



Marcin Hernes

Uniwersytet Ekonomiczny we Wrocławiu
Wydział Informatyki i Gospodarki Elektronicznej
Katedra Informatyki Ekonomicznej
marcin.hernes@ue.wroc.pl

INTEGRACJA WIEDZY NIEUSTRUKURALIZOWANEJ W WIELOAGENTOWYM SYSTEMIE WSPOMAGANIA DECYZJI FINANSOWYCH

Streszczenie: Agenty funkcjonujące w systemach wieloagentowych często podejmują decyzje na podstawie wyników przetwarzania wiedzy nieustrukturalizowanej. Ponieważ wiedza ta pochodzi z heterogenicznych źródeł, to przedstawiane przez agenty warianty decyzji mogą się różnić. Jednak w procesie decyzyjnym wymagana jest jedna, ostateczna decyzja, zatem wiedza ta powinna być automatycznie integrowana. Celem niniejszego artykułu jest opracowanie metody integracji wiedzy nieustrukturalizowanej z wykorzystaniem teorii consensusu w wieloagentowym systemie wspomaganie decyzji finansowych. W pierwszej części artykułu przedstawiono zagadnienia związane z reprezentacją wiedzy nieustrukturalizowanej agentów. Następnie opracowano algorytmy consensusu umożliwiające integrację wiedzy. W ostatniej części artykułu scharakteryzowano eksperyment badawczy przeprowadzony w celu weryfikacji opracowanej metody.

Słowa kluczowe: systemy wieloagentowe, decyzje finansowe, integracja wiedzy nieustrukturalizowanej.

Wprowadzenie

Rynek finansowy charakteryzuje się dużą zmiennością warunków funkcjonowania [Jajuga, 2007], dlatego też podejmowanie decyzji finansowych jest procesem ciągłym i wiąże się z wielowariantowością ze względu na ich wielokryterialny charakter, a kolejne sytuacje decyzyjne występują chronologicznie w czasie zbliżonym do rzeczywistego i są zawsze związane z ryzykiem.

Wielowariantowość związana z podejmowaniem decyzji pociąga za sobą potrzebę analizy i wartościowania dużej ilości informacji oraz wyciągania wniosków. Ponieważ są to czynności czasochłonne i praktycznie niemożliwe do zrealizowania przez decydenta w czasie rzeczywistym, dlatego też niezbędne stało się korzystanie z systemów wspomagających podejmowanie decyzji. Można wśród nich wymienić systemy wieloagentowe. Systemy te umożliwiają automatyczne, szybkie odnalezienie informacji o odpowiedniej wartości i wyciągnięcie na ich podstawie wniosków [Sobieska-Karpińska i Hernes, 2014]. Regułą jest, że każdy agent w systemie korzysta z innej metody wspomagania decyzji finansowych, jak również może analizować inny segment rynku. Pewna liczba agentów może np. podejmować decyzje z wykorzystaniem analizy technicznej¹, na podstawie różnego rodzaju wskaźników, co wiąże się z przetwarzaniem głównie wiedzy ustrukturalizowanej. Coraz częściej jednak agenci podejmują decyzje z wykorzystaniem analizy fundamentalnej² lub behawioralnej³, co wiąże się z przetwarzaniem opinii ekspertów lub inwestorów. Opinie te znajdują się między innymi na stronach internetowych biur maklerskich, banków lub też na portalach społecznościowych / blogach o tematyce finansowej. Zapisane są one głównie w języku naturalnym, co wiąże się z problemem przetwarzania wiedzy nieustrukturalizowanej.

Ponieważ agenci te charakteryzują się posiadaniem wiedzy pochodzącej z autonomicznych i rozproszonych źródeł, to przedstawiane przez nie warianty decyzji mogą się różnić. Wiedza agentów cechuje się więc pewnym poziomem heterogeniczności. W procesie decyzyjnym wymagana jest natomiast jedna, ostateczna decyzja. Zatem wiedza ta powinna być automatycznie integrowana. Może to być dokonane na przykład z użyciem pewnych kryteriów lub funkcji oceny wiedzy poszczególnych agentów. Jednakże w przypadku niewłaściwego lub mało precyzyjnego określenia tych kryteriów lub funkcji, zwiększa się poziom ryzyka wyboru wariantu, który nie gwarantuje odpowiedniego poziomu satysfakcji. Lepszym rozwiązaniem może być zastosowanie metod consensusu, które umożliwiają integrację wiedzy zarówno ustrukturalizowanej, jak i nieustrukturalizowanej. W consensusie każda ze stron jest brana pod uwagę, każda

¹ Analiza techniczna bazuje na aktualnym kursie i obrocie, punkcie odniesienia i danych z przeszłości, na podstawie których dokonuje się obliczeń różnego rodzaju wskaźników. Ten rodzaj analizy można stosować nawet nie znając nazwy emitenta papieru wartościowego.

² Analiza fundamentalna polega na badaniu kondycji finansowej spółki lub innego emitenta papierów wartościowych, jego silnych i słabych stron oraz perspektyw rozwoju. Odpowiada się w niej na pytanie, czy warto zainwestować w rozpatrywane papiery wartościowe (przy czym w przypadku notowań walutowych bada się całą gospodarkę danego kraju).

³ Analiza behawioralna polega na badaniu zachowania i nastrojów uczestników rynku finansowego.

ze stron konfliktu „traci” najmniej jak tylko to jest możliwe, każda ze stron wnosi swój wkład w consensus, wszystkie strony go akceptują i jest on reprezentacją wszystkich stron konfliktu. Decyzja wyznaczona za pomocą tych metod nie musi być więc decyzją wygenerowaną przez któregokolwiek z agentów, może być do niej bardzo zbliżona. Consensus umożliwia zatem integrację wiedzy w czasie rzeczywistym i gwarantuje osiągnięcie dobrego kompromisu przy niższym poziomie ryzyka, co w konsekwencji może prowadzić do wyznaczenia decyzji przynoszącej decydentowi satysfakcjonujące korzyści. Zastosowanie metod consensusu, w celu integracji wiedzy, pozwala również na skrócenie czasu podjęcia decyzji [Sobieska-Karpińska i Hernes, 2013], ponieważ nie ma potrzeby dokonywania przez człowieka analiz, wnioskowania oraz zastanawiania się nad wyborem docelowego wariantu decyzji spośród różnych decyzji generowanych przez agenty.

Celem niniejszego artykułu jest opracowanie metody integracji wiedzy nieustrukturalizowanej z wykorzystaniem teorii consensusu w wieloagentowym systemie wspomaganie decyzji finansowych (WSWDF). System ten zbudowany jest z użyciem architektury kognitywnego programu agentowego The Learning Intelligent Distribution Agent (LIDA) [Franklin i Patterson, 2006]. Integracja wiedzy umożliwi w konsekwencji wyznaczenie ostatecznej decyzji prezentowanej przez system użytkownikowi.

Artykuł został podzielony na cztery części. W pierwszej dokonano analizy pozycji literaturowych z zakresu analizowanej tematyki. Następnie przedstawiono zagadnienia związane z reprezentacją wiedzy nieustrukturalizowanej agentów funkcjonujących w systemie. W trzeciej części przedstawiono opracowane algorytmy consensusu umożliwiające integrację wiedzy. Ostatnia część artykułu przedstawia sposób przeprowadzenia i wyniki eksperymentu badawczego, mającego na celu weryfikację opracowanej metody.

1. Przegląd istniejących rozwiązań

Początkowo w wieloagentowych systemach wspomaganie decyzji wykorzystywano podejście tablicowe w celu ułatwienia wymiany poleceń między współpracującymi agentami oraz określenia reguł biznesowych [Luo, Liu i Davis, 2002]. W pracy [Sycara, Decker i Zeng, 2002] zaproponowano zaś system umożliwiający współpracę użytkownika z wieloma wyspecjalizowanymi agentami, które posiadają dostęp do różnego rodzaju modeli finansowych, analizują sytuację na rynku finansowym z uwzględnieniem kryteriów określonych przez

użytkownika. Natomiast w pracy [Westerhoff, 2011] opisano system, w którym agenty zostały podzielone na dwie grupy – agenty pierwszej grupy podejmują decyzje na podstawie metod analizy fundamentalnej, zaś agenty drugiej grupy podejmują decyzje na podstawie analizy technicznej. W artykule [Korczak, Hernes i Bac, 2014] przedstawiono zaś system wieloagentowy wspomagający inwestowanie na rynku wymiany walut Forex oraz metodę oceny strategii inwestycyjnych wybranych agentów.

W procesie przetwarzania wiedzy nieustrukturalizowanej, realizowanym również w wieloagentowych systemach wspomagania decyzji finansowych, stosowane są następujące metody:

- wyszukiwanie informacji,
- ekstrakcja informacji,
- eksploracja tekstu,
- przetwarzanie języków naturalnych (Natural Language Processing) [Potiopa, 2011; Baldoni i in., 2012].

Głównym celem **wyszukiwania informacji** jest znalezienie odpowiedzi na pytanie użytkownika wśród kolekcji dokumentów. **Ekstrakcja** polega na zidentyfikowaniu instancji pewnej predefiniowanej klasy zdarzeń, ich powiązań oraz wystąpień w dokumentach pisanych w języku naturalnym [Pham i Pham, 2012]. Celem **eksploracji tekstu** jest poznanie ukrytych w tekście informacji z wykorzystaniem metod dostosowanych do dużej liczby danych tekstowych [Łapczyński, 2010, s. 66]. **Przetwarzanie języków naturalnych** zawiera mechanizmy próbujące dokonać „zrozumienia” kontekstu tekstu. W metodach tych nie są obliczane wartości podobieństwa termów, ale przeprowadzane są następujące kategorie analiz tekstu:

1. Płytką analiza tekstu jest określana jako analiza tekstu, której efekt jest niepełny w stosunku do głębokiej analizy tekstu. Zwykle ograniczenie polega na rozpoznawaniu struktur nierekurencyjnych lub o ograniczonym poziomie rekurencji, które mogą być rozpoznane z dużym stopniem pewności.
2. Głęboka analiza tekstu jest procesem komputerowej analizy lingwistycznej wszystkich możliwych interpretacji i relacji gramatycznych występujących w tekście naturalnym. Taka pełna analiza może być bardzo złożona [Sołdacki, 2006].

W procesie analizy dokumentów tekstowych często wykorzystuje się semantyczne metody reprezentacji wiedzy, w tym sieci semantyczne [Dudycz, 2013]. Dzięki ich zastosowaniu możliwa jest szeroko rozumiana reprezentacja wiedzy, w której istotną kwestią jest zwrócenie uwagi na wzajemne zależności występujące pomiędzy obiektami.

Coraz częściej w celu przetwarzania wiedzy nieustrukturalizowanej wykorzystywane są kognitywne programy agentowe, które realizują funkcje poznawcze i decyzyjne, takie, jakie zachodzą w ludzkim mózgu, dzięki temu potrafią „zrozumieć” rzeczywiste znaczenie obserwowanych zjawisk i procesów biznesowych zachodzących m.in. na rynkach finansowych [Duch, 2010].

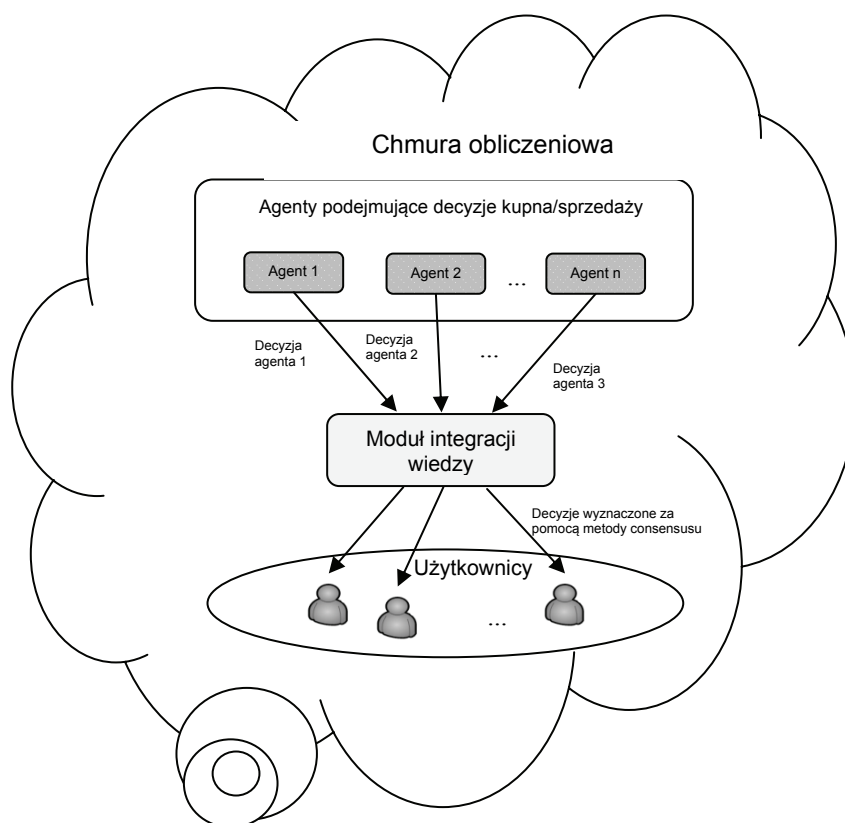
Prace nad wykorzystaniem metod consensusu w celu integracji wiedzy były prowadzone przez wielu autorów. Teoria consensusu dotyczy zarówno prostych struktur (porządek liniowy lub częściowy porządek), jak również struktur bardziej złożonych (podziały, hierarchie, n-drzewa) [Daniłowicz i Nguyen, 1988; Barthlemy, 1992; Maleszka, Mianowska i Nguyen 2013]. Opracowane zostały również metody consensusu dla struktur wielowartościowych i wieloatrybutowych [Nguyen, 2002; Hernes i Nguyen 2007]. W pracy [Nguyen, 2008] został zaproponowany formalny matematyczny model integracji wiedzy. Wykorzystuje on bazującą na modelu consensusu funkcję integracji wiedzy. Ta metodologia została wykorzystana do rozwiązywania konfliktów i nieścisłości wiedzy oraz jej integracji. Prowadzone były również prace nad wyznaczeniem consensusu w sytuacji wystąpienia zależności funkcyjnych pomiędzy atrybutami struktury wiedzy agenta [Zgrzywa, 2007].

Jednakże w dotychczasowych rozwiązaniach dotyczących integracji wiedzy w wieloagentowych systemach wspomaganie decyzji finansowych zajmowano się głównie wiedzą ustrukturalizowaną. Problem integracji wiedzy nieustrukturalizowanej nie był przedmiotem dotychczasowych badań.

2. Reprezentacja wiedzy nieustrukturalizowanej w WSWDF

Celem WSWDF jest wspomaganie inwestowania na rynkach finansowych poprzez podpowiadanie inwestorowi decyzji dotyczących kupna/sprzedaży. System składa się z następujących elementów (rys. 1):

1. Agenty podejmujące decyzje kupna/sprzedaży, przetwarzające wiedzę ustrukturalizowaną i nieustrukturalizowaną.
2. Moduł integracji wiedzy – funkcjonuje z wykorzystaniem metody consensusu; w module tym, na podstawie decyzji wygenerowanych przez agenty, wyznaczana jest decyzja ostateczna prezentowana użytkownikowi.
3. Użytkownicy – osoby inwestujące na rynku finansowym.



Rys. 1. Architektura WSWDF

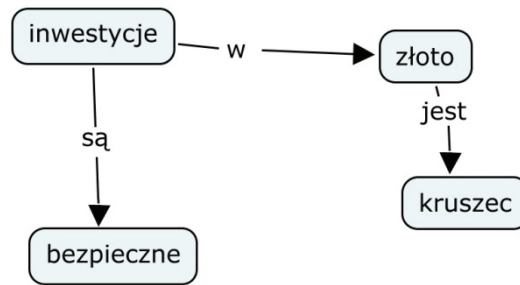
W budowie WSWDF wykorzystano architekturę The Learning Intelligent Distribution Agent (LIDA) opracowaną przez S. Franklina [Franklin i Patterson, 2006]. Zaletą tej architektury jest jej emergentno-symboliczny charakter, dzięki czemu możliwe jest przetwarzanie wiedzy zarówno ustrukturalizowanej (numerycznej i symbolicznej), jak i nieustrukturalizowanej (zapisanej w języku naturalnym). Przyjęto, że agent kognitywny składa się z następujących modułów:

- pamięć robocza,
- globalna pamięć robocza,
- pamięć sensoryczna,
- pamięć percepcyjna,
- pamięć epizodyczna,
- pamięć deklaratywna,
- pamięć sensoryczno-motoryczna,
- selekcja działań,
- bieżąca świadomość [www 1].

Agent kognitywny funkcjonuje w ramach cyklu kognitywnego, który dzieli się na trzy fazy: zrozumienia, świadomości oraz wyboru działań i uczenia się. Na początku fazy zrozumienia bodźce odbierane z otoczenia aktywują codelety (*codelets*)⁴ cech niskiego poziomu funkcjonujące w pamięci sensorycznej [www 1]. Wyjścia tych codeletów aktywują pamięć percepcyjną, gdzie codelety cech wysokiego poziomu zasilają bardziej abstrakcyjne wystąpienia, takie jak obiekty, kategorie, działania lub zdarzenia. Wyniki percepcji przekazywane są do pamięci roboczej, natomiast z użyciem pamięci epizodycznej i deklaratywnej tworzone są lokalne powiązania, a następnie, z wykorzystaniem wystąpień pamięci percepcyjnej, wygenerowany jest bieżący model sytuacyjny; innymi słowy agent posiada zdolność rozumienia zjawisk zachodzących w otoczeniu organizacji. Faza świadomości rozpoczyna się formowaniem koalicji najistotniejszych elementów modelu sytuacyjnego, które następnie przekazywane są do modułu bieżącej świadomości. Zawartość modułu bieżącej świadomości jest następnie przekazywana do globalnej pamięci roboczej, inicjalizując jednocześnie fazę wyboru działań agenta. W fazie tej możliwe schematy działań pobierane są z pamięci proceduralnej i przesyłane do modułu selekcji działań, gdzie konkurują o wybór w tym cyklu kognitywnym. Wybrane działania uruchamiają pamięć sensoryczno-motoryczną w celu utworzenia odpowiedniego algorytmu ich wykonania, co jest końcowym etapem cyklu kognitywnego [www 1]. Równoległe z poprzednimi działaniami realizowane jest uczenie się agenta.

Jako przedmiot niniejszych badań zostały wybrane agenty przetwarzające wiedzę nieustrukturalizowaną, zawartą głównie w dokumentach tekstowych. Metody analizy dokumentów tekstowych z wykorzystaniem agenta LIDA zostały scharakteryzowane w pracach [Bytniewski i Hernes 2014; Hernes, 2015] (ze względu na ograniczenie objętości tekstu, nie będą one przedmiotem analizy w niniejszym artykule). Z punktu widzenia opracowania metody integracji wiedzy najbardziej istotny jest sposób reprezentacji wiedzy nieustrukturalizowanej. W tym celu agent LIDA wykorzystuje sieć semantyczną, której przykład został przedstawiony na rys. 2.

⁴ W architekturze LIDA są to wyspecjalizowane, mobilne programy przetwarzające informację w modelu globalnej pamięci roboczej [www 1].



Rys. 2. Wycinek sieci semantycznej reprezentujący zdanie „bezpiecznie będzie zainwestować w złoto”

Formalna definicja sieci semantycznej⁵ przedstawia się następująco:

Definicja 1

Siecią semantyczną nazywamy trójkę:

$$SN = \langle N, I, L \rangle,$$

gdzie:

N – zbiór węzłów,

I – zbiór instancji,

L – zbiór łuków, czyli zbiór relacji zdefiniowanych na N .

Zdefiniowane zostaną teraz poszczególne elementy sieci semantycznej. Przyjęto, że wycinek rzeczywistości analizowanej przez agenta reprezentowany jest jako dwójka $\langle O, V \rangle$, gdzie O jest skończonym zbiorem obiektów, natomiast V jest dziedziną O , czyli V jest zbiorem wartości obiektów, przyjęto również, że:

$$V = \bigcup_{o \in O} V_o,$$

gdzie:

V_o jest dziedziną obiektu o .

Założono, że sieć semantyczna odnosi się do wycinka rzeczywistości $\langle O, V \rangle$.

Szczegółowe definicje sieci semantycznej przedstawione w dalszej części artykułu odnoszą się do trzech poziomów:

- poziom węzłów,
- poziom instancji,
- poziom łuków.

⁵ Sieć semantyczna (*semantic network*) w rozumieniu Quilliana [1963] jako zbiór obiektów powiązanych ze sobą różnymi relacjami.

Definicja 2

Węzłem sieci semantycznej opisującej wycinek rzeczywistości $\langle O, V \rangle$ nazywamy trójkę:

$$nd = \langle n, O^n, V^n \rangle,$$

gdzie:

n – niepowtarzalna nazwa węzła,

$O^c \in O$ – obiekt reprezentowany przez węzeł,

$V^c \in V$ – dziedzina obiektu:

$$V^c = \bigcup_{\langle o, v \rangle \in O^n} V_o.$$

Para $\langle O^n, V^n \rangle$ nazywana jest strukturą węzła nd . Wszystkie węzły należące do tej samej sieci semantycznej różnią się, jednakże mogą występować węzły o tej samej strukturze (np. “osoba”, “człowiek”).

Definicja 3

Instancja węzła nd jest opisana przez obiekty ze zbioru O^n o wartościach ze zbioru V^n i definiowana następująco:

$$\text{instance} = \langle i, v \rangle$$

gdzie i jest niepowtarzalnym identyfikatorem instancji w wycinku rzeczywistości $\langle O, V \rangle$ natomiast v jest wartością instancji obiektu O^n określoną funkcją:

$$v : O^n \rightarrow V^n,$$

taką, że $v(o, p) \in V_o$ dla każdego $\langle o, p \rangle \in O^n$.

Wartość v jest także nazywana opisem instancji obiektu. Węzeł może być również interpretowany jako zbiór wszystkich instancji opisywanych przez niego.

W celu zapisania faktu, że i jest instancją węzła nd można używać notacji $i \in nd$. Wszystkie instancje tego samego węzła powinny się różnić. Ta sama instancja może należeć do różnych węzłów i może mieć różne wartości.

Definicja 4

Niech dany jest zbiór węzłów N . Łukiem nazywana jest następująca relacja:

$$L = \langle nd, nd' \rangle.$$

gdzie:

$nd, nd' \in N$.

Wykorzystując definicje 1, 2, 3 i 4 w dalszej części artykułu przedstawiona zostanie metoda integracji wiedzy nieustrukturalizowanej w WSWDF.

3. Metoda integracji wiedzy

W ogólnym ujęciu integracja wiedzy nieustrukturalizowanej w WSWDF dotyczy sytuacji, w której istnieje pewna liczba sieci semantycznych reprezentujących ten sam wycinek rzeczywistości w określonym czasie (opinii ekspertów/ inwestorów dotyczących prognozy zachowania rynku finansowego w określonym czasie). Integracja polega na wyznaczeniu reprezentacji tych sieci semantycznych (jednej sieci semantycznej). W celu integracji wiedzy opracowano algorytmy wyznaczania consensusu, które odnoszą się do poziomów węzłów, instancji i łuków. Moduł integracji wiedzy w WSWDF funkcjonuje z wykorzystaniem tych algorytmów.

Biorąc pod uwagę poziom węzłów należy zauważyć, że te same węzły mogą posiadać różne struktury w różnych sieciach semantycznych. Przyjmijmy, że w sieci semantycznej SN_i węzeł ma strukturę $\langle O^i, V^i \rangle$ dla $i = 1, \dots, M$ (M – liczba ontologii, które wymagają integracji). Problem integracji na poziomie węzłów można przedstawić następująco:

Dla danego zbioru par: $X = \{\langle O^i, V^i \rangle\}$ (gdzie $\langle O^i, V^i \rangle$ jest strukturą węzła w sieci semantycznej SN_i) należy wyznaczyć dwójkę $\langle O^*, V^* \rangle$ reprezentującą dany zbiór. Zadanie to jest realizowane przez algorytm 1.

Algorytm 1

Dane: Struktury węzłów $\langle O^i, V^i \rangle$ dla $i = 1, \dots, M$.

Wyjście: Dwójka $\langle O^*, V^* \rangle$ jako wynik integracji danych dwójek.

START

1. Zbiór $O^* = \bigcap_{i=1}^M O^i$.
2. Dla każdej pary obiektów a, b ze zbioru O^* :
Jeżeli $a \rightarrow b$ oraz b nie występuje w relacji z innym obiektem ze zbioru O^* to zbiór $O^* = O^* \setminus \{b\}$.
3. Dla każdego obiektu a ze zbioru O^* wyznacz jego dziedzinę V_a jako sumę jego dziedzin w dwójkach $\langle O^i, V^i \rangle$.

STOP

Algorytm odnoszący się do poziomu instancji można opracować przyjmując następujące założenia:

1. Dany jest zbiór wartości $X = \{v_1, \dots, v_M\}$, gdzie v_i jest wartością instancji obiektu O_i oraz funkcja odległości d pomiędzy wartościami. Funkcje odległości określone są w odniesieniu do poszczególnych typów struktur danych. Prostym przykładem może być odległość w przestrzeni euklidesowej. Zdefiniowano również odległości w odniesieniu do bardziej złożonych struktur wiedzy. Na przykład w pracy [Nguyen, 2002] określono funkcje odległości dla podziałów i pokryć uporządkowanych, natomiast, w pracy [Maleszka, Mianowska i Nguyen, 2013] określono funkcję odległości dla drzew hierarchicznych. Na potrzeby eksperymentu badawczego przedstawionego w niniejszym artykule (punkt 4) wykorzystano funkcję odległości dla struktury decyzji finansowej (określonej jako zbiór instrumentów finansowych) zdefiniowaną w pracy [Hernes i Sobieska-Karpińska, 2016]. Zgodnie z tą miarą odległością pomiędzy dwoma strukturami decyzji finansowych jest liczba operacji dodawania/eliminacji elementów struktury, które należy wykonać w celu przekształcenia pierwszej struktury w drugą strukturę. Jako odległość pomiędzy fragmentami tekstu wykorzystano odległość kosinusową (można również wykorzystać odległość Hamminga).
2. Należy znaleźć wartość v danego typu, która najlepiej reprezentuje zbiór X .

Algorytm 2

Dane: zbiór $X = \{v_1, \dots, v_M\}$.

Wynik: Wartość v jako wynik integracji.

START

1. Wyznacz v , takie, że:

$$\sum_{i=1}^M d(v, v_i) = \min_{v \in \text{type}(O)} \sum_{i=1}^M d(v, v_i)$$

gdzie: $O = \bigcup_{i=1}^M O_i$.

STOP

W odniesieniu do poziomu łuków, rozważmy sytuację, w której łuki pomiędzy węzłami nd i nd' różnią się w sieciach semantycznych, które powinny być zintegrowane. Problem integracji można zdefiniować następująco:

Dla danego zbioru

$X = \{L_{SN_i}(nd, nd')\}$ dla $i = 1, \dots, M$ należy wyznaczyć łuki $L(nd, nd)$, które najlepiej reprezentują zbiór X .

Algorytm 3

Dane: Zbiór $X = \{L_{SN_i}(nd, nd')\}$ dla $i=1, \dots, M$.

Wyjście: Łuk $L(nd, nd')$ jako wynik integracji.

START

1. Zbiór $L(nd, nd') = \bigcup_{i=1}^M L_{SN_i}(nd, nd')$.

2. Dla każdego łuku $l \in L(nd, nd')$ wyznacz liczbę $\tau(l)$ jego wystąpień w zbiorze $L_{SN_i}(nd, nd')$.

3. Dla każdego łuku oblicz:

$$\nu(l) = \sum_{(l, l')} \tau(l').$$

4. Wyznacz taki l , że liczba $\nu(l)$ jest maksymalna, oraz

$$L(nd, nd') = L(nd, nd') \setminus \{l\}.$$

STOP

Na podstawie opracowanych algorytmów funkcjonuje moduł integracji wiedzy w WSWDF. W dalszej części artykułu przedstawione zostały wyniki eksperymentu badawczego przeprowadzonego w celu weryfikacji opracowanej metody integracji wiedzy.

4. Eksperyment badawczy

Analizę efektywności opracowanej metody przeprowadzono na danych z przedziału H4 notowań z rynku walutowego Forex. W celu dokonania tej analizy przeprowadzono test, w którym przyjęto następujące założenia:

1. Wykorzystano notowania pary EUR/USD, z trzech, losowo wybranych, następujących okresów:
 - 9.11.2015 g. 0:00 do 12.11.2015 g. 23:59,
 - 16.11.2015 g. 0:00 do 18.11.2015 g. 23:59,
 - 24.11.2015 g. 0:00 do 26.11.2015 g. 23:59.
2. Przy weryfikacji wykorzystano opinie czterech ekspertów pozyskane z portalu finansowych. Opinie te zostały przeanalizowane przez agenta kognitywnego i na podstawie wyników analizy agent ten określił, jakie decyzje powinny być podejmowane w badanych okresach (sygnały kup-wartość 1, sprzedaj-wartość -1, pozostaw bez zmian-wartość 0). Pozyskane opinie zostały zintegrowane w module integracji wiedzy i na podstawie wyników tego procesu również zostały określone decyzje kupna-sprzedaży.

3. Przyjęto, że jednostką miar efektywności (miar względnych) jest pips (zmiana ceny o jeden „punkt” na rynku Forex określana jest jako pips).
4. Nie uwzględnia się kosztów transakcji.
5. Money management – założono, że w każdej transakcji inwestor angażuje 100% posiadanego kapitału. Strategia zarządzania kapitałem może być ustalona przez użytkownika. Analizę efektywności przeprowadzono z wykorzystaniem następujących miar (wskaźników):
 - stopa zwrotu (wskaźnik x_1),
 - liczba transakcji,
 - największy zysk (wskaźnik x_2),
 - największa strata (wskaźnik x_3),
 - całkowity zysk (wskaźnik x_4),
 - liczba transakcji zyskownych (wskaźnik x_5),
 - liczba transakcji zyskownych pod rząd (wskaźnik x_6),
 - liczba transakcji stratnych pod rząd (wskaźnik x_7),
 - wskaźnik Sharpe’a (wskaźnik x_8)

$$S = \frac{E(r) - E(f)}{|O(r)|} \cdot 100\%,$$

gdzie:

$E(r)$ – średnia arytmetyczna stopy zwrotu,

$E(f)$ – średnia arytmetyczna stopy zwrotu wolnej od ryzyka,

$O(r)$ – odchylenie standardowe stóp zwrotu.

- przeciętny współczynnik zmienności (wskaźnik x_9) jest to stosunek odchylenia przeciętnego do średniej arytmetycznej pomnożony przez 100% i jest wyrażony w jednostkach niemianowanych:

$$V = \frac{s}{|E(r)|} \cdot 100\%,$$

gdzie:

V – przeciętny współczynnik zmienności,

s – odchylenie przeciętne stopy zwrotu,

$E(r)$ – średnia arytmetyczna stopy zwrotu.

- Value at Risk (wskaźnik x_{10}) – miara określana jako wartość narażona na ryzyko – czyli maksymalna strata rynkowej wartości portfela lub instrumentu finansowego możliwa do poniesienia w konkretnym horyzoncie czasowym i przy założonym poziomie ufności.

$$VaR = P * O * k$$

gdzie:

P – wartość kapitału początkowego,

O – zmienność – odchylenie standardowe stóp zwrotu w badanym okresie,

k – odwrotność standardowego skumulowanego rozkładu normalnego (przyjmując poziom ufności 95% wartość k wynosi 1,65).

6. Wykorzystano następującą funkcję oceny [Korczak, Hernes i Bac 2014]:

$$y = (a_1 x_1 + a_2 x_2 + a_3 (1 - x_3) + a_4 x_4 + a_5 x_5 + a_6 x_6 + \dots + a_7 (1 - x_7) + a_8 x_8 + a_9 (1 - x_9) + a_{10} (1 - x_{10})) ,$$

gdzie: x_i oznacza wartości znormalizowane wskaźników wymienionych w pkt. 6 od x_1 do x_{10} .

W badaniu przyjęto, że współczynniki a_1 do $a_{10} = 1/10$. Zaznaczmy, że współczynniki te mogą być modyfikowane z wykorzystaniem np. metody ewolucyjnej lub też ustalane przez użytkownika (inwestora) zgodnie z jego preferencjami (np. użytkownik może określić, czy zainteresowany jest wyższą stopą zwrotu przy jednoczesnym wyższym poziomie ryzyka, czy też niższym poziomem ryzyka, ale jednocześnie zgadza się na niższą stopę zwrotu). Funkcja przyjmuje wartości z zakresu $[0 \dots 1]$, a efektywność jest wprost proporcjonalna do wartości funkcji.

7. Wyniki uzyskane na podstawie poszczególnych opinii i wiedzy zintegrowanej porównane zostały z wynikami strategii Buy-and-Hold (inwestor podejmuje decyzję kupna na początku okresu, a decyzję sprzedaży na końcu okresu).

Badania efektywności metody przeprowadzono w następujący sposób:

1. Na podstawie opinii dotyczących pierwszego okresu agent określał, kiedy kupić, a kiedy sprzedać walutę EUR/USD.
2. W kolejnym kroku, biorąc pod uwagę wyniki uzyskane na podstawie poszczególnych opinii, funkcjonowania modułu integracji wiedzy oraz metody Buy-and-Hold, dla każdej operacji kupna sprzedaży określono wartość posiadanego kapitału oraz określono stopę zwrotu.
3. W końcowym etapie obliczono wartość wskaźników efektywności w odniesieniu do stóp zwrotu wynikających ze wszystkich decyzji (nie tylko z końcowych stóp zwrotu, ale ze wszystkich stóp zwrotu obliczonych po każdej decyzji sprzedaży). Obliczono również funkcje oceny.
4. Następnie kroki od 1 do 3 powtórzono wykorzystując dane z kolejnych okresów. Tabela 1 przedstawia otrzymane wyniki w poszczególnych okresach.

Tabela 1. Wyniki eksperymentu badawczego

Wskaźnik	Ekspert 1			Ekspert 2			Ekspert 3			Ekspert 4			Consensus			B&H		
	okres 1	okres 2	okres 3	okres 1	okres 2	okres 3	okres 1	okres 2	okres 3	okres 1	okres 2	okres 3	okres 1	okres 2	okres 3	okres 1	okres 2	okres 3
Stopa zwrotu	-65	1920	232	22	240	310	-35	311	420	-34	-94	411	15	430	538	-220	730	295
Liczba transakcji	11	18	20	6	12	7	9	10	14	7	5	8	4	5	6	1	1	1
Największy zysk	209	884	239	158	845	358	198	911	680	119	145	539	186	834	735	0	730	295
Największa strata	180	203	109	140	698	120	220	729	456	159	175	368	120	468	532	-220	0	0
Całkowity zysk	186	2485	487	320	1024	485	256	1124	1230	238	1000	1230	234	1123	1431	0	0	0
Liczba transakcji zyskowych	6	15	14	4	7	6	5	6	8	3	2	5	3	4	4	0	1	1
Liczba transakcji zyskowych pod rząd	3	6	6	2	3	5	2	3	3	1	1	3	2	3	3	0	1	1
Liczba transakcji stratnych pod rząd	2	1	3	2	2	1	1	1	2	1	1	1	1	1	1	1	0	0
Wskaźnik Sharpa	0,32	2,75	1,25	1,56	1,32	3,14	0,79	1,09	2,00	3,56	1,02	2,14	3,20	2,00	3,38	0,00	0,00	0,00
Przeciętny współczynnik zmienności [%]	1,80	69,73	1,75	0,76	1,60	9,87	1,01	0,87	1,30	1,45	90,70	2,34	0,17	0,35	0,45	0,00	0,00	0,00
Wartość narazona na ryzyko	6,00	106,67	116,00	4,40	48,0	42,86	3,12	64,00	95,00	5,10	184,0	44,35	1,20	19,50	21,10	-220,00	730,00	295,00
Wartość funkcji oceny (y)	0,18	0,43	0,17	0,32	0,31	0,21	0,23	0,33	0,28	0,19	0,11	0,22	0,28	0,44	0,54	0,12	0,23	0,18

Generalizując wyniki analizy efektywności można zauważyć, że w rozpatrywanych okresach podejmowane decyzje generowały zarówno zyski, jak i straty. W ocenie efektywności należy więc brać pod uwagę nie tylko stopę zwrotu, lecz także inne wskaźniki, uwzględniając również poziom ryzyka związanego z inwestycją, co umożliwia funkcja oceny. Można zauważyć, że ranking ocen decyzji wygenerowanych na podstawie opinii ekspertów, wyników modułu integracji wiedzy (consensus) i metody B&H różni się w poszczególnych okresach. W pierwszym okresie najlepsze decyzje podejmowane były generowane w wyniku integracji wiedzy, natomiast decyzje generowane na podstawie opinii ekspertów uzyskały wyższą ocenę niż ocena benchmarku B&H. W drugim okresie decyzje generowane na podstawie opinii Eksperta 2, Eksperta 3 i Eksperta 4 otrzymały ocenę niższą niż consensus, jednakże najwyżej ocenione zostały decyzje generowane na podstawie opinii Eksperta 1. Benchmark B&H uzyskał ocenę wyższą od decyzji generowanych na podstawie opinii Eksperta 4. W trzecim zaś okresie, najlepszą ocenę uzyskały decyzje otrzymane w wyniku integracji wiedzy. Najniższą ocenę w tym okresie otrzymał benchmark B&H.

Biorąc pod uwagę wszystkie rozpatrywane okresy można stwierdzić, że najczęściej (2 z 3 okresów) najwyżej oceniany były decyzje generowane w wyniku integracji wiedzy, z wykorzystaniem algorytmów wyznaczania consensusu, mimo że nie zawsze stopa zwrotu z tych decyzji była najwyższa. Ocena ta wynika jednakże z niskiego poziomu ryzyka związanego z inwestowaniem na podstawie tych decyzji. Z kolei najczęściej nisko oceniane były decyzje wyznaczone na podstawie opinii Eksperta 4, ponieważ przy stosunkowo dużym poziomie ryzyka, generował niewielkie stopy zwrotu.

Podsumowanie

W niniejszym artykule podjęto problem integracji wiedzy nieustrukturalizowanej w wieloagentowym systemie wspomagania decyzji finansowych. W celu rozwiązania tego problemu wykorzystano teorię consensusu. Opracowane algorytmy wyznaczania consensusu umożliwiają integrację wiedzy w sytuacji wystąpienia różnych opinii ekspertów dotyczących prognozowanej sytuacji na rynkach finansowych. W konsekwencji możliwe jest przedstawienie użytkownikowi jednej, satysfakcjonującej go decyzji, na podstawie której dokonywane są transakcje kupna-sprzedaży. Opracowane algorytmy wyznaczania consensusu pozwolą dodatkowo ułatwić pracę twórcom systemów wieloagentowych wspoma-

gających podejmowanie decyzji finansowych, ponieważ mogą być bezpośrednio implementowane jako moduł integracji wiedzy w tego typu systemie.

Dalsze prace badawcze dotyczyć będą m.in. opracowania algorytmów consensusu w celu integracji wiedzy nieustrukturalizowanej reprezentowanej za pomocą sieci semantycznej z poziomami aktywacji węzłów i luków (nazwanej w literaturze przedmiotu „slipnet”) oraz przeprowadzenia eksperymentów badawczych z wykorzystaniem opinii większej liczby ekspertów.

Literatura

- Baldoni M., Baroglio C., Patti V., Rena P. (2012), *From Tags to Emotions: Ontology-Driven Sentiment Analysis in the Social Semantic Web*, „Intelligenza Artificiale”, Vol. 6, No. 1.
- Barthelemy J.P. (1992), *Dictatorial Consensus Function on n-Trees*, “Mathematical Social Science”, No. 25.
- Bytniewski A., Hernes M. (2014), *Analiza opinii klientów o produkcji dokonywana w kognitywnym zintegrowanym systemie informatycznym zarządzania* [w:] T. Porębska-Miąc, H. Sroka, *Systemy wspomaganie organizacji*, Wydawnictwo Uniwersytetu Ekonomicznego, Katowice.
- Daniłowicz C., Nguyen N.T. (1988), *Consensus-Based Partition in the Space of Ordered Partitions*, „Pattern Recognition”, No. 21.
- Duch W. (2010), *Architektury kognitywne, czyli jak zbudować sztuczny umysł* [w:] R. Tadeusiewicz (red.), *Neurocybernetyka teoretyczna*, Wydawnictwa Uniwersytetu Warszawskiego, Warszawa.
- Dudycz H. (2013), *Mapa pojęć jako wizualna reprezentacja wiedzy ekonomicznej*, Wydawnictwo Uniwersytetu Ekonomicznego, Wrocław.
- Franklin S., Patterson F.G. (2006), *The LIDA Architecture: Adding New Modes of Learning to an Intelligent, Autonomous, Software Agent* [w:] *Proceedings of the International Conference on Integrated Design and Process Technology*, Society for Design and Process Science, San Diego, CA.
- Hernes M. (2015), *Performance Evaluation of the Customer Relationship Management Agent's in a Cognitive Integrated Management Support System* [w:] N.T. Nguyen (ed.), *Transactions on Computational Collective Intelligence*, No. XVIII, Springer, Berlin, ss. 86-104.
- Hernes M., Nguyen N.T. (2007), *Deriving Consensus for Hierarchical Incomplete Ordered Partitions and Coverings*, “Journal of Universal Computer Science”, Vol. 13(2).
- Hernes M., Sobieska-Karpińska J. (2016), *Application of the Consensus Method in a Multiagent Financial Decision Support System*, “Information Systems and e-Business Management”, Vol. 14 (1).

- Jajuga K. (2007), *Podstawy inwestowania na Gieldzie Papierów Wartościowych*, Gielda Papierów Wartościowych S.A., Warszawa.
- Korczak J., Hernes M., Bac M. (2014), *Performance Evaluation of Decision-Making Agents' in the Multi-Agent System* [w:] M. Ganzha, L. Maciaszek, M. Paprzycki (red.), *Proceedings of Federated Conference Computer Science and Information Systems (FedCSIS)*, Annals of Computer Science and Information Systems, Vol. 3, PIPS Press, Warsaw.
- Luo Y., Liu K., Davis D.N. (2002), *A Multi-Agent Decision Support System for Stock Trading*, „IEEE Network”, Vol. 16 (1), ss. 20-27.
- Łapczyński A. (2010), *Systemy ekstrakcji informacji*, Zeszyt Naukowy, nr 814, Wydawnictwo Uniwersytetu Ekonomicznego, Kraków.
- Maleszka M., Mianowska B., Nguyen N.T. (2013), *A Method for Collaborative Recommendation Using Knowledge Integration Tools and Hierarchical Structure of User Profiles*, „Knowledge-Based Systems”, Vol. 47, ss. 1-13.
- Nguyen N.T. (2002), *Metody wyboru consensusu i ich zastosowanie w rozwiązywaniu konfliktów w systemach rozproszonych*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław.
- Nguyen N.T. (2008), *Advanced Methods for Inconsistent Knowledge Management*, Springer Verlag, London.
- Pham L.V., Pham S.B. (2012), *Information Extraction for Vietnamese Real Estate Advertisements*, Fourth International Conference on Knowledge and Systems Engineering (KSE), Danang.
- Potiopa P. (2011), *Metody i narzędzia automatycznego przetwarzania informacji tekstowej i ich wykorzystanie w procesie zarządzania wiedzą*, „Automatyka”, nr 15/2, s. 409-419, <http://journals.bg.agh.edu.pl/AUTOMATYKA/2011-02/Auto40.pdf> (dostęp: 14.11.2015).
- Quillian R. (1963), *A Notation for Representing Conceptual Information: An Application to Semantics and Mechanical English Para-Phrasing*, SP-1395, System Development Corporation, Santa Monica.
- Sobieska-Karpińska J., Hernes M. (2013), *The Postulates of Consensus Determining in Financial Decision Support Systems* [w:] M. Ganzha, L. Maciaszek, M. Paprzycki (red.), *Proceedings of Federated Conference Computer Science and Information Systems (FedCSIS)*, Annals of Computer Science and Information Systems, Vol. 1, PIPS Press, Warsaw.
- Sobieska-Karpińska J., Hernes M. (2014), *Identification of the Knowledge Conflicts' Sources in the Architecture of Cognitive Agents Supporting Decision-Making Process* [w:] M. Ganzha, L. Maciaszek, M. Paprzycki (red.), *Proceedings of Federated Conference Computer Science and Information Systems (FedCSIS)*, Annals of Computer Science and Information Systems, Vol. 3, PIPS Press, Warsaw.
- Soldacki P. (2006), *Zastosowanie metod płytkiej analizy tekstu do przetwarzania dokumentów w języku polskim*, Rozprawa doktorska, Politechnika Warszawska, Warszawa.
- Sycara K.P., Decker K., Zeng D. (2002), *Intelligent Agents in Portfolio Management* [w:] N. Jennings, M. Wooldridge (eds.), *Agent Technology*, Springer, Berlin.

- Westerhoff F.H. (2011), *Multiasset Market Dynamics*, "Macroeconomic Dynamics", No. 8.
- Zgrzywa M. (2007), *Consensus Determining with Dependencies of Attributes with Interval Values*, "Journal of Universal Computer Science", Vol. 13 (2).
- [www 1] <http://ccrg.cs.memphis.edu/> (dostęp: 12.11.2016).

INTEGRATION OF UNSTRUCTURED KNOWLEDGE IN A MULTI-AGENT FINANCIAL DECISION SUPPORT SYSTEM

Summary: Agents running in multi-agent systems more often make decisions based on the unstructured knowledge, acquired from heterogonous sources, therefore decisions presented by these agents' may differ. The decision-making process requires one, the final decision, therefore this knowledge should be automatically integrated. The purpose of this paper is to develop method for the integration of unstructured knowledge in multi-agent financial decision support system. The first part of article presents the issues related to agents' unstructured knowledge representation are described. Next, the developed consensus algorithms have been presented. The last part presents research experiment aimed to verification of the developed method.

Keywords: multi-agent systems, financial decisions, unstructured knowledge integration.