

Włodzimierz Gogołek
Wlodzimierz@Gogolek.pl
Instytut Dziennikarstwa
Uniwersytet Warszawski
Warszawa

Spolegliwość komputerów – personalizacja

1. Wstęp

Znaczącym uwarunkowaniem efektywnego wykorzystania potencjału komputerów i Sieci (Internetu) jest sprawna komunikacja pomiędzy ludźmi i maszynami. W znacznej części o sprawności wymiany informacji pomiędzy ludźmi, a światem nieożywionym decyduje dostosowywanie komputerowego przekazu do informacji opisujących ich odbiorcę i jego otoczenie. Chodzi tu o wykorzystywanie, poprzez stosowanie odpowiednich, zautomatyzowanych procedur, analiz rosnącej liczby danych charakteryzujących odbiorcę informacji – człowieka. Jest to spełnieniem oczekiwań tego, iż mimo że *każdy użytkownik Internetu jest unikatowy, a jego sposób korzystania z Internetu – specyficzny*, informacje, które przekazuje komputer człowiekowi, uwzględniają jego indywidualność [Krug, 2006, s.140]. Spolegliwość komputerów stała się kontenerem informacyjnym sprawnie przekazującym treści z pamięci maszyn do umysłu ludzkiego.

2. Składowe personalizacji

Personalizacja – tak potocznie nazywana jest procedura aktywnego udziału maszyn podczas komunikacji z użytkownikami – obejmuje rosnącą ilość szerokiego spektrum informacji opisujących uczestnika procesu komunikacji oraz stosowny do nich wybór operacji wykonywanych przez maszyny. Na personalizację procesu wymiany informacji pomiędzy Siecią a jej użytkownikami składa się coraz większy zakres informacji – od bardzo ogólnych, neutralnych etycznie do intymnych. Budowanie profilu użytkownika, jego sylwetki, jest oparte na:

- informacjach formalnych; są to:
 - informacje wyboru – przekazywane świadomie przez użytkownika do komputera w celu dostosowywania przesyłanych do niego informacji będących przedmiotem aktu komunikacji; są one uzyskiwane w wyniku wyboru dokonanego przez użytkownika, m.in. formy odczytywanych ze stron informacji, np. krój, kolor, wielkość liter, układ strony – rozmieszczenie i wybór modułów tworzących stronę,
 - informacje ankietowe – pozyskiwane przez komputer od użytkownika za pośrednictwem standardowych narzędzi typu ankiet, formularz (np. płeć, wiek, zainteresowania, miejsce zamieszkania);
- informacjach analitycznych uzyskiwanych w wyniku zabiegów technologicznych, najczęściej bez wiedzy i bez zgody użytkownika; są to:
 - informacje formalne – pora dnia, pora roku, lokalizacja geograficzna odbiorcy (geolokalizacja, geotargeting); adresy IP, mimo dostępnych narzędzi służących

anonimowym wędrówkom po wirtualnym świecie, pozwalają w 99% określić kraj, w którym znajduje się użytkownik Sieci, także region i miasto [Goldsmith, 2006, s. 61].

- informacje behawioralne – pozyskiwane za pomocą różnego rodzaju procedur śledzenia zachowań użytkownika (zazwyczaj bez jego wiedzy) oraz analizy tych zachowań; dotychczas najczęściej są to dane wolnozmiennie (zainteresowania, preferencje), ale także mogą być nimi dane szybkozmiennie, wskazujące np. na aktualny stan psychofizyczny użytkownika,
- informacje kontekstowe – wydobywane z treści przekazywanych przez użytkowników do maszyn, np. analiza pytań wpisywanych do wyszukiwarek, analiza treści korespondencji elektronicznej i blogów użytkownika,
- dane statystyczne – zbiorcze informacje uzyskane na podstawie rejestracji zachowań użytkowników w Sieci; są one poddane obróbce statystycznej w celu wyszczególnienia różnych od siebie grup użytkowników (np. płeć, wiek) oraz prawidłowości odbierania (np. zainteresowanie lub brak zainteresowania określoną tematyką wiadomości) wyróżnionych przez te grupy form i treści informacji.

Zasygnalizowane wyżej informacje są wykorzystywane w procedurach opisanych specjalnymi algorytmami. Dzięki nim można z dużym prawdopodobieństwem przewidzieć kolejne zachowania określonej osoby, np. podczas nauczania wspomaganego komputerem czy kampanii reklamowej lub wyborczej [Baker, 2008, s. 14]. Służy temu tak zwana personalizacja serwisów WWW, która polega na *wykorzystywaniu znanych profili preferencji do dynamicznego dostosowywania zawartości serwisu do potrzeb poszczególnych użytkowników. Adaptatywne serwery WWW automatycznie odkrywają typowe schematy zachowań użytkowników, analizując informacje o użytkowaniu serwisu zawarte w logu serwera technikami eksploracji danych* [Wojciechowski, 2002, s. 141-152].

3. Model personalizacji w nauczaniu wspomaganym komputerem

W procesie nauki wspomaganą komputerem celowe jest, poza wspomnianymi informacyjnymi uwarunkowaniami, przystosowanie procedury nauczania do aktualnego stanu psychofizycznego ucznia. Teoretyczne podstawy tak rozumianej personalizacji w odniesieniu do procesu nauczania zostały opracowane już w 1976 roku [Gogołek, 1976]. Udowodniono możliwość wykorzystania zarówno danych wolnozmiennych (doświadczenia i cechy organizmu), jak i szybkozmiennych – opisujących stan psychofizyczny uczącego się (np. komputerowa analiza fal EEG, pobudzenia obwodowego (GSR), czas reakcji na określone bodźce). Pomijając procedurę przekazywania treści merytorycznych, bieżąca kontrola tych parametrów pozwala oceniać aktualny stan psychofizyczny ucznia (świadczący o gotowości do nauki) oraz wpływ, jaki wywierają na niego komunikaty (wzmocnienia instrumentalne). Wielokrotne korzystanie z owych komunikatów pozwala na ich klasyfikację. W efekcie system się uczy, jakie komunikaty i w jakich okolicznościach wywołują zamierzony skutek – np. stosowne zwiększenie poziomu aktywności emocjonalnej uczącego się (by proces nauki przebiegał efektywnie). Dzięki autonomii maszyn system ocenia potrzebę (wielkość i kierunek) zmiany aktywności emocjonalnej i odpowiednio do niej dobiera odpowiednie bodźce (inne dla różnych osób, stosownie do ich stanu

psychofizycznego oraz wcześniej opisanych informacji wolnozmiennych). Idąc dalej w poszukiwaniu precyzyjnej oceny stanu psychofizycznego użytkownika maszyny, wykorzystywany jest m.in. tzw. neuromarketing. Jego przedmiotem jest badanie reakcji psychofizjologicznych konsumentów (ocenianych za pomocą EEG i fMRI – funkcjonalnego magnetycznego rezonansu jądrowego), które warunkują np. kupno określonego produktu lub poprawę znajomości określonej marki produktu (badania takie w zakresie kosmetyków prowadził Christian Dior) [Lindstrom, 2008, s. 195]. Towarzysząca temu wymiana informacji, z udziałem maszyn, stanowi przedmiot zainteresowania ekspertów medycyny, technologii oraz marketingu i z powodzeniem może być wykorzystana w procesie nauki wspomaganej komputerami.

Personalizacja jest zatem wynikiem umiejętnej analizy i interpretacji danych mówiących o osobie współpracującej z maszyną. Jest także wykorzystaniem swojego doświadczenia – danych otrzymywanych podczas procesu wymiany informacji pomiędzy człowiekiem a komputerem (Siecią): bodziec – reakcja. Chodzi tu o śledzenie reakcji użytkowników na bodźce (komunikaty) generowane przez maszyny. Dzięki ogromnej liczbie pomiarów takich par informacji, a także dostępnymi informacjom o użytkowniku, odpowiednie mechanizmy programowe coraz trafniej dobierają bodźce, które mają spowodować oczekiwaną reakcję użytkownika (np. zwiększenie koncentracji ucznia podczas nauki wspomaganej maszyną). Zarysowany proces: bodziec – reakcja, wskazuje na widoczną autonomię maszyn, która w dużej części polega na korzystaniu z efektu sprzężenia zwrotnego (maszyna *uczy się* skuteczności bodźców), a w konsekwencji na dobieraniu coraz bardziej efektywnych komunikatów przekazywanych użytkownikowi. Użytkownik Sieci otrzymuje coraz większą część informacji w postaci łatwo absorbowanej pigułki informacyjnej. Dzięki personalizacji następuje dostosowanie tej pigułki do sylwetki odbiorcy informacji – do jego indywidualnych oczekiwań (predyspozycji).

Sztandarowym przykładem potencjału personalizacji są zasoby informacyjne Google'a, który dysponuje ogromnymi zbiorami danych o milionach swoich użytkowników. Ich źródłem, poza historiami wyszukiwań w Sieci, są także treści korespondencji elektronicznej przechowywanej w tym portalu, wypełnione arkusze kalkulacyjne i albumy fotograficzne [Gold, 2008].

4. Dane behawioralne

Ogromną wartość w kształtowaniu spolegliwości maszyn mają informacje pozyskiwane od odbiorcy bez jego wiedzy podczas czynnego korzystania z Sieci – tak zwane dane behawioralne – zawierające informacje o jego zachowaniach w Internecie. Wskazują one, między innymi, na jego zainteresowania czy okresy aktywności intelektualnej [Chen, 2005]. Wykorzystywane są zainteresowania internautów, z precyzją obejmującą takie szczegóły, jak np. w odniesieniu do filmowych entuzjastów: komedie, dramaty, filmy niezależne, aktorzy itd., a do kobiet obejmują m.in.: *seks i miłość, zdrowie, jedzenie, karierę i pieniądze, rozrywkę, kosmetyki, życie domowe i astrologię* [Yahoo..., 2008].

W efekcie maszyny dysponują zasobami informacyjnymi, które po niewielkiej obróbce sprawiają, iż Sieć posiada porównywalną, a nawet większą wiedzę o jej użytkownikach niż ich najbliżsi. Zaufanie, jakim darzona jest Sieć, której użytkownicy

powierzają nawet bardzo intymne informacje o sobie, wynika z braku wyboru i jednocześnie nadziei, iż moje dane giną wśród milionów innych.

5. Analiza sentymentów

Trafność personalizacji to także wykorzystanie umiejętnej diagnozy emocjonalnych odczuć adresata przekazu. Służą temu analityczne narzędzia mierzące sentyment użytkownika Sieci (*sentiment analysis*) do określonych osób, zdarzeń, informacji, marek towarowych oraz towarów, a nawet wskazywania dni, w których internauci są najbardziej podatni na określone przekazy informacyjne [Pang Bo, Lee Lillian, 2009].

Doświadczenia korzystania z coraz szerszego spektrum danych o użytkownikach Internetu dowodzą, iż personalizacja wprowadza znaczące zmiany w komunikacji maszyna – człowiek. Stwarza to nowe możliwości spolegliwego porozumiewania się Sieci z użytkownikiem, potrzebę uzupełniania wiedzy z tym związanej oraz wagę świadomości zagrożeń, które niesie ze sobą korzystanie z bogactwa informacji o zachowaniach ludzi (także emocjonalnych) przez maszyny.

6. Adaptatywne systemy hipermedialne

Zasygnalizowana problematyka personalizacji, łatwa do wykorzystania w edukacji wspomaganą maszynami, stanowi od dawna przedmiot adaptatywnych systemów hipermedialnych i innych systemów webowych (*adaptive Web systems – AWS*) tworzących klasę systemów adaptacyjnych indywidualnego użytkownika (*user-adaptive software systems*). Owe adaptacyjne systemy posługują się tak zwanym modelem użytkownika (sformalizowanym zapisem jego profilu). Tworzy go zbiór informacji o indywidualnym użytkowniku, który jest podstawą dla funkcjonowania adaptacyjnych systemów, a w efekcie do personalizacji. Ich funkcjonowanie polega m.in. na odmiennym zachowaniu się systemu w odniesieniu do każdego użytkownika. Jeśli na przykład użytkownik wyszukuje w Sieci określoną informację, system jest w stanie odpowiednio do jego profilu wybrać i dokonać priorytetyzacji uzyskanych wyników poszukiwań. Gdy użytkownik przegląda te wyniki, system może w sposób uzasadniony doświadczeniami (na podstawie wyników analizy statystycznej podobnych zdarzeń) manipulować kolejnością sprawdzania przez użytkownika kolejnych wyników (ukrywać, sortować), dostarczając tym samym personalizowaną pomoc podczas odbioru pozyskanych z Sieci informacji. W sytuacji gdy system odszuka wskazaną przez użytkownika stronę, może stosownie do informacji opisujących użytkownika (adaptacyjnie, np. czcionki, kolory) zaprezentować treść tej strony. W celu budowy i aktualizowania na bieżąco modelu użytkownika, a także zarządzania nim, zaawansowane systemy zbierania informacji w sposób ciągły rejestrują wszelkie dostępne w Sieci formalne i analityczne informacje o takim użytkowniku. Mogą one pochodzić z serwisów społecznościowych (np. Nasza Klasa, Grono, Facebook, Twitter), a także być pozyskiwane z mobilnych telefonów (np. o zmianach miejsca pobytu śledzonej osoby) [Brusilovsky, Kobsa, Nejd, 2007; Berkovsky, Kuflik, Ricci, 2008].

7. Nadzieje i obawy

Nadzieje i obawy związane z możliwością wykorzystywania rosnącej liczby informacji o użytkownikach Sieci w szczególności dotyczą młodych ludzi. Z jednej strony, uzyskana

dzięki super personalizacji spolegliwość maszyn daje trudne do wyobrażenia możliwości skutecznego przekazywania informacji, także w procesie nauki. Z drugiej strony, informacja gromadzona w Sieci dostępna jest wielu instytucjom, każda z nich ma inne zasady operowania zasobami informacyjnymi i ich wykorzystywania. Nigdy wcześniej tak wiele informacji o przeciętnym obywatelu nie było tak łatwo dostępne tak wielu osobom.

Literatura

Baker S.: *The Numerati*. New York 2008

Berkovsky S., Kuflik T., Ricci F.: *Mediation of user models for enhanced personalization in recommender systems. User Modeling and User – Adapted Interaction*. „Dordrecht” 2008, vol. 18, iss. 3

Brusilovsky P., Kobsa A., Nejd W.: *The Adaptive Web: Methods and Strategies of Web Personalization*. Berlin–Heidelberg 2007

Chen S. Y., Magoulas G. D.: *Adaptable and Adaptive Hyper-Media Systems*. IRM Press 2005

Gogolek W.: *O pewnym modelu komputeryzacji dydaktyki*. [W:] *Cybernetyka społeczna*. Red. J. Kossecki. Wydaw. WSP w Kielcach, Kielce 1976

Gold H.: *Launching Multicultural Media Campaigns*. The ClickZ Network, ClickZ Experts Agency Media Strategies, Jun 24, 2008

Goldsmith J., Wu T.: *Who Controls the Internet?* Oxford University Press, Oxford 2006

Krug S.: *Nie każ mi myśleć*. Helion, Gliwice 2006

Lindstrom M.: *Buyology: Truth and Lies About Why We Buy*. New York 2008

Pang Bo, Lee Lillian: *Opinion Mining and Sentiment Analysis*. Publishers Inc. Hanover, USA, 2008. United States Gross National Happiness. http://apps.facebook.com/usa_gnh/ [dostęp wrzesień 2009]

Wojciechowski M., Zakrzewicz M.: *Automatyczna personalizacja serwerów WWW z wykorzystaniem metod eksploracji danych*. Wydawnictwa Naukowo-Techniczne, Warszawa-Szczyrk 2002

Yahoo Launches Site Focused on Women. “The Associated Press” 2008, March 31