Table of Contents

International Journal of Digital Crime and Forensics

Volume 9 • Issue 1 • January-March-2017 • ISSN: 1941-6210 • eISSN: 1941-6229

An official publication of the Information Resources Management Association

Research Articles

COPYRIGHT

The *International Journal of Digital Crime and Forensics* is indexed or listed in the following: ACM Digital Library; Applied Social Sciences Index & Abstracts (ASSIA); Bacon's Media Directory; Cabell's Directories; Compendex (Elsevier Engineering Index); DBLP; GetCited; Google Scholar; INSPEC; JournalTOCs; Library & Information Science Abstracts (LISA); MediaFinder; Norwegian Social Science Data Services (NSD); SCOPUS; The Index of Information Systems Journals; The Standard Periodical Directory; Ulrich's Periodicals Directory; Web of Science Emerging Sources Citation Index (ESCI)

International Journal of Digital Crime and Forensics
Volume 9 • Issue 1 • January-March 2017

Age Detection Through Keystroke Dynamics from User Authentication Failures

Ioannis Tsimperidis, Democritus University of Thrace, Komotini, Greece
Shahin Rostami, Bournemouth University, Poole, UK
Vasilios Katos, Bournemouth University, Poole, UK

ABSTRACT

In this paper an incident response approach is proposed for handling detections of authentication failures in systems that employ dynamic biometric authentication and more specifically keystroke user recognition. The main component of the approach is a multi layer perceptron focusing on the age classification of a user. Empirical findings show that the classifier can detect the age of the subject with a probability that is far from the uniform random distribution, making the proposed method suitable for providing supporting yet circumstantial evidence during e-discovery.

KEYWORDS

Age Classifier, Age Detection, Biometrics, Multi Layer Perceptron, Offender Profiling

INTRODUCTION

In today's modern digital communications, user authentication failures may lead to a series of incidents where if a user's identity is misrepresented, it could cause, among other issues, false accusations, leading to misconduct of justice. In its simplest form, consider a scenario of an unattended laptop where an adversary sends an insulting email by spoofing the legitimate owner. If traditional forensics are used, the laptop could be identified, seized and analysed to prove that the particular device was the email's origin. However, in digital forensics it is of utmost importance to "put fingers on keyboard", that is, to identify the physical person that performed the action. Another scenario relates to systems that offer seamless, continuous, and transparent biometric authentication (Flior and Kazimierz, 2010). In a typical setting, once authentication fails, the user – hopefully the perpetrator and not the legitimate user, as can happen in many cases due to the Type I errors such authentication systems exhibit – would be denied access. In our proposal, if user authentication fails at some point in time, the system should have the ability to record the user's typing behaviour in order to gather his/her characteristics such as gender, handedness, age, and so forth. As this is acknowledged to be a challenging problem (Uludag et al., 2004) with relatively poor success rates, we consider the use of epistemic uncertainty handling frameworks (Katos and Bednar, 2008). More specifically, in this paper we propose a system that is based on keystroke duration measurements in order to provide circumstantial evidence for supporting a forensic analyst to make an informed judgement on the identification of a suspect. Clearly, as it is very common in forensic science practice, the information gathered from the proposed system would need to be correlated with other sources (such as GSM geolocation data (Lessard and Kessler, 2010), security cameras, as well as physical evidence if applicable), in order to allow further reduction of the

DOI: 10.4018/IJDCF.2017010101

uncertainty surrounding the crime scene. Accepting that the evidence is circumstantial, we can relax the requirements of high success rates, but nevertheless, it is expected that the proposed solution will provide results that will be significantly diverging from a uniform distribution. The main contribution and novelty of this paper is on the application of neural networks for improving forensic readiness by providing indication and circumstantial evidence on an author's age. An age detection component has a number of applications, both as a proactive security measure (e.g. alerting a minor in the case they believe they are chatting on a social networking site with a young person whereas they may be chatting with a potential predator), following thus the security by design approach, as well as a post mortem tool for performing attribution on a potentially offensive content.

The rest of the paper is organized as follows. A survey of the literature is presented followed by the description of the data acquisition method and approach to parameter extraction required to perform keystroke dynamics. The paper continues by developing the classifier employed in the proposed method and the empirical evaluation.

RELATED WORK

Artificial Neural networks were used in the past to classify computer users or computer items into some categories. For instance, Clark et al. (2003) presented an artificial neural network based system for automated e-mail filing into folders and anti-spam filtering. Nogueira et al (2005) and Auld et al (2007) dealt with Internet traffic. More specifically, Noguiera et al. (2005) proposed the classification of Internet users into groups according to their average transfer rate. Auld et al. (2007) classified flows based on header-derived statistics, and this is feasible even when the IP (host) address and application port number are not known.

A number of diverse approaches on artificial neural networks for age classification exist in the literature. Hewahi et al. (2010) for instance proposed neural network classifiers to determine the age using image analysis techniques. They examined 4 age classes, each of them divided into two subcategories, used 68 landmark points taken from face images and achieved a success rate of 78.4%. Similarly, Choobeh (2012) used the same number of landmark points, applied 130 attributes to each of the artificial neural network classifiers that involved and calculated the mean absolute error of exact age between 4.85 and 5.85 years. Yi et al (2014) reduced the mean absolute error further to 3.63 years, which is an acceptable value to supply a more stable age estimator for practical applications.

Determining a blog's author age group is another perspective in this field. Schler et al. (2006) divided users into three age classes and based on the authoring style, but also the content of the blogs, achieved a classification success rate of about 76%. Rangel et al (2013) presented the results of the first International Author Profiling Task in which, researchers examined blogs written in two different languages: English and Spanish. Several different features were selected and the best approach achieved a classification accuracy of 64%.

The wide expansion of social networks has motivated the researchers to develop techniques for determining a user's attributes when these are not disclosed on their personal profile. Rao et al (2010) proposed an approach to automatically discover a number of user attributes by examining their status messages, the social network structure and communication behaviour of the users. They created a data set from 2000 users, who were divided into two classes, users who are under 30 and users who are over 30, used a Support Vector Machine, SVM, classifier and attained a classification accuracy rate of about 74%. The work by Pennacchiotti and Popescu (2011) and Bergsma et al. (2013), also dealt with social networks. In this case parameters like linguistic content and contacts between users were considered and used to predict ethnicity, political affiliation and origin.

This paper builds upon the concept of user classification using keystroke dynamics. The underpinnings of the proposed approach have roots to the Halstead-Reitan Battery, which is a sequence of neuropsychological tests that include a finger-tapping test (Strauss et al., 2006). This test measures the rate at which a subject repeatedly pushes a button. Empirical results show correlation of this simple task with a subject's age, gender, dominant hand, overall health, and even whether he or she is putting maximal effort into the task. Since using a keyboard is a lot like pushing a button repeatedly, perhaps these same traits manifest in different typing rhythms.

Initially, keystroke dynamics were used for user authentication in order to replace the obsolete way of using passwords as a login mechanism. For example, Joyce and Gupta (1990) used digram latencies to distinguish a legitimate user from an impostor and the results showed that there were very few times that the impostor was able to successfully authenticate. Similarly, Monrose and Rubin (2000) collected data from 63 users, developed a toolkit for analysing the data, used 3 classifiers, and achieved classification accuracy between 83% and 94% depending on the approach being used.

User authentication regards a specific number of users, usually a limited number of them, with "known" typing patterns, that is the authentication system has prior knowledge of the users' typing behaviour through some training phase. Therefore, this approach can't be applied in the case of the Internet user base, where no data is available. However, in Tsimperidis and Katos (2013), users were classified according to a device that used for writing a text, keystroke dynamics can be employed for user classification at about 75% classification accuracy. Furthermore, Tsimperidis et al. (2015) showed that it is possible to predict the gender of the author of a text from the way he or she types and not from the content, with a classification accuracy of approximately 75%.

Finally, Idrus et al. (2013) in a research similar to this paper attempt to extract information from keystroke dynamics templates with the ability to recognise, amongst others, the age category of a user when he or she types a given password or passphrase on a keyboard. They collected the data from 110 users who were asked to type 5 phrases of a length between 17 to 24 characters, using an SVM classifier, separated the users into two age classes, those who were under 30 and those who were over 30, and achieved a classification accuracy between 65% and 82%.

High-Level Description of the Proposed Incident Response System

In the context of this paper user classification is defined as the process in which a person who uses an electronic device is recognised and placed into a group of particular attributes. Typical attributes include the handedness, gender, the input device (laptop, desktop, etc.), and of course, age.

Consider a biometric authentication system based on keystroke authentication. Such system will have a database containing all distinct typing patterns used for the continuous and transparent authentication of the user, as analysed and studied in the current research literature. Consider our example scenario where a user leaves her workstation and a perpetrator gains control of the keyboard. In addition, we assume that the perpetrator is not a user with registered attributes on the authentication database since in this case the system would be in the position to identify the perpetrator. Assuming that the biometric authentication system detects the fact that a different, unknown user is typing on the keyboard, there are two potential alternatives:

1. The authorisation system locks the user out;
2. The authorisation system adopts a honeypot approach by allowing the user to continue typing in order to collect circumstantial evidence in order to be used later on for the revelation of the identity of the perpetrator.

The proposed work focuses on option 2. Naturally, the exact response of the authorization system would depend upon a number of factors, including the underlying security policy and the associated risks of allowing an adversary to have access to a production system; such aspect is outside the scope of this work as it relates to different research priorities and objectives.

A state diagram of the proposed incident response system is shown in Figure 1. The major states of the system are training, "normal" operation (i.e. legitimate use of system) and incident response mode, where the system has detected that a physical user who is not an authorised member of the system has gained access to the keyboard.

DATA COLLECTION AND ATTRIBUTES EXTRACTION

Data acquisition for a keystroke dynamics analysis type of research is typically performed by deploying a keylogger on a volunteer's computer device. The volunteer may either be requested to type a specific, fixed text, or free text. The latter is preferred as it integrates better with the subject's regular typing activities, it is less intrusive and allows a larger amount of data to be collected.

For the reasons listed above, a free text keylogger - called "IRecU" - was designed and developed. This can be installed in desktops, laptops, tablets and smartphones with any version of MS Windows. During the first execution of the keylogger, among other attributes, volunteers had to provide information of the age group they belong to. For this project four age groups were chosen as follows: 18-25, 26-35, 36-45, and 46+.

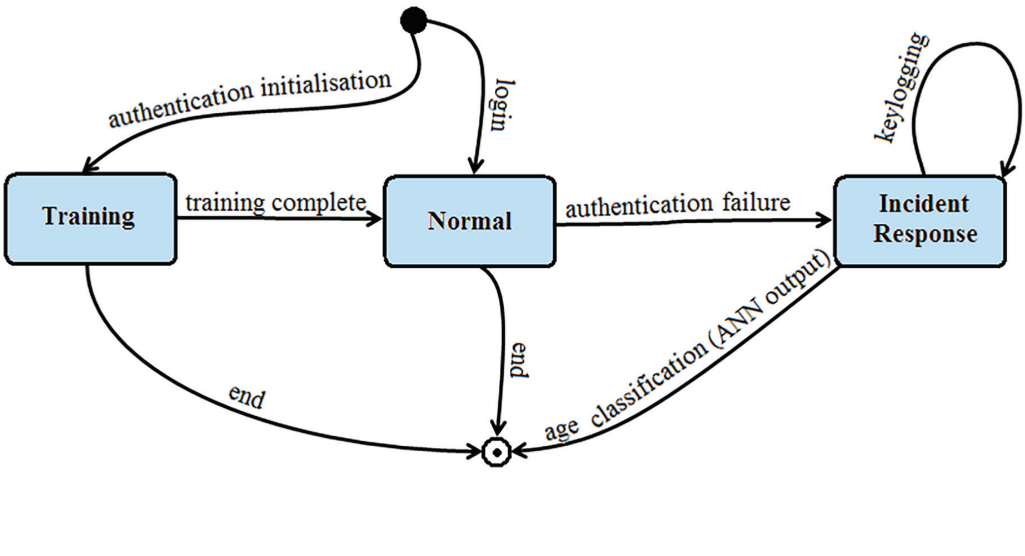
The continuous recording of a volunteer's typing during over a prolonged period of time inevitably introduces risks, due to the possibility of disclosing passwords and reading personal messages by a third person. To remedy this, the volunteers had the clear option to execute "IRecU" whenever he or she wished. Additionally, a consent form was signed by the volunteers together with a signed declaration by the researchers in which they stated that the collected data would be used only for the purposes of this project and will not be given to any persons other than the researchers.

For each volunteer, "IRecU" program create a delimited txt file with the following format:

```
75,#2014-05-06#,44449582,"dn"
75,#2014-05-06#,44449675,"up"
65,#2014-05-06#,44449732,"dn"
76,#2014-05-06#,44449810,"dn"
65,#2014-05-06#,44449831,"up"
76,#2014-05-06#,44449907,"up"
```

Each line is a record of a volunteer action, with comma-delimited fields. The first field represents the virtual key code of the key which the volunteer pressed or released. The second field depicts the date the action took place (in the format year-month-day). The next field is the elapsed time from the beginning of that day in milliseconds, and the forth field is the action, "dn" for key-press and "up" for key-release.

Figure 1. Main states of the incident response



From this data it is possible to extract all keystroke dynamics parameters. For instance, keystroke duration is calculated from the subtraction of ms that correspond to the "up" action minus the ms that correspond to the "dn" action, for the same key. In a similar way we can calculate all digram latency representations, i.e. press-press, release-press, press-release and release-release digram latency.

"IRecU" was distributed to a number of volunteers, during the period from 20/02/2014 until 27/12/2014. At the end for the period the users returned their data in log files. The number of log files per age group is presented in Table 1.

The log files varied in size from 170 KB to 271 KB and contained data from 2800 to 4500 keystrokes.

The next step was to extract the parameters that were examined. As already established by the literature, in keystroke dynamics there are plenty of parameters that can be used to characterise a user. A system that involves all parameters, or a big selection of them, will be rather complex and will require a substantial training effort, resulting from extensive and high volume data sets which are rather impractical and infeasible to acquire. For this reason, in keystroke dynamics research, there is a parameter reduction procedure. As evidenced by Ordal et al (2005) the keystroke durations are much less a characteristic of a user than digram latencies, while Zhao (2006) observes that the usage of latencies as parameters, bring better results than any other N-gram, for N≥4. The above, in conjunction with the fact that the digrams have larger appearance frequency in a text than any others N-grams, led us to choose the digram latencies as parameters for user age classification.

If it is assumed that an average computer keyboard has 100 keys, then the number of digrams is 4950, which is still a big number of parameters to investigate. However, some of the keys are barely used and too many digrams appear sparsely. By observing the log files, it was shown that the majority of the volunteers used 120 digrams frequently. These digrams were used as parameters to perform the classification.

To extract parameters from the log files another software component was developed, namely "ISqueezeU". This software read text files created by "IRecU" and it would calculate the average value of latency.

An example of digram latencies average values extraction by "ISqueezeU" program for specific log files and digrams is shown in Table 2:

The question mark (?) denotes that at the specific log file, the specific digram appears less than 5 times and therefore the value was designated as unknown.

AN ARTIFICIAL NEURAL NETWORK CLASSIFIER

The performance of several classification algorithms was considered to identify which classifier was most suitable for the extracted data and parameters. The classifiers which achieved a success rate greater than 45% have been listed in Table 3 along with the "Best Results" for their success rate and F-measure.

Table 1. Number of log files per age group

Age Group	Files
18-25	32
26-35	102
36-45	90
46+	15
total	239

Table 2. Examples from "IsqueezeU" output

Digram	LogFile 008	LogFile 027	LogFile 056	LogFile 065	LogFile 096	LogFile 155	LogFile 235
32-77	127.89	?	537.71	481.94	762.43	709.57	404.89
37-39	545.00	485.27	?	1533.10	1135.00	687.40	?
40-38	477.20	453.14	796.00	?	707.57	?	1857.71
65-32	121.19	177.35	216.53	302.44	225.00	389.30	440.07
69-73	118.33	142.86	130.76	321.53	160.29	204.14	346.11
82-79	163.86	345.18	186.65	297.83	190.88	162.69	319.57
83-65	?	178.20	261.40	213.75	160.14	174.86	171.40
84-72	93.58	272.93	139.57	177.37	155.15	182.00	251.67

Table 3. Success rates and F-measures of the tested classifiers

Classifier	Best Results	
	Success Rate	F-Measure
OneR	45.6%	0.442
Best First Decision Tree	49.0%	0.451
C4.5 Decision Tree	50.2%	0.472
Naïve Bayes	50.2%	0.488
Naïve Bayes / Decision Tree Hybrid (Kohavi, 1996)	54.8%	0.539
Simple Logistic (LogitBoost algorithm)	55.7%	0.552
Support Vector Machine	56.5%	0.545

The term "Best Results" refers to the highest achieved classification accuracy obtained through the optimization of classifier parameters. Most of the above results were not considered satisfactory and moreover some of these classifiers which achieve relatively high success rates present problems. For example, SVM requires longer training time since the number of variables used in SVM equals the number of training data. SVM is optimality, where in an optimal model, the parameters used must be chosen carefully such that it includes the kernel selection and tuning its parameters, margin value, gamma value, etc. Such processes could be time consuming and it might cause model generalization problem (Abe, 2010). For the reasons listed above we focused on Artificial Neural Networks.

Artificial Neural Networks (ANNs) are a division of statistical learning algorithms within the field of computational intelligence, inspired by the behaviour of biological neurons located in the brain and central nervous system (Nelles, 2013; Rojas, 2013). ANNs employ a set of self-adaptive weights and biases which are adjusted by a learning algorithm to capture the highly complex and non-linear underlying models of the data which they are applied to. Due to their self-adaptive nature, ANNs can detect complex relationships between both dependent and independent variables without prior or auxiliary knowledge (Tu, 1996).

ANNs have been applied to a broad domain of pattern recognition and classification problems. In contrast to classical classification techniques, such as discriminant analysis, which require an understanding of the underlying statistical models of the system that produced the data, ANNs are instead a "black-box" technique capable of adapting to the underlying models (Rostami et al, 2015). This has resulted in the successful application of ANNs in fields such as decision support for concealed weapon detection (Zhang, 2000), where their self-adaptive characteristics, in particular in high-dimensional datasets, has overcome many of the difficulties in model building associated with conventional classification techniques such as decision trees and k-nearest neighbour algorithms (Dreiseitl and Ohno-Machado, 2002). The perceptron was one of the first ANNs created for the purpose of pattern recognition (Nelles, 2013), we have employed this approach for our experiments.

In our case, the perceptron has 120 inputs and these are the digram latencies of the 120 selected digrams. Each input, is multiplied by a weight, which is typically a number between -1 and 1. The initial weight assignment is random. Then the products summed, and the sum is applied to a function, called transfer function. Figure 2 shows this procedure.

If the inputs are denoted as x_i, for i = 1 to 120, then the output y of a perceptron can be written as explained in detail by Bishop (1995) as:

$$y = f\left(\sum_{i=1}^{120} w_i \cdot x_i\right) \tag{1}$$

However, a perceptron can only solve linearly separable problems. Our problem was more complicated, because we tried to classify users according to their age, using more than one hundred digram latencies, and it was unlikely to be divided linearly. To succeed we had to create an ANN from perceptrons consisting of multiple layers called to a Multi-Layer Perceptron (MLP), as shown in Figure 3.

Whilst a number of techniques exist for the optimisation of weights in Artificial Neural Networks (ANNs) (such as the Widroff-Hoff least mean squares algorithm and back propagation), there are no hard and fast rules for choosing the structure of an ANN - in particular for choosing both the size (in terms of number of neurons) and number of hidden layers used in the network. However, this internal structure is one of the key factors in determining the efficiency of the network and the accuracy of the classification (Rostami, 2014). A preliminary empirical study was conducted to determine the optimal number of neurons to be used in the hidden layer of our neural network. In this study, the performance of the neural network was evaluated for different configurations of neurons on the hidden layer, starting from 10 with an incremental increase up to 240 (double the number of inputs). The results (illustrated in Figures 4 and 5) indicated that 62 neurons in the hidden layer offered the most robust performance in terms of the accuracy, precision, and sensitivity of the classifier.

The transfer function *f* is chosen to have a number of properties which either enhance or simplify the network containing the neuron. Crucially, for instance, any multi-layer perceptron using a linear transfer function has an equivalent single-layer network, while a non-linear function is necessary to gain the advantages of a multi-layer network. Therefore, we chose the sigmoid function, which is written:

$$f\left(x\right) = \frac{1}{1 + e^{-x}} \tag{2}$$

It takes values between 0 and 1. When x tends to -∞, *f*(x) tends to 0, and when x tends to +∞, *f*(x) tends to 1.

In general, ANNs are used for their simplicity, good ability to handle nonlinear and multidimensional problems. They are considered to be both fast and adaptive learners, in the sense that the network changes its structure based on the available information during the training process (Mallios et al, 2011; Wang and Musa, 2014).

Figure 2. The output of a perceptron

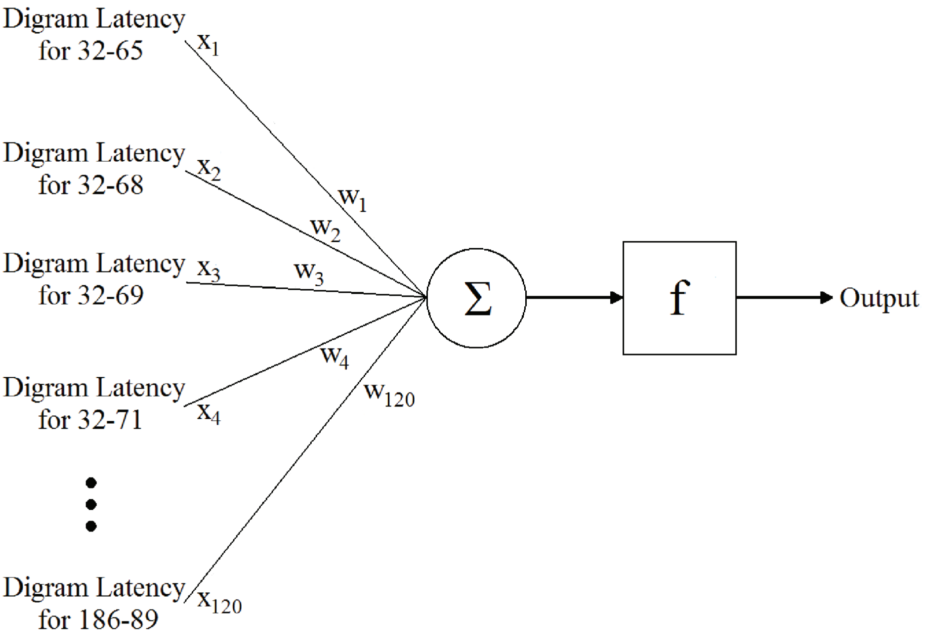


Figure 3. The multi-layer perceptron

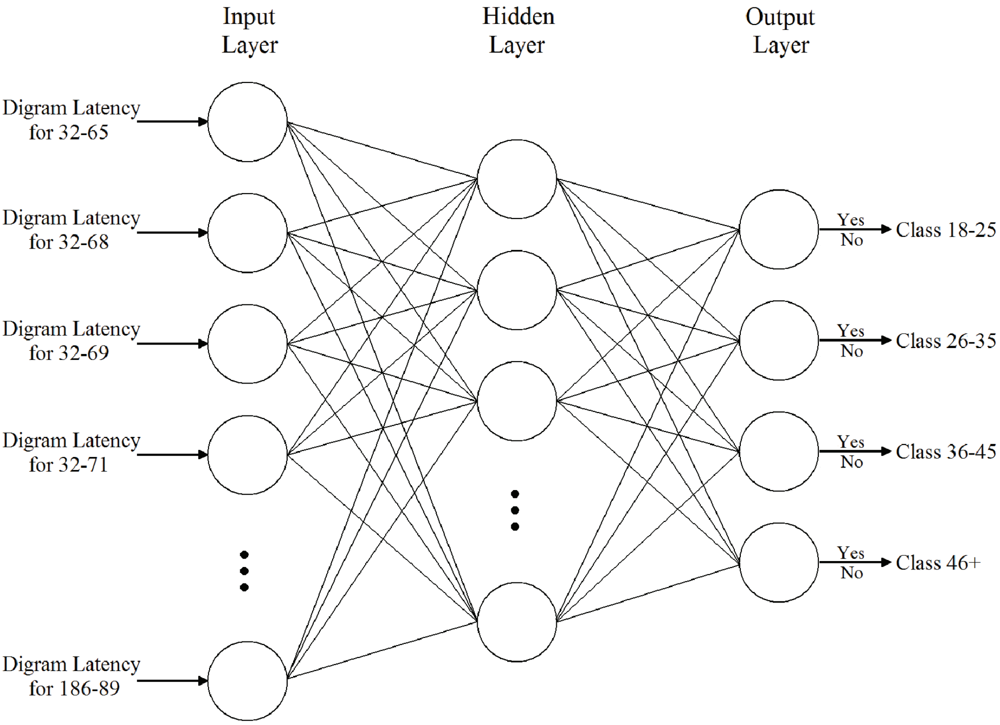


Figure 4. The standard deviation of accuracy of the system in relation to the number of neurons in hidden layer

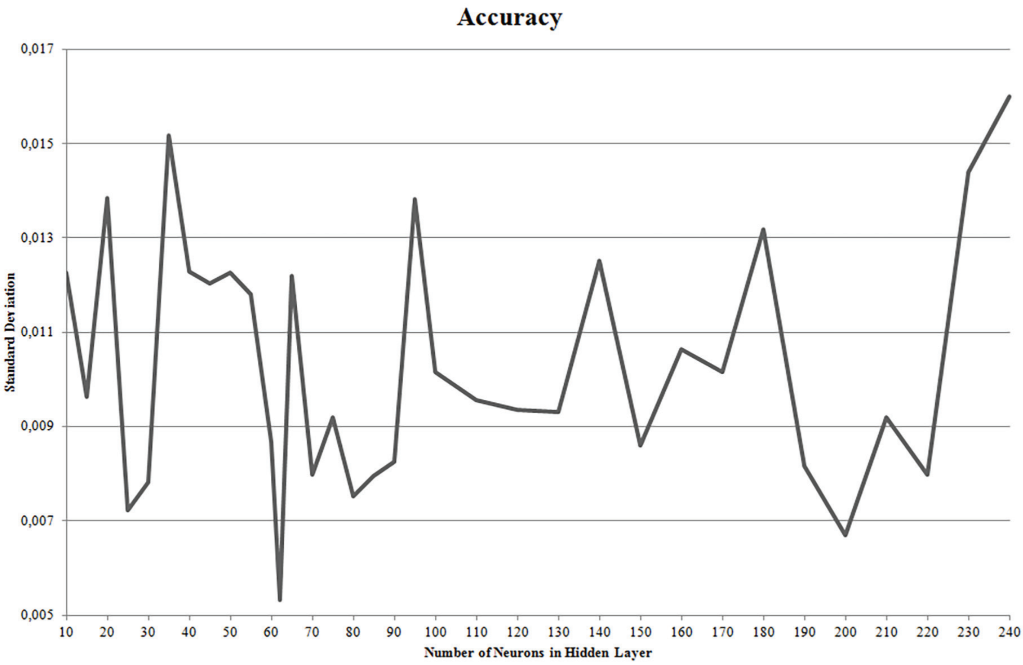
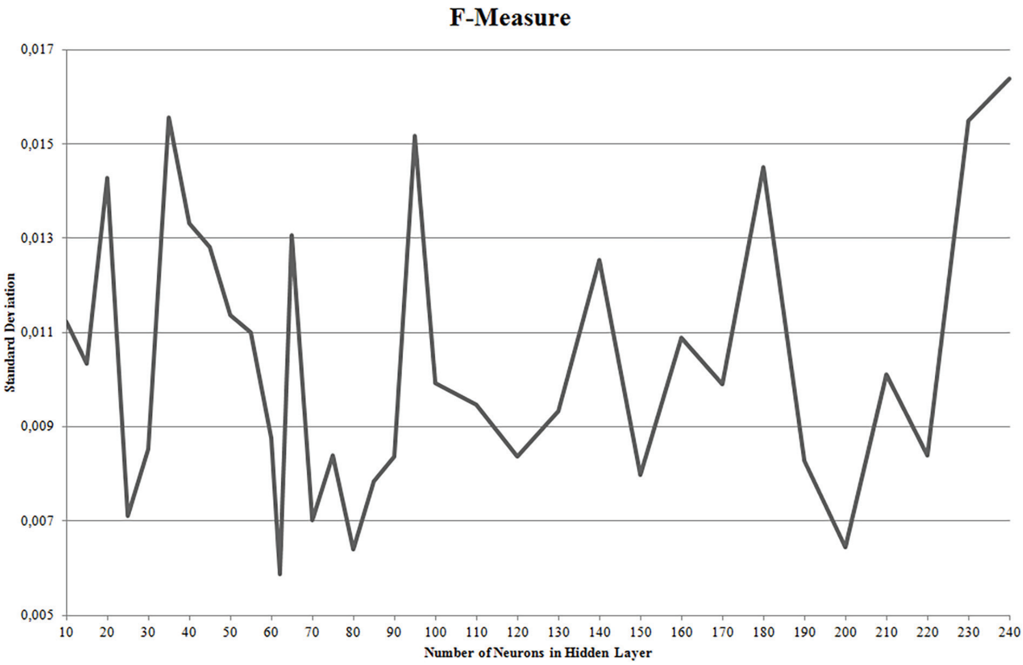


Figure 5. The standard deviation of F-measure of the system in relation to the number of neurons in hidden layer



RESULTS

Initially the stability of the proposed approach and its independence from data and classifier parameters was demonstrated. For this purpose cross validation was used. In this technique, data is split into K exclusive folds of equal size. Each part is used as a test set, with the rest used for training. The overall performance is measured by averaging the K runs results. The advantage of this method is that it matters less how the data gets divided. Every data point gets to be in a test set exactly once, and gets to be in the training set K-1 times. The variance of the resulting estimate is reduced as K is increased. As there were more than 200 samples available, we used 10 folds. Furthermore, we ran the experiment for a range of values for the neural network classifier parameters L (learning rate) and M (momentum). The learning rate ranged from 0.3 to 0.7, with a step of 0.05, and the momentum ranged from 0.2 to 0.6, with a 0.1 step. In addition the experiment was conducted 10 times, for every (L, M) pair. In every run the data were cross-validated using a 10-fold cross validation approach. The success rates for every pair of values are in Table 4.

It can be seen from the results that the success rate of the system barely escapes from 61%-62%, for a wide range of L and M values. This is a strong indication that the system is independent from data and classifier parameters.

For a more detailed view of results, we chose a specific configuration of the classifier as presented in Table 5.

The results obtained after conducting the experiment, again with the cross validation technique, are in Table 6.

As shown, 153 out of 239 files were successfully predicted, i.e. 64%, which is considerably better than the 25% uniform selection. The computed F-measure of the above experiment was 0.639.

At this point it is worthwhile mentioning that an age classification exercise is not explicit and this consequently distorts further the prediction results. For example, the separation between male and female in gender classification is clear cut, but in case of age classification a 26 year old person is more likely to have more in common with a 25 year old person than a 35 year old. However, according to the construction of classes in this paper, the 26 year old person is belongs to the same class as the latter. This is an inherent limitation of this methodology and it becomes evident in Table 6, where

Table 4. Success rates for L and M value pairs

		L								
		0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7
M	0.2	61.2%	61.3%	61.3%	61.3%	61.2%	61.3%	61.1%	61.3%	61.5%
	0.3	61.1%	61.2%	61.1%	61.3%	61.2%	61.1%	60.8%	61.1%	61.2%
	0.4	61.7%	61.5%	61.3%	61.1%	61.4%	61.3%	61.0%	61.3%	61.2%
	0.5	61.7%	61.1%	60.9%	61.4%	61.3%	61.7%	61.6%	61.0%	61.6%
	0.6	61.4%	61.4%	61.1%	61.6%	62.0%	62.1%	61.7%	61.8%	62.6%

Table 5. The multi-layer perceptron properties

Classifier Properties	
Learning Rate	0.7
Momentum	0.2
Training Time	500 ms

Table 6. Classification results for 4 classes

	Predicted as				
Age Group	**18-25**	**26-35**	**36-45**	**46+**	**Success Rate per Class**
18-25	17	8	6	1	53.1%
26-35	4	72	25	1	70.6%
36-45	9	24	56	1	62.2%
46+	0	1	6	8	53.3%
Success Rate per Prediction	56.7%	68.6%	60.2%	72.7%	

it is shown that the two outlier classes (“18-25” and “46+”) present a significant prediction rate for the neighbouring class. That is, 25% of “18-25” class log files were predicted as “26-35” and 40% of “46+” class log files were predicted as “36-45”.

Based on the results, in Table 6, it can be suggested that the proposed system can predict the age group of unknown users to a satisfactory level of accuracy. It should be noted that the proposed approach does not depend on auxiliary data such as content analysis, analysis of facial features (Schler et al., 2006; Garera and Yarowsky, 2009), or social network group memberships (Rao et al., 2010).

In order to examine the robustness of the approach, the classes were reduced to three by merging the “46+” class with the “36-45” class and the test was repeated using the same classifier. The results are shown in Table 7.

In the case of three classes, 166 out of 239 log files were successfully predicted, that is 69.5%, which is better than the 33.3% random uniform selection. The obtained F-measure value was 0.694. Once again the misclassifications were popular in the neighbouring classes. Specifically, a percentage of about 31% of “18-25” class log files were predicted that belong to “26-35” class and a percentage of about 19% of “36+” class log files were predicted that also belong to “26-35” class.

Finally, the samples of “18-25” and “26-35” classes were combined and the test was repeated for two classes this time, those who were up to 35 and those who were over 35. The results are shown in Table 8.

In the case of two classes, 173 out of 239 log files were successfully predicted. That is, the classifier produced a success rate of 72.4%, which is higher than 50% of the random, uniform selection (and with an F-measure of 0.724).

The results obtained from the execution of 4-classes, 3-classes and 2-classes experiments, are comparable to the work by Irdus et al. (2013). One significant difference is that our work uses free text data.

Table 7. Results for 3 classes

Age Group	**Predicted as**			**Success Rate per Class**
	18-25	**26-35**	**36+**	
18-25	16	10	6	50.0%
26-35	3	76	23	74.5%
36+	11	20	74	70.5%
Success Rate per Prediction	53.3%	71.7%	71.8%	

Table 8. Classification results for 2 classes

Age Group	Predicted as		Success Rate per Class
	18-35	36+	
18-35	101	33	75.4%
36+	33	72	68.6%
Success Rate per Prediction	75.4%	68.6%	

Improving the Success Rate

For the sake of completeness a number of meta-algorithms was employed in order to explore ways of improving the success rates. To this end, AdaBoost, MultiBoost, Random-correction-code and Exhaustive-correction-code were employed. The latter two were not applied in the 2-classes case, since they relate to multiclass problems.

The results obtained from these experiments are shown in Table 9. The notable success rate improvements are noted with bold and underlined font.

Figure 6 visualizes the results of Table 9, showing the difference in success rate of the system between before and after the use of meta-algorithms.

This improvement of success rate in 4-classes problem means that 21 more people in one thousand would be predicted in correct age class, giving more reliable information to forensic agents or to unsuspicious, benign users. Similarly, in the 2-classes problem, the age of 12 more people in one thousand is expected to be predicted correctly.

CONCLUSION AND OUTLOOK

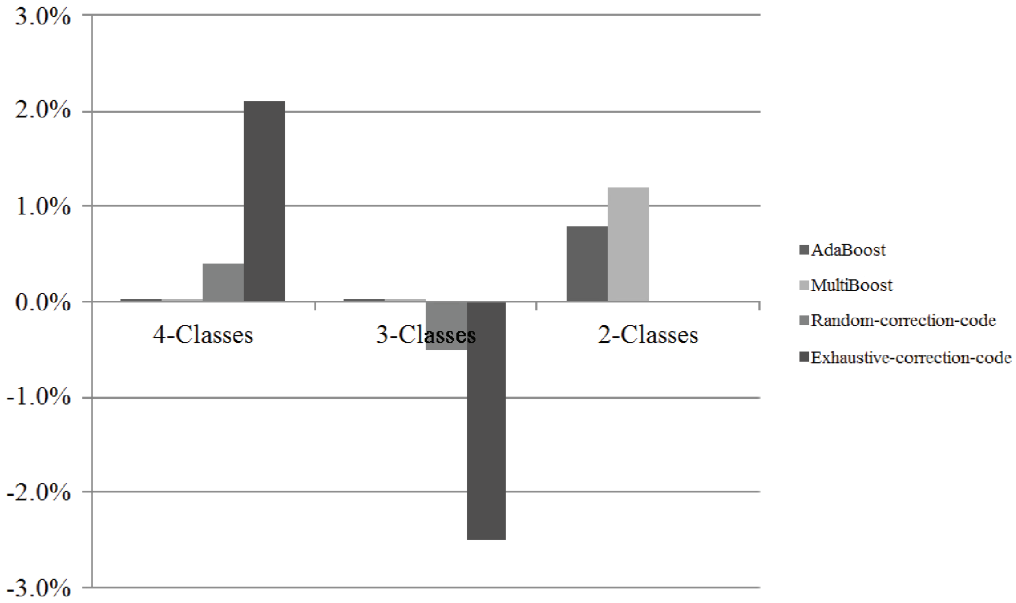
Age detection via keystroke analysis can provide circumstantial evidence for supporting the identification of a suspect in a case of account, identity theft or a masquerade attack. Moreover, the proposed approach can be used as part of a warning system for age discrepancies of users participating in online chats. This is particularly useful for detecting for example child grooming where an unsuspecting minor may have a conversation with a much older person who falsely declares to be of a younger age. Overall the ability to guess one's age apart from the offender profiling domain, may have other applications such as the validation of the information provided during a registration process.

An extension of this work involves the increased granularity of the classes, which can be done by generating a considerably bigger dataset. Furthermore, the proposed method can use multiple

Table 9. Success rate, time complexity and F-measure with and without using meta-algorithms for 4-classes, 3-classes and 2-classes' cases

	4-Classes			3-Classes			2-Classes		
	Succ. Rate	TBM	F-score	Succ. Rate	TBM	F-score	Succ. Rate	TBM	F-score
Without meta-algorithm	64,0%	43,99	0,639	69,5%	43,25	0,694	72,4%	42,94	0,724
AdaBoost	64,0%	87,92	0,639	69,5%	88,91	0,694	**73,2%**	91,58	0,732
MultiBoost	64,0%	89,36	0,639	69,5%	87,81	0,694	**73,6%**	83,94	0,737
Random-correction-code	**64,4%**	341,87	0,645	69,0%	255,81	0,692	-	-	-
Exhaustive-correction-code	**66,1%**	397,63	0,658	67,0%	127,09	0,670	-	-	-

Figure 6. Comparison between boosting meta-algorithms



classifiers, in parallel. In this case the results will need to be fed to an evidence-handling framework, such as Dempster-Shafer's theory of evidence, to allow the fusion of the classifier "views" in order to produce the final outcome.

REFERENCES

Abe, S. (2010). *Support Vector Machines for Pattern Classification*. New York: Springer. doi:10.1007/978-1-84996-098-4

Auld, T., Moore, A. W., & Gull, S. F. (2007). Bayesian neural networks for Internet traffic classification. *IEEE Transactions on Neural Networks*, *18*(1), 223–239. doi:10.1109/TNN.2006.883010 PMID:17278474

Bergsma, S., Dredze, M., Van Durme, B., Wilson, T., & Yarowsky, D. (2013). Broadly improving user classification via communication-based name and location clustering on twitter. *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*NAACL HLT '13 (pp. 1010-1019).

Bishop, C. (1995). *Neural Networks for Pattern Recognition*. Oxford: Clarendon Press.

Choobeh, A. K. (2012). Improving automatic age estimation algorithms using an efficient ensemble technique. *International Journal of Machine Learning and Computing*, *2*(2), 118–122. doi:10.7763/IJMLC.2012.V2.99

Clark, J., Koprinska, I., & Poon, J. (2003). A neural network based approach to automated e-mail classification. *Proceedings ofIEEE/WICInternational Conference on Web Intelligence* (pp. 702-705). doi:10.1109/WI.2003.1241300

Dreiseitl, S., & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: A methodology review. *Journal of Biomedical Informatics*, *35*(5), 352–359. doi:10.1016/S1532-0464(03)00034-0 PMID:12968784

Flior, E., & Kazimierz, K. (2010). Continuous biometric user authentication in online examinations.*Proceedings of 7th International Conference on Information Technology: New Generations ITNG '10*, Las Vegas, Nevada, USA (pp. 482-492). IEEE Computer Society. doi:10.1109/ITNG.2010.250

Garera, N., & Yarowsky, D. (2009). Modeling latent biographic attributes in conversational genres.*Proceedings of Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing ACL-IJCNLP '09* (Vol. 2, pp. 710-718). doi:10.3115/1690219.1690245

Hewahi, N., Olwan, A., & Tubeel, N., EL-Asar, S., & Abu-Sultan, Z. (2010). Age estimation based on neural networks using face features. *Journal of Emerging Trends in Computing and Information Sciences*, *1*(2), 61–67.

Idrus, S. Z. S., Cherrier, E., Rosenberger, C., & Bours, P. (2013) 'Soft biometrics for keystroke dynamics. *Proceedings of the International Conference on Image Analysis and Recognition ICIAR '13*, Povoa de Varzim, Portugal (pp. 11-18). doi:10.1007/978-3-642-39094-4_2

Joyce, R., & Gupta, G. (1990). Identity authorization based on keystroke latencies. *Communications of the ACM*, *33*(2), 168–176. doi:10.1145/75577.75582

Katos, V., & Bednar, P. M. (2008). A cyber-crime investigation framework. *Computer Standards & Interfaces*, *30*(4), 223–228. doi:10.1016/j.csi.2007.10.003

Kohavi, R. (1996). Scaling up the accuracy of Naive-Bayes classifiers: A Decision-Tree hybrid.*Proceedings of the Second International Conference on Knowledge Discovery and Data MiningKDD '96* (pp. 202-207).

Lessard, J., & Kessler, G. (2010). Android Forensics: Simplifying Cell Phone Examinations. *Small Scale Digital Device Forensics Journal*, *4*(1), 1–12.

Mallios, N., Papageorgiou, E., & Samarinas, M. (2011). Comparison of machine learning techniques using the WEKA environment for prostate cancer therapy plan. *Proceedings of the 20th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises WETICE '11* (pp. 151-155). doi:10.1109/WETICE.2011.28

Monrose, F., & Rubin, A. D. (2000). Keystroke dynamics as a biometric for authentication. *Future Generation Computer Systems*, *16*(4), 351–359. doi:10.1016/S0167-739X(99)00059-X

Nelles, O. (2013). *Nonlinear System identification: From Classical Approaches to Neural Networks and Fuzzy Models*. Springer Science & Business Media.

Nogueira, A., de Oliveira, M. R., Salvador, P., Valadas, R., & Pacheco, A. (2005). Classification of Internet users using discriminant analysis and neural networks.*Proceedings of the 1st Conference of Next Generation Internet Networks NGI '05*, Rome, Italy (pp. 341-348). doi:10.1109/NGI.2005.1431686

Ordal, P., Ganzhorn, D., Lu, D., Fong, W., & Norwood, J. B. (2005). Continuous identity verification through keyboard biometrics. *Journal of Undergraduate Research*, *4*(1), 20–24.

Pennacchiotti, M., & Popescu, A. (2011) 'A machine learning approach to twitter user classification. *Proceedings of the 5th International AAAI Conference on Weblogs and Social Media ICWSM '11* (pp. 281-288).

Rangel, F., Rosso, P., Koppel, M., Stamatatos, E., & Inches, G. (2013). Overview of the author profiling task at PAN 2013.*Proceedings of Conference and Labs of the Evaluation Forum PAN at CLEF '13*, Valencia, Spain.

Rao, D., Yarowsky, D., Shreevats, A., & Gupta, M. (2010). Classifying latent user attributes in twitter. *Proceedings of the 2nd International Workshop on Search and Mining User-Generated Contents SMUC '10* (pp. 710–718).

Rojas, R. (2013). *Neural Networks: A Systematic Introduction*. Springer Science & Business Media.

Rostami, S. (2014) Preference focused many-objective evolutionary computation [PhD thesis]. Manchester Metropolitan University, Manchester, United Kingdom.

Rostami, S., OReillya, D., Shenfieldb, A., & Bowringa, N. (2015). A novel preference articulation operator for the Evolutionary Multi-Objective Optimisation of classifiers in concealed weapons detection. *Information Sciences*, *295*, 494–520. doi:10.1016/j.ins.2014.10.031

Schler, J., Koppel, M., Argamon, S., & Pennebaker, J. (2006). Effects of age and gender on blogging. *Proceedings of theSymposium on Computational Approaches to Analyzing WeblogsAAAI Spring '06* (pp. 199-205).

Strauss, E., Sherman, E., & Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms and Commentary* (3rd ed.). New York: Oxford University Press.

Tsimperidis, I., & Katos, V. (2013). Keystroke forensics: Are you typing on a desktop or a laptop?*Proceedings of the 6th Balkan Conference in Informatics BCI '13*, Thessaloniki, Greece (pp. 89-94). doi:10.1145/2490257.2490272

Tsimperidis, I., Katos, V., & Clark, N. (2015). Language independent gender identification through keystroke analysis.Information and Computer Security, 23(3), 286-301.

Tu, J. V. (1996). Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. *Journal of Clinical Epidemiology*, *49*(11), 1225–1231. doi:10.1016/S0895-4356(96)00002-9 PMID:8892489

Uludag, U., Pankanti, S., Prabhakar, S., & Jain, A. K. (2004). Biometric cryptosystems: Issues and challenges. *Proceedings of the IEEE*, *92*(6), 948–960. doi:10.1109/JPROC.2004.827372

Wang, X., & Musa, A. (2014). Advances in neural network based learning. *International Journal of Machine Learning and Cybernetics*, *5*(1), 1–2. doi:10.1007/s13042-013-0220-2

Yi, D., Lei, Z., & Li, S. Z. (2014). Age estimation by multi-scale convolutional networks.*Proceedings of 12th Asian Conference of Computer Vision ACCV '14*, Singapore (pp. 144-158).

Zhang, G. P. (2000). Neural networks for classification: A survey. *IEEE Transactions on Systems, Man and Cybernetics. Part C, Applications and Reviews*, *30*(4), 451–462. doi:10.1109/5326.897072

Zhao, Y. (2006). 'Learning user keystroke patterns for authentication' – Proceeding of World Academy of Science. *Engineering and Technology*, *14*, 65–70.

Ioannis Tsimperidis received the Bachelor Degree from Aristotle University of Thessaloniki in 1997, and the MSc Degree from Democritus University of Thrace in 2002. Currently he is PhD student at the Department of Electrical and Computer Engineering of Democritus University of Thrace. His main areas of research interest are keystroke analysis, user classification and digital forensics. He works as Secondary Education teacher since 2006.

Shahin Rostrami's research interests lie within the areas of Computational Intelligence and Machine Learning, ranging from theory to their application to complex real-world problems. Recently, he has focused on the development of better techniques for Many-Objective (Evolutionary) Optimisation, Neuroevolution, and the optimisation of complex real-world engineering problems. He has also collaborated actively on research projects with researchers in the fields of physics, engineering, human computer interaction (HCI), and computer science.

Vasilios Katos is Professor and Head of Computing at Bournemouth University. Prior to his current post he was Professor of Information and Communications Systems Security at the Department of Electrical and Computer Engineering of Democritus University of Thrace in Greece, Principal Lecturer at the School of Computing at the University of Portsmouth where he participated in the development of the interdisciplinary Masters course MSc in Forensic IT. He has worked in the industry as a security consultant and expert witness in information systems security. His research interests are in information security, privacy, digital forensics and incident response.

Call for Articles

International Journal of Digital Crime and Forensics

Volume 9 • Issue 1 • January-March 2017 • ISSN: 1941-6210 • eISSN: 1941-6229
An official publication of the Information Resources Management Association

MISSION

The mission of the **International Journal of Digital Crime and Forensics (IJDCF)** is to provide and foster a forum for advancing research and development of the theory and practice of digital crime prevention and forensics. IJDCF addresses a broad range of digital crimes and forensic disciplines that use electronic devices and software for crime prevention and investigation. This journal informs a broad cross-sectional and multi-disciplinary readership ranging from the academic and professional research communities, to industry consultants and practitioners. IJDCF publishes a balanced mix of high quality theoretical and empirical research articles, case studies, book reviews, tutorials, and editorials.

COVERAGE AND MAJOR TOPICS

The topics of interest in this journal include, but are not limited to:

Computational approaches to digital crime preventions • Computer virology • Crime scene imaging • Criminal investigative criteria and standard of procedure on computer crime • Cryptological techniques and tools for crime investigation • Data carving and recovery • Digital document examination • Digital Evidence • Digital signal processing techniques for crime investigations • Identity theft and biometrics • Information Warfare • Machine learning, data mining, and information retrieval for crime prevention and forensics • Malicious codes • Network access control and intrusion detection • Policy, standards, protocols, accreditation, certification, and ethical issues related to digital crime and forensics • Practical case studies and reports, legislative developments, and limitations, law enforcement • Small digital device forensics (cell phones, smartphone, PDAs, audio/video devices, cameras, flash drives, gaming devices, GPS devices, etc.) • Steganography and steganalysis • Terrorism knowledge portals and databases • Terrorism related analytical methodologies and software tools • Terrorist incident chronology databases • Watermarking for digital forensics

ALL INQUIRIES REGARDING IJDCF SHOULD BE DIRECTED TO THE ATTENTION OF:

Wei Qi Yan, Editor-in-Chief • IJDCF@igi-global.com

ALL MANUSCRIPT SUBMISSIONS TO IJDCF SHOULD BE SENT THROUGH THE ONLINE SUBMISSION SYSTEM:

http://www.igi-global.com/authorseditors/titlesubmission/newproject.aspx

IDEAS FOR SPECIAL THEME ISSUES MAY BE SUBMITTED TO THE EDITOR(S)-IN-CHIEF

PLEASE RECOMMEND THIS PUBLICATION TO YOUR LIBRARIAN

For a convenient easy-to-use library recommendation form, please visit:
http://www.igi-global.com/IJDCF