

# The Use of Case-Based Reasoning for the Monitoring of Financial Fraud Transactions

Stelios Kapetanakis<sup>1</sup>, Georgios Samakovitis<sup>2</sup>, P.V.G Buddhika Dinesh Gunasekera<sup>2</sup>, Miltos Petridis<sup>1</sup>

<sup>1</sup> School of Computing, Engineering and Mathematics, University of Brighton, Moulsecoomb Campus, Lewes road, Brighton BN2 4GJ, UK,  
email: {s.kapetanakis, m.petridis}@brighton.ac.uk

<sup>2</sup> School of Computing and Mathematical Sciences, University of Greenwich, Maritime Greenwich Campus, Old Royal Naval College, Park Row, Greenwich, London SE10 9LS, UK,  
email: {g.samakovitis, gp102}@gre.ac.uk

**Abstract.** Financial transaction fraud constitutes an acute problem domain for detection and early warning systems. Despite that different branches of AI have been addressing the problem since the late 90s, CBR approaches have seldom been applied. This paper provides a proof of concept approach and experimental investigation for the use of a CBR Intelligent Monitoring System for detecting abnormal patterns in financial transaction flows. The representation of workflow related knowledge in this research using graphs is explained. The workflow process is orchestrated by a software system using BPEL technologies within a service-oriented architecture. Workflow cases are represented in terms of events and their corresponding temporal relationships. The matching and CBR retrieval mechanisms used in this research are explained and a simple evaluation of the approach is provided using simulation data. Further work on the system and the extension to a full intelligent monitoring and process optimisation system is finally presented.

**Keywords:** Abnormal financial transactions, Fraud detection, Case-based Reasoning, Business Workflows, Temporal Reasoning, Graph Similarity.

## 1 Introduction

Financial Fraud Detection (henceforth FFD) has been a prevalent topic in Expert Systems and Knowledge Discovery research for more than a decade [1, 2, 3, 4, 5]. Specialist areas such as Data Mining and Artificial Intelligence, among others, have contributed approaches to support reactive and proactive fraud identification and deterrence [6, 7, 8]. The fields of Financial Statement Fraud, Insurance Fraud and Credit Card Fraud, ([5, 7, 9, 10, 11, 12, 13, 14, 15, 16, 18, 17], among others) are the ones having drawn most of the attention of related literature, while, interestingly, bank-transfer transaction fraud per se has seldom been addressed.

Research in the field of FFD has focused on data mining and other classification techniques (see for instance [19, 16, 20], as also outlined in [5]), however, there is strikingly thin evidence of use of CBR in FFD [4, 21, 22]. Our literature review and analysis suggests that, despite its pronounced superior performance in systems [4, 41, 42, 43] possible reasons for that scarcity may well include (i) the focus of the relevant literature on optimising existing approaches; (ii) the lack of maturity of CBR research with reference to the transactions application scope; and (iii) the established view of the FFD problem as one seeking precision optimisation, rather than seeking new ways of identifying and representing activity patterns. Significant scope therefore exists for investigating the performance of CBR techniques for detecting Financial Fraud.

In addition, established views of the problem, as provided in much of the aforementioned literature, address Financial Fraud activity at the individual transaction level and use variants of clustering and classification methodologies for activity profiling and labelling of suspicious instances [16, 17, 19, 22]. As a result, limitations are introduced in assessing patterns of transactions, thus missing out on the opportunity to track and monitor transaction sequences that may be indicative of other suspicious activity relevant, for instance, to money laundering, insider trading or other illicit actions involving FFD. For that reason, the necessity arises for addressing financial transaction streams as assessment units. Streams are only reportedly used in [29, 30] where, however, the focus is placed on modelling as opposed to detection. Evidently, fraud involving financial transactions other than Financial Statement Fraud or Insurance Fraud (such as stock-trading, money transfer or remittance) appear to be suitable for treatment as sequences or flows. This provides strong motivation for exploring the usefulness of Workflow CBR approaches in FFD.

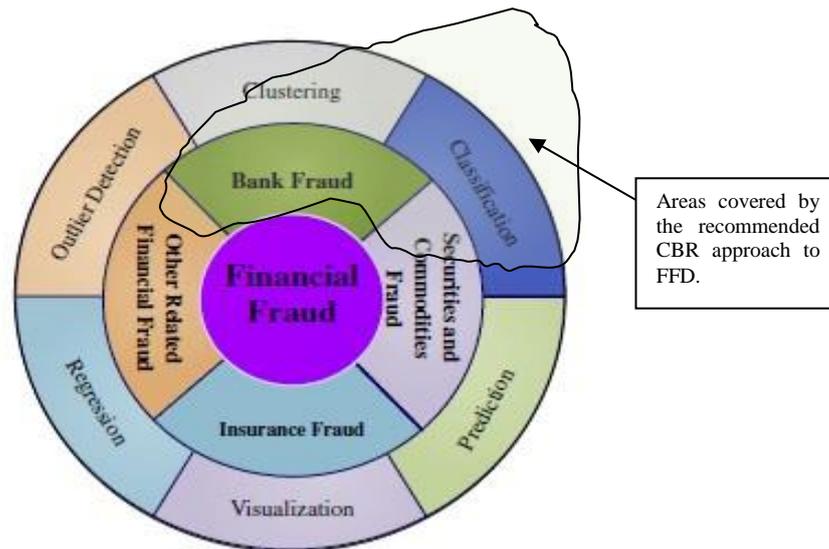
This paper will present a workflow approach to FFD with the use of Case-based Reasoning techniques. The contribution of this paper is to demonstrate usability of CBR workflow systems in FFD by means of presenting a proof-of-concept approach that operates within a proposed FFD framework. We first discuss the business domain and outline the literature scope and fields that have dealt with approaches to FFD. We then surface the relevance and significance of using a CBR Workflow approach to at least certain types of fraud detection. The proposed system is then outlined in order to explain the adopted approach. The paper then follows with presenting the model for Transaction Representation and offers the set of Similarity Measures that the model will use. A discussion based on a set of preliminary experiments is finally provided.

## **2 The Business Domain: Intelligent Approaches to Financial Fraud Detection**

Financial Fraud accounted for annual losses in the UK alone of £38bn (2011) rising to £73bn in 2012 [23, 24, 25], while it reached \$2,75trn in 2009 globally [26], averaging 4.5% of total expenditure [27]. Because of its sheer size, Financial Fraud offers significant scope for Decision Support optimisation, promising substantial economic benefits. Because of the nature of electronic money transfer, accurate labelling of financial transactions as genuine or fraudulent is paramount in ensuring customer trust, especially in light of the prevalence of user-driven electronic banking [28]. On that account, misidentified fraud instances (false positives) are equally detrimental to customer trust as are unidentified instances of original fraud (false negatives). This adds significantly to the precision requirements of FFD processes and systems, and partly hints to the current reluctance to adopt unified industry-wide approaches, as highlighted below.

Despite its significant and widely recognised economic and policy-related importance [23, 24, 25, 26], no explicit FFD framework is widely established. Transaction-handling entities (such as banks and financial services providers) address the problem at the firm level, while authorities and independent bodies (such as the FSA, SOCA, BAI, SFO, NFA among others) only address cases of large-impact financial fraud, typically linked to wider criminal activity [23, 25]. While certain individual banks and Financial Services providers boast bespoke rule-based monitoring systems to address the problem [29, 30], the main approach to fraud involving small amounts is mainly addressed through fraud protection insurance. The situation is therefore indicative of the absence of robust and reliable Decision Support that will facilitate fraud identification across the board.

In their literature investigation, Ngai et al. (5) offer a useful conceptual framework for a literature classification of FFD types and existing approaches in Data Mining & Artificial Intelligence to address Financial Fraud. In their analysis, they identify four different broad classes of financial fraud, represented in the inner circle of Figure 1. They also distinguish between Clustering, Classification, Prediction, Outlier Detection, Regression and Visualisation approaches (outer circle) as these make their presence in the literature on FFD.



**Figure 1: Data Mining Techniques applied in Financial Fraud Detection (from Ngai et. al. (2011)): the figure is adapted to reflect the remit and target scope of the Workflow CBR approach.**

Figure 1 is adapted to represent the FF classes and technique categories that our Workflow CBR approach covers. The highlighted area covers a Classification / Clustering approach to mainly Bank Fraud, partially spanning to ‘Other’ Financial Fraud types.

The logic of the approach presented in this paper, named CBR-FTIMS (CBR - Financial Transactions Intelligent Monitoring System) is introduced in the next section, followed by an analysis of the similarity measures used. The paper then addresses our proof-of-concept experiment and provides a preliminary evaluation of its results.

## 2.1 CBR Financial Transactions Intelligent Monitoring System

For the needs of the identification of fraudulent behaviours among financial transactions, a software prototype has been developed. Its aim is to notify the senior workflow stakeholders regarding suspicious behaviours as well as provide past available experience to justify any conducted case judgement. The software (CBR - FTIMS) has been designed to work on top of existing infrastructures, relating to actual knowledge as mined from their available data repositories. The latter could comprise transactions log files, databases with past information, etc. Since this knowledge resides in the format of sequential transactions, the software should be able to extract it. Additionally it should be able to represent their temporal flow, find the similarities among the transaction sequences and be able to present the available knowledge regarding an investigated case.

CBR-FTIMS has access to the data available in an existing repository and uses them as a knowledge case base for future reference. The transaction data along with their related classification can be used to solidify a possible suggestion for an investigation of a new case with unknown status. This status suggestion can simply be either “fraudulent” or not and can be added to an existing case base upon authorisation from a human expert.

The implemented prototype was developed to deal with transaction data and apply monitoring to the sequences of transactions in order investigate what is the current status. This monitoring relates to an “a posteriori” approach, relating to an investigated case and the available past cases.

At the current stage the system was developed to test whether the CBR approach is effective within the concept of the financial transactions’ fraud identification. Therefore, a simple evaluation was designed and performed in order to test its feasibility in the

operational domain of financial transactions. However, since this has to be applied over a number of existing financial systems, the prototype was not designed to deal with a specific system or an existing application. Its underlying philosophy was: i) to be able to accommodate workflows of an imported financial business processes annotated in WS-BPEL [35] format and represented with XPDL [36]; ii) to be able to understand the executed transactions from these business processes; iii) to establish workflow monitoring by applying CBR and using past available experience for classification.

In CBR-FTIMS the CBR cycle is being used as defined from Aamodt & Plaza [37] and similarity measures are being applied in order to find the nearest neighbours to an investigated case. The similarity measures are based on a graph representation of the available temporal knowledge and are used for classifying a case based on its neighbours. The following section refers in more details to the case representation in the prototype.

### 3 Case Representation and Similarity Measures

In CBR-FTIMS, the business process is being mapped using BPEL technologies, whereas UML is being used to define the available workflow actions in the business process.

When a workflow is being executed, the transaction details are usually being recorded in a log. The log can then be processed in order to identify the events that took place and reconstruct the actual flow of transactions.

By reverse-engineering the transactions available in a financial system's log, there is a considerable potential for identifying possible fraud attempts, safe transactions or incomplete ones due to a number of reasons (system failures, authentication errors, etc.). A system that deals with representation of transactions should be able to present their temporal sequences as well as their relations regarding key attributes. For financial transactions these have to do with the unique identifiers for bank accounts, the names of the account holders, etc.

For the needs of the transaction notation and the case construction, a temporal representation has been used in CBR-FTIMS. For the representation of the temporal sequences the general time theory of Ma & Knight [34] was used because of its applicability in workflow domains [38], its effectiveness in representing workflow executions [33] and its efficiency in monitoring and diagnosis [33, 38].

For the addressed domain the transaction representation has been conducted using a number of attributes (Sender IBAN, Receiver IBAN, Transaction Timestamp, etc.). These are extracted from the original transaction. The transaction format could vary in terms of the international formats (for instance ISO 8583, AS2805, etc.).

For the need of the similarity measures attention has been drawn to the persistence format of the transactions which, although represented in terms of workflows, are formatted in XML. Dorneles et al. [40] suggest that when coming to a collection of XML elements, similarity measures can be calculated by either using metrics for atomic values (MAVs) or metrics over complex values (MCVs). For the purpose of the evaluation of the suggested proof of concept we use a weighting algorithm based on the concept of MCVs, taking into account the variance of attributes of similar type that occur in each instance of transactions.

The algorithm, when used to estimate the similarity among two transaction instances, takes into account the available attributes after applying a certain weight-filtering on them. The filtering is based on the importance of the attributes. This algorithm has been used to calculate the similarity among individual transaction instances applying means of *jackardSim* upon Strings, *DateSim* across Dates, *diffNumberSim* upon Numbers, etc. (for instance see [40]).

The similarity upon collections of attributes can be represented mathematically as:

$$collectionSim(\varepsilon_g, \varepsilon_{g'}) = \frac{\sum_{\varepsilon_g^i, \eta = \varepsilon_{g'}^j, \eta \text{ and } i=j} (sim(\varepsilon_g^i, \varepsilon_{g'}^j))}{\max(m, n)}$$

where  $\varepsilon_g$  is a node in  $G$  and  $\varepsilon_{g'}$  is a node in  $G'$ ,  $n$  and  $m$  are the children of  $\varepsilon_g$  and  $\varepsilon_{g'}$ , respectively,  $\varepsilon_g.\text{score} = \text{collectionSim}(\varepsilon_g, \varepsilon_{g'})$  and  $(1 \leq i \leq n), (1 \leq j \leq m)$ . The  $\text{collectionSim}$  asserts that  $\varepsilon_g^i$  in  $G$  is compared with  $\varepsilon_{g'}^j$  in  $G'$  if and only if  $i = j$ . In this way it is assured that the similarity metrics are applied to the same fields.

## 4 Experiments & results (evaluation)

In order to evaluate the suitability of the approach, an experiment was designed and conducted using the proprietary software. For the needs of the feasibility evaluation a simplified workflow approach was used, constituted from simulated transaction data.

Financial institutions seldom are sharing their available transaction repositories since there are a number of reasons, including ethical limitations regarding the handling of the data. Therefore, for the evaluation of the suggested approach a number of simulated transactions was used in order to be able to construct a case base for the experimentation overall.

In order to evaluate the feasibility of the adopted approach, a selection of 600 simulated transactions were used to constitute its case base. The transactions were classified based on their status in terms of fraud. This could fit into three categories: “safe”, “fraudulent” or “cannot classify with certainty” in the following numbers: 520 as safe, 50 as fraudulent and 30 as cannot classify with certainty. The classification was based on a number of indicative fraud attributes such as the volume of the amount and the frequency of transactions, among others.

A transaction pattern can be presented using the following simplified representation:

$T_{\text{identification transaction number}} = (\text{Sender IBAN}, \text{Receiver IBAN}, \text{timestamp}, \text{transaction amount}, \text{transaction indication})$

The transaction indication is an artificial field that comprises the transaction references and was used to give a pre-classification to the available transactions.

When estimating the similarity among transactions the similarity algorithm, as presented in section 3, can be applied as follows in the case of two transactions  $T_1, T_2$  which share the same Receiver IBAN:

$T_1 = (\text{AD12 0001 2030 2003 5110 0900}, \text{AT61 1409 3002 3457 3201}, 28 \text{ April } 2012, 2583, \text{salary payment})$

$T_2 = (\text{AD12 0001 2030 2003 5110 0009}, \text{AT61 1409 3002 3457 3201}, 1 \text{ May } 2012, 1500, \text{rent})$

For the similarity calculation the  $\text{collectionSim}$  is being applied to both transactions having a  $\text{jaccardSim}(\varepsilon_p^1, \varepsilon_d^1) + \text{jaccardSim}(\varepsilon_p^2, \varepsilon_d^2) + \text{DateSim}(\varepsilon_p^3, \varepsilon_d^3) + \text{diffNumberSim}(\varepsilon_p^4, \varepsilon_d^4) + \text{jaccardSim}(\varepsilon_p^5, \varepsilon_d^5)$ . For the above example the  $\max(m, n) = 5$ . Therefore the final result will be the score of the above applied measures divided by 5.

For sequences of transactions the algorithm is applied sequentially on sums of individual MCVs. In that case a sequence  $G$  is compared with a  $G'$  following the above example in a larger scale.

For the needs of the evaluation experiment the 600 cases were split randomly into one case base of 560 cases and a target set of 40 cases. Using the KNN algorithm with  $K=3$ , the 3 nearest neighbours of each case were found, as shown in Figure 2 below were used to classify the transactions as “safe”, “fraudulent” or “cannot classify with certainty”.  $K$  was set to 3 due to the volume of the case sample. Previous CBR experiments that used larger values for  $K$  did not present any significantly different results [38].

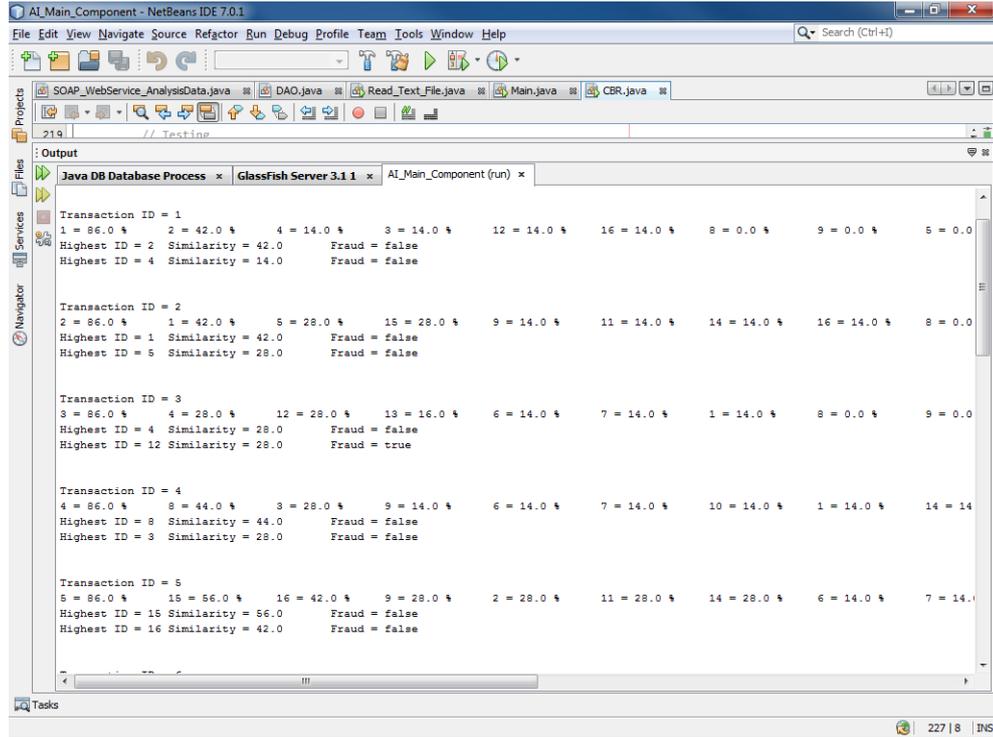


Figure 2: Transaction neighbours

The produced results were compared afterwards with the already known classification of the cases to evaluate the effectiveness of the proposed approach. For better precision the experiments ran for 10 times and the final classification results were averaged over the runs overall. The selection of 10 iterations was regarded sufficient since after this number there is no differentiation in the produced results. The following table shows the results of the evaluation runs.

	Number of cases		Percentage (%)	
Cases correctly classified	25	20 as "safe"	50	62.5
		4 as "fraudulent"	10	
		1 as "cannot classify with certainty"	2.5	
Missed positives	4	0 as "safe"	0	10
		1 as "fraudulent"	2.5	
		3 as "cannot classify with certainty"	7.5	
False positives	11	2 as "safe"	5	27.5
		5 as "fraudulent"	12.5	
		4 as "cannot classify with certainty"	10	
Total	40	40	100	100

Table 1: Case classification results

As it can be seen from Table 1 by applying CBR on the available case base there was a successful identification of cases that fit into the available patterns. From this first approach the system seems to overlook a small percentage of “fraudulent” transactions and identifies several cases as falsely “fraudulent” or “cannot classify with certainty”. A reason for that could be the shortage in terms of prevailing patterns since the similarity measures were focused on a quantitative approach in terms of calculating the similarity among workflow instances. Additionally the graphs created for this experiment were of a specific length and direction only, reducing considerably the calculation time but failing to reveal hidden patterns or any significant information from the existing data.

## 5 Conclusions

Evidence from academic research on FFD and observation of established industry practice demonstrate a large disconnect between academic approaches to intelligent monitoring for FFD, as recorded in the literature, and actual practice, as applied by practitioners and occasionally disclosed in the public domain. While this observation is consistent with past research on technologies in the UK banking sector [28, 39], it is indicative of the absence of unified industry-wide approaches to treating Financial Fraud, at both technology and policy levels. This, in turn, calls for significantly increasing integration between practitioner and research approaches to the FFD problem.

Our literature review suggests that CBR approaches to FFD rarely appear in the academic fields of Expert Systems, Data Mining and Artificial Intelligence. Similarly, workflow treatment of financial transactions is scarce. This provides significant scope for further work on CBR Workflow approaches to FFD.

The present research provided a proof-of-concept CBR Financial Transactions Intelligent Monitoring System (named CBR-FTIMS), aiming to demonstrating the use of a CBR workflow approach in identifying abnormal financial transactions.

Results of our experiments suggest that CBR can be applied successfully over a case base of transactions, where classification has been applied in advance, and contribute to the ranking of an unknown cluster of cases. Additionally it has been shown that by applying simplified CBR the number of false positives is high, something that has to be considered in future work on the monitoring process.

In line with the conclusions of this paper, further work calls for the extension of the CBR-FTIMS research and architecture development to accommodate real-time monitoring and proactive treatment of abnormal financial transactions. A number of areas have also been pointed out for future work in terms of the architectural approach in existing financial systems as well as the presentation of the monitoring results to financial stakeholders, not necessarily in the area of artificial intelligence.

This research has shown that there can be preliminary application of fraud detection measures to financial transactions with the use of CBR and workflow representation to classify cases of unknown status. More experiments in the area can throw light in terms of the prevailing techniques and the optimisation margins that should be applied to maximize classification efficiency.

## References

- [1]. Mieke, J., van der Werf, J.M., Lybaert, N., Vanhoof, K.: A business process mining application for internal transaction fraud mitigation, *Expert Systems with Applications*, Volume 38, Issue 10, pp 13351-13359 (2011)
- [2]. Cerullo, M.J. and Cerullo, V.: Using neural networks to predict financial reporting fraud: Part 2, *Computer Fraud & Security*, Volume, Issue 6, pp.14-17, (1999)
- [3]. Galliers, K.: AI techniques to counter financial fraud, *Computers & Security*, Volume 14, Issue 5, Page 426 (1995)
- [4]. Wheeler, R. and Aitken, S.: Multiple algorithms for fraud detection, *Knowledge-Based Systems*, Volume 13, Issues 2–3, pp.93-99 (2000)

- [5]. Ngai, E.W.T., Hu, Y., Wong, Y.H. Chen, Y. Sun,X.: The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature, *Decision Support Systems*, Volume 50, Issue 3, pp. 559-569 (2011)
- [6]. Ravisankar, P., Ravi, V., Raghava Rao, G. Bose, I.: Detection of financial statement fraud and feature selection using data mining techniques, *Decision Support Systems*, Volume 50, Issue 2, pp. 491-500 (2011)
- [7]. Jha, S., Guillen, M. Westland, J.C.: Employing transaction aggregation strategy to detect credit card fraud, *Expert Systems with Applications*, Volume 39, Issue 16, pp. 12650-12657 (2012)
- [8]. Debreceeny, R.S. Gray, G.L.: Data mining journal entries for fraud detection: An exploratory study, *International Journal of Accounting Information Systems*, Volume 11, Issue 3, pp. 157-181 (2010)
- [9]. Panigrahi, S., Kundu, A., Sural, S., Majumdar, A.K.: Credit card fraud detection: A fusion approach using Dempster–Shafer theory and Bayesian learning, *Information Fusion*, Volume 10, Issue 4, pp. 354-363 (2009)
- [10].Sanchez, D., Vila, M.A., Cerda, L., Serrano, J.M.: Association rules applied to credit card fraud detection, *Expert Systems with Applications*, Volume 36, Issue 2, Part 2, pp. 3630-3640 (2009)
- [11].Quah, J.T.S., Sriganesh, M.: Real-time credit card fraud detection using computational intelligence, *Expert Systems with Applications*, Volume 35, Issue 4, pp. 1721-1732 (2008)
- [12].Leonard, K.J.: Detecting credit card fraud using expert systems, *Computers & Industrial Engineering*, Volume 25, Issues 1–4, pp. 103-106 (1993)
- [13].Kirkos, E., Spathis, C., Manolopoulos, Y.: Data mining techniques for the detection of fraudulent financial statements, *Expert Systems with Applications* Volume 32, Issue 4, pp. 995–1003 (2007).
- [14].Rezaee, Z.: *Financial Statement Fraud: Prevention and Detection*: John Wiley and Sons, Inc., New York (2002) ISBN 0-471-09216-9
- [15].Zhou, W., Kapoor, G.: Detecting evolutionary financial statement fraud, *Decision Support Systems*, Volume 50, Issue 3, pp. 570-575 (2011)
- [16].Ravisankar P., Ravi, V. Raghava Rao, G. Bose, I.: Detection of financial statement fraud and feature selection using data mining techniques, *Decision Support Systems*, Volume 50, Issue 2, pp. 491-500 (2011)
- [17].Šubelj, L., Furlan, S. Bajec, M.: An expert system for detecting automobile insurance fraud using social network analysis, *Expert Systems with Applications*, Volume 38, Issue 1, pp. 1039-1052 (2011)
- [18].Derrig, R.A.: Insurance fraud, *The Journal of Risk and Insurance* Volume 69, Issue 3, pp. 271–287 (2002)
- [19].Glancy, F.H., Yadav, S.B.: A computational model for financial reporting fraud detection, *Decision Support Systems*, Volume 50, Issue 3, pp. 595-601 (2011)
- [20].Hand, D.J., Crowder, M.J.: Overcoming selectivity bias in evaluating new fraud detection systems for revolving credit operations, *International Journal of Forecasting*, Volume 28, Issue 1, pp.216-223 (2012)
- [21].Phua, C., Lee, V., Smith, K., Gayller, R.: *A Comprehensive Survey of Data Mining-based Fraud Detection Research*, *Artificial Intelligence Review*, (2005)
- [22].Kou, Y., Lu, C., Sirwongwattana, S., Huang, Y.: Survey of Fraud Detection Techniques, *Proceedings of the 2004 IEEE International Conference on Networking, Sensing & Control*, Taipei, Taiwan, March 21-23, pp. 749-754 (2004)
- [23].Ash, D.: The UK fraud landscape for financial services, *Computer Fraud & Security*, Volume 2011, Issue 4, pp. 16-18 (2011)
- [24].National Fraud Authority, Annual Fraud Indicator, January 2011, <http://www.homeoffice.gov.uk/publications/agencies-public-bodies/nfa/annual-fraud-indicator/annual-fraud-indicator-2011?view=Binary> , last accessed 10 Oct 2012
- [25].National Fraud Authority, Annual Fraud Indicator, January 2011, <http://www.homeoffice.gov.uk/publications/agencies-public-bodies/nfa/annual-fraud-indicator/annual-fraud-indicator-2012?view=Binary> , last accessed 10 Oct 2012
- [26].The Global Financial Cost of Fraud, *Finance Week*, 16 November 2009, <http://www.financeweek.co.uk/topic/global-financial-cost-fraud> , last accessed 10 Oct 2012
- [27].The Financial Cost of Fraud, Centre for Counter-Fraud Studies, University of Portsmouth, 2011, [http://pkfemail.co.uk/ukassets/images/460/Downloadfiles/The Financial Cost of Fraud WEB.pdf](http://pkfemail.co.uk/ukassets/images/460/Downloadfiles/The_Financial_Cost_of_Fraud_WEB.pdf) , last accessed 4 Oct 2012
- [28].Samakovitis, G.: *Technology Investment decision making: An Integrated analysis in UK Internet Banking*, PhD Thesis, The University of Edinburgh, (2006)
- [29].Edge, M. E., & Sampaio, P. R. F.: A survey of signature based methods for financial fraud detection. *Computers & Security*, Volume 28, Issue 6, pp. 381–394 (2009)

- [30].Edge, M. E., & Sampaio, P. R. F.: The design of FFML: A rule-based policy modeling language for proactive fraud management in financial data streams, *Expert Systems with Applications*, Volume 39 pp. 9966–9985 (2012)
- [31].SAS. (2011). SAS: The power to know, [www.sas.com](http://www.sas.com) , last accessed 12 Oct 2012
- [32].SPSS/IBM. (2011). IBM SPSS Statistics 19. [www.spss.com](http://www.spss.com) , last accessed 12 Oct 2012
- [33].Kapetanakis, S.: Intelligent Monitoring of Business Processes using Case-based Reasoning, PhD Thesis, The University of Greenwich (2012)
- [34].Ma, J., Knight, B.: A General Temporal Theory, *the Computer Journal*, Volume 37, Issue 2, pp. 114-123 (1994)
- [35].OASIS: BPEL, The Web Services Business Process Execution Language Version 2.0, [https://www.oasis-open.org/committees/tc\\_home.php?wg\\_abbrev=wsbpel](https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wsbpel), last accessed 7 Oct 2012
- [36].Workflow Management Coalition (WfMC): XPD L 2.1 Complete, <http://www.wfmc.org/xpdl.html> , last accessed 7 Oct 2012
- [37].Aamodt, A and Plaza, E.: Case-based reasoning: foundational issues, methodological variations, and system approaches. *AI Communications*. Volume 7, Issue 1, pp. 39–59 (1994)
- [38].Kapetanakis, S., Petridis, M., Knight, B., Ma, J., Bacon, L.: A case based reasoning approach for the monitoring of business workflows. In: Bichindaritz, I., Montani, S. (eds.) 18th International Conference on Case-Based Reasoning, ICCBR 2010, LNCS (LNAI), vol. 6176, pp. 390 – 405. Springer, Heidelberg (2010)
- [39].Samakovitis, G. and Fleck, J.: Practitioners, Observers and the Community of Received Wisdom: The Actor-based Approach to Technological Investment Decisions, In: 13th European Conference on Information Technology Evaluation, Genoa, Italy, pp. 28-29 (2006)
- [40].Dorneles, C. F., Heuser, C. A., Lima, A. E. N., da Silva, A. S. & de Moura, E. S.: Measuring similarity between collection of values, In: WIDM '04, Proceedings of the 6th annual ACM inter-national workshop on Web information and data management, ACM Press, New York, NY, USA, pp. 56–63 (2004)
- [41].MacCarthy B.L., Jou, P.: Case-based reasoning in scheduling. In: M.K. Khan and C.S. Wright (eds.), Proceedings of the Symposium on Advanced Manufacturing Processes, Systems and Techniques (AMPST96), pp. 211-218 (1996)
- [42].Petrovic, S., Qu, R.: Case-Based Reasoning as a Heuristic Selector in a Hyper-Heuristic for Course Timetabling Problems, In: Proceedings of the 6th International Conference on Knowledge-Based Intelligent Information Engineering Systems and Applied Technologies (KES'02), pp. 336—340 (2002)
- [43].Likhachev, M., Kaess, M., Kira, Z., Arkin, R.C.: Spatio-Temporal Case-Based Reasoning for Efficient Reactive Robot Navigation. Mobile Robot Laboratory, College of computing, Georgia Institute of Technology (2005)