

**МЕТОДИ И СОФТУЕР ЗА ИЗВЛИЧАНЕ НА ЗНАНИЯ ЗА ОЦЕНКА НА  
ФИРМЕНА КОНКУРЕНТОСПОСОБНОСТ****DATA MINING METHODS AND SOFTWARE FOR FIRM  
COMPETITIVENESS EVALUATION**

**Боряна Делийска**  
*Лесотехнически университет*

**Георги Маноилов**  
*CSC-Bulgaria*

**Аделина Иванова**  
*Лесотехнически университет*

**Abstract**

*The problem of firm competitiveness evaluation is inextricably linked with IT methods and applications. It is decided using diverse statistical methods and software. But considering necessity of large firm datasets processing and their different content depending on business field, for the purpose data mining methods and software are examined. On the base of data mining application overview, in the article workflow model of firm competitiveness evaluation is developed using RapidMiner tools. This model includes preprocessing, modeling and visualization operations on firm data stored in datawarehouse.*

**Keywords:** firm competitiveness, data mining, data warehouse, workflow.

**1. ВЪВЕДЕНИЕ**

Конкурентоспособността, която има различни измерения – глобална, регионална, фирмена, стокова и т.н., винаги е била трудна и спорна концепция. Няма единство относно нейната точна дефиниция, както относно системата от критерии, мерки и индикатори, така и в интерпретациите на резултатите от изчисляването ѝ. Показателите и движещите сили на конкурентоспособността имат много измерения и сложни отношения.

*Фирмената конкурентоспособност* показва способността на фирмата да произвежда и продава по-добри продукти от тези, предлагани от конкурентите. Превъзходството се оценява по фактори като цена, качество, технологични предимства и др. [2]. Фирмената конкурентоспособност може да се интерпретира и като “асиметрия или разлика между фирмите относно всякакви сравними величини, което позволява на една фирма да се конкурира по-добре от нейните съперници” [7].

Според други автори [8], обобщените финансовите показатели, като

възвръщаемост на продажбите или (оперативна) норма на печалба (return on sales, ROS); възвръщаемост на активите (return on assets, ROA); възвръщаемост на собствения капитал (return on equity, ROE) са най-важни показатели за нивото на фирмената конкурентоспособност.

Изчислението и анализа на фирмената конкурентоспособност са невъзможни без прилагането на съвременни информационни и комуникационни технологии (ИКТ). Целта на настоящото изследване е да се направи преглед на използваните за тази цел ИКТ и да се предложи нов метод и модел за извличане на знания за оценка на фирмената конкурентоспособност.

В следващия раздел е направен преглед и анализ на ИКТ и известни методи за изчисление за фирмена конкурентоспособност. В раздел 3 се предлага нов подход и модел за управление и извличане на знания за нея. В заключението се обсъждат перспективите и проблемите за реализирането на интелигентна система за управление и извличане на знания за фирмената конкурентоспособност.

## 2. ПРЕГЛЕД НА МЕТОДИ И ИКТ ЗА ИЗЧИСЛЕНИЕ НА ИНДЕКС НА ФИРМЕНА КОНКУРЕНТОСПОСОБНОСТ

Най-общо, фирмената конкурентоспособност се оценява с индекс, който е функция от множество променливи величини, наричани още стойностни драйвери (value drivers), показатели или критерии. Всеки критерий съдържа различен брой *индикатори* [4]. Критериите и индикаторите имат *тегловни коефициенти* (weights). За изчисляване на индекса на фирмената конкурентоспособност се прилагат различни статистически методи и софтуер, но няма единна методика за подбор на критериите и индикаторите, участващи в този индекс, както за и техните тегловни коефициенти. Практически, подборът им зависи силно от националните особености на икономиката и от конкретния бизнес сектор.

Индексът на фирмена конкурентоспособност *IFC* [8] се изчислява като линейна зависимост:

$$IFC = \sum_{i=1}^{NC} w_i C_i \quad (1)$$

където:

$w_i$  – тегловен коефициент на  $i$ -тия критерий  $C_i$ , където  $i = 1, 2, \dots, NC$ ;  $NC$  – брой критерии;

$C_i = \sum_{j=1}^{NI_i} IN_j$ , където  $j = 1, 2, \dots, NI_i$ ;  $NI_i$  – брой индикатори в  $i$ -я критерий  $C_i$ ;

Посоченото уравнение не е единственото намерено в специализираната литература, но всички останали се свеждат до него.

Очевидно, голямото разнообразие от фактори, влияещи върху фирмената конкурентоспособност, затруднява нейната обективна оценка. Различните автори дават предимство на едни или други критерии и индикатори, но е необходимо прилагането на формален подход за обосновка на избора им. За избора на критерии и за определяне на тегловните им коефициенти се използват качествени методи (метод "Делфи", пазарни проучвания, аналогия на жизнен цикъл и груповата експертна оценка) и количествени математически и статистически методи (на линейното програмиране, невронните мрежи, дървото

на решенията, дисперсионния и регресионния анализ и др.)

За изчисление на индекса за фирмена конкурентоспособност се прилагат софтуерни продукти за статистика с универсално предназначение, като STATISTICA, SPSS, STATGRAPHICS, R, GRETL, EViews, QstatLab и др.

Известни са специализирани разработки, патентовани в САЩ, например: метод и алгоритъм за класификация на конкурентни предимства [3], онлайн система и метод за кредитна оценка [6], онлайн анализ на конкурентоспособността в реално време [5] и др.

От 1990 г насам има редица изследвания и проекти за прилагане на методите и инструментите за извличане на знания (data mining) и генериране на решения (decision making) за създаване на системи за анализ и откриване на скрити закономерности в големи, натрупани с времето, масиви от данни за фирмено управление. Това са системи с изкуствен интелект, в които се прилагат и технологиите на Семантичната мрежа.

Например за прогнозиране на конкурентоспособността на фирма в [9] е разработена система за извличане на знания и генериране на правила, въз основа на различни комбинации от 10 критерия като: интелигентно управление, професионална компетентност, иновативност, почтеност, динамика, дързост, надеждност, капацитет и др., на които са присвоени стойности – много добър (1), добър (0,83), среден (0,5) и т.н.

Подобна е и системата, създадена по проект [1] за Европейско виртуално предприятие.

От друга страна, съществуват софтуерни пакети за извличане на знания с общо предназначение, като WEKA, STATISTICA Data Miner, SPSS, Knime, SAS, DMQL, 80legs, RapidMiner and Rapid-I, TextSTAT, AntConc, ConcApp, Microsoft SQL Server Data Mining Add-Ins, IBM InfoSphere Warehouse, Oracle Data Mining и др. Считаме, че те биха могли да се използват във фирменото управление, включително и за изследване на фирмената конкурентоспособност.

### 3. МЕТОД И МОДЕЛ ЗА ИЗВЛИЧАНЕ НА ЗНАНИЯ ЗА ФИРМЕНА КОНКУРЕНТОСПОСОБНОСТ

В настоящата работа е формулиран собствен метод и модел за решаване на проблема за подбор на критерии и индикатори, както и тегловните им коефициенти, участващи във формулата за изчисляване на индекса на фирмена конкурентоспособност, чрез извличане на знания от хранилища на фирмени данни. Подходът се състои в следните стъпки:

- избор на потребителски критерии и събиране на експертни мнения за конкурентоспособността на фирмени продукти;

- дефиниране на набор от възможни критерии и индикатори за фирмена конкурентоспособност, подлежащи на оценка;

- съставяне на последователност от операции за извличане на знания от хранилище за фирмени данни, включително:

- определяне на стойностите на тегловните коефициенти (в интервала  $[0,1]$ ) на избраните критерии чрез създаване и изпълнение на заявки за извличане на знания;
- подбор на критерии, чиято сума от тегловни коефициенти е близка до 1 (например  $\geq 0.95$ );
- дефиниране на формула за изчисление на индекса на фирмена конкурентоспособност;
- ранжиране на фирмите по индекса за конкурентоспособност.

и др.

#### 2.1. Избор на потребителски критерии

Потребителската оценка на конкурентоспособността на един продукт (стока) на пазара може да включва различни критерии, от които за настоящия модел избираме три: цена (C), качество (Q) и поддръжка (сервиз) (S). При това, оценката  $O_{iPj}$  на един продукт  $Pj$  ( $j = 1, 2, \dots, n_p$ ), произведен от фирма  $Fk$  ( $k = 1, 2, \dots, n_f$ ), дадена от  $i$ -я експерт ( $i = 1, 2, \dots, n_e$ ), се изчислява по формулата:

$$O_{iPjFk} = w1C_{iPjFk} + w2Q_{iPjFk} + w3S_{iPjFk}, \quad (1)$$

където  $w1$ ,  $w2$ ,  $w3$  са тегловни коефициенти, чиято сума  $w1+w2+w3 = 1$ ;

$C_{iPjFk}$ ,  $Q_{iPjFk}$ ,  $S_{iPjFk}$  са оценките на отделните критерии за всеки продукт по избрана бална система, например, шестстепенна – отлична (1), много добра (0.8), добра (0.6), средна (0.4), слаба (0.2) и лоша (0). В случая общата потребителска оценка:

$$O_{Pj} = \frac{\sum_{i=1}^{n_e} O_{iPjFk}}{n_e} \quad (2)$$

на продукта е в интервала  $[0,1]$ .

Експертните мнения за определен набор от продукти  $P1, P2, \dots, Pn$ , произведени от фирмите в един бизнес сектор, се събират чрез анкета и се записват в хранилището за фирмени данни.

#### 2.2. Дефиниране на критерии и индикатори за фирмена конкурентоспособност

Отделните автори посочват различни набори от критерии, съществени за фирмената конкурентоспособност. Сред тях най-често са [10]: *производителност на труда* (с индикатори: разходи за заплати, размер на печалбата и амортизационни отчисления), *финансови резултати* (коефициент на рентабилност, коефициент на обща ликвидност и коефициент на финансова автономност), *растеж* (нарастване на обема на продажбите, на относителния пазарен дял спрямо основния, на дълготрайните активи и на персонала), *иновативност* (брой на въведените в производство и продажба нови продукти, брой на реализираните проекти за подобряване на техническата база и проекти за организационно-управленски подобрения), *производствена и маркетингова гъвкавост* (продуктова гъвкавост на производството, обемна гъвкавост на производството и маркетингова гъвкавост), *адаптивност към пазара* (способност на предприятието да обновява пазарите си, средно време за разработване, производство и въвеждане на нов продукт или на модифициран продукт) и др. Според Стоенчев [11], съществените критерии са: *производителност на труда, разходи за активи, налични капитали,*

разходи за иновации, разходи за маркетинг. Съществуват и други варианти, но в повечето публикации се посочват гореизброените критерии.

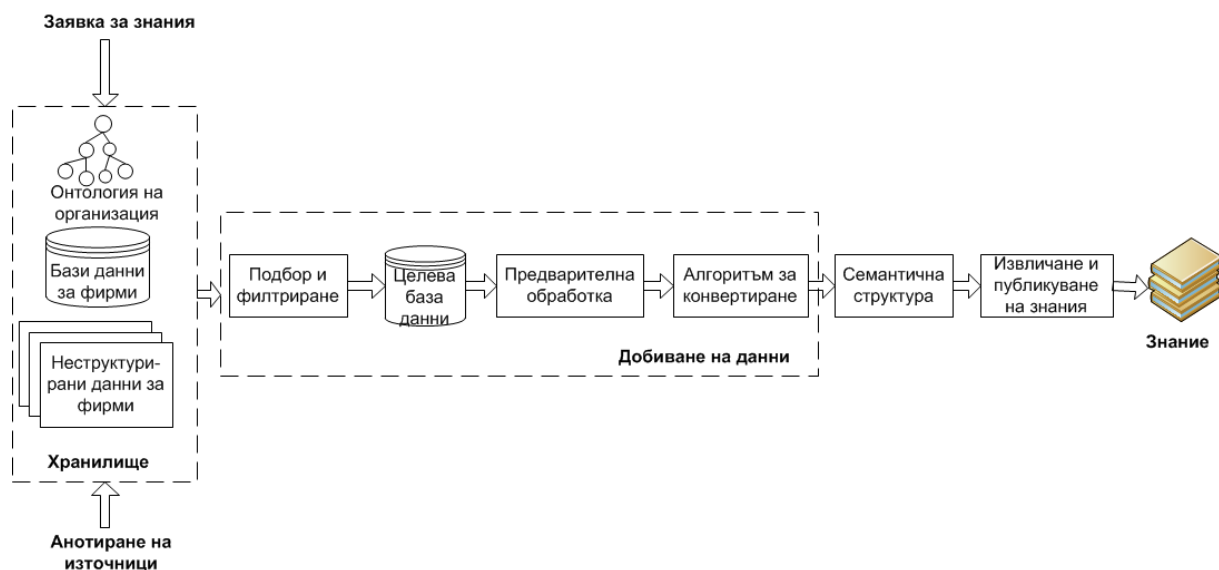
Приемаме, че стойността на критерия *продуктова конкурентоспособност*, който се оценява от експертите, се съпоставя с останалите критерии от избрания набор с цел установяване на тегловните им коефициенти. Стойностите на критериите се поддържат във фирмените бази данни.

### 2.3. Създаване на последователност от операции за извличане на знания

В тях за целта се организират потоци от операции (workflow), включително:

- съставяне и въвеждане на заявка за семантично търсене;
- анализ и конвертиране на заявката на език, разбираем от хранилище за събиране, съхраняване и поддържане на данни за предметната област;
- извличане на данни от хранилището и изграждане на целева база данни в съответствие със заявката;
- конвертиране на данните и изграждане на семантична структура на намерените в тях концепции и характеристики;
- извличане на знания от структурата;
- разпространение на знанията.

Тази последователност, конкретизирана за изчисление на фирмената конкурентоспособност, е представена схематично на фиг. 1.



Фиг. 1. Схема на извличане на знания за фирмена конкурентоспособност

Заявките извличат 3 основни типа знания:

- Прогнозиране на изменението на индекса на фирмената конкурентоспособност, предвид стойностите на критериите за различни интервали от време;

- Извличане и описание на съществените критерии за изчисляване на индекс на фирмена конкурентоспособност за определен интервал от време;

- ранжиране на фирми според конкурентоспособността им.

Заявките се въвеждат чрез интерфейса на системата и са от вида:

*Как влияе критерий ... върху оценката на продукт ... на фирма ... за период ...*

*Кои са критериите, влияещи 90% върху оценката на фирма ...*

*Как се класифицират фирмите според индекс на конкурентоспособност за период...*

и т.н.

Хранилището или складът (data warehouse) обхваща множество от полу-автономни бази данни или такива, организирани в многомерен куб за фирми в избран бизнес сектор. Те се събират и поддържат въз основа на предварителното им аотиране. Организацията на хранилището и съдържанието му са описани в онтология, която може да се разглежда от една страна като метаданни, а от друга – като ресурс за анализ на семантичната заявка.

В резултат от анализа на заявката за знания, се избират търсените обекти (таблицы) и намерените в тях данни се организират в т.нар. *целева база данни*, представляваща източник за процеса на добиване на знания. В нея се изпълнява предварителна обработка, включваща операции за изчистване, селектиране и трансформиране. Изчистването е свързано с откриване и премахване на некоректните и непоследователни данни. Селектирането отстранява данните, които не съответстват на входната заявка за знания. Трансформациите се отнасят до промяна на типа на данните (числови, текстови и др.), на мярката за стойностите и дименсиите им или до създаването на нови данни от съществуващите. Така създадената аналитична структура се конвертира в семантична чрез операции за търсене на закономерности като:

- Асоциация – между критерии, при които честотата на едновременна поява на определени техни стойности е сравнително висока;

- Класификация – функционална зависимост (в проценти) на експертна оценка на продукт от избрани критерии;

- Клъстериране/сегментиране – определят се групи на значимите за оценката критерии и техните тегловни коефициенти.

- Тенденция/регресия – прогнозиране на стойността на индекса на конкурентоспособност, на отделен критерий или група критерии, в зависимост от стойностите на друг критерий или за период от време.

Получената семантична структура представлява: правила за класификация, асоциативни правила, групи с определени характеристики или регресионна зависимост. Описва се с таблица, онтология, тезаурус, невронна мрежа и др.

Извлечените знания съдържат и оценка на достоверността им. Тя е във вид на точност (напр. процент на верните прогнози), квадратична абсолютна или относителна грешка, коефициент на корелация или регресия, в зависимост от типа на заявката.

## ЗАКЛЮЧЕНИЕ

Предимствата на семантичните методи и технологии се състоят във възможностите им за представяне в дълбочина на предметната област. При всяка заявка за семантично търсене се конструират различни целеви бази и семантични структури, въз основа на които се извличат знания. Съществен недостатък на семантичните методи и технологии понастоящем е сложното и скъпото им проектиране, разработване и внедряване. Но предвид големия им потенциал за практиката, те непрекъснато се усъвършенстват.

По-нататъшната ни работа е свързана с поэтапно разработване на конкретна система за управление и извличане на знания за фирмена конкурентоспособност в избран бизнес сектор, които биха били полезни за вземането на управленски решения.

## ЛИТЕРАТУРА

[1] Data Mining and decision support for business competitiveness: Solomon European Virtual Enterprise (SOL-EU-NET). IST-1999-11495 (2000-2002), in <http://SolEuNet.ijs.si>

[2] Deperru D., Cerrato D. (2005). Analyzing International Competitiveness At The Firm Level: Concepts And Measures. Working Paper n. 32, Dipartimento Scienze Sociali - Sezione Economia Aziendale, Universita Cattolica del Sacro Cuore, Piacenza

[3] Gary A. (2008). Competitive advantage rating method and apparatus. From: <http://www.faqs.org/patents/app/20080300960>

[4] Gelei A. (2004). Competitiveness: A match between value drivers and competencies in the Hungarian automotive supply chain. from [http://www.iimm.org/knowledge\\_bank/IFPSM/Andrea%20Gelei.pdf](http://www.iimm.org/knowledge_bank/IFPSM/Andrea%20Gelei.pdf)

[5] Hoffman P. (2001). Real-time on-line competitiveness analysis, United States Patent Application 20010034633. from: <http://www.freepatentsonline.com/y2001/0034633.html>

[6] Lee Eun-woo. (2003). On-line credit assessment system and method, United States Patent Application 20030140000. from: <http://www.freepatentsonline.com/y2003/0140000.html>

[7] Ma H. (2000). Of Competitive Advantage: Kinetic and Positional, *Business Horizons*, 43(1), 53-64

[8] Wang Man, Li Qian. (2007). *Empirical Analysis on Company's Competitiveness from Financial Index*. *Chinese Business Review*, Feb. 2007, Volume 6, No.2 (Serial No.44), ISSN1537-1506, USA

[9] Wütrich B. (1997). *Discovering Probabilistic Decision Rules*. *International Journal of Approximate Reasoning*, December 1997, pp. 269–277

[10] Велев Мл. (2004). *Оценка и анализ на фирмената конкурентоспособност*. София, ИК "Софттрейд" изучаване на фирмената конкурентоспособност, Изд. „Авангард прима”, София

[11] Стоенчев Н. (2010). *Възможности за приложение на статистически методи при изучаване на фирмената конкурентоспособност*. София Авангард Прима, 2010