



Grzegorz SŁOŃ

# ANALIZA WYBRANYCH ALGORYTMÓW UCZENIA W RELACYJNEJ ROZMYTEJ MAPIE KOGNITYWNEJ

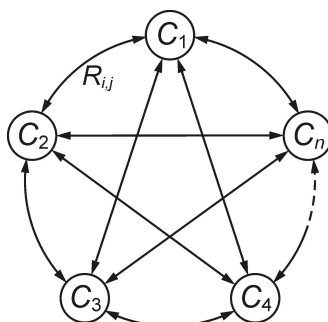
### *Streszczenie*

*Relacyjna rozmyta mapa kognitywna jest pewnym modelem matematycznym wykorzystującym metody inteligencji obliczeniowej, który może być stosowany do modelowania złożonych systemów charakteryzujących się nieprecyzyznością. Kluczowym problemem w tego rodzaju zagadnieniach jest adaptacja parametrów (uczenie) modelu w sposób, dzięki któremu jego praca będzie jak najdokładniej odwzorowywać zachowanie rzeczywistego obiektu. W artykule przedstawiono podstawowe informacje dotyczące problematyki uczenia modelu relacyjnej rozmytej mapy kognitywnej oraz efekty stosowania różnych podejść do uczenia w zależności od celów modelowania.*

### WSTĘP

Potrzeba modelowania pracy złożonych lub nieprecyzyjnych układów jest następstwem nowoczesnego podejścia do projektowania, nadzorowania i predykcji działania wszelkiego rodzaju systemów – ekonomicznych, technicznych, medycznych, społecznych, geograficznych, logistycznych i innych. Podejście to opiera się na ograniczeniu kosztownych badań kolejnych wersji obiektów rzeczywistych na rzecz opracowywania modeli i przeprowadzania badań symulacyjnych na tych modelach. Jest ono istotnie tańsze i bardziej bezpieczne, jednakże nie wszystkie rzeczywiste systemy łatwo się mu poddają. Dotyczy to zwłaszcza systemów złożonych, których struktura i wewnętrzne zależności nie są dobrze znane oraz systemów nieprecyzyjnych, w których określanie wartości poszczególnych parametrów oraz zależności pomiędzy nimi ma charakter subiektywny i zależy głównie od ocen obserwatorów. Prowadzone od lat próby modelowania takich systemów wykorzystują głównie metody inteligencji obliczeniowej, a główne ich trendy skupiają się na sztucznych sieciach neuronowych [1,8] i mapach kognitywnych z dodatkowymi aspektami rozmywania wartości parametrów [4,5,8,11]. Relacyjna rozmyta mapa kognitywna (RRMK) [3,9,10] stanowi rozwinięcie znanej z literatury rozmytej mapy kognitywnej (ang. *Fuzzy Cognitive Map – FCM*), w którym podjęto próbę rozwiązania głównego problemu FCM, jakim jest faktyczne odejście od modelu rozmytego podczas procesu uczenia modelu. Istota budowy modelu RRMK polega na wytypowaniu w badanym systemie pewnych parametrów kluczowych dla celów modelowania (noszą one nazwę czynników) oraz zaprojektowaniu ich wzajemnych oddziaływań (oddziaływania te nazywa się relacjami). Dzięki takiemu ujęciu można zniwelować problemy związane ze zbytnią złożonością lub niepełną wiedzą o obiekcie. Nieprecyzyzność uwzględnia się rozmywając zarówno wartości czynników jak i relacji pomiędzy nimi. W celu uniknięcia nieuprawnionego preferowania znaczenia

któregoś z czynników wszystkie ich wartości normalizuje się (do zakresu  $[-1, 1]$ ), podobnie jak moce relacji. Graficzna reprezentacja takiego modelu (przedstawiona na rys. 1 na przykładzie prostego systemu logistycznego) nie różni się w zasadzie od FCM. Różnica polega na stosowanym podejściu do procesu uczenia modelu.



**Rys. 1.** Ogólna reprezentacja graficzna relacyjnej rozmytej mapy kognitywnej na przykładzie systemu logistycznego, gdzie:  $C_1, \dots, C_n$  – czynniki rozmyte ( $C_1$  – kompetencje kadry zarządzającej,  $C_2$  – uśredniony czas realizacji zamówienia,  $C_3$  – rentowność,  $C_4$  – uśredniony poziom zapasów,  $C_5$  – uśredniona skala popytu);  $R_{i,j}$  – relacja rozmyta pomiędzy czynnikami  $i$ -tym a  $j$ -tym (strzałki pokazują kierunek oddziaływania);  $n$  – liczba czynników.

Czynnikami mogą być zarówno wielkości fizyczne występujące w obiekcie rzeczywistym, jak i abstrakcyjne – o charakterze pomocniczym. Relacje od początku tworzone są jako rozmyte. Główną zaletą tego typu podejścia jest brak konieczności uwzględniania wszystkich występujących w obiekcie wielkości. Model może zawierać niewielką liczbę czynników i ograniczoną liczbę relacji pomiędzy nimi (różną dla różnych potrzeb), a i tak, po odpowiednim nauczeniu, będzie działał z wystarczającą dokładnością. Oczywiście, podczas modelowania nie używa się czynników jako takich, a ich znormalizowanych wartości, które oznacza się jako:  $X_1, \dots, X_n$  i wtedy opis matematyczny modelu może być przedstawiony w następującej postaci:

$$\langle \mathbf{X}, \mathbf{R} \rangle \quad (1)$$

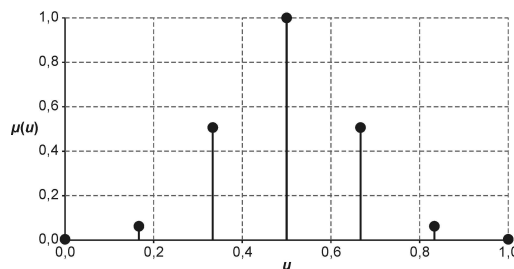
gdzie:  $\mathbf{X} = [X_1, X_2, \dots, X_n]^T$  – wektor wartości rozmytych czynników RRMK;  $\mathbf{R} = \{R_{i,j}\}$ , ( $i, j = 1, \dots, n$ ) – macierz relacji rozmytych pomiędzy czynnikami;  $n$  – liczba czynników RRMK.

W większości stosowanych obecnie rozwiązań FCM (m. in. [4,7,11]) trudności obliczeniowe występujące w zbiorach rozmytych omija się zastępując, na czas uczenia modelu, rozmyte wartości czynników i relacji liczbami rzeczywistymi. Tworzy się  $n$  wymiarowy wektor stanu, którego współrzędne (w postaci liczb rzeczywistych) reprezentują chwilowe wartości lingwistyczne poszczególnych czynników oraz  $n$ -wymiarową macierz kwadratową, której elementy (również o postaci liczb rzeczywistych) reprezentują relacje pomiędzy czynnikami, po czym przeprowadza uczenie na takim modelu. Na czas tego procesu działanie FCM, które początkowo definiuje się poprzez zestaw reguł typu IF-THEN, opisuje się za pomocą układu równań algebraicznych operujących liczbami rzeczywistymi. Takie podejście istotnie upraszcza procedurę uczenia, ale stwarza jednocześnie poważny problem, ponieważ traci się w ten sposób korzyści płynące ze stosowania lingwistycznych wartości opisujących właściwości i pracę obiektu. Podejście wykorzystywane w RRMK zakłada, że model jest rozmyty na wszystkich etapach: projektowania, uczenia i funkcjonowania. Aby to uzyskać nadaje się czynnikom formę liczb rozmytych [6], o odpowiednich funkcjach przynależności np. jak w (2):

$$\mu_{X_i}(u) = e^{-\left(\frac{u-\bar{X}_i}{\sigma_i}\right)^2} \quad (2)$$

gdzie:  $X_i$  – rozmyta wartość  $i$ -tego czynnika;  $u = \{u_1, \dots, u_k\}$  – nośnik (wspólny dla wszystkich czynników);  $k$  – liczba punktów próbkowania nośnika (ekwiwalent liczby wartości lingwistycznych);  $\bar{X}_i$  – centrum  $i$ -tego czynnika;  $\sigma_i$  – współczynnik rozmycia  $i$ -tego czynnika;  $i = 1, \dots, n$ ;  $n$  – liczba czynników.

Funkcja  $\mu(u)$  z równania (2), nazywana funkcją przynależności czynnika rozmytego, określa charakter czynnika, jednakże należy pamiętać, że w rzeczywistości wartość rozmyta czynnika ma dyskretną postać zbioru singletonów rozmytych, i może być przedstawiona graficznie jak na rys. 2.



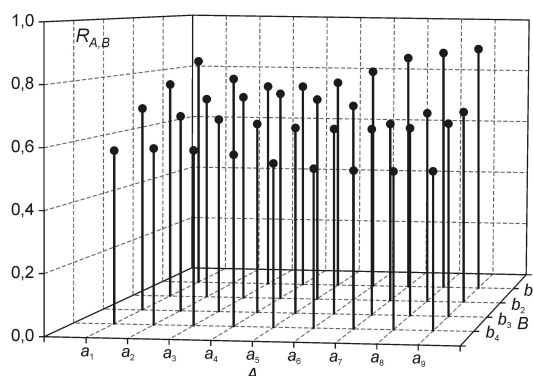
**Rys. 2.** Reprezentacja graficzna rozmytej wartości czynnika o funkcji przynależności zgodnej z (2).  $\bar{X} = 0,5$ ;  $\sigma = 0,2$ ; nośnik  $u$  o zakresie  $[0, 1]$ . Czarne punkty oznaczają położenie kolejnych singletonów rozmytych.

W modelu RRMK zależności pomiędzy czynnikami definiuje się (zamiast wielkościami lingwistycznymi) za pomocą relacji rozmytych [3,9,10], które mogą być obliczane jako zależności funkcyjne (np. z zastosowaniem funkcji przynależności (3)) lub inne.

$$\mu_{R_{i,j}}(u_i, u_j) = e^{-\left(\frac{u_j - r_{i,j}(u_i)}{\sigma_{i,j}}\right)^2} \quad (3)$$

gdzie:  $i, j$  – numery czynników powiązanych relacją;  $u_i, u_j$  – nośniki czynników powiązanych relacją;  $r_{i,j}(u_i)$  – współczynnik mocy relacji rozmytej pomiędzy czynnikami  $i$ -tym i  $j$ -tym (w podstawowej postaci wielkość ta ma postać funkcji liniowej ze współczynnikiem kierunkowym  $\bar{r}_{i,j}$ ;  $\sigma_{i,j}$  – współczynnik rozmycia relacji rozmytej pomiędzy czynnikami  $i$ -tym i  $j$ -tym;  $i = 1, \dots, n$ ;  $n$  – liczba czynników.

Tu również, podobnie jak w przypadku rozmytych wartości czynników, charakter relacji jest dyskretny, co pokazano na rys. 3, który przedstawia graficzną reprezentację przykładowej relacji rozmytej.



**Rys. 3.** Graficzna reprezentacja przykładowej relacji rozmytej  $R_{A,B}$  pomiędzy dwoma zbiorami  $A = \{a_1, \dots, a_9\}$  i  $B = \{b_1, \dots, b_4\}$ . Czarne punkty oznaczają moce powiązań pomiędzy poszczególnymi elementami zbiorów.

Punkty próbkowania nośników wartości czynników i relacji są, w pewnym sensie, odpowiednikami klasycznych wartości lingwistycznych.

Opis matematyczny dynamiki modelu RRMK również ma charakter dyskretny (z uwagi na rodzaj stosowanego nośnika rozmywania oraz wykorzystanie czasu dyskretnego). Pokazane wyżej podejście do budowy głównych elementów modelu powoduje, że do opisu jego dynamiki można wykorzystać arytmetykę liczb rozmytych [7], ze szczególnym uwzględnieniem maksyminowych (dyskretnych) operatorów arytmetycznych. Można stosować różne modele formalne, np. taki jak (4) [...], w którym wartości czynników w następnym kroku czasu dyskretnego są wyznaczana na podstawie ich bieżących przyrostów [3,9]:

$$X_j(t+1) = X_j(t) \oplus \bigoplus_{\substack{i=1 \\ i \neq j}}^n [(X_i(t) \ominus X_i(t-1)) \circ R_{i,j}] \quad (4)$$

gdzie:  $X_j(t-1)$ ,  $X_j(t)$ ,  $X_j(t+1)$  – rozmyte wartości  $j$ -tego czynnika w trzech kolejnych krokach czasu dyskretnego;  $R_{i,j}$  – relacja rozmyta pomiędzy czynnikami o numerach  $i$  oraz  $j$ ;  $t$  – czas dyskretny;  $n$  – liczba czynników;  $\oplus$ ,  $\ominus$ ,  $\circ$  – operatory arytmetyki rozmytej – odpowiednio: dodawania rozmytego, odejmowania rozmytego i maksyminowej kompozycji rozmytej.

Praktyczne stosowanie zaprezentowanego powyżej podejścia RRMK do modelowania złożonych systemów, wiąże się z pokonaniem dodatkowych trudności technicznych i arytmetycznych, jednakże pozwala w pełni wykorzystywać zalety rozmywania. W kolejnych rozdziałach przedstawiono rozważania na temat kilku kluczowych aspektów uczenia modelu RRMK zilustrowane prezentacją efektów wykorzystania kilku różnych algorytmów uczących.

## 1. GŁÓWNE ASPEKTY UCZENIA NADZOROWANEGO RRMK

Model (niezależnie od metody, którą go stworzono ani od sposobu działania) można uznać za poprawny jeżeli odwzorowuje pracę modelowanego systemu z wystarczającą dokładnością. Dokładność tę można szacować porównując wyniki działania modelu z wynikami działania modelowanego obiektu w porównywalnych warunkach (np. przy takich samych pobudzeniach wybranych czynników). Jeśli dokładność jest niższa od oczekiwanej, należy dokonać adaptacji parametrów modelu tak, aby tę dokładność poprawić. Proces taki nosi nazwę uczenia nadzorowanego. W przypadku RRMK zmienianymi parametrami mogą być właściwie tylko relacje rozmyte, przy czym trudno jest stosować analityczne metody optymalizacji, ponieważ relacje te mają dyskretny charakter i, dodatkowo, liczba punktów

próbkowania nośnika (liczba wartości lingwistycznych) również może podlegać modyfikacjom (z zastrzeżeniem, że musi ona pozostawać wspólna dla wszystkich komponentów modelu). Pewnym ułatwieniem może być stosowanie wyłącznie relacji wyznaczanych w oparciu o konkretne typy funkcji, takie jak np. (3), dzięki czemu można zawęzić liczbę modyfikowanych współczynników każdej relacji (jak to pokazuje (5), gdzie elementami macierzy relacji rozmytych  $\mathbf{R}$  są funkcje dwóch adaptowalnych parametrów), jednakże takie uproszczenie wciąż nie rozwiązuje problemu dyskretnego charakteru modyfikowanych wielkości.

$$\mathbf{R} = \begin{bmatrix} 0 & g(r_{1,2}, \sigma_{1,2}) & \cdots & g(r_{1,n}, \sigma_{1,n}) \\ g(r_{2,1}, \sigma_{2,1}) & 0 & \cdots & g(r_{2,n}, \sigma_{2,n}) \\ \vdots & \vdots & \ddots & \vdots \\ g(r_{n,1}, \sigma_{n,1}) & g(r_{n,2}, \sigma_{n,2}) & \cdots & 0 \end{bmatrix} \quad (5)$$

gdzie:  $\mathbf{R}$  – macierz relacji rozmytych RRMK;  $g$  – funkcja typu (3) definiująca zależność kształtu relacji rozmytej od dwóch adaptowalnych argumentów;  $n$  – liczba czynników.

W procesie uczenia można modyfikować model tak, aby służył on rozwiązywaniu różnych zadań: identyfikacji, prognozowania, sterowania lub filtracji. W niniejszej pracy skupiono się na identyfikacji, przy której proces uczenia sprowadza się w zasadzie do minimalizacji pewnej funkcji kryterialnej. W metodach uczenia nadzorowanego funkcja taka przeważnie wyznacza skumulowaną różnicę pomiędzy wartościami oczekiwanymi, a uzyskanymi w drodze symulacji modelowej. Może ona mieć następującą postać:

$$J(Q) = \Phi(\|\bar{X}_i(t) - Z_i(t)\|) \Rightarrow \min_Q \quad (6)$$

gdzie:  $J$  – kryterialna funkcja bliskości;  $\Phi$  – wybrana funkcja optymalizacji;  $\bar{X}_i$  – wyostrzona wartość  $i$ -tego czynnika;  $Z_i$  – zadana (w postaci liczby rzeczywistej) znormalizowana wartość  $i$ -tego czynnika;  $\|\cdot\|$  – wybrana norma;  $t$  – czas dyskretny;  $Q$  – wektor adaptowalnych parametrów zdefiniowany następująco:

$$Q = [\{\bar{r}_{i,j}\}, \{\sigma_{i,j}\}, k]^T \quad (7)$$

gdzie:  $\bar{r}_{i,j}$  – współczynnik kierunkowy funkcji mocy relacji rozmytej pomiędzy czynnikami  $i$ -tym a  $j$ -tym (jak w (3));  $\sigma_{i,j}$  – współczynnik rozmycia relacji rozmytej pomiędzy czynnikami  $i$ -tym i  $j$ -tym;  $k$  – liczba punktów próbkowania nośnika.

Funkcja  $\Phi$ , występująca w równaniu (6) może przybierać różne formy, zależnie od celów stawianych modelowi. Zasadniczo, model dynamiczny taki jak (4) może być użyty zarówno do odwzorowywania stanów stabilnych jak i dynamiki systemu (rozumianej jako proces dochodzenia do stanu stabilności po uprzednim pobudzeniu wybranych czynników). Dodatkowo, obserwacji może podlegać jeden (wybrany) lub więcej wyróżnionych czynników. Warunkuje to wybór formy funkcji  $\Phi$ .

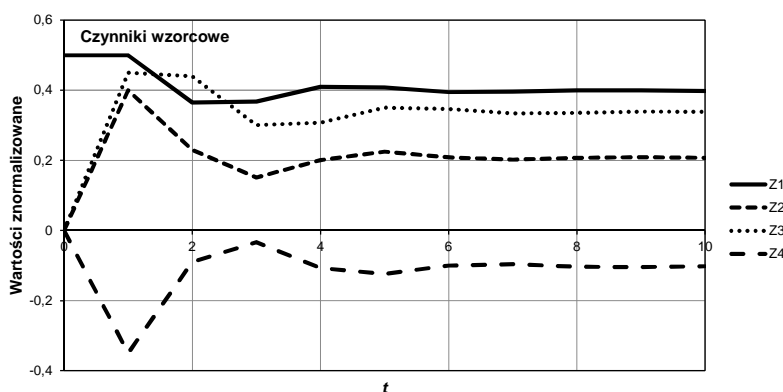
Funkcja bliskości  $J$  z równania (6) pozwala ocenić dopasowanie modelu do wymagań, nie określa ona jednak sposobu, w jaki mają być zmieniane parametry tego modelu w procesie uczenia, a ma to kluczowe znaczenie. Jak wspomniano wyżej, opracowanie analitycznej metody uczenia modelu RRMK jest trudne z uwagi na jego dyskretny charakter. W związku z tym na potrzeby uczenia można wykorzystać (z pewnymi modyfikacjami) jedną z wielu istniejących metod populacyjnych. Metody te, w skrócie, polegają na dokonywaniu drobnych zmian poszczególnych parametrów modelu i ocenianiu wpływu tych zmian na wartość

funkcji bliskości. Różnice pomiędzy poszczególnymi metodami skupiają się na dwóch zagadnieniach: sposobie dokonywania zmian oraz sposobie oceny przydatności dokonanej zmiany. Najprostszą z tych metod jest metoda różnicowa, w której moc i współczynnik rozmycia każdej relacji są zwiększane, a jeśli to nie przyniesie poprawy – zmniejszane o określoną niewielką wartość i proces ten powtarza się kolejno dla każdej relacji, badając pracę modelu po każdej takiej zmianie (oczywiście każda zmiana powoduje konieczność zbudowania danej relacji rozmytej od nowa w oparciu o nowe parametry). Cykle zmian powtarza się aż do spełnienia warunku stopu, którym jest wystąpienie jednego z trzech zdarzeń: osiągnięcie zadanej wartości funkcji bliskości, brak zmian tej funkcji w kolejnych cyklach zmian parametrów lub wykonanie zadanej liczby cykli uczenia. Metoda ta w swojej podstawowej postaci, choć skuteczna, jest czasochłonna i dodatkowo niesie ryzyko „utknięcia” w jakimś minimum lokalnym (co może łatwo się zdarzyć przy niewłaściwym doborze kroku z jakim zmieniane są parametry). Zamiast niej można zastosować inne metody heurystyczne, w których zastosowano mechanizmy umożliwiające „przeskoczenie” minimum lokalnego [7], np. metodę symulowanego wyżarzania (ang. *Simulated Annealing*), wielkiej powodzi (ang. *Great Deluge*) lub przeszukiwania tabu (ang. *Tabu Search*), bądź któryś z wielu algorytmów ewolucyjnych. Każda z tych metod działa na bazie populacji, której poszczególne składniki są na ogół zależne od pojedynczego parametru, tymczasem relacja rozmyta jest opisywana przez co najmniej dwa parametry (jak w (5)), więc wybraną populacyjną metodę uczenia należy odpowiednio zmodyfikować.

Celem niniejszej pracy jest analiza skutków wyboru celu uczenia (śledzenie jednego lub więcej czynników, odwzorowanie wartości w wybranym punkcie czasu dyskretnego bądź całego przebiegu czasowego), w związku z tym sam mechanizm uczenia ma wartość drugorzędą. Zastosowano metodę różnicową, w której zagrożenie zatrzymaniem algorytmu w minimum lokalnym zniwelowano stosując zmienny krok adaptacji parametrów. Ogólnie rzecz biorąc metoda ta polega na dokonywaniu sukcesywnych zmian kolejnych parametrów kolejnych relacji rozmytych (zwiększanie, a następnie zmniejszanie ich wartości), przy czym po każdej takiej zmianie model jest budowany od nowa i testowany pod kątem wartości funkcji bliskości (6). Jeśli wprowadzona zmiana powoduje zmniejszenie się funkcji bliskości, to jest ona utrwalana. Jeśli nie, to zmiana jest wycofywana. Działanie takie jest przeprowadzane kolejno wobec wszystkich aktywnych relacji rozmytych. Następnie cykl się powtarza i powtórzenia te kontynuuje się do spełnienia warunku stopu, którym jest osiągnięcie wcześniej założonej wartości funkcji (6) lub wykonanie określonej liczby cykli. Jeśli w trakcie kolejnych powtórzeń nie następuje zmiana wartości funkcji (6) to zmienia się krok zmian parametrów. Krok ten ustala się oddzielnie dla każdego rodzaju parametru, a jego wartość początkowa powinna być stosunkowo wysoka, co zmniejszy ryzyko zakończenia adaptacji w lokalnym minimum funkcji bliskości (6). Zmienianymi parametrami są  $\bar{r}_{i,j}$  oraz  $\sigma_{i,j}$  (z (5)). Ich wartości początkowe ustala się arbitralnie (na podstawie wiedzy ekspertowej) lub losuje z zakresu, którego granice ustala się w oparciu o wiedzę ekspertową. Liczbę punktów próbkowania nośnika  $k$  ustala się na poziomie zapewniającym akceptowalną dokładność – np. metodą omówioną w [2]. Dobór rozpiętości nośnika stanowi odrębne istotne zagadnienie wykraczające poza ramy niniejszej pracy. Z badań [3,9,10] wynika, że wystarczające jest stosowanie nośnika o zakresie  $[-2, 2]$  i taki też zakres wykorzystano w niniejszej pracy. Na potrzeby wyznaczania wartości funkcji bliskości (6) zastosowano wyostrzenie rozmytych wartości czynników metodą średniej ważonej.

## 2. ZALEŻNOŚĆ WYNIKÓW UCZENIA RRMK OD CELU MODELOWANIA

Dobór parametrów modelu RRMK zależy od celów modelowania oraz od akceptowalnego poziomu niedokładności modelu. Ten drugi czynnik ma wpływ głównie na liczbę wybranych czynników kluczowych (im jest mniejsza tym model jest prostszy i szybszy w działaniu, co jednak przekłada się na obniżenie dokładności) oraz gęstość próbkowania nośnika (im większa tym większa dokładność modelu, co jednak wydłuża czas jego działania). Głównym czynnikiem różnicującym sposób adaptacji parametrów jest cel modelowania, ponieważ to on decyduje o wyborze formy funkcji bliskości (6). Ogólna postać modelu RRMK ma charakter dynamiczny (zgodny z (4)), co oznacza, że model taki może służyć do odwzorowywania zarówno statycznych stanów obiektu, jak i dynamiki przebiegów czasowych wartości poszczególnych czynników. Zależnie od tak rozumianych celów modelowania można przyjmować różne formy funkcji bliskości (6). Dla zademonstrowania opisywanego podejścia zastosowano wspólną, kwadratową postać funkcji optymalizacji  $\Phi$  z (6). W celu uproszczenia opisu przyjęto także, że przedmiotem adaptacji będą indywidualne parametry poszczególnych relacji (tzn.  $\bar{r}_{i,j}$  oraz  $\sigma_{i,j}$ ), natomiast wspólny dla wszystkich parametr  $k$  będzie dobierany wstępnie inną metodą (dla ułatwienia porównań zachowano jednakową wartość  $k$  we wszystkich modelach). Przedmiotem analizy będzie przejście systemu od jednego stanu stabilnego do drugiego pod wpływem pobudzenia wybranej grupy czynników (zakłada się, że znormalizowane wartości czynników w stanie ustalonym wynoszą zero). W kolejnych podrozdziałach pokazano wyniki uczenia modelu dla dwóch różnych zakładanych celów modelowania. Przedmiotem modelowania jest obiekt, w którym wyodrębniono 4 czynniki, a następnie dokonano pobudzenia jednego z nich (nr 1) sygnałem o wartości 0,5. W efekcie wartości czynników obiektu zmieniały się tak, jak to obrazuje rys. 4.



**Rys. 4.** Przebiegi czasowe wartości czynników obiektu ( $Z_1, \dots, Z_4$ ) po pobudzeniu czynnika nr 1 sygnałem o wartości 0,5.

Przebiegi z rys. 4 stanowiły następnie wzorzec do uczenia modelu RRMK. W wyniku dodatkowych prób dobrano wstępnie wspólny dla obu uczonych modeli nośnik  $u$  o zakresie  $[-2, 2]$ , z liczbą punktów próbkowania  $k = 17$  (takie podejście ułatwia porównania). Wartości czynników były rozmywane zgodnie z funkcją przynależności (3) ze wspólnym dla wszystkich czynników współczynnikiem rozmycia  $\sigma = 0,5$ , natomiast relacje rozmyte budowano w oparciu o funkcję przynależności (4). Na potrzeby uczenia przyjęto wspólne dla wszystkich relacji początkowe wartości współczynników  $\bar{r}_{i,j} = 0$  oraz  $\sigma_{i,j} = 0,4$ . Następnie przeprowadzono uczenie metodą różnicową ze zmiennym krokiem zmian parametrów  $\bar{r}_{i,j}$  oraz  $\sigma_{i,j}$  (malejącym od 0,05 do 0,01) dla dwóch celów modelowania. Wyniki opisano poniżej.

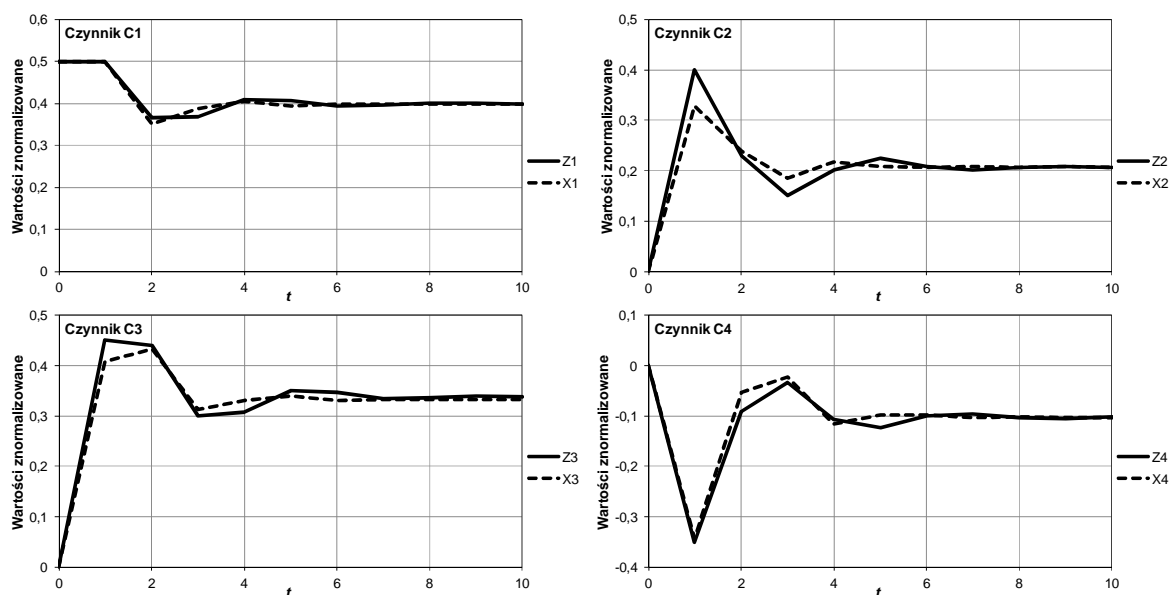
## 2.1. Uczenie modelu na potrzeby uzyskania założonych wartości wszystkich czynników w wybranym pojedynczym punkcie czasu dyskretnego

Podejście to ma zastosowanie w sytuacji, gdy przedmiotem zainteresowania są wartości wszystkich czynników w określonym kroku czasu dyskretnego (najczęściej po uzyskaniu stanu równowagi lub zbliżonego do równowagi). Wtedy równanie (6) przyjmie postać (8):

$$J(t) = \sqrt{\sum_{i=1}^n (\bar{X}_i(t) - Z_i(t))^2} \quad (8)$$

gdzie:  $J(t)$  – funkcja bliskości modelu w badanym kroku  $t$  czasu dyskretnego;  $\bar{X}_i(t)$  – wyostrzona wartość  $i$ -tego czynnika w badanym kroku  $t$  czasu dyskretnego;  $Z_i(t)$  – zadana (w postaci liczby rzeczywistej) znormalizowana wartość  $i$ -tego czynnika w badanym kroku  $t$  czasu dyskretnego;  $n$  – liczba czynników  $t$  – krok czasu dyskretnego, dla którego oblicza się wartość funkcji bliskości.

W wyniku uczenia uzyskano model, dla którego wartość funkcji bliskości (8) dla  $t = 10$  wyniosła  $J(10) = 0,000036$ . Przebiegi uzyskane przez model, porównane z przebiegami odniesienia, pokazano na rys. 5.



Rys. 5. Przebiegi czasowe wyostrzonych wartości czynników modelu ( $X_1, \dots, X_4$ ) w porównaniu z wartościami czynników obiektu ( $Z_1, \dots, Z_4$ ) po pobudzeniu czynnika nr 1 sygnałem o wartości 0,5 – po zakończeniu uczenia, którego celem było dopasowanie modelu do dokładnego odwzorowywania czynników w 10-tym kroku czasu dyskretnego (liczonym od momentu pobudzenia systemu).

## 2.2. Uczenie modelu na potrzeby uzyskania założonych przebiegów czasowych wartości wszystkich czynników

Takie podejście można stosować w sytuacji modelowania dynamiki systemu, gdy przedmiotem zainteresowania są przebiegi czasowe wartości wszystkich czynników począwszy od momentu zachwiania równowagi systemu przez określoną liczbę kroków czasu

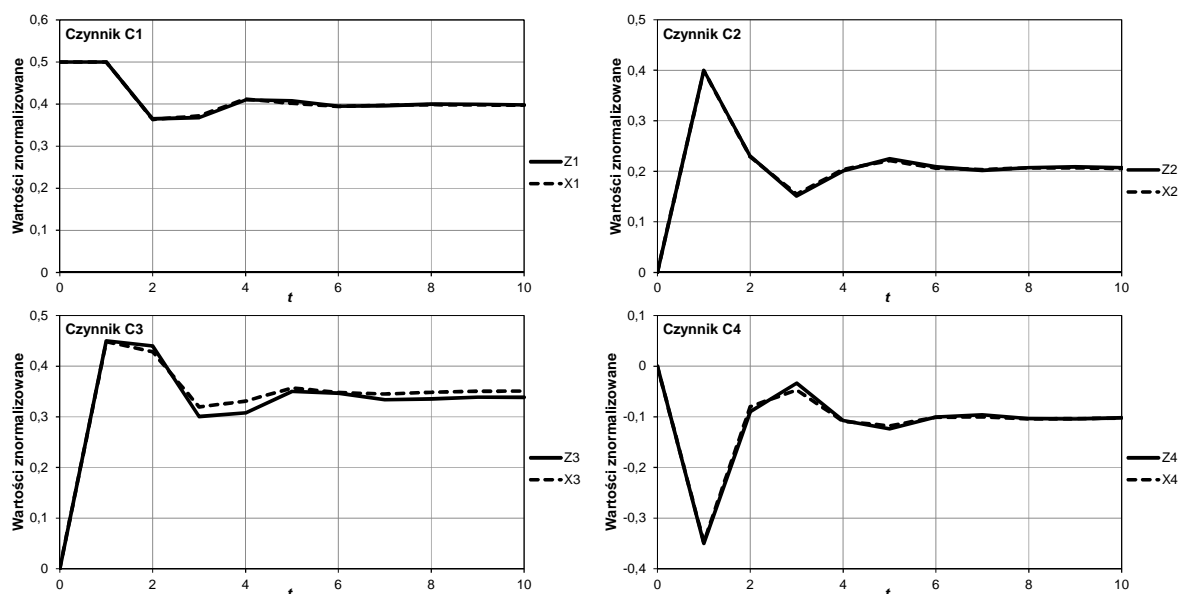


diskretnego (zwykle do uzyskania nowego stanu równowagi lub stanu zbliżonego do równowagi). Metoda taka może być wykorzystywana przy analizie obiektów technicznych lub gdy pośrednie wartości czynników – pomiędzy krańcowymi stanami równowagi – mają znaczenie (np. muszą się mieścić w określonych granicach). Wtedy równanie (6) przyjmie postać (9):

$$J = \sqrt{\sum_{t=0}^T \sum_{i=1}^n (\bar{X}_i(t) - Z_i(t))^2} \quad (9)$$

gdzie:  $J$  – funkcja bliskości modelu w badanym przedziale  $T$  czasu dyskretnego;  $\bar{X}_i(t)$  – wyostrzona wartość  $i$ -tego czynnika w kroku  $t$  czasu dyskretnego;  $Z_i(t)$  – zadana (w postaci liczby rzeczywistej) znormalizowana wartość  $i$ -tego czynnika w kroku  $t$  czasu dyskretnego;  $T$  – czas stabilizacji modelu dynamicznego (wyrażony w liczbie kroków czasu dyskretnego);  $n$  – liczba czynników.

W wyniku uczenia uzyskano model, dla którego wartość funkcji bliskości (9) dla  $T = 10$  wyniosła  $J = 0,00214$ . Przebiegi uzyskane przez model, porównane z przebiegami odniesienia, pokazano na rys. 6.



**Rys. 6.** Przebiegi czasowe wyostrzonych wartości czynników modelu ( $X_1, \dots, X_4$ ) w porównaniu z wartościami czynników obiektu ( $Z_1, \dots, Z_4$ ) po pobudzeniu czynnika nr 1 sygnałem o wartości 0,5 – po zakończeniu uczenia, którego celem było dopasowanie modelu do dokładnego odwzorowywania przebiegów czasowych czynników podczas 10 kroków czasu dyskretnego (licząc od momentu pobudzenia systemu).

## PODSUMOWANIE

Modele RRMK, uzyskane przy zastosowaniu dwóch podejść zależnych od celu modelowania, różnią się, chociaż odnoszą się do tego samego modelowanego obiektu. Bezpośrednie wyniki uczenia, w postaci końcowych wartości parametrów, przedstawiono w tab. 1.

**Tab. 1.** Wyniki uczenia modeli RRMK w zależności od celu modelowania. a) dopasowanie dot. wybranego punktu czasu dyskretnego (jak w podrozdz. 2.1); b) dopasowanie dot. kompletnych przebiegów czasowych (jak w podrozdz. 2.1).

a)					b)				
$\bar{r}_{i,j} / \sigma_{i,j}$	$X_1$	$X_2$	$X_3$	$X_4$	$\bar{r}_{i,j} / \sigma_{i,j}$	$X_1$	$X_2$	$X_3$	$X_4$
$X_1$	0,00 / 0,40	0,70 / 0,38	1,00 / 0,50	-0,71 / 0,42	$X_1$	0,00 / 0,40	0,95 / 0,46	1,00 / 0,25	-0,78 / 0,45
$X_2$	-0,24 / 1,32	0,00 / 0,40	0,15 / 0,36	0,54 / 0,51	$X_2$	-0,28 / 0,87	0,00 / 0,40	0,12 / 0,42	0,37 / 0,45
$X_3$	-0,17 / 1,27	0,00 / 0,49	0,00 / 0,40	0,32 / 0,46	$X_3$	-0,18 / 0,69	-0,27 / 0,58	0,00 / 0,40	0,30 / 0,48
$X_4$	0,12 / 1,33	0,25 / 0,44	0,07 / 0,89	0,00 / 0,40	$X_4$	-0,14 / 0,72	0,16 / 0,48	0,18 / 0,44	0,00 / 0,40

Model opisany w podrozdz. 2.1 charakteryzuje się wysoką dokładnością w wybranym punkcie czasu dyskretnego (sumaryczny błąd równy 0,000036), przy relatywnie niższej dokładności odwzorowywania pełnych przebiegów czasowych (sumaryczny błąd równy 0,01328). Z drugiej strony, model z podrozdz. 2.1 wykazuje ok. pięciokrotnie niższą dokładność w tym samym wybranym kroku czasu (sumaryczny błąd równy 0,000152) przy jednocześnie znacznie wyższej dokładności odwzorowywania przebiegów (sumaryczny błąd równy 0,00214). Dane te nie pozwalają na jednoznaczne wskazanie modelu, który byłby lepszy. Wybór taki zależy głównie od spełniania przez dany model wymagań stawianych w danym momencie. Wynika stąd, że, niezależnie od rodzaju inteligentnego podejścia zastosowanego do modelowania złożonego nieprecyzyjnego systemu, najlepszym rozwiązaniem jest tworzenie banku modeli i bieżące posługiwanie się tym spośród nich, który będzie najlepiej dostosowany do aktualnego zadania.

## ANALYSIS OF SELECTED LEARNING ALGORITHMS IN A RELATIONAL FUZZY COGNITIVE MAP

### *Abstract*

*Relational fuzzy cognitive map is a certain mathematical model which uses some methods of the computational intelligence and can be used to model complex systems characterized by imprecision. A key problem in this kind of issues is an adaptation of parameters of the model (learning) in the way in which his work will imitate as closely as possible the behavior of a real object. The article presents basic information about problems of the model of fuzzy relational cognitive map learning and the effects of different approaches to the learning in dependence on the modeling purposes.*

### BIBLIOGRAFIA

1. Haykin S., *Neural Networks: A Comprehensive Foundation*. 29<sup>th</sup> ed. Delhi, India: Pearson Education, Inc., 2005.
2. Jastrzebow A., Słoń G., *Obliczenia ziarniste w modelowaniu nieprecyzyjnych obiektów przy użyciu relacyjnych rozmytych map kognitywnych*. *Pomiary, Automatyka, Kontrola* vol. 56 nr 12/2010, str.: 1449-1452.
3. Jastrzebow A., Słoń G., *Parametric identification of dynamic imprecise models based on relational fuzzy cognitive maps*. in *Computer Applications in Electrical Engineering*, Nawrowski R., Ed. Poznań: Poznan University of Technology, 2011, vol. 9, pp. 234-244.
4. Kandasamy W. B. V., Smarandache F., *Fuzzy Cognitive Maps and Neutrosophic Cognitive Maps*. Xiquan, Phoenix, AZ, USA, 2003.

5. Kosko B., *Fuzzy cognitive maps*. Int. Journal of Man-Machine Studies, vol. 24, pp. 65-75, 1986.
6. Łachwa A., *Rozmyty świat zbiorów, liczb, relacji, faktów, reguł i decyzji*. Warszawa: Akademicka Oficyna Wydawnicza EXIT, 2001.
7. Papageorgiou E. I., *Learning Algorithms for Fuzzy Cognitive Maps – A Review Study*. IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews, vol. 42, no. 2, pp. 150-163, 2012.
8. Rutkowski L., *Metody i techniki sztucznej inteligencji*. Warszawa: PWN, 2005.
9. Słoń G., *Wybrane problemy projektowania rozmytych relacyjnych map kognitywnych*. in Technologie komputerowe w rozwoju nauki, techniki i edukacji, Jastriebow A., Kuźmińska-Sołśnia B., Raczyńska M., Eds. Radom: Wydawnictwo Naukowe Instytutu Technologii Eksploatacji - Państwowego Instytutu Badawczego, 2012, pp. 131-144.
10. Słoń G., Yastrebov A., *Optimization and Adaptation of Dynamic Models of Fuzzy Relational Cognitive Maps*. in Lecture Notes in Artificial Intelligence 6743, RSFDGrC 2011, Kuznetsov S. O., Ed. Berlin, Heidelberg, Germany: Springer-Verlag, 2011, pp. 95-102.
11. Stylios C. D., Groumpos P. P., *Fuzzy cognitive maps in modeling supervisory control systems*. Journal of Intelligent & Fuzzy Systems, vol. 8, no. 2, pp. 83-98, 2000.

**Autor:**

**dr inż. Grzegorz SŁOŃ** – Politechnika Świętokrzyska