

DOI: 10.5604/01.3001.0010.7365

ROZMYTA MAPA KOGNITYWNA JAKO INTELIGENTNY SYSTEM REKOMENDACYJNY ZASOBÓW STRONY INTERNETOWEJ

Aleksander Jastriebow, Łukasz Kubuś, Katarzyna Poczęta

Politechnika Świętokrzyska, Katedra Systemów Informatycznych, Zakład Zastosowań Informatyki

Streszczenie. Artykuł poświęcony jest budowie i analizie inteligentnego systemu rekomendacyjnego zasobów bazującego na rozmytej mapie kognitywnej. Opracowany system pozwala wskazać zasoby strony internetowej, którymi może być zainteresowany potencjalny użytkownik. Zasoby te są określone na podstawie aktywności innych użytkowników serwisu. Bazując na zbiorze anonimowo zebranych danych historycznych opracowano rozmytą mapę kognitywną, której czynniki odpowiadają poszczególnym zasobom strony internetowej. Wagi powiązań między nimi określono na podstawie liczby użytkowników odwiedzających poszczególne zasoby.

Słowa kluczowe: Sztuczna inteligencja, rozmyte mapy kognitywne, system rekomendacyjny

FUZZY COGNITIVE MAP AS AN INTELLIGENT RECOMMENDER SYSTEM OF WEBSITE RESOURCES

Abstract. This paper is devoted to the construction and analysis of the intelligent recommendation system for website resources based on fuzzy cognitive map. The developed system allows to identify resources, which may be interested in a potential user. These resources are determined on the basis of website users activity. Fuzzy cognitive map was developed using the dataset with anonymous collected historical data. The concepts of fuzzy cognitive map are identifiers of resources of website. Weights of the connection between them have been established based on the number of users visiting the resources.

Keywords: Artificial intelligence, fuzzy cognitive maps, recommender systems

Wstęp

Rozmyta mapa kognitywna (ang. *fuzzy cognitive map*, FCM) [6] to narzędzie pozwalające na budowę inteligentnego modelu dla wspomagania procesów decyzyjnych. FCM prezentuje zależności (powiązania) między kluczowymi czynnikami modelowanego zjawiska. Wagi powiązań między czynnikami mogą być reprezentowane przez wartości liczbowe z przedziału $[-1, 1]$ [13], reguły rozmyte [6] lub relacje rozmyte [11].

Niniejsza praca poświęcona jest zastosowaniu FCM do budowy systemu rekomendacyjnego dla serwisu internetowego. Systemy tego typu ułatwiają nawigację w serwisie internetowym poprzez prezentację treści (odnośników do zasobów) powiązanych z wybranym zasobem. Budowę i analizę symulacyjną systemu rekomendacyjnego bazującego na rozmytej mapie kognitywnej przeprowadzono przy pomocy oprogramowania opracowanego w technologii Java. Badania wykonano na podstawie historycznych danych pobranych z Repozytorium Uczenia Maszynowego UCI [15].

W artykule przedstawiono inteligentny system rekomendacyjny opracowany w formie rozmytej mapy kognitywnej. Rozdział pierwszy omawia podstawy teoretyczne FCM. W rozdziale drugim przedstawiono sposób budowy proponowanego systemu rekomendacyjnego, opisano analizowany zbiór danych historycznych oraz zaprezentowano wybrane wyniki analizy symulacyjnej opracowanego systemu. Ostatni rozdział stanowi podsumowanie pracy.

1. Rozmyte mapy kognitywne

Rozmyta mapa kognitywna to graf skierowany typu (1), którego zbiór wierzchołków X odpowiada czynnikom istotnym dla analizowanego zagadnienia. Czynniki te oddziałują na siebie z siłą opisaną przez macierz powiązań między nimi R [6]:

$$\langle X, R \rangle \quad (1)$$

gdzie: $X = [X_1, \dots, X_n]^T$ – wektor wartości czynników mapy, X_i – wartość i -tego czynnika mapy w chwili t , $X_i \in [0, 1]$, $R = \{r_{j,i}\}$ – macierz powiązań pomiędzy czynnikami, $r_{j,i}$ – waga powiązania pomiędzy j -tym i i -tym czynnikiem, $r_{j,i} \in [-1, 1]$, $i, j = 1, \dots, n$, n – liczba czynników.

Wartość czynników w kolejnych dyskretnych chwilach t wyznaczana jest przy pomocy odpowiedniego modelu dynamiki. Dobór modelu dynamiki oraz funkcji stabilizującej zależy od opracowywanego systemu. Parametry te mają duży wpływ na charakter odpowiedzi mapy.

Do budowy systemu został wybrany powszechnie stosowany nieliniowy model dynamiki, opisany zależnością (2) [6]:

$$X_i(t+1) = F \left(X_i(t) + \sum_{\substack{j=1 \\ j \neq i}}^n r_{j,i} \cdot X_j(t) \right) \quad (2)$$

gdzie: t – dyskretny czas, $t = 0, 1, \dots, T$, T – czas symulacji, $i = 1, 2, \dots, n$, n – liczba czynników, F – funkcja stabilizująca.

W pracy wybrano sigmoidalną funkcję stabilizującą opisaną równaniem (3) [1]:

$$F(x) = \frac{1}{1 + e^{-cx}} \quad (3)$$

gdzie: c – parametr, $c > 0$. Użyta funkcja stabilizacji wraz z parametrem $c = 5$ jest powszechnie stosowana [1, 13, 14] oraz działa poprawnie dla opracowanego systemu.

1.1. Geneza

Reprezentacja wiedzy w postaci mapy kognitywnej (ang. *cognitive map*, CM) została zaproponowana przez Roberta Axelroda w 1976 roku [2]. Model ten to graf prosty składający się z krawędzi oraz wierzchołków. Wierzchołki odpowiadają czynnikom (zagadnieniom) związanym z analizowanym problemem, natomiast krawędzie definiują wzajemne relacje (wpływy) między nimi.

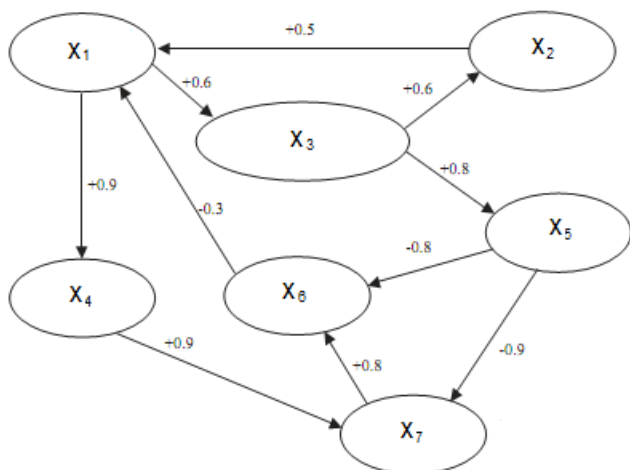
Z każdą krawędzią związany jest znak, który wyraża typ relacji. Jeżeli czynnik A pozytywnie wpływa na czynnik B (od wierzchołka A do wierzchołka B istnieje krawędź dodatnia) oznacza to, że wzrost wartości czynnika A powoduje wzrost wartości czynnika B, natomiast zmniejszenie wartości czynnika A skutkuje zmniejszeniem wartości czynnika B. Krawędź ujemna między dwoma wierzchołkami oznacza wpływ negatywny – wzrost wartości czynnika A powoduje zmniejszenie wartości czynnika B oraz zmniejszenie wartości czynnika A powoduje zwiększenie wartości czynnika B. Pozytywny lub negatywny wpływ czynnika A na czynnik B w mapie kognitywnej jest równoznaczny z takim samym wpływem czynnika B na czynnik A [13].

Przedstawiony sposób reprezentacji wiedzy ma swoje ograniczenia – pozwala jedynie na reprezentację dwóch skrajnych oddziaływań między czynnikami, co w wielu przypadkach jest niewystarczające [13]. Rozmyte mapy kognitywne są rozszerzeniem CM i zostały wprowadzone przez Barta Kosko w 1986 roku [6].

Najważniejszą różnicą między FCM, a CM jest sposób reprezentacji oddziaływania między czynnikami. CM pozwala na wykorzystanie tylko dwóch rodzajów relacji – pozytywnej i negatywnej. FCM umożliwia zdefiniowanie relacji jako wartości liczbowej (wagi). Podobnie jak w przypadku CM, z krawędzią grafu opisującą relację między dwoma wierzchołkami, poza znakiem związana jest również wartość liczbowa (waga). Wagi przyjmują zwykle wartości z przedziału $[-1, 1]$. Wartość relacji między dwoma czynnikami równa -1 oznacza całkowity wpływ negatywny, wartość 0 brak wpływu, a wartość $+1$ całkowity wpływ pozytywny. Wzajemny wpływ między dwoma czynnikami może być różny – wartość wpływu czynnika A na czynnik B nie musi być równa wartości oddziaływania czynnika B na czynnik A [6]. Ponadto, zdefiniowanie relacji między czynnikami w postaci wag pozwala na dokładniejszy opis siły oddziaływania między czynnikami oraz zastosowanie określeń rozmytych takich jak: słaby, średni, silny lub bardzo silny poza określeniami binarnymi (pozytywny lub negatywny) [6].

Opisany model FCM może być zatem przedstawiony w postaci zbioru wierzchołków (istotnych czynników problemu) oraz krawędzi (relacji przyczynowo-skutkowych między nimi) reprezentowanych przez wartości wagowe. Równoważnym modelem dla FCM często wykorzystywanym w celach obliczeniowych jest macierz kwadratowa zwana *macierzą relacji* zawierająca wartości wag dla wszystkich krawędzi grafu [6, 7].

Rys. 1 przedstawia graf reprezentujący przykład modelu FCM, natomiast tab. 1 przedstawia równoważną postać – macierz relacji dla tego samego modelu ($n = 7$).



Rys. 1. Graf reprezentujący model FCM [7]

Tabela 1. Macierz relacji odpowiadająca modelowi FCM przedstawionego w postaci grafu (rys. 1.) [7]

	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇
X ₁	0	0	0,6	0,9	0	0	0
X ₂	0,5	0	0	0	0	0	0
X ₃	0	0,6	0	0	0,8	0	0
X ₄	0	0	0	0	0	0	0,9
X ₅	0	0	0	0	0	-0,8	-0,9
X ₆	0,3	0	0	0	0	0	0
X ₇	0	0	0	0	0	0,8	0

Z FCM związane są również pojęcia takie jak: *gęstość* oraz *rozdzielczość (precyzja) wag*. Pojęcie gęstości oznacza stosunek liczby wag o wartości różnej od zera do liczby wszystkich wag modelu. Najmniejsza różnica między dwoma dowolnymi wartościami wag nazywana jest rozdzielczością lub precyzją wag. Mniejsza wartość precyzji pozwala na budowę dokładniejszych modeli, ale utrudnia proces określenia wartości relacji między czynnikami FCM [13, 14].

1.2. Budowa modelu FCM

Opracowanie modelu rozmytej mapy kognitywnej sprowadza się do wyboru istotnych czynników związanych z modelowanym zagadnieniem oraz określenia relacji między nimi. Najczęściej wybór czynników jest dokonywany przez eksperta lub grupę ekspertów. Kluczowym etapem budowy modelu FCM jest określenie siły wzajemności oddziaływania między czynnikami. Wartości te mogą być ustalone na podstawie wiedzy ekspertów lub w sposób automatyczny przy pomocy zgromadzonych danych historycznych.

Sposób pracy ekspertów jest zależny od opracowywanego modelu. Należy jednak zauważyć, że opracowywanie modeli składających się z dużej liczby czynników jest trudnym zagadnieniem, w szczególności określenie siły relacji między nimi. Z tego powodu prowadzone są badania dotyczące opracowania metod pozwalających w sposób automatyczny zdefiniować relacje między czynnikami – utworzyć macierzy wag FCM [1, 5, 8, 9, 12, 13].

Dane historyczne pozwalają na automatyczne ustalenie wartości relacji między czynnikami. Proces ten nazywany jest uczeniem rozmytych map kognitywnych. Polega on na wcześniejszym opracowaniu zebranych danych historycznych, w sposób pozwalający na ich wykorzystanie przez wybrany algorytm uczenia FCM. Najczęściej dane te muszą zostać znormalizowane oraz w zależności od algorytmu uczenia podzielone na pewne zbiory.

Algorytmy uczenia FCM można podzielić na trzy główne kategorie: nienadzorowane algorytmy oparte o regułę Hebba [11], nadzorowane algorytmy bazujące na metodzie gradientowej [10] oraz algorytmy uczenia inspirowane biologicznie. Do ostatniej kategorii zaliczamy algorytmy ewolucyjne (ang. *evolutionary algorithms*) [8, 13], sztuczne systemy odpornościowe (ang. *artificial immune systems*) [5] oraz metoda roju cząstek (ang. *particle swarm optimization*) [9]. Popularnym algorytmem ewolucyjnym wykorzystywanym jako algorytm uczenia FCM jest zmiennopopulacyjny algorytm genetyczny (ang. *real-coded genetic algorithm*) [3, 4, 8, 13]. Istnieją także rozszerzenia tradycyjnych genetycznych algorytmów uczenia FCM, które pozwalające na wybór tylko najistotniejszych powiązań między czynnikami [14].

Używane są również rozwiązania hybrydowe polegające na wstępnym określeniu macierzy relacji przez eksperta (grupę ekspertów), a następnie zastosowaniu wybranego algorytmu uczenia pozwalającego poprawić dokładność opracowanego modelu. Proces ten jest nazywany *douczeniem*. Może być stosowany w wyżej opisanym przypadku, jak również w sytuacji gdy opracowany model przestał być adekwatny do modelowanego zjawiska lub zostały zebrane nowe dane, które mogą być wykorzystane w procesie uczenia (douczenia) modelu. Pozwala to na aktualizację istniejącego modelu. Ten sposób aktualizacji nie wymaga ponownej analizy wszystkich zgromadzonych danych, a jedynie analizy nowych danych oraz wykorzystanie wcześniej opracowanego modelu.

2. System rekomendacyjny zasobów strony internetowej

Systemy rekomendacyjne ułatwiają nawigację w serwisie internetowym. Pozwalają na dostosowanie prezentowanej treści do odbiorcy lub grupy odbiorców na podstawie analizy historii przeglądanych zasobów przez użytkownika serwisu oraz aktywności wszystkich użytkowników.

2.1. Inteligentny system rekomendacyjny

Celem niniejszej pracy jest opracowanie i analiza inteligentnego systemu rekomendacji zasobów w postaci rozmytej mapy kognitywnej.

Budowę i analizę symulacyjną systemu rekomendacyjnego przeprowadzono na podstawie historycznych danych pobranych z Repozytorium Uczenia Maszynowego UCI [15]. Na podstawie logów internetowego serwisu Microsoft stworzony został zbiór zawierający anonimowe dane 38000 losowo wybranych użytkowników serwisu. Dane każdego użytkownika zawierają informację o odwiedzanych zasobach w okresie jednego tygodnia. Zasoby to obszary serwisu internetowego o wspólnym głównym katalogu wirtualnym obejmujące strony o danej tematyce. Każdy zasób jest opisany poprzez trzy atrybuty takie jak: identyfikator, nazwa oraz adres URL – główny katalog wirtualny. Opis wybranych zasobów zaprezentowano w tab. 2 [15].

Tabela 2. Lista wybranych zasobów wraz z atrybutami.

Identyfikator zasobu	Nazwa zasobu	Adres URL
1000	regwiz	/regwiz
1001	Support Desktop	/support
1002	End User Produced View	/athome
1003	Knowledge Base	/kb
1004	Microsoft.com Search	/search
1005	Norway	/norge
1006	misc	/misc
...
1297	Central America	/centroam

Powyższy zbiór danych pozwala na analizę preferencji użytkowników związaną z wyborem odwiedzanych zasobów. Został on wykorzystany do opracowania inteligentnego systemu rekomendacyjnego bazującego na rozmytej mapie kognitywnej, który na podstawie bieżącej aktywności użytkownika wskazywałby (rekomendował) inne zasoby serwisu.

Analiza oraz budowa systemu rekomendacyjnego jako FCM została wykonana za pomocą autorskiego oprogramowania napisanego w języku Java. Oprogramowanie jest w fazie ciągłego rozwoju, a jego główna koncepcja bazuje na budowie i analizie modeli rozmytych map kognitywnych w oparciu o dane historyczne oraz wiedzę ekspertową. Oprogramowanie pozwala m.in. na inicjalizację czynników mapy na podstawie dostępnych danych oraz określenie siły powiązań między nimi z zastosowaniem różnych metod uczenia FCM.

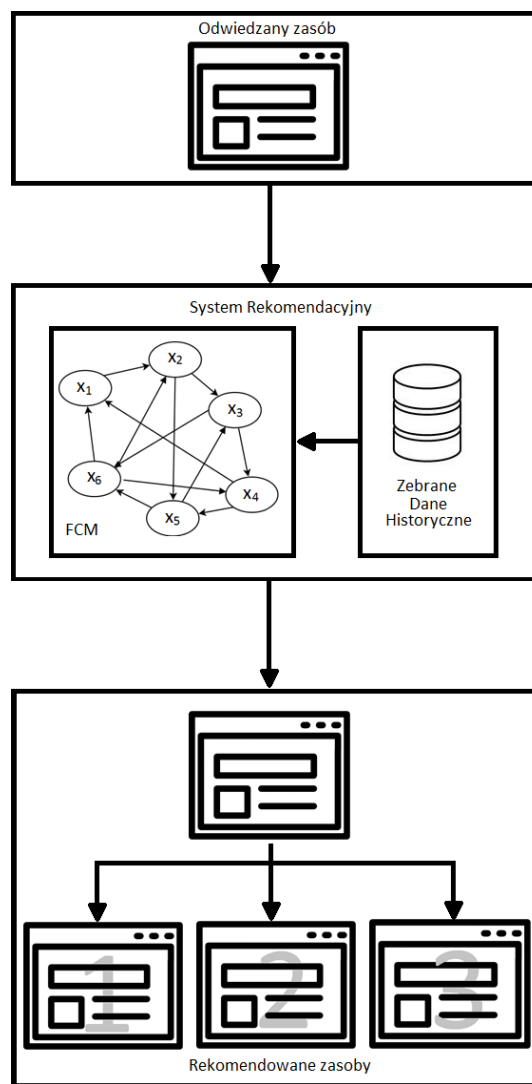
Rys. 2 prezentuje schemat działania inteligentnego systemu rekomendacyjnego.

Danymi wejściowymi dla systemu rekomendacyjnego jest obszar serwisu aktualnie przeglądany przez użytkownika. Odpowiedzią systemu (dane wyjściowe) jest lista zasobów, które mogą interesować użytkownika bardziej niż pozostałe. Źródłem wiedzy dla systemu rekomendacyjnego w formie rozmytej mapy kognitywnej są zebrane dane historyczne. Czynniki FCM odpowiadają zasobom serwisu, a macierz powiązań R reprezentuje zależności między nimi. Wartość powiązania między dwoma zasobami oznacza jak często użytkownik, który odwiedził pierwszy z nich, odwiedzi również drugi. Macierz wag zawiera wartości z przedziału $[-1, 1]$. Wartość -1 oznacza, że żaden użytkownik odwiedzający pierwszy obszar serwisu nie był zainteresowany obszarem drugim. Wartość 1 oznacza, że wszyscy użytkownicy, którzy odwiedzili pierwszy zasób serwisu, odwiedzili również zasób drugi. Wartości z przedziału od 0 do 1 oznaczają proporcję użytkowników, którzy zdecydowali się na odwiedzenie zasobu pierwszego i drugiego w stosunku do wszystkich użytkowników odwiedzających zasób pierwszy.

Dla prezentowanego systemu rekomendacyjnego macierz wag FCM została utworzona w oparciu o analizę danych historycznych. Analiza ta polegała na określeniu ilu użytkowników odwiedziło i -ty i j -ty obszar serwisu.

W pierwszym etapie analizy otrzymano macierz W o wymiarach $n \times n$, gdzie n to liczba zasobów serwisu (czynników zainicjalizowanej mapy). Każdy element macierzy W reprezentuje liczbę użytkowników odwiedzających i -ty i j -ty zasób.

Tab. 3 prezentuje fragment uzyskanej macierzy wraz identyfikatorami zasobów dodanych w pierwszej kolumnie oraz pierwszym wierszu tabeli.



Rys. 2. Schemat działania inteligentnego systemu rekomendacyjnego

Tabela 3. Macierz W określająca liczbę użytkowników odwiedzających zasoby

Identyfikator zasobu	1000	1001	1002	1003	1004	1005	1006	...	1279
1000	912	300	40	156	213	0	20	...	0
1001	300	4451	232	1806	1589	5	35	...	0
1002	40	232	749	127	244	0	11	...	0
1003	156	1806	127	2968	1281	6	11	...	0
1004	213	1589	244	1281	8463	5	42	...	0
1005	0	5	0	6	5	42	0	...	0
1006	20	35	11	11	42	0	135	...	0
...
1297	0	0	0	0	0	0	0	0	0

Następnie, na podstawie liczby użytkowników odwiedzających zasoby (macierz W) określono wagi powiązań między zasobami (macierz R), zgodnie z zależnością (4):

$$r_{i,j} = \begin{cases} -1 & \text{dla } w_{i,j} = 0 \text{ oraz } i \neq j \\ 0 & \text{dla } i = j \\ \frac{w_{i,j}}{w_{i,i}} & \text{dla } w_{i,j} > 0 \text{ oraz } i \neq j \end{cases} \quad (4)$$

gdzie: $W = \{w_{i,j}\}$ – macierz określająca liczbę użytkowników odwiedzających zasoby, $w_{i,j}$ – liczba użytkowników odwiedzających i -ty i j -ty zasób, $w_{i,i}$ – liczba użytkowników odwiedzających i -tą stronę, $w_{i,i} \geq w_{i,j}$ dla $\forall i, j = 1, 2, \dots, n$; $R = \{r_{i,j}\}$ – macierz wag powiązań między zasobami (czynnikami);

$r_{i,j}$ – waga relacji między i -tym i j -tym czynnikiem, $r_{i,j} \in [-1,1]$; $i, j = 1, \dots, n$; n – liczba czynników (zasobów).

Uzyskana macierz wag R została zaprezentowana w tab. 4.

Tabela 4. Macierz wag powiązań między czynnikami

Identyfikator zasobu	(X ₁) 1000	(X ₂) 1001	(X ₃) 1002	(X ₄) 1003	(X ₅) 1004	(X ₆) 1005	(X ₇) 1006	...	(X ₂₉₄) 1279
(X ₁) 1000	0.0	0.33	0.04	0.17	0.23	-1.0	0.02	...	-1.0
(X ₂) 1001	0.07	0.0	0.05	0.41	0.36	0.0	0.01	...	-1.0
(X ₃) 1002	0.05	0.31	0.0	0.17	0.33	-1.0	0.01	...	-1.0
(X ₄) 1003	0.05	0.61	0.04	0.0	0.43	0.0	0.0	...	-1.0
(X ₅) 1004	0.03	0.19	0.03	0.15	0.0	0.0	0.0	...	-1.0
(X ₆) 1005	-1.0	0.12	-1.0	0.14	0.12	0.0	-1.0	...	-1.0
(X ₇) 1006	0.15	0.26	0.08	0.08	0.31	-1.0	0.0	...	-1.0
...
(X ₂₉₄) 1297	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	0

Identyfikatory zasobów oraz czynników FCM (X₁-X₂₉₄) zostały zamieszczone w tabeli w postaci pierwszej kolumny oraz pierwszego wiersza. Wartość wagi równa -1 oznacza brak powiązania między i -tym oraz j -tym zasobem – żaden użytkownik odwiedzający i -ty obszar serwisu internetowego nie odwiedził równocześnie zasobu j -tego.

Utworzona macierz wag R , nieliniowy model dynamiki (2) wraz z sigmoidalną funkcją stabilizującą (3) pozwoliły na budowę inteligentnego systemu rekomendacyjnego opartego o FCM.

Dane wejściowe wczytane z dostępnego zbioru danych przekształcono do n -wymiarowego wektora początkowy V (5):

$$V = [v_1, v_2, \dots, v_n]^T \tag{5}$$

gdzie: V – wektor początkowy (inicjalizujący) wartości czynników FCM, $v_i \in \{0, 1\}$, n – liczba zasobów. Czynnikiem modelu reprezentuje odwiedzenie (wyświetlenie) zasobu serwisu internetowego. Wartość $v_i = 1$ oznacza, że i -ty obszar serwisu został odwiedzony przez użytkownika.

Wartości czynników mapy są przeliczane zgodnie z wybranym modelem dynamiki mapy oraz funkcją stabilizującą (6):

$$X_j(t+1) = F \left(X_j(t) + \sum_{\substack{i=1 \\ i \neq j}}^n r_{i,j} \cdot X_i(t) \right) \tag{6}$$

gdzie: $X_j(t+1)$ – wartość j -tego czynnika w chwili t , $t = 0, 1, \dots, T$, $X_j(t) \in [0, 1]$, $j = 0, 1, \dots, n$, n – liczba czynników, $F(x)$ – funkcja stabilizująca (3), $r_{i,j}$ – wartość wagi dla czynnika X_i i X_j , $X(0) = V$, V – wektor początkowy.

Wektor V' (7) stanowi odpowiedź systemu na wektor wejściowy V :

$$V' = X(T) \tag{7}$$

gdzie: $X(T)$ – wektor odpowiedzi FCM w chwili T .

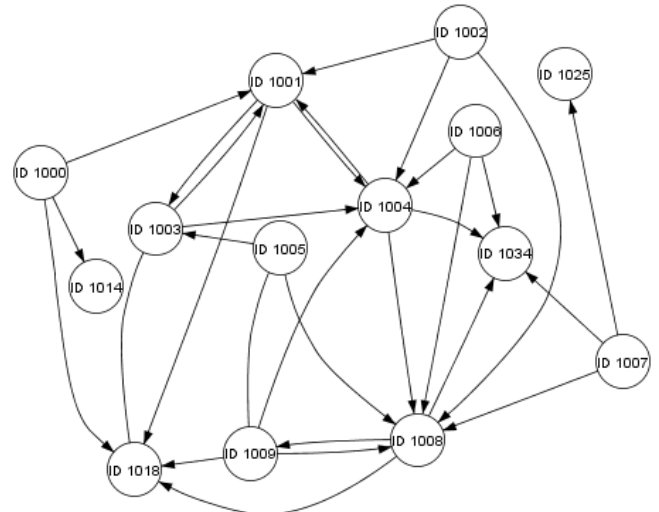
Wartości elementów wektora V' określają powiązanie zasobu podanego w wektorze inicjalizującym V z pozostałymi obszarami serwisu internetowego. Każdy element v'_j reprezentuje wartość powiązania j -tego zasobu serwisu z zasobem i -tym. Im większa wartość v'_j tym j -ty zasób jest mocniej powiązany z zasobem i -tym (podanym w wektorze inicjalizacji V). Oznacza to, że użytkownicy odwiedzający i -ty obszar serwisu, równie często odwiedzają obszar j -ty. Sortując wartości wektora odpowiedzi V' malejąco otrzymamy listę zasobów powiązanych z zasobem głównym.

Tab. 5 przedstawia efekt działania inteligentnego systemu rekomendacyjnego. Dla wybranych obszarów serwisu przedstawiono listę trzech zasobów z nim związanych. W nawiasach podano wartości odpowiedzi FCM (wartości wektora wyjściowego) dla każdego zasobu łącznie z tym, który inicjował odpowiedź mapy.

Rys. 3 obrazuje wynik działania opracowanego systemu dla wybranych zasobów.

Tabela 5. Najczęściej odwiedzane zasoby związane z wybranymi stronami – inteligentny system rekomendacyjny w formie FCM

Identyfikator głównego zasobu	Identyfikator powiązanego zasobu		
	1	2	3
1000(0,993307)	1014(0,885948)	1001(0,838891)	1018(0,802184)
1001(0,993307)	1018(0,90025)	1003(0,885948)	1004(0,858149)
1002(0,993307)	1008(0,858149)	1004(0,838891)	1001(0,824914)
1003(0,993307)	1001(0,954783)	1018(0,930862)	1004(0,895669)
1004(0,993307)	1008(0,759511)	1034(0,740775)	1001(0,721115)
1005(0,993307)	1008(0,785835)	1009(0,785835)	1003(0,668188)
1006(0,993307)	1034(0,90025)	1008(0,885948)	1004(0,824914)
1007(0,993307)	1008(0,912934)	1034(0,759511)	1025(0,622459)
1008(0,993307)	1034(0,920561)	1009(0,768525)	1018(0,75026)
1009(0,993307)	1008(0,939913)	1018(0,832018)	1004(0,802184)



Rys. 3. Graf reprezentujący najczęściej odwiedzane zasoby związane z wybranymi stronami

2.2. Porównawcza analiza statystyczna

W celu oceny efektywności opracowanego systemu rekomendacyjnego zasobów bazującego na rozmytej mapie kognitywnej przeprowadzono analizę statystyczną dostępnych danych. W pierwszym kroku pogrupowano zasoby na grupy zawierające zasób główny oraz zasoby z nim powiązane. Zasoby powiązane oznaczają obszary serwisu, które były wybierane najczęściej przez użytkowników odwiedzających zasób główny.

Dla każdego obszaru serwisu opracowano listę zasobów. Listy te zawierają wszystkie zasoby poza zasobem głównym. Przy każdym zasobie odnotowano ile osób odwiedziło ten zasób, pod warunkiem, że odwiedziło również zasób główny. Powstałe w ten sposób listy zostały posortowane malejąco względem liczby użytkowników, którzy je odwiedzili. Wybierając z posortowanej listy m pierwszych zasobów otrzymujemy liniowo uporządkowaną grupę zasobów powiązanych. Aktualizacja utworzonych list jest ograniczona, sprowadza się do ponownej analizy zebranych danych i utworzeniu nowych list.

Przyjęto, że dla każdego obszaru serwisu zaprezentowane zostaną trzy zasoby powiązane ($m=3$), które mogą być interesujące dla użytkownika odwiedzającego wybrany obszar serwisu. W wierszach tab. 6 znajdują się listy powiązanych obszarów serwisu dla wybranych zasobów głównych.

Pierwsza kolumna tab. 6 zawiera identyfikator obszaru serwisu (zasób główny), a pozostałe kolumny identyfikatory zasobów z nim związanych. W nawiasach podano liczbę użytkowników odwiedzających obszar serwisu. Dla zasobu głównego jest to liczba wszystkich użytkowników odwiedzających ten obszar serwisu, dla zasobów powiązanych jest to liczba użytkowników odwiedzających zarówno zasób główny oraz zasób powiązany.

Tabela 6. Najczęściej odwiedzane zasoby związane z wybranymi stronami – listy zasobów w oparciu o porównawczą analizę statystyczną

Identyfikator głównego zasobu	Identyfikator powiązanego zasobu		
	1	2	3
1000 (912)	1014 (373)	1001 (300)	1018 (254)
1001 (4451)	1018 (1944)	1003 (1806)	1004 (1589)
1002 (749)	1008 (269)	1004 (244)	1001 (232)
1003 (2968)	1001 (1806)	1018 (1534)	1004 (1281)
1004 (8463)	1008 (1977)	1034 (1743)	1017 (1632)
1005 (42)	1008 (11)	1009 (11)	1003 (6)
1006 (135)	1034 (59)	1008 (55)	1004 (42)
1007 (865)	1008 (403)	1034 (203)	1025 (90)
1008 (10836)	1034 (5260)	1009 (2549)	1018 (2390)
1009 (4628)	1008 (2549)	1018 (1472)	1004 (1283)

2.3. Podsumowanie

Porównując dane zawarte w tab. 5 i 6 można stwierdzić, że zarówno statystyczna lista powiązanych zasobów oraz inteligentny system rekomendacyjny w formie FCM wskazują te same najczęściej odwiedzane zasoby (obszary serwisu internetowego).

Ponadto inteligentny system rekomendacji pozwalać na aktualizację list powiązanych zasobów. Może zostać to zrealizowane przez modyfikację macierzy wag na podstawie nowo zebranych danych. Może to zostać zrealizowane poprzez zbudowanie drugiego modelu FCM, a następnie jego połączenie z modelem zbudowanym wcześniej. Połączenie dwóch (lub więcej) modeli FCM wykorzystujących ten sam model dynamiki oraz funkcję stabilizującą sprowadza się do uśrednienia macierzy wag. Dzięki temu nie jest wymagana ponowna analiza wszystkich zebranych danych, co może w znacznym stopniu ułatwić rozbudowę opracowanego systemu rekomendacyjnego zasobów.

3. Wnioski

Opracowany inteligentny system rekomendacyjny umożliwia rekomendację zasobów serwisu internetowego dla użytkownika indywidualnego na podstawie odwiedzanej części serwisu. System został zbudowany w formie rozmytej mapy kognitywnej za pomocą autorskiego oprogramowania opracowanego w języku Java. Implementacja systemu w postaci rozmytej mapy kognitywnej ułatwia jego aktualizację w oparciu o nowo zebrane dane. Przedstawiony został model rozmytej mapy kognitywnej oraz jego własności. Krótko omówiono tradycyjne metody uczenia powszechnie stosowane oraz opisano opracowaną metodę budowy rozmytej mapy kognitywnej bazującą na analizowanym zbiorze danych historycznych.

Zaprezentowano wybrane wyniki działania opracowanego systemu, które potwierdzają dostateczną efektywność zastosowania rozmytej mapy kognitywnej jako systemu rekomendacyjnego zasobów.

Literatura

- [1] Ahmadi S., Alizadeh S., Forouzi N., Yeh C., Martin R. L., Papageorgiou E.: ICLA: Imperialist Competitive Learning Algorithm for Fuzzy Cognitive Map. Proceedings of the 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Beijing, China, 2014.
- [2] Axelrod R.: Structure of Decision: The Cognitive Maps of Political Elites. Princeton University Press, Princeton, New York 1976.
- [3] Froelich W., Juszczyk P.: Predictive Capabilities of Adaptive and Evolutionary Fuzzy Cognitive Maps – A Comparative Study. Nguyen N.T., Szczerbicki E. (eds.): Intel. Sys. for Know. Management, SCI 252, Springer-Verlag, Heidelberg 2009, 153–174.

- [4] Froelich W., Papageorgiou E.I.: Extended Evolutionary Learning of Fuzzy Cognitive Maps for the Prediction of Multivariate Time-Series. Papageorgiou E.I.: Fuzzy Cognitive maps for Applied Sciences and Engineering – From fundamentals to extensions and learning algorithms. Springer, Intelligent Systems Reference Library 54, 2014, 121–131.
- [5] Kannappan A., Papageorgiou E.: A new classification scheme using artificial immune systems learning for fuzzy cognitive mapping. Fuzzy Systems (FUZZ), 2013 IEEE International Conference, 2013, 1–8.
- [6] Kosko B.: Fuzzy cognitive maps. International Journal of Man-Machine Studies 24(1)/1986, 65–75.
- [7] Lee K.C., Lee W.J., Kwon O.B., Han J.H., Yu P.I.: Strategic planning simulation based on fuzzy cognitive map knowledge and differential game. Simulation 71(5)/1998, 316–327.
- [8] Kubuś L., Poczęta K.: Learning Fuzzy Cognitive Maps using Evolutionary Algorithms – a comparative study. Transcom Proceedings 2015 section 3, 9–14.
- [9] Papageorgiou E.I., Parsopoulos K.E., Stylios C.S., Groumos P.P., Vrahtis M.N.: Fuzzy Cognitive Maps Learning Using Particle Swarm Optimization. Journal of Intelligent Information Systems 25(1)/2005, 95–121.
- [10] Poczęta K., Yastrebov A.: Analysis of Fuzzy Cognitive Maps with Multi-Step Learning Algorithms in Valuation of Owner-Occupied Homes. IEEE International Conference on Fuzzy Systems (FUZZIEEE), Beijing, China, 2014, 1029–1035.
- [11] Słoń G.: Application of Models of Relational Fuzzy Cognitive Maps for Prediction of Work of Complex Systems. 13th International Conference ICAISC 2014, Zakopane 2014, 307–318.
- [12] Stach W., Kurgan L., Pedrycz W.: Data-Driven Nonlinear Hebbian Learning Method for Fuzzy Cognitive Maps. IEEE Int. Conf. on Fuzzy Systems (FUZZ-IEEE), 2008, 1975–1981.
- [13] Stach W., Kurgan L., Pedrycz W., Reformart M.: Genetic learning of fuzzy cognitive maps. Fuzzy Sets and Systems 153(3)/2005, 371–401.
- [14] Stach W., Pedrycz W., Kurgan L.A.: Learning of fuzzy cognitive maps using density estimate. IEEE Trans. on Systems, Man, and Cybernetics Part B, 42(3)/2012, 900–912.
- [15] Breese J.S., Heckerman D., Kadie C.M.: Anonymous Microsoft Web Data Set, <http://mlr.cs.umass.edu/ml/datasets/Anonymous+Microsoft+Web+Data>, [16.04.2016]

Prof. dr hab. Aleksander Jastrzebow
e-mail: a.jastrzebow@tu.kielce.pl

Prof. dr hab. Aleksander Jastrzebow jest profesorem zwyczajnym w Katedrze Systemów Informatycznych Politechniki Świętokrzyskiej w Kielcach. Jest autorem oraz współautorem ponad 250 publikacji naukowych. Jego zainteresowania naukowe obejmują diagnostykę, sztuczną inteligencję, automatykę, zastosowania informatyki.



Mgr inż. Łukasz Kubuś
e-mail: l.kubus@tu.kielce.pl

Mgr inż. Łukasz Kubuś jest doktorantem na Politechnice Świętokrzyskiej w Kielcach. Jest autorem oraz współautorem 4 publikacji naukowych. Jego zainteresowania naukowe obejmują sztuczną inteligencję, algorytmy ewolucyjne, rozmyte mapy kognitywne oraz uczenie maszynowe.



Dr inż. Katarzyna Poczęta
e-mail: k.piotrowska@tu.kielce.pl

Dr inż. Katarzyna Poczęta jest adiunktem w Katedrze Systemów Informatycznych Politechniki Świętokrzyskiej w Kielcach. Jest autorem oraz współautorem 32 publikacji naukowych. Jej zainteresowania naukowe obejmują systemy wspomagania decyzyjnego, rozmyte mapy kognitywne, uczenie maszynowe oraz predykcję szeregów czasowych.

