Use of the Dempster-Shafer Theory for Fraud Detection: The Mobile Money Transfer Case Study

Luigi Coppolino, Salvatore D'Antonio, Valerio Formicola,
Carmine Massei, and Luigi Romano

Abstract. Security Information and Event Management (SIEM) systems are largely used to process logs generated by both hardware and software devices to assess the security level of service infrastructures. This log-based security analysis consists in correlating massive amounts of information in order to detect attacks and intrusions. In order to make this analysis more accurate and effective we propose an approach based on the Dempster-Shafer theory, that allows for combining evidence from multiple and heterogeneous data sources and get to a degree of belief that takes into account all the available evidence. The proposed approach has been validated with the respect to a challenging demonstration case, namely the detection of frauds performed against a Mobile Money Transfer service. An extensive simulation campaign has been executed to assess the performance of the proposed approach and the experimental results are presented in this paper.

1 Introduction

Frauds in the field of electronic payments continuously evolve as new payment technologies and platforms are introduced. Mobile Money Transfer (MMT) refers to payment services which allow to use virtual money in order to carry out payments, money transfers, and transactions through mobile devices. Such services are being increasingly adopted all over the world, particularly in developing countries where banking services and infrastructures are not so largely available as in developed countries and MMT solutions are being deployed to provide payment services to the so-called "'unbanked"' or "'underbanked"' people. Like any other money transfer service, this service is exposed to the risk of money laundering, i.e., the misuse

Luigi Coppolino · Salvatore D'Antonio · Valerio Formicola · Carmine Massei · Luigi Romano
University of Naples Parthenope, Department of Engineering Naples, Italy
e-mail: {luigi.coppolino,salvatore.dantonio,valerio.formicola,
lrom}@uniparthenope.it, carmine.massei@gmail.com

D. Camacho et al. (eds.), *Intelligent Distributed Computing VIII,*
Studies in Computational Intelligence 570, DOI: 10.1007/978-3-319-10422-5_48

consisting in disguising the proceeds of crime and illegal activities and transforming them into ostensibly legitimate money or other assets, or more generally to fraud risks that imply any intentional deception performed to gain financial profit. In this paper we propose a fraud detection system that relies on the Dempster-Shafer theory to spot evidence of ongoing security attacks against MMT systems. This theory is a data fusion technique that allows to combine multiple evidences and to compute a belief value.

The paper is organized as follows. Section II gives an overview of data fusion techniques, with focus on Dempster-Shafer theory; Section III describes the Mobile Money Transfer case study and the frauds considered in this paper; Section IV describes the proposed detection system applied to the MMT case study; in Section V experimental tests and results are shown; Section VI concludes by remarking achieved results and defining future works.

2 Data Fusion Techniques

Data fusion is a process whereby data from multiple sources are combined to yield improved accuracy and more inferences than those that could be achieved using a single source of information. Historically, data fusion has been used in military applications, like remote sensing [16] and target tracking [14]. Also several civil applications are progressively using data fusion techniques to improve the system security and reliability, like robotics [1], medicine [9] and financial infrastructures [10]. The most important problem in data fusion is the development of appropriate models of uncertainty associated with both the state and the observation process. There exist several methods for representing and reasoning about uncertainty, such as the Dempster-Shafer's Theory of Evidence and the Bayesan Inference. In this paper we used the Dempster-Shafer's Theory of Evidence since we do not have a good knowledge of the probabilistic distribution of the states and therefore we cannot calculate the probability a priori required by the Bayesian Inference. Dempster-Shafer's Theory of Evidence is a mathematical theory of evidence introduced in the 1960's by Arthur Dempster [3] and developed in the 1970's by Glenn Shafer [15]. In the Dempster-Shafer framework a proposition can be seen as subsets of a given set of hypotheses. For example, in a fraud detection system, we can consider the set of hypotheses as the set of categories of frauds. Each anomalous event is a subset of the frame of discernment Θ, hence the propositions of interest are in a one-to-one correspondence with the subsets of Θ. Furthermore the set of all propositions corresponds to the set of all subsets of Θ, which is denoted 2^{θ} and is called power-set. In other words we have a set of possible states of the system $\theta_{1...}\theta_N \in \Theta$ which are mutually exclusive and exhaustive. Our goal is to infer the true system state without having an explicit model of the system, but only relying on some observations $E_1 \dots E_M$. Based upon one evidence E_j we can assign a probability that supports a certain hypothesis H_j; in other words we assign a probability to an element of the power-set. A *basic probability assignment (bpa)* is a mass function m which assigns beliefs to a

hypothesis or, in other words, the measure of belief that is committed exactly to the hypothesis H. Therefore, a basic probability assignment is a function $m : 2^{\theta} \rightarrow [0, 1]$ such that $m(\varnothing) = 0$ and $m(H) \geq 0, \forall H \subseteq \Theta$ and $\sum_{H \subseteq \Theta} m(H) = 1$.

We assign two measures [5]:

- the *Belief* function Bel, describing the belief in a hypothesis H, as: $Bel(H) = \sum_{B \subseteq H} m(B)$. The belief corresponds to the lower bound on the probability or rather measures the minimum uncertainty value about a proposition. Its properties are: $Bel(\varnothing) = 0$ and $Bel(\Theta) = 1$.
- the *Plausibility* function of H, Pl(H), which corresponds to the upper bound on the probability and reflects the maximum uncertainty value about proposition H. The plausibility of H is defined as: $Pl(H) = \sum_{B \cap H \neq 0} m(B)$.

Therefore the true belief in the hypothesis H lies in the interval $[Bel(H), Pl(H)]$, while the degree of ignorance is represented by the difference $Bel(H) - Pl(H)$. The second important part of the Dempster-Shafer theory is a rule of combination that permits to combine two independent evidences $E_1$and E_2 into a single more informative hint:

$$m_{12}(H) = \frac{\sum_{B \cap C = H} m_1(B) m_2(C)}{\sum_{B \cap C = \emptyset} m_1(B) m_2(C)}$$

Based on this formula we can combine our observations to infer the system state based on the values of belief and plausibility functions. In the same way we can incorporate a new evidence and update our beliefs as we acquire new knowledge through observations. The theory of evidence allows to reason with uncertainty based on incomplete and also contradictory information extracted from a stochastic environment. Therefore, such theory does not need to know an "*a priori*" probability distribution on the system states like in the Bayesian approach [8].

3 The Mobile Money Transfer Case Study

The Mobile Money Transfer service is a system where virtual money is used to carry out various types of money transfers and financial transactions. For example, a customer can use his mobile phone to carry out financial operations, such as purchasing goods, receiving salary, paying bills, taking loans, paying taxes or receiving social benefits. MMT systems are experiencing rapid adoption. It is expected that mobile payment systems reach US$ 245B in value worldwide by 2014. At the same time, mobile money users are expected to be 340M, equivalent to 5% of global mobile subscribers [13]. The architecture of a MMT system is shown in Fig. 1.

Three classes of users (i.e., Customer, Retailer of *mMoney*, and Merchant) exist in a MMT scenario. They use their mobile phones to communicate with the operations server. Each user is an *mWallet holder*. An *mWallet* is an account hosted in the system allowing the *mWallet holder* to carry out various operations and transactions by using the *mMoney*. The users are connected to the Operations Server that is in

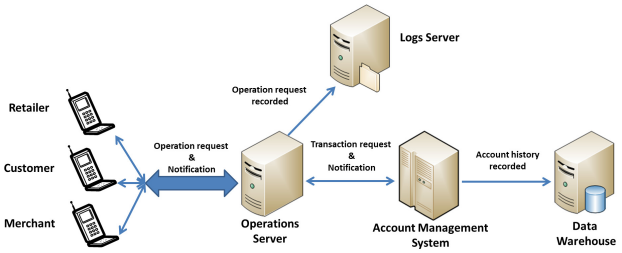

Fig. 1 Architecture of a Mobile Money Transfer system

charge of authenticating the users. It performs simple account management operations (like change the PIN code) and delivers notification messages. It is linked to the Account Management Server which manages the accounts (particularly, if the operation concerns the control of credit/debit). The Account Management Server also stores all the information regarding the user's behaviour.

The Operations Server is also linked to the Logs Server that collects the logs of the various operations that are carried out. The logs contain a wide range of information, such as requests for PIN modification, failed authentication, transaction request, transaction success notification. Historical account management data are stored in the Data warehouse and can be used to analyse customer behaviour. Both log information and account management data can be used to detect frauds. Therefore the input of the system is an operation request received from *mWallet holders*, while the output of the system is the notification of the operation's success/failure, the registration of transaction information and operation information, and the implementation of the requested operation.

Like any other money transfer service, this service is affected by security issues, such as money laundering, privacy protection, frauds, and credit and liquidity risks. Since the success of any payment system is based on ubiquity, convenience, and trust, it is necessary to address emerging risks in order to maintain public confidence in mobile money. To address the security issues of a MMT system, we used the model of such a system developed by the EU FP7 MASSIF ("*MAnagement of Security information and events in Service InFrastructures*"). The MASSIF project has investigated and developed several misuse cases [11]. In order to test the proposed approach we selected the use case named Account Takeover. In this misuse case a fraudster steals the mobile phone from its legitimate user and uses it to perform money transfer. In this misuse case it is very likely that the thief's behaviour differs from the original user's one. Therefore, in order to detect such a misuse case a learning stage is needed. In other terms, the fraud detection system has to be trained by feeding it with information on the user's habits and his usual behaviour. Since

user's data cannot be disclosed due to privacy reasons, we used the MMT simulator developed in the framework of the MASSIF project to generate synthetic data for the learning phase.

4 Fraud Detection through Data Fusion in a MMT System

Fraud detection is the identification of an actual or potential fraud within a system. It relies upon the implementation of appropriate processes to spot the early warning signs of a fraud and can help to uncover new frauds in action as well as historical frauds. It consists in identifying unauthorized activity once the fraud prevention has failed. With reference to the MMT scenario proposed in the MASSIF project [11] we simulated the account takeover misuse case where a fraudster steals the mobile phone from the legitimate user and uses it to perform money transfer. More precisely, once the fraudster has stolen the mobile phone, he attempts to find the pin related to the mobile payment application. Usually the fraudster makes ten attempts with false PIN code to enter the system. Once the fraudster has gained access to the mobile payment application, he tries to do small purchases. To do that he moves from one merchant to another one in order to buy goods. The time interval between two transactions ranges from 3 to 20 seconds. The fraudster performs up to 30 transactions with an amount between 31 and 50 €. In order to detect the Account Takeover misuse case we propose a Fraud Detection System (FDS), which implements a number of rules to analyse the deviation of each incoming transaction from the normal profile of the user and assign an initial belief to it. The initial belief values are combined to obtain an overall belief by applying the Dempster–Shafer theory. The overall belief is then compared with two thresholds in order to understand if the user's behaviour is to be considered fraudulent or genuine. The proposed FDS comprises the following three major components: a Rule Based Filter, a Dempster-Shafer combiner and an Analyser. The flow of events in the FDS has been depicted in the block diagram in Fig. 1.

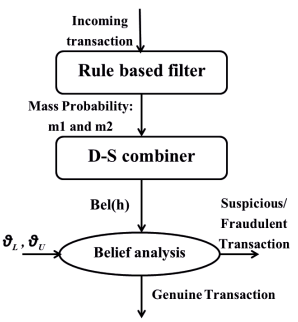

Fig. 2 Block diagram of the proposed FDS

4.1 The Rule Based Filter (RBF)

The RBF consists of rules which classify the transactions made by the user as fraudulent with a certain probability. It measures the extent to which the transaction's behavior deviates from the normal profile of the user. The rules used in our study are:

- Rule R1, authentication attempts: we analyse the time interval between the first and the last authentication attempt failed. If this time interval exceeds a given threshold (e.g. 15 seconds), then there is a high probability that the transaction is fraudulent.
- Rule R2, outlier detection: a user usually carries out similar types of transactions in terms of amount, which can be visualized as part of a cluster. Instead, a fraudster is likely to deviate from the customer's profile, so his transactions can be detected as exceptions to the cluster. This process is known as outlier detection. One of the most used algorithms used for detect cluster is the DBSCAN (*Density Based Spatial Clustering of Application with Noise*) [6]. Let $U^{'} = \{u_1 \dots u_n\}$ denote the clusters in a database *D* for a specific user of the MMT system U_k and $A = \{a_1 \dots a_n\}$ be the set of attributes used to generate the clusters. For any transaction the possible attributes are transaction amount, the date, the merchant involved in the transaction, etc. A transaction T^{c_k} is an outlier if it does not belong to any cluster in the set $U^{'}$. In this way we can understand if a transaction is fraudulent. The degree of outlierness allows to measure the extent of deviation of an incoming transaction. If the average distance of the amount *p* of an outlier transaction T^{c_k} from the set $U^{'}$ is v_{avg}, then its degree of outlierness is:

$$\begin{cases} d_{out} = \left(1 - \frac{\varepsilon}{v_{avg}}\right) & \mid N_\epsilon(p) \mid < MinPts \\ 0 & otherwise \end{cases}$$

  where *MinPts* is the minimum number of points required to form a cluster, while ε is the maximum radius of the cluster. As said earlier, to form a cluster we can use various attributes. In our study we used the amount of the transactions. Particularly, transactions with an amount between 31 and 50 €were considered as fraudulent.

An FDS is subjected to a large number of transactions, a high percentage of them being genuine. The RBF is an essential component since it separates out most of the easily recognizable genuine transactions from the rest.

4.2 The Dempster-Shafer Combiner (DSC)

The role of the DSC is to combine evidences from rules R1 and R2 and compute an overall belief value for each transaction. For the detection of fraud in the MMT system the Dempster-Shafer theory is more relevant as compared to other fusion

methods since it introduces a third alternative: "*unknown*". It provides a rule for computing the confidence measures of three states of knowledge: *{fraud, no fraud, suspicious}* based on data from new as well as old evidence. Furthermore, in DST, evidence can be associated with multiple possible events unlike traditional probability theory where evidence is associated with only one event. As a result, evidence can be more meaningful at a higher level of abstraction.

The part of DST that is of direct relevance is the Dempster's rule for combination [10]. In order to apply the Dempster-Shafer theory we need to define a frame of discernment U which is a set of mutually exclusive and exhaustive possibilities. With reference to the MMT fraud detection problem the frame of discernment is $U = \{\neg fraud, suspicious, fraud\}$. Hypothesis $F = \{fraud\}$ means that the transaction is fraudulent, hypothesis $N = \{\neg fraud\}$ is the hypothesis that the transaction is not fraudulent, and hypothesis $S = \{suspicious\}$ means that the transaction is suspicious. The mass probability assignments for the two rules R1 and R2 can now be given as follows:

- mass probability m_1: let *t* denote the time interval between the first and the last authentication attempt, we can consider the following assignments: if $t > 15$ seconds, then $[m_1(F) = 0.6, m_1(N) = 0, m_1(S) = 0.4]$. Instead, if $10 \leq t \leq 15$ seconds, then: $[m_1(F) = 0.4, m_1(N) = 0, m_1(S) = 0.6]$. Finally, if $t < 10$ seconds, then: $[m_1(F) = 0, m_1(N) = 0.6, m_1(S) = 0.4]$.
- mass probability m_2: for a transaction detected as an outlier we make the mass probability assignment using the degree of outlierness $d_{out} = 1 - \frac{\varepsilon}{v_{avg}}$ where ε is the credit limit that is the maximum amount of credit that a user can spend, while v_{avg} is the average distance of the amount of an outlier transaction from the set of the other transactions. Hence we consider the following assignment: $\left[m_2(F) = 1 - \frac{\varepsilon}{v_{avg}}, m_2(N) = 0, m_2(S) = 1 - \left(1 - \frac{\varepsilon}{v_{avg}}\right)\right]$.

As we can see in both cases the zero in the basic probability assignment for the hypothesis N does not imply impossibility. It means that neither of the rules R1 and R2 give any support to the belief that the set of transactions are genuine.

4.3 The Analyser

The two probability masses are combined using the Dempster-Shafer combiner to get the initial value of belief for the set of transactions made by the user. Particularly in our study we used the *Bel(F)*, i.e. the minimum probability that the event "*Fraud*" occurs. In our analysis we defined two thresholds: θ_L is the lower threshold, where $0 \leq \theta_L \leq 1$, and θ_U is the upper threshold, where $0 \leq \theta_U \leq 1$ and $\theta_L \leq \theta_U$.

If $Bel(F) < \theta_L$ the user behaviour is considered as genuine and is approved. On the other hand, if $Bel(F) > \theta_U$, then the user behaviour is declared to be fraudulent. In case $\theta_L \leq Bel(F) \leq \theta_U$, the user behaviour is labelled as suspicious.

The two thresholds and the other parameters can be chosen by observing the performance of the FDS over a large number of simulation trials.

5 Experimental Tests and Results

We demonstrated the effectiveness and performance of our FDS by conducting an extensive experimental campaign. Due to the unavailability of real data we used the simulator developed by the MASSIF project to generate synthetic transactions that represent the behaviour of genuine users as well as that of fraudsters [7]. We used standard metrics to evaluate the performance of the system under different test cases. True positives (TP) are the fraudulent users detected by the system and false positives (FP) are the genuine users with a normal behavior detected as fraudsters.

The effectiveness of the proposed system depends on θ_L and θ_U. If θ_U is set too high, then most of the frauds will go undetected, whereas if θ_U is set too low, then there will be a large number of false alarms. Similarly, high value of θ_L will let most of the frauds go through and low value of θ_L will lead to unnecessary investigation of a large number of genuine transactions. Hence, selection of the two thresholds has an associated tradeoff. We carried out our experiments to determine a good choice of these parameters.

In Fig.3 (left and right), we show how the mean values of TP and FP vary with each threshold value. Particularly, the values of TP strongly depend on the value of the upper threshold and this behaviour is especially noticeable for users who are victim of fraud. Instead, the values of FP depend on the value of the lower threshold and this behaviour is especially noticeable for users who are not a victim of fraud. It has to be noted that mean values of TP increase as θ_U increases. Good performance is attained with values of the upper threshold between 0.72 and 0.74. Instead, values of FP decrease as the θ_L increases, then good values for the lower threshold are under 0.35. The effectiveness of the FDS is also dependent on the two parameters, i.e. ε and *MinPts*. More precisely, the larger ε, the less is the number of clusters formed. In the limit, there will be only one large cluster. Also, the higher the value of MinPts, the less is the number of clusters formed. If it is set too high, there will be no clusters since the *MinPts* condition is not satisfied. However, if both the parameters are small, there can be a lot of clusters. If *MinPts* is set to 1, then each point in the database is

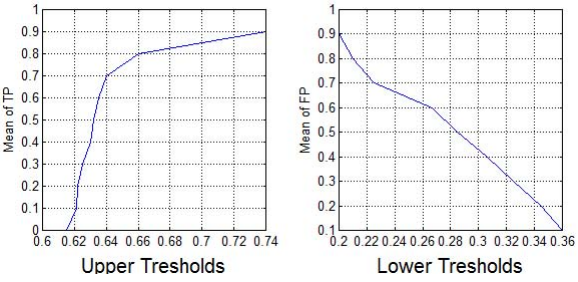

Fig. 3 Mean of True Positive (TP, left) and Mean of False Positive (FP, right) rates

treated as a separate cluster. In our study, after having studied the trend of the *Bel*, we decided to set $\varepsilon = 2\%$ of credit limit and $MinPts = 1$.

The experimental results show that the 95% of the users indicated by the simulator as victim of fraud is properly detected, while the 5% is detected as suspicious. Similarly, the 97% of the genuine users is properly detected, while the 3% is detected as suspicious.

6 Conclusions and Future Work

We have proposed a novel Fraud Detection System based on the integration of two approaches, i.e. the Dempster-Shafer theory and the rule-based filtering. Dempster's rule is applied in order to combine multiple evidences from the rule-based component for computation of belief about the transactions carried out by a user of the MMT system. This value of belief is compared with two thresholds in order to understand if the behaviour of the user is fraudulent, genuine or suspicious. Moreover the FDS has been designed as a modular architecture so that new rule-based filters can be added at a later stage using any other effective technique. The results of the simulation campaign show that the fraud detection system based on the Dempster-Shafer theory is able to detect frauds and suspicious behaviours of the MMT users. The system can be further improved by using a Bayesian approach for a more accurate assessment of the cases where the user is detected as suspicious. Finally, it would be interesting to compare the performance and accuracy of the Fraud Detection System based on the Dempster-Shafer theory with those of the FDS implemented by the MASSIF project and using the finite state machine technology. The latter approach has been already used in the scenario of an eHealth [4] infrastructure and of a dam infrastructure [2], [12].

Acknowledgements. The research leading to these results has received funding from the European Commission within the context of the Seventh Framework Programme (FP7/2007-2013) under Grant Agreement No. 313034 (Situation AWare Security Operation Center, SAWSOC Project) and under Grant Agreement No. 257644 (MAnagement of Security information and events in Service Infrastructures, MASSIF Project). It has been also partially supported by the TENACE PRIN Project (n. 20103P34XC) funded by the Italian Ministry of Education, University and Research.

References

1. Abidi, M.A., Gonzalez, R.C.: Data fusion in robotics and machine intelligence. Academic Press Professional, Inc. (1992)
2. Coppolino, L., D'Antonio, S., Formicola, V., Romano, L.: Enhancing siem technology to protect critical infrastructures. In: Hämmerli, B.M., Kalstad Svendsen, N., Lopez, J. (eds.) CRITIS 2012. LNCS, vol. 7722, pp. 10–21. Springer, Heidelberg (2013)
3. Arthur, P.: Dempster. A generalization of bayesian inference. Technical report, DTIC Document (1967)

4. Di Sarno, C., Formicola, V., Sicuranza, M., Paragliola, G.: Addressing security issues of electronic health record systems through enhanced siem technology. In: 2013 Eighth International Conference on vailability, Reliability and Security (ARES), pp. 646–653 (September 2013)
5. Durrant-Whyte, H.: Multi Sensor Data Fusion. Australian Centre for Field Robotics (2001)
6. Ester, M., Kriegel, H.-P., Sander, J., Xu, X.: A density-based algorithm for discovering clusters in large spatial databases with noise. In: KDD, vol. 96, pp. 226–231 (1996)
7. Gaber, C., Hemery, B., Achemlal, M., Pasquet, M., Urien, P.: Synthetic logs generator for fraud detection in mobile transfer services. In: 2013 International Conference on Collaboration Technologies and Systems (CTS), pp. 174–179. IEEE (2013)
8. Gros, X.: NDT data fusion. Elsevier (1996)
9. Jannin, P., Grova, C., Gibaud, B.: Medical applications of NDT data fusion. Springer, Heidelberg (2001)
10. 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 10(4), 354–363 (2009)
11. Rieke, R., Coppolino, L., Hutchison, A., Prieto, E., Gaber, C.: Security and Reliability Requirements for Advanced Security Event Management. In: Kotenko, I., Skormin, V. (eds.) MMM-ACNS 2012. LNCS, vol. 7531, pp. 171–180. Springer, Heidelberg (2012)
12. Romano, L., D'Antonio, S., Formicola, V., Coppolino, L.: Protecting the WSN zones of a critical infrastructure via enhanced SIEM technology. In: Ortmeier, F., Daniel, P. (eds.) SAFECOMP Workshops 2012. LNCS, vol. 7613, pp. 222–234. Springer, Heidelberg (2012)
13. Shen, S.: Market insight: The outlook on mobile payment (2010)
14. Smith, D., Singh, S.: Approaches to multisensor data fusion in target tracking: A survey. IEEE Transactions on Knowledge and Data Engineering 18(12), 1696–1710 (2006)
15. Srivastava, R.P.: The dempster-shafer theory: An introduction and fraud risk assessment illustration (2011)
16. Zhang, J.: Multi-source remote sensing data fusion: status and trends. International Journal of Image and Data Fusion 1(1), 5–24 (2010)