

SELEKCJA KRYTERIÓW OCENY JAKOŚCI SERWISÓW INTERNETOWYCH Z WYKORZYSTANIEM POJEMNOŚCI INFORMACYJNEJ HELLWIGA

Paweł Ziemba

Katedra Inżynierii Systemów Informatycznych
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie
e-mail: pziemba@wi.zut.edu.pl

Ryszard Budziński

Katedra Zarządzania i Inżynierii Wiedzy, Uniwersytet Szczeciński
e-mail: rbudzinski@wi.zut.edu.pl

Streszczenie: Jakość strony internetowej ma istotny wpływ na postrzeganie przez użytkowników serwisu instytucji, do której dana strona należy. Wobec tego istotną kwestią jest ocena jakości serwisów internetowych oraz wybór kryteriów tej oceny. W pracy zaproponowano metodę selekcji istotnych kryteriów oceny jakości serwisów internetowych, opartą na indywidualnym wskaźniku pojemności informacyjnej Hellwiga.

Słowa kluczowe: ocena jakości serwisów internetowych, selekcja kryteriów, pojemność informacyjna Hellwiga

WSTĘP

Na świecie, wg danych z 2012 roku, z niemal 677 mln stron internetowych korzysta 2,267 mld użytkowników [Internet World Stats, 2012] [Netcraft, 2012]. W sytuacji powstawania „społeczeństwa informacyjnego” czerpiącego informacje i „załatwiającego” sprawy urzędowe za pośrednictwem Internetu, jakość strony internetowej i możliwość „załatwienia” za jej pośrednictwem określonej sprawy bądź przeprowadzenia transakcji handlowych ma istotny wpływ na postrzeganie danej instytucji przez użytkowników serwisu i na liczbę użytkowników. Teza ta w różnej formie pojawia się w wielu publikacjach dotyczących oceny serwisów internetowych różnego rodzaju. Hwang stwierdza, że „zła jakość strony może powodować utratę zarówno potencjalnych sprzedaży jak i powtórnych odwiedzin” [Hwang i in., 2011]. Lin w kontekście jakości portali internetowych zaznacza, że

„użytkownicy mają wiele opcji, które potencjalnie redukują lojalność; oni mogą przenieść się gdziekolwiek indziej (do innego portalu) i nie wrócić” [Lin i in., 2005]. Z kolei, wg Grigoroudisa “nowoczesne strony prezentują znaczne zróżnicowanie cech i złożoność struktury. Podobnie jak we wszystkich systemach informacyjnych, ocena jest aspektem ich rozwoju i działaniem przyczyniającym się do maksymalnego wykorzystania zasobów zainwestowanych w witrynę” [Grigoroudis i in., 2008]. Ocena jakości serwisów internetowych, jest więc bardzo ważna. Należy zadbać o to, aby była ona jak najbardziej wiarygodna i w maksymalnym stopniu odzwierciedlała preferencje użytkowników. Jednocześnie powinno się ją uzyskiwać jak najmniejszym nakładem kosztów i pracy.

PROBLEM BADAWCZY

W literaturze wyróżnia się wiele modeli oceny jakości witryn WWW, m.in.: eQual [Barnes i Vidgen, 2006], Ahn, e-Servqual i liczne modele relacyjne, oparte np. na procedurze AHP [Lin, 2010]. Każdy ze stosowanych modeli różni się od pozostałych: wykorzystywanymi kryteriami opisu jakości, ich liczbą, skalą oceniania oraz metodyką pozyskiwania wartości ocen.

Podstawowy problem dotyczący modeli oceny dotyczy pozyskiwania ocen. Do tego celu można się posłużyć oceną ekspercką, ankietą lub śledzeniem śladów użytkowników. Śledzenie śladów użytkowników jest metodą bardzo kosztowną finansowo, czasowo i obliczeniowo. Wymaga to specjalistycznych urządzeń lub programów, śledzących np. ruch gałek ocznych użytkownika lub wskaźnika myszy. Zebranie danych w ten sposób wymaga znacznego czasu, więc dane do analizy pochodzą z bardzo małej próby badawczej. Dodatkowo przetworzenie tych danych wymaga skomplikowanych obliczeń. Jednocześnie stosowalność tej metody jest ograniczona do badania użyteczności serwisów internetowych, a nie ich jakości. Metody ankietowa i ekspercka wykorzystują dużą ilość kryteriów, więc są kosztowne czasowo. Ponadto ankietowani i eksperci często skarżą się właśnie na dużą liczbę kryteriów oceny, która może powodować niedbałość w ocenie i obniżać poprawność ocen.

Podobny problem dotyczy metodyki tworzenia modeli oceny jakości serwisów internetowych. Podczas tworzenia nowych modeli oceny, ich autorzy bazują zazwyczaj na istniejących modelach badania jakości serwisów, wybierając z nich kryteria oceny, które wydają się odpowiednie do zastosowania w ich autorskiej metodzie. Działanie to jest wykonywane przez autorów zazwyczaj w formie niesformalizowanej, bazującej na analizie literatury i własnych przemyśleniach [Lin, 2010]. Jest to podejście eksperckie, rzadko opierane na jasno sformułowanych metodologiach i analizach, a zazwyczaj charakteryzuje się ono pewną chaotycznością. Dodatkowo istnieje ryzyko, że w tworzonych w ten sposób modelach będą funkcjonowały kryteria o niewielkiej rzeczywistej wadze, które nie są istotne dla użytkowników serwisu internetowego. Występuje tutaj pewien rodzaj

tw. „przekleństwa wymiarowości”, polegający na problemie wyboru spośród wielu wskaźników jakości tylko takich, które są przydatne w ocenie jakości serwisów określonego typu.

Niektórzy autorzy, przy wyborze kryteriów stosują bardziej sformalizowane podejście: na podstawie badania ankietowego dokonują oni grupowania kryteriów poprzez metodę statystyczną, tj. analizę czynnikową. Eliminują oni kryteria, które według wyników analizy nie przynależą do żadnej z uzyskanych grup kryteriów, czy też do żadnego z uzyskanych wymiarów jakości [Ho i Lee, 2007][Ou i Sia, 2010][Hwang i in., 2011]. Wykorzystując analizę czynnikową otrzymują oni nowe klasy kryteriów, zawierające pewne podgrupy oryginalnych kryteriów. Podejście takie jest znacznie bardziej uzasadnione od opisanego wcześniej, jednak również tutaj istnieje pewien problem, dotyczący przypisywania wag kryteriom. Ze względu na wykorzystywane w analizie czynnikowej (analiza składowych głównych) rotacje i skalowania zmiennych, trudno jest ustalić, które z kryteriów są istotne. Przekształcenia te, mogą też wprowadzać znaczące zmiany w uzyskiwanych rezultatach [Yang i in., 2011]. W tym podejściu wagi kryteriów są określane na podstawie wyników ankiet, jako uśrednione wartości wag, podanych przez ankietowanych lub ekspertów. Określanie wag kryteriów na podstawie ankiet (tj. jawnie deklarowanych preferencji użytkowników) może generować duże błędy [Zenebe i in., 2010]. Potwierdzają to badania autorskie, z których wynika, że jawnie deklarowane preferencje użytkowników mogą się różnić od kryteriów, które są rzeczywiście wykorzystywane przez użytkowników w ocenie strony internetowej [Ziemia i Budziński, 2011].

W pracy [Ziemia, 2011] sformułowano procedurę selekcji istotnych kryteriów oceny serwisów i nadawania im wag, pozwalającą sformalizować proces doboru kryteriów dla metod oceny jakości serwisów internetowych. Procedura ta wykorzystuje techniki Data Mining, a konkretnie metody filtracyjne służące do selekcji cech, umożliwiające określenie pewnych cech w zbiorze danych, jako istotnych i odrzuceniu cech nadmiarowych. Efektem zastosowania opracowanej procedury jest zmniejszenie ilości kryteriów oceny jakości serwisów internetowych poprzez wybranie najważniejszych z nich. Jednocześnie zaletą wskazanej procedury jest to, że zredukowany za jej pomocą zbiór kryteriów daje wyniki oceny zbliżone do rzeczywistych ocen użytkowników. Wobec tego sformułowana procedura pozwala na określanie niejawnych preferencji użytkowników, którymi w rzeczywistości kierują się oni oceniając serwis, a które mogą różnić się od preferencji jawnie deklarowanych.

W w/w pracy stwierdzono, że spośród badanych metod filtracyjnych, najlepsze rezultaty daje zastosowanie metody symetrycznej niepewności. Metoda ta polega na badaniu swoistej korelacji między kolejnymi cechami (kryteriami), a poszczególnymi klasami obiektów (oceną końcową). Nie uwzględnia ona jednak wzajemnych korelacji między cechami. Może więc wystąpić sytuacja, gdy dwie cechy są silnie skorelowane wzajemnie między sobą oraz zachodzi ich silna korelacja z klasami obiektów. W przypadku zaistnienia takiej sytuacji, jedna z cech

może okazać się niepotrzebna, gdyż obie cechy będą warunkowały przypisanie do tej samej klasy obiektów. Dodatkowo, podczas oceny jakości na podstawie tych dwóch cech (kryteriów), ich współlistnienie w zbiorze kryteriów oceny może zawyżać wynik tejże oceny. Do rozwiązania tego problemu zdecydowano się zaadaptować metodę pojemności informacyjnej Hellwiga.

METODY POJEMNOŚCI INFORMACYJNEJ HELLWIGA

Metoda pojemności informacyjnej Hellwiga wywodzi się z ekonometrii, a polega ona na wyborze takich zmiennych objaśniających, które są silnie skorelowane z zmienną objaśnianą, a słabo skorelowane między sobą. W metodzie Hellwiga rozpatrywane są wszystkie niepuste kombinacje potencjalnych zmiennych objaśniających, tj. $2^m - 1$ kombinacji, gdzie m oznacza ilość zmiennych objaśniających. Pierwszym krokiem metody jest obliczenie macierzy współczynników korelacji liniowej Pearsona między zmiennymi objaśniającymi zawartymi w l -tej kombinacji, oraz wektora współczynników korelacji Pearsona między zmiennymi objaśniającymi i objaśnianymi zawartymi w l -tej kombinacji. Następnie obliczane są indywidualne wskaźniki pojemności informacyjnej zmiennych dla każdej kombinacji zgodnie z wzorem:

$$h_{ij} = \frac{r_j^2}{1 + \sum_{i=1, i \neq j}^m |r_{ij}|} \quad (5)$$

gdzie: h_{ij} – indywidualna pojemność informacyjna j -tej zmiennej w l -tej kombinacji, r_j – współczynnik korelacji j -tej zmiennej objaśniającej ze zmienną objaśnianą, r_{ij} – współczynnik korelacji między i -tą i j -tą zmienną objaśniającą, m – ilość zmiennych objaśniających w l -tej kombinacji. Następnie dla każdej z kombinacji obliczany jest integralny wskaźnik pojemności informacyjnej według wzoru:

$$H_l = \sum_{j=1}^m h_{lj} \quad (6)$$

Do modelu ekonometrycznego zalecane jest wybranie takiej kombinacji zmiennych, dla której uzyskano największą wartość integralnego wskaźnika pojemności informacyjnej [Gnat, 2006].

Jeżeli chodzi o zastosowanie metody Hellwiga, w sytuacji, gdy wykorzystywanych jest wiele cech, istnieje bardzo dużo ich możliwych kombinacji i ich obliczenie jest bardzo czasochłonne (np. dla 22 cech istnieje ponad 4 mln kombinacji). Wobec tego w pracy zdecydowano się zastosować tylko indywidualne pojemności informacyjne obliczane dla jednej kombinacji, zawierającej pełny zbiór cech. Wartości indywidualnych pojemności informacyjnych pozwoliły uzyskać ranking cech, wykorzystany w dalszych badaniach.

PROCEDURA BADAWCZA

Opracowana procedura selekcji kryteriów bazuje na założeniu, że ankietowa ocena wielokryterialna nie jest dokładna. Można tego dowieść przeprowadzając ankietę, w której ankietowani oceniają serwisy pod względem kolejnych kryteriów. Przypisują oni serwisom oceny względem kryteriów oraz oceny ogólne. Ocena jest dokładna, jeżeli oceny serwisów obliczone, jako średnie ważone ocen kryteriów, odpowiadają ocenom ogólnym tych serwisów. Drugie przyjęte założenie mówi, że: istnieje podzbiór kryteriów i ukryte wartości wag, których wykorzystanie w wielokryterialnej ocenie serwisów może dać rozwiązanie zbliżone do wartości oceny ogólnej serwisu określanej przez ankietowanych.

Autorzy sformułowali heurystykę, która mówi, że: *dla poszukiwanego „dobrego” podzbioru kryteriów istnieje model klasyfikacyjny (w sensie uczenia maszynowego), charakteryzujący się niskim stopniem sprzeczności przypadków uczących model.* Wobec tego budując modele klasyfikatorów z wykorzystaniem podzbiorów kryteriów, można wytypować podzbiory, dające rozwiązania bliskie optymalnym. W niniejszej pracy do typowania podzbiorów kryteriów wykorzystano metodę Hellwiga. Opracowaną procedurę przedstawiono wykorzystując model eQual.

Pierwszym etapem badań było zebranie ankiet użytkowników dotyczących oceny jakości trzech największych polskich portali informacyjnych, tj. wp.pl, onet.pl oraz interia.pl zgodnie z modelem eQual. W badaniach wykorzystano 85 ankiet dla każdego portalu, zebrano więc łącznie 255 ankiet. Użytkownicy poza ocenami każdego portalu względem kolejnych kryteriów eQual, podawali w ankietach także indywidualnie przez siebie przypisane wagi każdego kryterium oraz ogólną ocenę każdego serwisu. Kolejnym krokiem procedury było zastosowanie opracowanej metody selekcji cech wykorzystującej pojemność informacyjną.

Następnie walidowano zbudowany ranking poprzez zastosowanie klasyfikatora. Klasyfikator budowany był w wykorzystaniu zbioru kryteriów, z którego iteracyjnie eliminowano najmniej istotne kryterium. Wykorzystanym klasyfikatorem były drzewa decyzyjne CART stosujące do klasyfikacji miarę G-kwadrat oraz szacowanie a-priori prawdopodobieństwa przynależności do określonej klasy decyzyjnej zależnie od liczności przedstawicieli danej klasy w zbiorze treningowym [Rokach i Maimon, 2010][Webb, 2003]. Szacowanie a priori pozwoliło poprawić model klasyfikatora ze względu na to, że liczności poszczególnych klas decyzyjnych były różne. W celu stabilizacji wyników klasyfikacji, za każdym razem wykonywano 10-krotną walidację krzyżową [Rokach i Maimon, 2010]. Wpływ na wybór do klasyfikacji drzew decyzyjnych CART miał fakt, że klasyfikator ten przy pełnym zbiorze kryteriów najlepiej pozwalał uzyskać najlepsze wyniki klasyfikacji. Ponadto stosowane w drzewach decyzyjnych przycinanie pozwala uniknąć przeuczenia klasyfikatora. Minimalną

liczność w węźle drzewa ustalono na 5 przypadków. Podobnie, minimalna liczność węzła potomka, również wynosiła 5 przypadków.

Bazując na wynikach klasyfikacji, wybrano najlepiej rokujące podzbiory kryteriów. Następnie dla każdego serwisu internetowego obliczono jego średnią ocenę ogólną oceny serwisów według wzoru:

$$G_S = \left(\left(\sum_{i=1}^n s_i \right) / 7 \right) / n * 100\% \quad (7)$$

gdzie: s_i - ogólna ocena serwisu przypisana przez i -tego użytkownika (zakres 1-7), n - liczba użytkowników biorących udział w ankiecie. Dla każdego uzyskanego podzbioru kryteriów obliczono też średnie znormalizowane oceny serwisów, wykorzystując oceny kryterialne i wagi zgodnie z wzorem:

$$O_S = \left(\left(\sum_{i=1}^n \sum_{j=1}^m k_{ij} * w_{ij} \right) / \left(\sum_{i=1}^n \sum_{j=1}^m 7 * w_{ij} \right) \right) / n * 100\% \quad (8)$$

gdzie: k_{ij} - ocena serwisu przez i -tego użytkownika dla j -tego kryterium, w_{ij} - waga przypisana przez i -tego użytkownika dla j -tego kryterium. Ponieważ różne ilości kryteriów powodowały otrzymywanie przez serwisy różnej ilości punktów, więc punktacja dla każdego podzbioru kryteriów została znormalizowana do zakresu [0-1]. Jeżeli wybrany podzbiór kryteriów oceny dobrze odzwierciedla jakość serwisu internetowego, to między oceną ogólną serwisu a oceną kryterialną zachodzi zależność $G_S \approx O_S$. Porównanie wzorcowej wartości G_S i wartości O_S uzyskanych dla poszczególnych podzbiorów, pozwoliło wytypować optymalny podzbiór kryteriów służących ocenie serwisów informacyjnych. Porównanie tych wartości wykonano z wykorzystaniem miary bezwzględnego średniego odchylenia [Pham-Gia i Hung, 2001], według wzoru:

$$MAD = \sum_{i=1}^n \|O_i - G_i\| / n \quad (9)$$

gdzie: n - liczba badanych serwisów internetowych, G_i - średnia ocena ogólna i -tego serwisu internetowego, O_i - średnia ocena kryterialna i -tego serwisu internetowego. Optymalny podzbiór kryteriów wybrano bazując na wartości MAD i liczności podzbioru.

WYNIKI BADAŃ

W wyniku zastosowania metody Hellwiga otrzymano ranking kryteriów wraz z wagami. Uzyskany ranking kryteriów wykorzystano do klasyfikacji. Ranking ważności kryteriów (numery poszczególnych kryteriów i ich wagi) oraz uzyskane dla niego średnie i minimalne (dla najgorzej klasyfikowanych klas) wartości poprawnych klasyfikacji zawarto w tabeli 1. Średni poziom poprawnych klasyfikacji (C_{mean}) jest wartością procentową poprawnych przypisań do każdej

z klas decyzyjnych (1-7). Minimalny poziom poprawnych klasyfikacji (C_{\min}) odzwierciedla poziom poprawnych przypisań dla klasy, dla której ten poziom jest najniższy. Wyniki przedstawiono na tle rankingu tworzonego jawnie przez użytkowników oraz rankingu utworzonego z wykorzystaniem symetrycznej niepewności. W tabeli 1 zaznaczono podzbiory kryteriów wyznaczone metodą Hellwiga (Hellwig), uwzględnione w dalszym badaniu. Zaznaczono także podzbiory wyznaczone metodą symetrycznej niepewności (Symm. Uncert.) oraz oparte na jawnych opiniach użytkowników (Users), które uwzględniono w celach porównawczych.

Tabela 1. Istotności i współczynniki poprawnych klasyfikacji dla zbiorów kryteriów

Ilość kryt.		5	6	7	8	9	10	11	12	13	14	15
Hellwig	Num.	1	18	11	3	17	14	12	4	9	13	6
	Waga	0,055	0,053	0,051	0,050	0,050	0,049	0,046	0,045	0,044	0,044	0,043
	C_{\min} [%]	29,41	49,02	52,94	52,94	41,38	31,03	31,03	31,03	31,03	31,03	31,03
	C_{mean} [%]	58,04	68,24	68,63	68,63	72,55	75,29	73,73	72,94	74,90	75,29	75,29
Symm. Uncert.	Num.	5	17	1	11	9	3	6	12	10	4	19
	Waga	0,149	0,144	0,142	0,141	0,136	0,135	0,134	0,133	0,132	0,117	0,117
	C_{\min} [%]	42,86	43,14	37,93	37,93	44,83	34,48	34,48	37,93	57,14	57,14	57,14
	C_{mean} [%]	66,67	69,41	67,84	69,80	74,51	74,90	74,12	74,51	74,51	74,51	74,90
Users	Num.	18	4	12	17	2	13	1	14	22	15	8
	Waga	6,13	6,12	6,04	5,75	5,74	5,71	5,68	5,67	5,46	5,34	5,07
	C_{\min} [%]	34,48	41,67	41,67	41,67	41,38	41,38	41,38	55,17	55,17	55,17	41,18
	C_{mean} [%]	61,96	63,92	63,92	59,22	60,39	59,61	59,61	68,63	68,63	70,20	69,02

Źródło: opracowanie własne

Ostatni etap badania polegał na przypisaniu kryteriom, wag zawartych w tabeli 1 i obliczeniu średniego bezwzględnego odchylenia dla ocen obliczonych z użyciem wybranych podzbiorów kryteriów i ich wag. Oceny serwisów wyznaczone za pomocą kolejnych wybranych podzbiorów kryteriów i wartości średniego bezwzględnego odchylenia zawarte są w tabeli 2.

Analizując tabelę 2, należy zauważyć, że niemal każdy podzbiór kryteriów utworzony na podstawie opinii użytkowników charakteryzuje się wartością MAD większą od pełnego 22-elementowego zbioru kryteriów. Ponadto najniższe wartości MAD uzyskano dla podzbiorów, dla których istotność kryteriów wyznaczano za pomocą metody Hellwiga. Zbiory oparte na symetrycznej niepewności, charakteryzują się wartościami MAD większymi niż te, uzyskane dla zbiorów uzyskanych metodą Hellwiga.

Tabela 2. Bezwzględne średnie odchylenie uzyskane dla wybranych podzbiorów kryteriów

Podzbiór kryteriów	Onet.pl	Wp.pl	Interia.pl	MAD
Ocena ogólna	69,66 %	71,85 %	62,44 %	0,00 %
eQual 22	69,48 %	71,54 %	67,47 %	1,84 %
Users 14	70,87 %	73,00 %	68,62 %	2,84 %
Users 13	70,13 %	72,59 %	67,91 %	2,23 %
Users 12	70,47 %	73,14 %	68,28 %	2,64 %
Users 11	70,74 %	73,52 %	68,53 %	2,95 %
Users 10	69,80 %	72,48 %	67,97 %	2,10 %
Users 9	69,11 %	71,79 %	67,27 %	1,81 %
Users 8	68,43 %	71,11 %	67,06 %	2,20 %
Users 7	68,83 %	71,96 %	67,83 %	2,11 %
Users 6	70,01 %	72,55 %	68,74 %	2,45 %
SU 14	69,23 %	71,58 %	66,87 %	1,71 %
SU 13	68,76 %	71,06 %	66,41 %	1,89 %
SU 12	69,34 %	71,48 %	66,71 %	1,65 %
SU 11	69,97 %	71,75 %	67,08 %	1,68 %
SU 10	69,84 %	71,30 %	66,65 %	1,65 %
SU 9	69,33 %	70,65 %	66,12 %	1,74 %
SU 8	70,11 %	71,53 %	66,33 %	1,55 %
SU 7	69,08 %	70,07 %	64,95 %	1,62 %
SU 6	67,35 %	67,92 %	63,55 %	2,45 %
Hellwig 14	69,23%	71,41%	66,34%	1,59%
Hellwig 13	68,77%	70,92%	65,80%	1,73%
Hellwig 12	69,21%	71,49%	65,91%	1,43%
Hellwig 11	68,63%	70,85%	65,29%	1,63%
Hellwig 10	69,22%	71,07%	65,54%	1,44%
Hellwig 9	69,42%	71,31%	65,56%	1,30%
Hellwig 8	69,86%	72,03%	66,04%	1,33%
Hellwig 7	69,21%	71,30%	65,29%	1,29%
Hellwig 6	67,93%	69,62%	63,60%	1,71%

Źródło: opracowanie własne

Analizując tabelę 2, należy zauważyć, że niemal każdy podzbiór kryteriów utworzony na podstawie opinii użytkowników charakteryzuje się wartością MAD większą od pełnego 22-elementowego zbioru kryteriów. Ponadto najniższe wartości MAD uzyskano dla podzbiorów, dla których istotność kryteriów

wyznaczano za pomocą metody Hellwiga. Zbiory oparte na symetrycznej niepewności, charakteryzują się wartościami MAD większymi niż te, uzyskane dla zbiorów uzyskanych metodą Hellwiga.

PODSUMOWANIE

W artykule poruszono temat oceny jakości serwisów internetowych oraz redukcji kryteriów oceny serwisów i określania ich wag z wykorzystaniem metody pojemności informacyjnej Hellwiga. Wyniki badań porównano z wcześniejszymi rezultatami, uzyskanymi przy zastosowaniu selekcji kryteriów na podstawie opinii użytkowników oraz symetrycznej niepewności.

Opierając się na wynikach badań, w pracy wykazano, że zastosowanie pojemności informacyjnej Hellwiga pozwala uzyskać istotne kryteria oceny jakości serwisów internetowych. Wagi i kryteria wybrane za pomocą metody Hellwiga umożliwiają uzyskanie rozwiązania bliższego oczekiwanemu, niż pełny zbiór kryteriów i wagi przypisane przez użytkowników lub też symetryczna niepewność. Ponadto, jest to metoda dająca pełny ranking kryteriów, więc wykorzystując ją, ekspert zachowuje kontrolę przy wyborze podzbioru kryteriów.

Analizując uzyskane wyniki stwierdzić należy, że zbiory kryteriów uzyskane z wykorzystaniem metody Hellwiga dają lepsze wyniki klasyfikacji od zbiorów kryteriów utworzonych z pomocą użytkowników serwisów internetowych. Można z tego wnioskować, że preferencje użytkowników serwisów WWW są w pewnym stopniu niejawne. Mianowicie jawnie deklarują oni, że istotne dla nich są pewne określone cechy witryn internetowych, ale tak na prawdę przy ocenie jakości serwisów podświadomie kierują się nieco innymi kryteriami niż te, które zadeklarowali jako istotne.

BIBLIOGRAFIA

- Internet World Stats (2012) World Internet Users and Population Stats, <http://www.internetworldstats.com/stats.htm>
- Netcraft (2012) April 2012 Web Server Survey, <http://news.netcraft.com/archives/2012/04/04/april-2012-web-server-survey.html>
- Hwang J., Yoon Y.S., Park N.H. (2011) Structural effects of cognitive and affective responses to web advertisements, website and brand attitudes, and purchase intentions: The case of casual-dining restaurants, *International Journal of Hospitality Management*, No.30, s. 897-907.
- Lin, C.S., Wu, S., Tsai, R.J. (2005) Integrating perceived playfulness into expectation-confirmation model for web portal context. *Information & Management*, No.42, s. 683-693.
- Grigoroudis, E., Litos, C., Moustakis, V.A., Politis, Y., Tsironis, L. (2008) The assessment of user-perceived web quality: Application of a satisfaction benchmarking approach. *European Journal of Operational Research*, No.187, s. 1346-1357.

- Barnes S.J., Vidgen R.T. (2006) Data triangulation and web quality metrics: A case study in e-government, *Information & Management*, No.43, s. 767-777.
- Lin H.F. (2010) An application of fuzzy AHP for evaluating course website quality, *Computers & Education*, No.54, s. 877-888.
- Ho C., Lee Y. (2007) The development of an e-travel service quality scale, *Tourism Management*, No.28, s. 1434-1449.
- Ou C.X., Sia C.L. (2010) Consumer trust and distrust: An issue of website design, *International Journal of Human-Computer Studies*, No.68, s. 913-934.
- Yang Q., Shao J., Scholz M., Plant C. (2011) Feature selection methods for characterizing and classifying adaptive Sustainable Flood Retention Basins, *Water Research*, No.45, s. 993-1004.
- Zenebe A., Zhou L., Norcio A.F. (2010) User preferences discovery using fuzzy models, *Fuzzy Sets and Systems*, No.161, s. 3044-3063.
- Ziemba P., Budziński R. (2011) Dobór kryteriów dla oceny serwisów informacyjnych w portalach internetowych, *Studia i materiały PSZW*, No.37, s. 368-378.
- Ziemba P. (2011) Formal procedure of selecting the significant website assessment criteria based on data mining methods, *Young Scientists Innovations Forum 2011*, ISSN 2082-4831. www.fimb.p.lodz.pl.
- Gnat S. (2006) Przegląd niektórych metod doboru zmiennych do modeli jednorównaniowych, *Prace Katedry Ekonometrii i Statystyki*, No. 17, s. 165-178.
- Rokach L., Maimon O. (2010) Supervised Learning, Classification Trees, [In]: Maimon O., Rokach L. (eds.), *Data Mining and Knowledge Discovery Handbook*, Springer, s. 133-174
- Webb G.I. (2003) Association Rules, [In]: Ye N. (eds.), *The Handbook of Data Mining*, Lawrence Erlbaum Associates, s. 25-40.
- Pham-Gia T., Hung T.L. (2001) The Mean and Median Absolute Deviations, *Mathematical and Computer Modeling*, No. 34, s. 921-936.

CRITERIA SELECTION FOR THE WEBSITES QUALITY ASSESSMENT BASED ON THE INFORMATION CAPACITY OF HELLWIG

Abstract: Website quality is highly significant to the users' perception of an organisation. Therefore, the important issue is to assess the websites quality and selection of criteria for this assessment. The paper proposes a method for selection of important criteria assessing the websites quality, based on the individual indicator of the information capacity of Hellwig and Symmetrical Uncertainty.

Keywords: websites quality assessment, criteria selection, information capacity of Hellwig