

Grzegorz CHODAK*

PROPOZYCJA MODELU KLIENTA W SKLEPIE INTERNETOWYM

W artykule przedstawiono podstawowe założenia modelu klienta w sklepie internetowym. W pierwszej części omówiono źródła danych o zachowaniu klienta w sklepie, z uwzględnieniem analizy logów serwera. W dalszej części artykułu zaprezentowano model klienta wraz z przykładem. Omówiono także korzyści płynące z użycia tego modelu do symulacji zachowania klienta w celu optymalizacji zawartości sklepu. Pokazano także wady i ograniczenia proponowanego podejścia. W ostatniej części przedstawiono ważniejsze nie rozwiązane w artykule problemy i kierunki dalszych badań.

WSTĘP

Rozważania nad modelem klienta sklepu internetowego należałoby rozpocząć od zdefiniowania takiego klienta. Przez sklep internetowy autor rozumie wirtualny sklep, umieszczony na stronie www, umożliwiający klientowi składanie zamówienia on-line i dostarczający towar przy wykorzystaniu standardowych metod przesyłania towaru tj. poczty, przedsiębiorstw kurierskich lub, jeśli istnieje taka możliwość, za pomocą sieci komputerowej. Klientem sklepu jest każdy kto „wchodzi” do sklepu internetowego, a więc wysyła zapytanie do serwera, na którym umieszczony jest sklep internetowy. Przedmiotem dalszych rozważań będą sklepy typu B2C (ang. *business-to-consumer*), charakteryzujące się metodą zamawiania opartą na systemie koszykowym (ang. *e-cart system*).

Techniki analizy zachowania klienta sklepu internetowego są dynamicznie rozwijającą się gałęzią wiedzy z pogranicza informatyki i marketingu. Sklepy internetowe umożliwiają śledzenie zachowania klientów od momentu wejścia do sklepu, przez oglądanie towarów i ich wybór aż do opuszczenia sklepu po dokonaniu transakcji zakupu lub rezygnacji. Celem artykułu jest przedstawienie modelu opisującego zachowanie się klienta sklepu internetowego oraz omówienie korzyści oraz problemów jakie wiążą się z praktycznym wykorzystaniem tego modelu. W pierwszej części artykułu przedstawiono

* Instytut Organizacji i Zarządzania Politechniki Wrocławskiej, grzegorz.chodak@pwr.wroc.pl

źródła danych o zachowaniu się klienta w sklepie internetowym, jako podstawową bazę informacyjną dla zaprezentowanego modelu.

1. ŹRÓDŁA DANYCH O ZACHOWANIU KLIENTA W SKLEPIE INTERNETOWYM

Aby możliwe było zbudowanie modelu klienta, konieczne jest zgromadzenie danych na temat jego zachowania w sklepie internetowym. Dane te są gromadzone standardowo w dwóch miejscach: są to logi serwera www oraz bazy danych sklepu.

Handel internetowy ma pewne cechy specyficzne, umożliwiające śledzenie zachowania klienta w sklepie. Rejestrację akcji podejmowanych przez klienta w sklepie internetowym można porównać do nagrywania filmu przy pomocy kamery video podczas robienia zakupów w tradycyjnym supermarkecie. Chodząc po sklepie i wrzucając produkty do koszyka klient supermarketu jest anonimowy, a jego identyfikacja odbywa się przy kasie podczas płacenia kartą kredytową. Podobnie klient sklepu internetowego jest częściowo anonimowy podczas przeglądania zawartości sklepu i wybierania produktów, a jego pełna identyfikacja odbywa się po zalogowaniu się lub rejestracji w sklepie najczęściej „przy kasie”, czyli podczas finalizacji transakcji.

Zachowanie klienta jest rejestrowane zarówno w pliku logów serwera jak i w bazach danych sklepu. Dane do analizy można podzielić więc na dwie grupy: dane transakcyjne (ang. *transaction-based data*) oraz dane związane z konkretnym id klienta (ang. *consumer-based data*) [Theusinger 2000].

Obecnie na rynku dostępnych jest wiele programów do analizy logów serwera takich jak np. Awstats (www.awstats.org), Webalizer (www.webalizer.org), Analog (www.analog.cx), Report-Magic (www.reportmagic.org). Analiza danych zawartych w logu serwera www, określana w literaturze jako *web mining analysis*, jest gałęzią dziedziny wiedzy zwanej eksploracją danych (*data mining*). Analiza logów serwera ma jednak swoją specyfikę, która wynika głównie z charakteru danych i sposobu ich gromadzenia. Zapisy w logach serwera zawierają między innymi adres IP komputera, z którego nastąpiło odwołanie, nazwę pliku, jego wielkość i etykietę czasową. Techniki eksploracji danych wykrywają powtarzające się wzorce w ramach transakcji lub sekwencji transakcji użytkowników. W analizie logów serwera stosuje się takie techniki eksploracji danych, jak odkrywanie częstych ścieżek nawigacji (ang. *path traversal patterns*) oraz techniki ogólnego przeznaczenia tj. odkrywanie reguł asocjacyjnych (ang. *association rules*) i wzorców sekwencji (ang. *sequential patterns*) oraz klasyfikację (ang. *classification*) i grupowanie (ang. *clustering*) [Wojciechowski 1999].

Akcje podejmowane przez klienta mogą być także rejestrowane w specjalnie do tego celu skonstruowanej bazie danych. Standardowo w bazie danych zapisywana jest całość nawigacji dotyczących zamówień oraz wyświetlanych stron. Najczęściej małe i średnie

Propozycja modelu klienta w sklepie internetowym

sklepy wykorzystują relacyjne bazy danych oparte na serwerze MySQL, natomiast oprogramowanie dla dużych sklepów oparte jest na zaawansowanych serwerach baz danych Oracle czy Sybase.

Informacje zawarte w bazach danych powiązanych z analizatorem odwiedzin sklepu umożliwiają uzyskanie informacji na temat:

- oglądalności danego towaru – liczba otwartych stron zawierających informacje o danym towarze,
- liczby „wrzuceń” danego towaru do koszyka – liczba ta nie jest równa liczbie zamówień danego towaru, ponieważ klient może zrezygnować ze sfinalizowania zamówienia, bądź wyrzucić dany produkt z koszyka,
- liczby klientów wchodzących na stronę sklepu i rejestrujących się,
- liczby klientów wchodzących na stronę sklepu i nie rejestrujących się,
- statystyk odwiedzalności sklepu w ujęciu: godzin, dni tygodnia, miesięcy,
- powiązań danego towaru z innymi – zapamiętanie stron, na które przechodzi klient po obejrzeniu towaru,
- czasu przebywania klienta w sklepie,
- średniej liczby produktów oglądanych przez klienta,
- ścieżki określającej kolejność otwieranych przez klienta stron,

Warto również wymienić analizy związane z otoczeniem sklepu w sieci. Mogą one zawierać cenne informacje, np.:

- zapytania o strony sklepu z innych serwerów,
- wywołania banera reklamowego sklepu,
- zliczenie przekierowań z innych stron (np. w ramach programów partnerskich),
- statystykę słów wpisanych w wyszukiwarkach internetowych, po wpisaniu których klient trafił do sklepu (bardzo cenna informacja przy planowaniu kampanii reklamowej),
- statystykę słów wpisanych w wyszukiwarce sklepu,
- statystykę domen umożliwiającą między innymi śledzenie z jakich krajów klienci wchodzili do sklepu (kraj określany jest na podstawie numeru IP, z którego nastąpiło zapytanie o stronę).

Podstawowym celem eksploracji danych pochodzących ze sklepu internetowego jest klasyfikacja klientów oraz dostosowanie zawartości sklepu do oczekiwań poszczególnych klientów (ang. *customize*).

Sklepy internetowe pozwalają w lepszym stopniu na dopasowanie swojej zawartości do oczekiwań klientów niż tradycyjne formy handlu. W literaturze pojawiło się pod koniec lat dziewięćdziesiątych pojęcie Personalized Information Service rozumiane jako serwis internetowy dopasowany do konkretnego klienta lub grupy klientów [Sharma 2001].

Klasyfikacja klientów dokonana na podstawie ich zachowania w sklepie i ich pogrupowanie pozwala również na zaplanowanie odpowiednio dopasowanej akcji marke-

tingowej. Obecnie w sklepach internetowych stosuje się różnorodne techniki marketingowe, wpływające na proces decyzyjny klienta, tj.:

- polecenie produktów, które klient prawdopodobnie chciałby nabyć (ang. *recommendations*),
- wdrożenie systemu promocji związanych z uzależnieniem wielkości zamówienia od kosztu przesyłki towaru,
- utworzenie systemu promocji cenowej wiążącej grupy produktów, które klient chciałby nabyć.

2. MODEL ZACHOWANIA KLIENTA W SKLEPIE INTERNETOWYM

Dostosowanie zawartości sklepu internetowego do oczekiwań klientów może zostać osiągnięta na trzy sposoby: (1) poprzez użycie stereotypów, (2) przez zastosowanie wszelkiego rodzaju kwestionariuszy i ankiet, (3) dzięki zastosowaniu uczących się modeli (ang. *learned model*) [Langley, 2000]. W artykule podjęto próbę prezentacji modelu klienta, który został oparty na następujących założeniach:

- Klient podejmuje decyzję o podjęciu konkretnej akcji na podstawie zawartości strony, na której się znajduje.
- Na stronie można wyodrębnić zmienne, których wartości wpływają na decyzję klienta.
- Znajdując się na danej stronie sklepu, klient ma do wyboru określony zbiór możliwych do podjęcia akcji.
- Istnieje baza danych, na podstawie, której można przeprowadzić wnioskowanie na temat wpływu wartości zmiennych na zachowanie klienta, bądź pewnej klasy klientów.

Model klienta można określić jako zbiór trójek $\langle S_i, F_i, A_i \rangle$, gdzie:

$F_i: S_i \rightarrow A_i$

S_i – zbiór zmiennych, znajdujących się na i -tej stronie,

F_i – funkcja akcji, która przypisuje akcję określonym na i -tej stronie zmiennym.

A_i – zbiór możliwych do podjęcia przez użytkownika akcji na i -tej stronie,

i – znacznik strony www sklepu internetowego.

Ponieważ dla tych samych wartości zmiennych użytkownik może teoretycznie podjąć dwie różne akcje, dlatego zasadnym wydaje się użycie pojęcia **multifunkcji** zamiast funkcji. Z tego typu sytuacją będziemy mieli do czynienia np. w przypadku, gdy klient raz wchodzi do sklepu, aby zapoznać się z ofertą i drugi raz loguje się dokonując zakupu przy nie zmienionych wartościach zmiennych decyzyjnych. Jednak ze względu na problem identyfikacji multifunkcji pierwsze podejście do budowy modelu

ograniczone zostało do założenia, że ten sam zbiór wartości zmiennych decyzyjnych nie może powodować podjęcia przez klienta dwóch różnych akcji.

Każda odwiedzona po raz pierwszy przez klienta strona generować będzie taką trójkę. Każda kolejna wizyta na stronie nie powoduje utworzenia nowej trójki lecz jedynie ma wpływ na identyfikację funkcji akcji.

Maksymalna liczba trójek jest równa liczbie możliwych do wyświetlenia stron. Ponieważ funkcja akcji może zostać zidentyfikowana jedynie w przypadku, gdy klient lub grupa klientów odwiedziła wcześniej daną stronę przynajmniej raz, dlatego liczba trójek zostaje zredukowana do liczby stron, o które nastąpiło zapytanie.

Jednym z zadań jakie można postawić przed modelem klienta jest zidentyfikowanie zależności funkcyjnej między zmiennymi decyzyjnymi zawartymi na danej stronie a podjętą akcją. W przypadku prawidłowego określenia tych zależności można podjąć próbę optymalizacji wartości zmiennych w celu uzyskania jak najkorzystniejszej wartości funkcji, czyli pożądaną akcją.

Proponowany model zachowania klienta w sklepie internetowym można przybliżyć w postaci grafu opisującego możliwe do podjęcia przez niego akcje. Jako węzły grafu można przyjąć strony www, będące pewną instancją sklepu określającą środowisko decyzyjne klienta. O podjętej akcji decyduje funkcja akcji.

2.1. AKCJE PODEJMOWANE PRZEZ KLIENTA

Podejmowane przez klienta akcje można podzielić na trzy grupy:

1. Poszukiwanie towarów, zaznajamianie się z nimi oraz ze środowiskiem sklepu. W związku z powyższym, w danej chwili klient ma do wyboru następujące możliwości:

- przejście do innej strony zawierającej informacje o produktach lub informacje ogólne o sklepie, metodzie zamawiania itp., przez kliknięcie na link (tekstowy lub graficzny),
- nawiązanie kontaktu ze sprzedawcą przez:
 - komunikator internetowy,
 - pocztę elektroniczną,
 - chat,
- znalezienie informacji przy pomocy wbudowanej w architekturę sklepu wyszukiwarki,

2. Podjęcie decyzji o zakupie towaru. Tu klient ma do wyboru:

- wrzucenie produktu do koszyka,
- zalogowanie się (jeżeli procedura zamawiania tego wymaga),
- przejście do strony finalizującej zakup, gdzie klient musi wybrać:
 - formę wysyłki towaru,
 - formę płatności za towar,

3. Rezygnację z zakupów, polegającą na wylogowaniu się ze sklepu, bądź opuszczeniu go przez przejście do innej strony lub zamknięcie przeglądarki internetowej.

W literaturze spotyka się również inne klasyfikacje akcji klienta. Przykładowo Ansari i inni [2000] wyróżnili:

- dodanie lub usunięcie produktu z koszyka,
- rozpoczęcie procesu finalizacji zakupu (ang. *initiate checkout*),
- zatwierdzenie zakupu (ang. *finish checkout*),
- wyszukiwanie,
- rejestracja klienta w sklepie (ang. *register event*).

Każda z w/w akcji powoduje przejście do innej strony www, bądź pozostanie na tej samej stronie. Przykładem akcji pozostawiającej klienta na tej samej stronie (a więc w tym samym środowisku informacyjnym), a jednak w odmiennej sytuacji decyzyjnej, jest wrzucenie produktu do koszyka. Zawartość koszyka dodano więc jako dodatkowy zbiór zmiennych decyzyjnych.

Wybór konkretnej klasyfikacji akcji podejmowanych przez klienta, bądź konkretnego zbioru akcji, do budowy modelu, powinien być podyktowany decyzją przesądzałą o tym, analiza których akcji klienta wydaje się kluczowa i może nieść ze sobą najwięcej interesujących zarząd sklepu informacji.

2.2. ZMIENNE DECYZYJNE MODELU

Środowisko decyzyjne jest w proponowanym modelu stroną www zawierającą określone informacje, które powodują podjęcie przez klienta akcji. Strony www w sklepie internetowym są stronami tworzonymi dynamicznie. Najczęściej stosowanym rozwiązaniem informatycznym, służącym do dynamicznej prezentacji stron www jest wykorzystanie języka PHP połączonego z serwerem baz danych MySQL. Strony stworzone w języku PHP zawierają skrypty pobierające odpowiednie informacje z baz danych. Takie podejście pozwala między innymi na dynamiczne dostosowywanie wyświetlanych treści do podjętych przez klienta akcji.

Wybór argumentów funkcji, które wpływają na podjętą przez konsumenta akcję, zależy od strony, na której znajduje się klient. Proces składania zamówienia powinien być jak najbardziej uproszczony, jednak nie da się uniknąć pobrania od klienta podstawowych informacji dotyczących adresu wysyłki, wyboru formy wysyłki oraz formy płatności. Dlatego standardowo ciąg stron jakie musi odwiedzić klient, wygląda następująco:

1. strony z towarami, pozwalające na dokonanie wyboru zakupów,
2. strona logowania do sklepu,
3. strona wyboru formy wysyłki,
4. strona wyboru formy płatności,
5. strona finalizacji zamówienia.

Propozycja modelu klienta w sklepie internetowym

Klient również może się najpierw zarejestrować, a dopiero później dokonywać wyboru towarów, więc p. 1 można zamienić kolejnością z p. 2.

Jako osobny zbiór stron wymienić należy wszelkie informacyjne strony typu: informacje o przedsiębiorstwie będącym właścicielem sklepu, informacje kontaktowe, informacje dotyczące zwrotów, najczęściej zadawane pytania (FAQ) itp.

W przypadku, gdy klient znajduje się na stronie opisującej towar jako zmienne decyzyjne, można zaproponować:

- cenę,
- recenzje,
- dostępność towaru (czas realizacji zamówienia),
- wartość dodaną, rozumianą jako dostępne dodatkowe informacje o produkcie poza standardowym opisem i zdjęciem.

Na stronie logowania, jako zmienne decyzyjne można zaproponować liczbę obowiązkowych pól, które musi wypełnić klient.

Na stronie wyboru formy wysyłki można zaproponować: zbiór par typu (koszt wysyłki, czas dostarczenia towaru).

Na stronie wyboru formy płatności będzie to zbiór możliwych form płatności.

Jak wcześniej wspomniano, do zmiennych, które wpływają na podjętą przez klienta akcję zaliczyć należy zawartość koszyka. Nie jest to jednak zmienna decyzyjna, gdyż jej wartość zależy od klienta.

Dobór potencjalnych kandydatów na zmienne decyzyjne powinien wynikać z posiadanych danych na temat klienta lub klasy klientów oraz specyfiki sklepu.

2.3. FUNKCJA AKCJI

Najtrudniejszym elementem modelu jest zbiór funkcji akcji powiązanych z konkretnymi stronami. Istotnym problemem jest dobór metody identyfikacji funkcji.

W przypadku, gdy problem dotyczy jednoargumentowej funkcji wystarczające mogą okazać się klasyczne metody identyfikacji postaci funkcji tj. regresja liniowa.

W przypadku wieloargumentowej funkcji, konieczne jest zastosowanie bardziej zaawansowanych metod regresji wielowymiarowej, np. metoda LMS (ang. *least median of squares method*).

Algorytmy genetyczne (GA) umożliwiają identyfikację parametrów funkcji wielu zmiennych, praktycznie bez większych ograniczeń co do liczby parametrów. Jednak GA wymagają podania postaci funkcji, co w przypadku wieloargumentowych funkcji, związanych z charakterystyką towaru przedstawioną na stronie www, może okazać się metodą niewłaściwą.

Jeżeli dokładna postać funkcji akcji nie jest znana, konieczne wydaje się wykorzystanie technik sztucznej inteligencji, takich jak:

- sieci neuronowe,

- programowanie genetyczne.

Wykorzystanie sztucznych sieci neuronowych (ANN) wydaje się rozwiązaniem, które powinno w sposób efektywny pozwolić na określenie wpływu wartości zmiennych decyzyjnych na podejmowane akcje, nawet przy dużej liczbie tych zmiennych. Wykorzystanie ANN nie umożliwia określenia postaci funkcji akcji, pozwala jednak na symulację zachowania klienta przy określonych wartościach zmiennych decyzyjnych. W literaturze często spotyka się wykorzystanie ANN do optymalizacji zawartości stron internetowych. Przykładowo MacDonald i Silver [2001] wykorzystali sieć do identyfikacji kategorii produktów, które należałoby klientowi zasugerować do obejrzenia. Do uczenia sieci użyli wcześniejszej analizy ścieżek kategorii produktów przeglądanych przez klienta.

Programowanie genetyczne (*Genetic Programming* – GP) jest podobnie jak GA techniką optymalizacyjną opartą na mechanizmach ewolucyjnych. GP nie wymaga podania konkretnej postaci funkcji, ponieważ poszczególne osobniki mogą reprezentować dowolne wyrażenie algebraiczne.

Prawidłowa identyfikacja funkcji akcji może okazać się najtrudniejszym zadaniem i nie zawsze będzie możliwa. Jest to najslabszy punkt koncepcji prezentowanego modelu. Szczególnie istotny jest zasób posiadanych danych czyli stanów sklepu, rozumianych jako zbiory różnych wartości argumentów funkcji oraz reakcji konsumenta na wartości zmiennych. Aby możliwa była lepsza identyfikacja funkcji akcji, konieczne może być grupowanie użytkowników w klasy o podobnej charakterystyce, dzięki czemu próbka danych dla poszczególnej klasy będzie większa niż dla pojedynczego klienta.

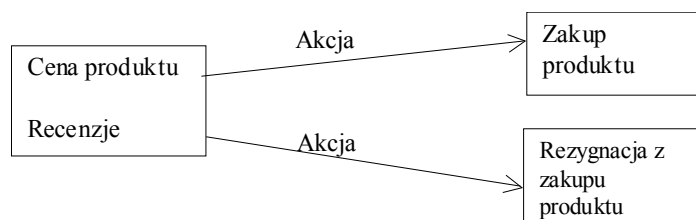
3. PROSTY PRZYKŁAD OBRAZUJĄCY MODEL UŻYTKOWNIKA

Przedstawiony w dalszej części artykułu prosty przykład ma na celu zobrazowanie procesu budowania modelu zachowania klienta oraz wyodrębnienie jego poszczególnych składowych.

Założmy, że informacja o produkcie składa się wyłącznie z ceny i recenzji (dla uproszczenia przykładu pominięty został aspekt koszyka klienta).

Akcja jaką może podjąć klient jest zakup bądź rezygnacja z zakupu (rys. 1).

$Akcja \in \{\text{zakup}, \text{rezygnacja z zakupu}\}$



Propozycja modelu klienta w sklepie internetowym

Rys. 1 Schemat drzewa decyzyjnego obrazującego zachowanie klienta

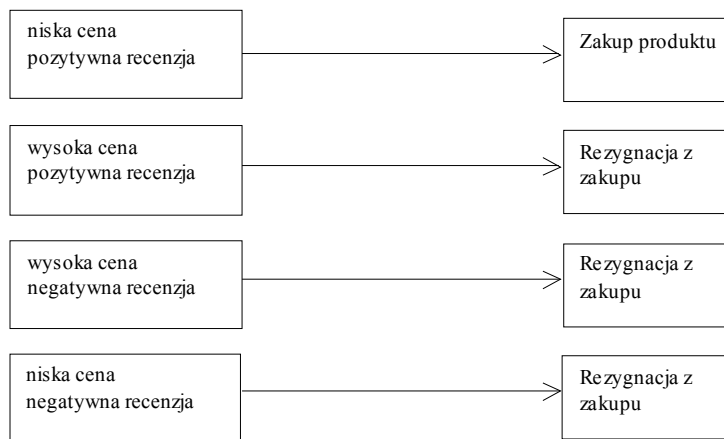
Decyzja o dokonaniu zakupu zostanie podjęta na podstawie wartości dwuargumentowej funkcji ceny i recenzji. Przyjęto uproszczoną charakterystykę argumentów funkcji:

$$cena \in \{niska, wysoka\}$$

$$recenzja \in \{pozytywna, negatywna\}$$

Identyfikacja funkcji może się odbyć na podstawie posiadanych informacji o wcześniejszych akcjach wykonanych przez klienta.

Załóżmy, że wcześniejsze zachowanie klienta opisują grafy skierowane przedstawione na rys. 2 (klient odwiedził jedynie jedną stronę, na której zmieniały się wartości zmiennych decyzyjnych).



Rys. 2. Grafy skierowane obrazujące podejmowane przez klienta decyzje

Budując proste drzewo decyzyjne można ustalić postać funkcji. Opis funkcji przedstawia się następująco:

$$Akcja(cena, recenzja) = Zakup \Leftrightarrow cena = niska \text{ i } recenzja = pozytywna$$

$$Akcja(cena, recenzja) = Rezygnacja \Leftrightarrow cena \neq niska \text{ lub } recenzja \neq pozytywna$$

Tak więc model klienta będzie trójką:

(1) zmienne	$cena \in \{niska, wysoka\}$ $recenzja \in \{pozytywna, negatywna\}$
(2) funkcja akcji	$Akcja(cena, recenzja) = Zakup \Leftrightarrow cena = niska \text{ i } recenzja =$ $pozytywna$ $Akcja(cena, recenzja) = Rezygnacja \Leftrightarrow cena \neq niska \text{ lub } recenzja \neq$ $pozytywna$

(3) wartości funkcji akcji	$Akcja \in \{zakup, rezygnacja\}$
----------------------------	-----------------------------------

Identyfikacja funkcji akcji odbyła się w zaprezentowanym przykładzie z wykorzystaniem prostego drzewa decyzyjnego. Było to możliwe jedynie dzięki przyjęciu bardzo uproszczonych dwuelementowych dziedzin zmiennych decyzyjnych. Do budowy rzeczywistego modelu sklepu należałoby jednak posłużyć się jedną z metod zaproponowanych w podrozdziale 2.3.

4. SYMULACJA ZACHOWANIA KLIENTA

Posiadając informacje na temat wpływu zmiany wartości zmiennych decyzyjnych na zachowanie klienta, można wykorzystać taki model do symulacji zachowania klienta, na zasadzie analizy *what-if*.

Obserwacja reakcji konsumenta na pojedyncze zmiany wartości zmiennych decyzyjnych (*ceteris paribus*) pozwala zaobserwować takie elementy jak:

- elastyczność cenową popytu, w tym reakcję klienta na promocje cenowe,
- wpływ wartości dodanej, liczby recenzji na wielkość sprzedaży,
- reakcję klienta na większą dostępność towaru w magazynie (krótszy czas realizacji zamówienia).

Analiza staje się znacznie ciekawsza, kiedy próbuje się obserwować reakcję klienta na zmianę kilku wartości zmiennych decyzyjnych. Można próbować wtedy odpowiedzieć na następujące pytania:

- Czy lepiej obniżyć cenę produktu, czy koszt przesyłki?
- Czy krótszy czas realizacji zamówienia rekompensuje wyższy koszt wysyłki?
- Czy wzrost liczby recenzji i pojawienie się wartości dodanej może zrekompensować wyższą cenę – określenie w jakim stopniu na popyt ma wpływ cena, a w jakim czynniki pozacenowe?

Znajomość modelu użytkownika pozwala na takie dostosowanie wartości argumentów funkcji akcji, czyli zmiennych decyzyjnych, by w efekcie uzyskać pożądaną akcję. Najczęściej pożądaną akcją będzie nakłonienie klienta sklepu internetowego do dokonania zakupów w sklepie. Należy jednak pamiętać, że celem działalności sklepu internetowego jest maksymalizacja zysku, który to cel nie jest tożsamy z maksymalizacją sprzedaży. Maksimum wielkości sprzedaży osiąga się (zakładając ujemne nachylenie krzywej popytu) określając cenę i koszt przesyłki na poziomie 0 zł. Konieczne wydaje się więc przekształcenie podjętej przez użytkownika akcji na wynik finansowy wynoszący:

- 0 w przypadku, gdy klient zrezygnował z zakupów,
- wartość różnicy między sumą ceny sprzedaży i pobieranej opłaty za przesyłkę, a

Propozycja modelu klienta w sklepie internetowym

sumą wartości towaru i kosztu opłaty za wysyłkę.

Przekształcenie decyzji klienta na wynik finansowy jest konieczne przy poszukiwaniu optymalnej wartości zmiennych decyzyjnych, jednak w samym procesie budowania modelu klienta nie jest to konieczne.

5. SŁABE STRONY MODELU KLIENTA W SKLEPIE INTERNETOWYM

Twierdzenie, że na podstawie wykonanych przez klienta czynności, polegających głównie na przechodzeniu między stronami, administrator sklepu posiada pełną wiedzę o zachowaniu klienta, jest dużym uproszczeniem. Analizuje się jedynie ułamek podejmowanych w danej chwili przez klienta akcji polegających na interakcji z komputerem, pomija się jego rzeczywiste zachowanie w trakcie wizyty w sklepie internetowym (np. równoczesne oglądanie telewizji, spożywanie posiłku i rozmowę z żoną).

Poszukiwanie funkcji odpowiedzialnej za proces decyzyjny klienta zakłada, że decyzja zależy wyłącznie od informacji zawartej na wyświetlanej stronie. Jest to oczywiście pewne uproszczenie, ponieważ decyzja mogła zostać podjęta już przed wejściem klienta do sklepu i w takim przypadku jest całkowicie niezależna od wyświetlanej treści. Decyzja jest także zależna od dokonanej wcześniej akcji. Z taką sytuacją mamy do czynienia, gdy klient nabywa pewne dobro, a następnie poszukuje dóbr komplementarnych (np. zakup drukarki i atramentu, aparatu cyfrowego i dodatkowej pamięci).

Proponowany model zakłada częstą zmianę zawartości stron, która umożliwia mierzenie wpływu zmian na zachowanie klienta. Im częstsze będą zmiany, tym większa jest szansa na poprawną identyfikację funkcji akcji. W przypadku statycznego sklepu, gdzie cena, opis produktu, koszt wysyłki, termin realizacji zamówienia nie ulegają zmianie, proponowany model nie dostarczy żadnych informacji decyzyjnych. Jest rzeczą oczywistą, że nie zmieniając ceny nie jest możliwe odgadnięcie wpływu zmiany ceny na popyt, jedynie na podstawie dotychczasowych zakupów klienta.

6. PROBLEMY DO ROZWIĄZANIA

Zaproponowany model klienta może przynieść wymierne korzyści tylko w przypadku, gdy prawidłowo zostanie zidentyfikowany wpływ wartości zmiennych decyzyjnych na podejmowane akcje. Dlatego niezwykle istotne jest prawidłowe wyróżnienie zmiennych decyzyjnych oraz zbioru akcji, a także metody identyfikacji funkcji akcji.

Osobnym bardzo istotnym problemem jest weryfikacja poprawności modelu. To w tym przypadku bardzo skomplikowane zadanie, ponieważ traktując model jako zbiór wszystkich trójek, związanych z poszczególnymi stronami może dojść do sytuacji, że

dla kilku stron funkcja akcji została zidentyfikowana prawidłowo, dla innych nieprawidłowo. Ponadto wraz z przybywaniem do bazy danych nowych informacji model powinien być cały czas iteracyjnie modyfikowany. Konieczna więc powinna być również iteracyjna weryfikacja modelu. Również sprecyzowanie tego jak określić czy funkcja akcji została zidentyfikowana prawidłowo lub nieprawidłowo pozostaje problemem nierozwiązanym i kierunkiem dalszych badań.

Jak wcześniej wspomniano, aby identyfikacja funkcji akcji mogła zostać przeprowadzona poprawnie, konieczne jest posiadanie jak największej liczby danych obrazujących wpływ konkretnych wartości zmiennych decyzyjnych na podjęte akcje. Niezbędne może okazać się grupowanie użytkowników w klasy o podobnej charakterystyce. Dobór kryteriów podziału oraz metod grupowania również pozostawiam jako kierunek dalszych badań nad modelem.

Istotną zmienną decyzyjną, która wpływa na zachowanie klienta jest liczba pozytywnych oraz negatywnych recenzji produktów. Pojawia się jednak problem w określeniu, które recenzje można uznać za pozytywne oraz w jakim stopniu. Jest to zadanie dla algorytmów specjalizujących się w analizie semantycznej języka naturalnego. Aby uprościć zagadnienie oceny recenzji, niektóre sklepy używają systemu „gwiazdkowego” – im więcej gwiazdek tym, lepszy produkt (opisany przykład można znaleźć w [Kulkarni 2005]). Pomińcie recenzji jako zmiennej decyzyjnej wydaje się nie do przyjęcia, stąd problem wymaga głębszej analizy.

Przedstawiony w artykule model może umożliwić lepsze dopasowanie zawartości sklepu do oczekiwań klienta. Efektywność przedstawionego modelu będzie uzależniona od zastosowania poprawnych algorytmów poszukiwania funkcji akcji, poprawnego grupowania klientów, a także wybrania właściwych zmiennych decyzyjnych i zbioru akcji. Proponowane podejście zawiera słabe punkty, jednak jego istotną zaletą jest prostota i niezależność od konkretnej implementacji informatycznej.

Dalszym etapem rozwoju modelu będzie zastąpienie funkcji akcji multifunkcją akcji, co może się wiązać również ze zmianą algorytmów identyfikacji multifunkcji.

LITERATURA

- ANSARI, S., KOHAVI, R., MASON, L., ZHENG, Z. 2000. *Integrating E-Commerce and Data Mining: Architecture and Challenges*. WEBKDD'2000 workshop: Web Mining for E-Commerce -- Challenges and Opportunities, <http://robotics.Stanford.EDU/~ronnyk/WBKDD2000/index.html>
- KULKARNI, K. 2005. *Why and how can we motivate amazon users to write more product reviews?* Final paper – Assignment 4 SI -688 Fundamentals of Human Behavior ClassID – 3113, <http://www-personal.si.umich.edu/~kirank/docs/688/688finalpaper.pdf> (pobrano 27.06.2005).
- LANGLEY, P. 2000. *User Modelling and Adaptive Interfaces*. Seventeenth National Conference on Artificial Intelligence, Daimler Chrysler Research and Technology Centre. (Za [MacDonald i Silver 2001]).

Propozycja modelu klienta w sklepie internetowym

- LINDEN, G., SMITH, B., YORK, J. 2003. *Amazon.com recommendations. Item-to-item Collaborative Filtering* IEEE Internet Computing, 03/2003.
- MACDONALD, R., SILVER, D. L. 2001. *Web-based User Profiling Using Artificial Neural Networks*, APICS Mathematics/Statistics and Computer Science Joint Conference.
- SHARMA, A. 2001. *A Generic Architecture for User Modeling Systems and Adaptive web services*, Workshop on E-Business & the Intelligent Web, Seattle, USA, August 5 2001.
- THEUSINGER, C., HUBER, K.P. 2000. *Analyzing the footsteps of your customers*, Case study by ASK/net and SAS Institute, Web Mining for E-Commerce - Challenges and Opportunities Workshop, Boston 2000.
- WOJCIECHOWSKI, M. 1999. *Odkrywanie wzorców zachowań użytkowników www*. Materiały konf. POLMAN'99, OWN, Poznań.