Zeszyty Naukowe
Wyższej Szkoły Bankowej w Poznaniu
Nr 40/2012

Information and communication technology w gospodarce opartej na wiedzy

Wybrane aspekty teoretyczne i aplikacyjne

The Poznan School of Banking
Research Journal
No. 40/2012

Information and Communication Technology in Knowledge Economy

Selected Theoretical and Application Aspects

scientific editor
Piotr Adamczewski

The Poznan School of Banking Press

Zeszyty Naukowe
Wyższej Szkoły Bankowej w Poznaniu
Nr 40/2012

Information and communication technology w gospodarce opartej na wiedzy

Wybrane aspekty teoretyczne i aplikacyjne

pod redakcją naukową
Piotra Adamczewskiego

Wydawnictwo
Wyższej Szkoły Bankowej w Poznaniu

Komitet wydawniczy / Editorial board
Przewodnicząca / Chairperson: *prof. nadzw. dr hab. Beata Filipiak*
Członkowie / Members: *prof. nadzw. dr hab. Władysław Balicki, dr Piotr Dawidziak*, *prof. nadzw. dr hab. Marek Dylewski, Grażyna Krasowska-Walczak* (dyrektor Wydawnictwa WSB w Poznaniu), *prof. nadzw. dr hab. inż. Tadeusz Leczykiewicz, Andrzej Małecki* (sekretarz), *prof. nadzw. dr hab. Ilona Romiszewska, prof. zw. dr hab. Janusz Sawczuk, prof. zw. dr hab. Stanisław Wykrętowicz*

Rada naukowa / Research council
prof. zw. dr hab. Przemysław Deszczyński, prof. nadzw. dr hab. Marek Dylewski, prof. nadzw. dr hab. Beata Filipiak, prof. nadzw. dr hab. Tadeusz Leczykiewicz, prof. zw. dr hab. Jan Szambelańczyk, prof. nadzw. dr hab. inż. Emília Zimková, prof. nadzw. dr hab. inż. Peter Krištofík, prof. nadzw. dr hab. Sergiy Gerasymenko, prof. dr Bernt Mayer, prof. dr Franz Seitz, prof. dr J. Michael Geringer

Czasopismo recenzowane według standardów Ministerstwa Nauki i Szkolnictwa Wyższego.
Lista recenzentów na stronie www.wydawnictwo.wsb.poznan.pl
oraz w ostatnim numerze czasopisma z danego roku.
The journal reviewed in compliance with the Ministry of Science and Higher Education.
The list of peer-reviewers is available at www.wydawnictwo.wsb.poznan.pl
and the most recent issue of the journal in the given year.

Redaktor naczelny czasopisma / Editor-in-chief
prof. nadzw. dr hab. Marek Dylewski

Redaktor naukowy / Scientific editor
dr Piotr Adamczewski

Weryfikacja streszczeń w języku angielskim / Summary reviews in English by
Krzysztof Sajon

Redakcja i korekta / Editing and proofreading
Wojciech Nowakowski

Redakcja techniczna i skład / Typesetting
Sebastian Surendra

Projekt okładki / Cover design
Jan Ślusarski

Wersja pierwotna – publikacja drukowana / Source version – printed publication

© Copyright by Wyższa Szkoła Bankowa w Poznaniu, 2012

ISSN 1426-9724

Wydawnictwo
Wyższej Szkoły Bankowej w Poznaniu
al. Niepodległości 2, 61-874 Poznań
tel. 61 655 33 99, 61 655 32 48
e-mail: wydawnictwo@wsb.poznan.pl, dzialhandlowy@wsb.poznan.pl
www.wydawnictwo.wsb.poznan.pl
Druk i oprawa / Printing and binding: ESUS Druk cyfrowy, Poznań

Spis treści

Wstęp 9

Piotr Adamczewski
E-logistyka jako czynnik rozwoju organizacji inteligentnych w gospodarce opartej na wiedzy 13

Łukasz Balicki
Rynkowe uwarunkowania modelu SaaS 29

Dariusz Ceglarek
Zastosowanie kompresji semantycznej w zadaniach przetwarzania języka naturalnego 39

Wojciech Fliegner
Standaryzacja elektronicznej sprawozdawczości finansowej 65

Jędrzej Musiał
Rozszerzony problem optymalizacji zakupów internetowych 79

Bogdan Pilawski
Narzędzia ETL w zasilaniu repozytoriów danych 91

Maciej Skała, Iga Stróżyk
Zarządzanie procesami i ryzykiem w organizacji z wykorzystaniem systemów informatycznych 105

Tomasz Cichowicz, Michał Frankiewicz, Filip Rytwiński, Jacek Wasilewski, Maciej Zakrzewicz
Odkrywanie anomalii w szeregach czasowych pochodzących z monitoringu systemów teleinformatycznych..... 115

Abstracts 131

Lista recenzentów współpracujących z czasopismem i redaktorzy statystyczni..... 135

Table of Contents

Introduction 9

Piotr Adamczewski
E-logistics as a factor in development of intelligent organization in knowledge society 13

Łukasz Balicki
Market conditions in SaaS-model 29

Dariusz Ceglarek
Applying semantic compression in Natural Language Processing tasks 39

Wojciech Fliegner
Standardization of electronic financial reporting 65

Jędrzej Musiał
Extended version of Internet shopping optimization problem 79

Bogdan Pilawski
Bringing data into the data repositories using ETL tools 91

Maciej Skała, Iga Stróżyk
Integrated process and risk management 105

Tomasz Cichowicz, Michał Frankiewicz, Filip Rytwiński, Jacek Wasilewski, Maciej Zakrzewicz
Anomaly detection in time series for system monitoring 115

Abstracts 131

List of reviewers collaborating with the journal and statistical editors 135

Wstęp

Ranga informacji we współczesnym świecie wynika w głównej mierze z technologii jej przetwarzania. Informacja, jako treść różnych przekazów komunikacyjnych, zawartość przekonań, poglądów, idei, naukowych praw i teorii, ale także sztuki czy nawet doświadczeń religijnych, stanowiła materiał i formę służenia człowiekowi na przestrzeni wieków. Etymologicznie informacja to coś, co pozostawia formę w czymś, czyli „in-forma-tio". Ale dopiero technika zapisu i przekazu myśli uczyniła z różnych tworów doświadczenia człowieka informacje *sensu stricto*. Obrazowo można stwierdzić, że wszystko zaczęło się od śladów ludzkiej ręki, która, idąc za myślą, pozostawiała różne znaki (jako dane) na kamieniu, glinie, papirusie czy papierze. Najpierw więc pojawiło się pismo odręczne, potem druk, aż w końcu cyfrowa postać informacji. To technologie informatyczne kodowania znaków i ich przetwarzanie algorytmiczne uczyniło informację zasobem strategicznym we współczesnym świecie.

Zgodnie z teorią amerykańskiego futurologa Alvina Tofflera ewolucję ludzkości można uporządkować w ramach tzw. trzech fal:

– pierwsza fala – rewolucja rolnicza,

– druga fala – rewolucja przemysłowa,

– trzecia fala – rewolucja informacyjna, w której zainicjowano budowę społeczeństwa informacyjnego.

Termin „społeczeństwo informacyjne" po raz pierwszy został użyty w 1963 r. przez japońskiego antropologa Tadao Umesamo, a spopularyzowany przez japońskiego futurologa Kenichi Koyamę w roku 1968. W końcu lat 70. dotarł do Europy, a w 80. – do Stanów Zjednoczonych. Według Tomasza Goban-Klasa i Piotra Sienkiewicza jest to społeczeństwo, które nie tylko dysponuje rozwiniętymi środkami przetwarzania informacji i komunikowania, lecz środki te są podstawą tworzenia dochodu narodowego i dostarczają źródła utrzymania większości społeczeństwa. Jego cechy można ująć następująco:

– informacja staje się podstawowym zasobem ekonomicznym, środkiem wzrostu i akumulacji dochodu, a także konkurencyjności (produkt cyfrowy),

– informacja w coraz większym stopniu staje się czynnikiem życia społecznego i politycznego,

– ludzie pochłaniają coraz więcej informacji jako konsumenci,

– rosnącą rolę informacji wymusza szybki rozwój sektora środków i usług komunikacyjnych; jednostki – podmioty polityczne oraz ekonomiczne – zużywają coraz więcej informacji, co z kolei determinuje rozrost tego sektora.

Aspekty ekonomiczno-społeczne społeczeństwa informacyjnego można syntetycznie określić następująco:

a) społeczeństwo charakteryzuje się nie tylko rozwojem nowych technologii (zwłaszcza informacyjnych), lecz także nowych sposobów zarządzania i organizacji pracy oraz nowych zawodów,

b) sektor informacji w gospodarce można podzielić na podsektory, zajmujące się odpowiednio:

– tworzeniem – tzw. pracownicy symboliczni,

– przetwarzaniem – sprzęt komputerowy, technologie i oprogramowanie,

– dystrybucją informacji – telekomunikacja, telewizja, światłowody, usługi udostępniania informacji,

c) zasadniczym wymiarem zmian gospodarczych w powiązaniu z rozwojem środków komunikacji jest proces globalizacji; to dzięki nim mogło dojść do „skurczenia się" czasoprzestrzeni, co jest istotą globalizacji (Bauman),

d) społeczeństwo informacyjne charakteryzuje się dehierarchizacją struktur ekonomicznych, co oznacza m.in. odejście od fordyzmu i tayloryzmu,

e) następuje proces decentralizacji produkcji i zarządzania,

f) pojawiają się organizacje wirtualne, np. banki,

g) rozwija się gospodarka sieciowa (*network economy*); logika sieci stała się ważnym elementem „ekonomii chwili" (*now economy*), charakteryzowanej przez zdolność do generowania wiedzy, przesyłania oraz zarządzania informacją i na jej podstawie podejmowania działań w czasie rzeczywistym w skali globalnej,

h) sieć jako forma ekonomicznej organizacji koncentruje się na realizacji specyficznych projektów biznesowych; jednostką procesu produkcyjnego nie jest firma, lecz projekt biznesowy,

i) logika sieci dotyczy także obszaru działań politycznych, związków społecznych oraz kontaktów międzyludzkich (tzw. indywidualizm sieciowy).

Społeczeństwo informacyjne składa się z takich samych elementów, jak społeczeństwa poprzednich epok. Tyle tylko, że systemy te nabierają obecnie innych cech, niż miały wcześniej. Przykładowo, w przeciwieństwie do społeczeństwa przemysłowego, produkcja nie jest związana z miejscem. Ma to różnorakie konsekwencje, np. dla sposobu funkcjonowania gospodarki czy charakteru relacji społecznych.

Powszechne stosowanie technologii teleinformatycznych w organizacjach gospodarczych zasadniczo zmieniło ich struktury wewnętrzne, funkcjonowanie oraz formy współpracy rynkowej. Stały się podstawą społeczeństwa i gospodarki opartej na wiedzy, jako naturalnej ewolucji w łańcuchu rozwoju. Społeczeństwo informacyjne to takie, w którym większość ludzi ma dostęp do informacji przez sieć globalną, jaką jest Internet. Z kolei społeczeństwo wiedzy to takie społeczeństwo, w którym większość ma dość wiedzy i umiejętności, aby z uzyskanej informacji umieć zrobić odpowiedni użytek, a w przypadku organizacji gospodarczych – przełożyć to na efektywniejsze funkcjonowanie w warunkach rynkowych.

Kluczem do transformacji dzisiejszego społeczeństwa w społeczeństwo i gospodarkę opartą na wiedzy jest edukacja realizowana przez całe życie. Wiedza przyrasta w ogromnym tempie, zależności rynkowe stają się coraz bardziej złożone i współzależne, co sprawia, że uczenie się jest konieczne przez całe życie. Kto tego nie zrozumie, nie nadąży za rozwojem społeczeństwa i gospodarki opartej na wiedzy, w wyniku czego może znaleźć się na jego marginesie – wśród tzw. wykluczonych.

Zbiór zawartych w zeszycie artykułów jest pokłosiem cyklicznych seminariów naukowych „Ku modelowi Gospodarki Opartej na Wiedzy GOW" akademickiego środowiska poznańskiego, jakie od 2005 r. odbywają się z inicjatywy Katedry Informatyki Stosowanej Wyższej Szkoły Bankowej w Poznaniu na tej uczelni. Idea tych spotkań zrodziła się z potrzeby holistycznego spojrzenia na problemy rozwiązań teleinformatycznych w ramach gospodarki nowego typu. Stąd wśród autorów znaleźli się zarówno przedstawiciele nauki, jak i reprezentanci praktyki gospodarczej, którzy na co dzień wdrażają, eksploatują i rozwijają takie rozwiązania. Wybrane wystąpienia odzwierciedlają różnorodność problematyki badawczej z zakresu informatyki stosowanej oraz ich aspekty aplikacyjne. Nie wyczerpują one całości zagadnień związanych z gospodarką opartą na wiedzy, ale mogą stanowić istotny przyczynek do jej funkcjonowania i dalszego rozwijania. Z taką nadzieją zeszyt trafia do rąk Szanownych Czytelników.

dr Piotr Adamczewski

Zeszyty Naukowe
Wyższej Szkoły Bankowej w Poznaniu
Nr 40/2012

Piotr Adamczewski
Wyższa Szkoła Bankowa w Poznaniu

E-logistyka jako czynnik rozwoju organizacji inteligentnych w gospodarce opartej na wiedzy

Streszczenie. *Celem artykułu jest ukazanie roli rozwiązań e-logistyki w rozwoju organizacji inteligentnych funkcjonujących w gospodarce opartej na wiedzy. Po ogólnej charakterystyce organizacji inteligentnej odniesiono się do zarządzania wiedzą i na tym tle ukazano istotę rozwiązań e-logistyki w zakresie wybranych rozwiązań informatycznych ze szczególnym uwzględnieniem sieci wartości oraz zastosowań rozwiązań informatycznych klasy ERP. W końcowej części ukazano perspektywy rozwojowe e-logistyki w budowaniu społeczeństwa opartego na wiedzy.*

Słowa kluczowe: *CRM, ERP, e-commerce, e-logistyk, łańcuch logistyczny, SCM, sieć wartości*

1. Wprowadzenie

Globalizacja gospodarki światowej oraz znoszenie barier handlowych, politycznych i ekonomicznych powoduje konieczność szybkich i efektywnych działań skutkujących dostosowaniem działalności organizacji do nowych warunków. Logistyka, będąca podstawowym czynnikiem konkurencyjności organizacji, jest szczególnie podatna na wprowadzanie wszelkiego typu innowacji i nowych idei, które – jeżeli odniosą sukces – mają szansę na zainteresowanie środowisk biznesowych i szybkie wdrożenie. Przekłada się to na finansowanie kolejnych badań nad nowymi technologiami i stanowi samonapędzający się mechanizm poszukiwania nowych rozwiązań innowacyjnych[1].

Współczesne mechanizmy rynkowe cechuje duża dynamika zmian otoczenia gospodarczego. Miarą ich dostosowania jest możliwość budowania przewagi konkurencyjnej organizacji inteligentnych z wykorzystaniem m.in. takich

[1] M. Dolińska, *Innowacje w gospodarce opartej na wiedzy*, PWE, Warszawa 2010.

czynników, jak wiedza czy kapitał intelektualny personelu, które pozwalają im na realizowanie swoich strategii rozwojowych. Kluczową rolę odgrywają tu zaawansowane rozwiązania dotyczące infrastruktury teleinformatycznej, bazującej na ICT (*Information and Communication Technology*), w zakresie wspomagania procesów logistycznych tych organizacji poprzez stosowanie rozwiązań organizacyjno-informatycznych, określanych jako e-logistyka[2]. Technologie te stanowią swoisty ekosystem informatyczny, umożliwiający wdrażanie i efektywne eksploatowanie systemów informatycznych, np. klasy ERP (*Enterprise Resource Planning*) oraz BI (*Business Intelligence*), jako atrybutów organizacji inteligentnych w gospodarce opartej na wiedzy.

2. Zarządzanie wiedzą w organizacji inteligentnej

Organizacja inteligentna to organizacja, która opiera swoją filozofię działania na zarządzaniu wiedzą[3]. Termin ten upowszechnił się w latach 90. XX w. za sprawą rosnącego rozwoju ICT, dynamicznie zmieniającego się otoczenia gospodarczego i wzrostu konkurencyjności rynkowej[4]. O organizacji inteligentnej można mówić wtedy, gdy jest to organizacja ucząca się, mająca zdolności do kreowania, pozyskiwania, organizowania i dzielenia się wiedzą oraz jej wykorzystywania w celu podniesienia efektywności działania oraz zwiększenia konkurencyjności na globalnym rynku[5]. Idea takiej organizacji zasadza się na podejściu systemowym, czyli traktowaniu jej jako złożonego organizmu opartego na istniejących strukturach i realizowanych procesach, ze szczególnym podkreśleniem roli wiedzy. W podejściu tym – nazywanym przez Petera Senge'a „piątą dyscypliną" – dzięki wiedzy i odpowiednim narzędziom wszystkie elementy składowe organizacji oraz jej personel potrafią umiejętnie współdziałać w realizacji określonych celów[6]. W efekcie cała organizacja funkcjonuje jak inteligentny organizm, dobrze sobie radzący w konkurencyjnym otoczeniu. Wyjaśnia on wzajemne związki

[2] Por.: P. Adamczewski, *Gospodarka oparta na wiedzy jako determinanta dla polskich przedsiębiorstw*, w: *Nauka dla gospodarki*, red. C.F. Hales, „Zeszyty Naukowe Uniwersytetu Rzeszowskiego «Nauka dla Gospodarki»" 2010, nr 1; *E-logistyka*, red. W. Wieczerzycki, PWE, Warszawa 2012; *Trendy rozwojowe inteligentnych organizacji w globalnej gospodarce*, PARP, Warszawa 2009.

[3] W.M. Grudzewski, I.K. Hejduk, *Kreowanie w przedsiębiorstwie organizacji intelektualnej*, w: *Przedsiębiorstwo przyszłości*, red. W.M. Grudzewski, J.K. Hejduk, Difin, Warszawa 2000; mfiles.pl/pl/index.php/organizacja_inteligentna [15.06.2012]; R. Orzechowski, *Budowanie wartości przedsiębiorstwa z wykorzystaniem IT*, Oficyna Wydawnicza SGH, Warszawa 2008.

[4] *Trendy rozwojowe inteligentnych organizacji w globalnej gospodarce*, op. cit.

[5] mfiles.pl/pl/index.php/organizacja_inteligentna, op. cit.

[6] P. Senge, *Piąta dyscyplina. Teoria i praktyka organizacji uczących się*, Oficyna Ekonomiczna, Kraków 2002.

między sposobami osiągania celów, ich rozumienia, metodami rozwiązywania problemów i komunikacji wewnętrznej oraz zewnętrznej[7].

Koncepcja organizacji inteligentnej zaczęła się kształtować jako odpowiedź na dynamicznie zmieniające się otoczenie gospodarcze, a w szczególności[8]:

– globalizację rynków, przemiany społeczno-gospodarcze oraz przyspieszenie dynamicznym postępem w zakresie ICT,

– rosnącą konkurencję rynkową, wymuszającą poszukiwanie efektywniejszych metod gospodarowania,

– wysokie tempo rozwoju techniczno-technologicznego,

– postępującą złożoność produktów,

– malejący cykl życia produktów.

Do najważniejszych atrybutów cechujących organizacje inteligentne można zaliczyć[9]:

– szybkość i elastyczność działania,

– umiejętność obserwowania otoczenia,

– zdolność do wczesnego diagnozowania sygnałów rynkowych i reagowania na zmiany w otoczeniu,

– umiejętność szybkiego wdrażania nowych rozwiązań opartych na wiedzy i osiągania dzięki temu korzyści ekonomicznych.

Rosnący wolumen informacji wykorzystywanych w organizacji inteligentnej idzie w parze ze wzrostem jej znaczenia. Już Peter Drucker wskazywał, że tradycyjne czynniki produkcji: ziemia, praca, kapitał, tracą na znaczeniu na rzecz kluczowego zasobu, jakim w kreatywnym funkcjonowaniu organizacji jest wiedza; stanowi ona niematerialne zasoby związane z ludzkim działaniem, których zastosowanie może być podstawą do zdobycia przewagi konkurencyjnej[10]. Wiedzę można traktować jako informację osadzoną w kontekście organizacyjnym i umiejętność jej efektywnego wykorzystania w funkcjonowaniu organizacji. Oznacza to, że zasobami wiedzy są dane o klientach, produktach, procesach, otoczeniu itp. w postaci sformalizowanej (dokumenty, bazy danych) oraz nieskodyfikowanej (wiedza pracowników)[11].

---

[7] I. Becerra-Fernandez, A. Gonzalez, R. Sabherwal, *Knowledge Management. Challenges, Solutions and technologies*, Upper Saddle River, Pearson-Prentice Hall, New York 2004.

[8] M.N. Aydin, M.E. Bakker, *Analyzing IT maintenance outsourcing decision from a knowledge management perspective*, „Information Systems Frontiers" 2008, t. 10; M. Dolińska, op. cit.; *Materiały firmowe firmy Raben*, materiał powielony, Poznań 2012.

[9] W.M. Grudzewski, I.K. Hejduk, op. cit.; *Trendy rozwojowe inteligentnych organizacji w globalnej gospodarce*, op. cit.

[10] P. Adamczewski, *Gospodarka oparta na wiedzy…*; I. Becerra-Fernandez, A. Gonzalez, R. Sabherwal, op. cit.; E. Waltz, *Knowledge Management in the Intelligence Enterprise*, Arteh House, Boston 2003.

[11] P. Adamczewski, *Transfer wiedzy dla wielkopolskiego sektora MSP w perspektywie strategii i-2010*, w: *Transfer wiedzy i funduszu europejskich do sektorów gospodarki krajów UE*, red. nauk. J. Stacharska-Targosz, J. Szostak, Wyd. WSB w Poznaniu, Poznań 2010; I. Becerra-Fernandez, A. Gonzalez, R. Sabherwal, op. cit.; P. Grajewski, *Procesowe zarządzanie organizacją*, PWE, Warszawa 2012.

Zarządzanie wiedzą określa się jako systematyczne i zorganizowane wykorzystywanie jej zasobów do usprawnienia funkcjonowania organizacji, a realizowane jest ono w ramach orientacji procesowej poprzez: lokalizowanie wiedzy, jej pozyskiwanie, gromadzenie, wzbogacanie i rozpowszechnianie. Praktyczny wymiar zarządzania wiedzą na poziomie organizacji inteligentnych może zatem przebiegać w ramach sekwencji procesów[12]:

– lokalizowanie wiedzy,
– selekcja wiedzy,
– kodyfikacja wiedzy,
– przetwarzanie i adaptacja wiedzy na potrzeby kierownictwa organizacji,
– transfer wiedzy,
– tworzenie nowej wiedzy,
– dzielenie się wiedzą,
– aktualizacja wiedzy.

Organizacje tradycyjne, które chcą stać się tzw. organizacjami inteligentnymi, muszą zmienić mentalność pracowników, uporządkować organizację i procesy biznesowe. Oznacza to, że wszystkie procesy (funkcje) organizacji inteligentnych powinny być objęte wysoce zintegrowanym systemem informacyjnym, przy czym nie wolno sprowadzać zagadnienia efektywnego zarządzania wiedzą tylko do wymiaru technologicznego – bardzo ważnego, ale nie decydującego o końcowej efektywności rozwiązań. Oprócz czynników „twardych", związanych z kwestiami techniczno-technologicznymi, równie ważne są czynniki „miękkie", opierające się na kreatywności i potencjale intelektualnym personelu, osadzone w racjonalnych strukturach organizacyjnych i efektywnie zorganizowanych procesach biznesowych.

W świetle powyższych ustaleń na system zarządzania wiedzą składają się następujące elementy[13]:

– strategia zarządzania wiedzą – wskazuje priorytety działań, określa rolę zarządzania wiedzą w realizacji celów strategicznych organizacji,

– ludzie i kultura organizacyjna – gotowość pracowników do dzielenia się wiedzą, wspierana przez kulturę organizacyjną,

– procesy biznesowe – orientacja procesowa organizacji pozwala efektywnie gromadzić, udostępniać i wyszukiwać wiedzę,

– technologia informacyjna – zapewnia użytkownikowi przyjazne gromadzenie, przetwarzanie i udostępnianie informacji.

---

[12] M.N. Aydin, M.E. Bakker, op. cit.; W.M. Grudzewski, I.K. Hejduk, op. cit.; A. Koronios, W. Yeoh, *Critical Success Factors for Business Intelligence Systems*, „Journal of Computer Information Systems", Spring 2010.

[13] A. Koronios, W. Yeoh, op. cit.; R. Magnier-Watanabe, D. Senoo, *The effect of institutional pressures on knowledge management and the resulting innovation*, „International Journal of Intelligent Enterprise" 2009, t. 1, z. 2.

W wymiarze praktycznym efektywne współdziałanie tych elementów oznacza konieczność wykorzystania zaawansowanych rozwiązań teleinformatycznych w ramach e-logistyki. Wykorzystuje ona zarówno innowacje techniczne, technologiczne, jak i organizacyjne, pojawiające się na przestrzeni ostatnich lat. Obejmują one niemal wszystkie sfery działalności logistycznej, począwszy od rozwoju środków transportu i wyposażenia, poprzez organizację i zarządzanie przepływem materiałów i surowców, aż do rozwoju struktur systemów realizujących procesy logistyczne. Ich obszarem działań jest realizacja wirtualnych procesów w środowisku rozległych sieci teleinformatycznych (platformą technologiczną jest najczęściej Internet), mających na celu koordynację i integrację partnerów biznesowych w łańcuchu dostaw.

3. E-logistyka w sieci wartości

Rozwój rozwiązań logistycznych jest uwarunkowany rozwojem systemów wspomagających działalność logistyczną, a w szczególności systemów informatycznych. W obecnym świecie biznesu, gdzie największe organizacje działają globalnie, istotą ich funkcjonowania jest sprawna komunikacja. Dlatego też rozwój technologii informatycznych i telekomunikacyjnych jest tak istotny dla branży logistycznej. Bez odpowiednich informacji żadne procesy logistyczne nie byłyby efektywne. Rosnące znaczenie gospodarki opartej na wiedzy w ramach rynków globalnych determinuje funkcjonowanie coraz bardziej rozbudowanych łańcuchów logistycznych[14].

Znaczenie e-logistyki wzrasta w zawrotnym tempie. Znajomość nowoczesnych technik i technologii zarządzania staje się nieodzowna. Wiedza z zakresu zarządzania połączona ze znajomością uwarunkowań specjalistycznych rozwiązań informatycznych daje synergiczny efekt, przekładający się na wzrost konkurencyjności organizacji.

Każda wdrażana innowacja poprawiająca jakość usług w istniejących łańcuchach logistycznych powinna charakteryzować się[15]:

– pewnością – dostawca winien spełniać wszystkie oczekiwania klienta zgodnie z zamówieniem,

– czasem realizacji – doprecyzowanie czasu działania mającego wpływ na koszty – często istotny czynnik wyboru operatora logistycznego,

– funkcjonalnością działania – uwzględnienie np. możliwości współpracy technicznej,

[14] *E-logistyka*, op. cit.; *Strategie i modele gospodarki elektronicznej*, red. C. Olszak, E. Ziemba, PWN, Warszawa 2007.

[15] Por. M. Dolińska, op. cit.; E. Waltz, op. cit.

– sprawną komunikacją – monitorowanie przepływu ładunków, materiałów, towarów, płatności, zarządzanie informacjami,

– uczciwością (rzetelne przedstawianie możliwości, a następnie wykonywanie usług zgodnie z deklaracjami).

Jako przykłady najważniejszych innowacji w tym zakresie można wskazać m.in.: system produkcji Toyoty, rozwiązania zorientowane na klienta – ECR (*Efficient Consumer Response*) oraz CPFR (*Continous Planning Forecasting and Replenishment*), kontener jako opakowanie zbiorcze, metodę optymalnej wielkości zamówienia – EOQ (*Economic Order Quantity*), linię montażową Forda, systemy monitorowania FedEx, metodę planowania zasobów dystrybucyjnych – DRP (*Distribution Resources Planning*) czy metody automatycznej identyfikacji (kody kreskowe i identyfikacja radiowa RFID)[16]. Rozwiązania te dzisiaj są standardami i obejmują zarówno technologię, jak i organizację przepływów materiałowych.

Jedną z najnowszych koncepcji rozwiązań biznesowych, opierających się na zastosowaniu nowoczesnych narzędzi informatycznych, są sieci wartości (*Value Nets*). Można je scharakteryzować następująco[17]:

– zbudowane są wokół klienta – poszczególne grupy klientów (czy, w pewnych przypadkach, nawet indywidualni klienci) otrzymują dostosowane do ich potrzeb rozwiązania,

– oparte są na współpracy – każda z operacji jest przypisana partnerowi potrafiącemu najlepiej ją zrealizować; znaczna część procesów przekazana jest wyspecjalizowanym usługodawcom (*outsourcing*),

– szybko dostosowują się do zmiennych warunków – mają zdolność do błyskawicznego reagowania na zmiany popytu, szybkiego wprowadzania nowych produktów czy też przekształcenia struktury sieci; wszystko to jest możliwe dzięki elastycznym i skalowalnym systemom zaopatrzenia, produkcji i dystrybucji,

– umożliwiają szybki przepływ produktów i informacji – przekłada się to na krótki czas realizacji zamówień; szybka realizacja idzie w parze z dogodnością warunków dostawy, a przede wszystkim z jej niezawodnością,

– wykorzystują nowoczesne technologie informatyczne, np. w postaci e-zaopatrzenia oraz e-sprzedaży (*e-commerce*), umożliwiające funkcjonowanie sieci wartości.

Przykładami takich sieci wartości są w praktyce funkcjonujące rozwiązania w firmach takich, jak: Cisco, Gateway, Zara, Biogen, Dell czy Apple Computer. Wprowadzone tam rozwiązania przyniosły m.in. obniżenie poziomu zapasów o 80%, a poziom zapasów mierzony w dniach sprzedaży zmalał z 27 do 2 dni[18].

---

[16] *E-logistyka*, op. cit.

[17] Por. J.B. Quinn, *Intelligent Enterprise*, Free Press, New York 1992; P. Senge, op. cit.

[18] R. Orzechowski, op. cit.; *Trendy rozwojowe inteligentnych organizacji w globalnej gospodarce*, op. cit.

Na skutek rozwoju organizacyjnego usługi logistyczne wykonywane przez wyspecjalizowane jednostki zewnętrzne zaczęły występować jako działania oferowane przez niezależnych dostawców w modelu ASP (*Application Service Provider*). W ten sposób ukształtował się rynek outsourcingowy, który nie obciąża producenta bądź przetwórcy kosztami tworzenia, utrzymywania i aktualizowania funkcjonalności aplikacji informatycznych obsługujących relacje biznesowe z partnerami w całym łańcuchu logistycznym. Poza usługami o charakterze zwartym pojawiły się nowoczesne rozwiązania skupiające w sobie szereg działań mających na celu koordynację i integrację sieci złożonej z producentów, hurtowników, detalistów, dystrybutorów oraz firm transportowych i spedycyjnych. Dostawcy usług e-logistycznych mogą organizować cały proces realizacji zamówienia (od jego złożenia po potwierdzenie i zrealizowanie dostawy) – określa się ich wtedy mianem integratorów 4PL (*Fourth Party Logistics*). Mogą również działać jako swoiste e-rynki, które kojarzą dostępne usługi w celu zaspokojenia potrzeb dostawców i odbiorców w łańcuchach dostaw towarów rynkowych. Integratorzy 4PL obsługują rynek B2B (*Business-to-Business*) i kontakty B2C (*Business-to-Customer*). W przypadku e-rynków istotne znaczenie ma możliwość kojarzenia popytu i podaży usług logistycznych w czasie rzeczywistym, na platformie ogólnodostępnych narzędzi internetowych. Stanowi to obecnie najbardziej rozwinięte rozwiązanie w zakresie e-logistyki[19].

Obszar logistyki w organizacjach jest szczególnie podatny na wprowadzanie usług świadczonych drogą elektroniczną. Wynika to z faktu, że stosowane obecnie rozwiązania wykorzystują szereg technologii zabezpieczających przekazywane informacje zarówno pod względem niezmienności ich treści, jak również dających możliwość potwierdzenia otrzymania dokumentu elektronicznego przez system informatyczny partnera biznesowego. Obok zaawansowanych rozwiązań bazujących na elektronicznej wymianie danych EDI (*Electronic Data Interchange*) pojawiło się wiele aplikacji komunikujących się z otoczeniem biznesowym za pośrednictwem standardowej przeglądarki internetowej i portali na stronach www. Portale internetowe oferują funkcjonalności, które stosują zasady właściwe dla giełd (np. giełda wolnych ładunków, pojazdów, szerokiej palety usług logistycznych), porównywarki cen (np. cen paliw, surowców) czy też serwisów informacyjnych (np. serwisy dla kierowców). Bardzo przydatna jest również usługa wyszukiwania połączeń komunikacyjnych, bazujących na systemie lokalizacji satelitarnej, np. GPS (*Global Positioning System*).

Dobrym przykładem ilustrującym takie zaawansowane rozwiązanie teleinformatyczne w ramach e-logistyki może być np. system Integrator, który umożliwia składanie zleceń *online* w firmie Raben. Ten operator logistyczny, mając na uwadze doskonalenie komunikacji z klientami, oddał do ich dyspozycji narzędzie

[19] *E-logistyka*, op. cit.

informatyczne do składania zleceń poprzez stronę internetową[20]. Rozwiązanie to nie tylko pozwala zaoszczędzić czas klienta, ale także umożliwia zdalny serwis i scentralizowane wsparcie. Działanie narzędzia jest proste, a dzięki technologii WEB nie ma problemów związanych z konfiguracją komputerów, wprowadzaniem zmian i aktualizacji. Po wypełnieniu zlecenia klient otrzymuje elektronicznie potwierdzenie jego przyjęcia przez system. Korzystając z Integratora, można wydrukować takie dokumenty, jak listy przewozowe czy etykiety. Istnieje również możliwość monitorowania drogi przesyłki oraz generowania różnorodnych statystyk, bowiem wszystkie niezbędne informacje, od momentu zlecenia odbioru aż po dostawę, znajdują się w jednym miejscu. Jest to niewątpliwa korzyść dla klientów. Istotnym rozwinięcie tego rozwiązania może być tzw. eSMS. Program ten umożliwi wysłanie do odbiorcy, drogą elektroniczną lub SMS-ową, krótkich informacji o przesyłce, np. potwierdzenia odbioru, aktualnego położenia itp. Rozwiązanie to oznacza, że odbiorca przesyłki ewentualne pytania może kierować bezpośrednio do przewoźnika, czyli firmy Raben, a nie do nadawcy przesyłki, który w tym przypadku musiałby najpierw wykonać telefon do firmy transportowej, a następnie kolejny do swojego odbiorcy. Zdecydowanie usprawnia to pracę klientów firmy jako nadawców przesyłek. Na tym jednak nie koniec innowacyjnych rozwiązań planowanych w Raben. W ciągu najbliższych miesięcy zostanie wdrożona „Z-ręczna dostawa" – usługa ręcznego rozładunku. Standardowo operator logistyczny odbiera przesyłki i dostarcza do wskazanego miejsca, jednak nie jest odpowiedzialny za ich ręczny rozładunek. Wspomniana usługa umożliwi zlecenie kierowcy rozładunku i dostarczenia przywiezionych rzeczy na wskazane miejsce, np. bezpośrednio do sklepu w centrum handlowym. Serwis ten jest szczególnie przydatny w tych punktach rozładunku, w których nie ma ramp czy wózków widłowych lub osoba przyjmująca towar nie jest w stanie go samodzielnie przenieść. Usługa „Z-ręczna dostawa" znajduje zastosowanie np. w centrach i galeriach handlowych, szpitalach lub aptekach.

4. Rola systemów ERP w e-logistyce

W coraz bardziej złożonych warunkach gospodarczych cenione są systemy informatyczne zwiększające przychody oraz optymalizujące koszty. Dlatego już od dawna dużym powodzeniem cieszą się systemy planowania zasobów organizacji klasy ERP (*Enterprise Resource Planning*), tak do obsługi klienta, jak i w obszarze zaplecza (*back-office*) nie mającego bezpośredniego przełożenia na procesy sprzedaży towarów i usług. Dobrze skonfigurowany system ERP może być źródłem oszczędności dla dowolnej organizacji, a dodatkowo pozwala szybciej

[20] *Materiały firmowe firmy Raben*, materiał powielony, Poznań 2012.

i w bardziej elastyczny sposób podejmować decyzje. W czasach dekoniunktury gospodarczej zmiany organizacyjne wynikające z prawidłowego wykorzystania zgromadzonych przez organizacje informacji o procesach i zasobach biznesowych mogą być najtańszą metodą ich rozwoju[21].

W ciągu ostatnich lat inwestycje w sprzęt ICT rosły bardzo dynamicznie, co oznacza, że wiele organizacji zdążyło się już wyposażyć w odpowiednią infrastrukturę informatyczną, która może wydajnie pracować przez kilka najbliższych lat. Teraz mogą więc one skupić się na zakupie oprogramowania biznesowego, takiego jak ERP. Podstawą osiągania sukcesu w biznesie jest umiejętność planowania i konsekwentnej realizacji celów biznesowych. Zadanie to jest tym trudniejsze, im szybciej rozwija się organizacja inteligentna. System klasy ERP to system informatyczny integrujący wszystkie aspekty działania przedsiębiorstwa. Zaawansowane systemy ERP umożliwiają nie tylko gromadzenie danych dotyczących bieżącej działalności, ale przede wszystkim przekształcanie ich w wiedzę niezbędną do podejmowania trafnych decyzji biznesowych. Z kolei te organizacje, które eksploatują już system ERP, powinny inwestować w moduły zwiększające jego możliwości. Wśród najczęściej wskazywanych są rozwiązania do zarządzania procesem sprzedaży oraz zakupami, bo pozwalają one na ujednolicenie procesu zakupów, a także skorzystanie z efektu skali, istotnego zwłaszcza w przypadku organizacji o rozproszonej infrastrukturze. Warto też skoncentrować się na lepszym wykorzystaniu i rozwoju modułów usprawniających zarządzanie finansami oraz funkcjonalności z zakresu CRM (*Customer Relationship Management* – zarządzanie kontaktami z klientami), SCM (*Supply Chain Management* – zarządzanie łańcuchem dostaw) i HRM (*Human Resource Management* – zarządzanie zasobami ludzkimi). Z drugiej strony – organizacje, które zdecydują się na odważne działania konkurencyjne, muszą dysponować narzędziami umożliwiającymi prowadzenie szczegółowych analiz informacji pochodzących z rynku[22].

Stosowanie narzędzi inteligencji biznesowej BI (*Business Intelligence*) pozwala na lepsze poznanie preferencji klientów oraz analizowanie wyników sprzedaży w celu eliminowania mniej dochodowych produktów i działań[23]. Analizy tworzone na podstawie informacji agregowanych przez systemy ERP są podstawą większości inicjatyw biznesowych w wielu organizacjach. Przydatne mogą okazać się też najprostsze nawet rozwiązania umożliwiające szacowanie ryzyka operacyjnego i ograniczanie ewentualnych zagrożeń, wynikających z problemów

---

[21] P. Adamczewski, *Holistyczne ujęcie uwarunkowań ICT w inteligentnych organizacjach społeczeństwa informacyjnego*, „Zeszyty Naukowe Uniwersytetu Rzeszowskiego" 2012 (w druku); E. Wang, C. Lin, J. Jiang, G. Klein, *Improving ERP fit to organizational process through knowledge transfer*, „International Journal of Information Management" 2007, nr 34, s. 134-153.

[22] P. Adamczewski, *Transfer wiedzy…*; M.N. Aydin, M.E. Bakker, op. cit.

[23] P. Adamczewski, *Holistyczne ujęcie uwarunkowań…*; J.N. Luftman, *Competing in the Information Age. Align in the Sand. Second Edition*, Oxford University Press, New York 2003.

organizacji znajdujących się w obrębie wspólnego łańcucha dostaw. Kryzys gospodarczy przyczyni się bowiem do zacieśnienia powiązań między przedsiębiorstwami skupionymi w ramach łańcuchów dostaw ze względu na konieczną wymianę usług i integrację procesów – przyniesie to dodatkowe korzyści w ramach efektu synergii. Analiza działalności organizacji jest kluczowym elementem zarządzania strategicznego. Dysponując pełną wiedzą, organizacja może podejmować trafne decyzje i w konsekwencji poprawiać swoją pozycję konkurencyjną. Dzięki błyskawicznemu dostępowi do aktualnych danych zarząd/dyrekcja dysponuje wiedzą pozwalającą podnosić efektywność pracy poszczególnych działów organizacji, a w sytuacji dużej konkurencji na danym rynku to właśnie decyzje z zakresu zarządzania wpływają na pozycję rynkową.

System ERP powinien być dopasowany do potrzeb organizacji, te zaś mogą być różne w zależności od wielkości przedsiębiorstwa i specyfiki branży. Mniejsze organizacje, np. z sektora MSP, czyli małych i średnich przedsiębiorstw, często potrzebują przystępnych cenowo narzędzi udostępniających najważniejsze funkcje analiz biznesowych. W takim przypadku niezwykle przydatne jest pełne zintegrowanie z wykorzystywanym oprogramowaniem biurowym, np. z pakietem MS Office czy kodami kreskowymi. Ułatwia to proces rejestracji i gromadzenia danych na poziomie wszystkich użytkowników systemu.

Prężnie rozwijające się przedsiębiorstwa przykładają większą wagę do elastycznych i nowoczesnych rozwiązań informatycznych o poszerzonych funkcjach analitycznych. Moduły analityczne powinny umożliwiać szybki dostęp do aktualnych danych, raportowanie i porównywanie wyników przedsiębiorstwa. Oznacza to, że systemy ERP muszą być wyposażone w standardowe raporty, ale również w łatwe ich generowanie z uwagi na potrzeby użytkownika końcowego. Istotną funkcjonalnością systemu powinno być także uzyskanie dostępu do kontekstowych informacji ważnych dla różnych użytkowników, co gwarantowałoby skoordynowanie codziennych działań logistycznych z ogólną strategią przedsiębiorstwa.

Rozważając wdrożenie nowoczesnego systemu ERP, należy brać pod uwagę zmiany, jakim podlega organizacja, choćby te związane z jej rozwojem, zatrudnieniem, rosnącymi wymaganiami czy poszerzaniem rynków zbytu. Dlatego warto decydować się na elastyczne systemy umożliwiające szybką modyfikację i dodawanie nowych komponentów pozwalających na dostosowanie się do indywidualnych oczekiwań użytkownika. Przemyślana decyzja dotycząca wybranego systemu ERP umożliwi znaczącą oszczędność w przyszłości, gdy wzrosną potrzeby przedsiębiorstwa w tym zakresie. Stąd wybrany system ERP powinien być wystarczająco skalowalny i elastyczny. Powinien też cechować się maksymalnie uproszczonym interfejsem obsługi, a najlepiej – być dostępnym przez dowolną przeglądarkę internetową. Wreszcie powinien dać się szybko wdrożyć i pozwalać na proste modyfikacje bez konieczności ingerencji w kod źródłowy. A to oznacza,

że powinien pochodzić od uznanego i sprawdzonego dostawcy, który zagwarantuje nie tylko dobry produkt, ale także metodologię sprawnego jego wdrożenia i dalszego rozwoju. W okresie pogłębiającego się globalnego kryzysu gospodarczego, a jednocześnie rozrastających się łańcuchów dostaw dla nowocześnie funkcjonujących organizacji, zdanie się na zaawansowane rozwiązania informatyczne staje się wręcz nakazem chwili.

Przed nowym wyzwaniem stają pozostałe technologie informatyczne, np. z zakresu automatycznej identyfikacji, łączności bezprzewodowej czy lokalizacji satelitarnej[24]. Analitycy branżowi oceniają, że właśnie zaawansowane rozwiązania informatyczne mogą odegrać istotną rolę w walce z kryzysem i jego skutkami. Powszechnie panująca moda na architekturę opartą na usługach SOA (*Service Oriented Architecture*), wirtualizację i WEB 2.0 może się okazać jednym z czynników rozwoju inwestycji dobrze powiązanych z procesami biznesowymi.

Sytuacja na rynkach finansowych oraz mało optymistyczne prognozy gospodarcze sprawiają, że wzrasta znaczenie optymalizacji infrastruktury i organizacji procesów biznesowych pod kątem zwiększania efektywności i redukowania kosztów prowadzenia działalności. Architektura zorientowana na usługi oraz wirtualizacja to rozwiązania mające coraz szersze zastosowanie. Jednak największe korzyści wynikają z odpowiedniego połączenia tego typu rozwiązań z procesami biznesowymi i kulturą organizacyjną. Po raz kolejny powraca zatem aspekt powiązania wymiaru technologii informatycznych i biznesu. Z dotychczasowych doświadczeń wdrożeniowych wynika, że największą barierą w skutecznym przekształcaniu architektury systemów w model usługowy jest brak zaangażowania ze strony pracowników odpowiedzialnych za kształtowanie biznesu. Można postawić tezę, że kryzys gospodarczy staje się dobrym pretekstem do zmiany podejścia do filozofii SOA[25].

Wirtualizacja znalazła stałe miejsce we współczesnej infrastrukturze informatycznej. Wirtualne serwery, dyski i sieci LAN (*Local Area Network*) zagościły w większości nowoczesnych przedsiębiorstw, dzięki czemu można optymalnie wykorzystać moce obliczeniowe. Przez wiele lat technologia i praktyka wymuszały zwiększanie liczby wykorzystywanych serwerów. Wynikało to z konieczności rozdzielenia aplikacji pomiędzy różne komputery ze względu na niekompatybilność i specyficzne wymagania dotyczące wersji systemu operacyjnego. Wymuszały to również względy bezpieczeństwa czy niezgodność wykorzystywanych aplikacji z nowymi wersjami systemów operacyjnych.

Wirtualizacja szturmem zdobywa nowe rzesze użytkowników: wprawdzie zazwyczaj nie prowadzi wprost do zmniejszenia liczby instalacji systemów operacyjnych, ale pozwala zmniejszyć liczbę wykorzystywanych serwerów oraz

[24] *E-logistyka*, op. cit.
[25] A. Koronios, W. Yeoh, op. cit.

zdecydowanie poprawić ich wydajność. Ponadto rozwiązuje problem niezgodności najnowszego sprzętu ze starymi wersjami systemów operacyjnych. Nic dziwnego, że wirtualizacja zasobów informatycznych jest postrzegana przez decydentów jako doskonała technologia umożliwiająca efektywniejsze prowadzenie biznesu. Argumenty nasuwają się same:

– dzięki wykorzystaniu maszyn wirtualnych służby informatyczne elastyczniej reagują na wymagania działów biznesowych (szybkie i łatwe wprowadzanie zmian w środowisku informatycznym),

– część aplikacji korzysta ze starych, niewspieranych wersji systemów operacyjnych, np. Microsoft Windows NT 4.0 Server, Novell NetWare 4.x, SCO Unix itp. Zdarza się, że aplikacje takie nie pracują poprawnie po zainstalowaniu nowej wersji systemu operacyjnego. Jeżeli system operacyjny nie jest już wspierany przez producenta, brakuje sterowników do nowych generacji serwerów. Gdy zachodzi konieczność przeniesienia takiej aplikacji na nową platformę sprzętową, to wirtualizacja jest jedynym rozwiązaniem,

– koszty utrzymania środowiska informatycznego zmniejszają się dzięki efektywniejszemu wykorzystaniu fizycznych serwerów. Daje to oszczędności na kosztach zasilania, klimatyzowania i wsparcia technicznego,

– odpowiednio zaprojektowane środowisko wirtualne może też skutecznie zabezpieczyć dostęp do danych i zmniejszyć ryzyko operacyjne.

Innym sposobem ograniczania kosztów związanych z utrzymaniem rozwiązań informatycznych jest m.in. zastosowanie energooszczędnych urządzeń i względnie taniego oprogramowania, dostępnego na zasadzie licencji programowania otwartego (*open source*). Wreszcie sposobem na zmniejszenie wydatków na ICT może okazać się *outsourcing*, tak usług, jak i oprogramowania, w modelu SaaS (*Software as a Service*), a nawet całych procesów biznesowych.

Już lata 90. dobitnie wykazały, że bez systemu klasy ERP nie ma nowoczesnego zarządzania w organizacji inteligentnej. Tradycyjnie rozumiane systemy ERP już nie wystarczają – ich podstawowa funkcjonalność została wzbogacona o moduły CRM (*Customer Relationship Management*), SRM (*Supplier Relationship Management*), SCM (*Supply Chain Management*) i PLM (*Product Lifecycle Management*)[26]. Zwłaszcza te ostatnie rozszerzenia zyskują na znaczeniu. Zarządzanie cyklem życia wyrobu obejmuje działania począwszy od momentu pojawienia się idei wyrobu aż po jego wycofanie z rynku. Składa się na to opracowanie koncepcji projektu, opracowanie technologii wytwarzania, zarządzanie

[26] P. Adamczewski, *Systemy ERP-BI w rozwoju organizacji inteligentnej*, w: *Kreatywność i systemy inteligencji biznesowej jako przedmiot badań ekonomicznych*, Wyd. Uniwersytetu Ekonomicznego w Katowicach, Katowice 2012; Idem, *Rozwinięte systemy klasy ERP w inżynierii wiedzy*, w: *Wiedza i komunikacja w innowacyjnych organizacjach. Systemy ekspertowe – wczoraj, dziś, jutro*, red. J. Gołuchowski, B. Filipczyk, Wyd. Uniwersytetu Ekonomicznego w Katowicach, Katowice 2010.

wytwarzaniem, zarządzanie dokumentacją i zamówieniami klientów. Istotnym elementem w systemie PLM jest obsługa zmian technicznych wyrobów w procesach produkcji i zaopatrzenia. W przypadku produkcji wielkoseryjnej z dużą liczbą wariantów, kiedy klient może określać własne życzenia co do modelu wyrobu i jego wyposażenia, istotne jest zastosowanie konfiguratora produktu. Pozwala on na tworzenie modelu produktu, dokumentacji wykonawczej i zestawień materiałów oraz szacowanie kosztów. Możliwe jest to za sprawą współdziałania z pakietami klasy CAD/CAM (*Computer Aided Design*/*Computer Aided Manufacturing*).

Najnowsze wersje ERP w pełni wykorzystują ostatnie rozwiązania technologii informatycznych, w tym również wspomnianą koncepcję SOA. Usługa jest tu rozumiana jako odrębny moduł funkcjonalny i traktowana na zasadzie elementu rozwiązania informatycznego realizującego konkretne zadanie. Niezależność takich usług pozwala na ich wykorzystywanie w ramach dowolnej platformy systemowej i języka programowania. Daje to niespotykane do tej pory możliwości w zakresie elastyczności działania i rozbudowy rozwiązań informatycznych. Powiązane ze sobą łańcuchami dostaw organizacje inteligentne obsługują strumienie materiałów i surowców, półfabrykatów i produktów gotowych oraz towarzyszących tym procesom informacji. Do realizacji tych zadań w sposób uporządkowany i powtarzalny wykorzystuje się systemy przepływu pracy (*workflow*), a wspomagane filozofią SOA pozwalają na urzeczywistnianie idei przedsiębiorstwa rozszerzonego w konwencji RTE (*Real-Time Enterprise*), czyli działającego w czasie rzeczywistym. Cele stawiane przed takimi rozwiązaniami można ująć następująco[27]:

– zarządzanie transakcjami w ramach branżowego łańcucha dostaw,

– planowanie i realizacja dostaw dokładnie na czas (*Just-in-Time*),

– spełnianie branżowych kryteriów łańcucha dostaw (monitorowanie produktów we wszystkich fazach jego powstawania),

– oferowanie szczegółowych analiz rentowności i obsługi klientów wraz z elastycznym raportowaniem.

Zgodnie ze wcześniejszymi zapowiedziami analityków branży informatycznej rośnie uznanie znaczenia w Polsce systemów klasy ERP w nowocześnie funkcjonujących organizacjach. Wyraża się to m.in. we wzroście sprzedaży tych systemów i liczbie ich efektywnych wdrożeń. Minione lata wyraźnie wskazują, że po zinformatyzowaniu wewnętrznych procesów logistycznych organizacje koncentrują się na informatycznym wspomaganiu kanałów dostaw i sprzedaży, a więc podążają w kierunku pełnej e-logistyki. Rosnąca skala wdrożeń systemów klasy ERP również w Polsce świadczy dobitnie, że hasło „ERP podstawą nowocześnie funkcjonującej firmy" przestało być tylko dyskutowane, ale stanowi decydującą determinantę sukcesów biznesowych w dobie gospodarki opartej na wiedzy.

[27] Por. *E-logistyka*, op. cit.; R. Magnier-Watanabe, D. Senoo, op. cit.

5. Kierunki rozwoju e-logistyki

Rozwój zaawansowanych systemów ERP rozbudza zapotrzebowanie na wspomaganie wspomnianych już informatycznych narzędzi analitycznych w zakresie inteligencji biznesowej. Rozwiązania te przekładają się już na efektywne wspomaganie procesów decyzyjnych. Coraz częściej mówi się o tzw. analityce biznesowej (*Business Analytics*)[28]. Obejmuje ona narzędzia i aplikacje do analizowania, monitorowania, modelowania, prezentowania oraz raportowania danych ułatwiających podejmowanie decyzji. W tym celu wykorzystuje się hurtownie danych, analizy operacyjne łańcuchów dostaw, analityczne systemy CRM, pogłębione analizy finansowe i wskaźniki wydajności organizacji inteligentnych. Użytkownikiem takich rozwiązań jest szczebel strategiczny organizacji, bazujących na pewnych agregatach danych. Wiąże się z tym problem integracji i synchronizacji danych. Integracja danych rozpoczyna się od możliwości wykorzystywania wielu źródeł danych – zarówno poprzez specjalne interfejsy, jak i przy użyciu standardowych mechanizmów typu ODBC (*Open DataBase Connectivity*). Źródłami danych mogą być relacyjne lub hierarchiczne bazy danych, pliki strukturalne, a także systemy ERP. Połączenia te powinny zatem umożliwiać nie tylko odczyt danych, ale także ich zapis i przetwarzanie. W większości organizacji występuje przypadek wielu środowisk informatycznych i mechanizmy dostępu powinny pozwalać na sięganie do danych znajdujących się na różnych platformach (w miarę możliwości bez stosowania plików pośrednich).

Oczekiwania wobec e-logistyki, wynikające z okresu dekoniunktury gospodarczej i działań naprawczych, można ująć następująco:

– nie ma w kryzysie „nowej" ekonomii bez „starej" ekonomii; pojawiają się określenia *new economy* oraz *now economy*, tłumaczone jako ekonomia chwili, stanowiąca kwintesencję działania organizacji inteligentnych w czasie rzeczywistym,

– „stara" ekonomia musi brać udział w tworzeniu docelowych rozwiązań e-logistyki: redukcja kosztów, ale to nie wszystko – wyzwaniem staje się redukcja czasu,

– istotna jest umiejętność transformacji procesów biznesowych na bazie zarządzania łańcuchem wiedzy KCM (*Knowledge Chain Management)*,

– docelowo konieczna jest pełna integracja procesów organizacji inteligentnej z procesami kontrahentów, czyli w całym łańcuchu dostaw SCM,

– organizacje inteligentne zdobywają przewagę konkurencyjną w społeczeństwie informacyjnym poprzez inwestowanie w zasoby niematerialne, tj. w wiedzę i kapitał intelektualny wspomagane zaawansowanymi rozwiązaniami informatycznymi,

[28] P. Adamczewski, *Holistyczne ujęcie uwarunkowań…*; R. Orzechowski, op. cit.

– w nowocześnie funkcjonujących organizacjach gra biznesowa toczy się w przestrzeni wyznaczanej przez wektory globalizacji, wirtualizacji oraz zarządzania wiedzą na poziomie zarządzania logistycznego wspomaganego e-logistyką,

– pod wpływem dynamicznego rozwoju e-logistyki konieczne staje się modyfikowanie dotychczasowych procesów i rekonfigurowanie modeli biznesu w całych łańcuchach dostaw,

– tworzenie rozwiązań e-logistyki staje się wyróżnikiem nowocześnie działających organizacja doby gospodarki opartej na wiedzy.

6. Podsumowanie

Zapotrzebowanie na zaawansowane technologie teleinformatyczne wspomagające procesy logistyczne jako podstawowe elementy e-logistyki będzie w dalszym ciągu wzrastało, bowiem organizacje inteligentne – z samej istoty działań gospodarczych – są zainteresowane optymalnym wykorzystywaniem swoich zasobów dla osiągnięcia maksymalnych korzyści z zainwestowanego kapitału. Coraz bogatsza oferta na polskim rynku rozwiązań ICT pozwala organizacjom dokonywać wyborów w zależności od potrzeb biznesowych i zasobności finansowej, a informatyczne wspomaganie całego łańcucha dostaw staje się już nie tylko wyzwaniem konkurującego rynku, ale wręcz koniecznością w celu sprostania coraz wyższym wymaganiom klientów w efektywnej ich obsłudze. Przy porównywalnych technologiach produkcyjnych i informacyjnych źródeł przewagi konkurencyjnej należy szukać w sprawnie zaprojektowanych i efektywnych łańcuchach e-logistyki organizacji inteligentnych, co nabiera szczególnego znaczenia przy rosnących wymaganiach mechanizmów rynkowych doby gospodarki opartej na wiedzy.

Literatura

Adamczewski P., *E-business applications in polish SME sector – condition and development*, „Studia Informatica" 2011, nr 2B(97), t. 32.

Adamczewski P., *Evolution in ERP – expanding functionality by BI-modules in Knowledge-based Management Systems*, w: *Information Management ICIM*, red. B. Kubiak, Gdansk University Press, Gdansk 2009.

Adamczewski P., *Gospodarka oparta na wiedzy jako determinanta dla polskich przedsiębiorstw*, w: *Nauka dla gospodarki*, red. C.F. Hales, „Zeszyty Naukowe Uniwersytetu Rzeszowskiego «Nauka dla gospodarki»" 2010, nr 1.

Adamczewski P., *Holistyczne ujęcie uwarunkowań ICT w inteligentnych organizacjach społeczeństwa informacyjnego*, „Zeszyty Naukowe Uniwersytetu Rzeszowskiego" 2012 (w druku).

Adamczewski P., *ICT in enterprise architecture of e-companies in light of studies on the sector of SME in Wielkopolska*, „AITM'08. Research Papers", nr 35, Wroclaw University of Economics, Wroclaw 2008.

Adamczewski P., *Rozwinięte systemy klasy ERP w inżynierii wiedzy*, w: *Wiedza i komunikacja w innowacyjnych organizacjach. Systemy ekspertowe – wczoraj, dziś, jutro*, red. J. Gołuchowski, B. Filipczyk, Wyd. Uniwersytetu Ekonomicznego w Katowicach, Katowice 2010.

Adamczewski P., *Strukturalne ujęcie ERP w systemie zarządzania wiedzą w organizacji*, w: *Technologie wiedzy w zarządzaniu publicznym '09*, red. J. Gołuchowski, A. Frączkiewicz-Wronka, Wyd. Akademii Ekonomicznej w Katowicach, Katowice 2009.

Adamczewski P., *Systemy ERP-BI w rozwoju organizacji inteligentnej*, w: *Kreatywność i systemy inteligencji biznesowej jako przedmiot badań ekonomicznych*, Wyd. Uniwersytetu Ekonomicznego w Katowicach, Katowice 2012.

Adamczewski P., *Transfer wiedzy dla wielkopolskiego sektora MSP w perspektywie strategii i-2010*, w: *Transfer wiedzy i funduszu europejskich do sektorów gospodarki krajów UE*, red. nauk. J. Stacharska-Targosz, J. Szostak, Wyd. WSB w Poznaniu, Poznań 2010.

Aydin M.N., Bakker M.E., *Analyzing IT maintenance outsourcing decision from a knowledge management perspective*, „Information Systems Frontiers" 2008, t. 10.

Becerra-Fernandez I., Gonzalez A., Sabherwal R., *Knowledge Management: Challenges, Solutions and technologies*, Upper Saddle River, Pearson-Prentice Hall, New York 2004.

Dolińska M., *Innowacje w gospodarce opartej na wiedzy*, PWE, Warszawa 2010.

E-logistyka, red. W. Wieczerzycki, PWE, Warszawa 2012.

Grajewski P., *Procesowe zarządzanie organizacją*, PWE, Warszawa 2012.

Grudzewski W.M., Hejduk I.K., *Kreowanie w przedsiębiorstwie organizacji intelektualnej*, w: *Przedsiębiorstwo przyszłości*, red. W.M. Grudzewski, J.K. Hejduk, Difin, Warszawa 2000.

Koronios A., Yeoh W., *Critical Success Factors for Business Intelligence Systems*, „Journal of Computer Information Systems", Spring 2010.

Luftman J.N., *Competing in the Information Age. Align in the Sand. Second Edition*, Oxford University Press, New York 2003.

Magnier-Watanabe R., Senoo D., *The effect of institutional pressures on knowledge management and the resulting innovation*, „International Journal of Intelligent Enterprise" 2009, t. 1, z. 2.

Materiały firmowe firmy Raben, materiał powielony, Poznań 2012.

mfiles.pl/pl/index.php/organizacja_inteligentna.

Orzechowski R., *Budowanie wartości przedsiębiorstwa z wykorzystaniem IT*, Oficyna Wydawnicza SGH, Warszawa 2008.

Quinn J.B., *Intelligent Enterprise*, Free Press, New York 1992.

Senge P., *Piąta dyscyplina. Teoria i praktyka organizacji uczących się*, Oficyna Ekonomiczna, Kraków 2002.

Strategie i modele gospodarki elektronicznej, red. C. Olszak, E. Ziemba, WN PWN, Warszawa 2007.

Trendy rozwojowe inteligentnych organizacji w globalnej gospodarce, PARP, Warszawa 2009.

Waltz E., *Knowledge Management in the Intelligence Enterprise*, Arteh House, Boston 2003.

Wang E., Lin C., Jiang J., Klein G., *Improving ERP fit to organizational process through knowledge transfer*, „International Journal of Information Management" 2007, nr 34, s. 134-153.

Zeszyty Naukowe
Wyższej Szkoły Bankowej w Poznaniu
Nr 40/2012

Łukasz Balicki
eZarządzanie Sp. z o.o. w Poznaniu

Rynkowe uwarunkowania modelu SaaS

Streszczenie. *W artykule omówiono sposób skutecznego zwiększania liczby klientów organizacji (a przez to i jej dochodów) dzięki wykorzystaniu informatycznego wsparcia w modelu SaaS (*Software as a Service*). Opisano etapy przygotowania i doboru strategii marketingowej oraz kluczowe aspekty skutecznego oferowania i utrzymania usług w architekturze SaaS. Szczegółowo omówiono technikę odwróconego ryzyka oferowania usług w tym modelu.*

Słowa kluczowe: *SaaS, outsourcing informatyczny, e-usługi, technika odwróconego ryzyka*

1. Istota i zakres modelu SaaS

Model SaaS odcisnął piętno na rynku organizacji i usług IT na świecie w ostatnim dziesięcioleciu. Jest on nierozerwalnie związany z coraz większą popularnością Internetu – nie tylko jako narzędzia do zdobywania informacji, ale również do prowadzenia biznesu. SaaS oznacza brak konieczności zakupu licencji i instalowania oprogramowania na stacji roboczej na korzyść wynajmu usługi dostarczanej za pośrednictwem globalnej sieci. W Polsce tendencji tej dodatkowo sprzyjają programy wsparcia ze środków finansowych Unii Europejskiej, czyli Program Operacyjny Innowacyjna Gospodarka 8.1, który jest niczym innym, jak wsparciem dla nowych organizacji uruchamiających aplikacje w modelu SaaS, nazwanych e-usługami[1].

Rozważając zapotrzebowanie na usługi w funkcjonowaniu każdej firmy, czas pracy zatrudnionych oraz niezawodność narzędzi pracy, takich jak urządzenia,

[1] http://mysteryshop.org [2.06.2012]; T. Kopczyński, *Outsourcing w zarządzaniu przedsiębiorstwem*, PWE, Warszawa 2010.

samochody czy oprogramowanie, przekłada się wprost na generowane przychody. Jak widać z pragmatyki gospodarczej, pełna dostępność usługi/oferenta to klucz do sukcesu rynkowego. W kontekście najbardziej rozpowszechnionych e-usług najlepiej sytuację można zilustrować na przykładzie bankowości elektronicznej. W razie awarii serwisu danego banku użytkownik domowy może zalogować się ponownie, np. wieczorem, i nie będzie ona dla niego dużym problemem. Natomiast w przypadku firmy niemożność wykonania na czas przelewu dla kontrahenta może skutkować karami umownymi, niewypłacenie wynagrodzeń na przełomie miesiąca oznacza duże zawirowania księgowe i niezadowolenie pracowników, brak płatności podatkowych, karne odsetki itd. Wspomniana charakterystyka definiuje jeden z możliwych sposobów podziału e-usług pod kątem ich potencjalnych odbiorców, ale także cech, jakimi muszą zostać opisane. Kryterium dostępności staje się jednym z wielu czynników koniecznych do określenia już podczas projektowania e-usługi. Jedynie właściwie przygotowana e-usługa przyniesie marketingowy sukces[2].

Za miarę powodzenia marketingowego projektu polegającego na wprowadzeniu na rynek usługi typu SaaS uważa się często liczbę korzystających z niej klientów. W celu wprowadzenia usługi na rynek trzeba wykonać wiele działań sprowadzających się do wywołania określonych zachowań u potencjalnych (uprzednio zdefiniowanych) klientów. Oczywisty jest fakt, że swoim działaniem możemy wywoływać pozytywne lub negatywne emocje[3]. Mechanizm ich działania jest różny, a negatywne emocje nie zawsze są złe. W skrócony sposób można to działanie podsumować jako konieczność ciągłego oddziaływania na klientów emocjami pozytywnymi oraz okresowym umiejętnym wywoływaniem u nich emocji negatywnych w celu ich mobilizacji (tab. 1).

Tabela 1. Emocje w zachowaniu klientów

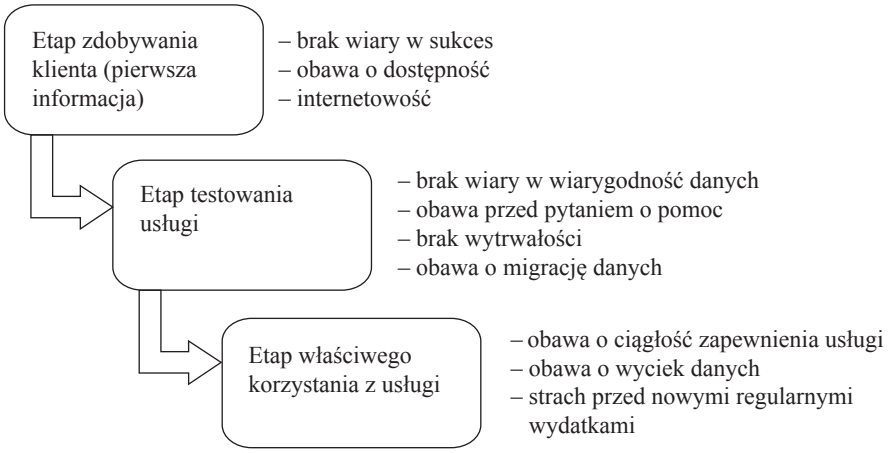
Emocje pozytywne	Emocje negatywne
Powodują chęć do ponownego przyjścia Powtarzanie tych samych czynności wywołuje uczucie nasycenia Są krótkotrwałe. Gdy ustaną, można szybko reagować	Powodują chęć przerwania aktywności Ograniczone i kontrolowane mogą wzmagać mobilizację Żyją „dłużej" niż pozytywne. Potrzeba więcej wysiłków, aby je przezwyciężyć

Ź r ó d ł o: opracowanie własne.

Wszystkie działania mające na celu zdobycie nowego klienta układają się w proces zwany ścieżką działania. Dla każdego typu prowadzonego biznesu ścieżka ta wygląda inaczej. Odnosząc się do organizacji oferującej usługi w modelu

[2] J.L. Bravard, R. Morgan, *Inteligentny outsourcing*, MT Biznes, Warszawa 2010.
[3] M. Kłos, *Outsourcing w polskich przedsiębiorstwach*, CeDeWu, Warszawa 2010.

SaaS, można wyróżnić trzy etapy: zdobywania klienta, testowania usługi oraz właściwego korzystania z niej. Na każdym z etapów pojawiają się różne emocje, powstrzymujące jej użytkowników przed korzystaniem z takich usług (rys. 1).



Rys. 1. Emocje na poszczególnych etapach zdobywania klienta

Ź r ó d ł o: opracowanie własne.

Znając powyższe obawy, wynikające z emocji, jakim podlega klient, można opracować właściwą strategię postępowania, różną na poszczególnych etapach pracy z klientem. Co ciekawe, argumenty finansowe pojawiają się na końcu, już podczas normalnej pracy z usługą. Oznacza to, że konkurowanie ceną oraz eksponowanie tego faktu jako największej korzyści nie powinno być kluczowym elementem przygotowywanej strategii sprzedaży.

2. Technika odwróconego ryzyka

Każdy konsument zadaje sobie pytania: Czy ta usługa jest dla mnie odpowiednia? Czy jest warta swojej ceny? Boi się, że straci czas, dane czy pieniądze. Występują w tym przypadku również czynniki psychologiczne i emocjonalne. Sposób rozwiązania tego problemu ilustruje tzw. technika odwróconego ryzyka, którą spotyka się na co dzień. Jej istota sprowadza się do następujących założeń[4]:

– klient otrzymuje gwarancję zwrotu pieniędzy w przypadku braku spełnienia oczekiwań w określonym czasie (np. 30 dni),

[4] M. Sobińska, *Zarządzanie outsourcingiem informatycznym*, Wyd. Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław 2011.

– klient towar otrzyma od nas gratis, jeżeli nie dostarczymy go w wyznaczonym terminie (np. 5 dni),

– klient czekający na dostawę dłużej niż zadeklarowany czas otrzymuje go od nas w standardzie usługi,

– gwarantowane są naprawy, doradztwo, darmowa pomoc w określonym przedziale czasu,

– istnieje możliwość wypróbowania danego towaru czy usługi z płatnością dopiero po określnym czasie,

– rekompensata – jeżeli klient znajdzie produkt taniej u konkurencji, to zwraca mu się dwukrotną różnicę ceny (tzn. klient zarobi) (por. tab. 2).

Tabela 2. Istota techniki odwróconego ryzyka

Technika odwróconego ryzyka	Standardowa transakcja
Firma podejmuje ryzyko Likwidacja bariery wejścia Promowanie wiarygodności marki, pewności działania i doświadczenia **Znaczne zwiększenie liczby nowych klientów**	Klient podejmuje ryzyko Niepewność co do poziomu usługi Widoczny brak pewności firmy

Ź r ó d ł o: opracowanie własne.

Stosując technikę odwróconego ryzyka w danej branży, przede wszystkim należy się zastanowić nad celem, jaki chcemy osiągnąć. Najprostszym sposobem wykorzystania tej techniki jest działanie według modelu darmowego testowania usługi przez określony czas. Dodatkowo należy pamiętać o pięciu zasadach zmaksymalizowania korzyści z jej wykorzystywania dla firmy[5]:

1. Im dłuższy czas darmowego użytkowania, tym dłuższy powinien być proponowany później okres trwania umowy.
2. Działanie świetnie sprawdza się w kontekście zapraszania przez obecnych klientów nowych użytkowników, którzy otrzymają prezent w postaci np. tygodniowego testowania. Korzystamy z potęgi marketingu szeptanego, reguły poleceń czy chęci ćwiczeń w gronie znajomych. Ten rodzaj działań nie przyniesie jednak skutku, gdy usługa staje się elementem przewagi konkurencyjnej w danej branży pomiędzy znającymi się podmiotami.
3. Pierwszy dzień, tydzień czy miesiąc za darmo powinien być standardem niezależnie od typu usługi.
4. Wszelkie działania przynoszące dodatkowe korzyści klientowi muszą być powiązane z uzyskaniem od niego pełnych danych osobowych oraz zgody na ich przetwarzanie, aby móc w przyszłości zastosować inne działania marketingowe wobec tej osoby.

[5] http://mysteryshop.org [2.03.2012].

Powyższy przykład dobrze ilustruje możliwości tej metody. Wydaje się jednak, że w kontekście danej branży można zastosować bardziej wyszukane warianty techniki odwróconego ryzyka, dodatkowo motywujące klientów korzystających z usługi. Przede wszystkim należy zdefiniować cele grupy docelowej. Na przykład w przypadku usługi dotyczącej cyklu treningów fizycznych połączonych z ich weryfikacją możliwe jest osiągnięcie jednego z trzech celów: zrzucenie wagi, poprawienie wydolności fizycznej oraz względy medyczne. Do osiągnięcia wszystkich z nich konieczna jest regularność ćwiczeń. Dobrym przykładem strategii danej usługi mogłoby być zapewnienie: „Jeżeli wykonując Nasz program ćwiczeń (trening *cardio* 2-3 razy w tygodniu) przez dwa miesiące nie schudniesz o 10% swojej wagi, zwrócimy wszelkie koszty poniesione na zakup usługi".

Główne zalety takiego podejścia:

1. Osoba przez dwa miesiące będzie regularnie trenowała, aby zachować możliwość zwrotu pieniędzy, co musi skutkować osiągnięciem celu, np. spadkiem wagi.
2. Pokazujemy gotową ścieżkę osiągnięcia celu, z gwarancją sukcesu.
3. Uczymy pozytywnych nawyków oraz zdobywamy klientów na lata.

Kolejnym przykładem pozytywnych działań jest oferowanie usług dodatkowych, uzupełniających. W ramach usługi dotyczącej treningów fizycznych możliwa jest np. promocja: „Do każdego rocznego pakietu treningów trzy odpowiednio dobrane diety z usługi X za darmo – oferta ważna tylko do końca roku".

Możliwe jest też zastosowanie różnych wariantów techniki odwróconego ryzyka podczas sprzedaży produktów materialnych, np. w zakresie zwrotu towaru. W tym przypadku można doskonale wykorzystać obowiązujące w Polsce prawo. Obecnie każdy produkt kupiony na odległość, np. przez Internet, można zwrócić w ciągu 10 dni bez podania przyczyny. Kładąc nacisk na satysfakcję klienta, jednocześnie zyskujemy u niego dużą wiarygodność[6].

Obawy wielu przedsiębiorców może budzić podejmowane ryzyko. Dotyczy ono przede wszystkim przypadku, gdy wielu użytkowników zechce skorzystać z darmowych miesięcznych okresów próbnych. Dzięki temu jednak zdobywamy niewielkim kosztem potencjalnych klientów, ale konieczne jest dobre przygotowanie procesu konwersji klientów potencjalnych na aktywnych. Nie powinniśmy się również obawiać, gdy wszyscy zaczną po paru dniach oddawać sprzedawany przez nas sprzęt. Średnia zwrotów mieści się zazwyczaj w granicach 1-2%.

Aby ta technika sprzedażowa przyniosła skutek, musimy być pewni, że świadczymy usługi na wysokim poziomie bezwzględnym oraz subiektywnie wyższym od relacji oferowanej usługi do ceny. Jeżeli obawiamy się, że zwrotów czy reklamacji może być dużo, to należy zweryfikować poziom świadczonych usług oraz szukać możliwości usprawnienia procesów. Działanie to powinien wykonać niezależny audytor.

---

[6] J.L. Bravard, R. Morgan, op. cit.

3. Zagrożenia związane z usługą SaaS

Możliwe zagrożenia/problemy związane z oferowaniem usługi typu SaaS nie ograniczają się jedynie do etapu jej projektowania oraz ułożenia właściwej strategii marketingowej, przekładającej się na dużą liczbę nowych klientów. Należy pamiętać, że w ramach usługi typu SaaS klient nie dokonuje inwestycji w oprogramowanie czy sprzęt, czekając na jej zwrot. W związku z tym ważne jest doprowadzenie do perfekcji wszystkich elementów procesu sprzedaży. Każda z możliwych do analizy usług typu SaaS jest unikalna i charakterystyczna, ponieważ działa w określonym otoczeniu, kierowana jest do określonych klientów i obsługiwana jest przez konkretnych pracowników. W związku z tym nie ma uniwersalnego rozwiązania dla wszystkich. Wszystkie zmienne wpływają na realizację określonej strategii, natomiast wszystkie działania są jej efektem.

Wśród pięciu głównych obszarów przyjętych w strategiach firm oferujących usługi typu SaaS dostrzega się następujące zagrożenia, które przekładają się wprost na obniżanie zysków:

A. Opieka nad klientem w zbyt małym zakresie – w tej kategorii występuje duża liczba niewielkich błędów skutkujących zagubieniem klienta.

1. Podczas pierwszego kontaktu:

– brak wiedzy pracowników na temat branży,

– brak rzeczywistej pomocy ze strony wdrożeniowców podczas wdrożenia i w późniejszym okresie,

– brak wyczucia nastrojów klienta przez pracowników,

– nieznana przeciętnemu człowiekowi terminologia używana w dokumentacji oraz mała przyjazność funkcjonalna systemu,

– niewykorzystywanie potencjału marketingu szeptanego.

2. Podczas kolejnych kontaktów:

– brak oferty sprzedaży dodatkowej, np. „Nowa dieta fitness”,

– zapominanie o monitorowaniu końca okresu wykupienia usługi,

– brak definiowania celów możliwych do osiągnięcia przez klienta oraz podkreślania sukcesów związanych z korzystaniem z usługi,

– brak dbania o właściwe korzystanie przez klienta z usług, zgodnie z regułą: „Jak już kupił karnet na miesiąc, to nie musi chodzić”, chyba że jest to element strategii przy tanich, długookresowych umowach.

B. Praktycznie brak opieki nad klientami korzystającymi z ofert zakupów grupowych przez Internet (typu Groupon). Osoba taka teoretycznie nie jest klientem firmy, ponieważ nie dokonała fizycznego zakupu, a jednak klub świadczy jej swoje usługi, będąc za to grupowo wynagradzanym. Czynności, do których ogranicza się firma, to pilnowanie autentyczności i poprawności kodów z wydruków klientów.

Podstawowy problem polega tu na odpersonalizowaniu klientów. Nie są bowiem prowadzone wobec nich akcje marketingowe ani żadne inne działania. W przypadku klientów usług typu Groupon powinniśmy próbować:

– pozyskać dane osobowe oraz zgodę na ich przetwarzanie,
– zebrać podpisy pod regulaminem korzystania z usługi,
– udostępnić np. internetowy panel klienta,
– zachęcać do ponownego skorzystania z usługi,
– uwzględniać ich w akcjach marketingowych (życzenia świąteczne, informacje o usługach, promocjach, imprezach itd.),
– sprzedawać usługi dodatkowe.

Działania te możemy podsumować jako pozyskanie stałego i lojalnego klienta.

C. Brak jasnego i prostego systemu motywacyjnego dla pracowników.

Każdy z członków zespołu musi mieć świadomość budowania wartości danego obiektu oraz wspólnego celu. Dodatkowo musi dostrzegać wymierne korzyści z zaangażowania. Nie ma innej możliwości realnej motywacji niż wypłacenie prowizji od ilości sprzedanych usług, wielkości sprzedaży, obrotów usług dodatkowych itd. Dodatkową zaletą motywacji jest likwidacja nadużyć w ramach sprzedaży (udostępnienia usług) poza obrotem klubowym, zgodnie z zasadą: po co oszukiwać, skoro można uczciwie zarobić na kolejnej sprzedaży. Jeżeli pracownik zrozumie, że sprzedając zarabia także na swoją premię (co jest dla niego istotnym dodatkiem do wynagrodzenia), to będzie to robił z przekonaniem i pełnym zaangażowaniem, co z pewnością przełoży się na jeszcze większą skuteczność.

D. Brak automatyzacji procesów, usprawniającej codzienną pracę.

Wiele działań, np. recepcji w klubie fitness, mogłoby się odbywać automatycznie, natomiast obsługa klienta mogłaby się skupić na kontakcie telefonicznym lub twarzą w twarz z klientem. Musimy starać się działać zgodnie z zasadą: „Jeżeli powtarzalną akcję nie wymagającą «inteligencji» możemy zautomatyzować, to tak, jak byśmy uwolnili dodatkowe godziny pracy naszych pracowników, które przyniosą dodatkowe korzyści". W przypadku usługi skierowanej do obiektów z branży *fitness* powyższą teorię możemy zastosować do następujących akcji:

– rezerwacje internetowe,
– informowanie i przypominanie o zajęciach, zmianach, odwołaniach,
– e-maile i SMS-y z życzeniami urodzinowymi,
– e-maile i SMS-y przypominające o płatnościach,
– e-maile i SMS-y przypominające o końcu umowy,
– automatyczne rozliczenia przelewów na konto,
– generowanie indywidualnych numerów kont,
– płatności internetowe.

E. Brak strategii odzyskiwania nieaktywnych klientów

Każda organizacja pozyskuje pewną liczbę nowych klientów w ramach prowadzonych działań marketingowych. Jest to bardzo ważne, jednak często nie przykładamy wystarczającej wagi do utrzymywania klientów zdobytych wcześniej – to trochę jak „ciągłe dolewanie wody do wiadra z coraz większymi dziurami". W dłuższej perspektywie woda i tak wyleci. Mówiąc o monitorowaniu, mamy na myśli:

– uprzedzanie momentu zakończenia umowy poprzez wcześniejszą sprzedaż kolejnej,

– kontakt e-mailowy, telefoniczny i SMS-owy z klientami, z którymi umowa wygasła,

– obserwacja i mobilizowanie klientów coraz rzadziej korzystających z opłaconych usług,

– automatyczne odnawianie umów na usługi,

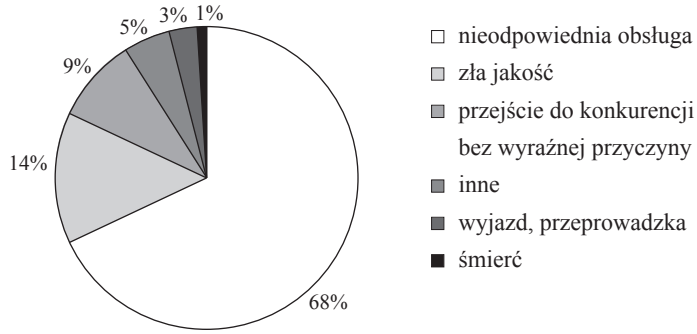
– promowanie *quasi*-umów na długi okres, gdzie motywatorem nie jest zapis umowy, lecz korzyść klienta w związku z regularnymi płatnościami,

– wskazywanie klientowi indywidualnych korzyści wynikających z regularnego korzystania z usługi,

– okresowe mailingi i akcje marketingowe adresowane do konkretnych grup dawnych klientów, np. pozyskanych podczas „Akcji Wiosna 2012".

Proces odzyskiwania klientów powinien być prowadzony każdego dnia.

Elementem każdego podejścia do zarządzania jest weryfikacja przyjętej strategii i dążenie do ciągłej poprawy działania. MSPA (Mystery Shopping Providers Association, http://mysteryshop.org) na podstawie swoich badań opublikowała najczęstsze powody odejść klientów (por. rys. 2).



Rys. 2. Typowy rozkład powodów odejść klientów

Ź r ó d ł o: http://mysteryshop.org [10.04.2012].

Każda organizacja powinna badać powyższe statystyki, przede wszystkim pod kątem odchyleń od średniej oraz zmian w czasie i efektów podejmowanych działań, ponieważ są to dane niepodważalne i nieobarczone subiektywnym postrzeganiem.

4. Elementy strony internetowej budujące zaufanie

Miejscem, gdzie potencjalny klient zapoznaje się najczęściej z oferowaną usługą, jest strona internetowa. Musi ona zostać tak skonstruowana, aby wzbudzała zaufanie. Jednym z często niedocenianych jej elementów są ilustracje, zdjęcia usług czy zdjęcia aktywnych klientów. Jeśli spojrzymy w statystyki odwiedzin Google Analytics firmowej strony www lub strony usługi SaaS i przefiltrujemy wyniki tylko dla nowych odwiedzających, to można zauważyć, że jedną z najczęściej odwiedzanych zakładek jest galeria. Klient może zobaczyć, jak wygląda przedmiotowa usługa, a przede wszystkim – przykładowi korzystający z niej klienci. Dobrze stworzona galeria zdjęć buduje zaufanie. W skład galerii powinny wchodzić profesjonalne zdjęcia:

– towarów i usług,

– całego zespołu wykonawczego; pamiętać należy jednak, by wstawić zdjęcia duże i wyraźne, przedstawiające uśmiechniętych ludzi, emanujących pozytywną energią,

– z klientami, wykonane np. podczas *eventu* lub burzliwych negocjacji. Zdjęcia te powinny być regularnie aktualizowane. Należy pokazać, że firma żyje oraz że pracują w niej ludzie z pasją wykonujący swoje zadania.

5. Podsumowanie

Omówienie w artykule uwarunkowań rynkowych działania organizacji oferujących usługi w modelu SaaS miało na celu lepsze zrozumienie funkcjonowania przedsiębiorstwa działającego w takim modelu. Różni odbiorcy końcowi usług, walka z obawami jako główną blokującą emocją czy technika odwróconego ryzyka to najważniejsze elementy, jakie należy brać pod uwagę podczas przygotowywania i realizowania strategii rozwoju przedsiębiorstwa oferującego usługi w modelu SaaS.

Literatura

Bravard J.L., Morgan R., *Inteligentny outsourcing*, MT Biznes, Warszawa 2010.

Foltys J., *Outsourcing w przedsiębiorstwach sektora MSP. Scenariusz aplikacyjny*, Wyd. Uniwersytetu Śląskiego, Katowice 2012.

Hale J., *Outsourcing. Training and Development*, John Wiley & Sons, New York 2009.

http://mysteryshop.org.

Kłos M., *Outsourcing w polskich przedsiębiorstwach*, CeDeWu, Warszawa 2010.

Kopczyński T., *Outsourcing w zarządzaniu przedsiębiorstwem*, PWE, Warszawa 2010.
Outsourcing w praktyce, red. D. Ciesielska, D. Radło, MT Biznes, Warszawa 2011.
Sobińska M., *Zarządzanie outsourcingiem informatycznym*, Wyd. Uniwersytetu Ekonomicznego we Wrocławiu, Wrocław 2011.
Vitasek K., Ledyard M., Manrodt K., *Zaangażowany outsourcing. Pięć zasad, które zmieniają outsourcing*, MT Biznes, Warszawa 2012.

Zeszyty Naukowe
Wyższej Szkoły Bankowej w Poznaniu
Nr 40/2012

Dariusz Ceglarek
Wyższa Szkoła Bankowa w Poznaniu

Zastosowanie kompresji semantycznej w zadaniach przetwarzania języka naturalnego

Streszczenie. *Kompresja semantyczna jest techniką pozwalającą uzyskać właściwą generalizację pojęć w zależności od kontekstu, dzięki czemu można znaleźć w różnych dokumentach tę samą myśl inaczej sformułowaną lub sformułowaną z użyciem innych pojęć. Rozwój koncepcji kompresji semantycznej i opracowanie nowych algorytmów pozwolił zastosować ją do klasyfikacji dokumentów i rozbudowy struktur reprezentacji wiedzy, takich jak sieci semantyczne. W artykule przedstawiono wyniki badań nad nowymi metodami i narzędziami kompresji semantycznej, które zostały przystosowane do zadań przetwarzania języka naturalnego.*

Słowa kluczowe: *kompresja semantyczna, ochrona własności intelektualnej, przetwarzanie języka naturalnego, reprezentacja wiedzy, sieć semantyczna*

1. Wprowadzenie

Kompresja semantyczna została opracowana z myślą o sytuacjach, gdy dwa lub więcej dokumentów zawiera wspólne fragmenty z pewnymi modyfikacjami przeprowadzonymi z użyciem słowników czy tezaurusów (np. poprzez użycie pojęć synonimicznych), ale które nie są podobne do siebie w sensie dosłownego porównania słowo po słowie, tak jak postępuje się w systemach wyszukiwawczych (*information retrieval systems* – IR). Z sytuacją taką mamy do czynienia w zadaniu wykrywania plagiaryzmów w określonych korpusach dokumentów[1]. Badania nad kompresją

[1] Zagadnienie to zostało opisane w: T. Ota, S. Masuyama, *Automatic plagiarism detection among term papers*, w: *Proceedings of the 3rd International Universal Communication '09*, ACM, 2009, s. 395-399; R. Lukashenko, V. Graudina, J. Grundspenkis, *Computer-based plagiarism detection methods and tools: an overview*, w: *Proceedings of the 2007 International Conference on Computer Systems and Technologies, CompSysTech '07. New York, USA*, ACM, 2007, s. 401-406;

semantyczną zostały zainicjowane podczas prac autora nad systemem ochrony własności intelektualnej SOWI[2].

Zadaniem systemu SOWI jest ochrona własności intelektualnej zawartej w dokumentach tekstowych, w tym wykrywanie zapożyczeń – sprawdzenie, czy w danym dokumencie tekstowym występuje odpowiednio duży fragment tekstu, który pokrywa się z treścią innego dokumentu w takim stopniu, że można mówić o zapożyczeniu treści i naruszeniu własności intelektualnej. Wdrożono tu autorskie algorytmy, co sprawia, że metody sprawdzania fraz wspólnych w dokumentach są niewrażliwe na zabiegi osób chcących ukryć fakt zapożyczenia fragmentów tekstu poprzez zmiany szyku tekstu oraz stosowanie w dokumencie synonimów czy pojęć bliskoznacznych. Zastosowanie w tych metodach kompresji semantycznej spowodowało, że możliwe jest wykrywanie nie tylko zapożyczeń w formie dosłownego skopiowania fragmentu tekstu, ale także polegających na przedstawieniu tej samej myśli za pomocą innych sformułowań. Stało się to możliwe również dzięki wykorzystaniu zaawansowanych struktur reprezentacji wiedzy o języku naturalnym, jakimi są sieci semantyczne. Autor dostrzegł możliwość zastosowania idei kompresji semantycznej również w innych sytuacjach, poprzez odpowiednie dostosowanie narzędzi i rozwiązań, zgodnie ze specyfiką rozmaitych zadań w ramach przetwarzania języka naturalnego.

Kompresja semantyczna może być także cennym narzędziem w zadaniach, w których głównym celem przetwarzania informacji jest przedstawienie użytkownikowi informacji dopasowanej do jego indywidualnych wymagań. Kompresję semantyczną można zatem zdefiniować jako skuteczną technikę uogólniania pojęć, która dopasowuje się do kontekstu i uwzględnia dodatkowo wymóg minimalizowania straty informacyjnej.

Powyższa definicja podkreśla potrzebę określenia właściwego kontekstu dla każdego pojęcia, które pojawia się w przetwarzanym dokumencie. Jest to zadanie trudne i jedynie osoba dysponująca odpowiednią wiedzą jest w stanie podać ze stuprocentową skutecznością właściwe znaczenie każdego pojęcia, gdyż w procesie tym należy uwzględnić również konotacje kulturowe danego pojęcia.

Autor wykazał, że kompresja semantyczna daje dobre wyniki, prawidłowo określając formy generalizujące pojęcia w języku naturalnym. Po raz pierwszy idea kompresji semantycznej pojawiła się w pracy N.N. Percovej[3], autorka nie podała jednak sposobu jej realizacji.

S. Burrows, S.M.M. Tahaghoghi, J. Zobel, *Efficient plagiarism detection for large code repositories*, „Software: Practice and Experience" 2007, t. 37, nr 2, s. 151-175.

[2] D. Ceglarek, *Koncepcja komponentowego systemu ochrony własności intelektualnej wykorzystującego semantyczne struktury informacji*, w: *Technologie informatyczne w zarządzaniu wiedzą – uwarunkowania i realizacja*, red. P. Adamczewski, M. Zakrzewicz, Wyd. WSB w Poznaniu, Poznań 2009.

[3] N.N. Percova, *On the types of semantic compression of text*, w: *COLING '82. Proceedings of the 9th conference on Computational linguistics*, t. 2, Academia Praha, 1982, s. 229-231.

Pierwotny pomysł kompresji semantycznej, która umożliwiałaby prawidłowe ujednoznacznianie pojęć wieloznacznych[4] podczas procesu ich generalizowania, został opracowany przez autora tego artykułu w postaci algorytmu kompresji semantycznej[5]. Algorytm został następnie zaimplementowany oraz przetestowany w szeregu eksperymentów, których efektem było wiele ulepszeń i rozszerzeń w stosunku do pierwowzoru. Istotne właściwości uzyskanego algorytmu to:

– zdefiniowanie i zaprezentowanie kompresji semantycznej jako technologii przydatnej w zadaniach przetwarzania języka naturalnego; skonstruowanie i zaimplementowanie w algorytmie słowników frekwencyjnych, które w przypadku występowania pojęć wieloznacznych wraz z algorytmami określającymi właściwe hiperonimy pojęć w zależności od kontekstu informacyjnego[6],

– przekształcenie sieci semantycznej WordNet[7] do postaci WiSENet, co spowodowało, że eksperymenty określające jakość kompresji semantycznej stały się możliwe zarówno dla dokumentów w języku polskim, jak i w angielskim[8],

– wysoce specjalizowany automat skończony pozwalający automatyzować budowę reguł, które wydobywają nowe pojęcia oraz nowe relacje leksykalne[9].

Niniejszy artykuł stanowi podsumowanie przeprowadzonych już badań, zatem jego struktura powinna konsekwentnie porządkować ich rezultaty. Dlatego następna sekcja poświęcona jest reprezentacji wiedzy, zwłaszcza strukturom tej

---

[4] M. Sanderson, *Word Sense Disambiguation and Information Retrieval*, w: *SIGIR '94. Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval*, red. W.B. Croft, C.J. van Rijsbergen, SIGIR, ACM/Springer, New York 1994, s. 142-151; J. Boyd-Graber, D.M. Blei, X. Zhu, *A Topic Model for Word Sense Disambiguation*, w: *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, Prague, June 2007*, s. 1024-1033.

[5] D. Ceglarek, K. Haniewicz, W. Rutkowski, *Quality of Semantic Compression in Classification*, w: *Computational Collective Intelligence, Second International Conference, ICCCI 2010, Kaohsiung, Taiwan, November 10-12, 2010. Proceedings*, cz. 1, red. J.-S. Pan, S.-M. Chen, N.T. Nguyen, Springer-Verlag, Berlin – Heidelberg 2010, „Lecture Notes in Computer Science" 2010, t. 6421, s. 162-171.

[6] R. Snow, D. Jurafsky, A.Y. Ng, *Learning syntactic patterns for automatic hypernym discovery*, w: *Advances in Neural Information Processing Systems (NIPS)*, 2005. Dokładny opis tego mechanizmu można znaleźć w: D. Ceglarek, K. Haniewicz, W. Rutkowski, *Semantic Compression for Specialised Information Retrieval Systems*, w: *Advances in Intelligent Information and Database Systems*, red. N.T. Nguyen, R. Katarzyniak, S.-M. Chen, Springer Verlag, Berlin – Heidelberg 2010, „Studies in Computational Intelligence" 2010, t. 283, s. 111-121.

[7] M. Miłkowski, *Automated Building of Error Corpora of Polish*, w: *Corpus Linguistics, Computer Tools, and Applications – State of the Art*, PALC 2007, red. B. Lewandowska-Tomaszczyk, Peter Lang, Frankfurt am Main 2008, s. 631-639.

[8] Przekształcenie to zostało opisane w: D. Ceglarek, K. Haniewicz, W. Rutkowski, *Quality of Semantic Compression…*

[9] D. Ceglarek, K. Haniewicz, W. Rutkowski, *Towards Knowledge Acquisition with WiSENet*, w: *New Challenges for Intelligent Information and Database Systems*, red. N.T. Nguyen, B. Trawinski, J.J. Jung, Springer Verlag, Berlin – Heidelberg 2011, „Studies in Computational Intelligence" 2011, t. 351, s. 75-84.

reprezentacji, ze szczególnym naciskiem na sieci semantyczne. Następnie przedstawiono proces przekształcenia najbardziej popularnej sieci semantycznej dla języka angielskiego WordNet do formatu SenecaNet, czego efektem jest sieć semantyczna WiSENet. Kolejna sekcja poświęcona jest globalnej i dziedzinowej kompresji semantycznej. Przedstawiono tu algorytmy i mechanizmy służące do jej utworzenia oraz przykłady zastosowań – m.in. pokazano, że kompresja semantyczna użyta w zadaniu klasyfikacji dokumentów tekstowych metodami analizy skupień podnosi jakość klasyfikacji dokumentów. Następny przykład pokazuje, jak kompresja semantyczna może być użyta do rozbudowy samej sieci semantycznej, przez co rozumie się odkrywanie nowych pojęć w celu ich dodania do sieci oraz odkrywanie nowych relacji leksykalnych pomiędzy konceptami już zgromadzonymi w sieci semantycznej. Ostatni przykład dotyczy zastosowania kompresji semantycznej do wspomagania rozumienia tekstu poprzez dopasowanie prezentowanych użytkownikowi pojęć zgodnie z jego poziomem rozumienia tekstów z danej dziedziny. Artykuł kończy się podsumowaniem, konkluzjami oraz wskazuje kierunki przyszłych badań.

2. Reprezentacja wiedzy

Analizą i automatycznym wyodrębnianiem prawidłowości w zbiorach dokumentów tekstowych i tekstowych bazach danych zajmuje się *text mining*, który jest multidyscyplinarną dziedziną, wykorzystującą m.in. metody statystyczne, metody systemów wyszukiwawczych (*information retrieval*) czy maszynowe uczenie. Ogólniejszą dyscypliną obejmującą problematykę sztucznej inteligencji i językoznawstwa, zajmującą się automatyzacją analizy, rozumienia, tłumaczenia i generowania języka naturalnego, jest przetwarzanie języka naturalnego (*Natural Language Processing* – NLP).

Metody *text miningu* składają się zazwyczaj z dwóch etapów: wygładzania tekstu (*text refining*) oraz wydobywania wiedzy (*knowledge discovery*). Na etapie wygładzania tekstu pozbawiony struktury dokument tekstowy przekształcany jest w formę pośrednią[10], tworzoną w celu wykrycia zależności między dokumentami (wydobycie wiedzy) w drugim etapie z wykorzystaniem metod charakterystycznych dla danego zadania *text miningu*[11].

W przypadku wszystkich zadań realizowanych w ramach zadań przetwarzania języka naturalnego niezbędne jest przeprowadzenie wygładzania tekstu, które pole-

[10] Forma pośrednia może mieć postać sekwencji cech, wektora cech lub grafu konceptualnego.

[11] Typowe zadania w ramach *text miningu* obejmują klasyfikację dokumentów (grupowanie, kategoryzację), automatyczne streszczanie dokumentów, grupowanie pojęć, wizualizację i nawigację w zbiorze dokumentów i ekstrakcję informacji.

ga na przekształceniu wyjściowego dokumentu tekstowego w strukturę zawierającą ułożone sekwencyjnie deskryptory pojęć (konceptów) występujących w wyjściowym dokumencie. Na wygładzanie tekstu składają się operacje: wyodrębnienia wyrazów (*tokenization*), usunięcia słów niemających znaczenia informacyjnego z tzw. stop-listy, identyfikacji pojęć wielowyrazowych oraz lematyzacji pojęć.

Ostatnią fazą wygładzania tekstu jest ujednoznacznienie pojęć (*disambiguation*), czyli wyznaczenie właściwych pojęć dla wieloznacznych termów[12], które wystąpiły w dokumencie (eliminowanie tzw. pozornego podobieństwa). Zjawisko wieloznaczności pojęć, zwane polisemią, dotyczy każdego języka naturalnego i oznacza, że jednemu konceptowi odpowiada wiele znaczeń, czyli że różne pojęcia nazywane są tak samo. Przykładem polisemii może być koncept „dysk", który ma takie znaczenia, jak: „komputerowy nośnik pamięci", „przyrząd lekkoatletyczny", „kość składowa kręgosłupa" lub „owalny kształt". Celem zastosowania ujednoznaczniania pojęć jest lepsze odwzorowanie termów występujących w dokumentach we właściwe pojęcia, a zatem lepsze dopasowanie informacji wywnioskowanej z dokumentów do potrzeb informacyjnych[13]. Istnieje wiele metod ujednoznaczniania pojęć, w tym metody oparte na semantycznej reprezentacji wiedzy. Skuteczną[14] metodę ujednoznaczniania pojęć (o skuteczności na poziomie dochodzącym do 82%), wykorzystującą sieć semantyczną dla języka polskiego, zaproponował autor niniejszego opracowania[15]. Wykorzystuje ona zależności semantyczne między pojęciami w sieci semantycznej SenecaNet dla języka polskiego, a jej działanie opiera się na wskazaniu najbardziej prawdopodobnego znaczenia pojęcia wieloznacznego – biorąc pod uwagę lokalny kontekst użycia owego pojęcia w badanym dokumencie.

2.1. Struktury reprezentacji wiedzy

Metody reprezentacji wiedzy są sposobem, w jakim wiedza o świecie jest przedstawiana wraz z metodami jej przetwarzania i wnioskowania (inferencji). Jest to ściśle określony język opisu wiedzy zaopatrzony w mechanizm jej przetwarzania.

[12] Termem jest słowo lub wielowyrazowy związek semantyczny (związek frazeologiczny, kolokacja), który wystąpił w dokumencie.

[13] Ch. Stokoe, M.P. Oakes, J. Tait, *Word Sense Disambiguation in Information Retrieval Revisited*, SIGIR, 2003.

[14] Skuteczne metody prawidłowo identyfikują od 70 do 75% znaczeń pojęć wieloznacznych. W pracy M. Sanderson, *Retrieving with Good Sense*, „Informational Retrieval" 2000, t. 2, nr 1, s. 49-69, pokazano, że wyłącznie metody analizy lingwistycznej są w stanie pokonać poziom 90% skuteczności ujednoznaczniania pojęć wieloznacznych.

[15] D. Ceglarek, *Zastosowanie sieci semantycznej do disambiguacji pojęć w języku naturalnym*, w: *Systemy wspomagania organizacji SWO 2006*, Wyd. AE w Katowicach, Katowice 2006.

Każdemu pojęciu odpowiada w języku naturalnym zapis w postaci wyrazu, kolokacji[16] lub związku frazeologicznego, który jest jego odzwierciedleniem. Zapis pojęcia w języku naturalnym nazywamy konceptem. Celem jest stworzenie przetwarzalnej przez system reprezentacji dokumentu, tak aby na podstawie treści dokumentu wyodrębnić jednostki odpowiadające znaczeniu informacyjnemu konceptów.

Popularne w systemach wyszukiwawczych i *text miningu* metody opierają się na prostej strukturze reprezentacji wiedzy, gdzie dokumenty reprezentowane są przez zbiory słów kluczowych, a najbardziej popularnym modelem zapytań kierowanych do systemu wyszukiwawczego z wykorzystaniem tej reprezentacji wiedzy jest model wektorowy (*vector space model*)[17].

Bardziej złożone struktury reprezentacji wiedzy to: słownik definicyjny (glosariusz), słownik dziedzinowy, sieć semantyczna i ontologia. Struktury te wprowadzają różne relacje leksykalne występujące pomiędzy przechowywanymi w nich konceptami.

Sieć semantyczna jest grafem skierowanym mającym koncepty (pojęcia) jako wierzchołki oraz krawędzie dla reprezentowania relacji leksykalnych między konceptami. Jest najlepszą strukturą reprezentacji wiedzy do odzwierciedlania powiązań semantycznych między konceptami w języku naturalnym[18], gdyż gromadzi całą wiedzę o semantyce pojęć ze względu na przechowywanie wszystkich relacji leksykalnych charakterystycznych dla języka naturalnego oraz brak nadmiernej złożoności (charakterystycznej dla ontologii).

Stąd wynika jej przydatność w systemach przetwarzających język naturalny. Wnioskowanie z wykorzystaniem sieci semantycznej odbywa się po krawędziach, które mogą posiadać wagi określające ich ważność. Wnioskowanie polega na przeszukiwaniu grafu, w którym, rozpoczynając poruszanie się od jednego węzła grafu (konceptu) i poruszając się po krawędziach (relacje między konceptami) wychodzących z węzła, docieramy do kolejnych węzłów, co odpowiada wnioskowaniu o właściwościach konceptów.

Korzyści wynikające ze stosowania sieci semantycznych w systemach przetwarzających język naturalny zostały opisane przez R.A. Baeza-Yatesa i B. Ribeiro-Neto[19]. Sieci umożliwiają przede wszystkim dostarczenie właściwych znaczeń pojęć, co skutkuje zwiększeniem precyzji odpowiedzi oraz wzrostem pełności odpowiedzi systemu. W zadaniach klasyfikacyjnych ta forma reprezentacji wiedzy podnosi jakość klasyfikacji i kategoryzacji[20].

---

[16] Kolokacja to związek semantyczny, którego znaczenie wynika z połączenia znaczeń kilku słów wchodzących w jego skład (np. „związek małżeński").

[17] R.A. Baeza-Yates, B. Ribeiro-Neto, *Modern Information Retrieval*, Addison-Wesley Longman Publishing, Boston 1999.

[18] S. Staab, A. Hotho, *Ontology-based text document clustering*, w: *IIS, Advances in Soft Computing*, red. M.A. Kłopotek, S.T. Wierzchoń, K. Trojanowski, Springer, 2003, s. 451-452.

[19] R.A. Baeza-Yates, B. Ribeiro-Neto, op. cit.

[20] M. Baziz, *Towards a Semantic Representation of Documents by Ontology-Document Mapping*, w: *Artificial Intelligence: Methodology, Systems, and Applications. 11th International Confer-*

Powszechnie stosowaną siecią semantyczną dla języka angielskiego jest WordNet[21]. Dla języka polskiego autor posłużył się zbudowaną w ramach projektu SeNeCa[22] siecią semantyczną o nazwie SenecaNet. Sieć ta jest od wielu lat rozbudowywana, a proces rozbudowy obejmuje dodawanie nowych konceptów oraz relacji leksykalnych pomiędzy konceptami. Pierwsze automatyczne metody rozbudowy sieci SenecaNet zaproponowali Ceglarek i Rutkowski[23]. Procedurę jej półautomatycznej rozbudowy z wykorzystaniem kompresji semantycznej opisano w rozdziale 4.1[24].

2.2. Sieć semantyczna SenecaNet

SenecaNet jest siecią semantyczną, która zawiera ponad 154 tysiące konceptów oraz przechowuje rozmaite relacje leksykalne pomiędzy konceptami dla języka polskiego (tab. 1). Sieć ta była pierwszą siecią semantyczną użytą do kompresji semantycznej, m.in. dzięki kilku jej specyficznym własnościom. Koncepty w tej sieci są przechowywane w postaci posortowanej topologicznie listy, co oznacza, że w definicji danego konceptu można odwołać się wyłącznie do konceptów wcześniej zdefiniowanych. W zbudowanej w ten sposób strukturze niemożliwe jest istnienie cykli, dlatego też algorytmy grafowe są wydajniejsze.

Przechowywanie definicji pojęć w postaci listy wynika z szeregu działań optymalizacyjnych, które są stosowane dla szybkiego jej przetwarzania. W tej formie sieć semantyczna może być traktowana jako struktura hierarchiczna, co jest znakomitym rozwiązaniem z obliczeniowego punktu widzenia. Definicja każdego konceptu w sieci SenecaNet zawiera deskryptor konceptu, listę możliwych form morfologicznych, a także hiperonimy konceptu[25], jego synonimy, holonimy, meronimy, konotacje oraz związane z konceptem relacje nienazwane. Notację w formacie sieci semantycznej SenecaNet ilustruje tabela 2.

---

ence, AIMSA 2004, Varna, Bulgaria, September 2-4, 2004. Proceedings, red. Ch. Bussler, D. Fensel, Springer, Berlin – Heidelberg 2004, „Lecture Notes in Computer Science" 2004, t. 3192, s. 33-43.

[21] Projekt Cognitive Science Laboratory Uniwersytetu Princeton jest dostępny pod adresem: http://wordnet.princeton.edu.

[22] Projekt SeNeCa (**Se**mantic **Ne**twork and **Ca**tegorization, http://seneca.kie.ae.poznan.pl) miał za zadanie automatyzację rozbudowy sieci semantycznej dla języka polskiego.

[23] D. Ceglarek, *Zastosowanie sieci semantycznej...*

[24] Szczegółowy opis w: D. Ceglarek, K. Haniewicz, W. Rutkowski, *Towards Knowledge Acquisition…*

[25] Każdy koncept może posiadać jeden lub więcej hiperonimów, dzięki czemu uzyskana struktura jest pod względem taksonomicznym heterarchią; zob. S. Staab, A. Hotho, op. cit.

Tabela 1. Porównanie sieci WordNet i SenecaNet

Parametry	WordNet	Sieć SenecaNet
Liczba konceptów	155 200	154 200
Liczba słów polisemicznych	27 000	21 300
Liczba synonimów		8400
Relacje hiperonimii, hiponimii	+	+
Relacje antonimii	+	–
Konotacje	+	+
Relacje nienazwane	–	+

Ź r ó d ł o: opracowanie własne.

Tabela 2. Format definicji pojęć w sieci SenecaNet

samochód → pojazd, &silnik
Chiny → kraj, :Azja
prowincja.n.01 → jednostka podziału administracyjnego
prowincja.n.02 → obszar geograficzny,*zacofany
prowincja → prowincja.n.01; prowincja.n.02
prowincja Guangdong → prowincja.n.01, :Chiny,
Guangzhou → miasto, :prowincja Guangdong, #stolica(prowincja Guangdong)
Canton → =Guangzhou, *starodawna nazwa chińska

Ź r ó d ł o: opracowanie własne.

2.3. Konwersja sieci semantycznej WordNet

Zastosowanie kompresji semantycznej dla języka angielskiego wymagało użycia sieci semantycznej o strukturze analogicznej do tej, którą posiada sieć semantyczna SenecaNet. Ze względu na trudność zadania utworzenia nowej sieci semantycznej dla języka angielskiego zdecydowano się wykorzystać istniejącą sieć semantyczną WordNet i poddać ją przekształceniu do takiej samej struktury reprezentacji pojęć jak w sieci SenecaNet[26]. Sieć WordNet jest bardo obszerną i dojrzałą siecią semantyczną, która osiągnęła swój obecny kształt dzięki wieloletniej pracy ogromnego zespołu ludzi, i jej przydatność została wykazana w szeregu badań i eksperymentów[27].

[26] Zadanie to zostało szczegółowo opisane w: D. Ceglarek, K. Haniewicz, W. Rutkowski, *Quality of Semantic Compression*…

[27] W pracy J. Rosenzweig, R. Mihalcea, A. Csomai, „*WordNet bibliography". Web page: a bibliography referring to research involving the WordNet lexical database*, http://lit.csci.unt.

Jednakże przekształcenie zorientowanej na synsety struktury sieci Word-Net w nieposiadającą cykli (w rozumieniu grafowym) strukturę operującą deskryptorami konceptów identyfikowanych w analizowanych tekstach okazało się zadaniem trudnym. Dlatego też został opracowany algorytm umożliwiający to przekształcenie, posługujący się zbiorami i bierze pod uwagę każdą lemmę przechowywaną w danym synsecie oraz synsety, które stanowią hiperonimy w stosunku do przetwarzanego synsetu.

Synset definiuje się jako grupę lemm (termów) mających takie samo znaczenie. Po dokładnym przestudiowaniu okazało się, że większość lemm zgromadzonych w jednym synsecie nie stanowi w stosunku do siebie idealnych synonimów. Mają one wspólne znaczenie, lecz poziom podobieństwa znaczeniowego jest różny. Każda lemma występująca w synsecie może być pojedynczym słowem lub związkiem semantycznym, który składa się z kilku słów[28].

Podczas przekształcania struktury sieci semantycznej ze zorientowanej na synsety w strukturę w pełni hierarchiczną niezbędne jest zmodyfikowanie sposobu wyboru konceptów opisujących dany synset – w celu uniknięcia występowania cykli w docelowym grafie konceptów. Najprostsza sytuacja występuje wtedy, kiedy lemma zawarta w deskryptorze synsetu występuje wyłącznie w tym synsecie, czyli lemma jest unikalnym deskryptorem synsetu (warunek unikalności). W innych przypadkach należy znaleźć inną lemmę z tego samego synsetu, która spełnia warunek unikalności. Przeprowadzone eksperymenty dotyczące rzeczowników wykazały, że postępowanie takie prowadzi do uzyskania sieci semantycznej, w której jest jedynie 25 000 konceptów z 86 000 istniejących w zawierających rzeczowniki synsetach WordNetu.

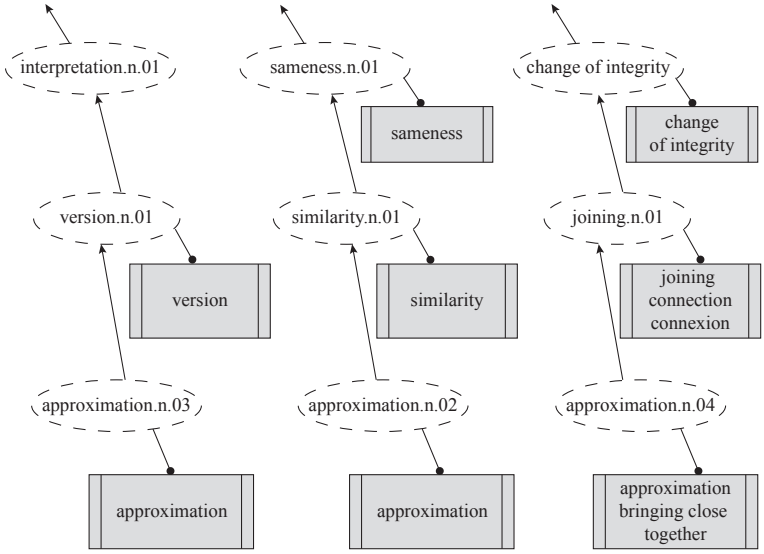
Wprowadzony został zatem „syntetyczny" deskryptor synsetów, który wynika z potrzeby uniknięcia występowania cykli.

Na rysunkach 1 i 2 przedstawiono wizualizację powyższego procesu przekształcenia na przykładzie lemmy „approximation", która występuje w kilku różnych synsetach. Z tego powodu nie może ona być deskryptorem synsetu. Na rysunku 2 można z łatwością zauważyć, że lemma „bringing close together" występuje dokładnie w jednym synsecie, a zatem może ona zastąpić syntetyczny deskryptor „approximation.n.04" (spełniając warunek unikalności). Procedurę przekształcenia struktury sieci WordNet do formatu SenecaNet opisano poniżej i pokazano w postaci pseudokodu w Algorytmie 1.

---

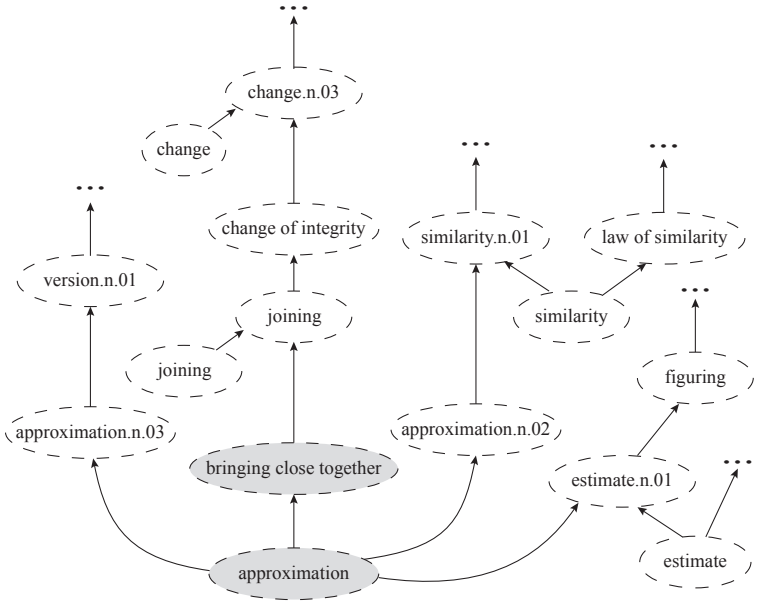
edu/%7Ewordnet [1.09.2007], przedstawiono 868 projektów wykorzystujących WordNet w zadaniach przetwarzania języka naturalnego.

[28] G.A. Miller, *Wordnet: a lexical database for English*, „Communications of the ACM" 1995, t. 38, nr 11.



Rys. 1. Struktura sieci WordNet zorientowana na synsety

Ź r ó d ł o: D. Ceglarek, K. Haniewicz, W. Rutkowski, *Quality of Semantic Compression in Classification*, w: *Computational Collective Intelligence, Second International Conference, ICCCI 2010, Kaohsiung, Taiwan, November 10-12, 2010. Proceedings*, cz. 1, red. J.-S. Pan, S.-M. Chen, N.T. Nguyen, Springer-Verlag, Berlin – Heidelberg 2010, „Lecture Notes in Computer Science" 2010, t. 6421, s. 162-171.



Rys. 2. Taksonomia konceptów w sieci SenecaNet

Ź r ó d ł o: jak przy rys. 1.

Algorytm 1. Algorytm przekształcenia sieci WordNet do formatu SenecaNet (efektem przekształcenia jest sieć WiSENet)

```
for all (d,S) ∈ WN do
  for all l ∈ S do
    F[l]++
  end for
end for
for all (d,S) ∈ WN do
  parsuj lemat z deskryptora w synsecie
  l ← split(d,".")[0]
  if F[l] = 1 then
    lemat może być użyty jako deskryptor synsetu
    d ← l
  else
    for all l ∈ S do
      if F[l] = 1 then
        d ← l
        exit
      end if
    end for
  end if
  SN[d] ← S
end for
```

WN – sieć WordNet w postaci listy synsetów identyfikowanych poprzez deskryptory *d*
S – synset zawierający wiele lemm *l*
F[*l*] – liczba synsetów zawierających lemmę *l*
SN – wynikająca z przekształcenia sieć semantyczna WiSENet

Pierwszym krokiem algorytmu jest zbudowanie słownika frekwencyjnego (F) dla lemm, który zawierać będzie liczbę synsetów zawierających daną lemmę. W tym celu algorytm rozpatruje i sumuje wszystkie synsety w sieci WordNet oraz wszystkie lemmy w synsetach. W następnym kroku algorytm pobiera deskryptor (w miarę możliwości lemmę) dla każdego synsetu. Następnie algorytm sprawdza, czy taka lemma występuje dokładnie w jednym synsecie – i jeśli odpowiedź jest pozytywna, to pobrana lemma może być użyta jako nowy deskryptor synsetu. W przeciwnym wypadku sprawdza pozostałe lemmy z analizowanego synsetu i sprawdza, czy istnieje taka, którą można wykorzystać jako deskryptor synsetu. Jeśli nie istnieje wśród nich żadna lemma spełniająca warunek unikalności, jako deskryptor synsetu zostaje wykorzystany oryginalny deskryptor z sieci WordNet.

Uzyskana w wyniku powyższego przekształcenia sieć semantyczna dla języka angielskiego WiSENet zawiera te same dane (koncepty i relacje leksykalne), które zawiera sieć semantyczna WordNet, jednakże ma hierarchiczną strukturę sieci SenecaNet.

3. Kompresja semantyczna

Zgodnie z definicją podaną na wstępie kompresja semantyczna jest techniką, która ma za zadanie dostarczyć bardziej ogólne koncepty w stosunku do konceptu, który wystąpił w analizowanym dokumencie lub w zapytaniu kierowanym do systemu. Konceptu istniejącego w danym dokumencie w pewnym kontekście, który decyduje o jego znaczeniu.

Gdy zadaniem algorytmu kompresji semantycznej jest wyznaczenie konceptu bardziej generalnego w stosunku do danego konceptu, algorytm musi precyzyjnie określić poziom tej generalizacji. Im wyższy jest poziom generalizacji, tym większa jest utrata informacji. W niektórych zastosowaniach może to stanowić pozytywne zjawisko (np. w zadaniu klasyfikacji dokumentów metodami analizy skupień[29]), ale wtedy, kiedy kompresja semantyczna ma służyć przekształceniu dokumentu w celu przedstawienia go odbiorcy będącemu człowiekiem, nie jest to zjawisko akceptowalne. W kompresji chodzi o wyznaczenie dla każdego pojęcia takiego pojęcia, które będzie jego reprezentantem (deskryptorem). Podstawą będzie liczba wystąpień w korpusie dokumentów (pojęcia częste będą reprezentowane przez same siebie, pozostałe pojęcia będą reprezentowane przez pojęcia nadrzędne w strukturze, których skumulowana liczba wystąpień jest duża).

Kompresja semantyczna powstała pierwotnie jako kompresja globalna. W związku z pojawieniem się licznych jej zastosowań i wykorzystaniem korpusów dokumentów z rozmaitych dziedzin została rozwinięta poprzez doskonalenie strategii generalizacji w zależności od dziedziny dokumentów, czego efektem jest kompresja dziedzinowa. Prezentuje to następna sekcja, w której zostały przedstawione również wyniki eksperymentów obrazujących skuteczność i efektywność kompresji.

3.1. Mechanizm kompresji semantycznej

Mechanizm kompresji semantycznej został zaprezentowany w 2010 r.[30] jako metoda podnosząca jakość i efektywność klasyfikacji dokumentów tekstowych.

[29] R. Nock, F. Nielsen, *On weighting clustering*, „The IEEE Transactions on Pattern Analysis and Machine Intelligence" 2006, nr 28(8), s. 1223-1235.

[30] Zob. D. Ceglarek, K. Haniewicz, W. Rutkowski, *Semantic Compression…*

Kompresja tekstu jest możliwa poprzez zastosowanie sieci semantycznej oraz danych o częstości wystąpień konceptów (w formie słowników frekwencyjnych). Efektem takiego postępowania jest redukcja liczby konceptów używanych do reprezentowania pojęć występujących w dokumentach tekstowych bez znaczącej straty informacji, co jest niezwykle istotne z perspektywy procesu przetwarzania języka naturalnego (zwłaszcza wtedy, kiedy stosuje się model wektorowy[31]).

Ponadto redukcja liczby konceptów pomaga w radzeniu sobie ze zjawiskami lingwistycznymi, które stanowią znaczne wyzwanie w zadaniach przetwarzania języka naturalnego[32]. Zjawiskiem, na które najczęściej zwraca się uwagę w zadaniach NLP, jest polisemia oraz synonimia[33]. Gdy używa się wielu termów jako określenia tego samego lub podobnego konceptu, to mogą one zostać zastąpione jednym, bardziej ogólnym konceptem. Przy zastosowaniu analizy statystycznej można sporządzić słownik frekwencyjny dopasowany do kontekstu danej dziedziny i wyznaczyć właściwy deskryptor dla konceptów polisemicznych. Uzyskana w wyniku zredukowania zbioru konceptów struktura jest wydajniejsza obliczeniowo i powoduje mniejszą utratę informacji niż rozwiązania, które nie stosują tej techniki.

Przyjmijmy, że posiadamy dokumenty poświęcone badaniom biologicznym i w związku z tym w dokumentach występują łacińskie nazwy gatunków zwierząt lub roślin. Podczas klasyfikacji takich dokumentów te łacińskie nazwy występujące w dokumentach poszerzają wektor termów opisujący dokumenty, co utrudnia obliczeniowo proces ich klasyfikacji oraz powoduje spadek jakości uzyskanej klasyfikacji. Zamiana nazwy łacińskiej określającej gatunek jakiejś rośliny na odpowiadający jej koncept skraca wynikowy wektor pojęć opisujących dokument oraz skutkuje minimalną stratą informacyjną. Naturalnie, taka zamiana może być dokonana dla specyficznego korpusu dokumentów, w którym nazwy łacińskie są stosunkowo rzadkie, a zatem możliwe do zastąpienia. Wybór konceptów w procesie generalizacji pojęć jest zależny od dziedziny dokumentów.

W ogólności, kompresja semantyczna umożliwia realizowanie zadań przetwarzania języka naturalnego, takich jak wyszukiwanie wzorców w tekstach, operując na poziomie konceptów, a nie pojedynczych słów. Osiąga się to nie tylko przez reprezentowanie termów przez ich wspólne znaczenie (rozwiązanie znane jako podejście zorientowane na synsety[34]), ale również poprzez zastępowanie długich fraz ich krótszymi odpowiednikami.

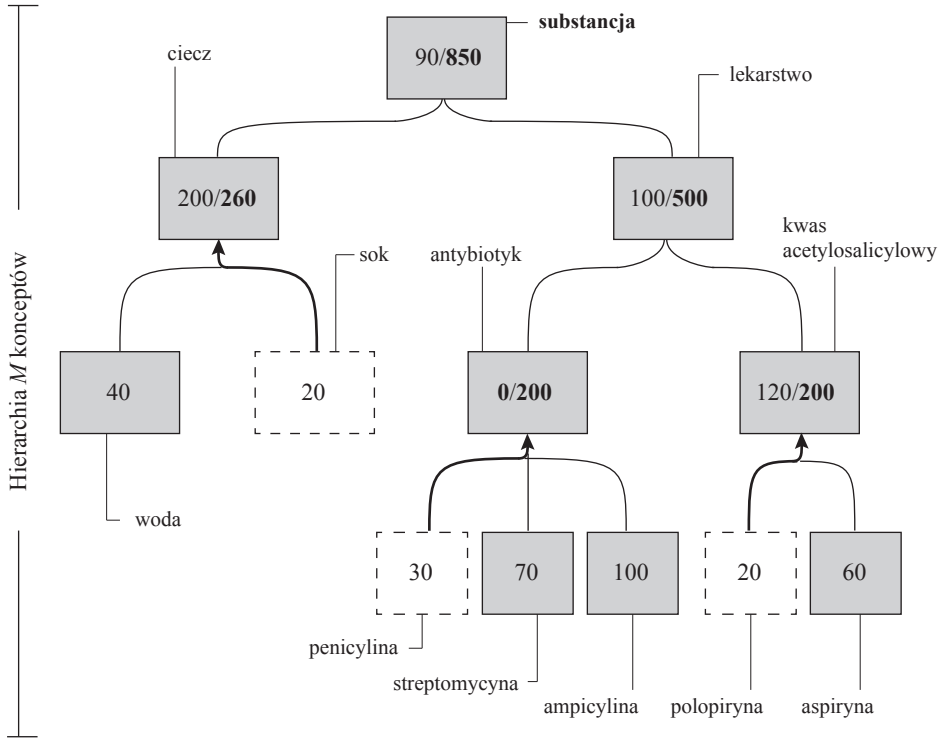
---

[31] R.A. Baeza-Yates, B. Ribeiro-Neto, op. cit.; K. Erk, S. Pad'o, *A Structured Vector Space Model for Word Meaning in Context*, w: *EMNLP '08 Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, Stroudsburg, PA, USA 2008, s. 897-906.

[32] R. Sinha, R. Mihalcea, *Unsupervised graph-basedword sense disambiguation using measures of word semantic similarity*, w: *International Conference on Semantic Computing ICSC 2007*, IEEE 2007, s. 363-369.

[33] R. Krovetz, W.B. Croft, *Lexical ambiguity and information retrieval*, „ACM Transactions on Information Systems" 1992, nr 10, s. 115-141.

[34] G.A. Miller, op. cit.



Rys. 3. Wybór konceptów o największej skumulowanej liczbie wystąpień spośród *M* konceptów w sieci semantycznej

Ź r ó d ł o: opracowanie własne.

Kompresja semantyczna pozwala na znajdowanie wspólnego znaczenia dla sentencji wyrażonych za pomocą różnych termów.

Mechanizm uogólniania pojęć nie jest zdolny do analizy zależności pomiędzy nimi oraz do dokonywania zmian w sentencjach zgodnie z zasadami gramatyki. Umożliwia jednak trafne wykrywanie, że różnie sformułowane sentencje niosą tę samą zawartość informacyjną. W zadaniu wykrywania plagiaryzmów w ramach omawianego systemu SOWI wykorzystywany jest algorytm „*bag of concepts*", którego wyróżniającymi cechami są: niewrażliwość na zmianę szyku termów w porównywanych dokumentach oraz niewrażliwość na stosunkowo krótkie niezgodności w sekwencjach termów[35]. Ceglarek i Haniewicz zaproponowali algorytm o tych samych własnościach, lecz mający znacznie mniejszą, bo logarytmiczną złożoność obliczeniową[36].

[35] Szczegółowy opis w: D. Ceglarek, *Koncepcja komponentowego systemu ochrony…*

[36] D. Ceglarek, K. Haniewicz, *Fast Plagiarism Detection by Sentence Hashing*, w: *Artificial Intelligence and Soft Computing. 11th International Conference, ICAISC 2012, Zakopane, Poland,*

Należy rozumieć, że kompresja semantyczna jest mechanizmem, którego zastosowanie oznacza utratę części informacji semantycznej, lecz utrata informacji jest nieznaczna, kiedy wybrane deskryptory pojęć bardziej generalnych są konceptami często występującymi w dokumentach tekstowych i znaczenie wybranych konceptów bardziej ogólnych jest podobne. Poziomem kompresji steruje się poprzez określenie liczby konceptów, które stają się deskryptorami używanymi do opisu tekstu dokumentów. Eksperymenty przeprowadzone w celu zmierzenia jakości metod w ramach zadań przetwarzania języka naturalnego pokazały, że redukcja liczby konceptów do ok. 4000 nie wpływa znacząco na utratę jakości w zadaniach klasyfikacji dokumentów. Idea wyboru konceptów zgodnie z ich częstością występowania w dokumentach pokazana została na rysunku 3. Opis samego algorytmu znajduje się w następnym punkcie artykułu.

3.2. Algorytm kompresji semantycznej

W korpusie dokumentów występuje M konceptów k_i, które można użyć do utworzenia M-elementowego wektora reprezentującego dokumenty. Określona jest również docelowa liczba konceptów N (gdzie $N < M$). W pierwszej kolejności należy obliczyć liczbę wystąpień $f(k_i)$ dla każdego konceptu k_i we wszystkich dokumentach. Następnie należy obliczyć skumulowaną liczbę wystąpień dla wszystkich konceptów – do liczby wystąpień danego konceptu dodaje się sumę wystąpień wszystkich jego hiponimów. Kolejnym krokiem jest włączenie informacji o liczbie wystąpień wynikających z relacji synonimii w ten sposób, że dla grupy konceptów połączonych relacją synonimii wybiera się koncept mający największą skumulowaną liczbę wystąpień i do jego skumulowanej liczby wystąpień dodaje się skumulowane liczby wystąpień wszystkich pozostałych synonimów.

Poruszając się w górę hierarchii konceptów, należy obliczyć skumulowaną liczbę wystąpień konceptów poprzez dodanie sumy skumulowanej liczby wystąpień hiponimów do danego konceptu (ich hiperonimu):
$cumf(k_i) = f(k_i) + \sum_j \left[(cumf(k_j) \right]$, gdzie k_i jest hiperonimem dla k_j (patrz Algorytm 2 oraz Algorytm 3).

Ostatnim krokiem algorytmu jest wybranie N konceptów o największej skumulowanej liczbie wystąpień, które stanowić będą listę deskryptorów (Algorytm 4). Opisana procedura kompresji semantycznej pozwala zredukować rozmiar wektora pojęć opisujących dokumenty o M-N konceptów.

April 29-May 3, 2012, Proceedings, t. 2, red. L. Rutkowski, M. Korytkowski, R. Scherer, R. Tadeusiewicz, L.A. Zadeh, J.M. Zurada, Springer-Verlag, Berlin – Heidelberg 2012, „Lecture Notes in Computer Science" 2012, t. 7268, s. 30-38.

Algorytm 2. Ustalenie uogólnionych konceptów, które staną się deskryptorami pojęć poprzez obliczenie skumulowanej liczby wystąpień w korpusie dokumentów C

```
//wybór synonimu reprezentującego grupę pojęć synonimicznych
max = 0
n = 0
sum = 0
for s ∈ S_v do
  sum = sum + l_s
  if l_s > max then
    max = l_s
    n = s
  end if
end for
l_s = l_s + sum
```

```
// obliczenie skumulowanej liczby wystąpień dla hiperonimów
for v ∈ V'' do
  p = card(H_v)
  for h ∈ H_v do
    l_h = l_h + l_v / p
  end for
end for
```

S_v – zbiór synonimów dla konceptu v
V – wektor konceptów przechowywany w sieci semantycznej
V' – topologicznie posortowany wektor V
V'' – odwrócony wektor V'
l_v – liczba wystąpień konceptu v w korpusie dokumentów C
H_v – zbiór hiperonimów konceptu v

Algorytm 3. Wybór m konceptów z sieci semantycznej w procedurze dziedzinowej kompresji semantycznej

```
for v ∈ V do
  if l_v ≥ f
    d_v = v
  else
    d_v = FMax(v)
  end if
end for
```

L – wektor przechowujący liczbę wystąpień konceptów w korpusie dokumentów C
L' – posortowany malejąco wektor L
f – liczba wystąpień m-tego konceptu w wektorze L'

Algorytm 4. Procedura FMax znajdująca dla danego konceptu *v* jego deskryptor (hiperonim o największej skumulowanej liczbie wystąpień)

```
FMax(v):
max = 0
x = ∅
for h ∈ H_v do
  if d_h ≠ ∅ then
    if l_{d_h} > max then
      max = l_{d_h}
      x = d_h
    end if
  end if
end for
return x
```

W tabeli 3 zamieszczone zostały dwa przykłady semantycznie skompresowanych fragmentów tekstu (przy 4000 deskryptorach konceptów) w języku angielskim. Po oryginalnych fragmentach (A, B) przytoczone są fragmenty skompresowane (A', B').

Tabela 3. Przykłady kompresji semantycznej dla języka angielskiego

A	The information from AgCam will provide useful data to agricultural producers in North Dakota and neighboring states, benefiting farmers and ranchers and providing ways for them to protect the environment.
A'	information will provide useful data economic producer american state adjective state benefit creator creator provide structure protect environment
B	Together the two groups make up nearly 70 percent of all flowering plants and are part of a larger clade known as Pentapetalae, which means five petals. Understanding how these plants are related is a large undertaking that could help ecologists better understand which species are more vulnerable to environmental factors such as climate change.
B'	together two group constitute percent group flowering plant part flowering plant known means five leafage understanding plant related large undertaking can help biologist better understand species more sensitive environmental factor such climate change.

Źródło: opracowanie własne.

3.3. Ocena jakości kompresji semantycznej

Przyjmijmy, że zadaniem systemu jest przetworzenie niezbyt specjalistycznego artykułu poświęconego najnowszym osiągnięciom w rozwoju antybiotyków. Po ustaleniu dziedziny dokumentu, co jest stosunkowo prostym zadaniem w ramach obecnie dostępnych systemów klasyfikacyjnych, można przystąpić do kompresji z wykorzystaniem sieci semantycznej. Przetwarzając ów artykuł, można zauważyć, że każde odniesienie do penicyliny lub streptomycyny jest miejscem stosownym do zastosowania kompresji. Sieć semantyczna zawiera relacje pozwalające na wywnioskowanie, że zarówno penicylina, jak i streptomycyna są antybiotykami. Rezultatem zastosowania kompresji jest skrócenie wektora opisującego dokument o dwa elementy poprzez zastąpienie konkretnych nazw antybiotyków ich generalizacją. Analogiczny proces może być zastosowany do kolejnych termów, których poziom specjalizacji odbiega od średniego poziomu artykułu.

Przygotowany został eksperyment mający określić, czy kompresję semantyczną można skutecznie zastosować w typowym dla *text miningu* zadaniu klasyfikacji dokumentów (zastosowano klasyfikację aglomeracyjną metodą Warda)[37]. W pierwszym przebiegu dokumenty klasyfikowane były w oryginalnej postaci, a w drugim przebiegu zostały poddane kompresji semantycznej. Do eksperymentu posłużyło 900 dokumentów tekstowych w języku angielskim z zakresu astronomii, biologii, ekonomii, kultury, medycyny, polityki, prawa oraz sportu. W celu zweryfikowania rezultatów wszystkie dokumenty zostały zaetykietowane manualnie listą kategorii, do których zostały zaliczone przez eksperta.

Klasyfikację metodą analizy skupień przeprowadzono ośmiokrotnie. Pierwszy przebieg został przeprowadzony bez zastosowania kompresji semantycznej: z użyciem ok. 25 000 konceptów. W następnych przebiegach algorytmu redukowano liczbę konceptów będących deskryptorami do 12 000, 10 000, 8000, 6000, 4000, 2000 – aż do kompresji semantycznej z użyciem 1000 deskryptorów konceptów.

Jakość klasyfikacji została obliczona poprzez porównanie dokonanej klasyfikacji z etykietami nadanymi dokumentom przez ekspertów. Uzyskane współczynniki jakości klasyfikacji zostały zaprezentowane w tabelach 4 oraz 5. W wynikach można zaobserwować, że nieznaczny spadek jakości klasyfikacji występuje dla silnej kompresji semantycznej (liczba konceptów zostaje zredukowana poniżej 4000).

---

[37] A. Hotho, S. Staab, G. Stumme, *Explaining Text Clustering Results Using Semantic Structures*, w: *Knowledge Discovery in Databases: PKDD 2003. 7th European Conference on Principles and Practice of Knowledge Discovery in Databases, Cavtat-Dubrovnik, Croatia, September 22-26, 2003. Proceedings*, red. N. Lavrač, D. Gamberger, H. Blockeel, L. Todorovski, PKDD, Springer Verlag, Berlin – Heidelberg 2003, „Lecture Notes in Computer Science" 2003, t. 2838, s. 217-228.

Tabela 4. Oszacowanie jakości klasyfikacji dokumentów oryginalnych (bez kompresji semantycznej) w proc.

Liczba cech / Liczba konceptów	1000	900	800	700	600	Średnia
bez kompresji	93,46	90,90	91,92	92,69	89,49	91,69
12 000 konceptów	91,92	90,38	90,77	88,59	87,95	89,92
10 000 konceptów	93,08	89,62	91,67	90,51	90,90	91,15
8000 konceptów	92,05	92,69	90,51	91,03	89,23	91,10
6000 konceptów	91,79	90,77	90,90	89,74	91,03	90,85
4000 konceptów	88,33	89,62	87,69	86,79	86,92	87,87
2000 konceptów	86,54	87,18	85,77	85,13	84,74	85,87
1000 konceptów	83,85	84,10	81,92	81,28	80,51	82,33

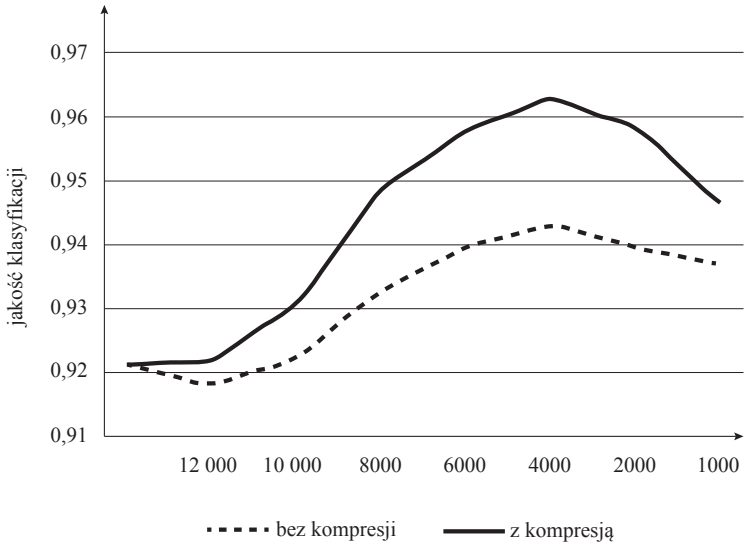
Ź r ó d ł o: opracowanie własne.

Tabela 5. Jakość klasyfikacji dokumentów przy zastosowaniu kompresji semantycznej w proc.

Liczba cech / Liczba konceptów	1000	900	800	700	600	Średnia
wszystkie koncepty	94,78	92,50	93,22	91,78	91,44	92,11
12 000 konceptów	93,56	93,39	93,89	91,50	91,78	92,20
10 000 konceptów	95,72	94,78	93,89	91,61	92,17	93,08
8000 konceptów	95,89	95,83	94,61	95,28	94,72	94,86
6000 konceptów	96,94	96,11	96,28	96,17	95,06	95,77
4000 konceptów	96,83	96,33	96,89	96,06	96,72	96,27
2000 konceptów	97,06	96,28	95,83	96,11	95,56	95,83
1000 konceptów	96,22	95,56	94,78	94,89	94,00	94,66

Ź r ó d ł o: opracowanie własne.

Rysunek 1 pokazuje jakość klasyfikacji uzyskanej w obu zadaniach. Utrata jakości klasyfikacji dokumentów jest nieznaczna dla kompresji, która redukuje liczbę deskryptorów konceptów do 4000. Natomiast silniejsza kompresja i związana z nią mniejsza liczba deskryptorów powoduje znaczne zmniejszenie jakości klasyfikacji, która jednak pozostaje na akceptowalnym poziomie.



Rys. 1. Jakość klasyfikacji dokumentów określająca w procentach prawidłowo sklasyfikowane dokumenty

Ź r ó d ł o: opracowanie własne.

4. Zastosowania kompresji semantycznej

Jednym z najważniejszych zastosowań kompresji semantycznej okazała się metoda półautomatycznej rozbudowy samej sieci. Inne zastosowanie polega na takim zaprezentowaniu użytkownikowi dokumentów w sposób uogólniony, aby dopasować uzyskane w wyniku uogólnienia pojęcia do stopnia kompetencji użytkownika w dziedzinie, której dotyczy dokument.

4.1. Rozbudowa sieci semantycznej. Algorytm regułowego wykrywania pojęć i relacji leksykalnych

Do wykrywania nowych pojęć skonstruowano algorytm, który został następnie użyty w eksperymencie z wykorzystaniem sieci semantycznej WiSENet. Pierwszym krokiem w algorytmie jest procedura rozwijająca reguły w zbiory hiponimów zawartych w sieci. Operacja ta jest kosztowna czasowo ze względu na konieczność przejścia po wszystkich możliwych krawędziach łączących wybrane koncepty oraz wszystkich konceptach końcowych w użytej sieci semantycznej.

Następnym etapem jest przeanalizowanie tekstów pod kątem pasujących fragmentów w tekstach. Operacja ta jest wykonywana z wykorzystaniem mechanizmu o nazwie *bag of concepts*, zaimplementowanego jako automat skończony wyposażony w zaawansowane metody wyzwalające zaprogramowane operacje. W każdym stanie automatu sprawdza on, czy którakolwiek z reguł wymagających sprawdzenia jest spełniona. Kompletny opis algorytmu znajduje się w pracy Ceglarka i wsp.[38], w pseudokodzie pokazany jest w Algorytmie 5.

Tabela 6. Przykład reguły i rezultatów działania automatu *bag of concepts*

Reguła: disease (wszystkie hiponimy), therapy (wszystkie hiponimy)
Znaleziony fragment: chemotherapy drug finish off remaining cancer **Koncepty spełniające regułę**: therapy → chemotherapy, disease → cancer **Ignorowane**: drug finish off remaining
Znaleziony fragment: gene therapy development lymphoma say woods **Koncepty spełniające regułę**: therapy → gene therapy, disease → lymphoma **Ignorowane**: development
***Znaleziony fragment**: cancer by-bid using surgery chemotherapy* ***Koncepty spełniające regułę**: therapy → chemotherapy, disease → cancer* ***Ignorowane**: by-bid using surgery*

Ź r ó d ł o: opracowanie własne.

Podany w tabeli 6 przykład pochodzi z eksperymentów przeprowadzonych na korpusie 2589 angielskich tekstów z dziedziny biologii i medycyny (łącznie dokumenty zawierały ponad 9 milionów słów). Rezultatem eksperymentu było znalezienie 471 nowych pojęć, które zostały następnie dodane do sieci WiSENet.

4.2. Mechanizm wspomagania rozumienia tekstu

Dziedzinowa kompresja semantyczna została sprawdzona również w zastosowaniu społecznościowym. Dla dokumentów w języku polskim (z dziedziny astronomii, biologii oraz astrobiologii) przeprowadzony został eksperyment z użyciem sieci semantycznej SenecaNet wraz z dodatkowym mechanizmem wykorzystującym analizator morfologiczny Morfologik[39]. Eksperyment polegał na dostosowaniu siły kompresji semantycznej do potrzeb użytkownika (w zależności od deklarowanego stopnia posiadanych kompetencji w danej dziedzinie). Jednocześnie zadaniem systemu było zaprezentowanie przekształconego tekstu w formie

[38] D. Ceglarek, K. Haniewicz, W. Rutkowski, *Towards Knowledge Acquisition…*

[39] M. Miłkowski, op. cit.

Algorytm 5. Algorytm automatu skończonego *bag of concepts* do wykrywania reguł z użyciem sieci semantycznej WiSENet

```
//przypisanie wyzwalaczy reguł do konceptów w sieci semantycznej
mapRulesToSemNet(SN, R[])
for all Rule ∈ R do
  for all Term, Relations ∈ Rule do
    N = SN.getNeighbourhood(Term, Relations)
    for all Term ∈ N do
      SN.createRuleTrigger(Term, Rule)
    end for
  end for
end for
//wygładzenie tekstu: tokenizacja, zastosowanie stop-listy, wykrywanie pojęć.
T = analyzeText(Input)
for each Term ∈ T
  if count(Bag) = size(Bag) then
    //deaktywowanie licznika wystąpień dla reguł związanych z termem Term.
    //wyjęcie termu ze zbioru Bag.
    oldTerm = pop(Bag)
  end if
  for all Rule ∈ SN.getTriggers(oldTerm) do
    Rule.unhit(Term)
    push(Bag, Term)
    for all Rule ∈ SN.getTriggers(Term) do
      //pobranie relewantnych reguł i aktywowanie licznika wystąpień termu.
      Rule.hit(Term)
      if Rule.hitCount = Rule.hitRequired then
        //wyświetlenie raportu informującego o spełnieniu reguły Rule
        Report(Rule, Bag)
      end if
  end for
end for
```

SN – sieć semantyczna WiSENet
R – zbiór reguł semantycznych
Bag – zbiór termów aktywowanych za pomocą automatu skończonego

Tabela 7. Przykład oryginalnego i skompresowanego fragmentu tekstu w języku polskim z wykorzystaniem analizatora morfologicznego Morfologik

Tekst oryginalny: „Zaćmienie księżyca"
O godzinie 19:42:06 Księżyc dotknie cienia Ziemi. Stopniowo od wschodniej strony nasz satelita będzie „pożerany" przez cień naszej planety. O godzinie 20:49:34 cień całkowicie pochłonie Księżyc. Jego barwa powinna stać się krwisto czerwona na skutek oświetlenia promieniami słonecznymi zagiętymi w ziemskiej atmosferze. Maksimum zaćmienia wypadnie o godzinie 21:20:36.
Tekst skompresowany (4000 deskryptorów dla konceptów z sieci)
O godzinie 19:42:06 Księżyc dotknie cienia Ziemi. Stopniowo od wschodniej strony nasz satelita będzie konsumowany przez cień naszej planety. O godzinie 20:49:34 cień całkowicie przyłączy Księżyc. Jego barwa powinna stać się kolorowo czerwona na skutek działania promieniami słonecznymi nierównymi w ziemskiej atmosferze. Maksimum zaćmienia **usunie** o godzinie 21:20:36.

Ź r ó d ł o : opracowanie własne.

zrozumiałej i poprawnej stylistycznie. Zastosowanie mechanizmu wykorzystującego Morfologik pozwoliło w sposób automatyczny dopasowywać formy deklinacyjne i koniugacyjne termów podlegających kompresji semantycznej[40]. W eksperymencie 95,5% transformacji zostało dokonanych poprawnie, uwzględniając wszelkie aspekty gramatyczne w języku polskim. Przykład ilustrujący uzyskane wyniki eksperymentu pokazany jest w tabeli 7.

5. Podsumowanie

Przeprowadzono szereg badań i eksperymentów, które miały na celu rozwinięcie koncepcji kompresji semantycznej i pokazanie jej rozmaitych zastosowań w dziedzinie przetwarzania języka naturalnego. Wyniki badań pokazały, że kompresja semantyczna może być z powodzeniem używana w rozmaitych zadaniach NLP. W pracy omówione zostały następujące istotne rezultaty przeprowadzonych badań:

– notacja SenecaNet dla sieci semantycznej,
– mechanizm globalnej i dziedzinowej kompresji semantycznej,

[40] Rozwiązanie zostało przedstawione w: D. Ceglarek, K. Haniewicz, W. Rutkowski, *Domain Based Semantic Compression for Automatic Text Comprehension Augmentation and Recommendation*, w: *Computational Collective Intelligence. Technologies and Applications. Third International Conference, ICCCI 2011, Gdynia, Poland, September 21-23, 2011, Proceedings*, t. 2, red. P. Jędrzejowicz, N.T. Nguyen, K. Hoang, Springer-Verlag, Berlin – Heidelberg 2011, „Lecture Notes in Computer Science" 2011, t. 6923, s. 40-49.

– mechanizm transformacji sieci semantycznej WordNet do formatu sieci SenecaNet,

– mechanizm łączący kompresję semantyczną z analizą morfologiczną do wspomagania rozumienia dokumentów w wybranych dziedzinach,

– automat skończony dla wyszukiwania nowych pojęć i nowych relacji leksykalnych.

W wyniku przeprowadzonych eksperymentów pokazano, że jakość klasyfikacji dokumentów z wykorzystaniem kompresji semantycznej wzrasta z 92,11% o dodatkowe 4,16%. Dzięki kompresji semantycznej możliwe stało się zbudowanie mechanizmu posługującego się stosunkowo ogólnymi regułami, które skutecznie wykrywają nowe pojęcia w dokumentach tekstowych. Autor zamierza wyszukać nowe zastosowania dla kompresji semantycznej. Dodatkowym zadaniem badawczym jest też udoskonalenie narzędzi i metod służących do w pełni automatycznej rozbudowy sieci semantycznej WiSENet.

Literatura

Baeza-Yates R.A., Ribeiro-Neto B., *Modern Information Retrieval*, Addison-Wesley Longman Publishing, Boston 1999.

Baziz M., *Towards a Semantic Representation of Documents by Ontology-Document Mapping*, w: *Artificial Intelligence: Methodology, Systems, and Applications. 11th International Conference, AIMSA 2004, Varna, Bulgaria, September 2-4, 2004. Proceedings*, red. Ch. Bussler, D. Fensel, Springer, 2004, „Lecture Notes in Computer Science" 2004, t. 3192, s. 33-43.

Boyd-Graber J., Blei D.M., Zhu X., *A Topic Model for Word Sense Disambiguation*, w: *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, Prague, June 2007*, s. 1024-1033.

Burrows S., Tahaghoghi S.M.M., Zobel J., *Efficient plagiarism detection for large code repositories*, „Software: Practice and Experience" 2007, t. 37, nr 2, s. 151-175.

Ceglarek D., *Zastosowanie sieci semantycznej do disambiguacji pojęć w języku naturalnym*, w: *Systemy wspomagania organizacji SWO 2006*, Wyd. AE w Katowicach, Katowice 2006.

Ceglarek D., *Koncepcja komponentowego systemu ochrony własności intelektualnej wykorzystującego semantyczne struktury informacji*, w: *Technologie informatyczne w zarządzaniu wiedzą – uwarunkowania i realizacja*, red. P. Adamczewski, M. Zakrzewicz, Wyd. WSB w Poznaniu, Poznań 2009.

Ceglarek D., Haniewicz K., Rutkowski W., *Quality of Semantic Compression in Classification*, w: *Computational Collective Intelligence, Second International Conference, ICCCI 2010, Kaohsiung, Taiwan, November 10-12, 2010. Proceedings*, cz. 1, red. J.-S. Pan, S.-M. Chen, N.T. Nguyen, Springer-Verlag, Berlin – Heidelberg 2010, „Lecture Notes in Computer Science" 2010, t. 6421, s. 162-171.

Ceglarek D., Haniewicz K., Rutkowski W., *Semantic Compression for Specialised Information Retrieval Systems*, w: *Advances in Intelligent Information and Database Systems*, red. N.T. Nguyen, R. Katarzyniak, S.-M. Chen, Springer Verlag, Berlin – Heidelberg 2010, „Studies in Computational Intelligence" 2010, t. 283, s. 111-121.

Ceglarek D., Haniewicz K., Rutkowski W., *Domain Based Semantic Compression for Automatic Text Comprehension Augmentation and Recommendation*, w: *Computational Collective Intelligence. Technologies and Applications. Third International Conference, ICCCI 2011, Gdynia, Poland, September 21-23, 2011, Proceedings*, t. 2, red. P. Jędrzejowicz, N.T. Nguyen, K. Hoang, Springer-Verlag, Berlin – Heidelberg 2011, „Lecture Notes in Computer Science" 2011, t. 6923, s. 40-49.

Ceglarek D., Haniewicz K., Rutkowski W., *Towards Knowledge Acquisition with WiSENet*, w: *New Challenges for Intelligent Information and Database Systems*, red. N.T. Nguyen, B. Trawinski, J.J. Jung, Springer Verlag, Berlin – Heidelberg 2011, „Studies in Computational Intelligence" 2011, t. 351, s. 75-84.

Ceglarek D., Haniewicz K., *Fast Plagiarism Detection by Sentence Hashing*, w: *Artificial Intelligence and Soft Computing. 11th International Conference, ICAISC 2012, Zakopane, Poland, April 29-May 3, 2012, Proceedings*, t. 2, red. L. Rutkowski, M. Korytkowski, R. Scherer, R. Tadeusiewicz, L.A. Zadeh, J.M. Zurada, Springer-Verlag, Berlin – Heidelberg 2012, „Lecture Notes in Computer Science" 2012, t. 7268, s. 30-38.

Erk K., Pad'o S., *A Structured Vector Space Model for Word Meaning in Context*, w: *EMNLP '08 Proceedings of the Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, Stroudsburg, PA, USA 2008, s. 897-906.

Hotho A., Staab S., Stumme G., *Explaining Text Clustering Results Using Semantic Structures*, w: *Knowledge Discovery in Databases: PKDD 2003. 7th European Conference on Principles and Practice of Knowledge Discovery in Databases, Cavtat-Dubrovnik, Croatia, September 22-26, 2003*, *Proceedings*, red. N. Lavrač, D. Gamberger, H. Blockeel, L. Todorovski, PKDD, Springer Verlag, Berlin – Heidelberg 2003, „Lecture Notes in Computer Science" 2003, t. 2838, s. 217-228.

Information Retrieval: Data Structures & Algorithms, red. W.B. Frakes, R.A. Baeza-Yates, Prentice-Hall, 1992.

Krovetz R., Croft W.B., *Lexical ambiguity and information retrieval*, „ACM Transactions on Information Systems" 1992, nr 10, s. 115-141.

Lukashenko R., Graudina V., Grundspenkis J., *Computer-based plagiarism detection methods and tools: an overview*, w: *Proceedings of the 2007 International Conference on Computer Systems and Technologies, CompSysTech '07. New York, USA*, ACM, 2007, s. 401-406.

Miller G.A., *Wordnet: a lexical database for English*, „Communications of the ACM" 1995, t. 38, nr 11.

Miłkowski M., *Automated Building of Error Corpora of Polish*, w: *Corpus Linguistics, Computer Tools, and Applications – State of the Art*, PALC 2007, red. B. Lewandowska-Tomaszczyk, Peter Lang, Frankfurt am Main 2008, s. 631-639.

Nock R., Nielsen F., *On weighting clustering*, „The IEEE Transactions on Pattern Analysis and Machine Intelligence" 2006, nr 28(8), s. 1223-1235.

Ota T., Masuyama S., *Automatic plagiarism detection among term papers*, w: *Proceedings of the 3rd International Universal Communication '09*, ACM, 2009, s. 395-399.

Percova N.N., *On the types of semantic compression of text*, w: *COLING '82. Proceedings of the 9th conference on Computational linguistics*, t. 2, Academia Praha, 1982, s. 229-231.

Rosenzweig J., Mihalcea R., Csomai A., *„WordNet bibliography". Web page: a bibliography referring to research involving the WordNet lexical database*, http://lit.csci.unt.edu/%7Ewordnet [1.09.2007].

Sanderson M., *Word Sense Disambiguation and Information Retrieval*, w: *SIGIR '94. Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval*, red. W.B. Croft, C.J. van Rijsbergen, SIGIR, ACM/Springer, New York 1994, s. 142-151.

Sanderson M., *Retrieving with Good Sense*, „Information Retrieval" 2000, t. 2, nr 1, s. 49-69.

Sinha R., Mihalcea R., *Unsupervised graph-basedword sense disambiguation using measures of word semantic similarity*, w: *International Conference on Semantic Computing ICSC 2007*, IEEE 2007, s. 363-369.

Snow R., Jurafsky D., Ng A.Y., *Learning syntactic patterns for automatic hypernym discovery*, w: *Advances in Neural Information Processing Systems (NIPS)*, 2005.

Staab S., Hotho A., *Ontology-based text document clustering*, w: *IIS, Advances in Soft Computing*, red. M.A. Kłopotek, S.T. Wierzchoń, K. Trojanowski, Springer, 2003, s. 451-452.

Stokoe Ch., Oakes M.P., Tait J., *Word Sense Disambiguation in Information Retrieval Revisited*, SIGIR, 2003.

Zeszyty Naukowe
Wyższej Szkoły Bankowej w Poznaniu
Nr 40/2012

Wojciech Fliegner
Wyższa Szkoła Bankowa w Poznaniu

Standaryzacja elektronicznej sprawozdawczości finansowej

Streszczenie. *Obecnie twórcy i odbiorcy sprawozdawczości finansowej mają do czynienia z wieloma różnymi formatami zapisu sprawozdań finansowych, co zwiększa czasochłonność oraz pracochłonność i tym samym koszty przetwarzania takich danych. Odpowiedzią na to wyzwanie ma być standard XBRL (*Extensible Business Reporting Language*). W artykule przedstawiono koncepcję XBRL, korzyści dla beneficjentów, stan prac i działania różnych krajowych organizacji zaangażowanych w rozwój i promocję tego standardu.*

Słowa kluczowe: *sprawozdawczość finansowa, XBRL, taksonomia*

1. Wprowadzenie

Coraz większy stopień globalizacji i wynikająca stąd wzmożona konkurencja wymagają od przedsiębiorstw doskonalenia systemów komunikowania się z otoczeniem. Jednym z narzędzi komunikowania się przedsiębiorstw z zewnętrznymi interesariuszami są sprawozdania finansowe generowane przez rachunkowość. Sprawozdawczość finansowa ma odpowiadać potrzebom informacyjnym licznych odbiorców: inwestorów, analityków finansowych, biur maklerskich, giełd papierów wartościowych, urzędów skarbowych, urzędów statystycznych, banków czy środowisk akademickich. Trwają wciąż prace zmierzające do standaryzacji i harmonizacji zasad rachunkowości w wymiarze międzynarodowym, tak by sprawozdania finansowe były bardziej zrozumiałe i ujednolicone bez względu na kraj ich przygotowywania. Z drugiej strony – istnieje pilna potrzeba ujednolicenia sposobu zapisu i przesyłania danych ze sprawozdań finansowych, aby dostęp, analiza i porównywanie tych danych było łatwiejsze i szybsze. Do niedawna brakowało

jednolitego standardu, który pozwoliłby sprawozdania finansowe zapisywać w postaci elektronicznej, przesyłać i przejmować przez inne aplikacje. W efekcie proces sprawozdawczości cechuje się różnorodnością standardów, zapisu danych oraz narzędzi służących do ich transferu, co zwykle oznacza dodatkowe koszty zarówno dla tworzących, jak i dla odbiorców sprawozdań.

Technologie informatyczne (w tym zwłaszcza Internet) wywierają istotny wpływ na proces komunikacji gospodarczej, w związku z czym w dziedzinie szeroko rozumianej sprawozdawczości finansowej pojawiły się rozwiązania określane jako: *web-based reporting*, *internet-based reporting*, *on-line reporting* oraz *real-time reporting*[1].

Badania dotyczące wytycznych oraz standardów regulujących informatyczne aspekty sprawozdawczości finansowej były początkowo związane z odwołaniami do technologii XML[2]. Do najważniejszych raportów z tego obszaru badań można zaliczyć publikacje: Charlesa Hoffmana i Carolyn Strand, Bryana Bergerona oraz Rogera Debreceny'ego i wsp.[3]

2. Charakterystyka standardu XBRL

Standard XBRL (*eXtensible Business Reporting Language*)[4] jest rozwiązaniem problemu elektronicznej sprawozdawczości finansowej, w szczególności w aspektach istotnych dla środowiska księgowych. XBRL to oparty na języku XML[5]

[1] Zob. C.E. Davis, C. Clements, W.P Keuer, *Web-based Reporting: a Vision for the Future*, „Strategic Finance", September 2003.

[2] Należy podkreślić, że formaty danych, takie jak PDF, XLS oraz HTML, nie rozwiązują problemu automatycznego przesyłu sprawozdań finansowych, bowiem dane przesyłane w tych formatach trzeba ponownie ręcznie wprowadzić do systemów informatycznych. Dopiero języki oparte na standardzie XML można traktować jako elektroniczne standardy sprawozdawczości, gdyż dane przesyłane w ten sposób mogą automatycznie być konsumowane przez systemy odbiorców.

[3] C. Hoffman, C. Strand, *XBRL Essentials*, American Institute of Certified Public Accountants, New York 2001; B. Bergeron, *Essentials of XBRL. Financial Reporting in the 21st Century*, Wiley, New Jersey 2003; R.S. Debreceny, C. Felden, B. Ochocki, M. Piechocki, M. Piechocki, *XBRL for Interactive Data. Engineering the Information Value Chain*, Springer, Heidelberg 2009.

[4] Historia powstania standardu XBRL sięga roku 1998 oraz pionierskich prac Charlesa Hoffmana w zakresie możliwości wykorzystania języka XML dla potrzeb raportowania finansowego. W sierpniu 1999 r. powołano Komitet Sterujący XBRL, który przekształcił pilotażowy projekt FR-XML (*Financial Reporting XML*) w międzynarodowy standard raportowania biznesowego oraz rozpoczął prace nad pierwszą oficjalną specyfikacją języka.

[5] Zgodność standardu XBRL z językiem XML została zapewniona poprzez współpracę konsorcjum XBRL International z organizacją World Wide Web Consortium (W3C). Konsorcjum XBRL zrzesza ponad 250 podmiotów, będących liderami branży finansowej oraz IT, a także regulatorów, tj. organizacje ustanawiające standardy rachunkowości, takie jak: Komitet Międzynarodowych Standardów Rachunkowości (International Accounting Standard Committee), Rada ds. Standardów

elektroniczny format wymiany zestandaryzowanych informacji finansowych, najczęściej w postaci sprawozdań finansowych.

Specyfikacja 2.1 języka XBRL[6] opublikowana w grudniu 2003 r. (wraz z późniejszymi erratami) – bazując na specyfikacji XML jako bardziej ogólnym języku znaczników – szczegółowo charakteryzuje składowe standardu XBRL.

2.1. Merytoryczne aspekty standardu

Zastosowanie standardu XBRL w praktyce oznacza publikowanie informacji w dokumencie elektronicznym z wykorzystaniem tzw. taksonomii, która definiuje strukturę sprawozdania (raportu) finansowego.

Taksonomie XBRL mają postać słowników tematycznych definiujących pojęcia (każdy z nich obejmuje od kilku do kilkunastu tysięcy pojęć), do których odnoszą się raportowane informacje. Informacje źródłowe dotyczące pojęć definiowanych w taksonomiach znajdują się w odpowiednich regulacjach prawnych (ustawie o rachunkowości, Międzynarodowych Standardach Sprawozdawczości Finansowej, Nowej Umowie Kapitałowej itp.), stąd standard XBRL nie stanowi sam w sobie regulacji prawnych, a jest jedynie ich elektronicznym odzwierciedleniem[7].

Poza definicją pojęć, taksonomia XBRL charakteryzuje szczegółowo każdy z elementów sprawozdania, określając jego typ (monetarny, procentowy, tekstowy itp.), wymiar czasowy (wartość za okres, na początek lub na koniec okresu sprawozdawczego) oraz charakter księgowy (winien lub ma). Co więcej – taksonomia XBRL pozwala na przypisanie danemu elementowi etykiet w wielu językach, umożliwiając przez to dostęp do danych finansowych użytkownikom różnych narodowości. Dzięki dokładnemu scharakteryzowaniu każdej pozycji sprawozdania finansowego dane przesyłane przy wykorzystaniu taksonomii XBRL mogą być w prosty sposób walidowane, przetwarzane oraz analizowane, co z kolei obniża prawdopodobieństwo wystąpienia błędu, pojawiające się w przypadku wielokrotnego, ręcznego wprowadzania danych.

Rachunkowości Finansowej (Financial Accounting Standard Board) i organizacje zawodowe zrzeszające audytorów (biegłych rewidentów), np. American Institute of Certified Public Accountants (AICPA).

[6] www.xbrl.org/Specification/XBRL-RECOMMENDATION-2003-12-31+Corrected-Errata-2005-11-07.htm#_6 [5.10.2012].

[7] Przykładowo, w taksonomii opartej na ustawie o rachunkowości zdefiniowane zostały takie terminy, jak: aktywa, zysk netto, podczas gdy taksonomia oparta na Międzynarodowych Standardach Sprawozdawczości Finansowej definiuje pojęcia *goodwill* (ogół niematerialnych składników przedsiębiorstwa, składających się na jego wartość rynkową) czy też *depreciation and amortisation* (amortyzacja).

Jednak sprawozdawczość, a w szczególności sprawozdawczość finansowa, to nie tylko lista zdefiniowanych pozycji. Przykładowo, w bilansie firmy występuje wiele zależności dotyczących prezentacji oraz obliczania danej pozycji, jak też jej nazewnictwa oraz umocowania w regulacjach prawnych. Tak więc w skład taksonomii, poza definicją samych pozycji, wchodzi pięć warstw: prezentacji, obliczeń, definicji, referencji oraz etykiet.

Warstwa prezentacji określa, w jaki sposób zdefiniowane pozycje finansowe powinny zostać umieszczone w sprawozdaniu. Na przykład pozycja zbiorcza bilansu „Aktywa razem" znajduje się pod pozycją „Aktywa trwałe". **Warstwa obliczeń** określa zależność między pozycjami „Aktywa trwale" i „Aktywa bieżące" oraz „Aktywa razem". Dwie pierwsze pozycje sumują się do pozycji ostatniej. **Warstwa definicji** pozwala na określenie dodatkowych powiązań między poszczególnymi pozycjami, które nie są regulowane przez warstwy prezentacji oraz obliczeń. Ostatnie dwie warstwy, **referencji** oraz **etykiet**, służą do wskazania aktu prawnego, w którym dana pozycja sprawozdania jest zdefiniowana, łącznie z jego numerem i paragrafem, a także do umieszczenia nazwy pozycji sprawozdania finansowego w danym języku w postaci etykiety[8].

Taksonomie dzielą się na dwa rodzaje: sprawozdawcze (FR – *financial reporting*) oraz księgi głównej (GL – *general ledger*). W grupach roboczych XBRL nieoficjalnie przyjęto, że taksonomie sprawozdawcze koncentrują się na raportach sprawozdawczych (forma dokumentów), natomiast GL na samych danych (bez formatu sprawozdania).

XBRL pozwala na ujęcie w sprawozdaniach pozycji specyficznych dla danej branży, a nawet poszczególnych przedsiębiorstw, poprzez stworzenie odpowiednich rozszerzeń do danej taksonomii bazowej[9] (zarówno zwiększających, jak i ograniczających zakres sprawozdawczy taksonomii bazowej).

2.2. Technologiczne aspekty standardu

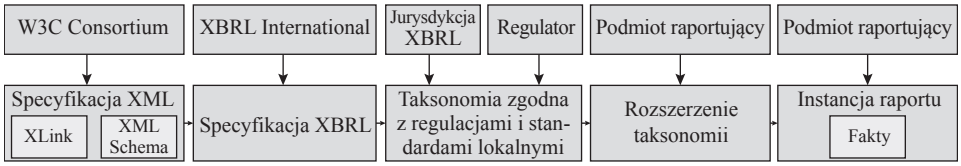
Generowanie sprawozdania finansowego w formacie XBRL odbywa się na podstawie taksonomii XBRL oraz danych finansowych (księgowych) dotyczących okresu obrachunkowego. Dokument XBRL składa się z instancji (*XBRL instance*), czyli zbioru raportowanych informacji (faktów), oraz z taksonomii (*taxonomy*), która jest słownikiem definiującym pojęcia (*concepts*), do których

[8] Zob. R.S. Debreceny, C. Felden, M. Piechocki, *New Dimensions of Business Reporting and XBRL*, DUV, Wiesbaden 2007.

[9] Zalicza się do nich takie taksonomie, jak: taksonomia IFRS (jest ona oparta na Międzynarodowych Standardach Sprawozdawczości Finansowej MSSF i Międzynarodowych Standardach Rachunkowości MSR – zob. C. Hoffman, *Financial Reporting Using XBRL – IFRS and US GAAP Edition*, Lulu Publishing House, Chicago 2005) oraz COREP (jest to taksonomia stworzona dla potrzeb nadzoru bankowego oraz uwzględniająca regulacje zawarte w Nowej Umowie Kapitałowej), a także wiele innych krajowych bądź branżowych taksonomii.

odnoszą się fakty (*facts*) i jednocześnie klasyfikacją[10] (czyli usystematyzowaniem – stąd słowo taksonomia) tych pojęć.

Relacje między specyfikacją XML oraz standardem XBRL w postaci specyfikacji, taksonomii i dokumentu elektronicznego przedstawia rysunek 1.

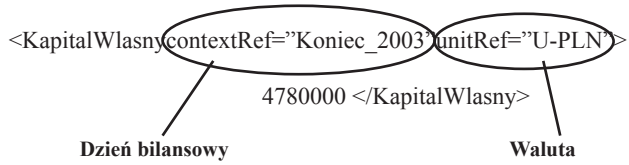


Rys. 1. Powiązanie XML z XBRL i dokumentem elektronicznym

Źródło: opracowanie własne.

W przypadku standardowych sprawozdań dokument XBRL zawiera podstawowe informacje na temat raportującego podmiotu (nazwa firmy, forma prawna, adres siedziby głównej itp.) oraz wartości finansowe poszczególnych pozycji sprawozdania osadzone w odpowiednim kontekście czasowym, walutowym oraz, jeśli jest to określone w taksonomii, wymiarowym. Jednoznaczne wskazanie źródłowej taksonomii zapewnia jednolitą interpretację faktów tworzących daną instancję, zarówno przez nadawców i odbiorców dokumentów XBRL, jak i przez aplikacje wykorzystywane do eksportowania i importowania tych dokumentów.

Zastosowanie języka XML przy tworzeniu standardu XBRL zmieniło podejście do klasycznego sprawozdania finansowego. Nie jest ono traktowane jako blok tekstu – tak jak w przypadku danych finansowych publikowanych na stronach internetowych bądź w drukowanych dokumentach – lecz każdemu faktowi towarzyszy odpowiedni znacznik (*tag*) wywodzący się z danej taksonomii (zob. rys. 2), co daje możliwość „inteligentnego" rozpoznawania poszczególnych informacji zawartych w dokumencie XBRL.



Rys. 2. Powiązanie XML z XBRL i dokumentem elektronicznym

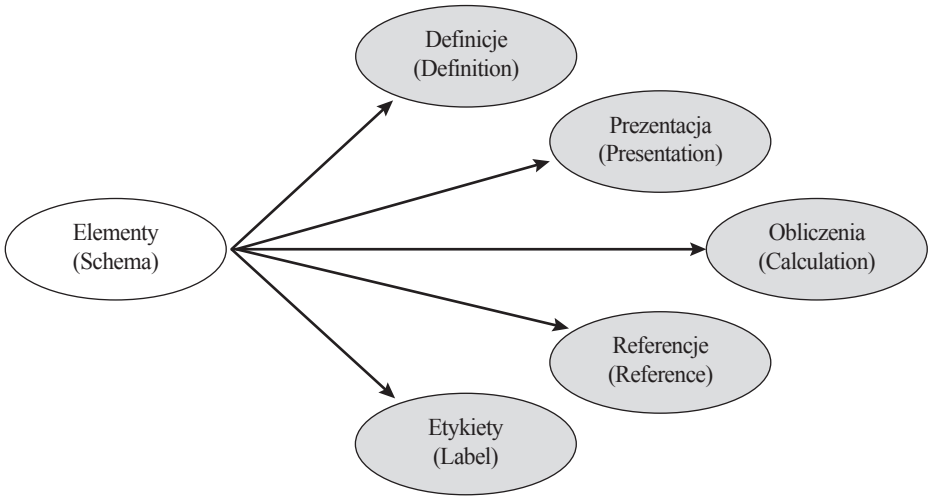
Źródło: opracowanie własne.

[10] Klasyfikacja ta, definiując relacje między pojęciami, określa również reguły kalkulacji wykorzystywane do weryfikacji danych.

Sama tylko obecność znaczników nie gwarantuje realizacji cechy dokumentów XML określanej jako samoopisywalność, o ile znacznikom nie towarzyszy zewnętrzny słownik, który ściśle definiuje ich znaczenie i zależności między nimi. Możliwości tworzenia powiązań w HTML są jednak mocno ograniczone. W sukurs przychodzą tu dwie dodatkowe rekomendacje W3C: XML Linking Language (XLink)[11] i XPointer[12]. XLink umożliwia definiowanie prostych i złożonych powiązań między pojęciami, mogącymi być zarówno elementami dokumentów XML, jak i zasobami zewnętrznymi. Dzięki temu jedne pojęcia mogą się wywodzić i odwoływać do innych. XPointer pozwala wskazać, o który element chodzi, czyli zapewnia składnię potrzebną w XLink do budowania wyrażeń wskazujących konkretny fragment dokumentu XML.

Tak więc wszystkie pojęcia taksonomii znajdują się w schemacie XSD (*XML Schema Definition*)[13], natomiast do zdefiniowania relacji między pojęciami służą tzw. warstwy powiązań (*linkbases*) związane ze specyfikacją XLink.

Architektura taksonomii została przedstawiona na rysunku 3.



Rys. 3. Architektura taksonomii

Ź r ó d ł o: opracowanie własne.

[11] Zob. *XML Linking Language (XLink) Version 1.0*, W3C Recommendation, 2001.

[12] Zob. *XPointer Framework*, W3C Recommendation, 2003.

[13] Schemat jest najważniejszym składnikiem taksonomii, ponieważ to on określa zawartość i strukturę instancji, czyli tej części dokumentu XBRL, która zawiera konkretne informacje – schemat realizuje to poprzez: a) definiowanie pojęć – i to jest jego główna rola, b) importowanie schematów związanych z taksonomiami wyższego poziomu (w przypadku taksonomii tworzonej na niskim poziomie, np. konkretnej firmy), c) definiowanie odwołań do warstw powiązań (*linkbases*).

Standardowo w taksonomii występują następujące warstwy powiązań:

– Labels Linkbase – zawiera etykiety pojęć, do wykorzystania przy wizualizacji instancji raportu. Umożliwia tworzenie wersji językowych taksonomii.

– Reference Linkbase – pokazuje odwołania do źródeł zewnętrznych, definiujących dane pojęcie, jak np. obowiązujące akty prawne.

– Definition Linkbase – definiuje relacje między pojęciami, takie jak: *general-special* (relacja generalizacji), *essence-alias* (oznaczająca inne ujęcie tego samego pojęcia), *requires-element* (pokazująca, że jeden element wymaga wystąpienia innego).

– Presentation Linkbase – określa, jak pojęcia są zorganizowane, czyli ich hierarchię, kolejność oraz organizację w postaci np. wierszy tabeli. Umożliwia późniejsze formatowanie raportu.

– Calculation Linkbase – definiuje podsumowania i proste wyliczenia poprzez przypisanie do każdego pojęcia wagi +1 (wielkość jest dodawana) lub –1 (odejmowanie) oraz wskazanie pojęcia nadrzędnego, do którego inne w grupie powinny się sumować.

Taksonomia XBRL jest więc rozszerzeniem XML Schema, ponieważ nie ogranicza się tylko do specyfikacji samych atrybutów poszczególnych znaczników i ich możliwych wartości, ale także definiuje relacje między nimi. Umożliwia to automatyczne weryfikowanie nie tylko syntaktycznej postaci dokumentu, ale i jego semantyki, np. poprzez sprawdzenie, czy podsumowania rzeczywiście są uzyskiwane z sumowania poszczególnych składników.

Podstawową składową instancji są fakty. Aby zapewnić ich jednoznaczną interpretację, towarzyszą im dodatkowe informacje, takie jak:

– kontekst – kontekstem jest czas, którego dotyczy raport, informacje o raportującym podmiocie (*context entity*) lub opcjonalnie scenariusz (*scenario*). Dla różnych typów raportów czasem może być konkretny dzień lub przedział czasu. Z kolei scenariusz mówi o tym, z jakiego typu faktami mamy do czynienia (*actual*, *budget*). Kontekst jest dla faktu identyfikowany poprzez atrybut contextRef,

– informacje o używanych jednostkach miar – każda wartość liczbowa musi być wyrażona w jakiejś jednostce. Informacja o jednostce jest obecna w elemencie tworzącym fakt w postaci atrybutu unitRef,

– dokładność – wszystkie fakty o wartościach liczbowych muszą mieć także określoną dokładność.

Fakty stanowią płaską strukturę, a ich układ w raporcie jest zdeterminowany przez warstwę powiązań *Presentation*.

Procedura tworzenia sprawozdań finansowych w standardzie XBRL jest zatem zestawem następujących działań:

1. Pozyskanie odpowiedniej taksonomii.
2. Transformacja pojęć ze sprawozdania finansowego na ich odpowiedniki w taksonomii. W przypadku zidentyfikowania w sprawozdaniu wielkości, które nie mają w taksonomii swoich odpowiedników, konieczne będzie rozszerzenie taksonomii.

3. Przypisywanie znaczników do źródeł danych (tagowanie).
4. Generowanie sprawozdania finansowego – tworzenie instancji dokumentu XBRL.

W powyższej procedurze istnieje kilka miejsc, w których możliwe jest uzyskanie wspomagania komputerowego – lista narzędzi XBRL dostępna jest na stronie http://xbrl.us/vendors/Pages/default-expand.aspx i liczy obecnie 30 pozycji. Ponadto coraz więcej aplikacji klasy ERP, które przetwarzają dane finansowe, zaczyna także wspierać XBRL – np. firma SAP oferuje aplikację SAP® BusinessObjects™ XBRL Publishing jako narzędzie do tworzenia i udostępniania sprawozdań w standardzie XBRL.

Istnieją także polskie rozwiązania programistyczne wspierające XBRL. Dla przykładu, firma Rodan Systems jako producent oprogramowania wspierającego zarządzanie informacją, opartego na własnej platformie produktowej OfficeObjects®, oferuje oprogramowanie OfficeObjects®e-Forms umożliwiające tworzenie i publikację formularzy elektronicznych oraz inteligentne gromadzenie i zarządzanie danymi z wykorzystaniem standardu XBRL.

3. Komunikacja w łańcuchu raportowania finansowego

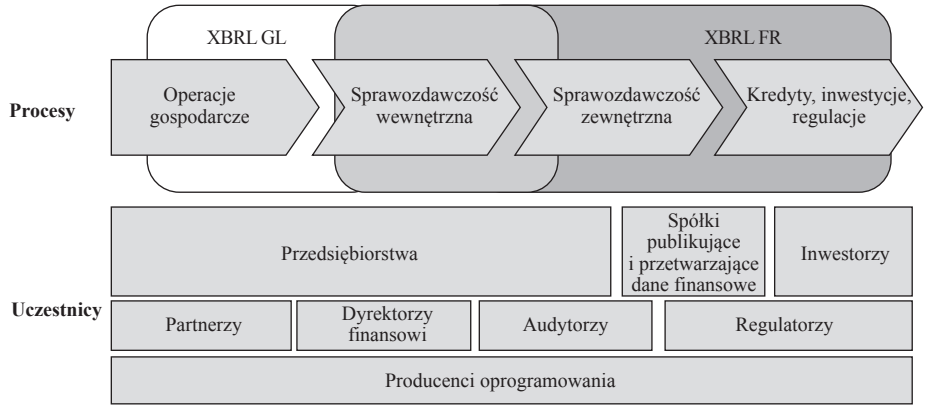
Koncepcja łańcucha sprawozdawczości finansowej[14] powstała jako bliźniacza koncepcja logistycznego łańcucha dostaw.

Łańcuch sprawozdawczości finansowej ma swój początek w momencie otrzymania przez podmiot informacji finansowych (np. faktury dotyczącej zakupu) związanych z prowadzonymi czynnościami gospodarczymi lub samym prowadzeniem tych czynności (przeniesienie materiałów do produkcji, wykonanie usługi i wystawienie faktury itp.). Dane te są gromadzone i przetwarzane w systemach księgowych, stanowiąc podstawy do generowania wewnętrznych i zewnętrznych raportów finansowych. Raporty wewnętrzne sporządzane są na potrzeby wąskiej grupy odbiorców podejmujących kluczowe decyzje w przedsiębiorstwie. W ramach łańcucha sprawozdawczości tworzone są także raporty zewnętrzne, w tym kwartalne i roczne sprawozdania finansowe, przeznaczone dla szerokiej i zróżnicowanej grupy odbiorców, do której należą m.in. instytucje kredytowe oraz instytucje publiczne. Dodatkowo, raportami spółek publicznych, których papiery notowane są na giełdach papierów wartościowych, zainteresowani są także inwestorzy (fundusze inwestycyjne, biura maklerskie, inwestorzy indywidualni) i organy regulujące rynki kapitałowe.

Automatyzacja poszczególnych procesów, w tym przede wszystkim transferu informacji w ramach łańcucha sprawozdawczości, może być wspomagana

[14] Por. B.M. Romney, P.J. Steinbart, *Accounting Information Systems*, Pearson Prentice Hall, Upper Sadle River 2006.

technologiami informatycznymi, w tym także standardem XBRL[15]. W pierwszym etapie może zostać zastosowana taksonomia księgi głównej XBRL. GL, która pozwala standaryzować takie obszary, jak: zestawienie kont, zapis listy płac, danych magazynowych itp., a więc informacje, które wykorzystywane są później do sporządzania wszelkiego rodzaju sprawozdań. W dalszych etapach powinny być dostępne taksonomie definiujące konkretne obszary sprawozdawczości w ramach taksonomii XBRL FR (rys. 4).



Rys. 4. Schemat łańcucha sprawozdawczości finansowej

Ź r ó d ł o: opracowanie własne.

Istotnymi cechami łańcucha sprawozdawczości finansowej są jego holistyczny aspekt oraz integracja różnych poziomów logicznych (uczestników, procesów, zasobów) dla celów uwzględnienia kompleksowej perspektywy sprawozdawczej.

Wśród uczestników łańcucha sprawozdawczości finansowej wyróżnić można następujące ich grupy:

- jednostki gospodarcze generujące dane oraz ich partnerów,
- zarząd, księgowych i audytorów,
- jednostki gromadzące i publikujące dane finansowe,
- inwestorów,
- instytucje rządowe i regulatorów,
- banki centralne,
- dostawców oprogramowania księgowego i finansowego.

[15] Por. N. Hannon, *Why Should Management Accountants Care about XBRL?*, „Strategic Finance", July 2004, s. 55-56 oraz A. Nut, M. Strauß, *eXtensible Business Reporting Language (XBRL). Konzept und praktischer Einsatz*, „Wirtschaftsinformatik" 2002, nr 44, s. 447-457.

Wszystkie wymienione wyżej jednostki biorą w różnym stopniu udział w różnorodnych procesach tworzenia, gromadzenia, analizy, przetwarzania i publikacji danych finansowych, a w szczególności danych sprawozdawczych.

Procesy generujące i przetwarzające dane sprawozdawcze można podzielić na (zob. rys. 4):

– operacje biznesowe,
– wewnętrzną sprawozdawczość finansową,
– zewnętrzną sprawozdawczość finansową,
– regulacje oraz polityki sprawozdawczości,
– regulacje związane z pozyskiwaniem źródeł kapitału oraz informowaniem inwestorów o wynikach gospodarczych jednostek.

Dzięki wykorzystaniu XBRL w łańcuchu sprawozdawczości finansowej można taniej, lepiej i szybciej zrealizować wszystkie etapy związane z przygotowaniem, rozpowszechnianiem i wykorzystywaniem sprawozdań finansowych[16]. W przekroju stron zaangażowanych w tym procesie oznaczać to może – przy spełnieniu określonych warunków – następujące korzyści:

a) dla producentów sprawozdań finansowych:

– obniżkę kosztów przygotowania i publikowania informacji – zastosowanie standardu XBRL wiąże się z realizacją zasady: „publikuj raz, wykorzystuj wiele razy",
– przyspieszenie i wzrost efektywności procesów decyzyjnych,
– dostarczanie informacji w czasie rzeczywistym wszystkim użytkownikom sprawozdań finansowych,
– zautomatyzowanie, a tym samym przyspieszenie procesu przekształcania informacji tworzonych w systemie rachunkowości w końcowy produkt tego systemu, którym jest zestaw sprawozdań finansowych,
– poprawę systemu informacji i kontroli wewnętrznej, wykorzystywanej w procesach zarządzania jednostką gospodarczą;

b) stronom wykorzystującym sprawozdania finansowe XBRL umożliwia:

– zwiększenie dostępu do informacji finansowych oraz obniżenie kosztów ich pozyskania,
– możliwość konsolidowania informacji pochodzących z różnych źródeł i systemów bez dodatkowych działań związanych z ich pozyskaniem,
– minimalizację wystąpienia błędu człowieka,
– wzrost tempa wykorzystania informacji, a zatem przyspieszenie podejmowania decyzji, które mogą być z tym związane;

c) dla pozostałych zainteresowanych dane finansowe stają się:

– bardziej dostępne i łatwiejsze do wykorzystania,
– szybsze do przetransferowania,
– bardziej wiarygodne w przypadku powiązania ich z podpisem elektronicznym.

[16] Szerzej na temat korzyści wynikających z zastosowania XBRL w: B.L. McGuire, S.J. Okesson, L.A. Watson, *Second – Wave Benefits of XBRL,* „Strategic Finance", December 2006.

Prace dotyczące przyszłego kształtu sprawozdawczości finansowej powinny koncentrować się na dwóch głównych problemach – rozwoju standardu XBRL oraz opracowaniu zestawu globalnych standardów rachunkowości, które zostałyby powszechnie zaakceptowane jako podstawa do tworzenia informacji prezentowanych w sprawozdaniach finansowych przedsiębiorstw, niezależnie od miejsca prowadzonej działalności gospodarczej.

4. Wdrożenia standardu XBRL w Polsce

W 2006 r. powstało Stowarzyszenie XBRL Polska jako organizacja odpowiedzialna za działania związane z promowaniem oraz rozwijaniem standardu XBRL, w tym związane z tworzeniem taksonomii zgodnej z polską ustawą o rachunkowości. Współpracuje ono z takimi instytucjami, jak: Stowarzyszenie Księgowych w Polsce, Krajowa Izba Biegłych Rewidentów, Narodowy Bank Polski, Ministerstwo Finansów oraz Stowarzyszenie Emitentów Giełdowych. Wspierają je również lub są członkami firmy audytorskie, takie jak: HLB, PwC oraz KPMG, a także producenci oprogramowania: BSB, Computerland, SAP, SAS Institute, Rodan. Na początku 2007 r. do Stowarzyszenia XBRL Polska zdecydowała się przystąpić Giełda Papierów Wartościowych w Warszawie.

Wdrażane obecnie zastosowanie standardu XBRL do gromadzenia danych w „Monitorze Polskim B”[17] powinno również usprawnić udostępnianie danych sprawozdawczych polskich podmiotów gospodarczych.

Istotnym krokiem we wdrażaniu standardu XBRL w sektorze bankowym była realizacja projektu COREP (*COmmon REPorting*), związana z zapisami Nowej Umowy Kapitałowej dotyczącymi raportowania informacji od banków komercyjnych do banków centralnych w UE. Z końcem października 2007 r. banki rozpoczęły przekazywanie swoich sprawozdań w standardzie XBRL. Zakres informacyjny tych sprawozdań przygotowany został przez Narodowy Bank Polski[18] na podstawie pakietów (taksonomii) FINREP[19] i COREP zdefiniowanych przez CEBS (Committee of European Banking Supervisors – Komitet Europejskich Nadzorców Bankowych).

Ponieważ w rozwój XBRL w Polsce angażują się najważniejsze instytucje i regulatorzy zajmujący się wieloma obszarami sprawozdawczości, a także

[17] Na początku lipca 2010 r. Centrum Obsługi Kancelarii Prezesa Rady Ministrów ogłosiło, iż została wybrana oferta na zakup systemu przesyłania sprawozdań finansowych drogą elektroniczną (zgodnie ze standardem XBRL) do „Monitora Polskiego B”.

[18] W części analitycznej systemu zaimplementowane zostały rozwiązania analityczne *Business Intelligence* przy wykorzystaniu narzędzi *Business Objects*.

[19] Taksonomia FINREP została rozszerzona przez Narodowy Bank Polski w celu umożliwienia polskiemu sektorowi bankowemu raportowania zakresu informacyjnego wymaganego przez NBP.

liderzy branż finansowej oraz IT, można oczekiwać ujednolicenia sprawozdawczości elektronicznej, co przyczyni się do obniżenia kosztów i optymalizacji procesu raportowania.

5. Podsumowanie

Zakres funkcjonalny standardu XBRL we wczesnej fazie jego tworzenia zamierzano ograniczyć tylko do potrzeb sprawozdawczości finansowej (stąd jego ówczesna nazwa: XFRML – *eXtensible Financial Reporting Markup Language*). Zakres ten został następnie rozszerzony (stąd jego obecna nazwa: *eXtensible Business Reporting Language*), aby umożliwić automatyzację procesu sprawozdawczego, eliminując przy tym konieczność manualnego i wielokrotnego wprowadzania danych na różnych etapach sprawozdawczości gospodarczej. Poza raportami finansowymi, takimi jak bilanse, rachunki zysków i strat czy rachunki przepływów, standard XBRL może służyć do standaryzacji zeznań podatkowych, raportów przesyłanych do GUS, raportów giełdowych, sprawozdań banków przesyłanych do organów nadzoru bankowego i wielu innych raportów. Co więcej, koncepcja *Enhanced Business Reporting*[20] ukazuje zastosowanie języka XBRL nie tylko do standaryzacji i przesyłu danych o charakterze ilościowym, ale także danych o charakterze jakościowym, które zajmują coraz istotniejsze miejsce w procesie podejmowania decyzji.

Nie bez znaczenia dla promocji standardu XBRL jest fakt, że jest to standard otwarty oraz wolny od opłat licencyjnych. Niewykluczone zresztą, że Komisja Papierów Wartościowych i Giełd wyda regulację wymuszającą stopniowe przejście na standard XBRL dla wszystkich firm raportujących, podobnie jak to uczynił jej amerykański odpowiednik, czyli SEC (Security and Exchange Commission), który zdecydował, że do 2011 r. wszystkie firmy swoje obowiązkowe raporty finansowe będą musiały przygotowywać przy użyciu i w formacie XBRL.

Zastosowanie XBRL w łańcuchu sprawozdawczości finansowej, czyli zaakceptowanie i zaadaptowanie tego rozwiązania przez wszystkich uczestników rynku globalnego, wprowadziłoby istotne zmiany w układzie: podmioty sporządzające sprawozdania finansowe (podaż informacji) – użytkownicy sprawozdań finansowych (popyt na informacje). Wynikałyby one przede wszystkim ze zwiększenia przejrzystości informacyjnej podmiotów gospodarczych oraz możliwości analizowania informacji finansowych poprzez porównywanie ich z podstawą odniesienia w skali branży, konkurentów na rynkach krajowych, regionalnych czy też globalnych, rozpatrywanych jednocześnie w różnych układach podmiotowo-przedmiotowo-przestrzenno-czasowych, bez konieczności korzystania z różnych

[20] Por. www.ebr360.com [5.10.2012].

dodatkowych źródeł. To z kolei dawałoby możliwość wywierania przez inwestorów i inne zainteresowane strony odpowiednio silnego nacisku na poprawę jakości informacji prezentowanych przez zarządy firm w sprawozdaniach finansowych i innych raportach przez nie przygotowywanych.

Literatura

Bergeron B., *Essentials of XBRL – Financial Reporting in the 21st Century*, Wiley, New Jersey 2003.

Davis C.E., Clements C., Keuer W.P., *Web-based Reporting: a Vision for the Future*, „Strategic Finance", September 2003.

Debreceny R.S., Felden C., Ochocki B., Piechocki M., Piechocki M., *XBRL for Interactive Data. Engineering the Information Value Chain*, Springer, Heidelberg 2009.

Debreceny R.S., Felden C., Piechocki M., *New Dimensions of Business Reporting and XBRL*, DUV, Wiesbaden 2007.

Hannon N., *Why Should Management Accountants Care about XBRL?*, „Strategic Finance", July 2004.

Hoffman C., *Financial Reporting Using XBRL – IFRS and US GAAP Edition*, Lulu Publishing House, Chicago 2005.

Hoffman C., Strand C., *XBRL Essentials*, American Institute of Certified Public Accountants, New York 2001.

McGuire B.L., Okesson S.J., Watson L.A., *Second – Wave Benefits of XBRL,* „Strategic Finance", December 2006.

Nut A., Strauß M., *eXtensible Business Reporting Language (XBRL) – Konzept und praktischer Einsatz*, „Wirtschaftsinformatik" 2002, nr 44, s. 447-457.

Romney B.M., Steinbart P.J., *Accounting Information Systems*, Pearson Prentice Hall, Upper Sadle River 2006.

XML Linking Language (XLink) Version 1.0, W3C Recommendation, 2001.

XPointer Framework, W3C Recommendation, 2003.

Zeszyty Naukowe
Wyższej Szkoły Bankowej w Poznaniu
Nr 40/2012

Jędrzej Musiał
Wyższa Szkoła Bankowa w Poznaniu

Rozszerzony problem optymalizacji zakupów internetowych

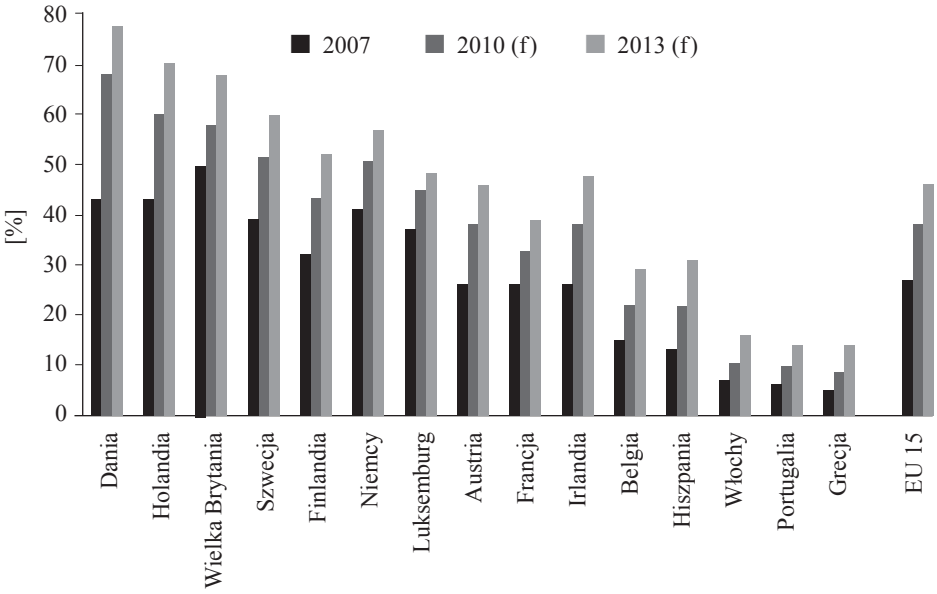
Streszczenie. *Problem optymalizacji zakupów internetowych (ISOP) dotyczy odpowiedzi na pytanie, w jaki sposób klient powinien dokonać zakupów określonych produktów spośród oferty sklepów internetowych. Z każdym sklepem i produktem związana jest oferta, a także możliwe jest zdefiniowanie dodatkowych wartości, jak np. koszt wysyłki (który może, ale nie musi być wartością stałą) czy funkcja określająca rabaty na zakup w danym sklepie. W pracy podano podstawową definicję problemu ISOP, a także zaprezentowano różne rozszerzenia tego problemu. Opisano zdefiniowany algorytm heurystyczny, a przeprowadzone badania eksperymentalne zostały skomentowane. Praca kończy się krótką dyskusją i propozycjami przyszłych badań.*

Słowa kluczowe: *zakupy internetowe, optymalizacja,* e-commerce*, ISOP, algorytmy*

1. Rozwój środowiska handlu elektronicznego

Bardzo ważnym aspektem informatyki w ostatnich latach jest rozwój sieci Internet. Liczba witryn internetowych nieustannie wzrasta. Jedną z przyczyn jest bez wątpienia ciągły wzrost liczby osób mających dostęp do sieci. Przez ostatnią dekadę wzrost wartości był stały i gwałtowny. Według najnowszych danych statystycznych[1] 32,7% (2267 milionów) mieszkańców świata ma dostęp do Internetu. Największa penetracji usługi jest zauważalna w Ameryce Północnej i wynosi 78,6%. Kolejnymi regionami (w kolejności malejącej) są: Australia i Oceania (67,5%), Europa (61,3%), Ameryka Południowa (39,5%), Bliski Wschód (35,6%), Azja (26,2%) i Afryka (13,5%). Najważniejszy jest fakt dynamicznego

[1] *Internet World Stats*, Internet Usage Statistics 2011, www.internetworldstats.com/stats.htm [19.05.2012].



Rys. 1. Odsetek klientów sklepów online

Ź r ó d ł o: *E-commerce across Europe – progress and prospects 2008*, The Future Fundation, London 2008.

wzrostu wskaźnika penetracji. W ostatnich 11 latach (2000-2011) wzrost liczby osób uzyskujących dostęp do Internetu wyniósł 528,1%, z liczby 360 mln użytkowników w roku 2000 do 2267 mln osób pod koniec roku 2011.

Główny Urząd Statystyczny[2] informuje, że poziom penetracji dostępu do sieci Internet w Polsce wynosi 57% (7,1 mln osób). Inne źródła, na podstawie wyników przeprowadzonych badań ankietowych, donoszą o wartości na poziomie 50%. Różnice mogą wynikać z innej metodologii czy szczegółowości przeprowadzonych badań, ewentualnie ze sposobu formułowania pytań ankietowych.

Warto zauważyć, jak ogromną wartość przedstawia rynek *e-commerce*, który jest ściśle powiązany z liczbą użytkowników uzyskujących dostęp do Internetu, jak i witryn *webowych*. Badania przeprowadzone w Stanach Zjednoczonych[3] pokazują, że wartość rynku *e-commerce* wzrosła z 7,4 mld dolarów w połowie 2000 r. do 34,7 mld w 2007 r.

Jednym z ważnych elementów rynku gospodarki elektronicznej są zakupy internetowe. Według prowadzonych badań[4] do 2013 r. prawie połowa mieszkańców

[2] *Główny Urząd Statystyczny*, 2010, www.stat.gov.pl/gus [10.10.2011].

[3] *Pew Internet & American Life Project*, On-line Shopping 2008, www.pewinternet.org/Reports/2008/Online-Shopping.aspx [20.05.2012].

[4] *E-commerce across Europe – progress and prospects 2008*, The Future Foundation, London 2008.

Europy ma dokonywać zakupów w sklepach internetowych. W 2006 r. notowano 26% odsetek klientów mających dostęp do Internetu (rys. 1).

2. Sklepy internetowe i porównywarki cen

Korzystanie z oferty sklepów internetowych/aukcji internetowych ma wiele zalet w porównaniu do zakupu w tradycyjnych sklepach. Najważniejszymi mogą być niższe ceny[5] i możliwość zakupów z bardzo odległych miejsc[6]. Należy jednak zauważyć, że zakupy związane są z dodatkowym kosztem wysyłki. Trzeba również brać pod uwagę konieczność poświęcenia czasu na znalezienie interesującej nas oferty. Zdarza się również, że przez stronę internetową nie można jednoznacznie określić, czy dane dwa produkty są identyczne. Szukanie najlepszej oferty znacznie ułatwiają serwisy umożliwiające porównywanie cen, tzw. porównywarki cen.

Porównywarki cen to serwisy internetowe oferujące stworzenie list rankingowych ofert na dany produkt. Konkurencyjność oferentów ma powodować obniżenie cen produktu, podobnie jak można to zaobserwować w przypadku aukcji internetowych. Na prezentowanej liście rankingowej znajdą się oferty sklepów (tych, które mają podpisane umowy z właścicielami serwisu) mających w ofercie szukany produkt. Warto zauważyć, że witryny oferujące porównywanie cen są obecnie bardzo popularne. Według zestawienia prezentowanego przez portal Alexa Rank[7] najbardziej znane porównywarki cen znajdują się na liście 1000 najpopularniejszych (o największej liczbie dziennych odwiedzin) witryn na świecie: nextag.com – 472. miejsce, shopping.com – 582. miejsce, bizrate.com – 821. miejsce.

Czynnikiem motywującym autora do rozpoczęcia pracy było zauważenie ułomności funkcjonalnej serwisów porównujących ceny produktów. Ich największą wadą jest możliwość porównania ceny tylko jednego produktu w danym momencie. Jeśli mamy zamiar kupić wiele produktów (np. książek, płyt audio itd.), otrzymujemy dużo oddzielnych list rankingowych. Zakupienie wszystkich żądanych produktów w najniższej sumarycznej cenie okazuje się zadaniem bardzo trudnym.

Załóżmy, że mamy do kupienia pewną ilość produktów w dostępnych sklepach. Każdy ze sklepów oferuje dany produkt (dla łatwości obliczeń i zasadności działania algorytmu można przyjąć, że oferta na produkt niedostępny w danym

[5] R. Hof, *More ways to price-shop online*, „BusinessWeek" 2003, nr 14(3851).

[6] W. Chu, B. Choi, M.R. Song, *The role of on-line retailer brand and infomediary reputation in increasing consumer purchase intention*, „International Journal of Electronic Commerce" 2005, nr 9, s. 115-127.

[7] *Alexa Rank*, www.alexa.com [20.08.2011].

sklepie jest nieproporcjonalnie wysoka – cena wielokrotnie wyższa niż w innych sklepach). Z każdym ze sklepów związany jest koszt wysyłki. Koszt wysyłki jest stały dla każdego sklepu, niezależnie od ilości zamawianych produktów.

Zadaniem jest dokonanie możliwie najtańszego zakupu wszystkich produktów spośród ofert wszystkich sklepów.

3. Problem optymalizacji zakupów internetowych

Zagadnienie optymalizacji zakupów internetowych to nowy problem zdefiniowany niedawno przez autora i innych (*Internet Shopping Optimization Problem* – ISOP)[8]. Zostało udowodnione, że problem należy do klasy problemów NP-trudnych[9]. Zaprezentowano również algorytm zachłanny[10], który uzyskiwał bardzo dobrą jakość wyników (kontra rozwiązanie optymalne), działając w bardzo krótkim czasie przy niewielkiej złożoności obliczeniowej.

Notacja używana w niniejszej pracy opisana jest następująco:

$N = \{1, \ldots, n\}$ – zbiór sklepów internetowych,

$M = \{1, \ldots, m\}$ – zbiór produktów do zakupienia,

d_i – koszt wysyłki wszystkich produktów zakupionych w sklepie i,

p_{ij} – podstawowa cena produktu j w sklepie i. $p_{ij} = p_j$, jeżeli cena produktu j jest taka sama we wszystkich sklepach,

N_i – podzbiór produktów zbioru N w sklepie i, $N_i \subseteq N$,

M_j – podzbiór sklepów zbioru M, w których można kupić produkt j, $M_j \subseteq M$,

S_i – podzbiór produktów wybranych przez klienta (algorytm) do zakupienia w sklepie i,

$N = \bigcup_{i=1}^{m} S_i$ oraz $S_i \bigcap S_j = \varnothing$, $i \neq j$, dla możliwego rozwiązania,

$T_i(S_i) = d_i + \sum_{j \in S_i} p_{ij}$ – łączny koszt dostawy i zakupów dokonanych w sklepie i za produkty $S_i \subseteq N_i$, jeżeli nie ma żadnych niejasności, notację S_i można pominąć w zapisie $T_i(S_i)$, używając T_i.

Problem optymalizacji zakupów internetowych można zapisać w postaci matematycznej jako:

---

[8] J. Blazewicz, M.Y. Kovalyov, J. Musial, A.P. Urbanski, A. Wojciechowski, *Internet shopping optimization problem*, „Applied Mathematics and Computer Science" 2010, nr 20(2), s. 385-390.

[9] M.R. Garey, D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP--Completeness*, Freeman, New York 1979.

[10] J. Musiał, *Problem optymalizacji zakupu wielu produktów w sklepach internetowych. Propozycja algorytmu heurystycznego*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego" 2010, nr 597, s. 585-592.

$$\min \sum_{i=1}^{m} \left(d_i y_i + \sum_{j \in N_i} p_{ij} x_{ij} \right)$$

$$p.\,o. \sum_{i \in M_j} x_{ij} = 1, \quad j = 1, ..., n,$$

$$0 \le x_{ij} \le y_i, \quad i = 1, ..., m, \quad j = 1, ..., n,$$
$$x_{ij} \in \{0, 1\}, \quad y_i \in \{0, 1\}, \quad i = 1, ..., m, \quad j = 1, ..., n.$$

3.1. Stałe koszty wysyłki, brak rabatów

Weźmy pod uwagę najprostszą odmianę problemu ISOP. Cel optymalizacji nie ulega zmianie i zadaniem jest dokonanie zakupów wybranych produktów spośród dostępnej oferty sklepów. Istotna jest ogólna kwota, którą należy zapłacić za produkty, i poszczególne koszty wysyłki. Warto natomiast zauważyć, że koszty wysyłki są wartością stałą dla każdego ze sklepów i nie ulegają zmianie podczas przygotowywania listy zakupów (kwota zakupów czy ilość zakupionych przedmiotów w danym sklepie nie wpływa na zmianę wartości atrybutu). Dla danego sklepu nie są również naliczane żadne dodatkowe rabaty, które wpływałyby na zmianę cen w zależności od wielkości atrybutu rabat (ten by się zmieniał dynamicznie w stosunku do wartości zakupów dokonanych w jednym sklepie).

Tak zaprezentowaną wersję problemu ISOP można zredukować do znanego problemu *Facility Location Problem* (FLP). Głównymi charakterystykami FLP są przestrzeń, w której się poruszamy, miara, z góry ustalone pozycje klientów i z góry określone możliwe pozycje otwarcia fabryk. Tradycyjna wersja problemu FLP polega na otwarciu (wskazaniu miejsc) dowolnej liczby fabryk w dowolnym miejscu w dostępnej przestrzeni (wersja ciągła problemu) lub spośród lokalizacji wskazanych w definicji (wersja dyskretna). Następnie należy przydzielić wszystkich klientów do fabryk, w taki sposób, aby suma kosztów połączeń (przesyłu) między wszystkimi klientami i fabrykami, plus suma kosztów związanych z otwarciem zakładów, była najmniejsza.

Można zauważyć redukcję do problemu FLP, jeżeli produkty zapiszemy jako klientów, a sklepy jako fabryki. Ceny produktów opiszemy jako odległość między fabrykami a klientami (między sklepami a produktami). Natomiast koszt wysyłki oznaczymy jako koszt otwarcia fabryk. Problem FLP jest bardzo szeroko opisywany w literaturze[11].

[11] H. Eiselt, C.L. Sandblom, *Decision analysis, location models, and scheduling problems*, Springer-Verlag, Berlin-Heidelberg 2004; C. Iyigun, A. Ben-Israel, *A generalized Weiszfeld method for the multi-facility location problem*, „Operations Research Letters" 2010, nr 38(3), s. 207-214;

Algorytmów do rozwiązania tej wersji problemu ISOP można szukać w programowaniu liniowym, relaksacjach Lagrange'a, relaksacjach w programowaniu liniowym, algorytmach genetycznych, podejściu opartym na teorii grafów, a także w metodzie *Tabu Search*[12].

Niestety, ze względu na określoną wartość odległości od danej fabryki do klienta (odległość nie może się zmieniać podczas dokonywania optymalizacji/obliczeń) redukcji można dokonać tylko z bardzo specyficznej wersji (subproblem) problemu ISOP, w której nie są zawarte żadne dodatkowe atrybuty, opisane w kolejnych rozdziałach.

3.2. Stałe koszty wysyłki i rabaty

W tym punkcie opisany zostanie przypadek, gdy dodatkowe atrybuty powodują zmianę problemu na tyle, że niemożliwe staje się zredukowanie choćby do wspomnianego wcześniej FLP. Zakłada się, że (chwilowo, dla ułatwienia definicji) koszt wysyłki jest stały dla każdego sklepu i niezależnie od ilości zakupionych produktów czy kwoty przeznaczonej na zakupy nie ulega zmianie. Drugim atrybutem, który w istotny sposób zmienia charakterystykę problemu, jest zdefiniowanie funkcji określającej rabaty obowiązujące dla sklepów. System rabatowy jest powszechnie stosowaną metodą pozyskiwania klientów bądź utrzymania dotychczasowych. Powoduje, że atrakcyjność sklepu wzrasta. Często spotykaną praktyką jest określenie wysokości rabatu w stosunku do kwoty wydanej na zakupy, przykładowo: „Zrób zakupy na kwotę 100 zł, a uzyskasz 2% rabatu, jeśli wydasz 200 zł, uzyskasz 4% rabatu, jeśli Twoje zamówienie przekroczy wartość 500 zł, przyznamy Ci 5% rabatu". Oczywiście rzędy wielkości, przedziały i wielkości rabatów są ustalane odrębnie przez danego sprzedającego, nie panują w tej materii żadne z góry ustalone zasady.

Zapis i notację zaprezentowaną na początku pracy należy rozwinąć o:

$f_i(T)$ – funkcję określającą rabaty dla sklepów, w których zostały dokonane zakupy.

Problem optymalizacji zakupów internetowych, w których występują rabaty, możemy zatem opisać w postaci matematycznej jako:

J. Krarup, D. Pisinger, F. Plastria, *Discrete location problems with push-pull objectives*, „Discrete Applied Mathematics" 2002, nr 123(1-3), s. 363-378; M.T. Melo, S. Nickel, F.S. da Gama, *Facility location and supply chain management – a review*, „European Journal of Operational Research" 2009, nr 196(2), s. 401-412; C.S. Revelle, H.A. Eiselt, M.S. Daskin, *A bibliography for some fundamental problem categories in discrete location science*, „European Journal of Operational Research" 2008, nr 184(3), s. 817-848.

[12] C.S. Revelle, H.A. Eiselt, M.S. Daskin, op. cit.

$$\min \sum_{i=1}^{m} f_i \left(d_i y_i + \sum_{j \in N_i} p_{ij} x_{ij} \right)$$

$$p.\,o. \sum_{i \in M_j} x_{ij} = 1, \;\; j = 1, \ldots, n,$$

$$0 \le x_{ij} \le y_i, \quad i = 1, \ldots, m, \;\; j = 1, \ldots, n,$$
$$x_{ij} \in \{0, 1\}, \quad y_i \in \{0, 1\}, \quad i = 1, \ldots, m, \;\; j = 1, \ldots, n.$$

Funkcja określająca rabaty może być dowolnie skonstruowana i wyglądać np.:

$$f_i(d_i + P) = \begin{cases} d_i + P, & \text{jeżeli } 0 < P \le 50, \\ d_i + 50 + 0{,}95(P - 50), & \text{jeżeli } P > 50. \end{cases}$$

3.3. Zmienne koszty wysyłki

Opisywane wcześniej rozszerzenia problemu optymalizacji zakupów internetowych zawierały stały koszt wysyłki bez względu na inne atrybuty i ich wartości. Niniejszy rozdział zawiera krótki opis i charakterystykę problemu, w którym mamy do czynienia ze zmiennymi kosztami wysyłki. Taka definicja jest zainspirowana przez obserwacje działalności realnych sklepów i sprzedaży *online* na rynku *e-commerce*. Często można zauważyć, że koszt wysyłki jest uzależniony od kwoty, jaką chcemy przeznaczyć na zakupy w danym sklepie (z takiej formy korzystają np. największe sieci handlowe *online* w Polsce). Im wyższą kwotę wydamy na zakupy, tym mniejszy będzie koszt dostawy. Czasami przy osiągnięciu danego progu wysyłka jest darmowa. Z marketingowego punktu widzenia dla sprzedawców jest to doskonałe posunięcie. Marża na wielu produktach jest naprawdę wysoka i sprzedawca woli zaoferować klientowi darmową dostawę, uzyskując przy tym dużo większą sprzedaż. Jako klienci działamy naturalnie w taki sposób, że nie lubimy płacić za usługę dostawy, a za przedmioty, które otrzymamy. Wiele osób decyduje się zatem kupić większą niż planowana początkowo liczbę przedmiotów (czasami takich, których nie miały zamiaru kupić) za znacznie wyższą łączną kwotę, tak aby nie dopłacać za wysyłkę towarów.

Poniżej przedstawiono notację matematyczną problemu ze zmiennym kosztem wysyłki i bez funkcji określającej rabaty:

$P = \sum_{j \in N_i} p_{ij} x_{ij}$ – wartość zakupów dokonanych w sklepie i,

$d_i(P)$ – funkcja określająca koszty wysyłki z danego sklepu, obliczana na podstawie wartości zakupów dokonanych w tym sklepie.

Problem optymalizacji zakupów internetowych, w których występują rabaty, można zatem opisać w postaci matematycznej jako:

$$\min \sum_{i=1}^{m} \left(d_i \left(\sum_{j \in N_i} p_{ij} x_{ij} \right) y_i + \sum_{j \in N_i} p_{ij} x_{ij} \right)$$

$$p.\,o. \sum_{i \in M_j} x_{ij} = 1, \;\; j = 1, \ldots, n,$$

$$0 \le x_{ij} \le y_i, \quad i = 1, \ldots, m, \;\; j = 1, \ldots, n,$$

$$x_{ij} \in \{0, 1\}, \quad y_i \in \{0, 1\}, \quad i = 1, \ldots, m, \;\; j = 1, \ldots, n.$$

Funkcja określająca koszty wysyłki może być dowolnie skonstruowana i mieć np. postać:

$$d_i(P) = \begin{cases} 20, & \text{jeżeli} \quad P < 100, \\ 10, & \text{jeżeli} \quad 100 \le P < 200, \\ 0, & \text{jeżeli} \quad P \ge 200. \end{cases}$$

4. Algorytmy heurystyczne rozwiązujące problemy ISOP z dodatkowymi charakterystykami

W jednej z wcześniejszych prac autora[13] opisano wyniki eksperymentalnych testów przeprowadzonych przy użyciu autorskiego algorytmu heurystycznego do rozwiązania problemu ISOP ze stałym kosztem wysyłki i bez funkcji rabatowej. Wyniki badań przedstawia tabela 1. Dokonano „zakupu" 10 produktów spośród ofert kolejno: 10, 20, 30, 40, 50 sklepów. Ponadto przedstawiono pomiary dla różnych przedziałów kosztów wysyłki. Łącznie dokonano ponad 1200 pomiarów.

Każdy test (dla zadanych danych, jak liczba sklepów, koszty wysyłki) przeprowadzono 50-krotnie, a następnie dokonano uśrednienia wyników. Wielokrotne przeprowadzanie testów pozwoliło zapobiec sytuacji pojawienia się wartości odbiegających, granicznych. Wyniki działania algorytmu A2OZ porównano z wynikami otrzymanymi przy zastosowaniu algorytmów porównywarek cen (zebranie wszystkich otrzymanych list rankingowych i sumowanie kosztów pojedynczych, najtańszych produktów) (PCS) oraz z wynikami obliczonymi przez porównywarki cen biorące pod uwagę koszty wysyłki (PCS+) (tab. 1).

Godny odnotowania jest fakt, że w każdym przypadku algorytm heurystyczny zaproponował lepsze rozwiązanie. Sumaryczny koszt koszyka zakupów (wszystkie produkty) dla algorytmu heurystycznego jest niższy od 2 do prawie 35%!

[13] J. Musiał, op. cit., s. 585-592.

Tabela 1. Wyniki otrzymane podczas działania algorytmów PCS, PCS+, A2OZ

Liczba sklepów	Liczba produktów	Wysyłka	PCS	PCS+	A2OZ
10	10	15-30	408	360	349
20	10	15-30	422	346	329
30	10	15-30	424	347	326
40	10	15-30	429	355	319
50	10	15-30	440	360	328
10	10	20-25	400	372	343
20	10	20-25	426	392	347
30	10	20-25	439	406	343
40	10	20-25	434	406	339
50	10	20-25	443	412	345
10	10	10-20	347	326	320
20	10	10-20	366	325	311
30	10	10-20	380	328	313
40	10	10-20	369	324	301
50	10	10-20	373	330	308
10	10	15-15	357	357	327
20	10	15-15	365	365	320
30	10	15-15	381	381	334
40	10	15-15	370	370	319
50	10	15-15	375	375	318

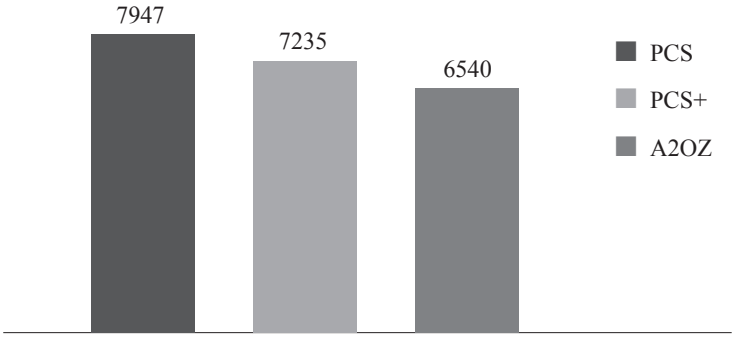
Ź r ó d ł o: opracowanie własne.

Analizując powyższe wyniki eksperymentów, można stwierdzić, że średnio algorytm A2OZ proponuje rozwiązanie o 21,8% tańsze niż rozwiązanie proponowane przez porównywarkę cen i o 11,4% tańsze niż rozwiązanie prezentowane przez porównywarkę cen analizującą koszty wysyłki.

Czas wykonania algorytmu heurystycznego nigdy nie przekroczył 2 sekund, działając *online* w środowisku rzeczywistym (aplikacja internetowa połączona z serwerem bazy danych).

Dla zobrazowania kosztów poniesionych na zakupy (sumując wszystkie wiersze tabeli) wyniki pracy algorytmów zaprezentowano na wykresie (rys. 1).

Do przeprowadzenia testów użyto algorytmu A2OZ i porównano wyniki z działaniem algorytmów PCS (porównywarka cen), PCS+ (ulepszony algorytm porównywarki cen, analizujący również koszty wysyłki) i wartością optymalna Opt. (tab. 2). Kolumny PCS, PCS+ i G opisują koszty zakupu produktów (dla każdej



Rys. 2. Koszty poniesione przy wykorzystaniu algorytmów PCS, PCS+, A2OZ

Źródło: opracowanie własne.

Tabela 2. Wyniki eksperymentu przeprowadzonego dla instancji problemu zawierających funkcję rabatową

Liczba sklepów	Liczba produktów	PCS	PCS+	G	Opt	G/PCS	G/PCS+	G/Opt
10	2	491,0	333,4	308,4	302,3	0,628	0,925	1,020
10	3	754,3	513,7	470,9	461,2	0,624	0,917	1,021
10	4	979,0	680,0	584,2	579,0	0,597	0,859	1,009
10	5	1304,3	905,2	676,6	663,8	0,519	0,747	1,019
15	2	561,9	378,9	346,6	334,0	0,617	0,915	1,038
15	3	865,0	487,3	430,1	421,3	0,497	0,883	1,021
15	4	987,6	675,0	556,2	541,3	0,563	0,824	1,028
15	5	1317,6	709,1	663,2	650,0	0,503	0,935	1,020
20	2	581,0	350,8	319,8	313,8	0,550	0,912	1,019
20	3	859,0	451,3	426,8	420,8	0,497	0,946	1,014
20	4	983,0	491,3	441,1	432,5	0,449	0,898	1,020
20	5	1325,4	765,8	630,9	615,7	0,476	0,824	1,025
25	2	583,0	305,2	274,9	271,3	0,472	0,901	1,013
25	3	781,9	487,0	390,3	382,3	0,499	0,801	1,021
25	4	1044,8	725,0	554,5	538,7	0,531	0,765	1,029
25	5	1142,8	765,2	675,3	652,6	0,591	0,883	1,035
30	2	579,0	302,4	278,7	269,4	0,481	0,922	1,035
30	3	818,5	512,5	472,3	456,0	0,577	0,921	1,036
30	4	949,4	590,8	484,6	472,5	0,510	0,820	1,026
30	5	1381,0	744,4	687,8	667,1	0,498	0,924	1,031

Źródło: opracowanie własne.

instancji problemu dokonano dziesięciokrotnego pomiaru). W trzech ostatnich kolumnach dokonano oceny jakości algorytmu A2OZ (oznaczonego w tabeli jako G) względem rozwiązań proponowanych przez porównywarki cen, jak i rozwiązania optymalnego. Model środowiska testowego zaczerpnięto z pracy C.S. Revelle i in.[14]

5. Podsumowanie

Problem optymalizacji zakupów internetowych (ISOP) jest nowym, bardzo interesującym zagadnieniem, nie tylko z matematycznego punktu widzenia, ale również istotnym w aspekcie ewentualnego wdrożenia w aplikacjach *online* działających na rynku *e-commerce*, zwłaszcza biorąc pod uwagę ciągły wzrost liczby osób dokonujących zakupów przez Internet.

Rozwijanie problemu o kolejne charakterystyki pozwala na jeszcze lepsze dopasowanie jego materii do realiów panujących na rynku handlu elektronicznego.

Kolejna praca z pewnością skupiona będzie na opracowaniu nowych, wydajniejszych (proponujących lepsze wyniki przy zachowaniu dużej szybkości obliczeń) algorytmów heurystycznych, a także na podjęciu próby konstrukcji efektywnego czasowo algorytmu dokładnego.

Literatura

Alexa Rank, http://www.alexa.com [20.08.2011].

Blazewicz J., Kovalyov M.Y., Musial J., Urbanski A.P., Wojciechowski A., *Internet shopping optimization problem*, „Applied Mathematics and Computer Science" 2010, nr 20(2), s. 385-390.

Chu W., Choi B., Song M.R., *The role of on-line retailer brand and infomediary reputation in increasing consumer purchase intention*, „International Journal of Electronic Commerce" 2005, nr 9, s. 115-127.

E-commerce across Europe – progress and prospects 2008, The Future Foundation, London 2008.

Eiselt H., Sandblom C.L., *Decision analysis, location models, and scheduling problems*, Springer-Verlag, Berlin – Heidelberg 2004.

Garey M.R., Johnson D.S., *Computers and Intractability: A Guide to the Theory of NP-Completeness*, Freeman, New York 1979.

Główny Urząd Statystyczny, 2010, www.stat.gov.pl/gus [10.10.2011].

Hof R., *More ways to price-shop online*, „BusinessWeek" 2003, nr 14(3851).

Internet World Stats, Internet Usage Statistics 2011, www.internetworldstats.com/stats.htm [19.05.2012].

Iyigun C., Ben-Israel A., *A generalized Weiszfeld method for the multi-facility location problem*, „Operations Research Letters" 2010, nr 38(3), s. 207-214.

Krarup J., Pisinger D., Plastria F., *Discrete location problems with push-pull objectives*, „Discrete Applied Mathematics" 2002, nr 123(1-3), s. 363-378.

[14] C.S. Revelle i in., op. cit.

Melo M.T., Nickel S., da Gama F.S., *Facility location and supply chain management – a review*, „European Journal of Operational Research" 2009, nr 196(2), s. 401-412.

Musiał J., *Problem optymalizacji zakupu wielu produktów w sklepach internetowych. Propozycja algorytmu heurystycznego*, „Zeszyty Naukowe Uniwersytetu Szczecińskiego" 2010, nr 597, s. 585-592.

Pew Internet & American Life Project, On-line Shopping 2008, www.pewinternet.org/Reports/2008/Online-Shopping.aspx [20.05.2012].

Revelle C.S., Eiselt H.A., Daskin M.S., *A bibliography for some fundamental problem categories in discrete location science*, „European Journal of Operational Research" 2008, nr 184(3), s. 817-848.

Zeszyty Naukowe
Wyższej Szkoły Bankowej w Poznaniu
Nr 40/2012

Bogdan Pilawski
Wyższa Szkoła Bankowa w Poznaniu
Bank Zachodni WBK

Narzędzia ETL w zasilaniu repozytoriów danych[1]

Streszczenie. *W artykule omówiono wybrane, podstawowe aspekty stosowania narzędzi ETL (Extract – Transform – Load) do zasilania repozytoriów danych. Na tle ewolucji tych repozytoriów przedstawiono charakterystykę narzędzi ETL oraz wskazano tendencje rozwojowe i formułowane wobec nich nowe wymogi. Całość uzupełniają praktyczne przykłady zastosowań tych narzędzi.*

Słowa kluczowe: *ETL, bazy danych, repozytoria danych, hurtownie danych*

1. Wprowadzenie

Stosowanie rozwiązań informatycznych w celu usprawnienia funkcjonowania przedsiębiorstw, instytucji i administracji w latach 70. i 80. XX w. ograniczało się w zasadzie do bieżącej obsługi. Obejmowało ono stosunkowo proste, ale masowe pod względem ilościowym czynności planowania oraz rejestracji i – wraz z rozwojem

[1] Opracowanie jest rozwinięciem i uzupełnieniem wystąpienia pod tym samym tytułem, jakie miało miejsce w ramach cyklu seminariów „Ku modelowi gospodarki opartej na wiedzy", organizowanego wspólnym staraniem Katedry Informatyki Stosowanej Wyższej Szkoły Bankowej w Poznaniu oraz działającego na tej uczelni Studenckiego Koła Informatyki Stosowanej. Celem wspomnianego wystąpienia było przede wszystkim przedstawienia stosunkowo mało znanych podstaw narzędzi ETL oraz ich miejsca i roli w szerszej dziedzinie hurtowni danych, utożsamianej często z obszarem zbierania danych i przekształcania ich w informacje w wyniku analizy. Przesądziło to o informacyjnym przede wszystkim zakresie tego wystąpienia i jego ograniczonym poziomie szczegółowości. Podobne cechy ma niniejsze opracowanie. W artykule uwzględniono m.in. doświadczenia autora z czynnego udziału w procesie wyboru narzędzi ETL na potrzeby dużej, międzynarodowej korporacji bankowej, zakończonego oceną i decyzją po kilkutygodniowych próbach, z udziałem ówczesnej światowej czołówki producentów takich narzędzi. Autor uczestniczył również w późniejszym opracowaniu strategii wdrożenia wybranych narzędzi.

technik i narzędzi – poszerzyło się z czasem o bieżącą obsługę transakcji. Silnym czynnikiem ograniczającym zakres tego stosowania były wówczas jego wysokie koszty, będące pochodną relatywnie wysokich cen sprzętu komputerowego[2]. Inną przyczyną tego stanu był niedorozwój narzędzi i metod gromadzenia i analizy danych.

Istotną zmianę przyniosło pojawienie się w drugiej połowie lat 70. XX w. tzw. minikomputerów oraz – na początku lat 80 – pierwszych komputerów osobistych. Masowość produkcji tych ostatnich, w połączeniu z rozwojem mikroelektroniki, spowodowała bardzo znaczny spadek cen sprzętu[3], czyniąc opłacalnymi wiele rozwiązań informatycznych, zupełnie nowych albo pozostających wówczas tylko w sferze koncepcji. Stan ten pozwolił m.in. na zwiększanie ilości danych pozostających w bezpośrednim dostępie (czyli – *de facto* – zapisanych w pamięci dyskowej), co pociągnęło też za sobą rozwój w zakresie oprogramowania dostęp ten obsługującego. W efekcie do dyspozycji pozostawały nie tylko bieżące dane transakcyjne, ale również tzw. dane historyczne, nie mające bezpośredniego związku z bieżącą działalnością, pozwalające jednak na ocenę i analizę działań przeszłych oraz – na jej podstawie – planowanie i wyznaczanie strategii na przyszłość.

2. Repozytoria danych i ich zasilanie

Początkowo wspomnianym działaniom analitycznym poddawano pliki i bazy danych wykorzystywane w zastosowaniach transakcyjnych, w czasie od nich wolnym. Typowym przykładem była praktyka jednej z brytyjskich sieci sprzedaży obuwia, która zaprezentowała swe rozwiązanie na początku lat 90., podczas dorocznej konferencji organizacji AMSU[4], odbywającej się na Uniwersytecie w Yorku. Polegało ono na wykonywaniu każdej doby analizy popytu, z podziałem na wzory, rozmiary oraz lokalizację, i kierowanie zaopatrzeniem sklepów według jej wyników. Analizę tę wykonywano na transakcyjnej bazie danych, po zakończeniu obsługi transakcyjnej i tzw. codziennego przetwarzania wsadowego[5].

[2] Cechy konstrukcyjne ówczesnego sprzętu komputerowego utrudniają dokładniejszy rachunek. Przykładowo: pojedynczy wymienny pakiet dyskowy o pojemności 30 MB, w systemie komputerowym zakupionym na początku lat 70. XX w. przez Zakłady H. Cegielski w Poznaniu, kosztował 210 GBP i wymagał dla działania napędu kosztującego ok. 10 000 GBP. Dla porównania – tani średniolitrażowy samochód kosztował wtedy w Wielkiej Brytanii (kraju producenta wspomnianego komputera) ok. 1200-1300 GBP (źródło: notatki autora).

[3] Mimo że spadek ten liczył się w rzędach wielkości, nie można go odnosić w równej mierze do sprzętu masowego (np. komputerów osobistych) i – będącego przedmiotem selekcji jakościowej – sprzętu stosowanego profesjonalnie.

[4] Association of Mainframe System Users – organizacja zrzeszająca użytkowników dużych komputerów produkcji brytyjskiej firmy ICL, przejętej później przez japońską firmę Fujitsu; w pracach i działaniach AMSU w latach 80. i 90. XX w. uczestniczyli liczni przedstawiciele użytkowników komputerów firmy ICL w Polsce.

[5] Źródło: notatki autora.

Praktyka taka była wówczas dość powszechna, miała jednak wiele wad. Jedną z nich była konieczność dysponowania odpowiednią ilością czasu, który każdej doby można było przeznaczyć na działania analityczne. Ich nieoczekiwane przedłużenie się stwarzało ryzyko opóźnienia w rozpoczęciu sesji obsługi transakcyjnej kolejnego dnia, co mogło oznaczać nawet brak możliwości obsługi bieżącej sprzedaży. Jej prowadzenie równolegle z działaniami analitycznymi było niemożliwe, gdyż te ostatnie angażowały niemal całość zasobów mocy obliczeniowej. Jeszcze większą przeszkodą okazała się wkrótce sama organizacja zapisów danych w plikach i bazach, dobrze uwzględniająca potrzeby przetwarzania transakcyjnego, ale niezbyt przydatna do złożonych, masowych działań analitycznych.

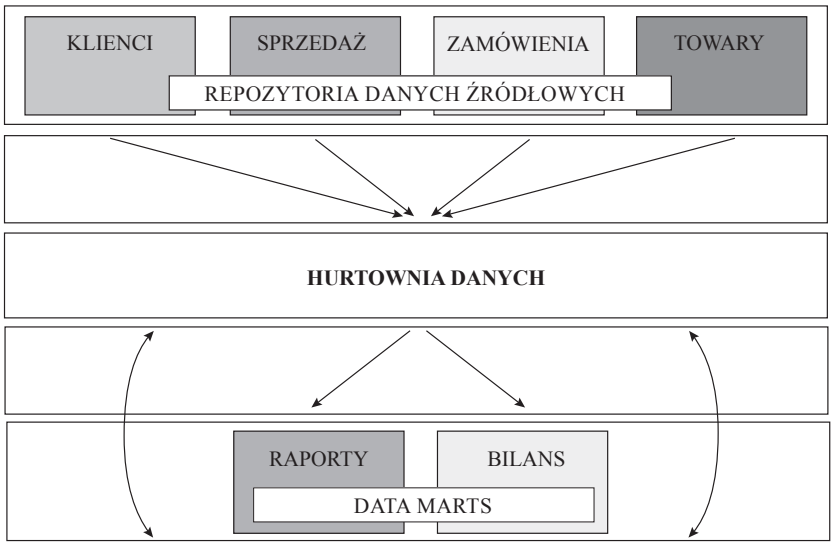
W efekcie, na przełomie lat 80. i 90. XX w., doprowadziło to do prób wyodrębniania repozytoriów danych przeznaczonych wyłącznie na potrzeby analityczne, na co wpływ miał również dalszy rozwój techniczny (zwiększanie pojemności pamięci dyskowych) i spadek cen sprzętu informatycznego. Repozytoria takie zaczęto określać mianem „hurtowni danych" (*data warehouse*), a za twórców ich podstaw uchodzą Bill Inmon i Ralph Kimball.

Hurtownie danych cechują się strukturą danych odmienną od znanej z transakcyjnych baz danych, gdzie przeważa „klasyczny" model relacyjny, chociaż – głównie w gałęziach przemysłu opartych na montażu – spotyka się również model hierarchiczny. Ta zasadnicza różnica w strukturze danych przesądza o fizycznej odrębności repozytoriów stanowiących hurtownie danych. Potrzebne do ich analizy znaczne moce obliczeniowe powodują też, że stosuje się do tego celu specjalizowane komputery, pracujące ze specjalistycznym oprogramowaniem.

W przypadku niektórych analiz wykonywanych na danych zgromadzonych w hurtowni zapotrzebowanie na moc obliczeniową jest tak duże, że praktycznie uniemożliwia równoległą realizację więcej niż jednej analizy. Przypadki takie powodują, że – przykładowo – analizy *ad hoc*, wykonywane w tym samym czasie, co zaplanowane, powtarzalne analizy rutynowe, utrudniają lub wręcz uniemożliwiają planowe wykonanie tych ostatnich. W celu zapobiegania takim sytuacjom stosuje się tzw. *data marts*[6], będące w istocie kopiami określonych podzbiorów hurtowni danych, przeznaczonymi do prowadzenia autonomicznych działań analitycznych o określonym zakresie.

Podzbiory takie, dla podkreślenia ich pochodzenia w całości od hurtowni danych, określa się mianem *dependent data marts*. Koncepcję tego rodzaju przedstawiono poglądowo na rys. 1.

[6] Termin *data mart* nie ma, jak dotąd, polskiego odpowiednika; samo słowo *mart* w języku angielskim oznacza *targowisko* i jest pewnego rodzaju pochodną, też zapożyczonego z handlu, terminu *hurtownia*, rozumianej jako jednostka umieszczona wyżej niż targowisko w hierarchii pośredników między producentem a konsumentem.



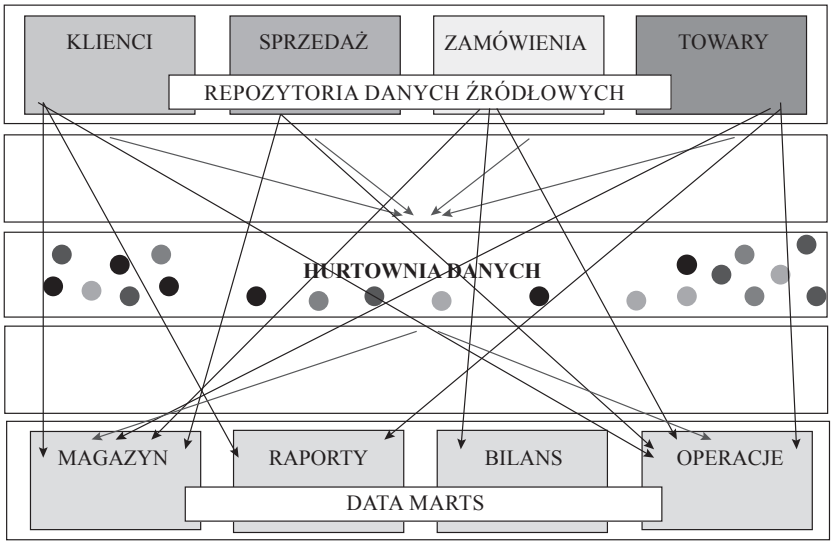
Rys. 1. Koncepcja *dependent data marts*

Ź r ó d ł o: opracowanie własne.

W licznych jednak dużych organizacjach, najczęściej z przyczyn związanych z ich złożoną przeszłością[7], ale również w wyniku sięgania po doraźne, uproszczone rozwiązania, wykształciła się praktyka korzystania z tzw. *independent data marts*. Polega ona na tworzeniu analitycznych *data marts* bezpośrednio z repozytoriów danych źródłowych, z pominięciem hurtowni danych bądź z jej tylko częściowym udziałem. Rozwiązania takie wydają się atrakcyjne dzięki szybkiemu przedstawianiu wyników, jednak – w dłuższej perspektywie – stwarzają ryzyko niespójności takich samych lub podobnych wyników, otrzymywanych z innych repozytoriów danych, również będących *independent data marts*, czy też – samej hurtowni danych. Wyniki takie mogą też być zaprzeczeniem sformułowanej przez B. Inmona zasady „jednej wersji prawdy" (*single version of truth*)[8]. Według tej zasady poleganie wyłącznie na danych pochodzących z ich hurtowni da zawsze takie same wyniki tych samych analiz. Pomijanie natomiast, czy nawet „obchodzenie" hurtowni grozi niespójnościami i rozbieżnościami w wynikach, które – przełożone na decyzje – mogą mieć negatywne skutki. Rozwiązanie z użyciem *independent data marts* przedstawia rysunek 2.

[7] Do takich przyczyn można zaliczyć łączenia się i podziały czy próby stosowania analitycznych rozwiązań informatycznych z okresu przed hurtowniami danych.

[8] Zob. www.b-eye-network.com/view/282.



Rys. 2. Koncepcja *independent data marts*

Źródło: opracowanie własne.

Skoro jednak hurtownia danych stanowi odrębną od pozostałych systemów informatycznych stosowanych w danej organizacji całość, pojawia się kwestia przenoszenia do niej danych z tych innych systemów. Nie stanowi to na ogół większego problemu tam, gdzie dane w hurtowni stanowią kopię danych z transakcyjnej bazy danych, a różnią się jedynie strukturą i organizacją. Przypadki tego rodzaju występują zazwyczaj tylko w niedużych organizacjach. Tam jednak, gdzie są liczne, zróżnicowane źródła danych, a ilość samych danych jest znaczna, terminowe i jakościowo poprawne zasilanie hurtowni danych nabiera szczególnego znaczenia i jest realizowane z udziałem specjalistycznego oprogramowania.

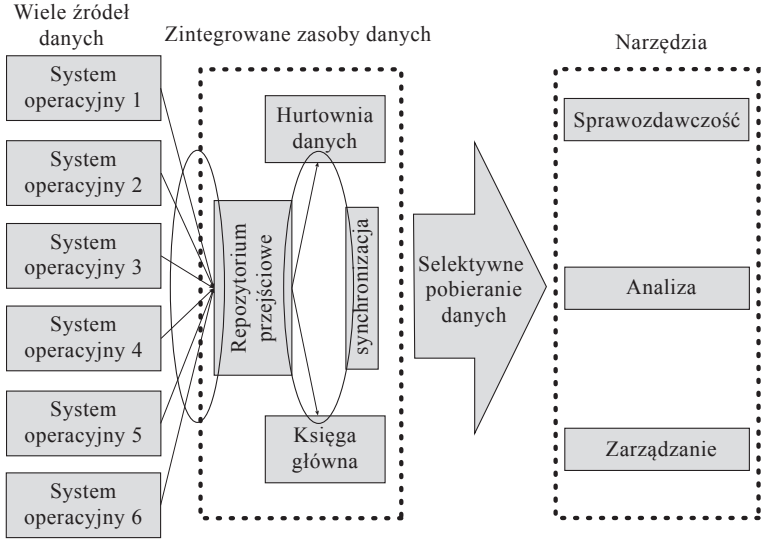
Wspomniane oprogramowanie jest określane symbolem ETL[9], który jest skrótem od angielskich terminów *Extract – Transform – Load*. Terminy te oddają istotę działania tego oprogramowania, które sprowadza się do trzech podstawowych czynności:

– pobrania danych z ich repozytoriów źródłowych (*Extract*),
– dokonania reorganizacji, przekształceń i agregacji danych (*Transform*),
– umieszczenia zreorganizowanych i przekształconych danych w hurtowni danych (*Load*).

[9] Krótki zarys historii narzędzi ETL w: Y. Montcheuil, C. Dupupet, *Third Generation ETL: Delivering the Best Performance*, Sunopsis Inc., Boston 2007.

3. Charakterystyka narzędzi ETL

Przykładowy, całościowy schemat stosowania hurtowni danych w banku przedstawia rys. 3. Po jego lewej stronie występują rozmaite repozytoria-źródła danych, pochodzących z systemów operacyjnych, zapewniających bieżącą obsługę działalności banku. Dane z tych repozytoriów są pobierane według określonych reguł (funkcja „Extract") i umieszczane w repozytorium przejściowym (*data stage*). W celu zagwarantowania, że dane repozytorium źródłowe odzwierciedla stan z określonego momentu (np. na koniec dnia w rozumieniu księgowym), na czas pobierania danych blokuje się możliwość aktualizacji zapisów w takim repozytorium. Blokada taka wyłącza dane repozytorium z działań bieżących, co, w niektórych warunkach, może ograniczać ich zdolność do działania[10]. Powoduje to dążenie do możliwie największego skrócenia operacji pobierania, czego wynikiem jest jej ograniczanie do samego tylko pobrania i przeniesienia danych. Tam, gdzie nie ma takiego ograniczenia, spotyka się rozwiązania, w których w trakcie przenoszenia danych dokonuje się również ich reorganizacji, przekształceń i agregacji. W przypadku wielu źródeł danych działania te mają charakter wstępny, czasem tylko kontrolny, a ich wyniki są podstawą do właściwych przekształceń, wykonywanych w samym już tylko repozytorium przejściowym (por. rys. 3).



Rys. 3. Schemat stosowania hurtowni danych w banku

Ź r ó d ł o: opracowanie własne.

[10] Ograniczenie takie wystąpi wtedy we wszystkich systemach działających w trybie określanym jako 7x24 (24 godziny na dobę, przez wszystkie dni tygodnia), czyli bez żadnych przerw.

Dwa etapy, w których najczęściej znajdują zastosowanie narzędzia ETL, oznaczono na rys. 3 owalami, umieszczonymi po obu stronach repozytorium przejściowego. Spotyka się jednak też rozwiązania, gdzie narzędzie te stosuje się po „drugiej" niejako stronie hurtowni danych, czyli do tworzenia z niej, wspomnianych już wcześniej, *data marts*. Nie jest to jednak zamierzony, główny cel powstania i istnienia tych narzędzi.

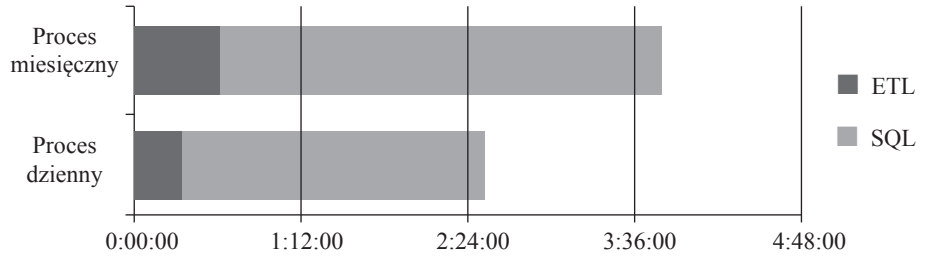
Portal internetowy o nazwie ETL Tools[11], zajmujący się różnymi aspektami praktycznymi związanymi z tymi narzędziami, pośród pożądanych cech narzędzi ETL wymienia następujące:

– obsługa podziału dużych tabel danych,
– obsługa bardzo dużych ilości danych,
– wykonywanie kontroli poprawności danych,
– graficzne odwzorowanie związków między repozytoriami i elementami danych,
– działanie z wieloma wersjami systemów operacyjnych i oprogramowania baz danych,
– obsługa metadanych,
– obsługa wykrywania błędów,
– obsługa gwiaździstych schematów organizacji danych,
– działanie wielowątkowe,
– kontrola wersji,
– harmonogramowanie zadań,
– graficzny interfejs użytkownika,
– interfejs przeglądarki internetowej.

Powyższą listę można uznać za dość wyczerpującą i opartą na szerokim doświadczeniu praktycznym, brak w niej jednak jednego szczególnie istotnego kryterium, jakim są tzw. interfejsy własne (*native interfaces*). Istotą tych interfejsów jest ich przygotowanie do współdziałania ze źródłowymi repozytoriami danych oparte na znajomości i wykorzystaniu ich wewnętrznych mechanizmów, zamiast sięgania po rozwiązania znormalizowane, typu język SQL. Żądania wykonania operacji na bazie danych, sformułowane w tym języku, przed właściwym wykonaniem każdorazowo wymagają konwersji i sprowadzenia do poziomu wspomnianych interfejsów własnych. Czynności te wydłużają znacznie czas trwania operacji pobierania danych, wydłużając w ten sposób okres niedostępności danego repozytorium dla innych działań. Przykład różnic między czasem trwania operacji na tych samych danych, raz wykonanych metodami tradycyjnymi (SQL), drugi raz – za pomocą narzędzi ETL z udziałem interfejsu własnego, przedstawia rys. 4[12].

[11] Zob. ww.etltools.com.

[12] Przykład z 2006 r. z jednego z polskich banków.



Rys. 4. Czas pobierania tych samych danych metodami SQL i ETL

Ź r ó d ł o: opracowanie własne.

Z danych przedstawionych na rys. 4 wynika, że różnica w czasie przetwarzania jest ponad pięciokrotna na korzyść interfejsów własnych. Po stronie wad tych interfejsów należy jednak wskazać to, że nie zawsze nadążają one za zmianami wprowadzanymi do obsługujących je mechanizmów przez producentów oprogramowania baz danych[13].

Powyższe nie oznacza jednak, że od narzędzi ETL nie oczekuje się „klasycznych" metod operowania danymi. Metody te obejmują nie tylko wspomniany już tu język SQL, ale również metody ODBC/JDBC, pliki jednowymiarowe (*flat files*) oraz usługi ESB[14].

Inną ważną właściwością narzędzi ETL jest ich zdolność do obsługi metadanych – zarówno po stronie systemów źródłowych, z których dane mają być pobierane, jak i po stronie hurtowni danych, gdzie narzędzia te potrafią działać z różnymi tzw. modelami danych. Modele takie to gotowe, wzorcowe struktury danych, przygotowane i udostępniane przez producentów hurtowni danych. Występują one w wielu wersjach, przeznaczonych dla różnych branż, z pewnym zakresem możliwości własnego kształtowania takiego modelu przez użytkownika.

Przygotowanie niektórych dostępnych na rynku narzędzi ETL do obsługi metadanych stanowi znaczne ułatwienie w korzystaniu z tych narzędzi i przyczynia się do ujednolicenia kategorii terminologicznych z zakresu nazewnictwa elementów danych. Ujednolicenie takie ułatwia istotnie porozumiewanie się służb biznesowych i informatycznych działających w danej organizacji. Pełne jednak ujednolicenie w zakresie metadanych jest ciągle bardzo odległą perspektywą, gdyż w praktyce

[13] Producenci ci nie zawsze informują wytwórców narzędzi ETL o wprowadzanych przez siebie zmianach i usprawnieniach, gdyż często sami dostarczają również takie narzędzia i nie leży w ich interesie usprawnianie działania narzędzi konkurentów.

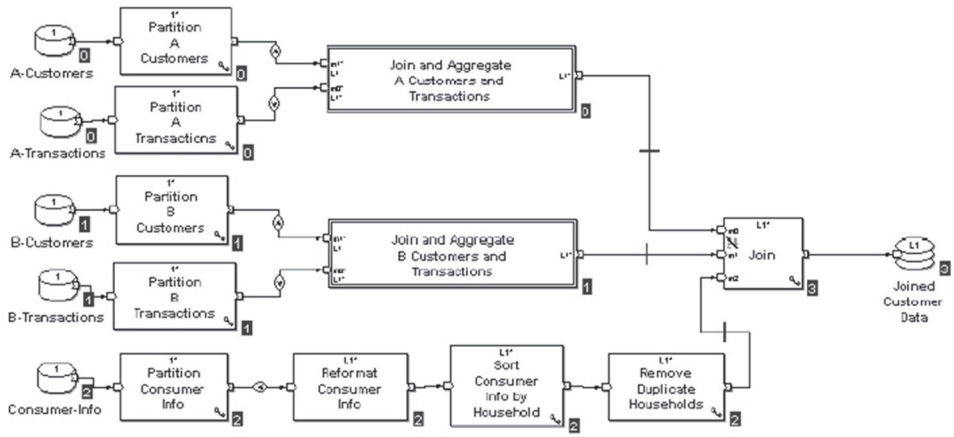
[14] Metody te wymieniane są m.in. w: P. Russom, *How To Evaluate Enterprise ETL*, Forrester Research, Cambridge (US), 2004, s. 8 oraz Y. Montcheuil, C. Dupupet, op. cit., s. 9 (rozdział „Data Access Technologies"); jako wymóg minimum konieczność ich obsługi przez narzędzia ETL wymieniają też autorzy opracowania: W. Eckerson, C. White, *Evaluating ETL and Data Integration Platforms*, The Data Warehousing Institute, Seattle 2003.

omawianej tu dziedziny występują obecnie trzy odmienne główne kategorie metadanych, a mianowicie metadane: biznesowe, operacyjne i techniczne. Pełne ich ujednolicenie dla wszystkich etapów – poczynając od pobierania danych z repozytoriów źródłowych, a kończąc na działaniach analitycznych – długo jeszcze nie będzie możliwe.

Z korzystaniem z metadanych wiąże się inna cecha, występująca pośród przytoczonych tu wcześniej właściwości, jakimi winny cechować się narzędzia ETL. Chodzi o tzw. interfejs graficzny, pozwalający projektować i wyznaczać związki między źródłowymi repozytoriami danych a zasobami hurtowni, w których dane te mają się znaleźć. Projektowanie to i wyznaczanie obejmuje również wskazywanie, posługując się metadanymi, jakie dane podlegają przenoszeniu i jakie transformacje, agregacje i czynności kontrolne mają przy tej okazji być na nich wykonane[15].

Przykład ekranowego interfejsu graficznego narzędzia ETL przedstawia rys. 5. Łączy się tam, przekształca i agreguje w jedną strukturę dane pochodzące z pięciu odrębnych repozytoriów źródłowych[16]. Patrząc od góry – dane z dwóch par repozytoriów są tam łączone wstępnie, podczas gdy dane z piątego repozytorium są poddawane przekształcaniu i reorganizacji, po czym dane z wszystkich pokazanych tam źródeł są łączone i umieszczane w jeszcze innym repozytorium.

Przykład stosowania reguł transformacji i kontroli danych ukazuje rys. 6. Odzwierciedla on jednocześnie przebieg testowy, jaki można wykonać w celach kontrolnych przy projektowaniu reguł pobierania i transformacji danych.

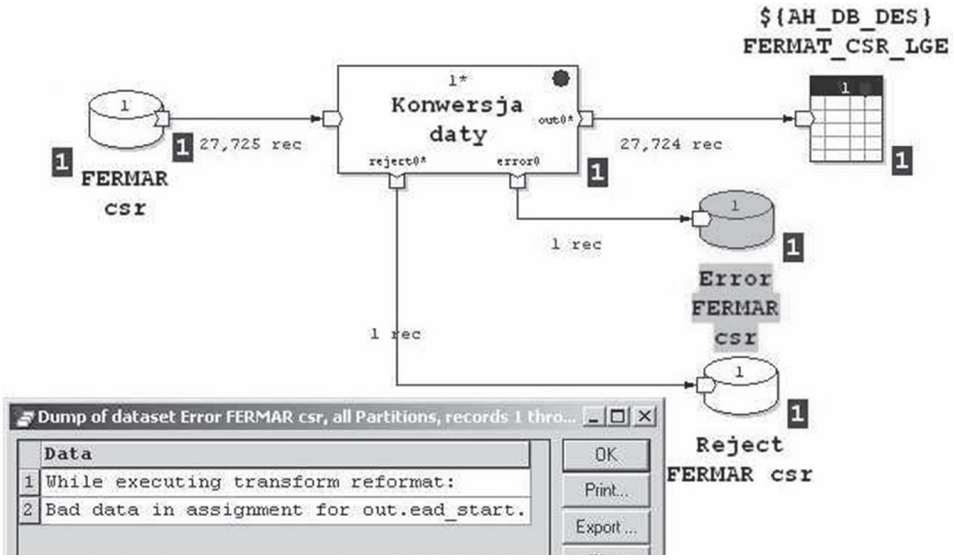


Rys. 5. Narzędzia ETL – graficzny interfejs użytkownika

Ź r ó d ł o: opracowanie własne (z praktyki).

[15] Brak interfejsu graficznego oznaczałby konieczność formułowania zadań w jakimś przeznaczonym do tego języku, co byłoby kłopotliwe w stosowaniu, bardzo pracochłonne i mało elastyczne.

[16] Wszystkie przytoczone tu przykłady pochodzą z systemu ETL o nazwie Co>Operation firmy Ab Initio.



Rys. 6. Narzędzia ETL – przebieg testowy

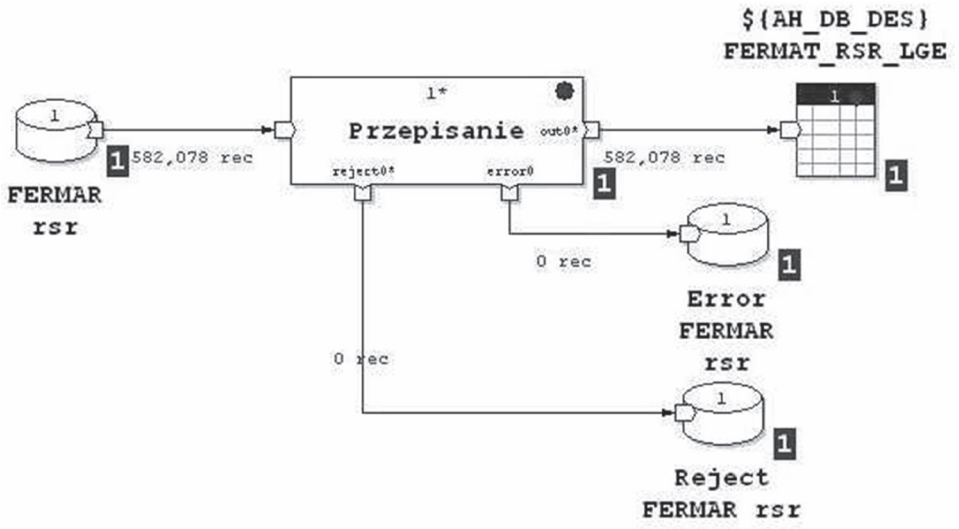
Ź r ó d ł o: opracowanie własne (z praktyki).

Ze schematu na rys. 6 widać, że z repozytorium o nazwie *FERMAR csr* pobrano 27 725 zapisów danych, z których jeden został zakwestionowany podczas wykonywania procedury o nazwie „Konwersja daty" i skierowany jednocześnie do dwóch repozytoriów pomocniczych: gromadzącego zapisy błędne (*Error FERMAR csr*) oraz zawierającego zapisy odrzucone w trakcie transformacji (*Reject FERMAR csr*).

Jeszcze inna istotna właściwość narzędzi ETL to realizacja funkcji kontrolnych z zakresu kompletności procesów przekształcania i przenoszenia danych. Hurtownia danych jest podstawą do sporządzania raportów, analiz oraz przygotowywania danych statystycznych. Wykonywanie wielu z tych czynności jest realizacją ustawowych obowiązków, a przekazanie – w ich ramach – błędnych danych może być nawet ścigane sądownie. Podejmowanie decyzji na podstawie błędnych czy chociażby tylko nieprecyzyjnych danych nie leży też w interesie organizacji, w której ma ono miejsce. Wszystkie przytoczone tu względy przemawiają więc za tym, aby przenoszone i przekształcane dane kontrolować również ilościowo, po to by mieć pewność, że uwzględniono wszystkie dane, które powinny być wzięte pod uwagę[17]. Przykład takiej kontroli przedstawia rys. 7. Przy symbolach repozytoriów: źródłowym (*FERMAR rsr*) i docelowym (*FERMAT_RSR_LGE*) widać liczniki przeniesionych zapisów (w obu przypadkach o stanie 582 078). Jedno-

[17] Ten aspekt praktyki stosowania narzędzi ETL poruszany jest m.in. w: P. Russom, op. cit.

cześnie ten sam schemat pokazuje, że żadne zapisy nie zostały w trakcie tego procesu zakwestionowane (zero zapisów w repozytoriach *Error FERMAR rsr* oraz *Reject FERMAR rsr*).



Rys. 7. Narzędzia ETL – kontrola ilościowa

Źródło: opracowanie własne (z praktyki).

Wymogi praktyczne, jakim winny odpowiadać narzędzia ETL zastosowane w konkretnej organizacji, mogą się różnić w szczegółach, ale można też odnieść do nich kilka podstawowych zasad, istotnych w większości sytuacji[18].

4. Narzędzia ETL a inne metody zasilania

Narzędzia ETL należą do najczęściej stosowanych w zasilaniu hurtowni danych w dane, ale nie są jedynym środkiem do tego celu. Ich wadą jest np. konieczność stosowania repozytorium pośredniego, co nie tylko jest źródłem dodatkowych kosztów, ale stanowi również istotny czynnik wydłużający cały proces zasilania. Narzędzia ETL nie radzą sobie też dobrze z przypadkami, w których potrzebne jest zasilanie ciągłe (systemy działające w trybie *on-line* i *quasi on-line*).

[18] Krótkie omówienie tych zasad w: *Managing Big Data: Building the Foundation for a Scalable ETL Environment*, Knightsbridge Solutions, Chicago 2002.

Rozwiązaniem mającym eliminować niektóre z wymienionych wad są narzędzia określane skrótem EL-T[19], których istota działania zakłada odwróconą, w stosunku do klasycznych narzędzi ETL, kolejność operacji: dane pobrane z repozytoriów źródłowych (funkcja *E-xtract*) najpierw są umieszczane w hurtowni danych (funkcja *L-oad*) i dopiero tam poddawane przekształceniu (funkcja *T-ransform*).

Inny postulat formułowany wobec narzędzi ETL dotyczy wydajności ich działania, gdzie stałym problemem są bardzo duże (i rosnące) ilości przenoszonych danych. Dla przestrzegania swoistej „czystości reguł" i w celu sprawowania właściwej kontroli nad procesami umieszczania danych w hurtowniach danych liczne dane są poddawane związanym z tym operacjom wielokrotnie. Działania takie pochłaniają zasoby infrastruktury informatycznej i pociągają za sobą koszty. Stan ten wpłynął na uzupełnienie narzędzi ETL o kolejną właściwość, określaną mianem *change data capture*[20]. Działanie w tym trybie polega na bieżącej, dokonywanej bezpośrednio w hurtowni danych aktualizacji tylko tych danych, które uległy zmianie. Jeszcze inne oczekiwania wobec narzędzi i metod ETL wiążą się z konceptem tzw. *Big Data* i – wiązanej m.in. z nim – metodyki przetwarzania *Hadoop*[21].

5. Podsumowanie

Mimo stosunkowo krótkiej historii hurtownie danych i metody ich zasilania danymi przeszły już długą ewolucję. Trwa ona nadal, gdyż wobec rozwiązań tych formułuje się coraz to nowe wymogi. Wiele z tych wymogów można spełnić również w wyniku rozwoju technicznego, powodującego, że konkretne rozwiązania, często od dawna przygotowane teoretycznie, znacznie później znajdują swe uzasadnienie ekonomiczne.

Liczne pierwsze zastosowania hurtowni danych zakładały, że wystarczy zebrać odpowiednio dużą ilość możliwie najbardziej szczegółowych danych, by w wyniku odnalezienia ukrytych w nich, często głęboko, prawidłowości uzyskać

[19] Przykład takiego rozwiązania i jego zalety przedstawiono w: *Is ETL Becoming Obsolete? Why a Business-Rules-Driven "E-LT" Architecture is Better*, Sunopsis Inc., Boston 2006.

[20] Metodykę tę bliżej omówiono m.in. w: *Augmenting ETL Systems With Real-Time Change Data Capture*, GoldenGate Software Inc., San Francisco 2007; tamże w związku z tym poddaje się również krytyce samą koncepcję „dnia operacyjnego" jako nieprzystającą do współczesnych potrzeb.

[21] Szeroki przegląd obecnego stanu wymogów wobec narzędzi ETL można znaleźć w: N. Yuhanna, *The Forrester Wave™: Enterprise ETL, Q1, 2012*, Forrester Research, Cambridge (US), 2012, tam też dokonano przeglądu i porównania bieżących możliwości tych narzędzi i zawarto wielokryteriową ocenę przodujących rozwiązań z tego zakresu.

zaskakujący efekt biznesowy, dający skokowy wzrost przewagi nad konkurentami. Przypominało to – w jakimś sensie – działania średniowiecznych alchemików, poszukujących dobrze ukrytego przez naturę sposobu na przemianę żelaza w złoto, o istnieniu którego byli przekonani.

Obecna praktyka hurtowni danych, jak każda inna dziedzina informatyki, rządzi się jednak realnymi i twardymi regułami, co nie oznacza, że nie może, w niektórych przypadkach, być źródłem spektakularnych sukcesów. Codzienność tej dziedziny to jednak żmudne, powtarzalne działania, których realizacja wymaga stałej dbałości o jakość danych i precyzję wyników. Istotną w tym rolę pełnią narzędzia ETL, które – podobnie jak cała dziedzina gromadzenia i analizy danych – podlegają ciągłej ewolucji i doskonaleniu.

Literatura[22]

Augmenting ETL Systems With Real-Time Change Data Capture, GoldenGate Software Inc., San Francisco 2007.

Data Integration: Moving Beyond ETL, DataFlux Corporation, Cary 2010.

Eckerson W., White C., *Evaluating ETL and Data Integration Platforms*, The Data Warehousing Institute, Seattle 2003.

Is ETL Becoming Obsolete? Why a Business-Rules-Driven "E-LT" Architecture is Better, Sunopsis Inc., Boston 2006.

Managing Big Data: Building the Foundation for a Scalable ETL Environment, Knightsbridge Solutions, Chicago 2002.

Montcheuil Y., Dupupet C., *Third Generation ETL: Delivering the Best Performance*, Sunopsis Inc., Boston 2007.

Russom P., *How To Evaluate Enterprise ETL*, Forrester Research, Cambridge (US), 2004.

Yuhanna N., *The Forrester Wave™: Enterprise ETL, Q1, 2012*, Forrester Research, Cambridge (US), 2012.

[22] Wykaz ten należałoby uzupełnić o liczne podręczniki systemu o nazwie Co>Operation firmy Ab Initio, czołowego producenta narzędzi ETL, które, mimo że nie są przytaczane w opracowaniu bezpośrednio, były wykorzystywane przy jego powstawaniu.

Zeszyty Naukowe
Wyższej Szkoły Bankowej w Poznaniu
Nr 40/2012

Maciej Skała, Iga Stróżyk
PBSG Sp. z o.o. w Poznaniu

Zarządzanie procesami i ryzykiem w organizacji z wykorzystaniem systemów informatycznych

***Streszczenie**. Celem artykułu jest przedstawienie możliwości wsparcia organizacji w procesie wdrażania nowych koncepcji zarządzania, takich jak zarządzanie procesami czy zarządzanie ryzykiem, poprzez zastosowanie odpowiednich systemów informatycznych. Wraz ze wzrostem popularności systemów zarządzania coraz więcej organizacji zdecydowało się na przejście z tradycyjnego zarządzania na zarządzanie procesowe. Artykuł opisuje możliwości wsparcia informatycznego organizacji w zakresie wdrożenia takiego podejścia do zarządzania organizacją. Podobnie wygląda kwestia zarządzania ryzykiem, które staje się nieodłącznym elementem funkcjonowania przedsiębiorstw. Obecnie dąży się do integracji zarządzania ryzykiem i zarządzania procesami, tak aby obejmowały one wszystkie aspekty organizacji (od strategii, poprzez cele, procesy i zadania). W celu wsparcia procesu wdrażania zintegrowanego zarządzania ryzykiem powstały również narzędzia informatyczne, których możliwości zastosowania zostały zaprezentowane w niniejszym artykule.*

***Słowa kluczowe:** zarządzanie ryzykiem, zarządzanie procesami, zintegrowane zarządzanie*

1. Wprowadzenie

Efektywne zarządzanie to obecnie nadrzędny cel funkcjonowania każdej organizacji. Coraz większą wagę przykłada się do wdrożenia właściwego i dedykowanego systemu zarządzania. Dzięki utrzymaniu takiego systemu organizacja ma pewność, że spełnia szereg kluczowych wymagań krajowych i międzynarodowych, które potwierdzają jej profesjonalizm i rzetelność[1]. Odpowiednio dobrane i skutecznie wdrożone systemy zarządzania pozwalają przekładać ogólną wizję kierownictwa organizacji na

[1] J. Brilman, *Nowoczesne koncepcje i metody zarządzania*, PWE, Warszawa 2002; *Rola znormalizowanych systemów zarządzania w zrównoważonym rozwoju*, red. J. Łańcucki, Wyd. Uniwersytetu Ekonomicznego, Poznań 2011.

działania operacyjne i indywidualne cele. Ich głównym zadaniem jest poprawienie jakości produktów i usług poprzez skuteczne zarządzanie procesami wewnętrznymi i zewnętrznymi. Dodatkowo pozwalają na wzrost produktywności, usprawniają realizowanie procesów oraz ograniczają koszty, wynikające z niesprawnego zarządzania organizacją. Skuteczność wdrożenia poszczególnych systemów zarządzania determinują takie aspekty, jak: adekwatność celów strategicznych i operacyjnych, zdolność do monitorowania i ciągłej weryfikacji realizowanych procesów, udział kompetentnej kadry oraz zaangażowanie najwyższego kierownictwa w ich utrzymanie.

Celem artykułu jest przedstawienie możliwości wsparcia organizacji w procesie wdrażania nowych koncepcji zarządzania, takich jak zarządzanie procesami czy zarządzanie ryzykiem, poprzez zastosowanie odpowiednich systemów informatycznych. W publikacji zostały opisane funkcjonalności takich systemów oraz korzyści, jakie wynikają z ich wdrożenia. W celu przedstawienia możliwości systemów informatycznych przeprowadzona została analiza funkcjonalności dostępnych na rynku narzędzi, ze szczególnym uwzględnieniem oprogramowania Smart oraz e-risk. Artykuł może być wykorzystany w procesie określania potrzeb organizacji w zakresie wdrożenia narzędzi informatycznych do zarządzania ryzykiem lub zarządzania procesami, jako źródło informacji na temat możliwych funkcjonalności systemów oraz usprawnień wynikających z ich implementacji[2].

2. Zarządzanie ryzykiem w organizacji

Praktyka dowodzi, że w większości przypadków organizacje przeprowadziły już działania w zakresie standaryzacji zadań, usprawnienia komunikacji czy poprawy organizacji pracy poprzez wdrożenie jednego z systemów zarządzania, np. najbardziej popularnego systemu zarządzania jakością ISO 9001. Obecnie przedsiębiorstwa podejmują działania polegające na wdrażaniu narzędzi w celu podnoszenia produktywności, usprawniania już zdefiniowanych procesów oraz ograniczania ich kosztów. Zatem zarządzanie procesami oraz ich optymalizacja (usprawnienie) stanowi kolejny poziom rozwoju organizacji. Etap ten jest ukierunkowany na realne obniżanie kosztów, poprawę efektywności i wdrożenie sprawnych mechanizmów zarządzania zmianami, mającymi zapewnić stały wzrost przyśpieszania realizacji poszczególnych zadań, a przez to procesów. Tak rozumiane zintegrowane zarządzanie procesami organizacji obejmuje przede wszystkim: reidentyfikację i usprawnienie procesów wewnętrznych, a także identyfikację i usprawnienie procesów zewnętrznych[3].

[2] *Zintegrowany System Zarządzania Ryzykiem – COSO II. Struktura ramowa*, PIKW i PIB, Warszawa 2007; Oprogramowanie Smart, http://oprogramowanie-smart.pl/smart [14.05.2012].

[3] *Zarządzanie procesami*, www.pbsg.pl/zarzadzanie-procesami.html [14.05.2012].

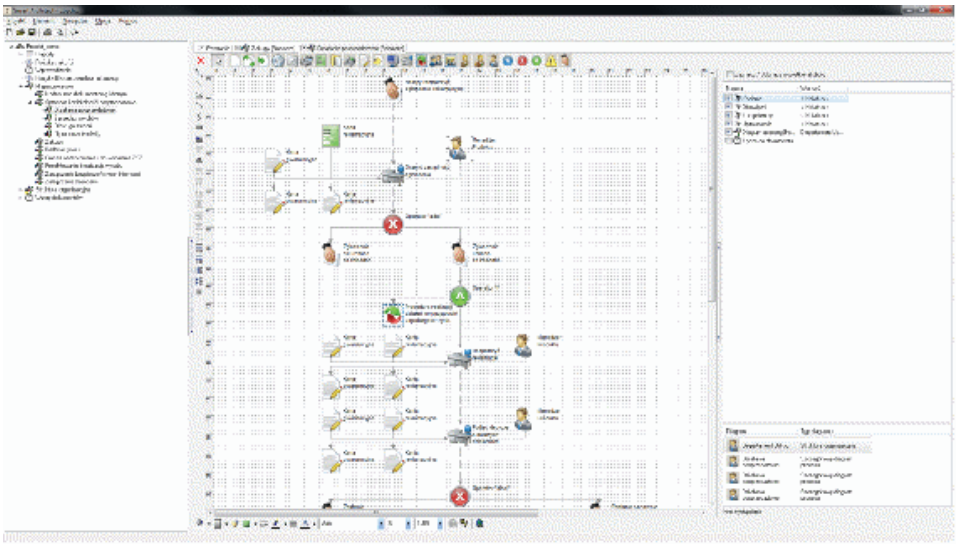
Zarządzanie procesami, w szczególności w ramach zintegrowanych systemów zarządzania, na podstawie międzynarodowych standardów, norm i modeli, np. ISO, ITIL, SOX, pozwala na określenie zadań w ramach realizowanych procesów oraz opisanie ich pod kątem konkretnych wymagań, celów i wskaźników. W związku z tym osoby zarządzające poszczególnymi etapami procesów mają możliwość rzetelnej oceny skuteczności i efektywności zarówno całego procesu, jak i jego części. Takie podejście przekłada się wprost na możliwość dynamicznego oddziaływania na procesy i bieżącego wdrażania działań korekcyjnych czy zapobiegających możliwościom wystąpienia niezgodności. Zintegrowane zarządzanie procesami pozwala ponadto na prowadzenie szczegółowych analiz związanych z przepływem informacji, zakresów odpowiedzialności i możliwości optymalizacji realizacji poszczególnych czynności w ramach procesów.

Potrzeba optymalizacji funkcjonowania organizacji i zachodzących w niej procesów wywołuje potrzebę szukania narzędzi i środków wspierających tego typu działania. Równocześnie faktem jest, że wraz z nieuniknioną modernizacją i wzrostem poziomu edukacji informatycznej powszechne staje się wdrażanie w przedsiębiorstwach profesjonalnego oprogramowania. W wyniku tego narzędzia informatyczne wspierające zarządzanie i utrzymanie procesów, umożliwiające integrację wielu systemów funkcjonujących w organizacji, stają się coraz bardziej powszechne w polskich przedsiębiorstwach i urzędach.

Przykładem narzędzia informatycznego wspierającego zarządzanie procesami jest oprogramowanie Smart. Jest to łatwe w użyciu i elastyczne narzędzie, zawierające szereg modułów upraszczających i przyspieszających codzienną pracę, monitorowanie efektywności oraz komunikację między pracownikami. Wysoki poziom skalowalności informacji pozwala na pomyślne wdrożenie systemu w każdej organizacji, dlatego Smart może być z powodzeniem stosowany w dużych, średnich i małych przedsiębiorstwach. Posiada on szereg funkcjonalności pozwalających na sprawne i efektywne zarządzanie procesami organizacji i jest podzielony na trzy zasadnicze moduły[4]:

– **Smart Architect** pozwala na tworzenie diagramów procesów (rys. 1). Jednym z podstawowych elementów modułu jest możliwość graficznego wizualizowania diagramów przy użyciu dostępnych ikon graficznych. Za pomocą tego modułu możliwe jest również tworzenie biblioteki dokumentów. Poszczególne formularze i szablony mogą być połączone z zadaniami w procesach i dostępne dla wszystkich użytkowników poprzez Smart Portal. Smart Architect daje administratorowi możliwość zarządzania użytkownikami programu – ich dostępem do całości lub jedynie wybranych elementów/poziomów programu. System wersjonowania pozwala użytkownikowi na efektywne zarządzanie wersjami elementów

[4] *Wdrożenie zarządzania ryzykiem*, www.pbsg.pl/wdrozenie-zarzadzania-ryzykiem.html [14.05.2012].



Rys. 1. Smart Architect

Ź r ó d ł o: materiały własne PBSG.

projektu. Po potwierdzeniu zmiany wskazane elementy są aktualizowane. Program automatycznie rejestruje zmiany merytoryczne i pozwala na ich późniejsze analizowanie.

– **Smart Audyt** to moduł do zarządzania audytami wewnętrznymi organizacji. Daje możliwość tworzenia planu audytów wewnętrznych, np. zintegrowanego systemu zarządzania, m.in. poprzez: przygotowanie planu audytu i jego etapów, definiowanie terminów i zakresu audytu, określanie zespołu audytorów, obszaru audytowanego (komórki, stanowiska, osoby), kryteriów audytu (procesy, dokumenty, ryzyka, punkty krytyczne itd.) oraz formułowanie pytań kontrolnych. Smart Audyt usprawnia także proces zarządzania działaniami poaudytowymi, w tym definiowanie i monitorowanie działań korygujących oraz zapobiegawczych wynikających z wykrytych niezgodności oraz ze spostrzeżeń z audytu.

– **Smart Portal** jest modułem powszechnie dostępnym dla wszystkich pracowników. Wyświetlane w nim diagramy procesów są w pełni interaktywne. Po wybraniu obiektu użytkownik widzi jego opis i atrybuty. W przypadku podłączenia innego elementu projektu do któregokolwiek z atrybutów po kliknięciu w obiekt otworzy się odpowiedni element (dokument, inny diagram). Poprzez portal użytkownik ma dostęp do plików zdefiniowanych przez Smart Architect (doc, xls, ppt, PDF itd.). Może w każdym momencie pobrać plik i zapisać go w dowolnym miejscu. Dodatkowo program pozwala na wysyłanie wiadomości

pomiędzy użytkownikami modułów Smart Architect a użytkownikami modułu Smart Portal.

Wdrożenie zarządzania procesami pozwala na zebranie szeregu informacji o organizacji i jest punktem wyjścia do przeprowadzenia wielu analiz, w tym m.in. analizy ryzyka czy analizy krytyczności procesów. Dzięki szczegółowemu zdefiniowaniu celów oraz zadań wykonywanych przez przedsiębiorstwo możliwe jest określenie zdarzeń zagrażających ich realizacji. Ma to istotne znaczenie w obliczu możliwych strat, jakie grożą przedsiębiorstwu, które nie wprowadzi stosownych mechanizmów ochrony przed negatywnymi zdarzeniami.

Znany jest przypadek Ericssona i Nokii – gdy w fabryce półprzewodników firmy Philips w Albuguergue wybuchł krótki, 10-minutowy pożar, obaj główni kontrahenci zareagowali zupełnie inaczej. Nokia dzięki szybkiemu przepływowi informacji natychmiast podjęła działania i wynegocjowała z Philipsem sprzedaż całości rezerw produkcyjnych, a także zwiększyła import z krajów azjatyckich dzięki dywersyfikacji dostaw. Jednocześnie, w wyniku natychmiastowej reakcji w sytuacji kryzysowej, umocniła swoją pozycję na rynku lidera wśród producentów telefonów komórkowych. Ericsson dowiedział się o zdarzeniu trzy dni po fakcie, kiedy nie było już możliwości negocjacji z Philipsem, co pociągnęło za sobą zmniejszenie produkcji, a jednocześnie spadek wartości akcji Ericssona. Powyższy przykład pokazuje, że w dzisiejszych czasach każde przedsiębiorstwo, aby osiągnąć zaplanowany cel, podejmuje ryzyko. Wolny rynek stwarza zarówno szanse na osiągnięcie ponadplanowych zysków, jak i ryzyko strat w wyniku niekorzystnych zmian w otoczeniu przedsiębiorstwa oraz błędów w zarządzaniu organizacją. Wszystkie decyzje biznesowe obarczone są ryzykiem, którego zmaterializowanie wiąże się z zaangażowaniem (często niemałych) środków finansowych. Dlatego coraz częściej firmy decydują się na wdrożenie systemu zarządzania ryzykiem.

Istnieje wiele modeli zarządzania ryzykiem, które zostały opisane w powszechnie obowiązujących standardach, m.in. SOX, ISO, COSO. Stosowane na świecie praktyki w tym zakresie nie narzucają konkretnych i jedynie właściwych sposobów ich wdrożenia. Dostarczają raczej pewnych prawidłowych ram do zorganizowania niezbędnych mechanizmów, przy zachowaniu właściwej i sprawdzonej koncepcji. Tym samym żaden standard nie wymusza konkretnego sposobu np. zbierania informacji o zdarzeniach, natomiast wskazuje, co taki mechanizm powinien zapewnić.

Wdrażając system zarządzania ryzykiem w organizacji, należy zadbać o odpowiednie kanały zbierania informacji oraz o opracowanie właściwych mechanizmów ochronnych. Pozyskane w ten sposób dane powinny być wykorzystywane w wybranej metodzie oceny ryzyka i służyć do podjęcia działań zaradczych w przypadku ryzyk nieakceptowanych. Monitorowanie wykonania tych działań, jak i zmieniającego się środowiska, ma pozwolić na zdobycie przez organizację określonej odporności na wystąpienie negatywnych zdarzeń.

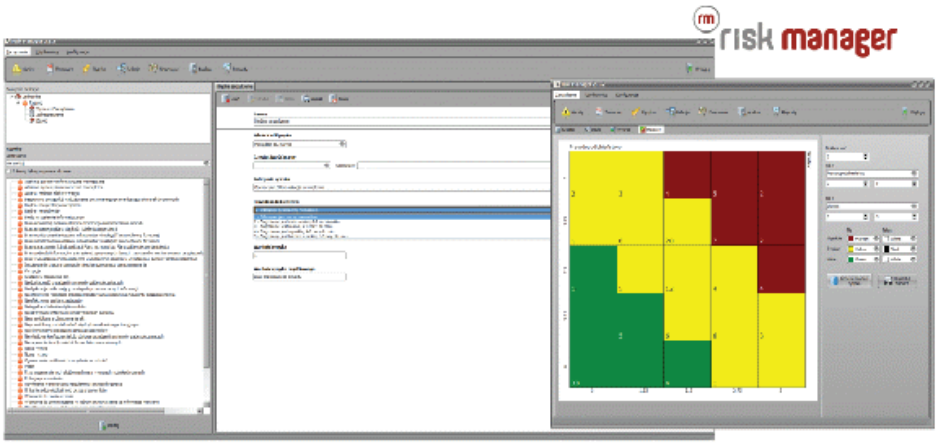
Większość organizacji decyduje się w pierwszej kolejności na wdrożenie zarządzania ryzykiem w jednym wybranym obszarze, np. bezpieczeństwa informacji czy ciągłości działania. Kolejnym krokiem, świadczącym o rozwoju i wzroście świadomości w zakresie budowania odporności organizacji na zagrożenia, jest wdrożenie zintegrowanego zarządzania ryzykiem (*Enterprise Risk Management* – ERM). Podejście to pozwala na integrację zarządzania ryzykiem z istniejącymi w organizacji procesami, co z kolei umożliwia określenie przyszłych zdarzeń, które mogą pozytywnie bądź negatywnie oddziaływać na prowadzoną działalność. Ponadto podejście to wymusza ocenę przyjętej strategii zarządzania organizacją w kontekście przygotowania jej na pojawienie się potencjalnych zagrożeń i obejmuje m.in. takie elementy, jak[5]:

– uzgodnienie apetytu na ryzyko i zweryfikowanie go pod kątem strategii,
– podjęcie decyzji w sprawie reakcji na ryzyko,
– ograniczenie strat wynikających z materializacji ryzyka oraz ze środków przeznaczanych na utrzymanie mechanizmów ochronnych,
– identyfikowanie i zarządzanie wieloma rodzajami ryzyka w przedsiębiorstwie,
– wykorzystywanie możliwości/szans organizacji.

Zintegrowane zarządzanie ryzykiem jest procesem dynamicznym oraz wielokierunkowym i interaktywnym, w którym prawie każdy komponent może wpływać na pozostałe. Wdrożenie efektywnego systemu zarządzania ryzykiem daje kierownictwu narzędzie do osiągania celów poprzez optymalizację realizowanych procesów. Kluczem do sukcesu są dwa podstawowe aspekty. Przede wszystkim należy opracować i wdrożyć skuteczny mechanizm umożliwiający kadrze kierowniczej na bieżąco oceniać ryzyko i dostarczać informacji do podejmowania decyzji. Z drugiej strony – niezbędne jest wsparcie najwyższego kierownictwa w kształtowaniu świadomości co do wagi zarządzania ryzykiem oraz korzyści z tego płynących. Takie podejście jest zalążkiem budowy nowocześnie zarządzanej organizacji.

Podobnie jak w zarządzaniu procesami, wraz z rozwojem systemów zarządzania ryzykiem pojawiła się potrzeba wsparcia informatycznego, które pozwoliłoby na skuteczne i efektywne zarządzanie ryzykiem organizacji. Przykładem takiego narzędzia może być oprogramowanie e-risk (rys. 2), które zostało opracowane w celu wsparcia procesu zarządzania ryzykiem, w tym procesu identyfikacji, analizy i monitorowania ryzyka w prowadzonej działalności. Może być ono wykorzystywane w zintegrowanym zarządzaniu ryzykiem, obejmującym wszystkie obszary organizacji.

[5] Z. Martyniak, *Nowe metody i koncepcje zarządzania*, Wyd. AE w Krakowie, Kraków 2002.



Rys. 2. e-risk manager

Ź r ó d ł o: materiały własne PBSG.

E-risk jest narzędziem elastycznym, pozwalającym zaimplementować dowolną metodykę zarządzania ryzykiem. Kładzie nacisk na możliwość budowania i wdrażania własnych metodyk, również poprzez powiązania danych wynikających z zarządzania procesami. Struktura oprogramowania pozwala dokonać analizy danych na temat ryzyk z wykorzystaniem rozbudowanego kreatora w dowolnym momencie pracy.

Program pozwala na wgląd w dane bieżące i historyczne oraz przedstawienie ich w formie wykresu lub macierzy ryzyka. Przygotowane pod metodykę szablony raportów umożliwiają zestawienie danych, które można wyeksportować do formatu RTF, PDF, JPG, XLS itd. Każdy użytkownik oprogramowania (np. właściciel ryzyka, menedżer ryzyka) ma w systemie przydzielony swój login i hasło oraz sprecyzowany zakres uprawnień. Zidentyfikowane ryzyka przydzielone są do poszczególnych osób, które ze wskazaną częstotliwością mają obowiązek dokonać oceny. Wykorzystanie oprogramowania umożliwia szybki przepływ informacji bez konieczności przekazywania dokumentów w formie papierowej.

Dzięki programowi e-risk zarządzanie – często ogromną – bazą ryzyk jest znacznie prostsze i uporządkowane. Wszystkie ryzyka znajdują się w jednym miejscu, a model zarządzania ryzykiem może być dowolnie konfigurowany. Dodatkowo program usprawnia złożony proces oceny i analizy ryzyka oraz wpływa na zmniejszenie kosztów zarządzania ryzykiem dzięki minimalizacji czasu potrzebnego na przeprowadzanie oceny, przygotowanie raportów i generowanie zestawień.

3. Podsumowanie

Większość organizacji stoi obecnie w obliczu podjęcia decyzji o rozwoju oraz optymalizacji funkcjonującego już podejścia do zarządzania procesami oraz zarządzania ryzykiem. Najbardziej korzystnym w tym przypadku rozwiązaniem jest zintegrowanie tych systemów, a także wykorzystanie profesjonalnego oprogramowania, będącego wsparciem dla tego procesu. Zintegrowane zarządzanie ryzykiem pozwala na przekształcenie powszechnie statycznego podejścia w zarządzanie proaktywne oraz budowanie wartości z wykorzystaniem funkcjonujących mechanizmów zarządzania procesami.

Dostępne obecnie na rynku narzędzia informatyczne wspierające zarządzanie i utrzymanie procesów umożliwiają integrację wielu systemów funkcjonujących w organizacji i stają się coraz bardziej powszechne w polskich przedsiębiorstwach i urzędach. Narzędzia te usprawniają zarządzanie procesami oraz zarządzanie ryzykiem poprzez umożliwienie wdrożenia dowolnej metodyki oraz automatyzację procesu prowadzenia analiz i monitorowania. Ponadto w zależności od stopnia zaawansowania aplikacji umożliwiają szybką wymianę informacji oraz pozbawioną błędów ich aktualizację. Wykorzystuje się je również jako portale informacyjne organizacji, do wskazywania pracownikom prawidłowego przebiegu procesów czy też do prowadzenia oceny ryzyka w całej organizacji.

Połączenie zarządzania procesami oraz zarządzania ryzykiem z narzędziami wspierającymi ten proces w przedsiębiorstwach staje się coraz bardziej powszechne. Korzyści płynące z takich wdrożeń wprost przekładają się na funkcjonowanie organizacji, jej skuteczność i efektywność. Niezależnie od przyczyn uruchomienia projektu wdrożenia narzędzi informatycznych w organizacji podstawowym efektem takiego przedsięwzięcia powinna być możliwość zintegrowanego zarządzania procesami, a także zagrażającymi im ryzykami. Poprawnie wymodelowane procesy i przypisane do nich ryzyka powodują, że menedżerowie otrzymują narzędzie do monitorowania każdego z procesów organizacji, a także otrzymują niezbędne informacje o możliwościach ich optymalizacji lub o konieczności wdrożenia stosownych mechanizmów ochronnych.

Literatura

Brilman J., *Nowoczesne koncepcje i metody zarządzania*, PWE, Warszawa 2002.
Martyniak Z., *Nowe metody i koncepcje zarządzania*, Wyd. AE w Krakowie, Kraków 2002.
Oprogramowanie e-risk, http://e-risk.pl [15.05.2012].
Oprogramowanie Smart, http://oprogramowanie-smart.pl/smart [14.05.2012].

Rola znormalizowanych systemów zarządzania w zrównoważonym rozwoju, red. J. Łańcucki, Wyd. Uniwersytetu Ekonomicznego, Poznań 2011.
Wdrożenie zarządzania ryzykiem, www.pbsg.pl/wdrozenie-zarzadzania-ryzykiem.html [14.05.2012].
Zarządzanie procesami, www.pbsg.pl/zarzadzanie-procesami.html [14.05.2012].
Zintegrowany System Zarządzania Ryzykiem – COSO II. Struktura ramowa, PIKW i PIB, Warszawa 2007.

Zeszyty Naukowe
Wyższej Szkoły Bankowej w Poznaniu
Nr 40/2012

Tomasz Cichowicz, Michał Frankiewicz, Filip Rytwiński, Jacek Wasilewski, Maciej Zakrzewicz

Politechnika Poznańska

Odkrywanie anomalii w szeregach czasowych pochodzących z monitoringu systemów teleinformatycznych

Streszczenie. *Zautomatyzowana analiza szeregów czasowych pochodzących z monitoringu systemów teleinformatycznych jest odpowiedzią na rosnącą złożoność topologiczną i techniczną współczesnych systemów. Jednym z trudniejszych zagadnień z zakresu analizy szeregów czasowych jest wykrywanie anomalii, sygnalizujących awarię lub niewłaściwe użycie systemu teleinformatycznego. W artykule omówiono kontekst wykrywania anomalii w szeregach czasowych pochodzących z monitoringu systemów teleinformatycznych, dokonano przeglądu dotychczasowych metod i algorytmów, zaproponowano dwie nowe metody wykrywania anomalii oraz zaprezentowano wyniki złożonych badań eksperymentalnych.*

Słowa kluczowe*: systemy teleinformatyczne, analiza szeregów czasowych*

1. Wprowadzenie

Złożoność topologiczna i techniczna współczesnych systemów teleinformatycznych wymaga ciągłego monitorowania sprawności i efektywności ich funkcjonowania. Każdy ze składników systemu teleinformatycznego – aktywne urządzenie sieciowe, serwer bazy danych, serwer aplikacji, urządzenie pamięci masowej, aplikacja itp. – dokonuje automatycznych obserwacji swojego stanu pracy, a wyniki tych obserwacji udostępnia zewnętrznym aplikacjom narzędziowym (*Network Monitoring Tools*). Bieżąca analiza wskaźników raportowanych przez składniki systemu teleinformatycznego umożliwia wczesne wykrywanie awarii, identyfikację działań „podejrzanych" (np. ataków typu *Denial of Service*, prób włamania), optymalizację wydajności i utylizacji systemów. Ze względu na rozmiary obecnych systemów i mnogość raportowanych przez nie wskaźników monitorowanie pracy systemów teleinformatycznych bezwzględnie wymaga daleko idącej automatyzacji.

Obecnie wykorzystywane aplikacje narzędziowe służące do monitorowania pracy systemów teleinformatycznych (np. IBM Tivoli, HP Network Node Manager, WhatsUpGold, Solar Winds Orion, Zenoss, Oracle Enterprise Manager) realizują zaledwie podstawowy funkcjonalny poziom automatyzacji. Skupiają się na wizualizacji obserwowanych wielkości w formie interakcyjnych wykresów graficznych, gromadzą odczyty historyczne w celu późniejszej analizy, automatycznie sygnalizują fakt przekroczenia statycznych poziomów alarmowych oraz znaczące odstępstwa od typowego poziomu wartości. Pomimo tych funkcjonalności wymienione narzędzia wymagają jednak zarówno wykonania złożonej wstępnej konfiguracji, jak i późniejszej ciągłej asysty ze strony doświadczonego administratora lub operatora systemu. W przypadku systemów bardzo dużych, w których liczba monitorowanych wielkości osiąga rząd tysięcy i dziesiątków tysięcy, nieprzekraczalną granicą stają się możliwości człowieka w zakresie jednoczesnego śledzenia wielu niezależnych zdarzeń.

Wśród odbiorców i użytkowników aplikacji narzędziowych służących do monitorowania pracy systemów teleinformatycznych coraz wyraźniej artykułowana jest potrzeba stosowania wysoce zautomatyzowanych algorytmów zarządzania, umożliwiających nienadzorowane wykrywanie zjawisk „nietypowych" na podstawie obserwacji wybranych cech statystycznych monitorowanych wielkości. Tak formułowany problem nazywany jest w piśmiennictwie wykrywaniem anomalii (*Anomaly Detection*), a w przedmiotowym obszarze zastosowań – wykrywaniem anomalii w szeregach czasowych (*Time Series Anomaly Detection*). Całkowicie zautomatyzowane wykrywanie anomalii jest zadaniem bardzo trudnym, wymagającym zapożyczeń z takich dziedzin naukowych, jak uczenie maszynowe, sztuczna inteligencja i eksploracja danych. Oczekuje się, że w niedalekiej przyszłości aplikacje narzędziowe służące do monitorowania pracy systemów teleinformatycznych będą wyposażane w takie rozwiązania.

W ogólności, problemu wykrywania anomalii nie zawęża się wyłącznie do analizy danych syntetycznych, generowanych przez urządzenia lub aplikacje komputerowe. Analogiczne wyzwania pojawiają się np. w epidemiologii, gdzie na podstawie statystyk zachorowań podejmuje się decyzje o ogłoszeniu epidemii, czy w sejsmologii, gdzie na podstawie zapisów amplitudy drgań gruntu w czasie przewiduje się nadejścia fal sejsmicznych. Zwykle jednak algorytmów wykrywania anomalii opracowanych dla innych dziedzin nauki nie można przenieść wprost na grunt monitorowania systemów teleinformatycznych – ze względu na inną charakterystykę szeregów czasowych lub inny minimalny czas reakcji (np. w epidemiologii – dni, w monitoringu systemów – sekundy).

W niniejszej pracy przedstawiono wyniki doświadczeń dotyczących zastosowania algorytmów wykrywania anomalii dla szeregów czasowych pochodzących z monitoringu systemów teleinformatycznych. W wyniku przeprowadzonych studiów literaturowych wyłoniono podzbiór algorytmów przystających do zdefinio-

wanych wymagań, a ponadto opracowano kilka specjalizowanych algorytmów własnych. Następnie zrealizowano złożony eksperyment obliczeniowy, w ramach którego badano skuteczność algorytmów wykrywania anomalii. W badaniach eksperymentalnych wykorzystywano rzeczywiste szeregi czasowe opisujące funkcjonowanie systemu klasy *e-commerce*.

2. Specyfika szeregów czasowych pochodzących z monitoringu systemów teleinformatycznych

Cechy charakterystyczne (zmienność, kształt, częstotliwości składowe itp.) szeregu czasowego pochodzącego z monitoringu są w znacznej mierze zależne od typu urządzenia, którego stan pracy podlega obserwacji. Podczas prowadzonych badań dokonano następującej klasyfikacji typów urządzeń/elementów:

a) **urządzenia sieciowe**, dokonujące retransmisji strumieni danych przysyłanych do systemu lub wysyłanych przez system do odbiorców zewnętrznych: karty sieciowe, routery, zapory sieciowe (*Firewall*), punkty dostępowe (*Access Point*). Mierzone wielkości to m.in. gęstość strumienia (rys. 1),

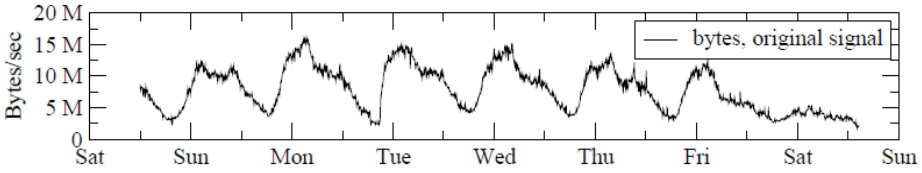
b) **urządzenia pamięci masowej**, zapisujące i odczytujące pliki użytkowników i aplikacji: macierze dyskowe, serwery plików, dyski sieciowe. Mierzone wielkości to m.in.: wykorzystana pojemność, średni czas dostępu, ilość odczytywanych/zapisywanych danych,

c) **oprogramowanie systemowe**, przetwarzające żądania użytkowników i aplikacji: systemy operacyjne, serwery baz danych, serwery aplikacji. Mierzone wielkości to m.in.: liczba równoczesnych sesji/połączeń, wskaźniki wydajnościowe (np. współczynniki trafień w bufory, liczby zdarzeń oczekiwania na zwolnienie blokady), liczba udanych i nieudanych logowań, zajętość pamięci operacyjnej, liczba procesów, liczba współbieżnych wątków,

d) **aplikacje biznesowe**, implementujące funkcje użytkowe: aplikacje ERP, aplikacje CRM, aplikacje *e-commerce* itp. Mierzone wielkości to m.in.: liczba żądań wykonania funkcji biznesowych, liczba równoczesnych sesji, średni czas odpowiedzi,

e) **procesory** wykonujące kod programowy aplikacji biznesowych i oprogramowania systemowego. Ta klasa urządzeń obejmuje zarówno procesory fizyczne, jak i wirtualne. Mierzone wielkości to m.in. chwilowe obciążenie procesora.

W zależności od typu urządzenia/elementu różnego znaczenia nabiera pojęcie anomalii w szeregu czasowym. O wystąpieniu anomalii w gęstości strumienia dostarczanego do karty sieciowej może świadczyć wzrost powyżej poziomu typowego (przeciążenie), ale też spadek do poziomu bliskiego zeru (awaria urządzenia poprzedzającego). Anomalią dla wykorzystanej pojemności macierzy dyskowej będzie zbliżenie się do poziomu 100% zapełnienia. Z kolei 100%



Rys. 1. Przykładowy szereg czasowy pochodzący z monitoringu karty sieciowej (gęstość strumienia wychodzącego, w bajtach na sekundę)

Ź r ó d ł o: P. Barford, J. Kline, D. Plonka, A. Ron, *A signal analysis of network trafic anomalies*, w: *IMW'02 Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurement*, ACM, New York 2002, s. 71-82.

obciążenie nie będzie traktowane jako anomalia w przypadku procesora, o ile nie będzie utrzymywać się przez dłuższy czas. Dla serwera aplikacji przejawem anomalii może być zbyt szybki wzrost liczby sesji (atak typu *Denial of Service*). W wielu przypadkach przejawem anomalii może być inny rozkład szeregu czasowego w skali dnia/tygodnia w odniesieniu do analogicznego okresu w przeszłości. Na przykład liczba 50 nieudanych prób logowania do aplikacji ERP może być naturalna podczas dnia roboczego, ale może świadczyć o anomalii, gdy zostanie odnotowana w niedzielę (próba włamania).

Dla wielu systemów informatycznych typowy jest powolny, stopniowy wzrost obciążenia i rozmiaru wynikający ze zwiększania się liczby użytkowników, klientów, obiektów przechowywanych w bazie danych itp. Zwykle skutkuje to pojawieniem się trendu wznoszącego w szeregu czasowym. Zjawisko takie powinno być uwzględniane przez algorytmy wykrywania anomalii – poprzez wznoszenie stosownych poziomów alarmowych. Podobnego traktowania wymagają zjawiska sezonowości (przejściowy wzrost obciążenia systemów *e-commerce* w okresach przedświątecznych).

W związku ze stosowanymi metodami pomiaru wielkości obserwowanych, opartymi na próbkowaniu czasowym, w analizowanych szeregach czasowych często występują składowe wysokiej częstotliwości, mogące zaburzać działanie algorytmów wykrywania anomalii. W celu eliminacji niepożądanych efektów stosuje się wstępne wygładzanie (*Smoothing*) szeregu czasowego. Intensywność wygładzania musi być jednak dostosowana do specyfiki badanego szeregu, aby nie spowodowała uszkodzenia informacji o występującej anomalii.

3. Dotychczasowe badania

3.1. Badane algorytmy wygładzania szeregu czasowego

Prowadzone badania objęły obserwację skuteczności działania opisanych w literaturze wybranych algorytmów wygładzania szeregu czasowego, takich jak:

średnia krocząca, wygładzanie wykładnicze, podwójne wygładzanie wykładnicze, potrójne wygładzanie wykładnicze, wygładzanie cepstralne.

Najprostszym algorytmem wygładzania jest **średnia krocząca**[1], powszechnie stosowana w finansach i analizie technicznej – średnia arytmetyczna wartości z ostatnich *n* próbek. Występuje w wielu odmianach, jak: ważona średnia krocząca, wykładnicza średnia krocząca, średnia krocząca poprawiona o wolumen, trójkątna średnia krocząca. Może służyć również jako metoda odcinania wartości skrajnych.

Wygładzanie wykładnicze ma na celu zmniejszenie wariancji źródłowego szeregu czasowego za pomocą ważonej średniej kroczącej z przeszłych wartości[2]. Wagi średniej maleją wykładniczo wraz z upływem czasu. Wygładzanie wykładnicze może być zastosowane w usuwaniu szumów oraz prognozowaniu przebiegów czasowych, gdzie stosunek między sygnałem a szumem jest niewielki oraz dane nie wykazują wyraźnego trendu i wahań sezonowych.

Podwójne wygładzanie wykładnicze[3], znane również jako wygładzanie wykładnicze Holta, jest udoskonaleniem modelu zwykłego wygładzania wykładniczego. Uwzględnia występowanie tendencji rozwojowych (trendów), jak i wahania przypadkowe. Opiera się na liniowym modelu Holta:

$$F_{t-1} = \alpha y_{t-1} + (1-\alpha)(F_{t-2} + S_{t-2})$$
$$S_{t-1} = \beta(F_{t-1} - F_{t-2}) + (1-\beta)S_{t-2}$$

gdzie: F_{t-1} – wygładzona wartość zmiennej prognozowanej na moment $t - 1$; S_{t-1} – wygładzona wartość przyrostu trendu na moment $t - 1$; α, β – parametry modelu o wartościach z przedziału [0, 1].

W przypadku niewystępowania trendu lub sezonowości najlepsze rezultaty wygładzania dawała metoda zwykłego wygładzania wykładniczego, natomiast w sytuacji, gdy w szeregach źródłowych pojawiał się trend wznoszący lub opadający, zwykłe wygładzanie wykładnicze wykazywało skłonność do opóźniania wygładzania. Jednakże mimo dobrych wyników wygładzania dla szeregów czasowych wykazujących trend wygładzanie Holta nie sprawdzało się przy szeregach czasowych ujawniających cechy sezonowości. Dla takich przebiegów czasowych bardziej atrakcyjną metodą było potrójne wygładzanie wykładnicze.

Potrójne wygładzanie wykładnicze[4], często nazywane wygładzaniem Holta-Wintersa, uwzględnia sezonowość w szeregu czasowym. Występuje w dwóch

[1] J. Durbin, *Efficient estimation of parameters in moving-average models*, „Biometrika" 1959, nr 3.

[2] *Averaging and exponential smoothing models*, www.duke.edu/~rnau/411avg.htm [01.2012].

[3] Ibidem; *OpenForecastAPI*, http://openforecast.sourceforge.net/docs [01.2012].

[4] *Averaging and exponential smoothing models*, op. cit.; *Triple exponential smoothing*, www.itl.nist.gov/div898/handbook/pmc/section4/pmc435.htm [01.2012]; *OpenForecastAPI*, http://openforecast.sourceforge.net/docs [01.2012].

wersjach modelu: addytywnej i multiplikatywnej. Dla wersji addytywnej równania modelu przedstawiają się następująco:

$$F_{t-1} = \alpha(y_{t-1} - C_{t-1-r}) + (1-\alpha)(F_{t-2} + S_{t-2})$$

$$S_{t-1} = \beta(F_{t-1} - F_{t-2}) + (1-\beta)S_{t-2}$$

$$C_{t-1} = \gamma(y_{t-1} - F_{t-1}) + (1-\gamma)C_{t-1-r}$$

natomiast dla wersji multiplikatywnej:

$$F_{t-1} = \alpha\frac{y_{t-1}}{C_{t-1-r}} + (1-\alpha)(F_{t-2} + S_{t-2})$$

$$S_{t-1} = \beta(F_{t-1} - F_{t-2}) + (1-\beta)S_{t-2}$$

$$C_{t-1} = \gamma\frac{y_{t-1}}{F_{t-1}} + (1-\gamma)C_{t-1-r}$$

gdzie: F_{t-1} – wygładzona wartość zmiennej prognozowanej na moment $t - 1$ po eliminacji wahań sezonowych; S_{t-1} ocena przyrostu trendu na moment $t - 1$; C_{t-1} – ocena wskaźnika sezonowości na moment $t - 1$; r – liczba okresów; α, β, γ – parametry modelu o wartościach z przedziału [0, 1].

Wadą tej metody jest wymagalność przynajmniej jednego zakończonego szeregu czasowego sezonowego do wyznaczenia początkowych estymat wskaźników sezonowości C. Kompletne dane sezonowe składają się z r okresów, ponieważ wymagana jest estymacja współczynnika trendu przy przejściu z jednego okresu do kolejnego. Zalecane jest wykorzystywanie dwóch zakończonych, kompletnych sezonów, tzn. $2r$ okresów – a w praktyce 5-6, gdyż umożliwia to modelowi skuteczniejszą adaptację do danych, a nie ślepe typowanie wartości lub poprawną estymację jedynie dla początkowych elementów. W badaniach wykorzystano model multiplikatywny oraz wymagano dwóch kompletnych cykli danych do inicjalizacji modelu.

Wygładzanie współczynnikami cepstralnymi bazuje na transformacie Fouriera przeniesionej w dziedzinę decybelową (cepstrum) i na oknie dolnoprzepustowym[5]. Kroki algorytmu są następujące:

1. Wykonywana jest szybka dyskretna transformata Fouriera na źródłowym szeregu czasowym.
2. Otrzymany wynik przekształcany jest tak, aby stał się widmem, w którym amplituda wyrażona jest w decybelach.

[5] J.O. Smith III, *MUS421/EE367B applications lecture b: Cross synthesis using cepstral smoothing or linear prediction for spectral envelopes*, https://ccrma.stanford.edu/~jos/SpecEnv/SpecEnv.pdf [01.2012]; *Cepstral smoothing*, https://ccrma.stanford.edu/~jos/SpecEnv/Cepstral_Smoothing.html [01.2012].

3. Za pomocą funkcji okna dolnoprzepustowego obcinane są mało znaczące składowe periodyczne, które z założenia zaburzają sygnał.

4. Wynik transformowany jest odwrotnie w szereg czasowy za pomocą odwrotnej transformaty Fouriera.

3.2. Badane algorytmy wykrywania wystąpienia anomalii

Prowadzone badania objęły obserwację skuteczności działania opisanych w literaturze wybranych algorytmów wykrywania wystąpienia anomalii, m.in.: metod finansowej analizy technicznej, metod ekstrakcji składowych sygnału, metod WSARE, metod opartych na klasyfikatorach decyzyjnych oraz metod trzech sigm.

Badaniom poddano trzy najpopularniejsze **modele analizy technicznej**: MACD (*Moving Average Convergence/Divergence*), Momentum (wskaźnik zmiany ROC) oraz wstęgę Bollingera. MACD[6] jest wskaźnikiem badającym zbieżność i rozbieżność średnich kroczących. Reprezentuje różnice wartości długoterminowej i krótkoterminowej średniej wykładniczej. Produktem tego modelu są dwie linie – MACD oraz linia sygnału (średnia z linii MACD). Moment przecięcia linii sygnału z linią MACD oznacza zmianę trendu, interpretowaną w badaniach jako prawdopodobne wystąpienie anomalii. Momentum[7] oznacza procentową zmianę wartości pomiędzy stanem aktualnym a stanem sprzed *k* punktów czasowych. Osiąganie ekstremum przez ten wskaźnik może być interpretowane jako wzmocnienie trendu – np. anomalia ataku przyrostowego na system lub anomalia zwiększenia wykorzystania systemu. Metoda wstęgi Bollingera[8] zakłada, że zmienność wartości obserwowanego parametru jest dynamiczna, a nie statyczna. Wstęga Bollingera składa się z: (1) wstęgi środkowej, będącej *n*-okresową średnią kroczącą, (2) wstęgi górnej, będącej *k*-krotnością *n*-okresowego odchylenia standardowego powyżej wstęgi środkowej, (3) wstęgi dolnej, która jest *k*-krotnością *n*-okresowego odchylenia standardowego poniżej wstęgi środkowej. Wstęga Bollingera tworzy swoisty korytarz, którego opuszczenie jest traktowane jako anomalia.

Metody ekstrakcji składowych sygnału traktują szereg czasowy jako opis próbek okresowego sygnału ciągłego, w stosunku do którego stosować można

[6] J.J. Murphy, *Technical analysis of the financial markets*, „Pennsylvania Dental Journal" 1999, nr 77(2); *Encyklopedia analizy technicznej*, www.wdsoftware.com/pl/encyklopedia-at/index.html [01.2012].

[7] *Encyklopedia analizy technicznej*, op. cit.; T. Fawcett, *An introduction to roc analysis*, „Pattern Recogn. Lett." 2006, nr 27, s. 861-874.

[8] *Encyklopedia analizy technicznej*, op. cit.; J. Bollinger, *Bollinger on Bollinger bands*, Mc-Graw-Hill, 2001.

techniki wyodrębniania składowych (np. sinusoidalnych). Przy wykorzystaniu takiego modelu przewidywane są przyszłe wartości sygnału, a następnie odnoszone do wartości faktycznie mierzonych – duże odstępstwo wskazuje na wystąpienie anomalii. Badane podejścia obejmowały: szybką transformację Fouriera, falki (*Wavelets*) i analizę głównych składowych (*Principal Component Analysis* – PCA)[9].

Interesującym podejściem jest WSARE (*What's Strange About Recent Events*), pierwotnie opracowane w celu wczesnego wykrywania zagrożeń epidemiologicznych na podstawie danych pochodzących z różnych źródeł, takich jak: przychodnie, szpitale, stacje meteorologiczne, dane o migracji ludności, ruchu ulicznym itp.[10] Głównym założeniem WSARE jest operowanie na dyskretnym, wielowymiarowym zbiorze danych i porównywanie wektora wartości teraźniejszych do danych historycznych, np. w postaci statystyk. W związku z docelowym zastosowaniem WSARE projektowano tak, aby bez względu na konkretne zastosowane algorytmy wykrywanie anomalii odbywało się szybko, a ogólna złożoność obliczeniowa była stała lub liniowa względem rozmiaru historii. Opublikowano trzy oficjalne implementacje (wersje) WSARE: 2.0, 2.5 i 3.0.

Metody wnioskowania probabilistycznego przy wykorzystaniu **klasyfikatorów decyzyjnych** opierają się na przewidywaniu prawdopodobieństwa wystąpienia określonej przyszłej wartości próbki w szeregu czasowym, a następnie porównania wartości przewidywanej z wartością faktycznie odnotowaną. Najbardziej rozpowszechnionym modelem pozwalającym określać prawdopodobieństwo zajścia pewnego ciągu zdarzeń są sieci Bayesa. Modelują one zależności przyczynowe poprzez tworzenie acyklicznego grafu skierowanego. Wierzchołki tego grafu reprezentują zdarzenia. W kontekście wykrywania anomalii w szeregu czasowym wierzchołkiem może być wartość badanej funkcji w ustalonym momencie czasu. Łuki natomiast modelują związki przyczynowe między zdarzeniami. Dzięki tak stworzonej sieci w łatwy sposób można wyznaczyć prawdopodobieństwo warunkowe zajścia konkretnych zdarzeń w systemie.

Naiwny klasyfikator Bayesa, będący obecnie jednym z popularniejszych klasyfikatorów, jest oparty na regule Bayesa, pozwalającej obliczać prawdopodobieństwo warunkowe zajścia zdarzenia[11]. Wykorzystuje on upraszczające obli-

[9] *Factor analysis*, www.psych.cornell.edu/Darlington/factor.htm [23.01.2012]; W.J. Krzanowski, *Principles of multivariate analysis: a user's perspective*, „Oxford statistical science series", Oxford University Press, Oxford 2000.

[10] W.-K. Wong, A. Moore, G. Cooper, M. Wagner, *What's Strange About Recent Events*, „Journal of Urban Health", czerwiec 2003, Supplement 1; W.-K. Wong, A. Moore, G. Cooper, M. Wagner, *What's Strange About Recent Events (WSARE): An algorithm for the early detection of disease outbreaks*, „Journal of Machine Learning Research" 2005, nr 6.

[11] S. Thrun, P. Norvig, *Online introduction to artificial intelligence*, www.ai-class.com/course/topic/6 [01.2012].

czenia założenie o wzajemnej warunkowej niezależności atrybutów opisujących próbkę względem zmiennej decyzyjnej. Mimo takiego uproszczenia modelowanej rzeczywistości algorytm daje w praktyce bardzo dobre rezultaty, m.in. przy wykrywaniu spamu.

Metoda trzech sigm opiera się na założeniu, że wartości szeregu czasowego przyjmują rozkład zbliżony do krzywej Gaussa. Zgodnie z tą metodą pojawienie się nowej próbki oznacza uaktualnienie średniej wartości dotychczasowej historii oraz odchylenia standardowego w obrębie tej historii. Następnie aktualna próbka jest porównywana z obliczoną średnią i w przypadku różnicy większej niż ustalona wielokrotność odchylenia standardowego (w badanym rozwiązaniu: trzykrotność) licznik sygnału anomalii jest zwiększany o 1. W przeciwnym przypadku licznik jest zerowany. Algorytm sygnalizuje anomalię z chwilą, gdy licznik przekroczy ustaloną wartość.

4. Propozycje nowych rozwiązań

4.1. Profile szeregów czasowych

Motywacją do opracowania metody profili była obserwacja cykliczności i wewnętrznego podobieństwa szeregów czasowych generowanych w zbliżonych warunkach, np. obciążenie serwera poczty elektronicznej we wtorek wykazuje podobną charakterystykę jak obciążenie tego samego serwera w poniedziałek (wysoka aktywność w godzinach pracy biura, niska aktywność w godzinach nocnych). Termin „profil szeregu czasowego" reprezentuje typowy kształt szeregu czasowego zobrazowanego w formie wykresu, wykorzystywany do analizy zachowania się systemu w analogicznych oknach czasowych w przyszłości. Znaczące odstępstwo od kształtu profilu jest sygnałem wystąpienia anomalii.

Profile szeregów czasowych generowano z wykorzystaniem algorytmu dwufazowego, obejmującego wygładzenie szeregu w oknie czasowym, a następnie ekstrakcję cech szeregu. Do wygładzania szeregu zastosowano algorytmy średniej kroczącej i wygładzania wykładniczego. Ekstrakcji cech dokonywano za pomocą funkcji: pochodnej (jako miary szybkości zmian wartości w szeregu czasowym), całki (jako sumy wartości szeregu czasowego), transformaty Fouriera (jako dekompozycji na częstotliwości składowe). Do wykrywania rozbieżności pomiędzy profilem historycznym a profilem bieżącym wykorzystano algorytm trzech sigm.

4.2. Tablice znamionowe

Metoda tablic znamionowych polega na generowaniu zestawów parametrów charakteryzujących zachowanie się szeregu czasowego w historycznych wąskich oknach czasowych, a następnie na porównywaniu tych parametrów z cechami szeregu aktualnego. Strukturę tablicy znamionowej przedstawiono w tabeli 1.

Tabela 1. Struktura tablicy znamionowej szeregu czasowego

Nazwa parametru	Opis
valueMin	minimalna wartość sygnału w oknie
valueMax	maksymalna wartość sygnału w oknie
valueAvg	średnia wartość sygnału w oknie
valueMed	mediana wartości sygnału w oknie
valueStd	odchylenie standardowe wartości sygnału w oknie
derivativeMin	minimalna wartość z pochodnej sygnału w oknie
derivativeMax	maksymalna wartość z pochodnej sygnału w oknie
derivativeAvg	średnia wartość pochodnej sygnału w oknie
derivativeMed	mediana wartości pochodnej sygnału w oknie
derivativeStd	odchylenie standardowe wartości pochodnej sygnału w oknie
fourierArea	pole pod wykresem widma fourierowskiego sygnału

Ź r ó d ł o: opracowanie własne.

5. Eksperymenty obliczeniowe

W celu oceny użyteczności i dokładności metod wykrywania anomalii w szeregach czasowych pochodzących z monitoringu systemów teleinformatycznych przeprowadzono eksperyment obliczeniowy z wykorzystaniem rzeczywistych danych pomiarowych, opisujących działanie m.in.: serwerów Microsoft Active Directory, serwerów poczty elektronicznej, serwerów baz danych, serwerów aplikacji i routerów Cisco. Dane pomiarowe prezentowały wartości takich wskaźników, jak: zużycie dysku, obciążenie procesora, ruch sieciowy przychodzący i wychodzący, zużycie pamięci operacyjnej maszyn wirtualnych Java, zużycie pamięci operacyjnej serwera. Badane algorytmy zostały zaimplementowane w formie wielowątkowej, modułowej aplikacji Java EE uruchomionej na platformie serwera aplikacji Glassfish 3.1. Dane źródłowe były utrwalone w bazie danych Oracle Database 11g. Całość środowiska obliczeniowego była osadzona na platformie Linux CentOS, wyposażonej w 8 rdzeni procesorów i 32 GB pamięci operacyjnej.

Podczas eksperymentów wyszukiwane były zarówno anomalie naturalne, jak i anomalie sztuczne dodane do rzeczywistych danych pomiarowych. Anomalie sztuczne były generowane jako trapezoidalne: (1) punktowe, (2) trójkątne krótkotrwałe, (3) długotrwałe. Na podstawie każdego eksperymentu odczytywano wartości czterech wskaźników:

1) liczba poprawnie wykrytych anomalii (TP – *True Positive*),
2) liczba niepoprawnie wykrytych anomalii (FP – *False Positive*),
3) liczba niewykrytych anomalii (FN – *False Negative*),
4) liczba poprawnie przeanalizowanych próbek bez anomalii (TN – *True Negative*).

Wskaźniki te posłużyły do wyprowadzenia wartości współczynników jakościowych: *czułości* i *swoistości*. Czułość (*Sensitivity*) przedstawia procent anomalii, które zostały poprawnie wykryte, wyrażony za pomocą formuły:

$$Sens = \frac{TP}{TP + FN}$$

Swoistość (*Specificity*) zdefiniowano jako procent poprawnie zdiagnozowanych przypadków, które nie były anomaliami:

$$Spec = \frac{TN}{TN + FP}$$

W celu ułatwienia prezentacji wyników eksperymentów czułość i swoistość złożyły się na *F*-wartość (*F-score*):

$$F = 2 \cdot \frac{Spec \cdot Sens}{Spec + Sens}$$

5.1. Profile szeregów czasowych

Wykonano ponad 60 000 testów opartych na permutacjach zestawu parametrów: 26 szeregów czasowych dla różnych rodzajów urządzeń, 8 rodzajów badanych anomalii, 5 długości okna (1, 2, 4, 12, 24 godziny), rodzaj algorytmu wygładzania (średnia krocząca, wygładzanie wykładnicze), 6 parametrów algorytmu wygładzania (okno o rozmiarze 2, 4, 6, współczynnik wygładzania wykładniczego 0,5, 0,6, 0,7), 2 rodzaje algorytmu ekstrakcji cech (pochodna, całka), 2 rodzaje algorytmu oceny wyniku (trzy sigmy, tolerancja procentowa), 4 poziomy tolerancji procentowej (10, 25, 50, 75).

Przykładowe najlepsze wyniki pomiarów przedstawiono w tabeli 2. Eksperyment został przeprowadzony z wykorzystaniem danych opisujących sieciowy ruch wychodzący z routera Cisco.

Tabela 2. Przykładowe najlepsze wyniki pomiarów dla metody profili szeregów czasowych – dane źródłowe opisujące ruch sieciowy wychodzący

Lp.	A	B	C	D	E	F	H	I
1	1,0000	0,8607	0,9251	wygł. wykł.	całka	alg. 3-sigma	1	0,6
2	1,0000	0,8607	0,9251	śr. krocząca	całka	alg. 3-sigma	1	2
3	1,0000	0,8607	0,9251	wygł. wykł.	całka	alg. 3-sigma	1	0,7
4	1,0000	0,8607	0,9251	śr. krocząca	całka	alg. 3-sigma	1	4
5	1,0000	0,8607	0,9251	wygł. wykł.	pochodna	alg. 3-sigma	1	0,6
6	1,0000	0,8607	0,9251	śr. krocząca	pochodna	alg. 3-sigma	1	4
7	1,0000	0,8607	0,9251	wygł. wykł.	całka	alg. 3-sigma	1	0,5
8	1,0000	0,8607	0,9251	wygł. wykł.	pochodna	alg. 3-sigma	1	0,7
9	1,0000	0,8607	0,9251	wygł. wykł.	pochodna	alg. 3-sigma	1	0,5
10	1,0000	0,8607	0,9251	śr. krocząca	pochodna	alg. 3-sigma	1	2

Objaśnienia: A – czułość, B – swoistość, C – *F*-wartość, D – algorytm wygładzania, E – algorytm ekstrakcji cech, F – algorytm wykrywania anomalii, H – szerokość okna, I – współczynnik wygładzania/długość okna wygładzania

Ź r ó d ł o: badania własne.

Na podstawie przeprowadzonych eksperymentów stwierdzono, że w zakresie wykrywania anomalii za pomocą metody profili szeregów czasowych najskuteczniejsze okazały się następujące kombinacje parametrów algorytmu:

a) badanie szeregu czasowego z użyciem długich okien (od 12 do 24 godzin), ekstrakcja cechy za pomocą funkcji pochodnej, a następnie porównywanie otrzymanych profili procentowo, z przyjętym progiem tolerancji: wysokim (50-75%) w przypadku szeregów czasowych obrazujących specyficzny ruch sieciowy, niskim (10-25%) w przypadku szeregów czasowych o charakterze podobnym do Apache Tomcat NonHeapMemoryUsage,

b) badanie szeregu czasowego z użyciem krótkich okien (godzinnych), ekstrakcja cechy za pomocą funkcji całki lub pochodnej, a następnie wykrywanie anomalii na różnicy otrzymanych profili przy użyciu algorytmu trzech sigm.

5.2. Tablice znamionowe

Wykonano ok. 5000 testów opartych na permutacjach zestawu parametrów: 26 szeregów czasowych dla różnych rodzajów urządzeń, 8 rodzajów anomalii, 6 długości okna (1, 2, 4, 8, 12, 24 godziny), 4 poziomy tolerancji procentowej (10, 25, 50, 75), 4 wartości współczynnika wygładzania dla metody średniej kroczącej (2, 4, 6, 10).

Przykładowe najlepsze wyniki pomiarów przedstawiono w tabeli 3. Eksperyment został przeprowadzony z wykorzystaniem danych opisujących obciążenie procesora.

Tabela 3. Przykładowe najlepsze wyniki pomiarów dla metody tablic znamionowych – dane źródłowe opisujące obciążenie procesora

Lp.	A	B	C	D	E	F
1	1,0000	0,3885	0,5596	0,75	1	4
2	1,0000	0,3869	0,5579	0,75	1	2
3	1,0000	0,3841	0,5550	0,75	1	4
4	1,0000	0,3836	0,5545	0,5	1	2
5	1,0000	0,3836	0,5545	0,5	1	4
6	1,0000	0,3825	0,5533	0,75	1	2
7	1,0000	0,3807	0,5515	0,75	1	4
8	1,0000	0,3793	0,5499	0,5	1	2
9	1,0000	0,3793	0,5499	0,5	1	4
10	1,0000	0,3791	0,5498	0,75	1	2

Objaśnienia: A – czułość, B – swoistość, C – *F*-wartość, D – próg tolerancji, E – szerokość okna, F – współczynnik wygładzania

Ź r ó d ł o: badania własne.

Na podstawie przeprowadzonych eksperymentów stwierdzono, że w zakresie wykrywania anomalii za pomocą metody tablic znamionowych dla szeregów czasowych najskuteczniejsze okazały się kombinacje parametrów: 75% próg tolerancji, okno o rozmiarze 1 godziny, współczynniki wygładzania o wartościach 2 i 4.

5.3. Naiwny klasyfikator Bayesa

Wykonano testy oparte na permutacjach zestawu parametrów: 26 szeregów czasowych dla różnych rodzajów urządzeń, 8 rodzajów anomalii, 8 długości okna uczącego (0,5, 1, 2, 4, 8, 12, 24 godziny, 7 dni).

Przykładowe najlepsze wyniki pomiarów przedstawiono w tabeli 4. Eksperyment został przeprowadzony z wykorzystaniem danych opisujących zużycie pamięci Apache Tomcat HeapMemoryUsage. Najwyższa skuteczność detekcji anomalii została odnotowana dla okna o szerokości 30 minut oraz 7 dni.

Tabela 4. Przykładowe najlepsze wyniki pomiarów dla metody naiwnego klasyfikatora Bayesa – dane źródłowe opisujące zużycie pamięci Apache Tomcat HeapMemoryUsage

Lp.	A	B	C	D
1	0,9500	1,0000	0,9744	10080
2	0,9500	1,0000	0,9744	10080
3	0,9500	0,9994	0,9741	30
4	0,9500	0,9994	0,9741	30
5	0,9500	0,9994	0,9741	30
6	0,9500	0,9994	0,9741	30
7	0,9500	0,9994	0,9741	30
8	0,9500	0,9994	0,9741	30
9	0,9500	0,9991	0,9739	30
10	0,9500	0,9991	0,9739	30

Objaśnienia: A – czułość, B – swoistość, C – *F*-wartość, D – szerokość okna (minuty)

Ź r ó d ł o: badania własne.

5.4. Wnioski końcowe z eksperymentów

Przeprowadzone eksperymenty pozwoliły wskazać, które z algorytmów wykrywania anomalii zapewniają najwyższą dokładność działania w odniesieniu do różnych kategorii szeregów czasowych. Wnioski końcowe były następujące:

– najwyższą skutecznością w wykrywaniu anomalii w szeregach czasowych opisujących zużycie pamięci operacyjnej serwera Apache Tomcat oraz zużycie pamięci fizycznej systemu operacyjnego cechowała się metoda profili szeregu czasowego,

– najwyższą skutecznością w wykrywaniu anomalii w szeregach czasowych opisujących ruch sieciowy wychodzący z serwera i zużycie dysku cechowała się metoda oparta na naiwnym klasyfikatorze Bayesa,

– metoda tablic znamionowych oferowała przeciętną jakość we wszystkich badanych kategoriach.

6. Podsumowanie

W artykule przedstawiono problem wykrywania anomalii w szeregach czasowych pochodzących z monitoringu systemów teleinformatycznych. Główną motywacją dla prowadzonych badań było zapotrzebowanie rynkowe na automatyczne algorytmy wspomagające pracę operatora/administratora dużych syste-

mów teleinformatycznych. Dokonano przeglądu istniejących metod wykrywania anomalii oraz zaproponowano dwie nowe metody: profili szeregu czasowego oraz tablic znamionowych profilu czasowego. W ramach eksperymentalnej ewaluacji potwierdzono ich skuteczność oraz wskazano najbardziej przystające obszary zastosowań. Dalsze planowane prace badawcze obejmą wykorzystanie korelacji pomiędzy różnymi szeregami czasowymi w celu wykrywania zachowań nietypowych.

Literatura

Averaging and exponential smoothing models, www.duke.edu/~rnau/411avg.htm [01.2012].

Barford P., Kline J., Plonka D., Ron A., *A signal analysis of network trafic anomalies*, w: *IMW'02 Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurement, ACM*, New York 2002, s. 71-82.

Bertsekas D., Tsitsiklis J., *Probabilistic systems analysis and applied probability*, http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-041-probabilistic-systems-analysis-and-applied-probability-fall-2010/lecture-notes [01.2012].

Bollinger J., *Bollinger on Bollinger bands*, McGraw-Hill, 2001.

Cepstral smoothing, https://ccrma.stanford.edu/~jos/SpecEnv/Cepstral_Smoothing.html [01.2012].

Chandola V., Banerjee A., Kumar V., *Anomaly detection. A survey. ACM*, „Comput. Surv.", lipiec 2009.

Durbin J., *Efficient estimation of parameters in moving-average models*, „Biometrika" 1959, nr 3.

Encyklopedia analizy technicznej, www.wdsoftware.com/pl/encyklopedia-at/index.html [01.2012].

Factor analysis, www.psych.cornell.edu/Darlington/factor.htm [23.01.2012].

Fawcett T., *An introduction to roc analysis*, „Pattern Recogn. Lett." 2006, nr 27, s. 861-874.

Gao J., Hu G., Yao X., Chang R.K.C., *Anomaly detection of network traffic based on wavelet packet*, APCC '06. Asia-Pacific Conference on Communications, 2006.

Generating mechanical forecasts from statistical models, www.mrp3.com/fcst_models.html [01.2012].

Krzanowski W.J., *Principles of multivariate analysis: a user's perspective*, „Oxford statistical science series", Oxford University Press, Oxford 2000.

Kumar N., Lolla N., Keogh E., Lonardi S., Ratanamahatana Ch.A., *Time-series bitmaps: a practical visualization tool for working with large time series databases*, SIAM 2005 Data Mining Conference, SIAM, 2005, s. 531-535.

Lin J., Keogh E., Lonardi S., Chiu B., *A symbolic representation of time series, with implications for streaming algorithms*, Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery, ACM Press, 2003.

Lo A.W., Mamaysky H., Wang J., *Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation*, „The Journal of Finance" 2000, nr 55(4), s. 1705-1770.

Murphy J.J., *Technical analysis of the financial markets*, „Pennsylvania Dental Journal" 1999, nr 77(2).

Naiwny klasyfikator Bayesa, www.statsoft.com.pl/textbook/stnaiveb.html [01.2012].

Ng A., *Machine learning*, www.ml-class.org/course/auth/welcome [01.2012].

OpenForecastAPI, http://openforecast.sourceforge.net/docs [01.2012].

Smith III J.O., *MUS421/EE367B applications lecture b: Cross synthesis using cepstral smoothing or linear prediction for spectral envelopes*, https://ccrma.stanford.edu/~jos/SpecEnv/SpecEnv.pdf [01.2012].

Stefanowski J., *Analiza szeregów czasowych*, www.cs.put.poznan.pl/jstefanowski/aed/TPtimeseries.pdf [01.2012].

Thrun S., Norvig P., *Online introduction to artificial intelligence*, www.ai-class.com/course/topic/6 [01.2012].

Triple exponential smoothing, www.itl.nist.gov/div898/handbook/pmc/section4/pmc435.htm [01.2012].

Wei L., Kumar N., Lolla V., Keogh E., Lonardi S., Ratanamahatana Ch.A., *Assumption-free anomaly detection in time series*, Proceedings of the 17th International Conference on Scientic and Statistical Database Management 2005, s. 237-242.

Wong W.-K., Moore A., Cooper G., Wagner M., *Bayesian network anomaly pattern detection for disease outbreaks*, Proceedings of the Twentieth International Conference on Machine Learning, Menlo Park, California, lipiec 2003, AAAI Press, s. 808-815.

Wong W.-K., Moore A., Cooper G., Wagner M., *What's Strange About Recent Events*. „Journal of Urban Health", czerwiec 2003, Supplement 1.

Wong W.-K., Moore A., Cooper G., Wagner M., *What's Strange About Recent Events (WSARE): An algorithm for the early detection of disease outbreaks*, „Journal of Machine Learning Research" 2005, nr 6.

Abstracts

Piotr Adamczewski

E-logistics as a factor in development of intelligent organization in knowledge society

E-logistics is based on organization-wide ICT-systems of interconnected solutions primarily related to finance, sales, and operations. By integrating these and other potentially critical business functions, e-logistics is a powerful tool for integrating and managing information to ultimately drive greater business performance and efficiency. But like so many other aspects of information technology, e-logistics is always evolving and successful ICT professionals are highly conscious of the need for credible information on the trends and innovations that are reshaping, and can and will reshape the landscape of e-logistics use and implementation.

Łukasz Balicki

Market conditions in SaaS-model

SaaS model applications are one of the best solutions to improve every organisation's performance and increase the number of customers. The article dicusses all aspects of the market implementation of the new SaaS service, its stages and critical points. A great emphasis was put on marketing. A few techniques as eg. reverse risk management were described in more detail. The aim of the article is to popularize the knowledge and practical aspects of successful SaaS implementation to the market model.

Dariusz Ceglarek

Applying semantic compression in Natural Language Processing tasks

Semantic compression is a new technique that enables to attain correct generalisation of terms in a given context. Thanks to this generalisation, some common thought can be detected in different documents. The rules governing the generalisation process are based on a data structure referred to as a domain frequency dictionary. Having established the domain for a given text fragment a disambiguation of possibly many hypernyms becomes a feasible task. Semantic compression, thus informed generalisation, is possible through the use of semantic networks as a knowledge representation structure. In the light of given overview, one can see that semantic compression makes possible a number of improvements in comparison to already established Natural Language Processing techniques. These improvements along with detailed discussion of various elements of algorithms and data structures necessary to make the semantic compression a viable solution are the core of this work. The semantic compression can be applied in a variety of scenarios. The original scenario for which the semantic compression was introduced was plagiarism detection. With the increasing effort spent on development of the semantic compression, new domains of application were discovered. Thanks to the remodeling of already existing data sources to match the algorithms enabling

the semantic compression, it became possible to use it as a base for an automaton. Thanks to the exploration of hypernymhyponym and synonym relations the automaton is capable of discovering new terms that may be included in the knowledge representation structures.

Wojciech Fliegner

Standardization of electronic financial reporting

Nowadays issuers and receivers of financial reporting have to deal with plenty of different formats of financial statements. The preparation of financial statements for the needs of investors, tax offices, statistics offices, exchange securities and for internal users is a burden for companies and makes comparisons and analyses of these statements for their users time consuming. It causes difficulties that bring higher costs as well as longer and less efficient decisions. A way of resolving this issue may be an introduction of one common format of preparation and presentation of financial statements in a global perspective – XBRL. This article presents the concept of XBRL, its benefits for companies and the state of works and activities of different international and national organizations involved in XBRL promotion and development.

Jędrzej Musiał

Extended version of Internet shopping optimization problem

A customer would like to buy a given set of products in a given set of Internet shops. For each Internet shop, standard prices for the products are known as well as a concave increasing discounting function of total standard and delivery price. The problem is to buy all the required products at the minimum total discounted price and with different variation of shipping cost. Computational complexity of various special cases is established. Properties of optimal solutions are proved and polynomial time and exponential time solution algorithms based on these properties are designed. Two heuristic algorithms are suggested and computationally tested.

Bogdan Pilawski

Bringing data into the data repositories using ETL tools

The article undertakes to discuss some selected, basic aspects of feeding the data repositories with ETL (Extract – Transform – Load) software tools. With the evolution of those repositories in the background, the characteristics of ETL tools are discussed, and also their trends of development, and newly emerging requirements and expectations. A few practical use cases complement the considerations.

Maciej Skała, Iga Stróżyk

Integrated process and risk management

The purpose of this article is to present the support which can be given to an organization in an implementation of new management concepts like process or risk management by IT systems. With the increase of popularity of management systems more and more organizations have decided

to change the traditional way of management to process management. The article describes possibilities of IT support while implementing such approach to an organization. A similar thing is risk management which has become an inherent element of organization. Currently the aim is to integrate risk management and process management, so they include all of the organization aspects (from the strategy, through objectives, processes and tasks). In order to support the implementation of integrated risk management new IT software has been made, the scope of which is presented in the article.

Tomasz Cichowicz, Michał Frankiewicz, Filip Rytwiński,
Jacek Wasilewski, Maciej Zakrzewicz

Anomaly detection in time series for system monitoring

Automated analysis of time series describing performance indicators is a common requirement for efficient monitoring of large, complex, distributed IT systems. One of the most challenging tasks in time series analysis is anomaly detection as the anomalies may indicate failures or misuse of an IT system. In this paper we focus on anomaly detection methods in time series describing key performance indicators of an IT system. We study the existing methods and propose two new approaches. Our results have been verified in a series of experiments.

Lista recenzentów współpracujących z czasopismem „Zeszyty Naukowe Wyższej Szkoły Bankowej w Poznaniu”

(List of reviewers collaborating with “The Poznan School of Banking Research Journal”)

Prof. nadzw. dr hab. Agnieszka Alińska – *Szkoła Główna Handlowa w Warszawie*
Prof. dr Artem Bardas – *National Mining University, Dnipropetrovsk, Ukraina*
Prof. zw. dr hab. Ewa Maria Bogacka-Kisiel – *Uniwersytet Ekonomiczny we Wrocławiu*
Prof. nadzw. dr hab. Jan Borowiec – *Uniwersytet Ekonomiczny we Wrocławiu*
Prof. zw. dr hab. Grażyna Borys – *Uniwersytet Ekonomiczny we Wrocławiu*
Prof. nadzw. dr hab. Stanisław Czaja – *Uniwersytet Ekonomiczny we Wrocławiu*
Prof. nadzw. dr hab. inż. Anna Beata Ćwiąkała-Małys – *Uniwersytet Wrocławski*
Prof. nadzw. dr hab. Waldemar Dotkuś – *Uniwersytet Ekonomiczny we Wrocławiu*
Prof. nadzw. dr hab. Józef Dziechciarz – *Uniwersytet Ekonomiczny we Wrocławiu*
Prof. zw. dr hab. Teresa Famulska – *Uniwersytet Ekonomiczny w Katowicach*
Dr Donald Finlay – *Coventry University Business School, Wielka Brytania*
Prof. zw. dr hab. Stanisław Flejterski – *Uniwersytet Szczeciński*
Dr Klaus Haberich – *Franklin University, USA*
Prof. nadzw. dr hab. Jerzy Ryszard Handschke – *Uniwersytet Ekonomiczny w Poznaniu*
Prof. dr hab. Eva Horvátová – *Ekonomická univerzita v Bratislave*
Prof. nadzw. dr hab. Maria Jastrzębska – *Uniwersytet Gdański*
Prof. zw. dr hab. Adam Kopiński – *Uniwersytet Ekonomiczny we Wrocławiu*
Prof. zw. dr hab. inż. Dorota Elżbieta Korenik – *Uniwersytet Ekonomiczny we Wrocławiu*
Prof. zw. dr hab. Stanisław Korenik – *Uniwersytet Ekonomiczny we Wrocławiu*
Prof. nadzw. dr hab. Maria Kosek-Wojnar – *Uniwersytet Ekonomiczny w Krakowie*
Prof. nadzw. dr hab. Mirosława Lasek – *Uniwersytet Warszawski*
Prof. zw. dr hab. Teresa Krystyna Lubińska – *Uniwersytet Szczeciński*
Prof. nadzw. dr hab. Bartłomiej Nita – *Uniwersytet Ekonomiczny we Wrocławiu*
Prof. zw. dr hab. Edward Nowak – *Uniwersytet Ekonomiczny we Wrocławiu*
Prof. zw. dr hab. Adam Nowicki – *Politechnika Częstochowska*
Prof. zw. dr hab. Walenty Ostasiewicz – *Uniwersytet Ekonomiczny we Wrocławiu*
Prof. nadzw. dr hab. Zbigniew Pastuszak – *Uniwersytet Marii Curie-Skłodowskiej w Lublinie*
Prof. zw. dr hab. Kazimierz Perechuda – *Uniwersytet Ekonomiczny we Wrocławiu*
Prof. zw. dr hab. Bogusław Pietrzak – *Szkoła Główna Handlowa w Warszawie*
Prof. nadzw. dr hab. Marzanna Poniatowicz – *Uniwersytet w Białymstoku*
Prof. zw. dr hab. Wanda Ronka-Chmielowiec – *Uniwersytet Ekonomiczny we Wrocławiu*
Prof. nadzw. dr hab. Henryk Salmonowicz – *Akademia Morska w Szczecinie*
Prof. nadzw. dr hab. Jadwiga Sobieska-Karpińska – *Uniwersytet Ekonomiczny we Wrocławiu*
Prof. zw. dr hab. Jerzy Sokołowski – *Uniwersytet Ekonomiczny we Wrocławiu*
Prof. dr Christopher Washington – *Franklin University, USA*
Prof. nadzw. dr hab. dr h.c. inż. Tadeusz Zaborowski – *Polska Akademia Nauk Oddział w Poznaniu*
Prof. nadzw. dr hab. Ewa Ziemba – *Uniwersytet Ekonomiczny w Katowicach*
Prof. zw. dr hab. Marian Żukowski – *Katolicki Uniwersytet Lubelski Jana Pawła II*

Redaktorzy statystyczni (Statistical editors)

Prof. nadzw. dr hab. Maria Chromińska
Dr Rafał Koczkodaj