

MODELLING OF DOMESTIC AND FOREIGN VISITORS' BEHAVIOUR AT COMMERCIAL BANK WEBSITE DURING THE RECENT FINANCIAL CRISIS

Martin Drlík, Anna Pilková, Michal Munk, Peter Švec

Received: April 11, 2013

Abstract

DRLÍK MARTIN, PILKOVÁ ANNA, MUNK MICHAL, ŠVEC PETER: *Modelling of domestic and foreign visitors' behaviour at commercial bank website during the recent financial crisis*. Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis, 2013, LXI, No. 7, pp. 2065–2070

The paper focuses on modelling of commercial bank website visitors' behaviour. The authors analyse domestic and foreign market participants' interests in mandatory financial information disclosure of a commercial bank during the recent financial crisis and try to answer the question whether the purposes of Basel 2 regulations under the Pillar 3 – Market discipline, publishing financial information, have been fulfilled. The authors analyse bank website logs files using web log mining methods to better understand the rate of using of web pages, where mandatory financial information about Basel 2 is published. After data pre-processing the authors use association rule analysis to identify the association among content categories of the website. The results show that there is small interest in mandatory financial information disclosure by the commercial bank in general. The foreign website visitors take more concern in mandatory financial information disclosure, and they take less interest in general information about bank than domestic ones.

Basel 2 Pillar 3, information disclosure, user behaviour modelling, web log mining, website analytics

Market discipline (MD) in banking represents the mechanism via which market participants monitor and discipline excessive risk-taking behaviour by banks (Stephanou, 2010). The importance of MD has been formally stressed through the latest architecture of the banking regulatory frameworks Basel II and Basel III (BCBS, 2006; BCBS, 2011). Within these frameworks market discipline is one of the three pillars, it constitutes the Pillar 3.

The European institutions only require Pillar 3 information to be disclosed publicly. In general, the disclosures are to be made to the market for the benefit of the market. The banks publish the Pillar 3 information on their website, which is currently the best way how to make information easily accessible to market participants.

We try to answer the question if this information disclosure according to Basel II, Pillar 3 is interesting

for different kinds of stakeholders. We model the behaviour of users browsing bank website where information is publicly available. We analyse the market participants' interest in mandatory disclosure of financial information by means of methods of web log mining. Such study has been not carried out yet.

The structure of the paper is as follows. We summarize the current status of the research in the second chapter. We introduce used methods and data resources in the third chapter. The fourth chapter describes the experimental results in detail. Finally, we discuss the obtained results in the last chapter.

Related Work

There are numerous definitions of market discipline (MD) in the literature. We define

Market discipline in banking as the mechanism via which market participants monitor and discipline excessive risk-taking behaviour by banks (Stephanou, 2010).

The importance of MD has been formally stressed through the latest architecture of the banking regulatory frameworks Basel II and Basel III (BCBS, 2006; BCBS, 2011). Within these frameworks market discipline is one of the three pillars (Pillar 3), along with capital regulation (Pillar 1) and supervision (Pillar 2). The BCBS and EU legislation (CRD) only require Pillar 3 information to be disclosed publicly, they do not define how it should be done. Neither Basel Committee nor the European Union related legislation asks for standardised format of information disclosure and frequency is limited to minimum one time per year (European Parliament and European Council, 2006a; European Parliament and European Council, 2006b). The banks publish the Pillar 3 information on their websites, which is currently the best way to make information easily accessible to market participants.

The National Bank of Slovakia (NBS) issued two decrees (National Bank of Slovakia, 2010; National Bank of Slovakia, 2011) that standardize required Pillar 3 accounting, risk and other information and frequency is stated quarterly, semi-annually and annually.

Web log mining has received extensive attention because of its significant theoretical background and powerful application potential. Many web usage mining approaches and methods were surveyed in (Srivastava, 2000; Pierrakos, 2003; Tao and Murtagh, 2000). The last comprehensive surveys on web log mining have been done by Koutri *et al.* (Koutri, Avouiris and Daskalaki, 2005) and Kosala and Blockeel (Kosala and Blockeel, 2000).

Web log data are pre-processed in the first phase in order to identify users, sessions, page views, and clickstreams (Chitraa and Davamani, 2010; Munk, Kapusta and Švec, 2009; Munk, Kapusta and Švec, 2010). According to Cooley *et al.* (Cooley, Mobasher and Srivastava, 1999) data preparation phase consists of data cleaning, data integration, data transformation, path completion and data reduction (Munk and Drlík, 2011a; Munk and Drlík, 2011b).

As we have mentioned earlier, the main objective of this paper is to analyse part of the commercial bank website related to Pillar 3 and find out stakeholders' behaviour. Stakeholders of the banks, as one of key building blocks, have been studied from different perspectives. However, not so much directly as far as their behaviour is in relation to usage of Pillar 3 information. European Banking Authority (EBA) in its regular survey on Pillar 3 implementation interviewed only rating agencies, credit institutions and analysts (European Banking Authority, 2012). The problem is, that they are not typical representatives of the stakeholders of the commercial bank operating in CEE, owned by a foreign group.

According to Song *et al.* (Song, 2007) the behaviour pattern discovering is one of the web log mining sub-areas that focus on finding out typical flow models of user actions from recorded events. Predominantly used methods are discovering of association and sequence rules, segmentation (cluster analysis, analogy-based methods, etc.) as well as classification (decision rules, decision trees, Bayesian classification, etc.). Pattern discovery of web usage mining which is the outcome of the proposed methodology is discussed in detail in (Shun-Hua, Miao and Guo-Hai, 2010; Song, 2007) and (Nina, 2009).

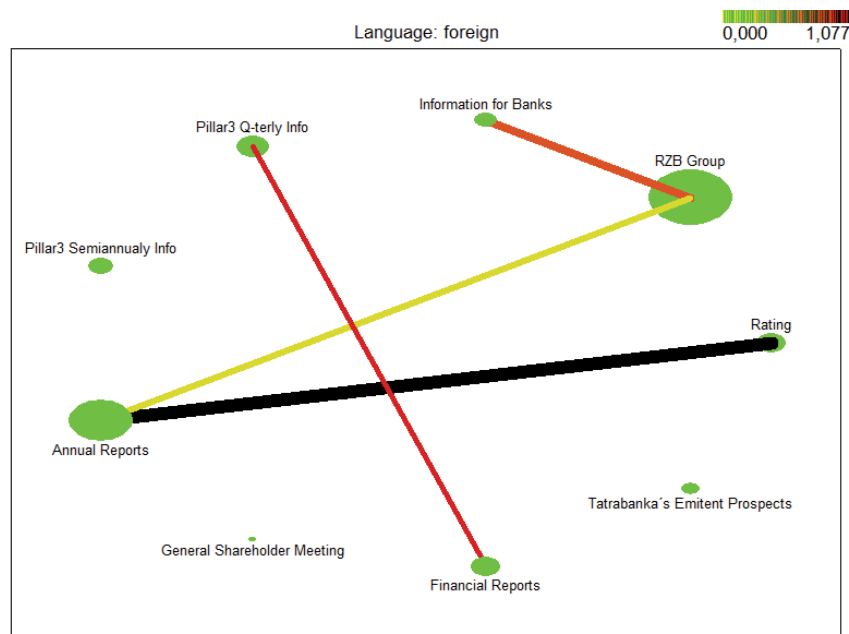
METHODS

The structure as well as character of data about web usage predetermines the application of specific methods for analysis – data modelling. Sometimes it is enough to slightly adapt the existing methods and procedures from the KDD area, another time it is necessary to radically change the steps of data pre-processing and analysis (Škorpil and Štastný, 2008).

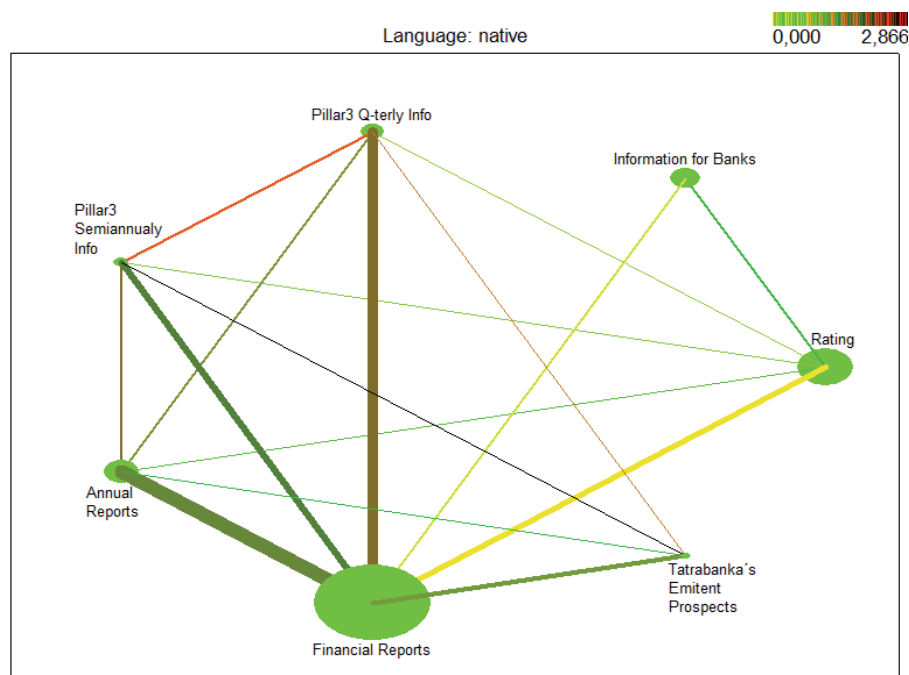
The users' behaviour patterns are most often sought in pre-processed data from the log file. For this purpose, a sequence rule analysis is used whose aim is sequence rules extraction. With the help of these rules we predict sequences of visit of various web parts by users. This method was derived from association rules and is an example of method taking into account specific features of the Internet. The difference between association and sequence rule analysis lies in the fact, that in case of association rule analysis we do not analyse sequences but transactions, that means that we do not include the time variable into the analysis (Munk, 2009). Similarly, as in the market basket analysis one transaction represents one purchase, in case of the web, one transaction is the set of pages visited by a user during one session, while we ignore the order in which the user visited individual web pages. Similarly, in case of virtual learning environments we do not investigate accesses to course pages but to course activities (Čápay *et al.*, 2011; Klocoková and Munk, 2011). It is important to be aware of this fact, mainly when interpreting the discovered association rules consisting of web pages on different hierarchical levels (Brečka *et al.*, 2011). The association rule analysis is mainly applied in portal visit rate analysis when we do not examine the accesses on portal pages but the broader categories of portal content.

RESULTS

We deal with the results of association rules analysis in more detail. This analysis represents the nonsequential approach to analysed data. We do not analyse sequences, but transactions, i.e. we do not regard the time variable into the analysis. The transaction thereafter represents the set of visited



1: Web graph: visualization of found rules for web pages in foreign languages



2: Web graph: visualization of found rules for web pages in Slovak language

web parts of *About Bank* content which is related to the Pillar 3.

Concerning data, we will consider of web parts of the commercial bank website related to the Pillar 3 (Category: *Pillar 3 related, Pillar 3 Disclosure Requirements*) which was visited by the stakeholder during one session as a one transaction.

We used statistical tool Sequence, Association and Link Analysis for identification of associations among web parts related to the Pillar 3. We extracted

association rules at first, then we calculated for these rules two variables: *support* and *lift*.

Web graph (Fig. 1) visualises found association rules. The size of a node represents *support* of this category item, line width represents support of the rule and line brightness stands for the *lift* of the rule. In other words, the *lift* represents the measure of interestingness and offers the most interesting results because it can be interpreted as how often times the categories of problems occur together than in the case if they were statistically insignificant.

Considering the web graph (Fig. 1) we can see that the most visited web parts of the portal belong to *RZB Group, Annual Reports, Pillar3 Q-terly Info and Financial Reports* (*support* > 10%). The combinations of these web parts in the identified sessions do not exist or exist with very low probability (*support* < 1%). We can see also that the web parts *Annual Reports* ==> *Rating* are more frequent together as apart in identified sessions (*lift* = 1.08). We can claim the same for web parts *Pillar3 Q-terly Info* ==> *Financial Reports* (*lift* = 1.03).

We can observe the greatest measure of interestingness (*lift*) in these cases, because the *lift* greater than one means that the selected pairs of web parts are more frequent together than apart in the set of web parts visited by the particular stakeholders. But, we have to be aware that the *lift* does not depend on the rule orientation. The measure of interestingness (*lift*) of the remaining found rules was less than one.

Considering the web graph (Fig. 2) we see that web parts *Financial Reports, Rating, Annual Reports, Information for Banks* and *Pillar3 Q-terly Info* (*support* > 10%) belong to the most visited parts of the portal. We found out that the combinations of parts *Annual Reports* ==> *Financial Reports* and *Pillar3 Q-terly Info* ==> *Financial Reports* have *support* > 10% at the same time.

The other interesting finding is that *Bank's Emitent Prospects* ==> *Pillar3 Semiannually Info, Pillar3 Semiannually Info* ==> *Pillar3 Q-terly Info, Annual Reports* ==> *Pillar3 Semiannually Info* and *Pillar3 Q-terly Info* ==> *Financial Reports* exist together more frequently than apart (*lift* > 1.6) in the identified sessions.

Another group of rules, *Pillar3 Q-terly Info* ==> *Bank's Emitent Prospects, Financial Reports* ==> *Pillar3 Semiannually Info, Financial Reports* ==> *Annual Reports, Annual Reports* ==> *Pillar3 Q-terly Info* and *Financial Reports* ==> *Bank's Emitent Prospects* attain also the *lift* > 1.1. The measure of interestingness (*lift*) of the remaining found rules was less than one.

DISCUSSION

Studied bank is a typical representative for its peer group in the Slovak market: it is a universal commercial bank. The bank has multiplicity of stakeholders. Among them key role play insured and uninsured depositors, mortgage bond holders, subordinated debtholders and minority shareholders.

Stakeholders of the bank presented very low interest in Pillar 3 required information (1.22%) and information related to Pillar 3 (5.05%) in 2008 and 2009. The highest interest, almost 94%, was on general information on the bank.

We have possible explanations for this kind of behaviour. First of all, we assume there was almost no impact of the crisis on the domestic banks as they were in a very good shape as far as capital and liquidity position are. Depositors as key stakeholders did not have serious reason to monitor

them. Secondly, deposit insurance fund coverage and tough competition among commercial banks operating in the market might create sufficient safety net that can be a reason not only for low market discipline but also for moral hazard in the stakeholders' behaviour. In the same time numerous minority shareholders have a lack of incentives to control the bank and prefer to rely on control executed by majority shareholder. Local regulator (NBS) is demanding and has a very good reputation. Finally, level of sophistication of the presented information (that is more or less standardized by Decree issued by the local regulator) is above of financial literacy and expertise of stakeholders. In this respect also the bank's complexity of operations have a significant impact on level of sophistication provided information (Rábová, 2010).

Foreign stakeholders vs. domestic ones presented higher interest in both Pillar 3 disclosure required information (3.15% vs. 0.91%) and Pillar 3 related information (11.82% vs. 3.94%). It is also evident from the analysis that both groups of stakeholders were more interested in information related to Pillar 3 than Pillar 3 disclosure required information. And above all they were more interested in this type of information provided quarterly than semi-annually. Further analysis discovered out that the highest frequency of visiting Pillar 3 quarterly disclosure required information were in the fourth quarter.

The highest interest of foreign stakeholders was at *RZB Group* information and usually in combination with *Information for Banks*, what indirectly confirms that these stakeholders might be analysts from partners' banks, rating agencies etc. The second highest frequency (the most significant) had *Annual reports* category which is tightly related to visit of *Rating* part. Those who visit *Annual Reports* visit also *RZB Group* information (and vice versa) but less frequently. Foreign stakeholders visit maximum two web parts in general.

The highest frequency of domestic stakeholders' visits was at *Financial Reports* part, *Rating, Annual Report* and *Information for Banks*. In comparison to foreign shareholders there are more subsequent visits from one web part than foreign shareholders do. Maximum is six visits from *Financial Reports* and the most often the highest occurrence is between *Financial Reports* and *Annual Reports, Financial Reports* and *Pillar 3-Quarterly info*.

The outcomes of web information analysis and modelling are patterns of stakeholders' behaviour. There is a difference in the pattern of behaviour between domestic stakeholders and foreign ones. Domestic stakeholders behaviour pattern looks like messy one – stakeholders are fishing after information without knowing which information they are looking for. Foreign stakeholders behaviour pattern are more “focus oriented”. It looks like they know which information they are looking for. This different behaviour can be explained by higher literacy of foreign stakeholders (they can be analysts of rating agencies, participating banks etc.).

We found out that Pillar 3 disclosure is fully formally implemented in line with BCBS and EBA requirements. However, lack of the commercial bank's stakeholders interest on the published information and their pattern of behaviour addressed problems and raised further research questions. These problems can belong to one of the following two groups: stakeholders have not been able to cope with Pillar 3 disclosed information or did not have any incentives. Lack of financial literacy of stakeholders, their information needs and mandatory standardized format of disclosed information by the regulator is the main areas of

the first group of problems. Financial safety net in the country, transparency and moral hazard in stakeholder behaviour belong to the second group of problems. In this context it is also important to study legal regulatory framework that determine content and frequency of presented information in Slovakia: whether it is inevitable to standardise format and also whether annual full set of Pillar 3 disclosure would not be sufficient instead of quarterly and semi-annually disclosed sets. Change of content and frequency might be beneficial for stakeholders and have a positive impact from time consuming and costs points of view for banks.

SUMMARY

The paper focuses on modelling of commercial bank website visitors' behaviour. The authors analyse domestic and foreign market participants' interests in mandatory financial information disclosure of a commercial bank during the recent financial crisis and try to answer the question whether the purposes of Basel 2 regulations under the Pillar 3 – Market discipline, publishing financial information, have been fulfilled. The authors analyse bank website logs files using web log mining methods to better understand the rate of using of web pages, where mandatory financial information about Basel 2 is published. After data pre-processing the authors use association rule analysis to identify the association among content categories of the website. The results show that there is small interest in mandatory financial information disclosure by the commercial bank in general. The foreign website visitors take more concern in mandatory financial information disclosure, and they take less interest in general information about bank than domestic ones.

Acknowledgement

This paper is supported by the project VEGA 1/0392/13 "Modelling of Stakeholders' Behaviour in Commercial Bank during the Recent Financial Crisis and Expectations of Basel Regulations under Pillar 3- Market Discipline".

REFERENCES

- BCBS, 2006: *International convergence of capital measurement and capital standards: A revised framework – comprehensive version*. Bank for International Settlements. Available online: <http://www.bis.org/publ/bcbs128.htm>.
- BCBS, 2011: *Basel III: A global regulatory framework for more resilient banks and banking systems – revised version*. Bank for International Settlements. Available online: <http://www.bis.org/publ/bcbs189.htm>.
- BREČKA, P., MAGDIN, M., KOPRDA, Š., 2011: Two-state Regulation in MATLAB for the Comparison of Some Parameters (Damage, Power Consumption) by PSD Regulation. *Advances in Computer, Communication, Control & Automation, Lecture Notes in Electrical Engineering*. Berlin: Springer-Verlag, 369–376.
- CÁPAY, M., BALOGH, Z., BOLEDOVIČOVÁ, M., MESÁROŠOVÁ, M., 2011: Interpretation of questionnaire survey results in comparison with usage analysis in e-learning system for healthcare. *Communications in Computer and Information Science*, 167, 2: 504–516. ISSN 1865-0929.
- European Banking Authority, 2012: *Follow-up review of banks' transparency in their 2011 Pillar 3 reports*. Available online: <http://www.eba.europa.eu/-/eba-publishes-follow-up-review-of-banks-transparency-in-their-2011-pillar-3-reports>.
- European Parliament and European Council, 2006a: Directive 2006/48/EC of 14 June 2006 relating to the taking up and pursuit of the business of credit institutions. *Official Journal of the European Union* 2006, L 177/1-197. Available online: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:2006L0048:20100330:EN:PDF>.
- European Parliament and European Council, 2006b: Directive 2006/49/EC of the European Parliament and of the Council of 14 June 2006 on the capital adequacy of investment firms and credit institutions. *Official Journal of the European Union* 2006, L 177/201. Available online: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:177:0201:0201:EN:PDF>.
- KLOCOKOVÁ, D., 2011: Integration of heuristics elements in the web-based learning environment: Experimental evaluation and usage analysis. *Procedia Social and Behavioral Sciences*, 2011, 15: 1010–1014. ISSN 1877-0428.
- KLOCOKOVÁ, D., MUNK, M., 2011: Usage analysis in the web-based distance learning environment in a foreign language education: Case study. *Procedia – Social and Behavioral Sciences*, 2011, 15: 993–997. ISSN 1877-0428.

- KOSALA, R., BLOCKEEL, H., 2000: Web mining research: a survey. *ACM SIGKDD Explorations Newsletter*, 1, 2: 1–15. ISSN 1931-0145.
- KOUTRI, M., AVOURIS, N., DASKALAKI, S., 2005: *A survey of web-usage mining: techniques for building web-based adaptive hypermedia systems*. Adaptable and Adaptive Hypermedia Systems. Pennsylvania, USA: IGI Global. 125–149. ISBN 1-59140-536-X.
- MUNK, M., 2009: Objavovanie znalostí na základe používania webu – metódy a aplikácie. *Forum Statisticum Slovaca*, 5, 1: 65–72. ISSN 1336-7420.
- MUNK, M., KAPUSTA, J., ŠVEC, P., 2009: Data preprocessing dependency for web usage mining based on sequence rule analysis. *IADIS Multi Conference on Computer Science and Information Systems, MCCSIS 2009*. Algarve, Portugal: IADIS Press, 179–181. ISBN 978-972-8924-97-3.
- MUNK, M., KAPUSTA, J., ŠVEC, P., 2010: Data preprocessing evaluation for web log mining: reconstruction of activities of a web visitor, *ICCS 2010, Procedia – Computer Sciences*, 2010, 1: 2273–2280. ISSN 1877-0509.
- MUNK, M., DRLÍK, M., 2011a: Impact of different pre-processing tasks on effective identification of users' behavioral patterns in web-based educational system. *Procedia Computer Science*, 2011, 4: 1640–1649. ISSN 1877-0509.
- MUNK, M., DRLÍK, M., 2011b: Influence of different session timeouts thresholds on results of sequence rule analysis in educational data mining. *Communications in Computer and Information Science*, 166, 1, part 1: 60–74. ISSN 1865-0929.
- National Bank of Slovakia, 2010: *Decree of the National Bank of Slovakia of 21 September 2010 No. 15/2010 on disclosures by banks and branches of foreign banks*. Available online: http://www.nbs.sk/_img/Documents/_Legislativa/_FullWordingsOther/EN_O_15_2010.pdf.
- National Bank of Slovakia, 2011: *Decree of the National Bank of Slovakia of 29 November 2011 No. 19/2011 amending decree of national bank of slovakia No 15/2010 on disclosures by banks and branches of foreign banks*. Available online: http://www.nbs.sk/_img/Documents/_Legislativa/_FullWordingsOther/EN_O_19_2011.pdf.
- NINA, S. P., RAHMAN, M., BHUIYAN, K.I., AHMED, K.E.U., 2009: Pattern discovery of web usage mining. *International Conference on Computer Technology and Development. ICCTD '09*. Washington DC: IEEE Computer Society, 499–503. ISBN 978-0-7695-3892-1.
- PIERRAKOS, D. et al., 2003: Web usage mining as a tool for personalization: a survey. *User Modeling and User-Adapted Interaction*, Hingham: Kluwer Academic Publishers, 13, 4: 311–372. ISSN 0924-1868.
- RÁBOVÁ, I., 2010: Formalizace podnikových pravidel pro informační systémy. *Acta Universitatis agriculturae et silviculturae Mendelianae Brunensis*, 58, 6: 389–399. ISSN 1211-8516.
- TAN, S-H., CHAN, M., YANG, G-H., 2010: User behavior mining on large scale web log data. *Apperceiving Computing and Intelligence Analysis (ICACIA), 2010 International Conference*. Chengdu, China: University of Electronics Science, 60–63. ISBN 978-1-4244-8025-8.
- SONG, J., LUO, T., GAO, F., 2007: The data preprocessing of behavior pattern discovering in collaboration environment. *Web Intelligence and Intelligent Agent Technology Workshops, 2007 IEEE/WIC/ACM International Conferences*. USA: Silicon Valley, CA, 521–525. ISBN 0-7695-3028-1.
- SRIVASTAVA, J., COOLEY, R., DESHPANDE, M., TAN, P-N., 2000: Web usage mining: discovery and applications of usage patterns from Web data. *SIGKDD ACM SIGKDD Explorations Newslette*. 1, 2: 12–23. ISSN 1931-0145.
- STEPHANOU, C., 2010: *Rethinking market discipline in banking: lessons from the financial crisis*. Policy research working paper 5227. The World Bank. Available online: <http://elibrary.worldbank.org/content/workingpaper/10.1596/1813-9450-5227>.
- ŠKORPIL, V., ŠTASTNÝ, J., 2008: Comparison of learning algorithms. *24th Biennial Symposium on Communications*, Kingston, Canada: Queens University Kingston, 231–234. ISBN 978-1-4244-1946-3.
- TAO, F., MURTAGH, F., 2000: Towards knowledge discovery from www log data. *Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'00)*, IEEE Computer Society Washington, DC, USA: IEEE Computer Society, 302. ISBN 0-7695-0540-6.

Address

Mgr. Martin Drlík, PhD., Department of Computer Science, Constantine the Philosopher University in Nitra, Tr. A. Hlinku 1, 949 74 Nitra, Slovak Republic, doc. RNDr. Michal Munk, PhD., Department of Computer Science, Constantine the Philosopher University in Nitra, Tr. A. Hlinku 1, 949 74 Nitra, Slovak Republic, doc. Ing. Anna Pilková, CSc., MBA, Department of Strategy and Entrepreneurship, Faculty of Management of CU, Odbojarov 10, P.O.BOX 95, 820 05 Bratislava 25, Slovak Republic, PaedDr. Peter Švec, Ph.D., Department of Computer Science, Constantine the Philosopher University in Nitra, Tr. A. Hlinku 1, 949 74 Nitra, Slovak Republic, e-mail: mdrlik@ukf.sk, mmunk@ukf.sk, anna.pilkova@fm.uniba.sk, psvec@ukf.sk