed since responsiveness began to drop.

An increase in responsiveness in the SI model can occur when there is no interaction between nodes. As we have mentioned, there are no interruptions in the SI model, so the following method was added for the universality of the algorithm and in terms of future research. Growth can be calculated using Formula (2):

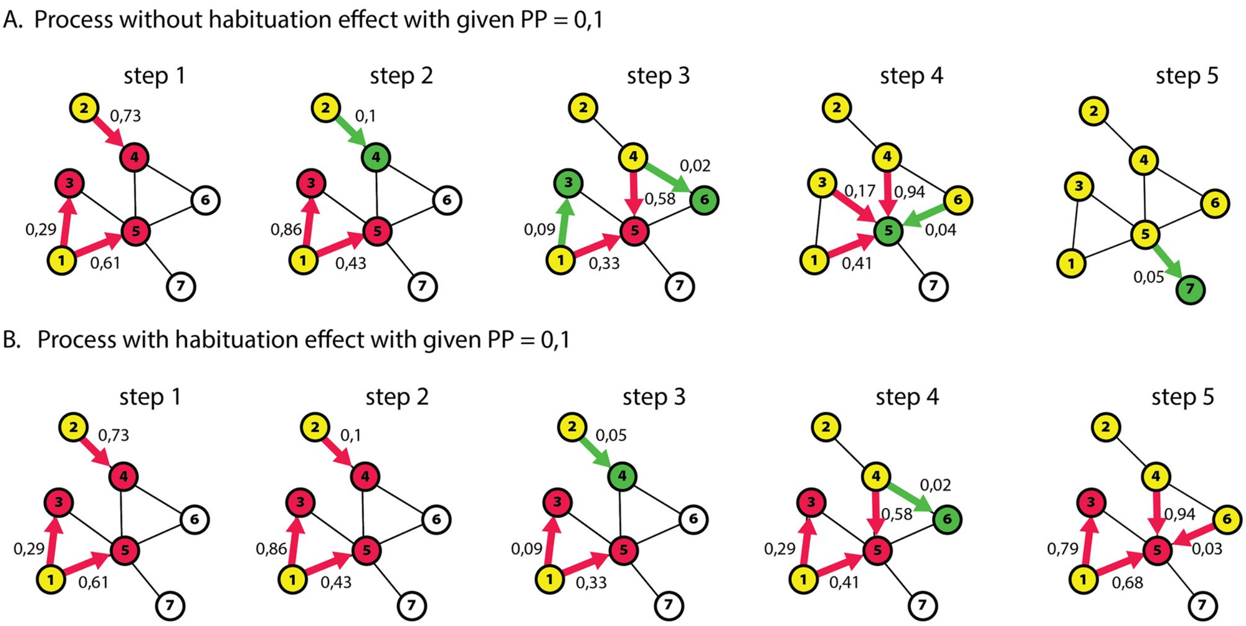
$$y = y_0 - (y_0 - y_1)\exp\left(\frac{-\alpha \cdot Cnt_{-1}}{\tau}\right) \tag{2}$$

where y_0 represents the initial responsiveness value, equal to 1.0, y_1 refers to the responsiveness value reached during the decrease periods, and t, in this case, represents the time passed from the beginning of the recovery process.

In both cases, time t does not represent simulation steps, but the number of contacts. In our study, the reason for the decrease in responsiveness is each failed message delivery attempt, not the time in the sense of a simulation step in which there is a variable number of attacks per node.

4. Illustrative example

In the following section, we will present a simplified process and how it differs from the basic SI model. The network slice shown consists of seven nodes connected by eight edges. Fig 3 is divided into two parts: A, corresponding to the process without the effect of habituation, and B, which assumes this effect. Both show five steps of the simulation. For both processes, the assumed propagation probability threshold is 0.1. In turn, the edge weights are randomized for a given edge at each successive contact. Each contact results in an attempt to pass the content on. In the case of Process 3.A, contacts resulting in failed attempts do not reduce the propagation probability in any way. For 3.B, each node is assigned an initial responsiveness level of 1. This level is reduced based on Formula (1) from the previous section. Under the figures with each step, there are tables with the current responsiveness levels (R LVL) and information about which relationship is affected (IN: infected node; SN: susceptible node). The probability of propagation is always multiplied by the current level of responsiveness of a given node.



IN	SN	R LVL
1	5	1,00 ↘ 0,82
1	3	1,00 ↘ 0,82
2	4	1,00 ↘ 0,82

IN	SN	R LVL
1	5	0,82 ↘ 0,67
1	3	0,82 ↘ 0,67
2	4	0,82 ↘ 0,67

IN	SN	R LVL
1	5	0,67 ↘ 0,55
1	3	0,67 ↘ 0,55
2	4	0,67 → 0,67

IN	SN	R LVL
1	5	0,55 ↘ 0,46
1	3	0,55 ↘ 0,46
4	5	0,46 ↘ 0,38
4	6	1,00 → 1,00

IN	SN	R LVL
1	5	0,38 ↘ 0,27
1	3	0,46 ↘ 0,38
4	5	0,27 ↘ 0,23
6	5	0,23 ↘ 0,19

Fig 3. Toy example showing the steps of the simulation for processes with and without the habituation effect.

https://doi.org/10.1371/journal.pone.0280266.g003

Infection occurs when the drawn edge weight is less than or equal to the assumed propagation probability. Selected parameters of the habituation process are $\tau = 5$ and $\alpha = 1.05$.

In Step 1, both Process A and B fail to infect any new node because the drawn edge weights do not allow it. However, in a process with the habituation effect superimposed, each ineffective contact resulted in a reduction in the levels of responsiveness, which are shown in the table under Fig 3.A.step1. This means that each attack in Step 2 will take place not with an assumed threshold of 0.1 but with a threshold of 0.1 X 0.82 = 0.082.

Step 2 marks the first time that a decrease in responsiveness affects the process. The drawn edge weight is equal to the assumed threshold, so in Fig 3.A.step2, the activation of Node 4 by Node 1 occurs. In Example Fig 3.B.step2, contagion will not occur because the responsiveness level will lower the assumed activation threshold to 0.082 as a result of a previous failed attempt. Since no new nodes have been infected, we know that, in the next round, all nodes will have their responsiveness reduced to 0.67, resulting in a PP of 0.067 for them instead of the given 0.1.

For a process with no habituation effect, two more infections and one ineffective contact occur in Step 3. The process with the superimposed effect succeeds in infecting the first node. Despite the reduced PP, the drawn weight is smaller: 0.06 is smaller than 0.1 reduced to 0.067. In Fig 3.B.step3, on the edge between Node 1 and 3, a situation like in Fig 3.B.step2 occurs between Node 2 and 4. The assumed threshold without reduced responsiveness would allow contagion on this edge as in Fig 3.A.step3 between the same nodes.

In Step 4, one infection occurs in both examples. In Fig 3.B.step4, infection occurs between Node 4 and Node 6. This happens at the first contact; as a result of this, the level of responsiveness of Node 6 does not change. As we mentioned in an earlier chapter, once infected, the level of responsiveness of such a node has no effect on whether it will pass the information on.

In Fig 3.A.step5, the last vulnerable node is infected. At the same stage, the process with the habituation effect covered less than 60% of the network. In addition, Node 5 has its responsiveness level reduced to 0.19. Such a node can only be infected if the minimum activation threshold at the next contact is method.

The graphic above shows how each successive contact generating another opportunity affects the length of the process. This can be understood as the cost of the next attempt. It is possible to try one more time to reach the customer, but this generates additional costs and can cause the opposite reaction, i.e., discouragement or exhaustion with the amount of content delivered.

5. Results

The main objective of this study was to integrate the effect of habituation within the model of information spreading processes and analyze its effect on the performance of spreading processes. The SI model was used as one of the basic models with many extensions and based on the repeated contact between nodes. It was initially proposed for the spread of epidemics and has also been implemented for information diffusion. The differences in the process with and without the habituation process were analyzed. Simulations were performed on the agent-based proposed model with the integration of the habituation process. All processes had the target of achieving a network coverage level of 80%. This decision was made to standardize the results, as when testing the longest intervals, the processing time needed for full coverage increased significantly, as will also be shown in the analysis below. The experiment was based on different parameters such as ranking types, networks, propagation probabilities, seed fraction characteristics for spreading models, and the habituation process parameters τ and α. All results for the individual parameters can be found in the Tables 3–6. In order to highlight the characteristics of the influence of the habituation effect on the spreading process, we decided to divide the graphs into two sections: network coverage and process duration. The network coverage section presents the results for each considered parameter after 5, 10, and 15 simulation steps. Process duration presents the values for half of the assumed threshold, in this case 40%, and after reaching the goal, i.e., infecting 80% of the network.

Impact of the habituation effect on network coverage in information-spreading processes

The average results for all simulation runs with all parameters for processes with and without the habituation effect are presented in Fig 4.A1. It can be clearly seen that, at each simulation run presented, the process with habituation performed relatively poorly. After five steps, the difference in favor of the process without habituation was 19.17%. After 10 steps, it decreased slightly and amounted to 18.76%. After 15 steps, the process without habituation reached the intended 80%, and its advantage over the "chasing" process with the habituation effect was equal to 14.2%. Parameters showed statistical significance. The Wilcoxon test showed p-values less than 0.05 for both τ and α. Fig 4.A2 shows the drops in coverage for all processes with habituation, sorted by the drop in coverage. It can be seen that the longer the simulation lasted, the smaller the decreases were, which was due to the fact that some processes did not even need seven steps to reach the target. The largest difference after five steps was 52.2%. After 10

Table 3. Coverage for spreading processes with and without habituation.

Parameter	Value	HAB						NO HAB		
		5 steps	p-value	10 steps	p-value	15 steps	p-value	5 steps	10 steps	15 steps
		43.74%	<2.2e-16	59.79%	<2.2e-16	66.77%	<2.2e-16	62.91%	78.55%	80.00%
NET	N1	42.74%	<2.2e-16	59.98%	<2.2e-16	67.09%	<2.2e-16	62.79%	79.81%	80.00%
	N2	52.75%	<2.2e-16	67.12%	<2.2e-16	72.71%	<2.2e-16	73.12%	80.00%	80.00%
	N3	37.93%	<2.2e-16	54.60%	<2.2e-16	62.37%	<2.2e-16	55.86%	76.04%	80.00%
	N4	41.53%	<2.2e-16	57.45%	<2.2e-16	64.93%	<2.2e-16	59.84%	77.02%	80.00%
ALPHA	1.05	42.21%	<2.2e-16	57.41%	<2.2e-16	64.56%	<2.2e-16	62.91%	78.55%	80.00%
	1.2	45.26%	<2.2e-16	62.16%	<2.2e-16	68.99%	<2.2e-16	62.91%	78.55%	80.00%
PP	0.05	31.59%	<2.2e-16	48.25%	<2.2e-16	57.47%	<2.2e-16	53.91%	75.67%	80.00%
	0.1	55.89%	<2.2e-16	71.32%	<2.2e-16	76.08%	<2.2e-16	71.90%	80.00%	80.00%
SP	0.01	41.41%	<2.2e-16	59.59%	<2.2e-16	67.06%	<2.2e-16	60.25%	78.70%	80.00%
	0.05	46.06%	<2.2e-16	59.98%	<2.2e-16	66.49%	<2.2e-16	65.56%	78.39%	80.00%
TAU	5	31.36%	<2.2e-16	45.60%	<2.2e-16	54.35%	<2.2e-16	62.91%	78.55%	80.00%
	10	40.64%	<2.2e-16	56.20%	<2.2e-16	63.92%	<2.2e-16	62.91%	78.55%	80.00%
	15	59.92%	<2.2e-16	62.31%	<2.2e-16	69.19%	<2.2e-16	62.91%	78.55%	80.00%
	20	49.29%	<2.2e-16	66.22%	<2.2e-16	72.31%	<2.2e-16	62.91%	78.55%	80.00%
	25	51.47%	<2.2e-16	68.61%	<2.2e-16	74.10%	<2.2e-16	62.91%	78.55%	80.00%

https://doi.org/10.1371/journal.pone.0280266.t003

steps, we observed a value of 45.51%; after 15 steps, it was 39.4%. The smallest differences also decreased over time, with 15.05%, 2.65%, and 0.28% after 5, 10, and 15 steps, respectively. Fig 4.A3 to 4.A5 shows all the runs sorted by the coverage of the processes without habituation. Each graph represents the situation after 5, 10, and 15 steps.

Considering the individual networks used for the simulations (Fig 4.B1), it can be seen that, when simulations without the effect of habituation performed better on a given network correspondingly, processes with applied habituation also showed improved performance compared

Table 4. Duration for spreading processes with and without habituation.

Parameter	Value	HAB				NO HAB	
		40% coverage	p-value	80% coverage	p-value	40% coverage	80% coverage
		6.15	<2.2e-16	39.83	<2.2e-16	3.30	9.46
NET	N1	6.33	<2.2e-16	33.14	<2.2e-16	3.51	8.5
	N2	4.27	<2.2e-16	26.48	<2.2e-16	2.58	7.03
	N3	7.52	<2.2e-16	53.33	<2.2e-16	3.71	11.60
	N4	6.48	<2.2e-16	46.37	<2.2e-16	3.40	10.70
ALPHA	1.05	6.53	<2.2e-16	45.40	<2.2e-16	3.30	9.46
	1.2	5.77	<2.2e-16	34.26	<2.2e-16	3.30	9.46
PP	0.05	8.51	<2.2e-16	58.14	<2.2e-16	3.84	11.91
	0.1	3.79	<2.2e-16	21.52	<2.2e-16	2.76	7.01
SP	0.01	6.59	<2.2e-16	35.13	<2.2e-16	3.76	9.16
	0.05	5.71	<2.2e-16	44.53	<2.2e-16	2.84	9.76
TAU	5	9.89	<2.2e-16	67.42	<2.2e-16	3.30	9.46
	10	6.55	<2.2e-16	46.84	<2.2e-16	3.30	9.46
	15	5.27	<2.2e-16	34.84	<2.2e-16	3.30	9.46
	20	4.66	<2.2e-16	27.23	<2.2e-16	3.30	9.46
	25	4.38	<2.2e-16	22.82	<2.2e-16	3.30	9.46

https://doi.org/10.1371/journal.pone.0280266.t004

Table 5. Coverage decrease for spreading processes with the habituation effect.

Parameter	Value	5 steps	p-value	10 steps	p-value	15 steps	p-value
NET	N1	31.93%	<2.2e-16	24.85%	<2.2e-16	16.14%	<2.2e-16
	N2	27.86%	<2.2e-16	16.10%	<2.2e-16	9.11%	<2.2e-16
	N3	32.10%	<2.2e-16	28.20%	<2.2e-16	22.04%	<2.2e-16
	N4	30.60%	<2.2e-16	25.41%	<2.2e-16	18.84%	<2.2e-16
ALPHA	1.05	32.90%	<2.2e-16	26.91%	<2.2e-16	19.30%	<2.2e-16
	1.2	28.06%	<2.2e-16	20.87%	<2.2e-16	13.76%	<2.2e-16
PP	0.05	41.40%	<2.2e-16	36.24%	<2.2e-16	28.16%	<2.2e-16
	0.1	22.27%	<2.2e-16	10.85%	<2.2e-16	4.90%	<2.2e-16
SP	0.01	31.27%	<2.2e-16	24.28%	<2.2e-16	16.17%	<2.2e-16
	0.05	29.74%	<2.2e-16	23.49%	<2.2e-16	16.89%	<2.2e-16
TAU	5	50.15%	<2.2e-16	41.95%	<2.2e-16	32.06%	<2.2e-16
	10	35.40%	<2.2e-16	28.45%	<2.2e-16	20.10%	<2.2e-16
	15	27.01%	<2.2e-16	20.67%	<2.2e-16	13.51%	<2.2e-16
	20	21.65%	<2.2e-16	15.70%	<2.2e-16	9.61%	<2.2e-16
	25	18.18%	<2.2e-16	12.65%	<2.2e-16	7.38%	<2.2e-16

https://doi.org/10.1371/journal.pone.0280266.t005

with the other runs for the other networks. *Net*2 obtained the best results for both types of processes. It has the highest mean degree (DG = 27.31), clustering coefficient (CC = 0.23) and a network density of two to three times that of the other networks (ND = 0.0223—see Table 1). The worst effects were observed for *Net*3. Comparing these extreme cases, we obtained 52.27% for *Net*2 and 37.93% for *Net*3 after 5 steps, and these values were 67.12% and 54.60% after 10 steps and were 72,71% and 62,37% after 15 step, respectively. Fig 4.B2 shows the percentage drops in coverage at the selected steps. For Network 2 and 3, the difference in coverage in favor of Network 2 was 4.24% after 5 steps, 12.1% after 10 steps, and 12.93% after 15 steps. Thus, looking at the specifics of the network, it can be concluded that the network density became more important over time in terms of reducing the effect of habituation. When the average response rate was already strongly reduced for the whole network, the number of

Table 6. Duration increase for spreading processes with the habituation effect.

Parameter	Value	40% coverage	p-value	80% coverage	p-value
NET	N1	80.34%	<2.2e-16	289.88%	<2.2e-16
	N2	65.50%	<2.2e-16	276.67%	<2.2e-16
	N3	102.70%	<2.2e-16	359.74%	<2.2e-16
	N4	90.59%	<2.2e-16	333.36%	<2.2e-16
ALPHA	1.05	97.88%	<2.2e-16	379.92%	<2.2e-16
	1.2	74.85%	<2.2e-16	262.16%	<2.2e-16
PP	0.05	121.61%	<2.2e-16	388.16%	<2.2e-16
	0.1	37.32%	<2.2e-16	206.99%	<2.2e-16
SP	0.01	75.27%	<2.2e-16	283.52%	<2.2e-16
	0.05	101.06%	<2.2e-16	356.25%	<2.2e-16
TAU	5	199.70%	<2.2e-16	612.68%	<2.2e-16
	10	98.48%	<2.2e-16	395.14%	<2.2e-16
	15	59.70%	<2.2e-16	268.29%	<2.2e-16
	20	41.21%	<2.2e-16	197.84%	<2.2e-16
	25	32.73%	<2.2e-16	141.23%	<2.2e-16

https://doi.org/10.1371/journal.pone.0280266.t006

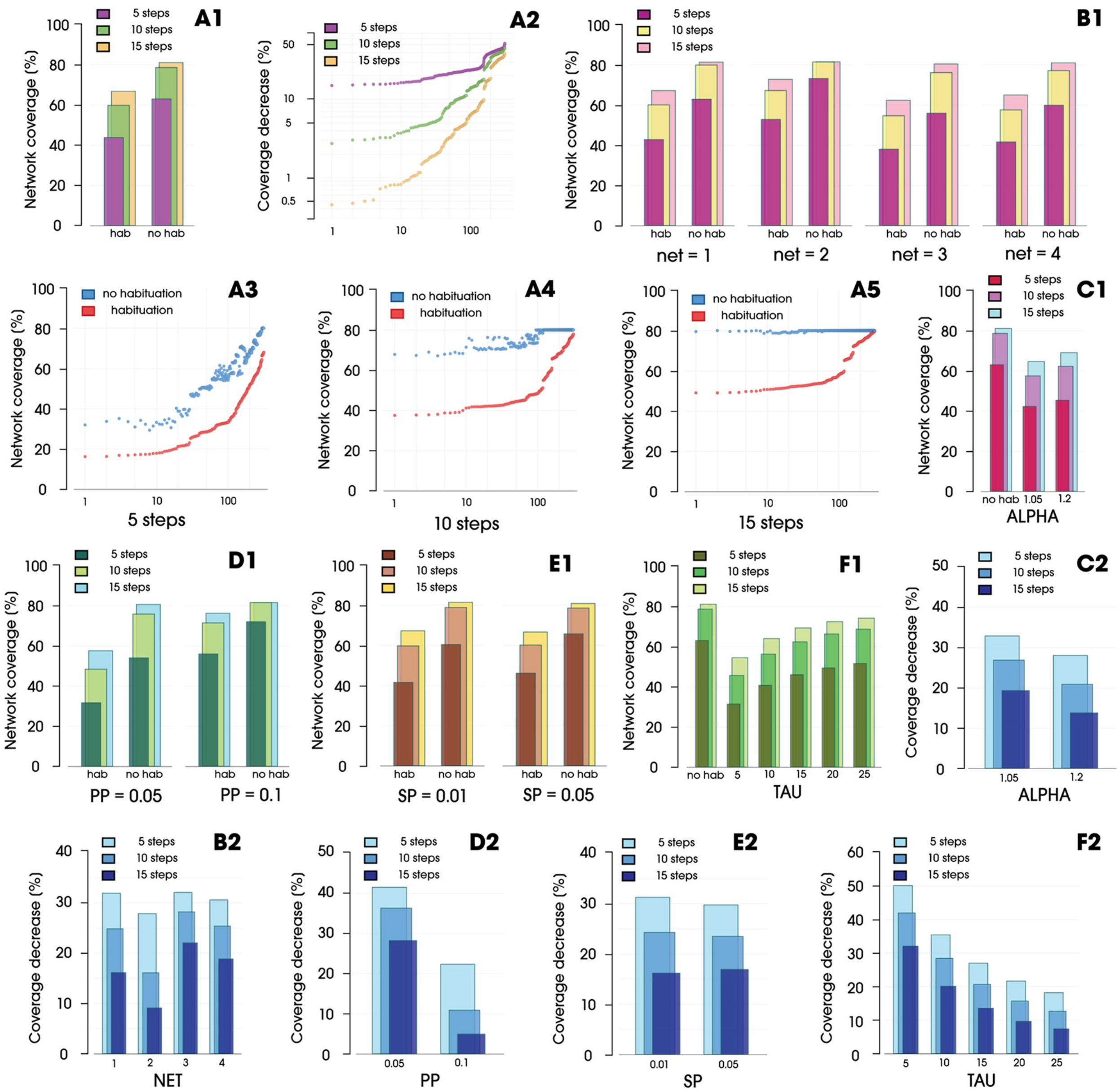


Fig 4. (**A1**) Coverage for spreading processes with and without habituation. (**A2**) Coverage decrease in processes with the habituation effect, compared to processes without habituation, sorted by the decrease in coverage. (**A3**--**A5**) Distances between coverage in simulations with and without the habituation effect, with results sorted by coverage without habituation and assigned corresponding results from processes with the habituation effect. (**B1**) Coverage for each network for spreading processes with and without habituation. (**B2**) Decrease in coverage for each network in relation to a process without habituation. (**C1**) Coverage for each alpha value for spreading processes with habituation compared to a process without habituation. (**C2**) Decrease in coverage for each alpha in relation to a process without habituation. (**D1**) Coverage for propagation probabilities for spreading processes with and without habituation. (**D2**) Decrease in coverage for each PP value in relation to a process without habituation. (**E1**) Coverage for each seeding percentage for spreading processes with and without habituation. (**E2**) Decrease in coverage for each SP in relation to a process without habituation. (**F1**) Coverage for each tau value for spreading processes with habituation in comparison with processes without habituation. (**F2**) Decrease in coverage for each tau in relation to a process without habituation.

https://doi.org/10.1371/journal.pone.0280266.g004

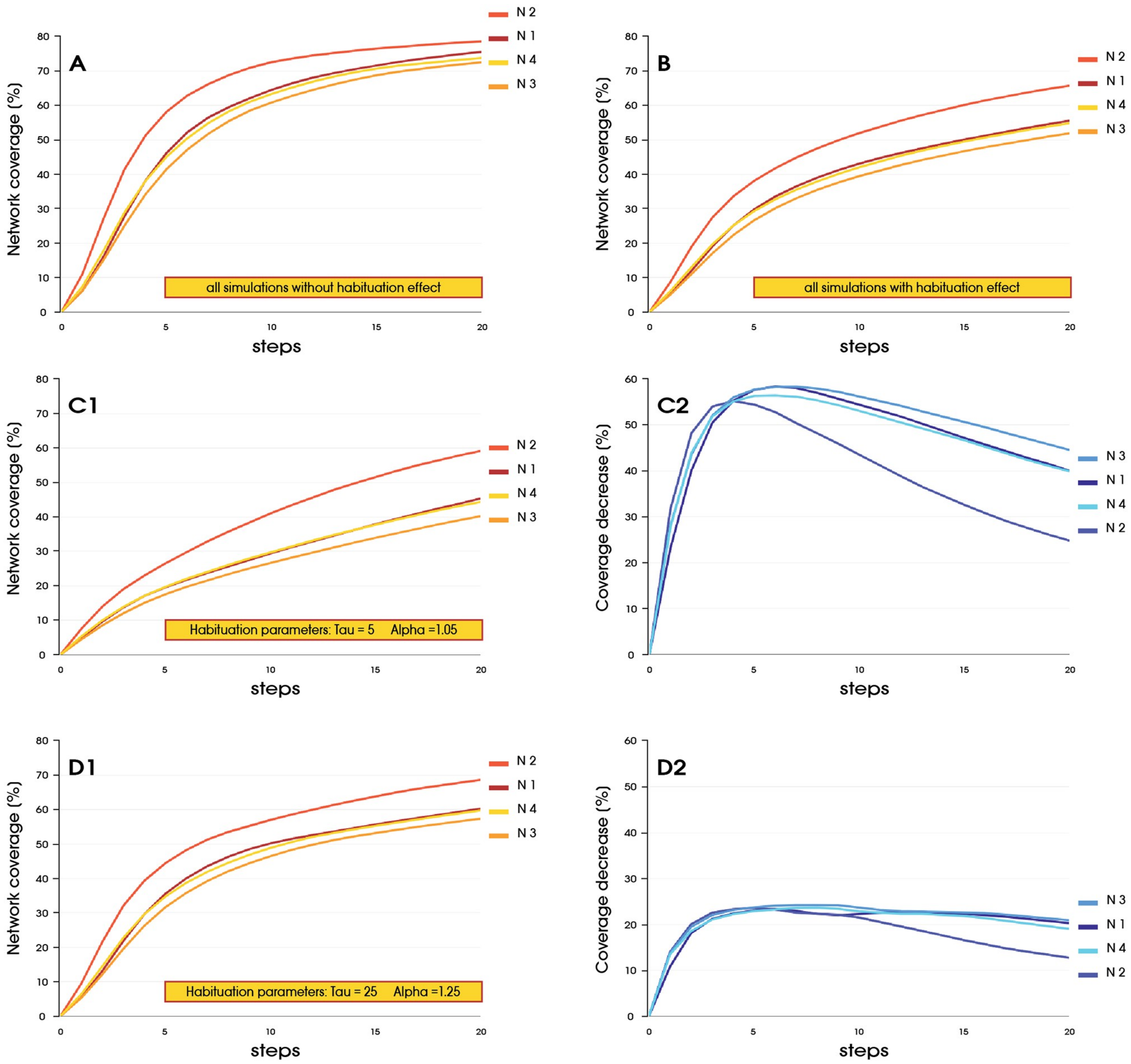


Fig 5. (**A**) Coverage for each network for all spreading processes without habituation. (**B**) Coverage for each network for all spreading processes with habituation. (**C1**) Coverage for each network for spreading processes with the "worst" set of habituation parameters. (**C2**) Decrease in coverage for each network in relation to a process without habituation with the "worst" habituation parameters. (**D1**) Coverage for each network for spreading processes with the "best" set of habituation parameters.(**D2**) Decrease in coverage for each network in relation to a process without habituation with the "best" habituation parameters.

https://doi.org/10.1371/journal.pone.0280266.g005

connections became more important. The model used, makes the number of possible. contacts the most important element affecting the speed of coverage. The impact of network heterogeneity is further shown in the Fig 5. Parameters such as modularity, eigenvector do not show significant interference with the processes within set of analysed networks. Fig 5.A and 5.B shows all runs for selected networks with and without the imposed habituation effect. Fig 5.C1 and 5.D1 presents the "worst" and the "best" set of habituation parameters, respectively. Fig 5.

C2 and 5.D2, on the other hand, show declines in the performance of these processes relative to simulations without the habituation effect applied.

Fig 4.C1 illustrates the comparison between simulations without habituation and simulations in terms of the α parameter. This parameter affects the way the response curve flattens out. The higher the value, the lower the decreases are and the higher the increases are during the recovery of a node. This results in minimally improved network coverage results for $\alpha = 1.2$. Fig 4.C2 shows the decreases in network coverage with respect to the process without habituation with respect to the analyzed α parameters. The difference in subsequent simulation steps were 4.84%, 6.04%, and 5.54% in favor of the parameter $\alpha = 1.2$.

In Fig 4.D1, we compare the individual propagation probabilities with and without the implemented habituation effect. In this case, there was a clear difference in favor of $PP = 0.1$. A higher propagation probability resulted in faster network coverage. Fig 4.D2, on the other hand, shows that, comparing the processes with the habituation effect, a higher *PP* also resulted in smaller drops compared to the processes without the effect of habituation; thus, after 5, 10, and 15 simulation steps, the difference in favor of $PP = 0.1$ was, respectively, 19.13%, 25.39%, and 23.26%. At the same time, in the last step in both simulations without habituation, the processes had already reached the assumed 80% coverage, so they did not increase their "advantage".

The impact of the number of initial nodes initiating the process is shown in Fig 4.E1. It is interesting to note that processes with a higher number of seeds perform worse in terms of network coverage over time. Although the process with $SP = 0.05$ covered 4.65% more of the network after five simulation steps, it was only 0.39% more after 10 steps. After 15 steps, the process with the smaller *SP* (equal to 0.01) achieved a network coverage of 67.06%, whereas the process with the larger SP covered 66.49%, losing 0.57 percentage points. It can be concluded from this that the initial higher number of seeds resulted in an increased number of contacts, as well as the unsuccessful ones, which results in a faster decrease in the level of responsiveness over the whole network, ultimately causing a decrease in the spreading speed. Fig 4.E2, in turn, shows that processes with more grains over time also lose their advantage in terms of a decrease in coverage compared to processes without habituation with the same *SP*. After five steps, the process with a larger *SP* showed a smaller decrease of 29.74%, whereas the processes with a smaller *SP* exhibited a decrease of 31.27%. After 10 steps, they reached 23.49% and 24.28%, respectively. After 15 steps, the situation reversed and the advantage was reached by processes with $SP = 0.01$, losing 16.7%, while processes with $SP = 0.05$ lost 16.89% to processes without the habituation effect.

The second habituation parameter, τ, is shown in Fig 4.F1. It is a time constant, and the smaller it is, the more rapidly habituation occurs. This is confirmed in the graphs, as $\tau = 5$ achieved the worst result at each stage of the simulation (31.36%, 45.60%, and 54.35%), whereas the largest considered $\tau = 25$ covered the network the best (51.47%, 68.61%, and 74.10%). When we look at the declines in coverage shown in Fig 4.F2, we can also notice that, as the τ parameter increases, the decline in coverage decreases faster over time relative to processes without habituation with lower values of this parameter.

Impact of the habituation effect on duration in information-spreading processes

Fig 6.A1 and 6.A2 show all simulation runs divided into processes with and without the implemented habituation effect, sorted in ascending order by length. In both examples, the tendency for the efficiency of the process to decrease (based on its duration) is even more pronounced. A comparison of the averages of all runs can be found in Fig 6.A3. To achieve half of the

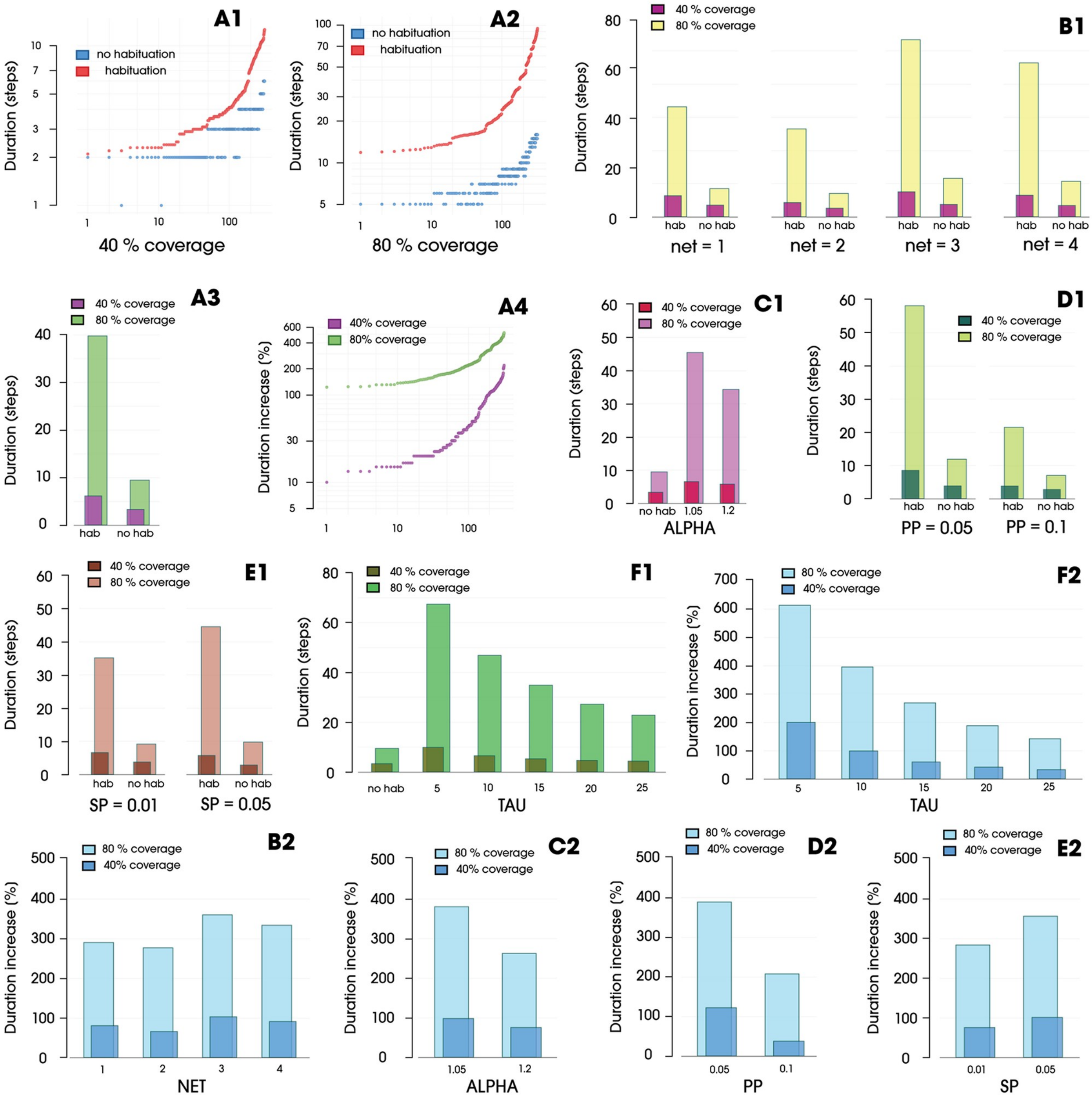


Fig 6. (**A1**-**A2**) Distances between the durations of simulations with and without the habituation effect, with results sorted by duration without habituation and assigned corresponding results from processes with habituation. (**A3**) Duration for spreading processes with and without habituation. (**A4**) Duration increases in processes with the habituation effect, compared to processes without habituation, sorted by coverage increase. (**B1**) Durations for each network for spreading processes with and without habituation. (**B2**) Increases in duration for each network in relation to a process without habituation. (**C1**) Durations for each alpha value for spreading processes with habituation compared to a process without habituation. (**C2**) Increases in duration for each alpha value in relation to a process without habituation. (**D1**) Durations for propagation probabilities for spreading processes with and without habituation. (**D2**) Increases in duration for each PP in relation to a process without habituation. (**E1**) Durations for each seeding percentage for spreading processes with and without habituation. (**E2**) Increase in duration for each SP in relation to a process without habituation. (**F1**) Durations for each tau value for spreading processes with habituation in comparison with processes without habituation. (**F2**) Increases in duration for each tau value in relation to a process without habituation.

https://doi.org/10.1371/journal.pone.0280266.g006

assumed coverage (40%), processes without habituation required 3.30 simulation steps on average, whereas processes with habituation required almost half as much time (86% more), namely 6.15 steps. Accordingly, for the 80% threshold, the results were as follows: 9.46 steps to 39.83. To achieve the same network coverage, the process with habituation required 321% more time. This was confirmed by the U Mann–Whitney test, which showed the significance ($p < 0{,}05$) of the impact of the habituation effect on the duration of the process. For the assumed 40% coverage, the habituation parameters show statistical significance at the level of 0.0466, whereas for the level of 80%, the parameters also showed statistical significance, which was higher, at 0.0098. This confirms the increasing influence of the habituation process on the duration of coverage of the assumed thresholds. Fig 6.A4 compares the magnitude of the duration increases for the earlier graphs (A1, A2). In the case of a 40% network coverage, the increases are in the range of 100%–350%; given an 80% coverage, we obtain a range of duration increases of 250%–650%.

If we look at the division in terms of adopted networks (Fig 6.B1), we can see a division into two groups: *Net*1 with *Net*2, and *Net*3 with *Net*4. The first group has a noticeably higher network density (ND), and degree (DG) values. The second group, on the other hand, is characterized by a significantly higher number of nodes. In the SI model in which the number of contacts is crucial, with low density and the habituation effect implemented, we observe a noticeable decrease in the speed of the processes taking place. The time increases presented in Fig 6.B2 show that, if the processes performed relatively well, they also performed well in relation to "their" collected processes without the habituation effect.

For the comparison of process duration (Fig 6.C1) in terms of parameter α, again the higher parameter ($\alpha = 1.2$) performed better for both a 40% coverage (5.77–6.53 steps) and an 80% coverage, with the difference increased even more (34.26–45.40 steps). Both parameters obtained significantly worse results than the processes without habituation, as can be clearly seen in Fig 6.C2. For a 40% network coverage, the duration increased for $\alpha = 1.2$ by 74.85% and for $\alpha = 1.05$ by 97.88%, whereas for an 80% coverage, the duration increased for these parameters by 262.12% and 379.92%, respectively.

Propagation probability (Fig 6.D1) was shown to be a much stronger factor influencing the duration of processes with habituation compared to without habituation. For an 80% coverage, processes without habituation took 69% longer with $PP = 0.05$ than with $PP = 0.1$ (11.91–7.01 steps). For the same coverage value, processes with habituation for a lower PP lasted 170% longer (58.14–21.52 steps). This is further evidence that, as time passes, the habituation process has an increasing effect on decreasing spreading efficiency, directly affecting the overall process's duration.

The fewer nodes infected in a given step, the longer the entire process takes (Fig 6.D2). To achieve the assumed network contagion threshold, processes with habituation when $PP = 0.05$ take 388.16% longer than processes without the habituation effect; when $PP = 0.1$, the increase is 206.99%. A trend can be noted that a doubled PP causes a two-fold decrease; i.e., it is inversely proportional. Reaching a 40% network coverage where $PP = 0.1$ for processes with habituation takes only 37.32% longer than processes without habituation; when $PP = 0.05$, this percentage is 121.61%. Reaching a 40% network coverage where $PP = 0.1$ for processes with habituation takes only 37.32% longer than for processes without habituation; when $PP = 0.05$, this percentage is, again, 121.61% for processes with habituation.

The effect of the number of nodes that initiate the entire process is shown in Fig 6.E1. For coverage up to the 40% level, habituation-aware simulations with $SP = 0.05$ reach the assumed threshold faster than processes with a lower initial seed count, $SP = 0.01$. On the other hand, reaching the 80% threshold for the same processes takes longer with a higher $SP = 0.05$. This could imply that the initial high number of infected nodes results in a higher number of ineffective attacks, which, after the early rapid coverage of the network, results in a concomitant, faster decrease in the average responsiveness of the entire network, which slows down the process, making it more difficult for subsequent infections to occur at the same rate. Fig 6.E2 shows the increase in the duration of processes with the habituation effect relative to processes without it. Processes with the higher initial number of infected nodes lasted longer when they reached both a 40% and 80% network coverage. As in the previous graph, we can observe that the processes with habituation and a higher *SP* perform worse at the initial stage due to a decrease in responsiveness compared with processes without habituation and thus an *SP* value. Again, the decrease in average network responsiveness at earlier stages of the process slows down the spreading from the very beginning.

The effects of individual values of τ on the increase in the duration of the network coverage process are shown in Fig 6.F1. As assumed, increasing the value of τ promotes process acceleration. Processes with a τ value of 5 take about three times as long to reach the assumed thresholds than processes with a τ value of 25. Fig 6.F2 shows the increases in the duration of processes with habituation relative to processes without it. Again, it can be observed that the more rapidly the habituation process is, the more quickly the responsiveness of the network decreases, which is a simple factor that leads to fewer infected nodes in subsequent steps. This means more failed infection attempts, which in turn again causes the responsiveness to drop, and so on. For a τ value of 5 for a 40% network coverage, the process duration increases by 200%; for a threshold of 80%, the time increases by over 600%. As an extreme comparison, a τ of 25 reaches the 80% threshold with an increase of less than 200%.

6. Conclusions

Most earlier works related to repeated contacts within social networks and information diffusion processes have assumed an increased potential response with each repetition. From that perspective, each contact may increase a node's interest in a discussed product or idea. As a result, the transmission rate increases.

From another perspective, each incoming message can be treated as a stimulus, and according to research related to habituation, the response to repeated stimuli decreases. However, a response also has the potential to recover if stimuli are not encountered over a certain timeframe. In this study, we evaluated the impact of habituation on spreading processes under an SI model with assumed repeated contact during the process.

Our experimental study revealed the existence of a relationship between process dynamics and selected habituation parameters. Many significant decreases in performance and the coverage of simulated processes were observed. We believe that failing to account for the habituation effect can result in significant performance degradations in real systems, despite the selection of appropriate seeds and individual process parameters to maximize impact. This maximization may turn out to be "lethal" for the intended effects of the campaign, as content overload will have the opposite effect.

One of the ways to avoid such situations may be the selection of an appropriate "dosage" of content so as not to lead to a decrease in responsiveness, which this study tries to prove by looking for causality in the habituation effect. The obtained results have several implications for practice and real campaigns. Spreading processes in social networks allows us to analyze the entire spectrum of events that affect the functioning of the world, including marketing, social and political campaigns, viral rumor marketing, and the entire flow of information, ending with the widely discussed current topic of the ongoing COVID-19 pandemic. The means of communication can generate different stimuli to change the user's behavior or opinion. Maximizing contact or influence and seed choice were considered when analyzing such an impact. Although increasing the number of contacts may be effective for many scenarios, repeated unwanted messages, instead of the intended maximization, may have the opposite effect.

An interesting example may be the ongoing large-scale vaccination campaigns all around the world. After very rapid initial growth, vaccination rates quickly began to decrease, many countries have not yet reached their intended thresholds. At the same time, the number of people who were skeptical about vaccination increased. Information that annoys people the first time around, when repeated persistently, will increase the level of aversion in the recipient and increase the likelihood that they will discourage additional people in their environment. At the same time, a "latent" disinformation campaign is underway in a form similar to whisper marketing, in which information is infrequent but highly emotionally tinged.

Several directions of future work can be taken. First of all, the need for methods allowing us to decrease the negative impact of habituation of information diffusion performance is strongly suggested. Appropriate timing and campaign intensities can be adjusted using computational models. Another exciting topic is related to competing processes and the impact of habituation on their interactions.

Acknowledgments

This work was supported by the National Science Centre of Poland, Decision No. 2017/27/B/HS4/01216

Author Contributions

Conceptualization: Piotr Bartków, Jarosław Jankowski.

Data curation: Piotr Bartków, Kamil Bortko.

Formal analysis: Piotr Bartków, Kamil Bortko, Jarosław Jankowski, Patryk Pazura.

Funding acquisition: Jarosław Jankowski.

Investigation: Piotr Bartków, Patryk Pazura.

Methodology: Kamil Bortko, Patryk Pazura.

Project administration: Jarosław Jankowski.

Resources: Piotr Bartków.

Software: Piotr Bartków.

Supervision: Piotr Bartków, Jarosław Jankowski.

Visualization: Piotr Bartków.

Writing – original draft: Piotr Bartków, Jarosław Jankowski.

Writing – review & editing: Piotr Bartków.

References

1. Jung AR, Heo J. Does cluttered social media environment hurt advertising effectiveness? The moderation of ad types and personalization. Journal of Research in Interactive Marketing. 2021;. https://doi.org/10.1108/JRIM-11-2020-0238
2. Nielsen JH, Huber J. The effect of brand awareness on intrusive advertising. In: Society for Consumer Psychology Conference, San Diego, CA, February 2009; 2009.

3. Ferreira C, Michaelidou N, Moraes C, McGrath M. Social media advertising: Factors influencing consumer ad avoidance. Journal of Customer Behaviour. 2017; 16(2):183–201. https://doi.org/10.1362/147539217X14909733609398
4. Lane DM. Limited Capacity, Attention. Information processing and decision making. 1982; 2:121.
5. Benway JP. Banner blindness: The irony of attention grabbing on the World Wide Web. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting. vol. 42. SAGE Publications Sage CA: Los Angeles, CA; 1998. p. 463–467.
6. Yang Q, Zhou Y, Jiang Y, Huo J. How to overcome online banner blindness? A study on the effects of creativity. Journal of Research in Interactive Marketing. 2021;. https://doi.org/10.1108/JRIM-12-2019-0212
7. Çiçek M,Eren-Erdoğmuş İ,Daştan İ. How to increase the awareness of in-app mobile banner ads: exploring the roles of banner location, application type and orientation. International Journal of Mobile Communications. 2018; 16(2):153–166. https://doi.org/10.1504/IJMC.2018.089757
8. Thompson RF. Habituation: a history. Neurobiology of learning and memory. 2009; 92(2):127. https://doi.org/10.1016/j.nlm.2008.07.011 PMID: 18703156
9. Thompson RF, Spencer WA. Habituation: a model phenomenon for the study of neuronal substrates of behavior. Psychological review. 1966; 73(1):16. https://doi.org/10.1037/h0022681 PMID: 5324565
10. Rankin CH, Abrams T, Barry RJ, Bhatnagar S, Clayton DF, Colombo J, et al. Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. Neurobiology of learning and memory. 2009; 92(2):135–138. https://doi.org/10.1016/j.nlm.2008.09.012 PMID: 18854219
11. Blumstein DT. Habituation and sensitization: new thoughts about old ideas. Animal Behaviour. 2016; 120:255–262. https://doi.org/10.1016/j.anbehav.2016.05.012
12. Horn G. Neuronal mechanisms of habituation. Nature. 1967; 215(5102):707–711. https://doi.org/10.1038/215707a0 PMID: 6059539
13. Stanley JC. Computer simulation of a model of habituation. Nature. 1976; 261(5556):146–148. https://doi.org/10.1038/261146a0 PMID: 179015
14. Innis N, Staddon J. What should comparative psychology compare? International Journal of Comparative Psychology. 1989; 2(3). https://doi.org/10.46867/C4F01X
15. Wang D, Arbib MA. Modeling the dishabituation hierarchy: The role of the primordial hippocampus. Biological cybernetics. 1992; 67(6):535–544. https://doi.org/10.1007/BF00198760 PMID: 1472577
16. Marsland S, Nehmzow U, Shapiro J. A model of habituation applied to mobile robots; 1999.
17. Balkenius C. Attention, habituation and conditioning: Toward a computational model. Cognitive Science Quarterly. 2000; 1(2):171–214.
18. Marsland S, Nehmzow U, Shapiro J. Novelty detection for robot neotaxis; 2000.
19. Marsland S. Using habituation in machine learning. Neurobiology of learning and memory. 2009; 92(2):260–266. https://doi.org/10.1016/j.nlm.2008.05.014 PMID: 18602487
20. Mintz Y, Aswani A, Kaminsky P, Flowers E, Fukuoka Y. Nonstationary bandits with habituation and recovery dynamics. Operations Research. 2020; 68(5):1493–1516. https://doi.org/10.1287/opre.2019.1918
21. Ueda K, Sekoguchi T, Yanagisawa H. How predictability affects habituation to novelty. PloS one. 2021; 16(6):e0237278. https://doi.org/10.1371/journal.pone.0237278 PMID: 34061853
22. Chen K. Simulating infant visual learning by comparison: an initial model. In: Proceedings of the 42nd Annual Conference of the Cognitive Science Society.; 2020.
23. Sekoguchi T, Sakai Y, Yanagisawa H. Mathematical model of emotional habituation to novelty: Modeling with Bayesian update and information theory. In: 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC). IEEE; 2019. p. 1115–1120.
24. Zhou C, Zhao Q, Lu W. Impact of repeated exposures on information spreading in social networks. PloS one. 2015; 10(10):e0140556. https://doi.org/10.1371/journal.pone.0140556 PMID: 26465749
25. Zhou C, Zhang P, Zang W, Guo L. Maximizing the cumulative influence through a social network when repeat activation exists. Procedia Computer Science. 2014; 29:422–431. https://doi.org/10.1016/j.procs.2014.05.038
26. Shan X, Chen W, Li Q, Sun X, Zhang J. Cumulative activation in social networks. Science China Information Sciences. 2019; 62(5):1–21. https://doi.org/10.1007/s11432-018-9609-7
27. Piedrahita P, Borge-Holthoefer J, Moreno Y, González-Bailón S. The contagion effects of repeated activation in social networks. Social networks. 2018; 54:326–335. https://doi.org/10.1016/j.socnet.2017.11.001

28. Read JM, Eames KT, Edmunds WJ. Dynamic social networks and the implications for the spread of infectious disease. Journal of The Royal Society Interface. 2008; 5(26):1001–1007. https://doi.org/10.1098/rsif.2008.0013 PMID: 18319209

29. Smieszek T, Fiebig L, Scholz RW. Models of epidemics: when contact repetition and clustering should be included. Theoretical biology and medical modelling. 2009; 6(1):1–15. https://doi.org/10.1186/1742-4682-6-11

30. Diekmann O, De Jong M, Metz JAJ. A deterministic epidemic model taking account of repeated contacts between the same individuals. Journal of Applied Probability. 1998; 35(2):448–462. https://doi.org/10.1239/jap/1032192860

31. Dietz K, Hadeler K. Epidemiological models for sexually transmitted diseases. Journal of mathematical biology. 1988; 26(1):1–25. https://doi.org/10.1007/BF00280169 PMID: 3351391

32. Leskovec J, Adamic LA, Huberman BA. The dynamics of viral marketing. ACM Transactions on the Web (TWEB). 2007; 1(1):5. https://doi.org/10.1145/1232722.1232727

33. Michalski R, Jankowski J, Kazienko P. Negative effects of incentivised viral campaigns for activity in social networks. In: 2012 Second International Conference on Cloud and Green Computing. IEEE; 2012. p. 391–398.

34. Bonell C, Michie S, Reicher S, West R, Bear L, Yardley L, et al. Harnessing behavioural science in public health campaigns to maintain 'social distancing'in response to the COVID-19 pandemic: key principles. J Epidemiol Community Health. 2020; 74(8):617–619. https://doi.org/10.1136/jech-2020-214290 PMID: 32385125

35. Jankowski J. Habituation effect in social networks as a potential factor silently crushing influence maximisation efforts. Scientific reports. 2021; 11(1):1–12. https://doi.org/10.1038/s41598-021-98493-9 PMID: 34561501

36. Kempe D, Kleinberg J, Tardos É. Maximizing the spread of influence through a social network. In: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining; 2003. p. 137–146.

37. Li M, Wang X, Gao K, Zhang S. A survey on information diffusion in online social networks: Models and methods. Information. 2017; 8(4):118. https://doi.org/10.3390/info8040118

38. Tsadikovich D, Kamble A, Elalouf A. Controlled information spread for population preparedness in disaster operations management. International journal of disaster risk reduction. 2020; 42:101338. https://doi.org/10.1016/j.ijdrr.2019.101338

39. Opsahl T. Triadic closure in two-mode networks: Redefining the global and local clustering coefficients. Social Networks. 2013; 35(2):159–167. https://doi.org/10.1016/j.socnet.2011.07.001

40. Adamic LA, Glance N. The political blogosphere and the 2004 US election: divided they blog. In: Proceedings of the 3rd international workshop on Link discovery; 2005. p. 36–43.

41. Kunegis J. Konect: the koblenz network collection. In: Proceedings of the 22nd international conference on world wide web; 2013. p. 1343–1350.

42. Opsahl T, Panzarasa P. Clustering in weighted networks. Social networks. 2009; 31(2):155–163. https://doi.org/10.1016/j.socnet.2009.02.002

A10.

Jankowski, J., Bartków, P., Pazura, P., Bortko, K. (2020). Evaluation of the costs of delayed campaigns for limiting the spread of negative content, panic and rumours in complex networks. In Computational Science–ICCS 2020: 20th International Conference, Amsterdam, The Netherlands, June 3–5, 2020, Proceedings, Part IV 20 (pp. 291-304). Springer International Publishing.

Evaluation of the Costs of Delayed Campaigns for Limiting the Spread of Negative Content, Panic and Rumours in Complex Networks

Jaroslaw Jankowski( ), Piotr Bartkow, Patryk Pazura, and Kamil Bortko

Faculty of Computer Science and Information Technology, West Pomeranian University of Technology, Szczecin, Poland
jjankowski@wi.zut.edu.pl

Abstract. Increasing the performance of information spreading processes and influence maximisation is important from the perspective of marketing and other activities within social networks. Another direction is suppressing spreading processes for limiting the coverage of misleading information, spreading information helping to avoid epidemics or decreasing the role of competitors on the market. Suppressing action can take a form of spreading competing content and it's performance is related to timing and campaign intensity. Presented in this paper study showed how the delay in launching suppressing process can be compensated by properly chosen parameters and the action still can be successful.

Keywords: Information spreading · Competing processes · Limiting the spread · Viral marketing · Misinformation · Rumours · Social networks

1 Introduction

Information spreading processes within the networks are usually analysed from the perspective of the performance with main goal to maximize their coverage. It can be achieved by a proper selection of initial nodes starting propagation among their neighbours. The problem defined as influence maximisation was presented together with greedy solution for finding set of nodes delivering results close to optimum [17]. Apart from greedy approach other possibilities use heuristics based on selection of nodes with high centrality measures [20]. Increasing the spread is central problem for viral marketing, diffusion of products and innovations [12].

The purpose of information spreading can be suppression of other processes [5] in case of the spread of misleading or harmful information, rumours and by marketing companies to compete with other products. One of directions is studying the factors affecting the dominance over other processes [3]. It can be used for spreading competing products, opinion, ideas or digital content in a form of memes, videos [29] or gifts [7]. Information spreading can be also used for stopping real epidemics, with the use of educational information, awareness [10] and related to the rumors

V. V. Krzhizhanovskaya et al. (Eds.): ICCS 2020, LNCS 12140, pp. 291–304, 2020.
https://doi.org/10.1007/978-3-030-50423-6_22

and panic [25]. For planning and launching competing processes, the habituation effect should also be taken into account, because of dropping response to multiple viral marketing messages received in a short time [6].

The performance of the suppressing process is related to various factors including proper timing and campaign parameters. Presented in this paper study shows relation between delays of suppressing actions and the costs required to make them successful. It is assumed that delayed action still can be successful, if the strength of the process is increased, when compared to the harmful process. However, it is related to increased costs of the action, for example better, more appealing content, the number of seeds, more effective way of selection seeds or incentives used.

The remainder of the paper is organised as follows: Sect. 2 includes the related work, the assumptions are presented within the Sect. 3 and are followed by empirical results within the Sect. 4. Obtained results are concluded in the last section.

2 Related Work

Online platforms based on social networks created infrastructures for information spreading processes [2]. Initially activated network members can spread information to their neighbours, in the next step they can spread content to their neighbours and so on. The process continues till saturation point, when new transmissions are not possible. At the end, the process reaches some fraction of nodes or whole network. In most cases the main goal is the increasing the spread of the content to achieve a high number of network nodes influenced [27].

Spread of information can be modelled with various models. Some of approaches use models derived from epidemiology like SIS and SIR with their further extensions [16] while other use branching processes [14]. From the perspective of network structures the most often and well studied models include Independent Cascade Model (ICM) and Linear Threshold Model (LT) [17]. Both approaches were used for various applications and further extensions taking into account time factors or network dynamics [15].

Apart from increasing the process dynamics, quite opposite goals can be taken, to use techniques to block spreading information [26]. To stop epidemics, spreading information about pathogens can be used [10]. For example, several studies were carried on to study spreading information in communication networks to increase awareness [4].

Not only pathogens can be treated as harmful with the need to block spreading. Recently more and more attention is put on spreading the information and digital content which can be potentially harmful or even dangerous for target users. Processes of this type can be based on misleading information and fake news, false medical information, panic and rumours which can negatively influence audiences or promote bad behaviours. It is important to spread alternative information to suppress the negative content and prevent harmful rumours spreading [13]. Similar situation takes place on the market where companies are trying to launch viral campaign or spread information competing with other campaigns [30].

Factors affecting competition were analysed in terms of network structures for multi-layer networks [4]. Another studies take into account immunization strategies based on vaccination of nodes [28]. Another aspects are related to intervals between marketing messages because the ability to process information is limited [19] and habituation effect takes place [24]. As a result high intensity of viral marketing messages received in a short time can be treated as a SPAM [6].

Together with the need of modeling multiple processes within the networks extension of models were proposed. For example Linear Threshold model was extended towards Competitive Linear Threshold Model (CLTM) [11]. The influence blocking maximization problem (IBM) was defined. Authors assumed that positive (+) and negative (−) information spreads within the network. Network nodes can be in three states inactive, +active, and -active. Second main spreading model, the Independent Cascade Model, was extended towards Multi-Campaign Independent Cascade Model (MCICM) [5]. It assumes two campaigns spreading simultaneously within the network with competition mechanism. One of processes is treated as the primary campaign and the secondary, treated as limiting campaign, is decreasing the dynamics and the coverage of the first one. Like in the ICM model activated node had the single chance to activate it's neighbours. The objective of the study was to protect nodes from activation by the first process by activation with the second one.

3 Research Assumptions and Propagation Model

In this paper we analyse the role of timing for suppressing campaign and the relation between the costs of delayed campaigns and their performance. In our research we assume that costs are related to propagation probability and the number of seeds. Propagation probability can be related to incentives, samples quality or others ways to motivate users to propagate the content. Number of seeds is related to fraction of target audience selected as seed and is directly related to campaign costs. The goal is to analyse the costs required to launch successful campaign (treated as a Positive Process) even the delay in the relation to Negative Process takes place. It is assumed that Positive Process is the reaction, that's why the Negative Process starts first.

Positive and negative propagation processes considered in this paper are modelled within network $N(V, E)$ based on vertex set $V = v_1, v_2, ..., v_m$ and edges set $E = e_1, e_2, ..., e_n$. According to the used Independent Cascade model [17] node $u \in V$ is contacting all neighbours, nodes with relation represented by edge $(u, v) \in E$, within the network N and has only one chance to activate node $v \in V$, in the step $t + 1$ with propagation probability $PP(u, v)$ under condition that node v was activated at time t. For our case, Independent Cascade Model is adapted to two concurrent cascades. Probability $PP_{NP}(u, v)$ denotes the probability that node u activates node v one step after node u is activated by a Negative Process. Probability $PP_{PP}(u, v)$ denotes the probability that node u activates node v one step after u is activated by a Positive Process. Two separate seed sets are used to initialise Negative Process and Positive Process. Seed set denoted by $S_{NP} \subseteq V$ is used to initialise the Negative Process. The ranking method R_{NP} is used to select number of seeds according to seeding fraction SF_{NP}. Seed set for Positive Process denoted by $S_{PP} \subseteq V$ such as $S_{NP} \cup S_{PP} = \varnothing$ is used to initialise the negative process. The ranking method R_{PP} is used to select a number of seeds according to seeding fraction SF_{PP}. Every seed node $s_{NP} \subseteq S_{PN}$ is activated in time $t_{NP} = t1$, Every seed node $s_{PP} \subseteq S_{PP}$ is activated in time $t_{PP} \subseteq T$, $T = \{t_1, t_2, ..., t_n\}$. Lets denote by A_{NP}, t the set of active nodes $A_{NP,t} \in V$ possessing the negative information at time t, activated in time point $t-1$ by a Negative Process, and by A_{PP}, t the set of active nodes possessing the positive information $A_{PP,t} \in V$ at time t, activated by a Positive Process in time $t-1$. Let's denote by set of not active nodes $A_{\varnothing,t} \in \{V - A_{NP,t} - A_{PP,t}\}$. Selection of nodes $a_{NP,t}$ newly activated by a Negative Process among all neighbours n_i such as $(n_i, v_i) \in E$ takes place according to the formula:

$$\bigvee_{v_i \in A_{NP,t}} a_{NP} \in \{n \in N(v_i) | n \in A_{NP,t}, n \in (A_{\varnothing,t} + A_{PP,t})\} \tag{1}$$

with probability PP_{NP}. Selection of candidates for activations with positive process a_{PP} takes place among all neighbours n_i such as $(n_i, v_i) \in E$ not active or activated by a Negative Process:

$$\bigvee_{v_i \in A_{PP,t}} a_{PP} \in \{n \in N(v_i) | n \in A_{PP,t}, n \in (A_{\varnothing,t} + A_{NP,t})\} \tag{2}$$

with probability PP_{PP}. The sequence of steps (1) and (2) is taken randomly to deliver equal chances to positive and negative process. All newly activated nodes are included in active nodes sets for the next time point $t+1$, respectively for Negative and Positive Process as $A_{NP,t+1} = a_{NP}$ and $A_{PP,t+1} = a_{PP}$. Process is continued until no more new activations are observed. Final results are represented by coverage $C_{\varnothing}$ with nodes not activated by any process, C_{NP} with nodes activated by a Negative Process and C_{PP} with nodes activated by a Positive Process such as $V = C_{\varnothing} \bigcup C_{NP} \bigcup C_{PP}$.

3.1 Illustrative Example

To clarify the approach, the toy example presents three scenarios. In Fig. 1 a Positive Process (green) is started in the same step like Negative Process (red) and has high chance to suppress it. In the second scenario presented in Fig. 2 Positive Process was started too late and parameters where not enough to dominate Negative Process. Third scenario in Fig. 3 shows that delayed process can survive but parameters like propagation probability should be properly adjusted. It is using a simplified graph with weights assigned to edges. Information flows only if weight on the edge is lower or equal to propagation probability $PP1$ for

Negative Process and $PP2$ for Process. Both $PP1$ and $PP2$ are the same to all nodes according to Independent Cascade Model. Graph has 10 nodes and 17 edges. Both processes are competing. Also the seeds selected for the spreading are given, for Negative Process it is node number 1 and for Positive Process node number 10. Last parameter is delay Di that is responsible for start suppression process (green). For this example process consists of simulation steps. Each step is divided into two stages: stage for Negative Process, and stage for Positive Process. Figure 1 - Figure 3 show how delay can affect the results of reaction depending om process parameters.

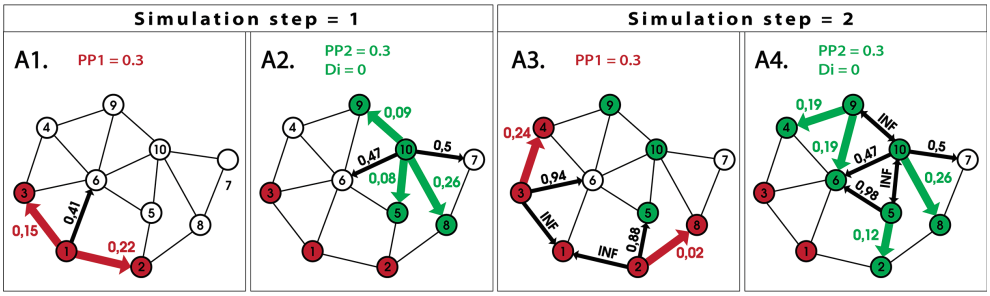


Fig. 1. Example for competing processes with same parameters. In $SIMULATION\ STEP = 1$ **(A1)** infected node 1 begins and try to infect all his neighbors $(2, 3, 6)$. Propagation probability $(PP1 = 0,3)$ allows to infect node 2 $(0,22 < 0,3)$ and 3 $(0,15 < 0,3)$ but prevent to infect node 6 $(0,41 > 0,3)$. **(A2)** node 10 starts suppressing campaign and tries to infect his neighbors (5, 6, 7, 8, 9). $PP2 = 0,3$ allows to infect nodes 5, 8 and 9. In $SIMULATION\ STEP = 2$ cycle repeats for every node with S1 **(A3)** and then with S2 **(A4)**. In this case process 2 begins to defeat process 1. In the next step suppressing process ($S2$) finishes with more nodes activated than process ($S1$). (Color figure online)

3.2 Assumptions for Experimental Study

Information spreading for both Negative Process (NP) and Positive Process (PP) are divided into simulation steps. In each step, each process has a chance to increase coverage with activated nodes contacting their neighbours. Simulation starts with choosing seed nodes according to seeds selection strategy. The spreading process starts with selection of seeds according to their ranking, R_{NP} for Negative Process and R_{PP} for Positive Process. Negative Process can be initialized by choosing random nodes (like in the real world, a disease itself cannot choose node while the carrier does), otherwise marketing strategies are usually based on strictly selected nodes. To try to suppress these two ways of contamination with competing process we choose three seeding strategies based on tree rankings: Random, Degree based and third Effective Degree. We choose selecting nodes by the most common centrality measure, the node degree, treated as a one of main and relatively effective heuristics for seed selection as well as a reference method [12]. Additionally we use effective version of degree, computed before launching Positive Process. It is not based on the total number of the neighbours, but on the number of nodes infected by Negative Process.

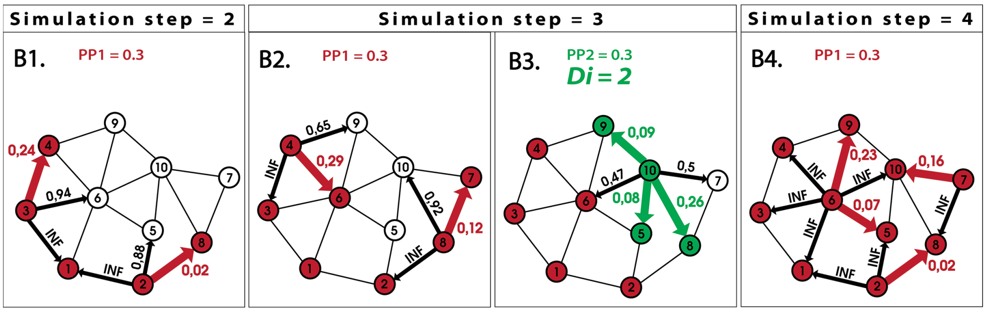


Fig. 2. Competing processes with the same Propagation probability but process no 2 starts with given $Di = 2$. **(B1)** In $SIMULATION\ STEP = 2$ (step = 1 for process no 1 is the same as in Fig. 1 **(A1)**) infected nodes 2 and 3 infect nodes 4 and 8. **(B2)** In $SIMULATION\ STEP = 3$ nodes infected in last step infect next nodes (6 and 7). **(B3)** $Di = 2$ allows process no 2 to start in $SIMULATION\ STEP = 3$ Propagation probability ($PP2 = 0,3$) allows to infect node (5, 8 and 9). **(B4)** shows that process no 1 will win in $SIMULATION\ STEP = 3$. Comparison with Fig. 1. shows that same Propagation probability do not guarantee success if we delay our reaction (Color figure online)

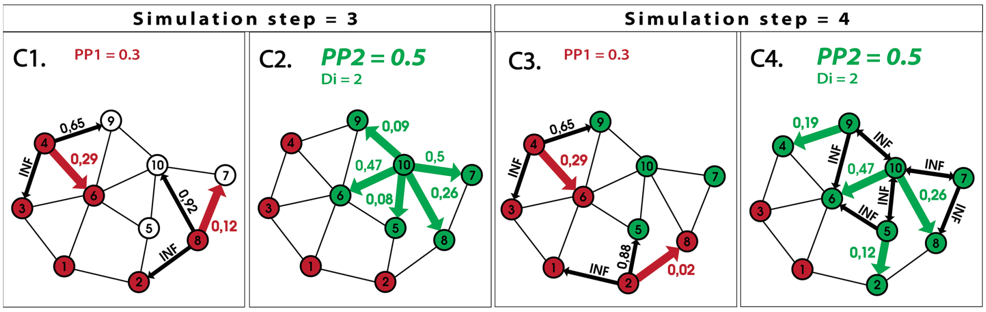


Fig. 3. According to examples above Fig. 3. shows that if we wait with reaction, we need to increase the cost of our campaign to success. **(C1)** In $SIMULATION\ STEP = 3$ (step 1 same as in Fig. 1. **(A1)** and 2 as in Fig. 2. **(B1)**) nodes 4 and 6 infect. In **(C2)** process no 2 starts with same Di as in Fig. 2. but with increased $PP1$ (from 0.3 to 0.5). This affects on possibility to infect nodes 6 ($0,47 < 0,5$) and 7 ($0,5 = 0,5$) which wasn't possible earlier with $PP2 = 0.3$. $SIMULATION\ STEP = 4$ repeats competing process **(C3** and **C4)** which results in victory of a process 2 in next step of a simulation. (Color figure online)

The experiment assumes five different sizes of seed set selected in each used network. Experiments verifies the efficiency of increasing number of seeds. Number of seeds, seeding fraction (SF_{NP}, SF_{PP}) is equal to 1%, 2%, 3%, 4% or 5% and represents the percentage of nodes selected as seeds. Suppressing process (Positive Process) will start with the given delay Di. It can start at the same step or, later (delay: 0–8), to test consequences of late reaction. And finally for both of the competing strategies we give propagation probability (PP_{PP}, (PP_{NP})) equal 0.1, 0.2, 0.3, 0.4 or 0.5. Propagation probability is responsible for the chance to infect node and represents propagation probability according to Independent Cascade Model [17]. During the process for each edge possible for transmission random value is dynamically generated. if it takes value lower or equal to propagation probability (different for Positive Process and Negative Process) activation of contacted node takes place. All values used for all parameters are presented in Table 1.

Table 1. Networks and diffusion parameters used in simulations for Positive Process (PP) and Negative Process (NP)

Symbol	Parameter	Values	Variants
R_{NP}	Ranking type for NP	2	Random, Degree
R_{PP}	Ranking type for PP	3	Random, Degree, Effective Degree
PP_{NP}	NP Propagation Probability	5	0.1, 0.2, 0.3, 0.4, 0.5
PP_{PP}	PP Propagation Probability	5	0.1, 0.2, 0.3, 0.4, 0.5
SF_{NP}	Seeds Fraction for NP	5	1%, 2%, 3%, 4%, 5%
SF_{PP}	Seeds Fraction for PP	5	1%, 2%, 3%, 4%, 5%
Di	Delay in PP initialisation	9	0, 1, 2, 3, 4, 5, 6, 7, 8
N	Network	5	Real networks

Performance of Positive Process can be measured with the use of several metrics. Performance Factor (PF) is represented by a total number of nodes activated by a positive process by the number of nodes activated by negative process for the same configuration parameters. Another metrics, Success Rate (SR), represents percentage of spreading processes with winning Positive Process.

4 Results from Empirical Study

4.1 Experimental Setup

Simulations were performed on five real networks $N1$–$N5$ UoCalifornia [23], Political blogs [1], Net science [21], Hamsterster friendships [18] and UC Irvine forum [22] available from public repositories, having from 899 to 1899 nodes and from 2742 to 59835 edges. We obtained total $R_{NP} \times R_{PP} \times PP_{NP} \times PP_{PP} \times SF_{NP} \times SF_{PP} \times Di \times N$ with the total number 168,750 of simulation configurations. For each of them ten runs were repeated and averaged. The main goal is to investigate the influence of increasing the efficiency of nodes in contaminating their environment. In order to gather necessary knowledge each combination of parameters to find the most successful way to suppress spreading potentially dangerous process as compared. The loop searches through nodes activated by Positive Process, for each infected node the script is looking through its neighbours and tries to infect every neighbour who is not infected with the same disease. Nodes that are infected in this step of contamination cannot spread the disease in the same step. If $Di = 0$, seeds for Negative Process are selected from 'healthy' nodes and the loop repeats spreading but searches through the nodes activated by Positive Process and tries to infect with positive content every node activated by negative process or neutral. If $Di > 0$, Positive Process starts after $Di + 1$ cycles of Positive Process spreading. In this case the simulation step ends when Positive Process step ends, until the suppressing process is activated. The competing lasts until one of the strategies defeat the competitor and spread all over the network or network states stabilize.

4.2 Overall Results

In this section, results from agent-based simulations are presented. During analysis we estimated costs of making effective delayed process. The main goal was answer the question to as far we need to increase propagation probability (PP) to obtain certain success rate (SR) of suppressing process under varying delay steps. Figure 4 shows significance of suppression process. As we can notice, cases with no delay ($Di = 0$) provides the best performance of Positive Processes. Overall, for cases with no delay suppression campaign achieved 31% coverage. Subsequently differences between Positive Process and Negative Process are grown. Negative Process reached the best performance when the Positive Process was most delayed. It was analysed for eight steps of delay ($Di = 8$), and for this case overall negative campaign performance is 62.4%.

For a more detailed evaluation of the diffusion of Positive Process we figured out three factors, presented in the Fig. 5. Used propagation probability (PP), seeds fraction (SF) and networks (N) were analysed. Propagation probability causes the biggest increment of coverage performance. Along with propagation probability, coverage performance is increasing as following: 2.00%, 7.88%, 12.10%, 15.98%, 21.44%. A similar relationships can be seen for seeds fraction (SF) values. However there the growth of performance isn't so drastic. The following results we obtained: 9.72%, 11.10%, 12.13%, 12.86%, 13.59%. In terms of coverage performance $N3$ achieves the highest performance 18.43%. The worst outcome was obtained within $N5$, it 7.39%. Therefore, it is evidence that effect are with relation to network topology.

4.3 Sensitivity Analysis

For sensitivity analysis to determine the key parameters affecting coverage of the positive process the meta-modeling based on the Treed Gaussian Process (TGP) was used. Briefly, TGP is one of the significant machine learning methods, developed by Gramacy [8]. Gramacy et al. further extended TGP to be suitable for sensitivity analysis [9]. Since then, after constructing TGP models, sensitivity analysis can be used to identify key variables in the models by using the variance-based method. We used here two sensitivity indexes: first order and total effects. The first order index represents the main effect and the contribution of each input variable to the variance of the output. The total effects include both main effects and interactions between input variables.

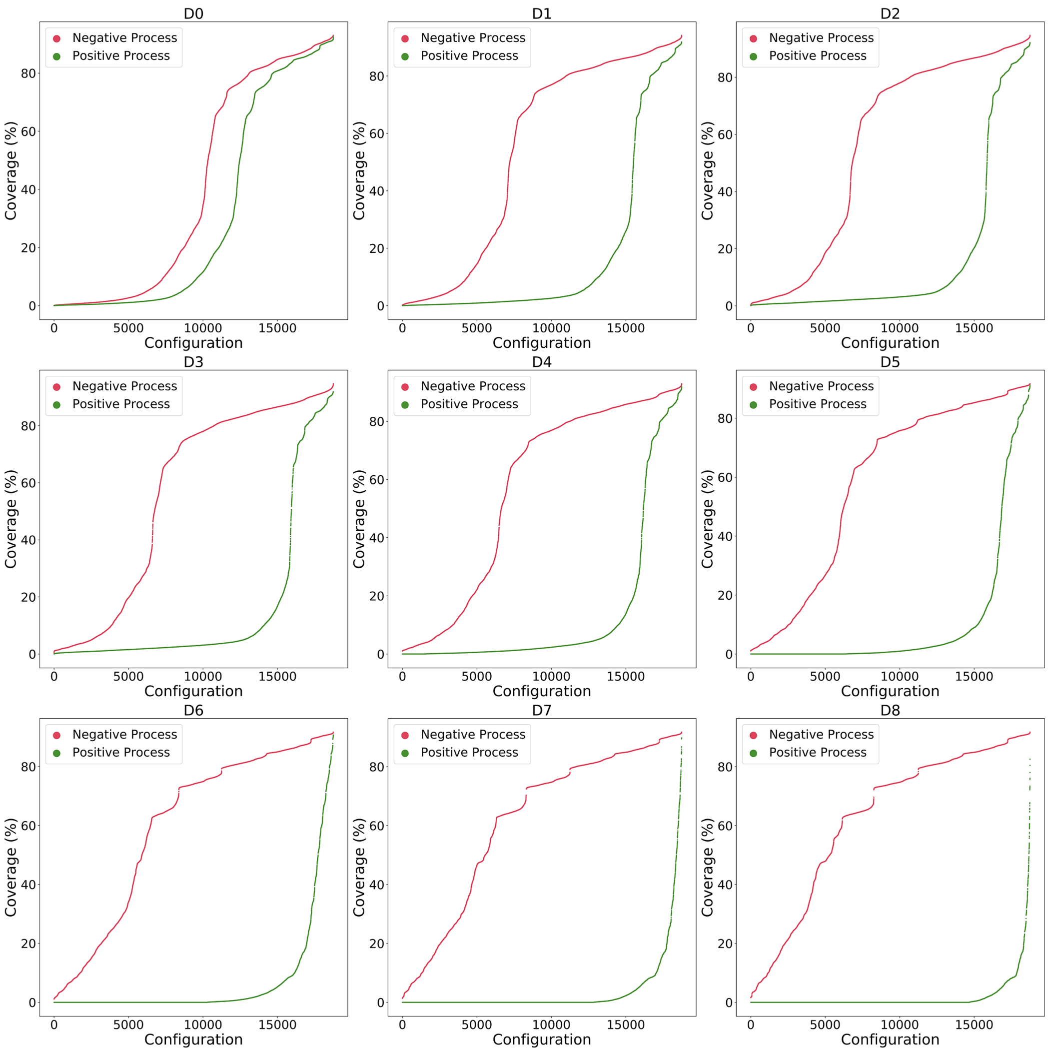


Fig. 4. Comparison of two spreading processes, negative (red) and positive (green) for each combination of configuration parameters. Figure presents relation between both processes, from delay equal to 0 up to delay equal to 8 steps. Along with steps of delay the significance of suppression decrease and distance between both processes grows. Together with increased delay of Positive Process, the Negative Process changes from s-shaped towards increased dynamics. (Color figure online)

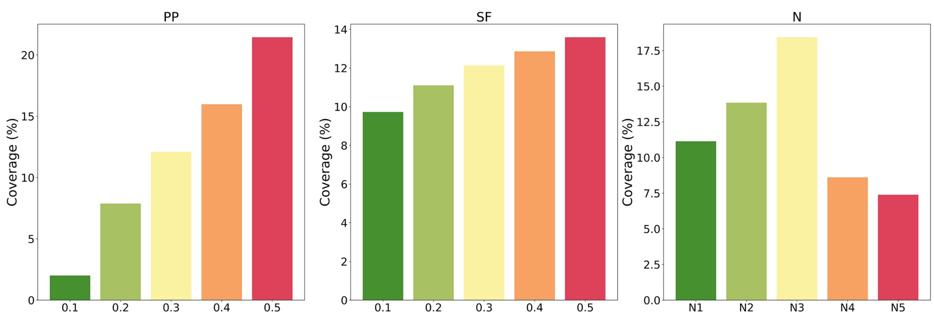


Fig. 5. Average coverage performance of spreading positive process figured for individual values of PP, SF and N respectively. (Color figure online)

The Fig. 6A shows the slopes of the various parameters used in simulations. It provides information on whether the output, performance of the Positive Process, is an increasing or decreasing function of the respective input data. Solid lines are mean values that are within the 95% confidence interval. It was observed that the changes of four parameters SF_{NP}, R_{NP}, R_{PP}, SF_{PP} had only small influence on the output. In addition, with the increase of SF_{PP} and R_{PP}, the main effects caused by the increase of these variables slightly increased, suggesting that the impact of individual differences between SF_{PP} and R_{PP} on changes in the examined networks was slightly improved in such conditions. However, a clear improvement can be seen with the increase of PP_{PP}. Here we see a clearly noticeable increase, which indicates a significant impact of the variable under the existing conditions. Impact of N shows that results were highly dependent on used networks. The increase in the R_{NP} and SF_{NP} variables indicates that the effects caused by their increase have worsened to a lesser extent. It can be clearly seen that increases in delay Di, and PP_{NP} negatively affect the coverage of positive process This means that the variables have a negative impact and their significance deteriorates under the circumstances.

In Fig. 6B first order sensitivity indicators quantify changes in output variables suitably caused by individual input variables while in Fig. 6C sensitivity indicators reflect the interactive effects of all input variables on the output variable. Figure 6B clearly shows that PP_{NP} is the main contributor to the network coverage of PP. Di and PP_{PP} are classified as second and third factors respectively contributing to network coverage of positive process. This differs to some extent from the individual effects shown in Fig. 6A, which can be explained by the combined effects of PP_{NP}, Di and PP_{PP}. The role of remaining variables is approximately the same, sharing small values of the network response coverage of positive process. The cumulative effects Fig. 6C increases when we consider the interactions between all variables, especially for PP_{NP}, to a slightly lesser extent for Di. The sensitivity indicators are not sum to one and it indicates that interactive effects between two or more variables are important for the individual assessment.

4.4 Evaluation the Costs of Delayed Suppressing Process

Another step of analysis includes analysis of Success Rate for different propagation probabilities of both processes. Figure 7 shows Success Rates for each pair of probabilities for Positive Process (PP) and Negative Process (NP). Success Rate (SR) for propagation probability 0.1 for both processes is marked with RED within the Delay 0 section. If delay takes place it was impossible to obtain same Success Rate without increasing Propagation Probability for PP. For example for Delay 4 it was possible to achieve similar Success Rate for $PP_{PP} = 0.2$ and for Delay 5 with $PP_{PP} = 0.5$ (reference cells marked with RED). Success Rate

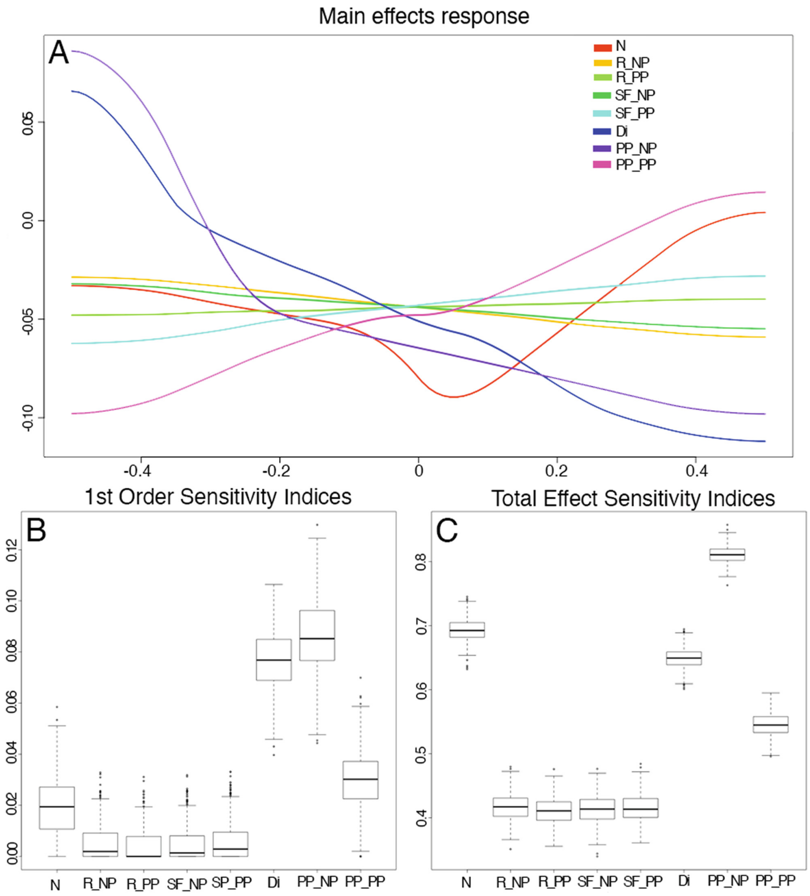


Fig. 6. Sensitivity analysis results for used parameters. **(A)** Main effects response, **(B)** 1st Order Sensitivity Indices and **(C)** Total Effect Sensitivity Indices. (Color figure online)

values marked with GREEN are related to processes competing with Negative Process spreading with $PP_{NP} = 0.2$. In similar way reference values are marked for $PP_{NP} = 0.3, 0.4, 0.5$ with colors BLUE, YELLOW and VIOLET respectively. If Positive Process is delayed one, two or three steps ($Di = 1$, $Di = 2$ and $Di = 3$) properly increased propagation probability makes possible obtaining high Success Rate above 80%. It Positive Process starts with delay four or five steps ($Di = 3$, $Di = 4$) even with high probability only in few cases Success Rate exceeded 50%. Results show that further delay with Positive Process is resulted dropping Success Rate to low ranges. If delay is longer than five steps it was impossible to obtain Success Rate higher than 15% even if probability of negative process was at the level 0.1 and the positive process was launched with probability 0.5%.

Delay 0

	0.1	0.2	0.3	0.4	0.5
0.1	38,93%	87,73%	97,73%	99,60%	99,87%
0.2	2,67%	33,73%	61,33%	81,73%	93,73%
0.3	0,53%	8,53%	32,80%	50,53%	68,93%
0.4	0,00%	0,93%	15,33%	32,00%	44,80%
0.5	0,00%	0,00%	3,60%	19,20%	31,73%

Delay 1

	0.1	0.2	0.3	0.4	0.5
0.1	12,93%	68,40%	90,00%	96,93%	98,27%
0.2	1,73%	8,40%	20,93%	45,07%	68,00%
0.3	0,13%	1,73%	6,67%	13,33%	28,93%
0.4	0,00%	0,13%	1,33%	4,13%	10,40%
0.5	0,00%	0,00%	0,13%	1,33%	4,00%

Delay 2

	0.1	0.2	0.3	0.4	0.5
0.1	4,13%	53,60%	81,60%	88,93%	88,93%
0.2	1,73%	6,67%	14,00%	35,47%	61,87%
0.3	0,13%	1,33%	4,27%	11,20%	24,93%
0.4	0,00%	0,00%	0,80%	2,40%	8,53%
0.5	0,00%	0,00%	0,13%	0,53%	2,00%

Delay 3

	0.1	0.2	0.3	0.4	0.5
0.1	0,53%	45,07%	72,40%	78,40%	80,27%
0.2	0,67%	4,80%	11,87%	30,27%	58,53%
0.3	0,00%	0,40%	3,07%	9,73%	18,53%
0.4	0,00%	0,00%	0,80%	1,87%	6,00%
0.5	0,00%	0,00%	0,00%	0,40%	0,53%

Delay 4

	0.1	0.2	0.3	0.4	0.5
0.1	0,00%	40,67%	57,47%	62,93%	62,00%
0.2	0,13%	2,80%	5,47%	21,73%	37,20%
0.3	0,00%	0,53%	1,87%	7,73%	12,93%
0.4	0,00%	0,00%	0,13%	0,93%	4,53%
0.5	0,00%	0,00%	0,00%	0,00%	0,80%

Delay 5

	0.1	0.2	0.3	0.4	0.5
0.1	0,00%	28,40%	36,00%	36,93%	37,07%
0.2	0,00%	0,67%	1,87%	9,87%	12,53%
0.3	0,00%	0,00%	0,67%	5,60%	8,53%
0.4	0,00%	0,00%	0,13%	0,80%	3,07%
0.5	0,00%	0,00%	0,00%	0,00%	0,13%

Delay 6

	0.1	0.2	0.3	0.4	0.5
0.1	0,00%	11,47%	12,93%	13,73%	14,53%
0.2	0,00%	0,00%	0,00%	0,67%	1,47%
0.3	0,00%	0,00%	1,07%	2,67%	5,60%
0.4	0,00%	0,00%	0,13%	0,53%	2,27%
0.5	0,00%	0,00%	0,00%	0,00%	0,00%

Delay 7

	0.1	0.2	0.3	0.4	0.5
0.1	0,00%	2,27%	2,27%	3,20%	2,67%
0.2	0,00%	0,00%	0,00%	0,00%	0,00%
0.3	0,00%	0,00%	0,13%	0,40%	2,13%
0.4	0,00%	0,00%	0,00%	0,13%	1,60%
0.5	0,00%	0,00%	0,00%	0,00%	0,00%

Delay 8

	0.1	0.2	0.3	0.4	0.5
0.1	0,00%	0,00%	0,00%	0,27%	0,00%
0.2	0,00%	0,00%	0,00%	0,00%	0,00%
0.3	0,00%	0,00%	0,00%	0,13%	0,13%
0.4	0,00%	0,00%	0,00%	0,13%	0,53%
0.5	0,00%	0,00%	0,00%	0,00%	0,13%

Fig. 7. Success Rate (SR) represented by a percentage of winning Positive Processes with corresponding Negative Processes for each pair of propagation probabilities (PP_{NP}, PP_{PP}) and the delay from $Di = 0$ to $Di = 8$. Rows are denoted with PP_{NP} and columns are denoted with (PP_{PP}). Colors for Delay 0 denote selected cases for propagation probability, and same colors in tables with other delays show when was possible to obtain same or higher SR with given PP_{NP} and what required PP_{PP} for delayed Positive Process. (Color figure online)

5 Conclusions

Information transmitted with the use of electronic media spreads with high dynamics. Apart from neutral or positive content online social networks can be used for information or content potentially harmful. Misleading information, rumour or textual information may cause panic and lead to bad behaviours. From that perspective suppressing information spreading processes is challenging and important direction in the area of network science, what was confirmed by earlier studies. Possible actions taken against negative content can be based on launching competing campaigns with the main goal for limiting the dynamics and coverage of negative processes. Later the action is taken it can be less effecting and stopping negative content can be problematic.

Presented study showed how delays in launching positive process is influencing its performance and ability to reduce negative process. It can be achieved by proper adjusting the parameters of limiting process when compared to negative process. Increasing propagation probabilities increases the ability to cope with the negative process, even if limiting action is taken with delay at the moment when large fraction of network is covered by negative content.

Future directions include the role of intervals between messages on campaign performance. Another possible areas include experiments within temporal networks and investigation of the role of changing network topology on performance of limiting actions.

Acknowledgments. This work was supported by the National Science Centre, Poland, grant no. 2017/27/B/HS4/01216.

References

1. Adamic, L.A., Glance, N.: The political blogosphere and the 2004 US election: divided they blog. In: Proceedings of the 3rd International Workshop on Link Discovery, pp. 36–43 (2005)
2. Bakshy, E., Rosenn, I., Marlow, C., Adamic, L.: The role of social networks in information diffusion. In: Proceedings of the 21st International Conference on World Wide Web, pp. 519–528 (2012)
3. Bharathi, S., Kempe, D., Salek, M.: Competitive influence maximization in social networks. In: Deng, X., Graham, F.C. (eds.) WINE 2007. LNCS, vol. 4858, pp. 306–311. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-77105-0_31
4. Bródka, P., Musial, K., Jankowski, J.: Interacting spreading processes in multilayer networks: a systematic review. IEEE Access **8**, 10316–10341 (2020)
5. Budak, C., Agrawal, D., El Abbadi, A.: Limiting the spread of misinformation in social networks. In: Proceedings of the 20th International Conference on World Wide Web, pp. 665–674 (2011)
6. Datta, S., Majumder, A., Shrivastava, N.: Viral marketing for multiple products. In: 2010 IEEE International Conference on Data Mining, pp. 118–127. IEEE (2010)
7. Goode, S., Shailer, G., Wilson, M., Jankowski, J.: Gifting and status in virtual worlds. Journal of Management Information Systems **31**(2), 171–210 (2014)
8. Gramacy, R.B., Taddy, M.A.: Tgp: Bayesian treed gaussian process models. R Package Version, 2–1 (2008)
9. Gramacy, R.B., Taddy, M.: Categorical inputs, sensitivity analysis, optimization and importance tempering with TGP version 2, an R package for treed gaussian process models. University of Cambridge Statistical Laboratory Technical report (2009)
10. Granell, C., Gómez, S., Arenas, A.: Competing spreading processes on multiplex networks: awareness and epidemics. Phys. Rev. E **90**(1), 012808 (2014)
11. He, X., Song, G., Chen, W., Jiang, Q.: Influence blocking maximization in social networks under the competitive linear threshold model. In: Proceedings of the 2012 SIAM International Conference on Data Mining, pp. 463–474. SIAM (2012)
12. Hinz, O., Skiera, B., Barrot, C., Becker, J.U.: Seeding strategies for viral marketing: an empirical comparison. J. Mark. **75**(6), 55–71 (2011)
13. Huang, J., Jin, X.: Preventing rumor spreading on small-world networks. J. Syst. Sci. Complexity **24**(3), 449–456 (2011)
14. Jankowski, J., Michalski, R., Kazienko, P.: The multidimensional study of viral campaigns as branching processes. In: Aberer, K., Flache, A., Jager, W., Liu, L., Tang, J., Guéret, C. (eds.) SocInfo 2012. LNCS, vol. 7710, pp. 462–474. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-35386-4_34

15. Jankowski, J., Michalski, R., Kazienko, P.: Compensatory seeding in networks with varying avaliability of nodes. In: 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2013, pp. 1242–1249. IEEE (2013)
16. Kandhway, K., Kuri, J.: How to run a campaign: optimal control of sis and sir information epidemics. Appl. Math. Comput. **231**, 79–92 (2014)
17. Kempe, D., Kleinberg, J., Tardos, É.: Maximizing the spread of influence through a social network. In: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 137–146 (2003)
18. KONECT: Hamsterster friendships network dataset (2016)
19. Lang, A.: The limited capacity model of mediated message processing. J. Commun. **50**(1), 46–70 (2000)
20. Mochalova, A., Nanopoulos, A.: On the role of centrality in information diffusion in social networks. In: ECIS, p. 101 (2013)
21. Newman, M.E.: Finding community structure in networks using the eigenvectors of matrices. Phys. Rev. E **74**(3), 036104 (2006)
22. Opsahl, T.: Triadic closure in two-mode networks: redefining the global and local clustering coefficients. Soc. Netw. **35**(2), 159–167 (2013)
23. Opsahl, T., Panzarasa, P.: Clustering in weighted networks. Soc. Netw. **31**(2), 155–163 (2009)
24. Vance, A., Kirwan, B., Bjornn, D., Jenkins, J., Anderson, B.B.: What do we really know about how habituation to warnings occurs over time? a longitudinal FMRI study of habituation and polymorphic warnings. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 2215–2227 (2017)
25. Wang, J., Zhao, L., Huang, R.: Siraru rumor spreading model in complex networks. Phys. A **398**, 43–55 (2014)
26. Wang, S., Zhao, X., Chen, Y., Li, Z., Zhang, K., Xia, J.: Negative influence minimizing by blocking nodes in social networks. In: Workshops at the Twenty-Seventh AAAI Conference on Artificial Intelligence (2013)
27. Wang, Z., Chen, E., Liu, Q., Yang, Y., Ge, Y., Chang, B.: Maximizing the coverage of information propagation in social networks. In: Twenty-Fourth International Joint Conference on Artificial Intelligence (2015)
28. Wang, Z., Zhao, D.W., Wang, L., Sun, G.Q., Jin, Z.: Immunity of multiplex networks via acquaintance vaccination. EPL (Europhys. Lett.) **112**(4), 48002 (2015)
29. Wei, X., Valler, N.C., Prakash, B.A., Neamtiu, I., Faloutsos, M., Faloutsos, C.: Competing memes propagation on networks: a network science perspective. IEEE J. Sel. Areas Commun. **31**(6), 1049–1060 (2013)
30. Wu, H., Liu, W., Yue, K., Huang, W., Yang, K.: Maximizing the spread of competitive influence in a social network oriented to viral marketing. In: Dong, X.L., Yu, X., Li, J., Sun, Y. (eds.) WAIM 2015. LNCS, vol. 9098, pp. 516–519. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-21042-1_53

A11.

Fornalczyk, K., Bortko, K., Disterheft A. Jankowski, J. (2023). Modeling the Impact of Video Dynamics on User Engagement and Eye Tracking Patterns. Procedia Computer Science.

Pd: KES2023 General & Invited Tracks : Upload Publication Files

Odebrane

k

kacper_fornalczyk@wp.pl <kacper_fornalczyk@wp.pl>
do kbortko

a

Przetłumacz wiadomość

Dnia 23 maja 2023 22:05 kesgen@prosemanager.com napisał(a):

Dear It Engineer Kacper Fornalczyk

The paper you submitted to KES2023 General & Invited Tracks has been accepted.

Please upload the Publication Files for the paper as follows:-
1. An ABSTRACT with NO title, authors names or details, just the summary itself
- as a TEXT file (named k23-614-abstract.txt), and
- as a word-processor file in MS Word (named k23-614-abstract.doc);
2. A word-processor version of the FULL PAPER in MS Word
(named k23-614.doc)
or a word-processor version of the FULL PAPER in LaTeX
(named k23-614.tex) etc
3. A PDF copy of the FULL PAPER
(named k23-614.pdf).

Please ensure
A. The paper DOES NOT have PAGE NUMBERS.
B. The PAGE SIZE is correct, if A4 is requested do not send the paper in US letter size.

Details of the paper are as follows:-
Paper Id : k23-614
Title : Modeling the Impact of Video Dynamics on User Engagement and Eye Tracking Patterns
Main Author : It Engineer Fornalczyk.
Other authors (if any) : Phd Kamil Bortko, Professor Jarosław Jankowski
Login URL: http://www.kes2023.prosemanager.com/login.asp

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2023) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

27th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2023)

Modeling the Impact of Video Dynamics on User Engagement and Eye Tracking Patterns

Kacper Fornalczyk[a], Kamil Bortko[a,*], Aneta Disterheft[a], Jarosław Jankowski[a]

[a] *West Pomeranian University of Technology, Faculty of Computer Science and Information Technology, Żołnierska 49, 71-210 Szczecin, Poland*

Abstract

Video content on the internet commands a substantial portion of online activities and holds immense value in driving effective marketing campaigns. The question is, how can we measure the potential of video content to capture user attention and how to improve video effectiveness without techniques that increase cognitive load and result in avoidance and skipping of advertisements. This study investigated how video dynamics represented by dedicated metrics is related to eye-tracking patterns and whether video dynamics can be a predictor of user engagement represented by the number of fixations and the duration of gaze on video content. The results showed that selected metrics can serve as predictors of eye movement patterns and that video dynamics have a real impact on user engagement with the viewed video content.

Keywords:
video message; efficiency; dynamics of a video message; eye tracking

1. Introduction

Online video content has become highly prevalent and dominant across the internet, serving as a popular medium for marketing purposes. This includes the utilization of in-stream ads and integration within social platforms, games, and portals, alongside written content. Content creators employ a range of techniques to enhance user engagement, employing strategies such as evoking emotions, utilizing visual effects, and crafting dynamic experiences. However, it is essential to strike a balance, as these techniques can impose cognitive load and divert users' attention from their primary objectives when browsing websites[10]. Consequently, users may swiftly skip advertising content that fails to immediately capture their interest[8]. This poses a challenge for content producers who strive to create video ads[15] that are less likely to be overlooked by consumers. One potential approach is to reduce the intrusiveness and dynamic elements of video content.

* Corresponding author. Kamil Bortko
E-mail address: kbortko@zut.edu.pl

This study aims to delve into the relationship between video dynamics, which are represented by specific metrics, and eye-tracking patterns. The primary objective is to examine whether video dynamics can effectively predict user engagement, as indicated by fixations. By doing so, researchers seek to uncover effective methods for producing videos with low dynamics and cognitive load, while still maintaining sufficient distinctiveness to sustain user attention.

2. Literature review

Although there has been previous research on television advertising, further exploration into online video ads is warranted due to their higher engagement and wider usage compared to static display or text ads [22]. Advertisers and advertising platforms need to understand when video ads become too distracting and result in ad blockers, emphasizing the importance of identifying factors that impact video ad performance [23][19]. This study focuses on measuring video ad dynamics as a performance factor[9], utilizing an algorithm that extracts features like video-level visual variance, scene-to-scene visual variance, and average scene cut frequency [17]. Consumer involvement plays a significant role in ad effectiveness, as higher engagement reduces ad skipping rates [14]. Ad avoidance is associated with cognitive factors linked to engagement, such as arousal levels [4][3]. Engagement level and ad length are critical factors influencing ad avoidance, with longer ads being more prone to skipping behavior[5]. Longer ads also disrupt goal-oriented search more significantly [13, 7]. Nowadays, shorter video ads of fifteen or six seconds align better with consumers' attention spans, allowing ad providers to increase content exposure rates without incurring higher costs [20]. However, longer ads may be more effective in enhancing brand recognition, although shortened fifteen-second ads can achieve comparable awareness and brand recall results as thirty-second spots [17]. Furthermore, this study aims to explore the relationship between video ad dynamics and user engagement by analyzing eye-tracking patterns [24][16]. By understanding how video dynamics influence users' attention and fixation, we can gain insights into optimizing video ad performance[7] [25][2]. Through this research, we strive to enhance our understanding of the factors that contribute to successful video ad campaigns and provide actionable insights for advertisers and advertising platforms. This study examines various factors to evaluate video performance, focusing on how video dynamics and scene differences impact eye-tracking patterns and user engagement at both the single video and intra-scene levels.

3. Conceptual framework

The data flow diagram and stages of processing within the system are presented in the Fig. 1. As depicted in the diagram, the system is capable of operating on various video file formats. Within the system, movies are seen as a collection of frames, from which, with appropriate initial parameters selected by the user of the system, a set of key frames is extracted. This set undergoes transformations and computations defined by visual variation and Agrawal's measures. The results of these operations are values that allow the classification of movies based on their level of dynamics. The knowledge derived from the system enabled conducting research on how the dynamics of video transmission affect user engagement, represented by the number of fixations and the duration of gaze focus on video content. It has been assumed that the user will be able to specify input parameters such as the percentage of key frames for determining dynamics and the number of analysis repetitions from which the dynamics value would be averaged. In accordance with the nature of the measures used, the determination of dynamics will be based on examining the magnitude of variability between consecutive frames. To accomplish this, a certain number of key frames must be extracted from the loaded video into the model while preserving their original order.

4. Experiment

In the presented model, two different measures were used. Each of them can be used separately to determine the dynamics, but in the described model, both measures were included to analyze which one better reflects the subjective assessment of video dynamics made by the human visual system. Both solutions are similar in that they involve calculating the difference between frames, which in turn is based on mathematical operations performed on numerical values corresponding to the colors of individual pixels in two different frames. Detailed information about these solutions is provided in the following section.

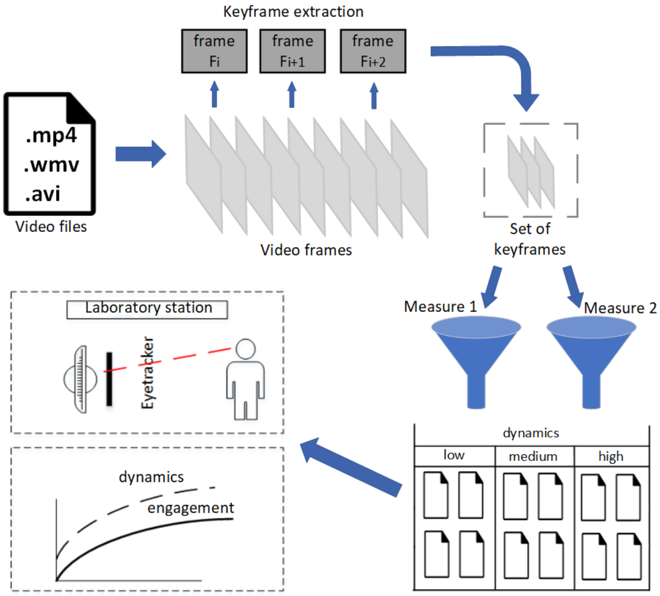


Fig. 1. Conceptual framework for study.

4.1. *Visual Variation Measure*

This measure was proposed by Xi Li, Mengze Shi, and Xin (Shane) Wang and was presented in their article titled ”Video mining: Measuring visual information using automatic methods.” It is a normalized measure of visual information changes in videos and differs from other static image measures in that it captures dynamic changes [18]. The process of determining visual variation for two frames can be defined in the following steps:

1. First, the images representing the selected frames are read as matrices.
2. Then, the images are converted from the RGB color space to grayscale by averaging the color components of each pixel.
3. The next step involves normalizing the values from the range from 0 to 255 to the range from 0 to 1. The authors justify this step as it compensates for any exposure differences [18].
4. Next, the distance between individual pixels of the two frames is calculated, where the distance is defined as the Manhattan norm. This norm can be expressed as follows:

$$d(x'_i, y'_i) = |x'_i - y'_i| \tag{1}$$

5. Having a matrix with dimensions equal to the resolution of the compared frames, where each position represents the distance between corresponding pixels, the last step is to calculate the average value of all distances.

This averaged value represents the information about the magnitude of visual variation between the two frames.

4.2. *Agrawal Measure*

In his guide titled ”A Beginners’ Guide to Image Similarity using Python” [1], Praatek Agrawal presents another, similar method for calculating visual difference between two digital images. Here are the steps involved in determining the difference:

1. Again, the images representing the frames are read as three-dimensional matrices.
2. The three-dimensional matrices representing the RGB channels are flattened into one-dimensional vectors, one for each image.

3. For each vector, a histogram needs to be generated. This can be done by creating arrays with indices from 0 to 255, where each index corresponds to the number of occurrences in the vector with a value equal to that index.
4. With the histograms obtained, the distance between them needs to be calculated, representing the difference between the input images. This can be done using the L2 norm or Euclidean distance.

Unfortunately, it was not possible to determine conclusively whether the method presented above is directly authored by P. Agrawal or if it is just a way and examples of its usage presented in his article.

Nineteen participants took part in the experiment, including 6 women and 13 men, ranging in age from 19 to 41. The participants were divided into two groups: one group consisting of 9 individuals and the other group consisting of 10 individuals. This division into two groups was necessary to ensure that no participant would be presented with an ad of the same content but different durations. For example, in one group, a 6-second ad was displayed at a normal pace, while in the other group, the same ad with a duration of 15 seconds was slowed down by a factor of 2.5. This way, each group of participants was assigned 6 ads from the previously prepared set of 12 variations. The experiment utilized a 27-inch Dell monitor and a Tobii Pro X3 eye tracker with a sampling rate of 120 Hz. The participants sat approximately 54 cm away from the monitor, and calibration was performed before each test. The average duration of the experiment was 8 minutes and 43 seconds.

The purpose of the experiment was to investigate two key research questions. The first question was to analyze how the dynamics of video communication, represented by dedicated metrics, are related to eye-tracking patterns, and whether user engagement with marketing video content depends on the degree of this dynamics. The second area of the experiment aimed to examine whether the results obtained by the proposed model, which analyzes the dynamics of video communication based on metrics that determine the difference between individual video frames using changes in pixel colors, align with the subjective assessment made by the human visual system.

To address these research questions, a specific structure of operation was adopted.

In the first step, a collection of existing video advertisements that met the following criteria was searched for and created:

- Advertisements with durations of 6 and 15 seconds.
- Advertisements generated computationally, meaning they do not depict real-life frames.
- The ad content does not involve humanoid objects.

The selection of 6-second ”bumper ads” and 15-second ads was based on their current popularity as ad formats in terms of duration [20]. Therefore, it was assumed that ads of such lengths would be used in the experiment. The two additional assumptions, that the ads were computationally generated and did not involve humanoid objects, were made to minimize the risk that the content of the ads would influence user engagement rather than their dynamics. It is widely known that people tend to direct their attention to objects resembling humans, especially faces [21]. This tendency may also apply to other elements in the real world. If one video contains a human face while another does not, there is a significant likelihood that the ad with the face will attract more attention, which could interfere with the research results. Additionally, it was planned to slow down the 6-second videos to 15 seconds and speed up the 15-second videos to 6 seconds. This additional step aimed to examine how the assessment of dynamics changes when comparing two different versions of videos with the same content but different durations. When accelerating or decelerating recordings containing humanoid objects, their movement may appear unnatural, which can also impact user attention. Hence, the selection of ads was limited to those that solely presented products without humanoid objects.

The next step involved conducting expert evaluations of the selected video ads. For this purpose, an electronic survey was prepared, allowing the playback of the previously selected video recordings, and it was distributed to individuals knowledgeable about video dynamics. Each expert had the task of assessing the dynamics of each video on a scale from 1 to 10. Finally, the survey results for each video were averaged.

Next, the same videos were analyzed using the created model. The analysis was performed using both measures (visual variation and P. Agrawala measure), utilizing the key frame extraction method through random selection. The analysis started with extracting 5% of key frames from each video, and measurements were repeated 20 times.

The subsequent stage involved dividing the videos into six groups based on the expert evaluation and the results obtained from the model. These groups included videos with low, medium, and high dynamics lasting 6 seconds, as well as corresponding groups for 15-second videos. From each group, one video was selected where the expert assessment and model results were most consistent.

The final step before the main part of the experiment involved preparing additional versions of the selected videos. The 6-second videos were slowed down by a factor of 2.5 to reach a duration of 15 seconds, while the 15-second videos were accelerated to 6 seconds (also by 2.5-fold). This additional step aimed to prepare variations of the selected videos to examine how the assessment of dynamics changes when comparing different versions of videos with the same content but different durations. Finally, the main experiment involved 12 video recordings: the original 6-second and 15-second videos, as well as their modified versions.

With the selected research material prepared, it was possible to proceed with the experiment in a specially designed environment.

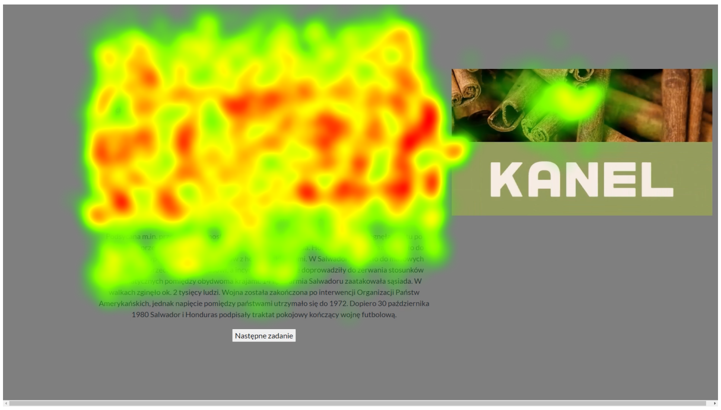


Fig. 2. Heat map obtained from the experiment (source: Own work).

After successful calibration with the eye-tracker, users were shown a welcome screen. This screen could contain information about the purpose of the experiment, instructions on how to proceed during the study, and any other necessary information. The purpose of the welcome screen was to greet the user, ensure their understanding of the research objective, and provide necessary guidelines for conducting the experiment. Depending on the specific context and design of the application, the welcome screen may have included additional elements such as a welcome message, graphics, or navigation options. After the initiation, the user was presented with text displayed on the left side of the screen. On the right side of the text, a randomly selected video advertisement was played in a loop. The looping advertisement was displayed from the moment the text appeared until the user proceeded to the questions related to the read content. The advertisements were shown in a random order without repetitions. As mentioned earlier, each group was assigned 6 videos and corresponding 6 texts.When the user finished reading the text, they would proceed to the questions regarding the read content by using a button placed below the text. After providing their answers, the user would proceed to the next text view by using a button, and a new advertisement would be randomly displayed to them. At no point was the user time-limited. They could read the texts and provide answers for as long as they needed.

5. Results

The results of the video transmission dynamics analysis, conducted using the algorithm developed for this purpose are presented in Tables 1 and 2.

Table 1 presents the results for 3 advertisements lasting 15 seconds (columns ”15s”) and their accelerated six-second versions (columns ”6s sped up”). The results reflect the subjective evaluation of experts (selection of 3 videos with low, medium, and high dynamics). The video rated as low in dynamics for both versions and each measure obtained the lowest dynamic values determined by the system. The video rated as the most dynamic achieved the highest values. The video intended to present content with medium dynamics obtained results close to the average values of the high and low dynamic videos, at least for the non-accelerated versions (15-second videos). However, for

Table 1. Analysis of dynamics for 15-second ads and their sped up 6-second counterparts.

	Agrawal		Visual Variation	
Dynamics	15s	6s mod	15s	6s mod
Low	19391.2018	36582.1904	0.0067	0.0139
Medium	132790.981	180905.27	0.0715	0.1085
High	263933.5143	432426.2827	0.107	0.1572

Table 2. Analysis of dynamics for 6-second ads and their slowed-down 15-second counterparts.

	Agrawal		Visual Variation	
Dynamics	6s	15s mod	6s	15s mod
Low	19475.0796	14534.6453	0.01826	0.0126
Medium	95310.1196	42419.3689	0.0858	0.0431
High	283323.4738	139258.5001	0.1808	0.1026

Table 3. Average concentration time, ratio, number of glances, and fixations with a division into ad dynamics and their pace (15-second videos and their equivalents accelerated to 6 seconds).

	Average time		Ratio		Gaze number		Fixations	
Dynamics	15s	6s mod	15s	6s mod	15s	6s mod	15s	6s mod
Low	1,095091	0,609112	0,032138	0,018887	3	3	6	5
Mid	1,354828	1,578909	0,038139	0,041082	3	4	6	7
High	1,9315	1,911031	0,052796	0,055113	4	5	10	10

the accelerated videos, the differences in values are not as similar. This can be observed, for example, in the ”6s sped up” column for the visual variation measure. The obtained value for the medium dynamic video is higher than that for the low dynamic video by 0.0946, while the value for the high dynamic video is only higher than that for the medium dynamic video by 0.0487. As can be seen, the first difference is nearly twice as large. It can also be noticed that the system does not always reflect a 2.5-fold acceleration in its calculations, which could be expected. For example, for a 15-second video with low dynamics, accelerating it to 6 seconds resulted in an increase of 1.8865 times in the dynamic value according to the Agrawala measure and 2.0746 times according to the visual variation measure. However, for a video with medium dynamics, the increase was only 1.3623 times according to the Agrawala measure and 1.5175 times according to the visual variation measure.

Similarly, Table 2 presents the results for 3 advertisements lasting 6 seconds (columns ”6s”) and their slowed-down fifteen-second versions (columns ”15s slowed down”). Here, too, the results more or less reflect the subjective evaluation of video dynamics made by the experts. In each column, it can be observed that the values resulting from the analysis increase monotonically with the previously determined video dynamics. Here, a similar effect can be observed as in the case of 15-second videos and their accelerated versions. Slowing down a video by 2.5 times does not result in an equal decrease in its dynamics. For example, the value obtained for a video with low expert-rated dynamics decreased by only 1.3399 times according to the Agrawala measure. Slowing down better reflects the dynamics for a video with medium dynamics. Here, the value resulting from the dynamic analysis decreased by 2.2469 times for the same measure. In both cases, a similar pattern can be observed for the visual variation measure.

Tables 3 and 4 contain aggregated data for all study participants regarding the time of gaze fixation on video advertisements, the ratio of time spent watching the advertisement to time spent reading the text, the number of glances at the video advertisement (how many times users left and returned to the area where the advertisement was displayed), and the number of eye fixations on the advertisement. All of this data was recorded for all 12 advertisements, divided by video speed (natural speed and modified) and by dynamics (low, medium, high).

The results presented in Table 3 refer to 15-second videos and their accelerated versions. As can be observed, the average time of gaze fixation on the video increased with the increase in dynamics. This applies to both the unmodified and accelerated versions. However, it is difficult to determine definitively which variant (unmodified or accelerated) captured participants’ gaze for a longer average time. For low-dynamic advertisements, the average time devoted to the accelerated video is 0.486 seconds shorter than its 15-second version. Conversely, for medium-dynamic

Table 4. Average concentration time, ratio, number of glances, and fixations with a division into ad dynamics and their pace (6-second videos and their equivalents slowed down to 15 seconds).

	Average time		Ratio		Gaze number		Fixations	
Dynamics	6s	15s mod	6s	15s mod	6s	15s mod	6s	15s mod
Low	1,104244	0,575244	0,028971	0,015977	4	2	7	3
Mid	0,508769	1,26747	0,014121	0,033762	2	3	3	6
High	1,813391	0,797668	0,04494	0,026128	4	3	9	4

Table 5. Average results from the experiment for different levels of dynamics without grouping.

Dynamics	Average Time	Ratio	Gazes	Fixations
Low	0.8459	0.024	3	5.25
Medium	1.18	0.0318	3	5.5
High	1.61	0.0447	4	8.25

Table 6. Average results from the experiment divided by ad length and tempo.

	15s	6s mod	6s	15s mod
Average time	1.4605	1.3664	1.1421	0.8801
Ratio	0.0410	0.0384	0.0293	0.0253
Gazes	3.3333	4.0000	3.3333	2.6667
Fixations	7.3333	7.3333	6.3333	4.3333

advertisements, participants spent, on average, 0.2241 seconds more on the accelerated version. For low-dynamic videos in both variants, the results are very close to each other. A similar pattern is reflected in the average ratio of time spent watching the advertisement to time spent reading the text. For low-dynamic videos, the ratio is better in their 15-second version, while for medium-dynamic videos, it is better in the accelerated version. Highly dynamic videos have a slightly different ratio between the two variants. When it comes to the number of glances participants made at the videos (moving between the advertisement area and the rest of the interface), the results are low and difficult to interpret. The most noticeable finding is that for highly dynamic videos, the number of returns to the video was higher for both variants compared to low and medium-dynamic advertisements, where the results were nearly identical. The same pattern occurs for the number of recorded fixations. The quantity of fixations clearly outweighs the rest for highly dynamic advertisements. Whether the advertisement is accelerated or not does not seem to have a significant impact on the number of fixations.

Table 4 presents the same type of results but for 6-second video advertisements and their slowed-down counterparts. In this case, the medium-dynamic video in both variants seems to disrupt the expected results. In all categories (average fixation time, ratio, glances, and fixations), the medium-dynamic advertisement breaks the monotonous increase in results. For low and highly dynamic videos, the average time spent watching is respectively 1.9197 and 2.2733 times shorter. However, users spent, on average, 2.4912 longer on the slowed-down version of the medium-dynamic video. In the column representing the ratio, it can be observed that shorter natural-paced advertisements perform better, at least for low and high levels of dynamics. In this case, once again, the results for the medium-dynamic advertisement are opposite – the slowed-down variant achieved a better ratio. The same trend can be seen in the number of glances at the video and the number of fixations.

Table 5 presents the average results in each of the four previously mentioned categories, without dividing them by video speed and length but only by content dynamics. The results are averaged for all advertisements with the same dynamics. Here, it can be observed that in each category, the obtained results increase with the increase in advertisement dynamics. The average time spent watching the video increases by an average of 0.3821 seconds with the increase in dynamics level. Similarly, the ratio increases by an average of 0.0104. Regarding the number of glances, it is similar for each dynamics level, but high dynamics can be distinguished by the fact that users returned to advertisements with this dynamics level on average one more time. If we look at the number of fixations, for low and medium-dynamic videos, the number hovers around 5-6 fixations, while for highly dynamic videos, it ranges from 8-9 fixations. Table 6 presents the average results in the same categories, divided by video speed and duration.

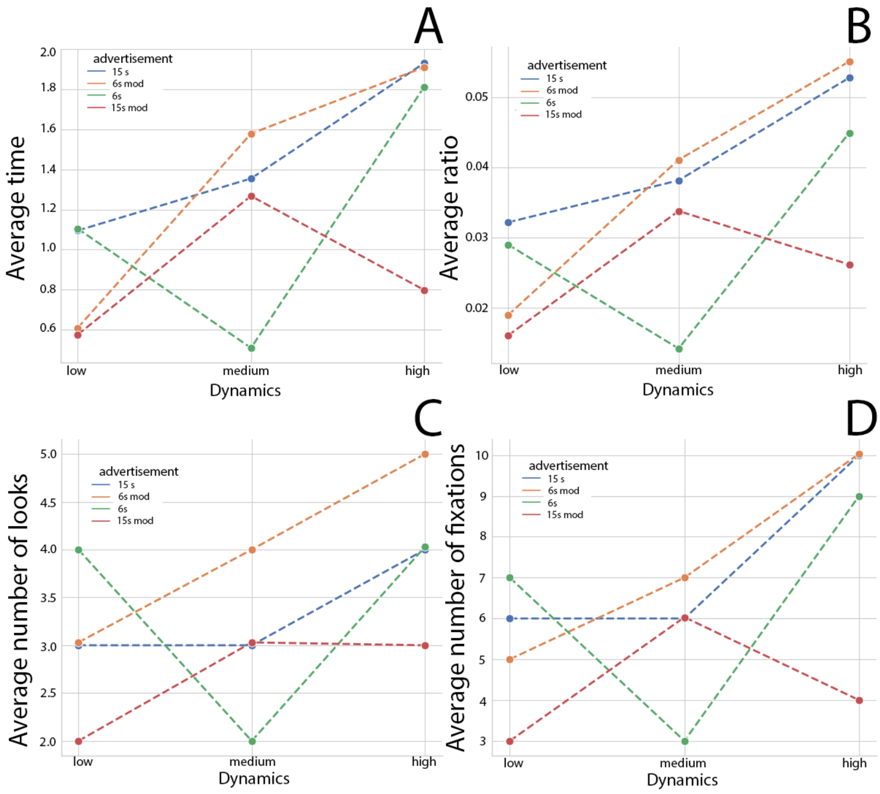


Fig. 3. Charts present a comparison of all ad variants for each video dynamic. The horizontal axis represents the level of video dynamic, while the vertical axis represents the average time spent watching the ad, the average ratio of viewing time to text reading time, the average number of gazes, and the average number of fixations on the ad, respectively.

The results are averaged for all 3 advertisements in each variant, regardless of dynamics. From the results, it can be inferred that participants spent the most time on average watching 15-second advertisements at normal speed. The slow-motion 15-second advertisements performed the worst in this regard, with an average viewing time of less than one second. The same applies to the average ratio of viewing time to reading text for both sets. Regarding the number of visits to the area where the advertisement was displayed, participants were most likely to return to the 6-second advertisements that were sped up. Once again, the slow-motion 15-second videos performed the worst in this aspect. In terms of the average number of fixations, the highest score was achieved by both the 15-second advertisements and their accelerated 6-second variants. Once again, the group of videos slowed down to 15 seconds performed the worst. Figures 3 present a comparison of all advertisement variants for each level of dynamics. The horizontal axis represents the level of dynamics, while the vertical axis represents the average time spent watching the advertisement, the average ratio of viewing time to reading text, the average number of glances, and the average number of fixations on the advertisement, respectively.

Figure 3A illustrates that for the group of 15-second videos and their shortened 6-second versions, the amount of time spent watching the advertisement increases with the level of dynamics. However, this pattern does not hold true for the second group of advertisements. Slowing down the videos to 15 seconds for low and high dynamics resulted in participants spending less time on average on those advertisements. For the medium level of dynamics in this group, the opposite effect can be observed.

The same pattern can be observed in the case of the average ratio of viewing time to reading text, as shown in Figure 3B. A similar pattern can be seen in the results of the average number of glances, as presented in Figure 3C. However, in this case, the only group of videos where the number of returns to the advertisement area does not increase or maintain with the level of dynamics is the group of 6-second videos at a natural pace. In this case, the advertisement selected for medium dynamics exhibited a double decrease in the average number of visits to the video area. For the average number of fixations, as shown in Figure 3D, it can be observed that the number of fixations increases with the level of dynamics for the group of 15-second videos and their accelerated 6-second versions. An exception is observed for 15-second videos with average and low dynamics, where the results are equal. For the group of 6-second videos, we can see that ads with low and high dynamics exhibit a high number of fixations. However, the video with medium dynamics in this group has the lowest average number of recorded fixations. The group of the same videos slowed down to 15 seconds shows the worst results in terms of fixation count for a given dynamics, except for the moderately dynamic advertisement.

6. Conclusions

Based on the results obtained from the conducted study, it can be concluded that as the level of video dynamicity increases, user engagement in content reception also increases. This was demonstrated by the results obtained for the 15-second advertisements and their accelerated 6-second versions. The research showed that in both groups, the viewing time, the ratio of viewing time to time spent on reading the texts, the number of visits, and the number of fixations on the videos increased with the level of dynamicity. Although the analysis of the second group of advertisements (6-second ads and their slowed-down versions) did not confirm this through the average dynamicity advertisement, the regularity can be observed in the results obtained for low and high dynamicity (the results are always better for high dynamicity level). A weak point of this experiment was that, in order to maintain a relatively short and non-fatiguing duration (below 10 minutes), experts selected only one advertisement for each dynamicity level, video length, and tempo. The small research group (19 people) also did not allow for selecting more than one advertisement per variant and dividing the experiment into multiple research groups, which would have allowed for maintaining a short duration for the entire study. Therefore, there are indications to suspect that the average dynamicity advertisement in the 6-second video group may have been inaccurately selected, resulting in results that do not converge with those obtained for the 15-second video group.

The results also do not clearly indicate that slowing down or accelerating the video, while maintaining the entirety of the original content, significantly affects the average viewing time without considering the advertisement's dynamicity. The same can be said for the other categories. The number of fixations, visits, and the ratio do not seem to be dependent on the length and tempo of the videos. The only thing that seems noticeable is that the slowed-down videos (from 6 seconds to 15 seconds) with low and high dynamicity, on average, achieve results that are half as good in all categories except the number of visits. However, the film with average dynamicity did not achieve similar results, which, as mentioned earlier, may have been inappropriately selected for its function. Nevertheless, the results gathered in Table 5 show that the slowed-down videos consistently performed the worst in all categories compared to the other variants. However, this area would require further and deeper research to unequivocally determine whether user engagement actually decreases when the video is slowed down.

Analyzing the results of all advertisements without dividing them into any other groups than classifying them based on three levels of video dynamicity, it can be observed that in all categories, the results improve with an increase in dynamicity level. This is evidenced by the results presented in Table 4.

From the conducted studies, one general conclusion can be drawn, stating that video dynamicity has a real impact on user engagement. The results indicate that as the level of dynamicity increases, user engagement, perceived as both viewing time and the number of visits to the video display area and fixations, seems to increase. However, studies showing that the average level of visual stimulation is perceived as the most satisfying ([6], [12]) may suggest that excessive dynamicity in video content should be avoided, as it may have the opposite effect to increased user engagement.

Effective marketing video content requires the integration of several elements to enhance user engagement based on emotional appeal, dynamics, and attention-grabbing techniques[11]. Despite well-designed content, the intensive use of video content for marketing purposes leads to avoidance behaviors, skipping, or blocking video ads. The user's acceptable video length has been reduced to a few seconds. In such conditions, it is worthwhile to explore methods that allow for the design of effective content without resorting to techniques that spoil the user experience. In the proposed approach, it has been shown that video content creators can attempt to control consumer engagement using scenes with different levels of dynamism.

Future research aims to identify not only video features based on color differences but also object features within the frame to enable the assessment of differences based on scene elements rather than just visual intensity.

7. Acknowledgements

This work was supported by the National Science Centre of Poland, the decision no. 2017/27/B/HS4/01216.

References

[1] Agrawal, P.: A beginners' guide to image similarity using python (accessed on 21/05/2023), https://www.analyticsvidhya.com/blog/2021/03/a-beginners-guide-to-image-similarity-using-python/

[2] Arantes, M., Figueiredo, F., Almeida, J.M., Prates, R.O.: An investigation of user actions and experiences when exposed to youtube video ads. In: Proceedings of the 24th Brazilian Symposium on Multimedia and the Web. pp. 283–290 (2018)

[3] Belanche, D., Flavián, C., Pérez-Rueda, A.: Understanding interactive online advertising: Congruence and product involvement in highly and lowly arousing, skippable video ads. Journal of Interactive Marketing **37**, 75–88 (2017)

[4] Belanche, D., Flavián, C., Pérez-Rueda, A.: User adaptation to interactive advertising formats: The effect of previous exposure, habit and time urgency on ad skipping behaviors. Telematics and Informatics **34**(7), 961–972 (2017)

[5] Belanche, D., Flavián, C., Pérez-Rueda, A.: Brand recall of skippable vs non-skippable ads in youtube: Readapting information and arousal to active audiences. Online Information Review **44**(3), 545–562 (2020)

[6] Berlyne, D.E.: Novelty, complexity, and hedonic value. Perception Psychophysics **8**(5), 279—-286 (1970)

[7] Campbell, C., Mattison Thompson, F., Grimm, P.E., Robson, K.: Understanding why consumers don't skip pre-roll video ads. Journal of Advertising **46**(3), 411–423 (2017)

[8] Edwards, S.M., Li, H., Lee, J.H.: Forced exposure and psychological reactance: Antecedents and consequences of the perceived intrusiveness of pop-up ads. Journal of advertising **31**(3), 83–95 (2002)

[9] Frade, J.L.H., de Oliveira, J.H.C., Giraldi, J.d.M.E.: Advertising in streaming video: An integrative literature review and research agenda. Telecommunications Policy **45**(9), 102186 (2021)

[10] Goodrich, K., Schiller, S., Galletta, D.: Intrusiveness of online video advertising and its effects on marketing outcomes. International Journal of Advertising (2011)

[11] Goodrich, K., Schiller, S.Z., Galletta, D.: Consumer reactions to intrusiveness of online-video advertisements: do length, informativeness, and humor help (or hinder) marketing outcomes? Journal of Advertising Research **55**(1), 37–50 (2015)

[12] Hebb, D.O.: Drives and the cns (conceptual nervous system). Psychological Review **62**(4), 243 (1955)

[13] Hegner, S.M., Kusse, D.C., Pruyn, A.T.: Watch it! the influence of forced pre-roll video ads on consumer perceptions. In: Advances in Advertising Research (Vol. VI), pp. 63–73. Springer (2016)

[14] Jeon, Y.A.: Skip or not to skip: Impact of empathy and ad length on viewers' ad-skipping behaviors on the internet. In: International Conference on Human-Computer Interaction. pp. 261–265. Springer (2018)

[15] Kim, N.Y.: The effect of ad customization and ad variation on internet users' perceptions of forced multiple advertising exposures and attitudes. Journal of Interactive Advertising **18**(1), 15–27 (2018)

[16] Kim, S.: Effects of ad-video similarity, ad location, and user control option on ad avoidance and advertiser-intended outcomes of online video ads. Ph.D. thesis, University of Minnesota (2015)

[17] Li, H., Lo, H.Y.: Do you recognize its brand? the effectiveness of online in-stream video advertisements. Journal of advertising **44**(3), 208–218 (2015)

[18] Li, X., Shi, M., Wang, X.S.: Video mining: Measuring visual information using automatic methods. International Journal of Research in Marketing **36**(2), 216–231 (2019)

[19] Munaro, A.C., Hübner Barcelos, R., Francisco Maffezzolli, E.C., Santos Rodrigues, J.P., Cabrera Paraiso, E.: To engage or not engage? the features of video content on youtube affecting digital consumer engagement. Journal of consumer behaviour **20**(5), 1336–1352 (2021)

[20] Raditya, D., Gunadi, W., Setiono, D., Rawung, J.: The effect of ad content and ad length on consumer response towards online video advertisement. The Winners **21**(2), 119–128 (2020)

[21] Saeideh Bakhshi, David A. Shamma, E.G.: Faces engage us: Photos with faces attract more likes and comments on instagram. in conference on human factors in computing systems. pp. 965–974 (2014)

[22] Schwenzow, J., Hartmann, J., Schikowsky, A., Heitmann, M.: Understanding videos at scale: How to extract insights for business research. Journal of Business Research **123**, 367–379 (2021)

[23] Semerádová, T., Weinlich, P.: The (in) effectiveness of in-stream video ads: Comparison of facebook and youtube. In: Research Anthology on Strategies for Using Social Media as a Service and Tool in Business, pp. 668–687. IGI Global (2021)

[24] Smith, K., Moller, N., Cooper, M., Gabriel, L., Roddy, J., Sheehy, R.: Video counselling and psychotherapy: A critical commentary on the evidence base. Counselling and Psychotherapy Research **22**(1), 92–97 (2022)

[25] Wooley, B., Bellman, S., Hartnett, N., Rask, A., Varan, D.: Influence of dynamic content on visual attention during video advertisements. European Journal of Marketing **56**(13), 137–166 (2022)

A12.

Bortko, K., Bartków P., Jankowski, J. (2023). Modeling the Impact of Habituation and Breaks in Exploitation Process on Multi-Armed Bandits Performance. Procedia Computer Science, 176, 2756-2763.

KES2023 General & Invited Tracks : Upload Publication Files

Odebrane

kesgen@prosemanager.com **przez** **zutedupl.onmicrosoft.com** wt., 23 maj, 22:27

do kbortko

Przetłumacz wiadomość

Wyłącz dla następującego języka: angielski

Dear Phd Kamil Bortko

The paper you submitted to KES2023 General & Invited Tracks has been accepted.

Please upload the Publication Files for the paper as follows:-
1. An ABSTRACT with NO title, authors names or details, just the summary itself
- as a TEXT file (named k23-613-abstract.txt), and
- as a word-processor file in MS Word (named k23-613-abstract.doc);
2. A word-processor version of the FULL PAPER in MS Word
(named k23-613.doc)
or a word-processor version of the FULL PAPER in LaTeX
(named k23-613.tex) etc
3. A PDF copy of the FULL PAPER
(named k23-613.pdf).

Please ensure
A. The paper DOES NOT have PAGE NUMBERS.
B. The PAGE SIZE is correct, if A4 is requested do not send the paper in US letter size.

Details of the paper are as follows:-
Paper Id : k23-613
Title : Modeling the Impact of Habituation and Breaks in Exploitation Process on Multi-Armed Bandits Performance
Main Author : Phd Bortko.
Other authors (if any) : Phd Piotr Bartków, Professor Jarosław Jankowski
Login URL: http://www.kes2023.prosemanager.com/login.asp

Log in to PROSE using the User Id and Password you have been allocated.

If you do not have your login details, access this link, and your details will be emailed to you.

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 00 (2023) 000–000

Procedia Computer Science

www.elsevier.com/locate/procedia

27th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2023)

Modeling the Impact of Habituation and Breaks in Exploitation Process on Multi-Armed Bandits Performance

Kamil Bortko[a,*], Piotr Bartków[a], Jarosław Jankowski[a]

[a]*Faculty of Computer Science and Information Technology, West Pomeranian University of Technology, Szczecin, Poland*

Abstract

Habituation is a common phenomenon in learning, where the response to a repeated stimulus decreases over time. In Multi-Armed Bandits (MAB) algorithms used for visual content delivery optimization, habituation can lead to suboptimal performance. For example, if an agent becomes habituated to a suboptimal arm, it may continue to choose that arm even if better options are available. Habituation can be modeled as a form of "forgetting", where the agent gradually loses confidence in its estimates of the reward probabilities for each arm as time passes. Proposed approach allows updating estimates according to habituation model, while also exploiting the arm with the highest estimated reward probability. The results showed that it is a very good solution to introduce breaks in the multi-armed bandit habit model, which was implemented and tested in our study.

Keywords: MAB alghoritm; Habituation; Computational; Breaks in Habituation

1. Introduction

Habituation is a process in which an organism becomes accustomed to a repeated, unchanging stimulus, leading to a decrease in responsiveness over time[20]. This phenomenon is widely observed in animals, including humans, and is thought to be an adaptive mechanism that allows organisms to filter out irrelevant or redundant stimuli and allocate their attention and resources to more important or novel information[23]. Habituation is a form of non-associative learning and is characterized by a decline in responsiveness that is specific to the habituated stimulus, as well as the recovery of responsiveness after a period of non-exposure[22]. Habituation is a basic form of learning and is thought to be the foundation of more complex forms of learning such as classical and operant conditioning. Habituation is used in many fields as a tool to understand the mechanisms of learning, memory, attention and neural plasticity.There are several ways to address habituation in MAB algorithms. One approach is to use an adaptive exploration strategy that takes into account the agent's past behavior[18]. For example, a strategy called "optimism in the face of uncer-

* Corresponding author. Kamil Bortko
E-mail address: kbortko@wi.zut.edu.pl

tainty" starts with high estimates for the expected rewards of all arms and updates them as the agent explores. This approach encourages the agent to explore new arms and reduces the chance of becoming habituated to a suboptimal arm[16]. Another approach is to use a "restless bandit" algorithm, which periodically resets the agent's estimates of the expected rewards for all arms. This approach also encourages exploration and reduces the chance of habituation.As proposed in this paper habituation-based exploration in multi-armed bandits is a strategy that incorporates the concept of habituation into the multi-armed bandit (MAB) problem[10]. In this strategy, the algorithm adapts to the repeated selections of the same arm by decreasing its responsiveness, encouraging the exploration of other arms. As stated by Auer[5] in their seminal paper on the subject, "Habituation can be seen as a form of exploration, as the agent is forced to try other actions when the current action becomes ineffective." Habituation-based exploration in MAB algorithms has been shown to improve performance by balancing exploration and exploitation and achieving a better trade-off between exploration and exploitation than traditional MAB algorithms. The article was divided into an experiment setup where it was divided into two parts: MAB algorithm with habituation and MAB algorithm with habituation and breaks. Then, the results for both groups of analyzes are presented. Finally, a conclusion was presented with guidelines for future research.

2. Literature review

Habituation, a form of non-associative learning, is defined as the gradual decrease in responsiveness to a repeated, unchanging stimulus[14]. It is a widely observed phenomenon across a range of organisms, including humans, and is thought to be an adaptive mechanism that allows organisms to filter out irrelevant or redundant stimuli and allocate their attention and resources to more important or novel information[21]. As stated by Donald Hebb, a Canadian psychologist, "Habituation is the first and most elementary form of learning" and it is considered as the foundation of more complex forms of learning such as classical and operant conditioning[12]. Habituation is characterized by a decline in responsiveness that is specific to the habituated stimulus, as well as the recovery of responsiveness after a period of non-exposure[20]. Habituation has been widely studied in many fields, including neuroscience, psychology, and biology, to understand the mechanisms of learning, memory, attention, and neural plasticity. Other approach is to use a "contextual bandit" algorithm, which takes into account the current context of the agent, such as its location, time of day, or task, when making its arm selections[13]. This approach can help to reduce habituation by providing the agent with a more diverse set of options to choose from[18]. It's important to note that this is a general overview and further research and studies are needed for a more detailed and complete paper. One of the key challenges in MAB research is the trade-off between exploration and exploitation[9][25]. An agent must explore different options in order to learn about their expected rewards, but it must also exploit the options that it has learned are the most promising[5][11]. Habituation can occur when an agent becomes too focused on exploiting a suboptimal option, and neglects to explore other options[6]. One approach to addressing habituation in MAB algorithms is to use a method called "ε-decreasing" exploration[17]. This method starts with a high value of ε, which represents the probability of the agent choosing to explore an arm, rather than exploit an arm[15]. As the agent continues to play, the value of ε is gradually decreased[2], thus encouraging the agent to explore less and exploit more[4][1]. Next approach is to use a method called "Thompson Sampling" exploration[24]. Thompson sampling algorithms sample from the posterior distribution of the expected rewards for each arm. The agent then chooses the arm with the highest sample. As the agent plays more, the posterior distribution for each arm becomes more certain, thus encouraging the agent to explore less and exploit more. As noted by Bengio et al., habituation is a major factor to take into account in order to properly balance exploration and exploitation in reinforcement learning. They explain that "when a reward is repeatedly received from a given arm, an agent will tend to become less interested in this arm over time," which can be beneficial if the arm's reward is low, but detrimental if the reward changes over time[7].

3. Experiment Setup

3.1. MAB algorithm without habituation

The classic MAB algorithm typically involves the following steps:

1. Initialize the estimates or statistics for each arm, such as the average reward or expected value.

2. Based on a selection strategy, the algorithm decides which arm to choose in each iteration. Common strategies include epsilon-greedy (balancing exploration and exploitation based on a predefined exploration rate),

3. The algorithm selects an arm and receives a reward based on the underlying reward distribution associated with that arm. The reward could be stochastic or deterministic, depending on the problem setting.

4. After receiving a reward, the algorithm updates the estimates or statistics for the selected arm, incorporating the new information. This update helps refine the estimates of the arm's expected reward over time.

5. Steps 2-4 are repeated iteratively for a predefined number of interactions or until a stopping condition is met.

The purpose of the classic MAB algorithm is to strike a balance between exploration and exploitation to maximize the cumulative reward obtained over time. It is used in various applications, such as online advertising, clinical trials, recommendation systems, and resource allocation, where the agent needs to make sequential decisions with limited information. By exploring different arms and learning from the observed rewards, the algorithm gradually improves its ability to select the most rewarding arm while minimizing regret (the difference between the expected reward and the maximum achievable reward).

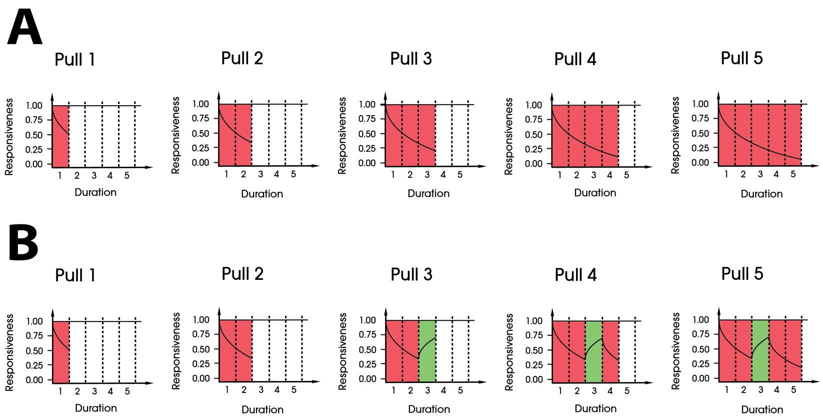


Fig. 1. Example depicting the shape of the responsiveness decline curve with continuous exposure (**A**) and with intermittent exposure breaks (**B**).

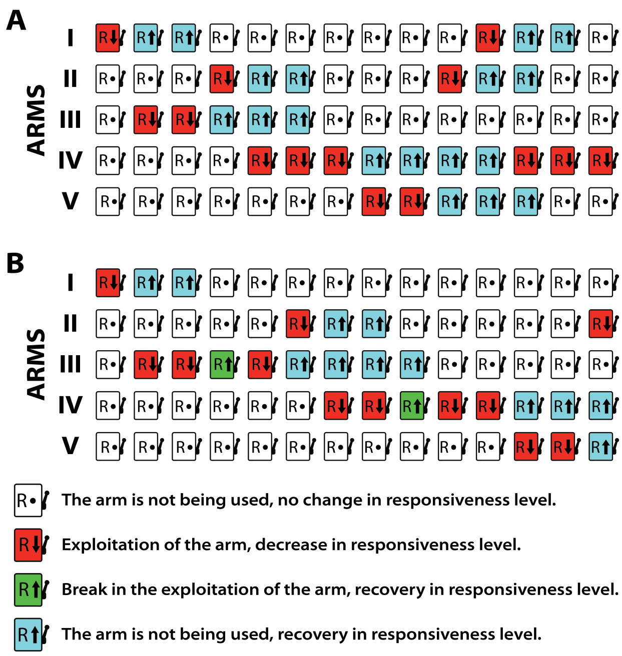


Fig. 2. A graphic illustrating the performance of a Multi-Armed Bandit algorithm in our experiment with the added effect of habituation, with continuous exposure(**A**) and with intermittent exposure breaks(**B**).

3.2. MAB algorithm with habituation

In the context of MAB with habituation, habituation refers to the phenomenon where an organism's response to a stimulus decreases over time with repeated exposure to the same stimulus, when the stimulus is no longer relevant [8]. In MAB algorithms, habituation can be modeled by assigning lower values to arms that have been played more frequently, so that they become less likely to be selected in the future, while arms that have been played less frequently will have higher values and are more likely to be selected.

To mitigate the negative effects of habituation on an agent's performance, various strategies have been developed to balance exploration and exploitation in MAB problems. One popular strategy is ε-greedy, which involves choosing the arm with the highest expected reward with probability (1 - ε) and a random arm with probability ε . Therefore, it is important to consider these factors when designing MAB algorithms that are robust to habituation. To integrate habituation into MAB, each action is assigned a reactivity level of R, which is used to calculate the probability of selecting an arm at a given stage of the simulation. The responsiveness decreases at each stage, which is taken into account by reducing the sampling range according to the habituation curve. If $R < 1.0$, then for a given P_s, its new value for a given action is calculated according to the formula $P_s = P * R$.

By balancing exploitation (selecting the arm with the highest expected reward) and exploration (trying arms with potentially higher reward), the algorithm can converge towards the optimal arm over time.

When an unsuccessful attempt is made, the responsiveness is calculated according to Formula (1):

$$y = y_0 - \frac{S}{\alpha}(1 - \exp(\frac{\alpha \cdot Cnt_{+1}}{\tau})) \tag{1}$$

where y_0 represents the initial habituation value. For inactive contacted arms, it is valid for each discreet time point t. S represents stimulus exposition and in this experiment always takes the value of 1 because of the number of actions in the current time step. α is responsible for the recovery rate. τ is a constant influencing habituation process. t is valid for the time that has passed since responsiveness began to drop.

An increase in responsiveness in the MAB can occur when there is no interaction between the drawn value of the probability of each arms. As we have mentioned, there are no interruptions in the MAB, so the following method was added for the universality of the algorithm and in terms of future research. Growth can be calculated using Formula (2):

$$y = y_0 - (y_0 - y_1)\exp(\frac{-\alpha \cdot Cnt_{-1}}{\tau}) \tag{2}$$

where y_0 represents the initial responsiveness value, equal to 1.0, y_1 refers to the responsiveness value reached during the decrease periods, and t, in this case, represents the time passed from the beginning of the recovery process.

The algorithm discussed in the article belongs to the class of multi-armed bandit algorithms, which are closely related to machine learning and probability theory. The problem of multi-armed bandit algorithms is to distribute limited resources among possible, competing decisions in order to maximize expected gain[19]. The gain parameters of a given decision are not known or are only partially known before the distribution of resources begins. The gain parameters of a given decision are learned as resources are allocated to it. There are two phases of resource distribution among decisions: exploration and exploitation.

During the exploration phase, available resources are distributed among different decisions in order to better understand the gain parameters and the profitability of the decisions. During the exploitation phase, available resources are used only on one decision with the highest expected gain. The resources used only result in better understanding of the exploited decision's parameters

The algorithm discussed in the article is based on the assumptions of the ε-greedy strategy, in which in each iteration, with probability (1 - ε), the decision with the highest expected gain is selected, and with probability ε, a random decision is selected (with the same probability).

The algorithm has been extended with a mechanism for changing the responsiveness of decisions, which can be compared to habituation, involving becoming accustomed to a repeated, unchanging stimulus. This mechanism is intended to affect the length and distribution of the exploration and exploitation phases of decisions. By reducing the use (reducing the length of the exploitation phase) of the decision with the highest expected gain, and increasing the use (increasing the number of exploration phases and the length of the exploitation phase) of decisions with lower expected gains.

This modification, in addition to the ε parameter, also uses the alpha and tau parameters. Each decision, in addition to the gain parameters, also stores its responsiveness level, number of consecutive uses, and number of consecutive non-uses. In each iteration, one decision is used whose responsiveness level drops depending on the number of consecutive uses and the alpha and tau parameters. For the remaining unused decisions, the responsiveness level increases depending on the number of consecutive non-uses and the alpha and tau parameters.

Habituation-based exploration in multi-armed bandits is a strategy that incorporates the concept of habituation into the multi-armed bandit (MAB) problem. In this strategy, the algorithm adapts to the repeated selections of the same arm by decreasing its responsiveness, encouraging the exploration of other arms. As stated by Auer et al. in their seminal paper on the subject, "Habituation can be seen as a form of exploration, as the agent is forced to try other actions when the current action becomes ineffective." [3]. Habituation-based exploration in MAB algorithms has been shown to improve performance by balancing exploration and exploitation and achieving a better trade-off between exploration and exploitation than traditional MAB algorithms.

Habituation in Multi-Armed Bandits (MAB) can be represented conceptually as a process where a player becomes less sensitive to rewards from an arm as they repeatedly pull it what is shown on Figure 1 (B). The normal behavior of the algorithm under ideal conditions reflects Figure 1 (A). The player's behavior and choices are shaped by the rewards they receive, and over time, as the rewards from a particular arm become more predictable, the player will tend to invest less effort into exploring that arm and instead focus on other arms that might offer higher rewards. Figure 3 (A) illustrating the performance of a Multi-Armed Bandit algorithm with the added effect of habituation, with continuous exposure.

3.3. MAB algorithm with habituation and breaks

Figure 3 (B) shows performance of a Multi-Armed Bandit algorithm with the added effect of habituation with intermittent exposure breaks which significantly increase the total reward compared to the approach without interruptions.

In a Multi-Armed Bandit (MAB) problem, the agent must choose which arm to pull in order to maximize the expected reward. However, as the agent pulls the same arm repeatedly, they can become habituated to it, causing a decrease in the level of responsiveness and a decrease in the reward obtained. The consequence of this is that the agent may switch to a new arm, but this process can also be repeated if the content is not properly dosed.

To address this issue, the solution is to take breaks in exposure to the stimulus, even if it is desired by the user. These breaks reduce the level of arm switching, reducing the level of exploration, while causing a simultaneous increase in the level of exploitation. The breaks in exposure are introduced as a parameter to reduce the impact of the habituation phenomenon.

In the MAB problem, every pull has a chance of yielding "information". In the example given, the first two pulls resulted in unsuccessful attempts to upload content, leading to a decrease in responsiveness. On the third pull, the user was not exposed to the stimulus, partially recovering the curve. However, further unsuccessful attempts in the following pulls caused a further decrease in responsiveness. The breaks in exposure are introduced to reduce the impact of habituation by controlling the habituation process.

Breaks in habituation in the context of the multi-armed bandit problem can be represented mathematically as follows:

Let p be a variable representing the probability of a break in the habituation process, a variable that describes the chances of a break . Then, the probability of selecting an arm for the next pull, taking into account the possibility of a breakout, can be expressed by the formula:

$$P(a_t = i) = (1 - p)q_t(i) + p * (1/K) \tag{3}$$

where a_t is the arm selected for the $t - th$ pull, i represents the index of the arm, $q_t(i)$ is the estimated value of the $i - th$ arm at time t, K is the total number of arms, and $(1 - p)q_t(i)$ is the probability of selecting the $i - th$ arm without a break, while $p * (1/K)$ is the probability of selecting any arm at random during a break.

This formula allows for a controlled implementation of breaks in habituation, with the probability of a break being randomly determined at each pull according to the specified range. By introducing breaks, the impact of habituation on the responsiveness of the arms can be reduced, leading to better overall performance in the MAB algorithm This is a chance to take a break while restoring the responsiveness of the other unused arms. The breaks are implemented as a condition that randomly selects from the range of 0 to 1 (e.g. 0.2, 0.3, etc.), determining the probability that the operated arm will rest at a given moment. Observations show that breaks improve results in relation to the implementation of habituation. This may be due to a wider reconstruction of the responsiveness of given arms and a smaller impact of habituation alone on the process.

In the study, the researchers introduced several breaks: 10%, 20%, 30%, 40%, and 50%, to verify their impact on the process. The breaks in exposure can help to control the habituation process and improve the overall performance of the MAB algorithm.

4. Results

The following figure displays a portion of the results obtained from our experiments, illustrating the mean number of decisions made in favor of the option with the highest reward, with a step of 100 over a range of 0 to 3000 iterations.

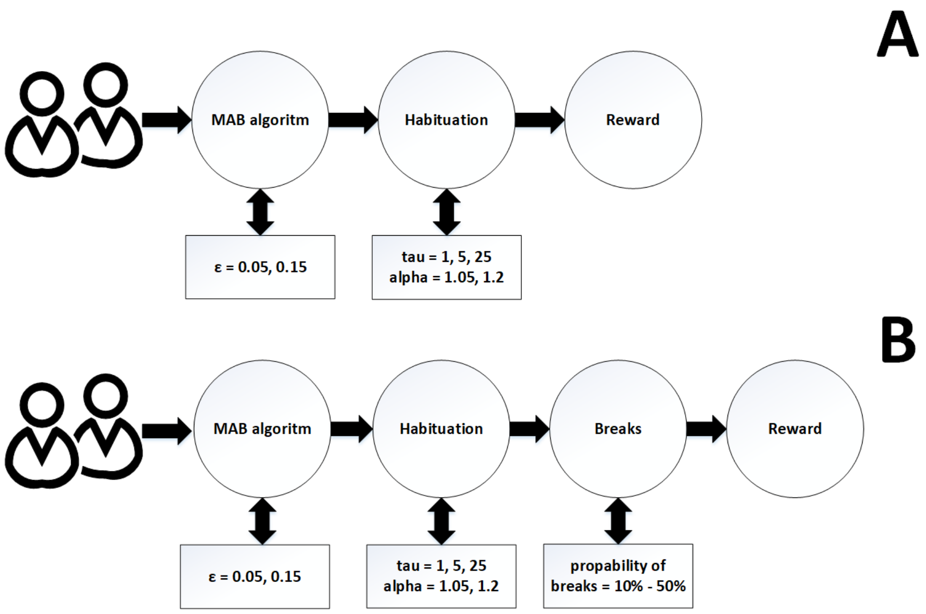


Fig. 3. The flowchart presents two paths of the experiment. Figure (**A**) depicts the experiment with the implemented habituation method in MAB. On the other hand, figure (**B**) illustrates the course of the experiment assuming breaks.

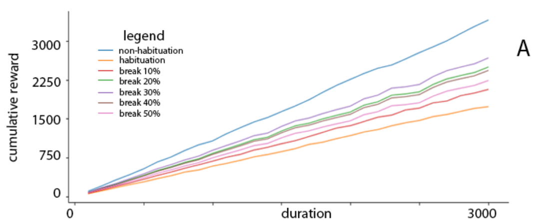


Fig. 4. Figure (**A**) shows average cumulative reward for four curves: without habituation, with 10% breaks to 50% breaks and averaged curves for all combinations of tau and alpha.

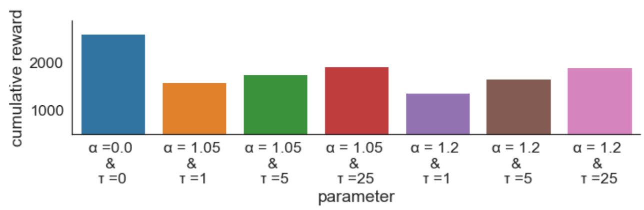


Fig. 5. Bar graph with a division into alpha and tau parameters in various variants in combination with pairs of parameters.

The results vary depending on the chosen values of ε, alpha, and tau parameters, as well as the presence or absence of habituation. The selected parameter values fall within the recommended ranges found in the literature.

Changing the value of ε does not significantly alter the shape of the graph, but increasing its value does lead to lower adopted values. The same holds true for the use of habituation. Here is a general overview of all the curve variations. In the following sections, we will discuss individual groups and divisions.

4.1. MAB algorithm with habituation

It can be observed that changing the accepted values of the alpha parameter with tau remaining unchanged has a greater impact on the values when the value of tau is smaller. Increasing the alpha value results in higher obtained values. The obtained results increase with an increase in tau, but the increase is not proportional.

The results indicate that the habituation mechanism in combination with the ε-greedy strategy reduces the degree of exploitation of the most profitable decision, which increases the exploration and exploitation of decisions with a lower gain.

Figure 5, we can see a curve without habituation with two types of $\varepsilon = 0.05$. We use different habituation parameters: alpha values of 1.05 and 1.2, and tau values of 1, 5, and 25.The parameter pairs shape the results, and the best configuration is achieved by using alpha = 1.05 and tau = 25. The worst pair is alpha = 1.2 and tau = 1. The general trend that can be observed is that it is better to choose a lower alpha, such as 1.05, and tau in direct proportion, meaning the lowest value.

4.2. MAB algorithm with habituation and breaks

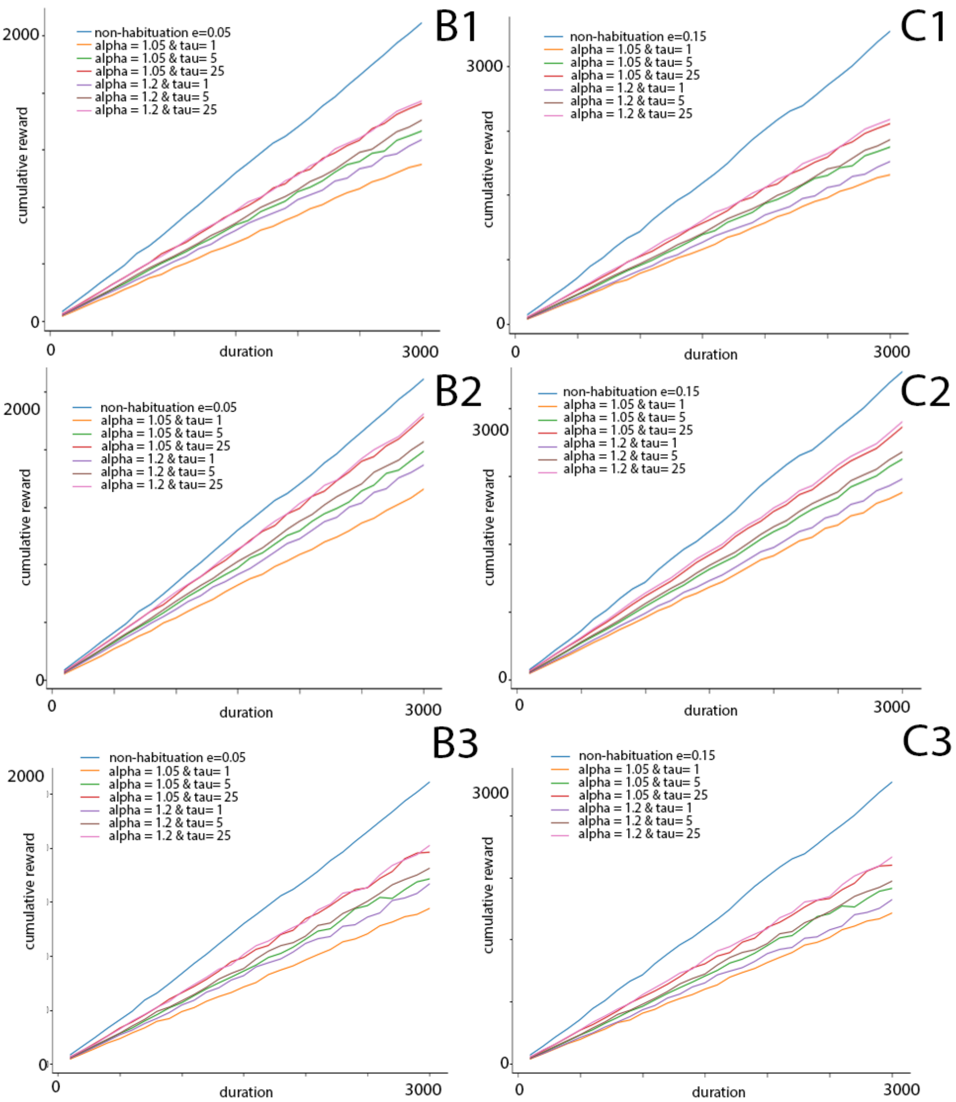


Fig. 6. Figure B1-3 shows all combinations for ε=0.05 - non-habituation, habituation and breaks 30% and 50%. Figures C1-3 shows all combinations for ε=0.15 same as the previous group.

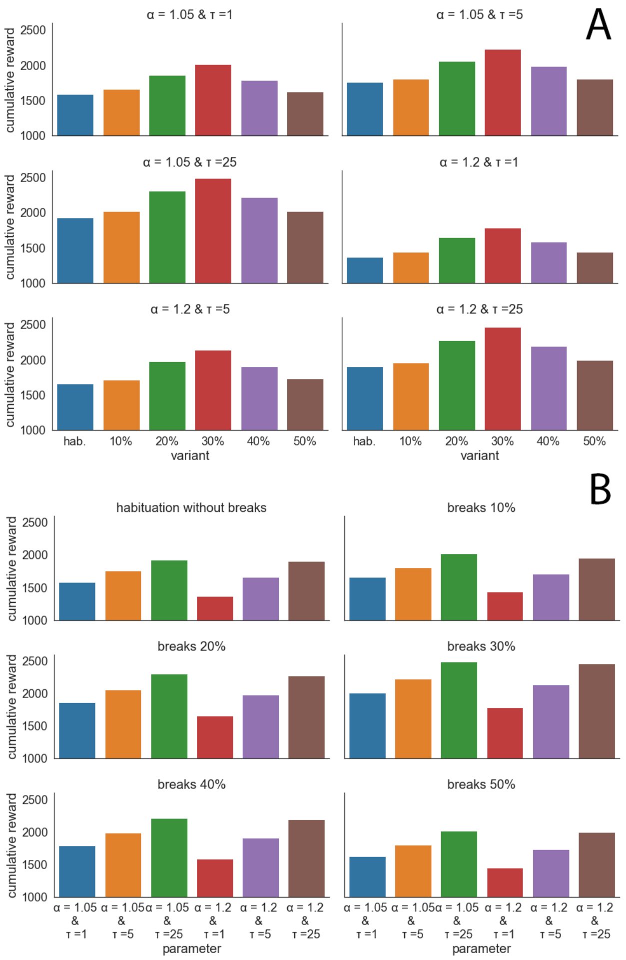


Fig. 7. Bar chart with the division into alpha and tau parameters in different variants of gaps (A) and variants of gaps in combination with pairs of parameters (B) for epsilon = 0.05.

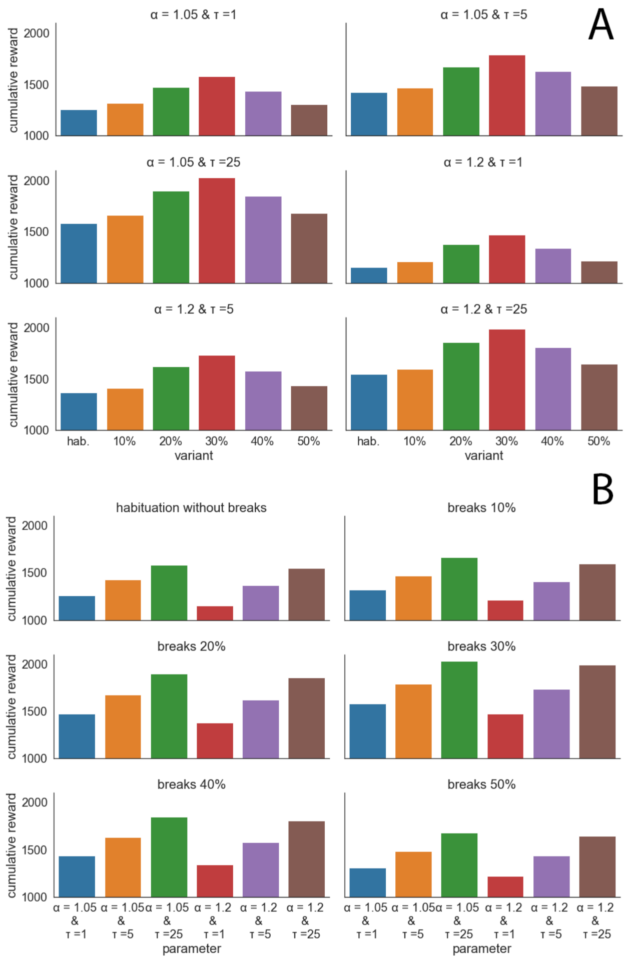


Fig. 8. Bar chart with the division into alpha and tau parameters in different variants of gaps (A) and variants of gaps in combination with pairs of parameters (B) for epsilon = 0.15.

Each combination with the habituation parameters shows the cumulative reward for each option. Cumulative reward with the right parameter will never reach the expected most desirable variant, which is ideal. The true reflection of the approach is the implemented habituation mechanism, which shows how the user behaves naturally.

Figure with imposed gaps indicate an increase in cumulative reward with each approach towards the curves with habituation. Breaks have a positive effect, allowing us to make up for the lost score with a significant number of pulls. The breaks allow for better results compared to the regular habituation approach for individual parameters.

Figure 4 (A) shows the overall average trend of each curve from each approach that we examined. We see here the approach without habituation, the approach with habituation (averaged rewards from each combination of tau and alpha parameters), and few curves with 10%, 20%, 30%, 40% and 50% breaks. We can clearly see the impact of breaks on the non-habitation approach.

The 10% breaks result in approximately 15% gain compared to habituation approaches, while the approach with 20% breaks shows an increase of almost 38% between the habituation curve and the non-habituation curve. The best result is achieved with a 30% break, which is nearly 50% better. This means that we can recover the ”lost” cumulative reward by introducing breaks. Slightly more frequent breaks lead to better results. On the other hand, we observe a downward trend with 40% breaks, where the cumulative value remains at a similar level to the 20% breaks. 50% breaks seem to converge towards the 10% break.

Figure 6 with intervals overlaid indicates a cumulative reward increase with each approach towards the curves with habituation. The intervals have a positive impact, allowing us to recover lost cumulative reward with a significant number of pulls. The charts with habituation parameters show a close to 30-40% drop in cumulative reward compared to the non-habitation MAB approach. As seen in Figure 6 (B1), the parameters themselves do not differ significantly from each other. The intervals allow for better results compared to the regular habituation approach for individual parameters. Figures 6 (B2) and 6 (C2) show 10% intervals. The results improved by almost 12% in each iteration compared to the interval-free runs. Figures 6 (B3) and (C3) show 20% breaks. Here, there is a change of almost 35% in cumulative reward results compared to the curves with habituation parameters. The 10% and 20% intervals have a significant impact, bringing us closer to the expected growth and desired results.

Figure 7 (A) and (B) shows the cumulative reward bars for 6 different variants, divided into 6 different groups of alpha and tau parameters. They represent the approach with implemented habituation without breaks and habituation with breaks ranging from 10% to 50%. Upon examining the results, we can conclude that the best parameters for each of the variants are the set of alpha=1.05 and tau=25, followed by alpha=1.2 and tau=25. The worst result was obtained with the set of parameters alpha=1.05 and tau=1, and alpha=1.2 and tau=1. If we divide them into specific variants, we can see that increasing tau and decreasing the alpha parameter is the most optimal here.

By analyzing the approach with a division into specific sets of tau and alpha parameters, a division into break groups was made. We can see that in each case, a 30% break is the most optimal for each approach. 30% breaks achieve a cumulative reward balance of almost twice as good as other break levels.

By analyzing the sets of parameter pairs and comparing them with the break variants, we can unambiguously conclude that the best option when choosing a combination is to use a 30% break in combination with the set of parameters alpha=1.05 and tau=25, followed by a slightly worse pair of alpha=1.2 and tau=25.

Referring to the Figure 8 (A) and (B) approach, a similar trend to the previous epsilon 0.05 is observed. The cumulative reward bars for six different variants, divided into six different groups of alpha and tau parameters, are depicted in the Figure 8. The variants represent the approach with implemented habituation without breaks and habituation with breaks ranging from 10% to 50%. Upon examining the results, it can be concluded that the optimal parameters for each of the variants are the set of alpha=1.05 and tau=25, followed by alpha=1.2 and tau=25. The worst result was obtained with the set of parameters alpha=1.05 and tau=1, and alpha=1.2 and tau=1. When dividing the variants into specific sets, it is observed that increasing tau and decreasing the alpha parameter is the most optimal approach.

The approach was further analyzed by dividing it into specific sets of tau and alpha parameters, and grouping them based on breaks. It is observed that in each case, a 30% break is the most optimal for each approach. 30% breaks achieve a cumulative reward balance of almost twice as good as other break levels.

By analyzing the sets of parameter pairs and comparing them with the break variants, it can be unambiguously concluded that the best option when choosing a combination is to use a 30% break in combination with the set of parameters alpha=1.05 and tau=25, followed by a slightly worse pair of alpha=1.2 and tau=25.

5. Conclusion

Overall, habituation is a complex phenomenon that can affect the performance of MAB algorithms. However, by using adaptive exploration strategies, restless bandit algorithms, or contextual bandit algorithms, researchers can reduce the chance of habituation and improve the performance of MAB algorithms. It's important to note that this is a general overview and further research and studies are needed for a more detailed and complete paper. In conclusion, habituation is a key aspect of Multi-Armed Bandit (MAB) algorithms, as it helps the agent to balance exploration and exploitation. By modeling habituation as a form of "forgetting", the agent is able to continue updating its estimates of the reward probabilities for each arm while also exploiting the arm with the highest estimated reward probability. Overall, habituation is an important factor to consider in the design and implementation of MAB algorithms to achieve optimal results. In addition to reducing the impact of habituation, breaks in exposure to stimuli in the MAB problem have several other benefits. Firstly, they allow the user to access other arms of the algorithm, which can lead to the discovery of more advantageous decisions. Secondly, breaks prevent the quick depletion of the best arm and prevent the algorithm from getting stuck in suboptimal solutions. In the case of recommendation systems, breaks in habituation can also improve the user experience by preventing the display of the same recommendations too often and increasing the diversity of proposed content. This can have a positive impact on the length of user sessions and increase engagement with the platform.

Finally, applying breaks in exposure to stimuli can improve the stability and efficiency of the MAB algorithm, reducing the impact of randomness on results and increasing the precision of decisions. Overall, breaks in habituation are an important tool in the MAB problem and can help achieve better results and user experiences.

In the real world, such as the advertising world he goal of the advertising campaign is to maximize revenue from displaying ads. The advertiser earns money each time a user clicks on the offer. Similar to the case of MAB , there is a trade-off between exploration, which aims to gather information about the effectiveness of the ad using click-through rates, and exploitation, in which we stick to the ad that has been the best so far. Future research should focus on testing different models of multi-armed bandits and test various combinations of tau and alpha.

Acknowledgements

This work was supported by the National Science Centre of Poland, the decision no. 2017/27/B/HS4/01216

References

[1] Audibert, J.Y., Bubeck, S., Munos, R.: Best arm identification in multi-armed bandits. In: COLT. pp. 41–53 (2010)
[2] Audibert, J.Y., Munos, R., Szepesvári, C.: Exploration–exploitation tradeoff using variance estimates in multi-armed bandits. Theoretical Computer Science **410**(19), 1876–1902 (2009)
[3] Auer, P.: Using confidence bounds for exploitation-exploration trade-offs. Journal of Machine Learning Research **3**(Nov), 397–422 (2002)
[4] Auer, P., Cesa-Bianchi, N., Fischer, P.: Finite-time analysis of the multiarmed bandit problem. Machine learning **47**, 235–256 (2002)
[5] Auer, P., Cesa-Bianchi, N., Freund, Y., Schapire, R.E.: The nonstochastic multiarmed bandit problem. SIAM journal on computing **32**(1), 48–77 (2002)
[6] Bastani, H., Bayati, M., Khosravi, K.: Mostly exploration-free algorithms for contextual bandits. Management Science **67**(3), 1329–1349 (2021)
[7] Bengio, Y., Louradour, J., Collobert, R., Weston, J.: Curriculum learning. In: Proceedings of the 26th annual international conference on machine learning. pp. 41–48 (2009)
[8] Dubey, R., Griffiths, T.L., Dayan, P.: The pursuit of happiness: A reinforcement learning perspective on habituation and comparisons. PLoS computational biology **18**(8), e1010316 (2022)
[9] Galichet, N., Sebag, M., Teytaud, O.: Exploration vs exploitation vs safety: Risk-aware multi-armed bandits. In: Asian Conference on Machine Learning. pp. 245–260. PMLR (2013)
[10] Greenewald, K., Tewari, A., Murphy, S., Klasnja, P.: Action centered contextual bandits. Advances in neural information processing systems **30** (2017)
[11] Grover, A., Markov, T., Attia, P., Jin, N., Perkins, N., Cheong, B., Chen, M., Yang, Z., Harris, S., Chueh, W., et al.: Best arm identification in multi-armed bandits with delayed feedback. In: International Conference on Artificial Intelligence and Statistics. pp. 833–842. PMLR (2018)
[12] Hebb, D.O.: What psychology is about. American Psychologist **29**(2), 71 (1974)
[13] Hillel, E., Karnin, Z.S., Koren, T., Lempel, R., Somekh, O.: Distributed exploration in multi-armed bandits. Advances in Neural Information Processing Systems **26** (2013)
[14] Jankowski, J.: Habituation effect in social networks as a potential factor silently crushing influence maximisation efforts. Scientific Reports **11**(1), 19055 (2021)
[15] Karnin, Z., Koren, T., Somekh, O.: Almost optimal exploration in multi-armed bandits. In: International Conference on Machine Learning. pp. 1238–1246. PMLR (2013)
[16] Killian, J.A., Lalan, A., Mate, A., Jain, M., Taneja, A., Tambe, M.: Adherence bandits (2023)
[17] Lu, T., Pál, D., Pál, M.: Contextual multi-armed bandits. In: Proceedings of the Thirteenth international conference on Artificial Intelligence and Statistics. pp. 485–492. JMLR Workshop and Conference Proceedings (2010)
[18] Mintz, Y., Aswani, A., Kaminsky, P., Flowers, E., Fukuoka, Y.: Nonstationary bandits with habituation and recovery dynamics. Operations Research **68**(5), 1493–1516 (2020)
[19] Pike-Burke, C., Grunewalder, S.: Recovering bandits. Advances in Neural Information Processing Systems **32** (2019)
[20] Rankin, C.H., Abrams, T., Barry, R.J., Bhatnagar, S., Clayton, D.F., Colombo, J., Coppola, G., Geyer, M.A., Glanzman, D.L., Marsland, S., et al.: Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. Neurobiology of learning and memory **92**(2), 135–138 (2009)
[21] Slivkins, A., et al.: Introduction to multi-armed bandits. Foundations and Trends® in Machine Learning **12**(1-2), 1–286 (2019)
[22] Stanley, J.C.: Computer simulation of a model of habituation. Nature **261**(5556), 146–148 (1976)
[23] Thompson, R.F.: Habituation: a history. Neurobiology of learning and memory **92**(2), 127 (2009)
[24] Thompson, W.R.: On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. Biometrika **25**(3-4), 285–294 (1933)
[25] Zhou, L.: A survey on contextual multi-armed bandits. arXiv preprint arXiv:1508.03326 (2015)