

Paweł Ziemba

Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

ZASTOSOWANIE KORELACJI I METODY WIELOKRYTERIALNEGO PODEJMOWANIA DECYZJI PRZY DOBORZE KRYTERIÓW OCENY SERWISÓW INTERNETOWYCH

Streszczenie: W artykule poruszono problem selekcji kluczowych kryteriów oceny jakości serwisów internetowych. Wygenerowano tutaj rankingi kryteriów z wykorzystaniem miar korelacji i metody Promethee GDSS. Rankingi te odzwierciedlają stopień, w jakim każde z kryteriów opisuje ogólną jakość serwisów internetowych. Następnie z każdego z rankingów sekwencyjnie usuwano kolejne kryteria i badano wpływ tego działania na poprawność oceny serwisów. W pracy porównano też wpływ na poprawność oceny serwisów wywierany przez rankingi utworzone z wykorzystaniem miar korelacji oraz metody Promethee GDSS i wpływ na ocenę wywierany przez rankingi utworzone przez badanie ankietowe użytkowników.

Słowa kluczowe: Promethee GDSS, badanie ankietowe, selekcja kryteriów oceny, miary korelacji.

1. Wstęp

Na całym świecie dostępnych jest niemal 285 mln serwisów internetowych [*February 2011... 2011*], przy czym w sieci funkcjonuje już prawie 2 mln domen „.pl” [*Rynek nazw... 2011*]. Przy takiej liczności witryn zdobycie popularności przez nowy serwis jest bardzo trudne. Nowa witryna, obok działań marketingowych podejmowanych w celu zdobycia popularności, powinna również pozytywnie wyróżniać się jakością postrzeganą przez użytkowników. Badanie serwisów funkcjonujących od długiego czasu w Internecie jest stosunkowo łatwe, gdyż można tego dokonywać, bazując na danych ilościowych dostępnych w sieci, np. takich jak liczba użytkowników, liczba odsłon, zasięg itp. Dużo trudniejsza jest ocena serwisów dopiero projektowanych bądź modyfikowanych. Ocena taka musi bazować na zupełnie innych kryteriach, najczęściej o charakterze jakościowym. Istnieje wiele modeli oceny serwisów, które wykorzystują tego typu kryteria, m.in.: eQual [Barnes, Vidgen 2006], metoda Ahn [Ahn 2007], e-SERVQUAL [Parasuraman 2005]. W pracy [Ziemba, Budziński 2010] przedstawiono metodę integracji różnorodnych modeli oceny jakości serwisów internetowych, w wyniku której uzyskano złożoną strukturę wielu

kryteriów oceny. Struktura ta ze względu na licznosc zawartych w niej kryteriów jest bardzo rozbudowana, a w ostatecznej wersji ma ona sluzyc pomocą decydentowi przy doborze kryteriów do oceny serwisów określonego typu. Wobec tego potrzebna jest metoda wiarygodnego selekcjonowania kryteriów, które będą przydatne przy ocenie serwisów wskazanej kategorii. Obecnie trwają prace autora nad ostatecznym określeniem takiej metodyki.

Celem niniejszej publikacji jest zbadanie hipotezy mówiącej o tym, że miary korelacji i wielokryterialna metoda grupowego wspomagania decyzji Promethee GDSS znajdują zastosowanie w zadaniu selekcji użytecznych kryteriów oceny jakości serwisów internetowych. Wyniki selekcji weryfikowano, obliczając stopień poprawnych klasyfikacji wykonanych z wykorzystaniem wyselekcjonowanych podzbiorów kryteriów. Klasyfikatorem zastosowanym w badaniach były drzewa decyzyjne. Badania i ogólną metodykę selekcji oraz redukcji liczby kryteriów przedstawiono na przykładzie kryteriów jednego z najczęściej stosowanych modeli oceny, tj. modelu eQual.

2. Model oceny jakości serwisów internetowych eQual

Narzędziem służącym do zweryfikowania hipotezy badawczej, na którego kryteriach oparto badanie, był model eQual. Jego wybór jest uzasadniony z tego względu, że jest to najlepiej sformalizowany model oceny spośród rozważanych przez autora, a jego wielokrotne stosowanie w literaturze wskazuje na jego wysoką uniwersalność. Był on z powodzeniem użyty m.in. do oceny serwisów: e-commerce [Barnes, Vidgen 2005a, 2005b], e-government [Barnes, Vidgen 2003] i witryn uczelnianych [Barnes, Vidgen 2000].

Tabela 1. Kryteria występujące w modelu eQual

Lp.	Kryterium
1	2
1	Mysle, ze latwo mozna nauczyc sie operowac ta strona
2	Interakcja ze strona jest dla mnie pewna i zrozumiala
3	Mysle, ze strona jest latwa w nawigacji
4	Mysle, ze strona jest latwa w uzyciu
5	Strona ma atrakcyjny wyglad
6	Projekt jest odpowiedni dla typu strony
7	Strona wydaje sie profesjonalnie wykonana
8	Strona wywoluje we mnie pozytywne doznania
9	Dostarcza precyzyjnych informacji
10	Dostarcza wiarygodnych informacji
11	Dostarcza aktualnych informacji

Tabela 1, cd.

1	2
12	Dostarcza trafnych informacji
13	Dostarcza łatwych do zrozumienia informacji
14	Dostarcza informacji odpowiednio szczegółowych
15	Dostarcza informacji w odpowiedniej formie
16	Ma dobrą reputację
17	Czuję, że transakcje przeprowadzane za jej pomocą są bezpieczne
18	Moje osobiste dane przesyłane do tej strony są bezpieczne
19	Wywołuje wrażenie personalizacji
20	Wydaje się, że strona tworzy społeczność wirtualną
21	Komunikacja z twórcami strony jest łatwa
22	Towary/usługi oferowane na stronie zostaną mi dostarczone w przypadku ich zamówienia
23	Ogólna ocena strony

Źródło: [Barnes, Vidgen 2006].

Pierwszą wersję modelu eQual opracowano w 2000 r., a sam model nazywany był WebQual. Został on stworzony na podstawie funkcji dostosowania jakości (*quality function deployment*) dostarczającej metod identyfikacji i wprowadzania opinii użytkowników o jakości produktu na kolejnych etapach jego tworzenia [Barnes, Vidgen 2000]. Oceniający podają tutaj stopień zgodności każdego z kryteriów ze stanem faktycznym, wykorzystując skalę punktową o wartościach z zakresu od 1 do 7. Sama metodologia uzyskiwania opinii o serwisie internetowym, poza uzyskiwaniem wartości poszczególnych kryteriów dla określonej witryny, umożliwia także uzyskiwanie wag, przy czym wagi punktowane są również w zakresie 1-7. Poza punktami przyznanymi dla każdego z kryteriów podawana jest również ogólna ocena witryny [Barnes, Vidgen 2005b]. Poszczególne kryteria wykorzystywane w modelu eQual zawarte są w tab. 1.

3. Współczynniki korelacji

Korelacja jest miarą powiązania pomiędzy dwiema zmiennymi lub większą ich liczbą. Współczynnik korelacji liniowej Pearsona określa stopień liniowych powiązań wartości dwóch zmiennych. Jest to korelacja parametryczna, która nie daje poprawnych wyników przy zależnościach nieliniowych. W takim przypadku wykorzystywane są korelacje nieparametryczne, tj. R Spearmana, tau Kendalla i współczynnik gamma. Współczynnik R Spearmana można traktować podobnie jak współczynnik korelacji liniowej Pearsona, tj. w kategoriach procentu wyjaśnianej zmienności. Współczynnik tau Kendalla określa różnicę między prawdopodobieństwem tego, że dwie zmienne układają się w tym samym porządku w obrębie obserwowanych da-

nych, a prawdopodobieństwem, że ich uporządkowanie się różni. Statystyka gamma z kolei jest zalecana w przypadkach, gdy dane zawierają wiele powiązanych obserwacji. Wyjaśnia się ją jako różnicę między prawdopodobieństwem, że uporządkowanie dwóch zmiennych jest zgodne, a prawdopodobieństwem, że jest niezgodne, podzieloną przez 1 minus prawdopodobieństwo występowania obserwacji powiązanych [Internetowy podręcznik...].

4. Metoda Promethee GDSS

Metoda Promethee GDSS wywodzi się i korzysta bezpośrednio z procedury Promethee II. Metoda Promethee II wykorzystuje porównania parami i relację przewyższania w celu wybrania najlepszej alternatywy decyzyjnej. Wykorzystywane są w niej pozytywne i negatywne przepływy preferencji określające, jak bardzo dany wariant przewyższa inne i jak bardzo jest przewyższany przez inne warianty [Peng 2010]. Procedura Promethee II złożona jest z pięciu kroków:

- 1) porównania parami wariantów decyzyjnych względem kolejnych kryteriów decyzyjnych,
- 2) zastosowania wybranej dla każdego kryterium funkcji preferencji,
- 3) wyznaczenia indeksu preferencji wariantów zgodnie z przyjętymi wagami kryteriów,
- 4) wyznaczenia pozytywnych i negatywnych przepływów preferencji dla wariantów,
- 5) wyznaczenia przepływów preferencji netto [Behzadian 2010].

W metodzie Promethee II decydent może wybierać spośród sześciu funkcji preferencji wykorzystujących: zwykle kryterium, *quasi*-kryterium z progiem równoważności, kryterium z liniową preferencją i progiem preferencji, kryterium poziomu z progiem równoważności i preferencji, kryterium z liniową preferencją i obszarem obojętności, kryterium Gaussa [Brans, Mareschal 2005].

Indeks preferencji wariantów wyznaczany jest zgodnie z wzorem (1):

$$\pi(a_i, b_j) = \frac{\sum_{k=1}^n w_k * \phi_k(a_i, b_j)}{\sum_{k=1}^n w_k}, \quad (1)$$

gdzie j_k oznacza wskaźnik zgodności dla pary wariantów porównywanych względem kryterium k zgodnie z przyjętą funkcją preferencji. Pozytywne i negatywne przepływy preferencji obliczane są z wykorzystaniem wzorów (2) i (3).

$$\varphi^+(a_i) = \sum_{j=1}^n \pi(a_i, b_j), \quad (2)$$

$$\varphi^-(a_i) = \sum_{j=1}^n \pi(b_j, a_i). \quad (3)$$

W końcowym etapie realizacji procedury Promethee II wyznaczany jest całkowity porządek wariantów zgodnie z przepływem preferencji netto opisanym wzorem (4):

$$\varphi(a_i) = \varphi^+(a_i) - \varphi^-(a_i). \quad (4)$$

W metodzie tej wyróżnia się relacje równoważności i preferencji w szerokim sensie:

- 1) wariant a_i przewyższa wariant b_j ($a_i L b_j$), gdy $\varphi(a_i) > \varphi(b_j)$,
- 2) wariant a_i jest równoważny wariantowi b_j ($a_i I b_j$), gdy $\varphi(a_i) = \varphi(b_j)$ [Ghaffghazi 2010].

Procedura Promethee GDSS rozszerza funkcjonalność Promethee II o zagadnienie grupowego podejmowania decyzji. Końcowa agregacja ocen poszczególnych decydentów odbywa się tutaj za pomocą metody Promethee II [Brans 1998]. Procedura GDSS złożona jest z trzech etapów, takich jak:

- 1) wyznaczenie alternatyw decyzyjnych i kryteriów,
- 2) indywidualna ocena wariantów przez każdego decydenta,
- 3) globalna ocena wariantów przez grupę [Escobar-Toledo, Lopez-Garcia 2005].

W pierwszym kroku decydenci powinni wspólnie ustalić rozpatrywane warianty decyzyjne i kryteria, według których alternatywy będą rozpatrywane [Morais, de Almeida 2007]. Drugi etap oceny grupowej polega na tym, że każdy z decydentów indywidualnie ocenia wszystkie warianty zgodnie z wybraną procedurą [Leyva-Lopez, Fernandez-Gonzalez 2003]. W trzecim kroku procedury GDSS wyniki oceny uzyskane przez każdego decydenta są rozpatrywane zgodnie z metodą Promethee II, przy czym kryteria oceny w macierzy decyzyjnej są zastępowane przez decydentów [Morais, de Almeida 2007]. Zaleca się tutaj użycie liniowej funkcji preferencji z progiem preferencji $p = 2$ [Kodikara 2008; Brans, Mareschal 2005].

5. Procedura badawcza

Pierwszym krokiem analizy było zebranie ocen użytkowników dotyczących portali informacyjnych. Ocenianymi przez użytkowników serwisami były Interia.pl, Wp.pl i Onet.pl, a opiniowali oni wskazane portale za pomocą 22 kryteriów eQual z wykorzystaniem skali 7-punktowej. Dodatkowo w celu weryfikacji wyników eksperymentu zebrano od użytkowników szacowane przez nich wagi kryteriów i oceny ogólne przyznawane przez każdego z nich poszczególnym serwisom. Ankiety zostały zebrane wśród studentów i uczniów wyższych klas szkół średnich, wobec czego wyniki badania można interpretować jako preferencje grupy wiekowej 18-24 lata. Łącznie zebrano 100 zestawów ankiet, z których każdy zawierał oceny trzech serwisów oraz wagi kryteriów. Ze względu na to, że 15 ankiet było wypełnionych niepoprawnie (niekompletne ankiety; ankiety, w których podano w każdym polu tę samą wartość i z tego względu uznano je za niewiarygodne), zebrano w sumie 255 opinii, na bazie których przeprowadzono badania.

Celem dalszej pracy było wyznaczenie współczynników korelacji między: ocenami cząstkowymi przyznawanymi przez użytkowników dla portali względem kolejnych kryteriów jakości oraz ocenami końcowymi przyznawanymi dla portali. Bazując na obliczonych współczynnikach korelacji, wyznaczono rankingi kryteriów. Następnie kolejno eliminowano kryteria, które były najmniej skorelowane z ocenami końcowymi. Wpływ usunięcia kolejnych kryteriów na przewidywaną ocenę końcową serwisu analizowano przez klasyfikację opinii użytkowników do jednej z 7 klas oceny końcowej serwisu. W ten sposób starano się uzyskać najmniej liczne zbiory kryteriów pozwalające rozróżnić warianty oceny końcowej, a tym samym zredukować liczby wymiarów oceny i pozostawić tylko kryteria istotne. Jako klasyfikator w tym wypadku wykorzystano drzewa decyzyjne C&RT wykorzystujące do klasyfikacji: miary G-kwadrat oraz szacowane *a priori* prawdopodobieństwa przynależności do określonej klasy decyzyjnej zależnie od liczności przedstawicieli danej klasy w zbiorze treningowym [Waheed 2006]. Dodatkowo, w celu stabilizacji wyników klasyfikacji, wykonywano 10-krotną walidację krzyżową [Weiss, Zhang 2003]. Wykorzystanymi w badaniu współczynnikami korelacji były: korelacja liniowa Pearsona, rang Spearmana, Gamma i tau Kendalla, przy czym dla wyników każdej z miar korelacji zachowano odpowiednie wielkości współczynników istotności. Następnie przeprowadzono procedurę grupowego wspomaganie decyzji Promethee GDSS. Współczynniki każdego typu korelacji zostały tutaj wykorzystane jako cząstkowe opinie decydentów zgodnie z procedurą Promethee GDSS odgrywające rolę kryteriów decyzyjnych, a poszczególne kryteria pełniły funkcję wariantów decyzyjnych.

Ponadto, aby zweryfikować uzyskane wyniki, porównano je z efektami eliminacji kryteriów przebiegającej zgodnie z kolejnością uzyskaną na podstawie wag przyznawanym kryteriom przez użytkowników.

6. Wyniki badań

Przeprowadzone badania statystyczne pozwoliły uzyskać współczynniki korelacji oceny serwisu internetowego względem każdego kryterium z oceną końcową tego serwisu. Z kolei przeprowadzenie zmodyfikowanej procedury Promethee GDSS, w której zamiast wynikami oceny przez ekspertów posłużono się korelacjami, dało w wyniku przepływy preferencji netto.

W tabeli 2 przedstawiono kryteria w kolejności wyznaczonej przez poszczególne współczynniki korelacji, przepływy preferencji netto oraz zgodnie z kolejnością uzyskaną po uśrednieniu wag przyznanych poszczególnym kryteriom przez ankietowanych. Jak wynika z tab. 2, kolejność kryteriów przeznaczonych do eliminacji jest bardzo podobna dla wyników każdej z metod korelacyjnych. Podobnie jest również z kolejnością kryteriów uzyskaną zgodnie z procedurą Promethee GDSS, gdyż generuje ona w tym przypadku pewnego rodzaju uśrednienie wyników każdej z korelacji. Najmniej istotny według wyników badań jest zbiór kryteriów $Z \in \{22, 21\}$, na

kolejnych pozycjach pojawiają się różnice pomiędzy wynikami, a za najważniejsze kryteria w każdym przypadku zostaje uznany zbiór $Z \in \{16, 8\}$. Za najmniej istotne spośród badanych kryteriów powinny więc być uznane zagadnienia: komunikacji z twórcami witryny i handlu elektronicznego, a najbardziej istotnymi są: reputacja serwisu i przyjemność korzystania z witryny. Należy przy tym zauważyć, że ranking kryteriów wygenerowany na podstawie uśrednionych wag przypisywanych kryteriom przez użytkowników nadaje kryteriom zupełnie inną kolejność. Według jawnych opinii ankietowanych, najbardziej istotne są kryteria związane z jakością informacji. Rozbieżności pojawiają się tutaj również przy ocenie ważności: atrakcyjności wyglądu witryny, łatwości operowania stroną i zrozumienia interakcji z serwisem. Najbardziej zauważalną różnicą jest jednak postrzeganie przez użytkowników wagi kryteriów w wyniku przeprowadzonych badań uznanych za najważniejsze, tj.: reputacji serwisu i przyjemności korzystania z witryny. Użytkownicy w swoich jawnych opiniach nie przyznają tym kryteriom szczególnego znaczenia, wobec czego wyraż-

Tabela 2. Współczynniki korelacji, przepływy preferencji netto i średnia wag przyznanych przez użytkowników

Korelacja liniowa Pearsona																						
Kryterium	8	16	5	1	3	18	11	12	17	14	9	4	19									
Wsp. kor.	0,61	0,60	0,55	0,54	0,52	0,52	0,51	0,51	0,51	0,50	0,50	0,49	0,47	0,46	0,45	0,45	0,45	0,34	0,34	0,29	0,26	
Korelacja rang Spearmana																						
Kryterium	8	16	14	18	12	17	1	9	5	3	11	19	4									
Wsp. kor.	0,62	0,58	0,53	0,53	0,51	0,50	0,49	0,49	0,49	0,48	0,48	0,47	0,45	0,44	0,44	0,43	0,40	0,35	0,33	0,29	0,26	
Korelacja gamma																						
Kryterium	8	16	14	1	12	18	11	9	17	5	3	13	4									
Wsp. kor.	0,63	0,62	0,57	0,54	0,53	0,53	0,52	0,51	0,51	0,50	0,50	0,48	0,47	0,47	0,47	0,46	0,45	0,42	0,35	0,35	0,30	0,27
Korelacja tau Kendalla																						
Kryterium	8	16	14	18	12	1	17	9	11	5	3	19	4									
Wsp. kor.	0,52	0,50	0,45	0,43	0,43	0,42	0,41	0,41	0,41	0,41	0,40	0,39	0,37	0,37	0,37	0,36	0,36	0,33	0,29	0,27	0,24	0,22
Przepływy preferencji netto Promethee GDSS																						
Kryterium	8	16	14	18	1	12	5	17	9	3	19	4	13									
ϕ netto	0,0787	0,0676	0,0347	0,0302	0,0283	0,0252	0,0219	0,0202	0,0185	0,0179	0,0152	0,0028	0,0000	-0,0039	-0,0069	-0,0097	-0,0101	-0,0242	-0,0584	-0,0635	-0,0851	-0,0994
Średnia waga przyznanych przez użytkowników																						
Kryterium	11	10	9	3	18	4	12	17	2	13	1	14	22									
Śr. waga	6,49	6,43	6,39	6,16	6,1	6,08	6,05	5,77	5,76	5,72	5,7	5,7	5,48	5,36	5,09	4,98	4,92	4,85	4,83	4,77	3,93	3,39

Źródło: opracowanie własne.

nie widoczna jest tutaj rozbieżność pomiędzy jawnie deklarowanymi preferencjami użytkowników, a tym, czym w rzeczywistości według przeprowadzonych badań kierują się użytkownicy podczas oceny jakości serwisów informacyjnych.

W wyniku przeprowadzonych klasyfikacji na zbiorach kolejno redukowanych kryteriów uzyskano wyniki przedstawione w tab. 3.

Tabela 3. Średnie wyniki klasyfikacji dla wszystkich klas

Liczba kryteriów	Odsetek poprawnych klasyfikacji (średnio)					
	według wag użytkown.	liniowa Pearsona	rang Spearmana	gamma	tau Kendalla	Promethee GDSS
22	73,73	73,73	73,73	73,73	73,73	73,73
21	76,86	73,73	73,73	73,73	73,73	73,73
20	76,86	73,33	73,33	73,33	73,33	73,33
19	73,33	76,86	73,33	73,33	73,33	73,33
18	69,80	76,86	76,86	76,86	76,86	76,86
17	69,80	75,29	75,29	75,29	75,29	75,29
16	69,80	75,29	75,29	75,29	75,29	75,29
15	69,02	75,29	75,29	75,29	75,29	75,29
14	70,20	75,29	74,90	75,29	75,29	75,29
13	68,63	74,90	74,90	75,29	74,90	74,90
12	68,63	75,29	74,90	75,29	74,90	74,90
11	59,61	75,29	75,29	75,29	75,29	75,29
10	59,61	69,41	74,12	75,29	75,29	75,29
9	60,39	67,45	74,12	64,31	64,31	69,41
8	59,22	67,84	64,31	64,31	64,31	68,24
7	63,92	66,67	61,96	62,35	61,96	70,20
6	63,92	66,27	60,39	58,43	58,43	58,43
5	61,96	57,25	59,22	56,86	59,22	57,65
4	56,86	56,08	58,43	54,90	58,43	58,43
3	50,59	59,22	53,73	53,73	53,73	53,73
2	47,06	50,59	50,59	50,59	50,59	50,59
1	40,39	45,10	45,10	45,10	45,10	45,10

Źródło: opracowanie własne.

Dokonując analizy tab. 3, łatwo zauważyć, że redukcja kryteriów według kolejności wyznaczonej na podstawie wag użytkowników daje zdecydowanie najgorsze efekty. Poziom poprawnych klasyfikacji spada tutaj ostatecznie poniżej wartości 70% już dla liczby 13 wykorzystywanych kryteriów i niemal stale utrzymuje się poniżej poziomu uzyskiwanego dla zbiorów kryteriów redukowanych zgodnie z ran-

kingiem uzyskanym innymi metodami. Jeżeli chodzi o redukcję kryteriów zgodnie z kolejnością wyznaczoną przy wykorzystaniu korelacji, najlepsze wyniki uzyskuje się w badanym przypadku dla korelacji rang Spearmana. Pozwala ona utrzymać poziom poprawnych klasyfikacji powyżej 70% nawet dla zbioru 9 wykorzystywanych kryteriów. Z kolei klasyfikacje według rankingów uzyskanych zgodnie z współczynnikami korelacji gamma i tau Kendalla dają bardzo podobne efekty, a spadek współczynnika poprawnych klasyfikacji poniżej 70% ma miejsce przy liczbie 9 kryteriów. Niewiele gorsze efekty uzyskano dla klasyfikacji z redukcją kryteriów zgodną z współczynnikiem korelacji liniowej Pearsona. Z kolei „uśrednienie” rankingów korelacyjnych przez wykorzystanie metody wielokryterialnego wspomaganie decyzji grupowych Promethee GDSS dało zaskakujący efekt. Działanie to pozwoliło utrzymać współczynnik poprawnych klasyfikacji w granicach 70% nawet dla liczby tylko 7 wykorzystywanych kryteriów decyzyjnych. Efekt ten został uzyskany, mimo że w procedurze Promethee GDSS jako dane źródłowe przetwarzane były współczynniki korelacji poszczególnych kryteriów uzyskane dla metod mierzenia korelacji, dla których wyniki poprawnych przyporządkowań spadły poniżej 70% znacznie wcześniej, tj. przy ilościach 10, 8 i 9 wykorzystywanych kryteriów. Można z tego wywnioskować, że pewnego rodzaju „uśrednienie” uzyskanych za pomocą korelacji rankingów może w pewnych sytuacjach dawać pozytywne efekty. Drugim narzucającym się wnioskiem jest fakt mówiący o tym, że nawet niewielka zmiana kolejności eliminowanych kryteriów może znacznie wpływać na „dobroć” uzyskiwanego rozwiązania.

7. Podsumowanie

W artykule poruszono problem wielowymiarowości oceny jakości serwisów internetowych. Przedstawiono tutaj zarys metody redukcji kryteriów oceny uwzględniający formalizację procedur: doboru kryteriów do redukcji oraz weryfikowania uzyskiwanych wyników. Sformułowany zarys metody redukcji wymiarów oceny został zbadany na przykładzie redukcji kryteriów oceny modelu eQual.

Jeżeli chodzi o rankingi kryteriów wygenerowane z wykorzystaniem różnych współczynników korelacji, kolejność kryteriów przeznaczonych do eliminacji jest bardzo podobna dla każdego z rankingów. Podobnie jest również z kolejnością kryteriów uzyskaną zgodnie z procedurą Promethee GDSS, gdyż generuje ona w tym przypadku pewnego rodzaju uśrednienie wyników każdej z korelacji. Z kolei ranking kryteriów wygenerowany na podstawie uśrednionych wag przypisywanych kryteriom przez użytkowników nadaje kryteriom zupełnie inną kolejność. Weryfikacja uzyskanych rankingów z wykorzystaniem klasyfikatora w postaci drzewa decyzyjnego pokazuje, że rankingi wygenerowane z wykorzystaniem metod statystycznych i metody wielokryterialnej analizy decyzyjnej są bardziej precyzyjne od rankingów użytkowników. Kryteria wybrane zgodnie z tymi rankingami lepiej odzwierciedlają ogólne postrzeganie jakości serwisów przez użytkowników. Wyraźnie widoczna jest

wobec tego rozbieżność pomiędzy jawnie deklarowanymi preferencjami użytkowników a tym, czym w rzeczywistości według wyników przeprowadzonych badań kierują się użytkownicy podczas oceny jakości serwisów informacyjnych.

Nawiązując do rankingu uzyskanego za pomocą metody Promethee GDSS, należy zaznaczyć, że w badanym przypadku daje on najlepsze akceptowalne wyniki klasyfikacji dla niewielkiej liczby kryteriów. „Uśrednienie” rankingów korelacyjnych przez wykorzystanie metody Promethee GDSS pozwoliło utrzymać współczynnik poprawnych klasyfikacji w granicach 70% nawet dla liczby tylko 7 wykorzystywanych kryteriów decyzyjnych. Stało się tak, mimo że w procedurze Promethee GDSS jako dane źródłowe przetwarzane były wyznaczone i zbadane wcześniej rankingi korelacyjne, a kolejność kryteriów była tutaj bardzo podobna do kolejności każdego z rankingów korelacyjnych. Świadczy to o tym, że nawet niewielka zmiana kolejności eliminowanych kryteriów może znacznie wpływać na „dobroć” uzyskiwanego rozwiązania.

Literatura

- Ahn T. (red.), *The impact of Web quality and playfulness on user acceptance of online retailing*, „Information & Management” 2007, no 44.
- Barnes S.J., Vidgen R.T., *Data triangulation and web quality metrics: A case study in e-government*, „Information & Management” 2006, no 43.
- Barnes S.J., Vidgen R.T., *The eQual Approach to the Assessment of E-Commerce Quality: A Longitudinal Study of Internet Bookstores*, [w]: *Web Engineering: Principles and Techniques*, red. W. Suh, Idea Group Publishing, 2005a.
- Barnes S.J., Vidgen R.T., *WebQual: An Exploration of Web Site Quality*, Proceedings of the Eighth European Conference on Information Systems, 2000.
- Barnes, S.J., Vidgen, R.T., *Assessing the quality of a cross-national e-government web site: a case study of the forum on strategic management knowledge exchange*, „Industrial Management & Data Systems” 2003, no 103(5).
- Barnes S.J., Vidgen R.T., *Data Triangulation in action: using comment analysis to refine web quality metrics*, Proceedings of the 13th European Conference on Information Systems, 2005b.
- Behzadian M. (red.), *PROMETHEE: A comprehensive literature review on methodologies and applications*, „European Journal of Operational Research” 2010, no 200.
- Brans J.P. (red.), *Combining multicriteria decision aid and system dynamics for the control of socio-economic process. An iterative real-time procedure*, „European Journal of Operational Research” 1998, no 109.
- Brans J-P., Mareschal B., *Promethee Methods*, [w]: *Multiple Criteria Decision Analysis*, red. J. Figueira, Springer, Boston 2005.
- Domagała I. (red.), *Rynek nazw domen Q4 2010. Szczegółowy raport NASK za czwarty kwartał 2010 roku*, NASK, Warszawa 2011.
- Escobar-Toledo C.E., Lopez-Garcia B., *The use of multicriteria decision aid system in the information technology (it) allocation problem*, „Operational Research: An International Journal” 2005, vol. 5, no 2.
- February 2011 Web Server Survey*, <http://news.netcraft.com/archives/2011/02/15/february-2011-web-server-survey.html>.

- Ghafghazi S. (red.), *A multicriteria approach to evaluate district heating system options*, „Applied Energy” 2010, no 87.
- Internetowy podręcznik statystyki*, <http://www.statsoft.pl/textbook/stathome.html>.
- Kodikara P.N., *Multi-Objective Optima Operation of Urban Water Supply Systems*, Victoria University, 2008.
- Leyva-Lopez J.C., Fernandez-Gonzalez E., *A new method for group decision support based on ELECTRE III methodology*, „European Journal of Operational Research” 2003, no 148.
- Morais D.C., de Almeida A.T., *Group decision-making for leakage management strategy of water network*, „Resources, Conservation & Recycling” 2007, no 52.
- Parasuraman A. (red.), *E-S-QUAL a multiple-item scale for assessing electronic service quality*, „Journal of Service Research” 2005, vol. 7, no 10.
- Peng Y. (red.), *User preferences based software defect detection algorithms selection using MCDM*, „Information Sciences” 2010.
- Waheed T. (red.), *Measuring performance in precision agriculture: CART-A decision tree approach*, „Agricultural Water Management” 2006, no 84.
- Weiss S.M., Zhang T., *Performance Analysis and Evaluation*, [w]: *The Handbook of Data Mining*, red. Y. Nong, Lawrence Erlbaum Associates, New Jersey 2003.
- Ziemba P., Budziński R., *Ontologies in Integration of Websites Evaluation Methods*, [w]: *Information Systems in Management VI*, red. P. Jałowiecki, A. Orłowski, WULS Press, Warszawa 2010.

APPLYING CORRELATION AND MULTI-CRITERIA DECISION MAKING METHOD TO SELECT THE WEBSITES ASSESSMENT CRITERIA

Summary: The article deals with the selection of criteria which are essential for the websites quality evaluation. The rankings of criteria were generated using correlations and Promethee GDSS method. These rankings take into account the degree in which every criterion explains the complete quality evaluation of websites. Then criteria were removed from every ranking and the influence of the removal on the correctness of the evaluation was examined. The rankings gained using correlations and Promethee GDSS method were compared with the rankings of criteria formulated by users.

Keywords: Promethee GDSS, surveys, selection of assessment criteria, correlations.