Fuzzy System-based Suspicious Pattern Detection in Mobile Forensic Evidence

Konstantia Barmpatsalou[1], Tiago Cruz[1], Edmundo Monteiro[1], and Paulo Simoes[1]

Center for Informatics and Systems of the University of Coimbra, Pólo II-Pinhal de Marrocos, 3030-290 Coimbra, Portugal
{konstantia,tjcruz,edmundo,psimoes}@dei.uc.pt

Abstract. Advances in Artificial Intelligence and Soft Computing have increased the probabilities of implementing mechanisms that are able to predict or -in the most optimistic scenarios- imitate human behaviour. One of the fields that benefits more from the particular improvements is the one of criminal investigation and more precisely the subdomain of Digital Forensics. Criminal activity involving digital devices and particularly smartphones shows interesting behavioural variations that can be tacked and distinguished from application and system activity. In this paper, the authors create a technique that analyzes smartphone users' activity and recognizes potentially suspicious patterns according to predefined expert knowledge in actual use case scenarios. Mamdani-type fuzzy systems are tested as detection mechanisms in existing datasets and different configurations are applied. Lastly, all the solutions are evaluated for their accuracy against the ground truth in order to select the most appropriate settings for each case. Since the experiments are conducted with successful outcomes, the current paper results can be used as a springboard for future research concerning the performance testing of other soft computing methods, such as neurofuzzy networks.

Key words: Mobile Forensics, Fuzzy Systems, Membership Functions

1 Introduction

Digital criminal investigation involving computers, mobile devices and networks, and its respective automation is a field that continues evolving steadily. New forensic techniques are emerging and the investigators' tasks are facilitated [1]. Especially when the approaches concern tasks directly involving digital devices, such as the cases of data acquisition and malware identification, the progress is exponential. Recent advances in the Hard Computing field, such as efficient machine learning methods played a crucial role towards this evolution. Nevertheless, when human behaviour is involved and and uncertainty in actions increases, Hard Computing techniques are not as compelling and different methods need to be adopted.

The majority of the traditional machine-learning approaches, such as *Support Vector Machines (SVM)*, *Linear Regression*, etc. support binary output states,

which signify that either a feature has a specific characteristic or it does not. However, there are problems with higher complexity than plain binary feature classification. Problems depending on multiple data sources and occurring in a vague contextual base render binary approaches rather inflexible, due to higher degrees of uncertainty [2].

When an individual is prompted to investigate different data types acquired from mobile devices for suspicious patterns, each type has a different level of importance towards the calculation of an overall suspiciousness rate for certain criminal actions. Thus, it is easily perceivable that single actions or data patterns cannot be strictly characterized as innocent or guilty, but they would rather need a more detailed type of classification in order to determine different levels of suspiciousness.

By regarding evidence, actions and potential outcomes from a fuzzy theory-related point of view, investigators can take advantage of the multiple states and create more realistic event combinations. However, fuzzy theory by itself is lacking the ability of learning from previous conditions and states, which is one of the major drawbacks concerning its application to systems that are in great need of memory so as to be efficient, other than being limited to representation and basic functionality. Moreover, while fuzzy systems make use of a comprehensive decision-making, they have severe adaptation limitations.

The particular drawbacks served as inspiration, not only for the current paper, but also for future work. The authors did not only need to define how specific combinations derived from forensically acquired data form potentially suspicious patterns, but also to generalize this rule base into a global pattern recognition mechanism, which would be able to identify suspiciousness levels and adapt their parameters according to the behaviour of different data sources, so as to optimize the results at the maximum possible level.

For the specific reason, *Neuro-Fuzzy systems (NFSs)* were considered a suitable future candidate. They "provide powerful and flexible universal approximations with the ability to explore interpretable *IF-THEN* rules" [3]. The consolidation of fuzzy systems and neural networks offers adequate pattern recognition capabilities in an uncertain universe of discourse with consistent justification, derived from the solid fuzzy rule base source.

This paper is the first part of a two-step approach aiming to create a semi-automated consulting, prediction and decision-making methodology for mobile forensic investigation purposes. Firstly, expert knowledge in the field of mobile criminal investigation is used in order to create the ground truth. Moreover, the authors proceed to the generation of assumptions and suspicious patterns concerning the outcome of various user actions in different data types that can be retrieved during a forensic acquisition. Afterwards, the knowledge is diffused to the creation of fuzzy systems and their equivalent rules, one for each data type present. Finally, the fuzzy systems are validated with actual data and their performance is evaluated against the ground truth equivalent. The current paper elaborates the methodology used up to this point. However, the schema will be complete in the second step, which consists of the use and performance evalua-

tion [4] of a NFS atop of the fuzzy systems. Profiting from both the advantages of neural networks and fuzzy systems is promising for an efficient criminal pattern -agent infiltration, drug trafficking, arson or murder- recognition procedure.

The rest of the paper is presented in the following manner. Section 2 contains the related work in the field, while Section 3 presents the respective methodology the authors followed. Section 4 demonstrates the proposed mechanism, Section 5 performs the results evaluation and Section 6 enumerates the conclusions after the research conduction.

2 Related Work

To the best of the authors' knowledge, noteworthy research has been conducted in the area of fuzzy and Neuro-Fuzzy data analysis for Mobile Forensics and similar disciplines, such as the one of Intrusion Detection, which provided essential and useful insights towards the completion of the current work. However, the amount of research papers concerning Mobile Forensics might be smaller than the ones referring to Intrusion Detection, mainly due to the fact that the former discipline is more recent.

Based on the "*Autonomous Agents for Intrusion Detection (AAFID)*" [5] architecture, "Fuzzy Intrusion Recognition Engine (FIRE)" [6] used fuzzy agents in order to detect and determine the severity level of various attack types in computer networks.

An *Intrusion Detection System (IDS)* based on a hybrid neural network and fuzzy logic-based implementation was designed by Chavan et al. [7]. During the training procedure, the authors used SNORT in order to capture traffic data over an IP network. Afterwards, "a signature pattern database was constructed using protocol analysis and a Neuro-Fuzzy learning method" [7]. The system was then evaluated using the 1998 *DARPA Intrusion Detection Dataset* [8] for known attack patterns, achieving relatively high classification accuracy scores.

Data deriving from traditional criminal investigation procedures provided a starting research point for Stoffel et al. [9]. The authors applied the fuzzy sets theory (clustering-membership function extraction-rule inference) to evidence deriving from criminal activity in Switzerland and proved that their methodology is appropriate for "inferring expert-system-like rules from a forensic database" [9].

The paper by Islam and Verma [10] is more oriented towards handling of data deriving from mobile devices. More specifically, a fuzzy model is used so as to perform a privacy risk analysis of *Short Message Service (SMS)* texts in 3G networks based on the senders' identity and the relationship to the user. However, the authors did not present any experimental data so as to further evaluate their initial hypothesis.

In order to effectively detect *Denial of Service (DoS)* attacks in a computer network infrastructure, Arun Raj Kumar and Selvakumar [11] profited from the combination of the precise rule definition of fuzzy systems and the automatic rule acquisition of neural networks.

Automatic rule definition by a Neuro-Fuzzy system was also successful in cases of Android malware detection patterns [12] according to raw data retrieved from devices, such as CPU usage, permissions granted per application and functions called.

It is notable that the majority of the research conducted in the field of hybrid intelligent systems (neural networks, fuzzy systems, genetic algorithms) is primarily related to intrusion detection, secondarily to algorithms concerning the performance of network forensics techniques and only fewer works have been dedicated to other research areas related to Digital Forensics. This deficiency is also a result of the fact that datasets concerning actual evidence from mobile devices are rather limited. The next section describes the methodology the authors followed in order to develop the fuzzy part of the Neuro-Fuzzy system for detecting suspicious patterns in mobile data.

3 Methodology

This section presents the proposed methodology concerning suspicious pattern detection from mobile dataset. The procedure consists of the construction of a use case scenario, where the problem and the meaning of suspiciousness are contextualized. Afterwards, the authors proceed to the rule inference and the ground truth generation with the aid of expert knowledge. Further details concerning the used datasets are provided and the fuzzy systems for the use case are configured.

3.1 Use Case Scenario

One of the fundamental steps that need to be taken in order to proceed with the fuzzy system creation is structuring a scenario of potential criminal activity occurrences and inferring the respective rules according to existing expert knowledge concerning the particular situation. The authors used the FP 7 Project SALUS D2.3 publicly available deliverable [13] so as to determine a Use Case Scenario that would fit their needs. The use cases concerned three different events that required the presence of *Public Protection and Disaster Relief (PPDR)* systems; "public order demonstration or riot, Olympic-style sporting event and heavy flooding due to prolonged periods of rain" [13]. The first option, public order demonstration or riot, was considered as the most suitable for the research purposes, due to the high probability of occurrence of unfortunate events involving mobile devices belonging to PPDR officers.

One of these events is the backbone of the scenario the authors constructed. The current paper examines the case of PPDR officers infiltrating the rioting forces and how this fact can be proved by their device seizure upon suspicion. The forensic investigation authorities capture an image of the device at a given moment after the rioting incident, which is used as the base for further investigation. The fuzzy system will be tested for its efficiency against this part of

information. However, no assumptions can be made without the presence of expert knowledge, which is elaborated in detail below.

3.2 Expert Knowledge

Conducting a research strongly correlated to actual criminal investigation would be impossible without prior expert knowledge available. The knowledge base encountered in the current paper is a hybrid compilation of incidents the use cases provided in the SALUS FP7 Project deliverables [14] and of on-field investigation practices provided by an officer of the *Greek Police Escort Teams Department (GPETD)*. After collecting all the essential insights, the authors have been able to structure the rules of each fuzzy system present in the research. Due to space limitations, one use case will be examined (PPDR officers infiltrating for protesters) and the example of SMS data deriving from three devices will be presented.

Another challenge that the authors faced was the lack or unavailability of actual evidence retrieved from devices involved in criminal activities. As a result, delinquent actions had to be simulated and injected in the datasets as standalone patterns. The a-priori expert knowledge served as a solid background for the rule generation, which is analyzed in the following subsection.

3.3 Rule Inference

Using the expert knowledge mentioned in the previous section, the authors created respective rules concerning the data categories for the use cases. The rules were formed from a combination of the available data and the investigation directives for the use case. If the use case changes, the rules are as well altered. For the scenario of the rioting infiltration by PPDR officers, the following setup was created.

Sent SMS texts retrieved from a device of a potential infiltrator may have the following attributes:

- If officers are infiltrators, they will use their devices to communicate with their accomplices only in cases of extreme necessity. As a result, the rate with which a sent message will appear is going to be very low.
- Most of the accomplices may use one-time payphones, which are equipped with SIM modules from the same country the incidents occur. Recipients with local numbers are considered more suspicious.
- According to the *GPETD* experts, messages exchanged during rioting or right before similar incidents are very short in length.
- As a result, the sent SMS pattern with the combination (very low appearance frequency-very short length-local country code source) is considered the most suspicious.

Nonetheless, the rule inference procedure needs a functioning dataset that is able to cover the research requirements in size and content. The following subsection covers in detail the challenges the authors faced in the quest of a suitable data source.

3.4 Datasets and Ground Truth Generation

One of the main limitations the authors have encountered was the dataset availability and suitability. Due to the increased sensitivity of data deriving from mobile devices and the limitations this fact provokes in terms of data distribution and ownership, there are not many available sources of mobile device images or database entries. Most of the existing, purely experimental datasets consist of limited entries, fact that rendered them inappropriate for the current research, since its purpose cannot be based on a small amount of data.

A more appropriate alternative was the "Device Analyzer Dataset" [15], a collection of real-time usage data from Android devices, provided by the University of Cambridge. However, the dataset availability is not the only restriction to be taken into account. Data anonymity should also be preserved and no attempts to infer names or other entities from the set should be performed. As an additional step to privacy preservation, the current work does not contain names of entities in their encoded format, but swaps their presence to their appearance frequency instead.

Each dataset is a compilation of snapshots belonging to a certain device and contains lists of attributes such as call logs, SMS texts, network usage statistics, location data, alarms, application settings, etc., retrieved during a considerable period of time. All the information is stored in a Comma Separated Value (.csv) file and each row consists of the data type header, alongside with the existing data. Pre-processing is essential in order to separate the data types and adjust the information to the needs of the research. In this paper, adapted information from three different mobile devices, namely (Dev.1, Dev.2 and Dev.3) is used for SMS data.

The data are formatted in a three-column .csv file and each column represents one attribute; message length, receivers' appeance frequency and receivers' localization. Each row is an SMS text with its equivalent characteristics. In the rest of the paper, this row will be referred to as a pattern. The proposed methodology can be expanded to every data type with the appropriate pre-process. According to the previous paragraph, the SMS data type can be represented as follows:

$$\texttt{SMS}(\texttt{Appearance\_Frequency}, \texttt{Length}, \texttt{Country\_Source}) \tag{1}$$

It can be safely deduced that for any given data type, each pattern belonging to the same dataset will have a similar format to the following:

$$\texttt{DataType}(\texttt{Attribute1}, \texttt{Attribute2}, ..., \texttt{AttributeN}) \tag{2}$$

which can be formally represented as:

$$\mathtt{X_i} = (\mathtt{X_{i1}}, \mathtt{X_{i2}}, ..., \mathtt{X_{in}}) \tag{3}$$

Having determined the rule combinations that render a pattern suspicious in subsection 3.3, the authors' next step is the generation of ground truth data. Every tuple of attributes (see equation 1) corresponds to a suspiciousness numerical value in a scale from zero to one, where zero is the lowest and one is the highest value. Let S be the discrete suspiciousness subset which contains the following indicative values corresponding to different suspiciousness levels:

$$\texttt{S} = \{\texttt{0.15} : \texttt{VERY.LOW}, \texttt{0.25} : \texttt{LOW}, \texttt{0.5} : \texttt{MEDIUM}, \texttt{0.75} : \texttt{HIGH}, \texttt{1} : \texttt{VERY.HIGH}\} \quad (4)$$

Since the datasets were not originally created for digital forensic analysis purposes and the existence of potentially suspicious patterns is unlikely, the authors injected the datasets with suspicious attribute combinations so as to have a complete view of the future system performance.

3.5 Fuzzy System Configuration

In order to proceed to the creation of the fuzzy part of the system, the authors of this paper followed the guidelines provided by Fuller [16], which are summarized in the following key points and equivalent justifications. One of the first factors to be taken into consideration is if fuzzy systems are the appropriate solution for the given problem. In theory, suspiciousness (the output) can be depicted as a fuzzy variable in the following manner:

$$\texttt{Suspiciousness} = [\texttt{VERY.LOW}, \texttt{LOW}, \texttt{MEDIUM}, \texttt{HIGH}, \texttt{VERY.HIGH}] \quad (5)$$

which can be formally represented in the following way. Let $\overline{Y}$ be a fuzzy variable and y one of its instances. Every instance will always belong to a discrete set of values with specific ranges, upper and lower bounds.

$$\forall y \in \overline{Y}, a \leq y \leq b, \{\overline{Y} = [a, b] | \forall z \in \overline{Y}, \exists U : \overline{Y} \cap U = \{z\}\} \quad (6)$$

Moreover, all input variables should be described approximately or heuristically. Table 1 represents the fuzzy approximation of all the system inputs.

Input Variable	**Fuzzy Approximation**
Length	VERY SHORT, SHORT, MEDIUM, LONG, VERY LONG
Appearance Frequency	VERY LOW, LOW, MEDIUM, HIGH, VERY HIGH
Country Source	FOREIGN, UNDEFINED, LOCAL

Table 1. Fuzzy variable ranges

The first column represents the variable, whereas the second shows the ranges attributed to each variable. The theoretically defined rules in Subsection 3.3 have to be represented in a formal manner [17] and be placed in the appropriate system section so as to become structural elements of the rule base. An example of a rule concerning suspicious patterns is presented below.

```
IF (Appearance==Very_Low)&&(Length == Low)&&(Country_Source==Local) THEN
                    (Suspiciousness==Very_High)
```

The rest of the rules are formed in a similar manner, with different values and have the generic following format. Let f be the fuzzy variable, V its equivalent value, S the overall suspiciousness and V_s its value.

$$\texttt{IF } (\texttt{f}_1\texttt{==V}_1)\texttt{\&\&}(\texttt{f}_2\texttt{==V}_2)\texttt{\&\&...\&\&}(\texttt{f}_n\texttt{==V}_n)\texttt{ THEN } (\texttt{S==V}_s)$$

Afterwards, during designing the system, the authors reviewed and verified the criteria for "readability and interpretability of the variables and the rules that are deriving from them" [18], as they were presented in the papers by Guillaume and Charnomordic [19] and Gacto et al. [20].

- While aiming to maintain a high degree of semantic cohesion, every fuzzy set should represent a well-defined and non-vague concept. The fuzzy sets and the value range of each variable participating in the current research have specific meanings; fact that can be proven by consulting Table 1.
- Each fuzzy variable should not exceed the 7 ± 2 range fields, which is defined as the threshold for human perception capabilities [18]. In the current paper, the maximum number of different value ranges is 5, number that falls between the aforementioned limit.
- There is no point within the system's universe of discourse that does not belong to at least one fuzzy set.
- A fuzzy set should be normal; in a fuzzy system $\overline{F}$, there should always exist at least one χ, the membership degree (height) of which should be equal to 1.
- It is obligatory that "all fuzzy sets should overlap in a certain degree" [18].

After concluding the fuzzy system configuration phase, the structure of the proposed mechanism is completed. The next section presents its role and contribution.

4 Proposed Mechanism

The mechanism proposed in the current paper is depicted in Figure 1 and has a dual purpose.

- It is a proof-of-concept that fuzzy systems are a satisfactory solution for evaluating forensic data deriving from rather uncertain user behaviour that cannot be described in the form of specific and predictable norms, such as the cases of malware identification and attack detection.
- Its effectiveness enables the perspective of further evolution so as to achieve a higher investigation automation level and a lower dependence degree from expert knowledge. Such an effort will be undertaken by the authors of the current paper in a future phase, when neuro-fuzzy networks will be used in order to optimize the existing system parameters.

The upper half of the figure presents the solution described in the current paper. Data concerning evidence acquired from mobile devices are split into various categories and can be generally related to "user information, application-generated content or system default settings" [1]. Due to the nature of this research, the first type is preferred. Afterwards, each data type attribute is becoming a fuzzy variable. If the number of attributes is greater than five, application of feature selection methods is advised, so as to avoid an increased degree of complexity during the rule inference phase [21, 22].

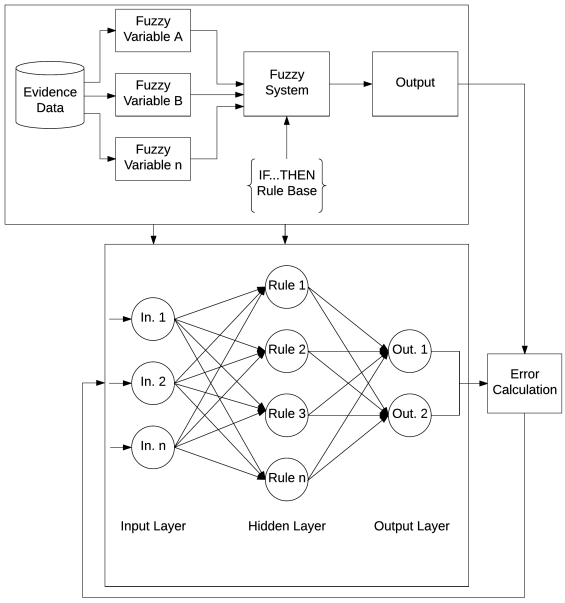

Fig. 1. Proposed mechanism outline

In the next step, the appropriate membership functions have to be selected. The procedure followed is rather inverse. The general practice guidelines suggest the membership function choice as a combination of the researchers' intuition and the means to the best system parameter performance [16]. For the specific case, the authors of this paper lean to the latter option. They tested the Mamdani Fuzzy System functionality with different membership functions and selected the most appropriate, according to various metrics concerning the system performance. The procedure is elaborated in detail in Section 5.

The lower half of the figure depicts the future phase of the current research. A 3-layer neural network accepts the fuzzy inputs as its input layer, the fuzzy rules as the hidden layer and produces the respective outputs, situated in the output layer. A back-propagation algorithm is used in order to compare the neural network outputs to the fuzzy ones and re-configure the parameters for the next run. Such a procedure would not be achievable without a successful fuzzy system evaluation procedure, which is the foundation for optimal output value comparisons.

5 Evaluation

The fuzzy system evaluation and simultaneous membership function selection was a rather complicated procedure. In order to select the appropriate setup for each dataset assigned to the respective fuzzy system, the authors followed

an evaluation methodology based on the comparison of the fuzzy systems' output and the ground truth values. With the ground truth considered the target and the fuzzy output being the feature variable, the fuzzy output values of five fuzzy systems configured with different membership functions (Triangular, Trapezoidal, Bell, Gauss and Gauss2) were classified into five different groups of suspiciousness using the *Nearest Neighbour*, *SVM*, *Naive Bayes*, *AdaBoost* and *Random Forest* classification techniques.

M.F.	Algorithm	AUC	Accuracy	Precision	Recall	FPR
Triangular	kNN	0.583	0.267	0.811	0.267	0.175
	SVM	0.578	0.809	0.800	0.809	0.169
	Naive Bayes	0.567	0.805	0.649	0.805	0.174
	AdaBoost	**0.592**	**0.815**	**0.842**	**0.815**	**0.164**
	Random Forest	0.592	0.814	0.840	0.814	0.164
Trapezoidal	kNN	0.573	0.808	0.799	0.808	0.172
	SVM	0.573	0.808	0.799	0.806	0.172
	Naive Bayes	0.561	0.802	0.648	0.802	0.176
	AdaBoost	**0.574**	**0.808**	**0.846**	**0.808**	**0.171**
	Random Forest	**0.574**	**0.808**	**0.846**	**0.808**	**0.171**
Bell	kNN	0.923	0.951	0.951	0.9512	0.029
	SVM	0.748	0.824	0.825	0.824	0.102
	Naive Bayes	0.904	0.872	0.910	0.872	0.035
	AdaBoost	**0.974**	**0.981**	**0.981**	**0.981**	**0.009**
	Random Forest	0.945	0.963	0.964	0.963	0.021
Gauss	kNN	0.908	0.952	0.952	0.952	0.037
	SVM	0.858	0.864	0.889	0.864	0.058
	Naive Bayes	0.858	0.852	0.880	0.852	0.055
	AdaBoost	**0.925**	**0.960**	**0.961**	**0.960**	**0.030**
	Random Forest	0.915	0.956	0.956	0.956	0.032
Gauss2	kNN	0.924	0.961	0.961	0.961	0.0299
	SVM	0.884	0.871	0.903	0.871	0.0481
	Naive Bayes	0.882	0.865	0.893	0.865	0.0450
	AdaBoost	0.926	0.963	0.963	0.963	0.0305
	Random Forest	**0.931**	**0.963**	**0.963**	**0.963**	**0.0276**

Table 2. Evaluation metrics per membership function for the SMS Dev. 1 dataset

The respective confusion matrices were created and the following metrics were calculated in average for all the groups of suspiciousness (See equation 4); Area Under Curve (AUC) ("ability of a classifier to rank a randomly chosen positive test example higher than a negative one" [23]), Accuracy (amount of correctly classified patterns over the total amount of patterns), Precision (positive predictive value, ratio of True Positive (TP) values over the sum of TP and False Positives (FP)), Recall (TP rate or sensitivity, ratio of TP over the sum of and False Negative (FN) values) and False Positive Rate (FPR)(ratio of FP values over the sum of FP and True Negative (TN) values). An analytical presentation of the metrics formulae can be found in the following equations.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

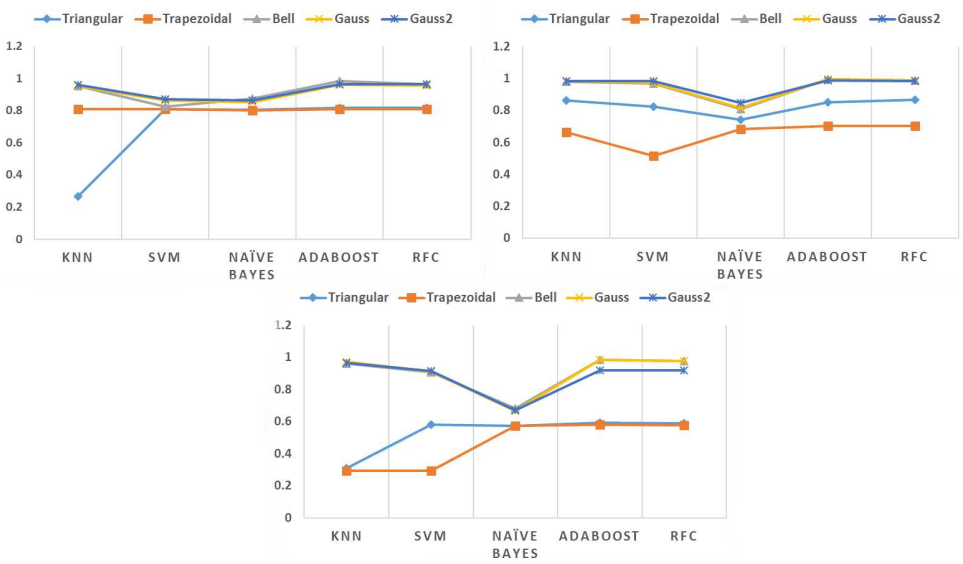

Fig. 2. Average accuracy per dataset, membership function and classification technique

$$Precision = \frac{TP}{TP + FP} \tag{8}$$

$$Recall = \frac{TP}{TP + FN} \tag{9}$$

$$FPR = \frac{FP}{FP + TN} \tag{10}$$

Table 2 contains the cumulative results for all the candidate membership functions and their respective metrics after every classification type. After evaluating all the three datasets –which can be found in Appendix A–, the following observations were made:

- Triangular and Trapezoidal membership functions perform worse than the rest of the other candidates in every dataset and under every classification algorithm.
- The Bell membership function shows the best performance rates in every dataset; in the third dataset, its performance is equal to the one of the Gauss2 membership function.
- In the majority of the tests, the AdaBoost and Random Forest classification algorithms showed the best performance rates. On the contrary, kNN, SVM and Naive Bayes showed the poorest performance.
- The performance difference among the Bell, Gauss and Gauss2 membership function is very low and they can be considered as efficient alternatives.

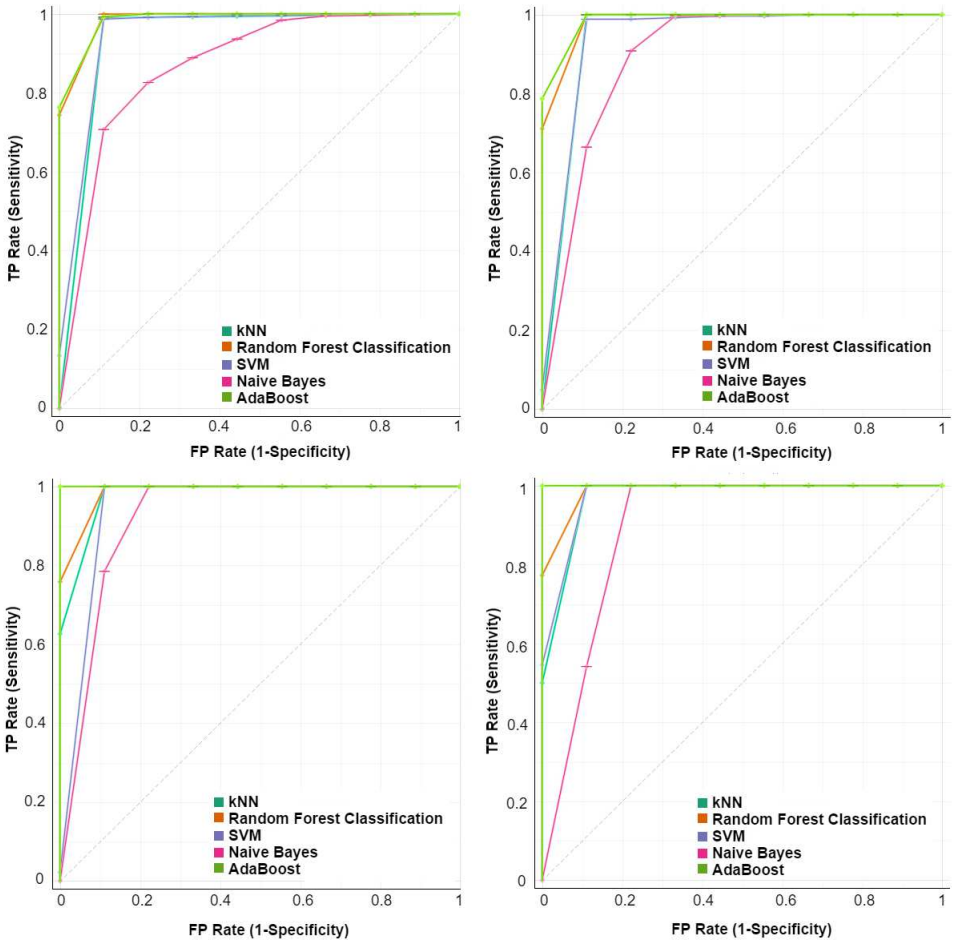

Fig. 3. ROC curves for the Dev.3 dataset

Figure 2 summarizes the aforementioned claims by depicting the average accuracy of the fuzzy systems per dataset, membership function and classification algorithm. The overall better suitability of the Bell, Gauss and Gauss2 membership functions is observable by the equivalent curves. Finally, Figure 3 depicts the Receiver Operating Characteristic (ROC) Curves for four out of the five suspiciousness values of Equation 4 for the Dev.3 dataset and the Bell membership function in the following order: (upper part: S=0.25, S=0.5; bottom part: S=0.75, S=1). The effectiveness of the system is significantly higher for the higher suspiciousness values.

6 Conclusions

The evaluation procedure of the proposed methodology was concluded successfully. The most appropriate parameters for the fuzzy systems were selected and the detection of potentially suspicious patterns was rather successful, with a small number of missclassified patterns. Despite the satisfactory results, the aforementioned procedure revealed the need for a mechanism that will be able to optimize the parameters of a fuzzy system, so as to replicate the proposed methodology and achieve the replacement of trial and error methods by automatic parameterization.

One of the biggest advantages of the method used in the current paper is that the fuzzy systems can provide adequate results without the need of directly accessing sensitive data, such as recipient identities. Testing the fuzzy systems in many different datasets and obtaining the same type of results is an encouraging factor towards the proof of their suitability for detection of behavioural-based criminal activity.

Moreover, there are some points that need to be taken into consideration and to be examined more extensively. Accessing actual data concerning the use case circumstances would be the best approach for evaluating the fuzzy systems' efficiency. Moreover, there is a considerable probability that the fuzzy systems will behave in a different way for shrunk or extended ranges of values (a very high span of appearance frequency rates), a characteristic that is in need of generalization and proper adjustment of parameters. The upcoming stage of the authors' work comprises the experimentation with different data types and the development of an appropriate Neuro-Fuzzy network that will co-operate with the fuzzy systems, tune their existing parts, such as variable ranges and membership functions and aims to complete the current contribution.

Acknowledgments

This work was partially funded by the ATENA H2020 EU Project (H2020-DS-2015-1 Project 700581). We also thank the team of FP7 Project SALUS (Security and interoperability in next generation PPDR communication infrastructures) for the fruitful discussions and feedback and the GEPTD officer Nikolaos Bouzis for the insights on common in-field investigation practices.

References

1. Barmpatsalou, K., Damopoulos, D., Kambourakis, G., Katos, V.: A critical review of 7 years of Mobile Device Forensics. Digital Investigation 10, 4, 323–349 (2013)
2. Gegov, A.: Fuzzy Networks for Complex Systems: A Modular Rule Base Approach. Springer Berlin Heidelberg (2011)
3. Kar, S., Das, S., Ghosh, P.K.: Applications of neuro-fuzzy systems: A brief review and future outline. Applied Soft Computing 15, 243-259 (2014)

4. Siddique, N., Adeli, H.: Computational intelligence: synergies of fuzzy logic, neural networks and evolutionary computing. John Wiley & Sons (2013)
5. Balasubramaniyan, J., Garcia-Fernandez, J., Spafford, E., Zamboni, D.: An architecture for intrusion detection using autonomous agents. Technical report, COAST Laboratory, Purdue University (1998)
6. Dickerson, J.E., Dickerson, J.A.: Fuzzy network profiling for intrusion detection. In: PeachFuzz 2000. 19th International Conference of the North American Fuzzy Information Processing Society - NAFIPS (Cat. No.00TH8500), pp. 301–306 (2000)
7. Chavan, S., Shah, K., Dave, N., Mukherjee, S., Abraham, A., Sanyal, S.: Adaptive neuro-fuzzy intrusion detection systems. In: International Conference on Information Technology: Coding and Computing - ITCC 2004, vol. 1, pp. 70-74 (2004)
8. MIT Lincoln Laboratory: DARPA intrusion detection evaluation. Technical report, MIT Lincoln Laboratory (1998)
9. Stoffel, K., Cotofrei, P., Han, D: Fuzzy methods for forensic data analysis. In: 2010 International Conference of Soft Computing and Pattern Recognition, pp. 23-28 (2010)
10. Islam, M.,Verma, V.: Fuzzy logic based risk model for SMS threats in 3G systems. International Journal of Advanced Research in Computer Science and Software Engineering 2, 2, 265-276 (2012)
11. Arun Raj Kumar, P., Selvakumar, S.: Detection of distributed denial of service attacks using an ensemble of adaptive and hybrid neuro-fuzzy systems. Comput. Commun. 36, 3, 303-319 (2013)
12. Shalaginov, A., Franke, K.: Automatic rule-mining for malware detection employing neuro-fuzzy approach. In: Norsk informasjonssikkerhetskonferanse (NISK) 2013, (2013)
13. Nyanyo, A., Marques, H., Wickson, P., Brouwer, F., Blaha, M., Jelenc, D., Brouet, J., Junittila, K., Kolundzija, B.: Deliverable 2.3: SALUS use cases final. Technical report, SALUS Consortium (2014)
14. SALUS Consortium: SALUS security and interoperability in next generation PPDR communication infrastructures, https://www.sec-salus.eu
15. Wagner, D.T., Rice, A., Beresford, A.R.: Device analyzer: Understanding smartphone usage. In: Mobile and Ubiquitous Systems: Computing, Networking, and Services: 10th International Conference, MOBIQUITOUS 2013, Tokyo, Japan, December 2-4, 2013, Revised Selected Papers (Cham, 2014), I. Stojmenovic and S. Cheng, Zixueand Guo, (eds.), Springer International Publishing, pp. 195-208 (2014)
16. Fuller, R: Neural fuzzy systems. Abo (1995)
17. Mamdani, E.H.: Application of fuzzy algorithms for control of simple dynamic plant. In: Proceedings of the Institution of 121 Electrical Engineers, vol. 12, pp. 1585–1588 (1974)
18. de Lima, H.P., de Arruda Camargo, H.: A methodology for building fuzzy rule-based systems integrating expert and data knowledge. In: 2014 Brazilian Conference on Intelligent Systems, pp. 300-305 (2014)
19. Guillaume, S., Charnomordic, B.: Fuzzy inference systems: An integrated modeling environment for collaboration between expert knowledge and data using FISPRO. Expert Systems with Applications 39, 10, 8744–8755 (2012)
20. Gacti, M., Alcala, R., Herrera, F.: Interpretability of linguistic fuzzy rule-based systems: An overview of interpretability measures. Special Issue on Interpretable Fuzzy Systems. Information Sciences 181, 20, 4340–4360 (2011)
21. Casillas, J., Cordon, O., Jesus, M.D., Herrera, F.: Genetic feature selection in a fuzzy rule-based classification system learning process for high-dimensional prob-

lems. Information Sciences. Recent Advances in Genetic Fuzzy Systems 136, 14, 135 157 (2001)
22. Chandrashekar, G., Sahin, F.: A survey on feature selection methods. Computers & Electrical Engineering 40, 1, 16–28 (2014)
23. Japkowicz, N., Shah, M.: Evaluating Learning Algorithms: A Classification Perspective. Cambridge University Press (2011)

Appendix A SMS Datasets Evaluation Metrics

The appendix contains the analytical metrics for all the datasets tested in Section 5 as supplementary resources. Table 3 corresponds to the dataset of the second device (Dev.2), whereas Table 4 refers to the dataset of the third device (Dev.3).

M.F.	Algorithm	AUC	Accuracy	Precision	Recall	FPR
Triangular	kNN	0.888	0.864	0.885	0.864	0.045
	SVM	0.875	0.822	0.840	0.822	0.052
	Naive Bayes	0.791	0.740	0.691	0.740	0.078
	AdaBoost	**0.897**	**0.850**	**0.870**	**0.850**	**0.043**
	Random Forest	0.890	0.867	0.888	0.867	0.045
Trapezoidal	**kNN**	**0.801**	**0.665**	**0.850**	**0.665**	**0.082**
	SVM	0.587	0.514	0.307	0.514	0.168
	Naive Bayes	0.727	0.684	0.606	0.684	0.107
	AdaBoost	0.742	0.704	0.647	0.704	0.102
	Random Forest	0.741	0.703	0.646	0.703	0.102
Bell	kNN	0.984	0.980	0.977	0.980	0.005
	SVM	0.976	0.968	0.966	0.968	0.008
	Naive Bayes	0.846	0.809	0.743	0.809	0.054
	AdaBoost	**0.998**	**0.997**	**0.997**	**0.997**	**0.001**
	Random Forest	0.991	0.989	0.986	0.989	0.004
Gauss	kNN	0.987	0.984	0.982	0.984	0.004
	SVM	0.980	0.972	0.9709	0.972	0.007
	Naive Bayes	0.850	0.815	0.746	0.815	0.052
	AdaBoost	**0.995**	**0.994**	**0.991**	**0.994**	**0.001**
	Random Forest	0.991	0.989	0.986	0.989	0.002
Gauss2	kNN	0.986	0.983	0.981	0.983	0.004
	SVM	0.988	0.984	0.982	0.984	0.003
	Naive Bayes	0.880	0.848	0.781	0.848	0.040
	AdaBoost	**0.989**	**0.986**	**0.983**	**0.986**	**0.003**
	Random Forest	0.988	0.984	0.982	0.984	0.003

Table 3. Evaluation metrics per membership function for the SMS Dev. 2 dataset

M.F.	Algorithm	AUC	Accuracy	Precision	Recall	FPR
Triangular	kNN	0.619	0.310	0.857	0.310	0.158
	SVM	0.611	0.582	0.508	0.582	0.159
	Naive Bayes	0.604	0.573	0.365	0.573	0.160
	AdaBoost	**0.617**	**0.591**	**0.651**	**0.591**	**0.156**
	Random Forest	0.617	0.590	0.610	0.590	0.157
Trapezoidal	kNN	0.608	0.294	0.571	0.294	0.143
	SVM	0.609	0.294	0.571	0.294	0.143
	Naive Bayes	0.600	0.571	0.365	0.571	0.162
	AdaBoost	**0.606**	**0.579**	**0.371**	**0.579**	**0.160**
	Random Forest	0.605	0.578	0.371	0.579	0.161
Bell	kNN	0.971	0.963	0.963	0.962	0.010
	SVM	0.937	0.906	0.922	0.906	0.025
	Naive Bayes	0.722	0.682	0.527	0.682	0.102
	AdaBoost	**0.990**	**0.986**	**0.986**	**0.986**	**0.004**
	Random Forest	0.983	0.978	0.978	0.978	0.033
Gauss	kNN	0.979	0.971	0.972	0.971	0.008
	SVM	0.940	0.909	0.975	0.975	0.025
	Naive Bayes	0.713	0.666	0.519	0.666	0.191
	AdaBoost	**0.990**	**0.986**	**0.986**	**0.986**	**0.006**
	Random Forest	0.981	0.975	0.975	0.975	0.006
Gauss2	**kNN**	**0.975**	**0.967**	**0.968**	**0.967**	**0.009**
	SVM	0.944	0.915	0.931	0.915	0.023
	Naive Bayes	0.716	0.671	0.521	0.671	0.108
	AdaBoost	0.949	0.920	0.935	0.920	0.022
	Random Forest	0.946	0.917	0.932	0.917	0.022

Table 4. Evaluation metrics per membership function for the SMS Dev. 3 dataset