2016

A hybrid behaviour recognition and intrusion detection method for mobile devices

Ashley Woodiss-Field
School of Science, Edith Cowan University

DOI: 10.4225/75/58a6a8f60feea
Woodiss-Field, A. (2016). A hybrid behaviour recognition and intrusion detection method for mobile devices. In Johnstone, M. (Ed.). (2016). *The Proceedings of 14th Australian Information Security Management Conference, 5-6 December, 2016, Edith Cowan University, Perth, Western Australia.* (pp.37-47).

A HYBRID BEHAVIOUR RECOGNITION AND INTRUSION DETECTION METHOD FOR MOBILE DEVICES

Ashley Woodiss-Field
School of Science, Edith Cowan University, Perth, Australia
awoodiss@our.ecu.edu.au

Abstract

Behaviour-based authorisation is a technique that assesses the user of a device for authenticity by comparing their activities to previously established behaviour profiles. Passwords and other point of entry authorisation techniques are often inadequate for protecting mobile device security as they only provide an initial barrier to usage and do not operate continuously. Behaviour-based authorisation continuously assesses user authorisation, using the device owner's profile for authentication. This research improves upon behaviour-based authorisation performance by applying a hybridised intrusion detection method. The constituent intrusion detection methods that were applied include context-awareness and self-correction. Performance of a behaviour-based authorisation method can be measured as either an increase in intrusion detection, without significantly increasing false positives or a decrease in false positives without totally compromising intrusion detection. This research found that an increase in performance can be achieved by the addition of intrusion detection components.

Keywords
Behaviour-based authorisation, intrusion detection, context-awareness, self-correction, hybrid intrusion detection, artificial neural network, mobile security

INTRODUCTION

Mobile devices facilitate a range of multimedia applications, many of which are driven by user data. This user data has become sensitive in nature, including personal information, login credentials, and business data (Li, Clarke, Papadaki, & Dowland, 2013). Mobile device theft rates have been recorded:

- Approximately 1.6% of mobile phone owners experienced mobile phone theft in England and Wales during 2012/2013 (Home-Office, 2014).
- Collected US law enforcement data combined with FBI crime data indicates an estimate of 1/10th of all theft for 2013 in the US is associated with the theft of a mobile device. According to a consumer report's survey, 3.1 million smartphone thefts occurred in 2013 (Federal-Communications-Commission, 2014; Tapellini, 2014).

Credent (cited by Li et al., 2013) found that 40% of participants in a survey failed to utilise personal identification numbers (PIN). PIN-based approaches are also often misused when they are weak, rarely changed, or shared with others. A fundamental weakness is that it does not seek to validate a user once they know the PIN (Li et al., 2013).

Behaviour-based authorisation is a technique that assesses the user of a device for authenticity by comparing their activities to previously established behaviour profiles. A benefit of behaviour-based authorisation is active authentication wherein if an intruder somehow bypassed standard point of entry techniques, such as PINs or passwords, they will still be under scrutiny (Li et al., 2013).

The research goal was to ascertain whether or not a combined intrusion detection method can be applied to behaviour-based authorisation on mobile devices, with the purpose of performance improvement. The performance improvement was measured by an increase in intrusion detection accuracy without significantly increasing false positives or a reduction of false positives without totally compromising the intrusion detection rate. Intrusion detection methods that were experimented with include contextual-awareness, adapting a system's assessment metrics based on the apparent situation, and self-correction, altering assessment results through internal determination based on known information without additional external assistance.

RESEARCH BACKGROUND

Approaches to Behavioural Profiling and Authorisation

Multiple approaches to behaviour profiling, including such approaches for the purpose of authorisation, have been made. These approaches include different types of artificial neural networks, rule-based methods, specially designed frameworks, and clustering methods.

Artificial Neural Networks

Artificial neural networks (ANNs) are dynamic systems of interconnected parts. They are made up of artificial neurons which each take in a set of inputs. Inputs are regulated by weight values and activation functions. A training algorithm is applied to an ANN and the weights for all inputs are adjusted to improve the accuracy of the output. ANNs are able to detect non-linear relationships and interactions among variables (Anandarajan, 2002; Wasserman, 1989b).

Li et al. (2013) experimented a radial basis function (RBF) ANN. The activation functions in an RBF ANN are RBFs, which are used to determine the Euclidean distance of the inputs from certain points, known as centres. Using this, the mapping relationship between the input and desired output can be obtained (Chang, Cheng, & Yu-Feng, 2010; Li et al., 2013).

Li et al. (2013) also ran experiments with a feed-forward multi-layered perceptron (FF MLP) network. FF MLP ANNs pass values in one direction, from input to output, typically using the sigmoid function or hyperbolic tangent function as their activation function. The activation functions scale the outputs of each neuron which are passed to further neurons in the network (Anandarajan, 2002; Li et al., 2013; Wasserman, 1989b).

A simple recurrent network (SRN) was experimented with by Anandarajan (2002) for the purpose of classifying user behaviour in terms of workplace internet activity. An SRN has a small amount of additional neurons in the input layer, known as context units that receive feedback signals from the hidden layers, allowing the ANN to record the outputs of the hidden layers. This allows for past values to influence present values going into the ANN meaning that the ANN can learn faster than a standard feed-forward ANN (Anandarajan, 2002; Krenker, Kos, & Bešter, 2011).

It was found that ANN based systems, particularly RBF ANNs, would deliver a greater performance in terms of correctly profiling user behaviour, at the cost of being more resource intensive, using twice the amount of computational power of that of a rule-based method (Li et al., 2013).

Rule-Based and Fuzzy Systems

A dynamic rule-based approach was devised by Li et al. (2013) based on the idea that recent historical usage can be used to predict the probability of a present event. As shown in Equation 1 the approach would provide a mechanism to have all recorded outputs bounded between 0 and 1 to help determine a threshold. If a singular output breaches the threshold, then the event is deemed to have breached the rule-set (Li et al., 2013).

$$1 - \frac{\sum_{i=1}^{N}\left(\frac{O_{ix}}{\sum_{x}^{M} O_{ix}}\right)}{N} \geq t(1)$$

Equation 1 Where: i = the features of one chosen application, x = the value of $Feature_i$, M = total number of values for $Feature_i$ N = total number of features, t = predefined Threshold, O_{ix} = feature occurrence (Li et al., 2013).

Sood, Mehmi, and Dogra (2015) conducted experimentation with a fuzzy system applying a user profiling system for cloud computing. Fuzzy systems work by representing data in degrees rather than as booleans in a linguistic way. Fuzzy systems are ideal for solutions where the problem is non-linear or hard to define. Fuzzy logic is rule-based and determines what degree values are in relation to their variables. Variables that have had their fuzzy values determined are weighted and aggregated to develop a crisp output (Sood et al., 2015).

Rule-based methods appear to be able to distinguish user activity from different users through behaviour profiling, but at a lesser rate of some machine learning methods as RBF ANNs (Li et al., 2013).

Other methods of user profiling

Majeed, Jing, Novakovic, and Ouazzane (2014) applied K-means clustering as a means to establish a user profile with which to compare future user behaviour for legitimacy. K-means clustering is used to create a usage profile from the input data gathered by a feature extractor. Once the usage profile is created the feature extractor compares new user data to determine deviation. If a persistent deviation occurs, an alert is made (Majeed et al., 2014).

Shi, Niu, Jakobsson, and Chow (2011) developed a machine learning framework based around modelling independent features. The user model is developed as a combination of the probability density of all features, based on time of occurrence. Selected features were based on good events, bad events and location. Good events would be those conducted with known contacts, bad events are those with unknown contacts. When in training, frequency and time between certain events would be used to establish feature probability. Using the Gaussian mixture model (GMM) the probability of user being at a certain location at a certain time would also be calculated (Shi et al., 2011).

The research conducted by Shi et al. (2011) provide a set of concise features that behaviour-based authorisation can be tested on, including calls made, texts made, internet activity, and user location (Shi et al., 2011).

Of the methods of behaviour-based profiling and authorisation, ANNs and rule-based methods appear to produce effective results, being able to identify most intrusive activity records during various experiments. However, the problem of false positives remains among these methods. Work done by Li, Wheeler, and Clarke (2014) based on that from Li et al. (2013) found that while illegitimate users were denied application access 95.83% of the time, their developed framework allowed legitimate users access only 87.09% of time (Li et al., 2014).

Intrusion detection methods for reducing false positives

False positive reduction is important to intrusion detection as false positives make it more difficult to identify intrusive activity and can cause resource intensive false alarms that discourage the use of the intrusion detection system altogether.

Cluster-based Intrusion Detection Methods

Yassin, Udzir, Muda, and Sulaiman (2013) used K-means clustering and Naïve Bayes classification methods combined to minimise false alarms generated by anomaly-based intrusions detection systems. K-means clustering is used to separate activity data, normal or intrusive, into separately identifiable partitions. Naïve Bayes Classification uses a set of attributes assigned to a set of classes and calculates the probability of activity belonging to one class (normal activity) to another (intrusive activity) by assessing the occurrence of attributes with the given activity. The combination of K-means clustering and Naïve Bayes classification performed greater, in terms of reducing false positives, than either of the two methods standalone, demonstrating the advantages of hybrid approaches (Yassin et al., 2013).

Hybrid frequency and relation-based Intrusion Detection Methods

Spathoulas and Katsikas (2010) discuss and experiment with a system with three components:

- Neighbouring Related Alerts (NRA)
- High Alert Frequency (HAF)
- Usual False positives (UFP)

The NRA method relies on the assumption that most attacks have a group of alerts related to it. NRA works by counting the amount of neighbours that exist in a time window, that have the same source and destination addresses (Spathoulas & Katsikas, 2010).

HAF is based on how often an alert with a certain signature appears within a certain time window. Each existing signature is given an average frequency which establishes the general distribution of signatures. After that, every

time a signature occurs, the minimum amount of time for reoccurrence is established. If reoccurrence of a signature appears before it should several times it will cause an alert (Spathoulas & Katsikas, 2010).

UFP is designed based on the idea that patterns of common false positives are related to topology problems or misconfigured services. The frequency for the signature during an attack-free period is made. During actual deployment, if a number of common frequencies occur more than expected, it may be a true positive and an alert is made if an established limit is breached.

Each component acts independently to one another before combining their individual verdicts on whether or not a given alert is true. This allows for the strengths and weaknesses of each component to complement and mitigate each other (Spathoulas & Katsikas, 2010).

System security status level

Part of the research conducted by Li et al. (2013) involved the development of a framework that could facilitate behaviour-based authorisation. The framework was developed to take into account different error rates experienced by different sets of application usage data. A System security status (SSS) level is kept between -3 and +3 where -3 is low security and +3 is high security. SSS level is determined by the performance of a given application based on equal error rate (EER) and the verification result, which can be a pass or fail. A verification test is when user behaviour is matched to what is expected. If the verification is successful, it is added to the SSS level. If verification fails, it is subtracted. The SSS also decreases over time through lack of usage. The SSS framework allows for a given base behaviour-based authorisation method to produce initially incorrect results but for corrections to be made if necessary based on further results (Li et al., 2013).

CHOSEN BEHAVIOUR-BASED AUTHORISATION METHOD

Artificial Neural Network

The behaviour-based authorisation method chosen as a baseline for this research was an ANN. The ANN was built using pybrain, a python module that allows customisable ANNs to be built and applied to datasets. The ANN was built with a configuration of four inputs, one output neuron and a variable amount of hidden neurons. The input values were the day of the week for a given activity record, the time of the day for a given activity record, the type of activity recorded and the details of the activity record. The output of the ANN would be a single number, where the closer to one the value was, the more likely the record would be that of a legitimate user.

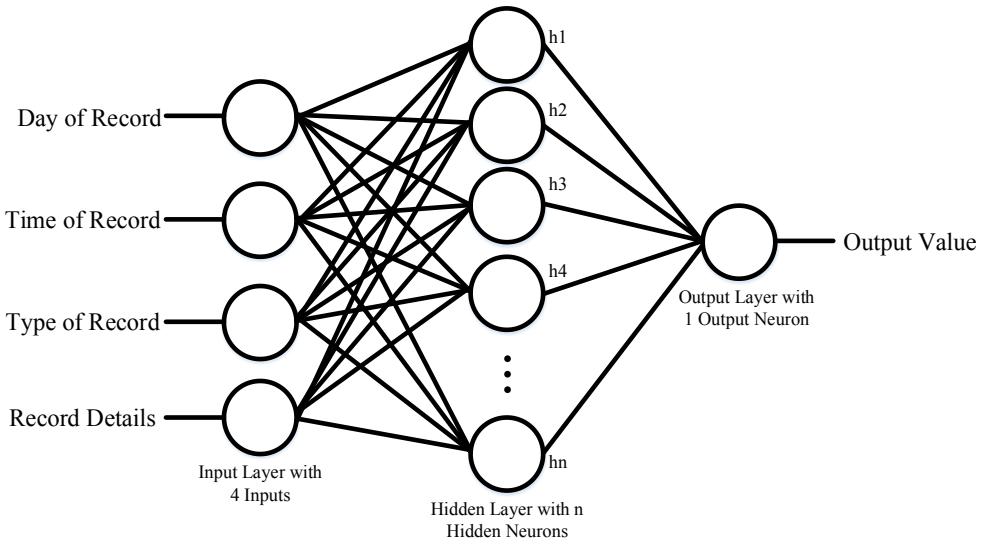

Figure 1 ANN configuration for Behaviour-Based Authorisation Diagram, the number of hidden neurons (n) is a configurable amount that had been experimented on.

The ANN configuration in Figure 1 receives four input values in the input layer and produces one output. The output value is a number that is tested by examining whether or not it comes between an established pair of

threshold values. If the output value falls outside the threshold values, it is designated as an intrusion, if it falls between the threshold values it is designated as legitimate activity. The ANN was trained using the supervised method of backwards propagation, as provided by the pybrain module (Wasserman, 1989a). Only legitimate user behaviour could be used to train the ANN as intrusive behaviour would be unknown.

Chosen Behavioural Attributes

Partly based on the research conducted by Shi et al. (2011), for this experiment the following mobile device attributes were used as independent/control variables to measure accuracy of the improved behaviour based authorisation method:

- calls made to known contacts
- calls made to unknown contacts
- text messages sent to known contacts
- texts messages sent to unknown contacts
- packet data sent to and from the device
- location of the mobile device
- day of the week and time of day activity occurs

INTRUSION DETECTION COMPONENTS

Self-correction using surrounding data

Self-correction provides a form of improvement that doesn't require user intervention (Patel et al., 2011). Because self-correction must be conducted without external intervention, it must be able to determine a correction only using internal factors (Schmeck, Müller-Schloer, Çakar, Mnif, & Richter, 2010; Yang et al., 2013). Surrounding data describes the records that occur before and after a given timeline record. Using surrounding data, a given record can be reassessed and potentially corrected if required. This works by assessing the surrounding data of a record. If a certain amount of the surrounding data records are assessed as intrusive, then it is concluded that the surrounded record is also intrusive. If a certain amount of the surrounding data records are not intrusive, the surrounded record will be determined to be a legitimate record.

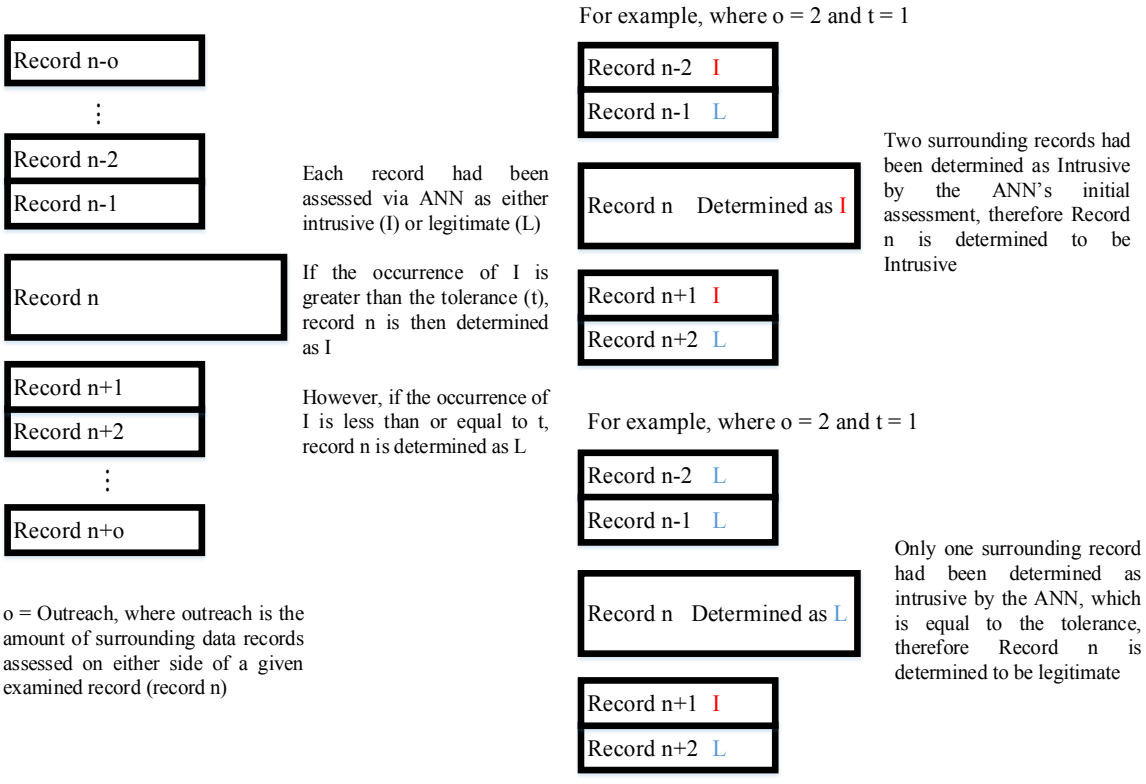

Figure 2 Diagram describing self-correction via surrounding data records

The self-correction process works by examining records that occur before and after a current record, based on a configurable outreach, and adjusts an initial assessment of a record, whether it is intrusive or legitimate, by examining the initial assessments of the surrounding records. If the numbers of records surrounding the examined record, within the given outreach, that is different to the examined record exceed a threshold value, which is also configurable, the examined record is changed to match the different records.

Adjusting self-correction through context awareness

Different conditions that include erratic user behaviour, overtrained ANNs or less distinguishable intruders require different approaches to mitigate (Piotrowski & Napiorkowski, 2013; Tetko, Livingstone, & Luik, 1995). A dynamic approach that can apply changes to assessment when required will allow a given problem to be mitigated without causing another problem.

Context awareness was applied to self-correction to allow for the tolerance of potential false assessments to be adjusted. If false positives occur frequently, tolerance was increased so that false positives that would otherwise remain incorrect can be changed. In the same way that false positives can be accounted for, false negatives that would evade the self-correction process can be caught by adjusting the tolerance to be lower. The context-awareness component adjusts tolerance whenever a false positive or false negative was identified and corrected using self-correction. Whenever a record was not identified as false, the tolerance was changed to be closer to the default.

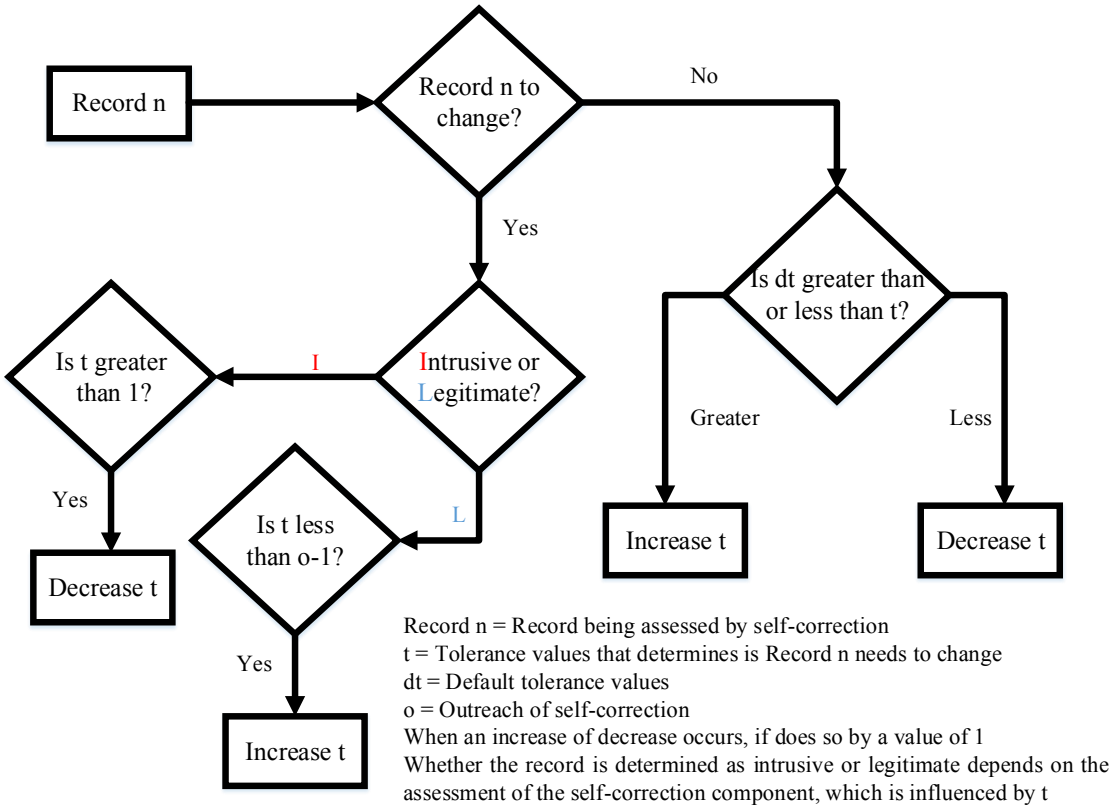

Figure 3 Flowchart of the context-aware component working with the self-correction component

The context-aware component adjusts the threshold of the self-correction component to be either more restrictive when false negative has recently been detected and changed or less restrictive when a false positive has recently been detected and changed. If no changes had been made due to caught false positives or false negatives, the context-aware component either does nothing or changes to be closer to the default self-correction threshold.

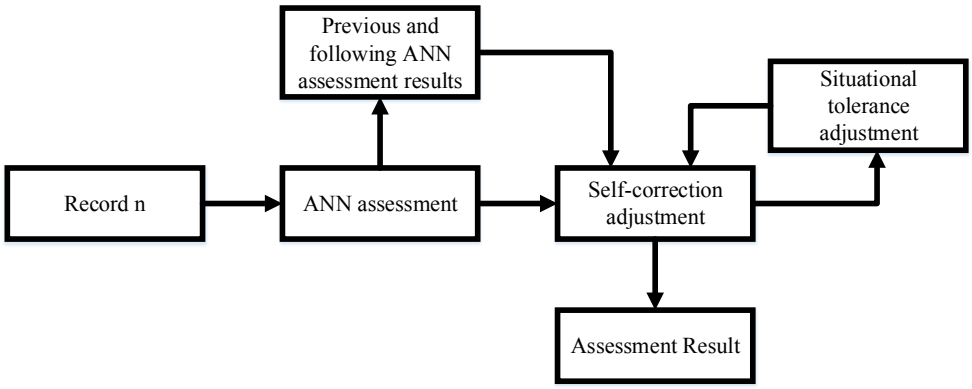

Figure 4 Overview of a record being processed through ANN combined with both intrusion detection components.

The overall process, as described by Figure 4, examines each record which are assessed by the ANN, previous and following assessments are also retrieved. The self-correction component adjusts the assessment of the ANN based on the previous and following record assessment, the ANN's assessment and the threshold adjustment made by the context-aware component. The self-correction component sends the result to the context-aware component which, based on the situation, will make a tolerance threshold adjustment to the self-correction component for the next record. The result from the self-correction component is also the final result for the record.

EXPERIMENT RESULTS

To test the capabilities of the applied intrusion detection methods of improving behaviour-based authorisation performance, three systems were developed. The base ANN (treatment script 1) served as the baseline for experimentation and the ANN with a self-correction component and context-awareness (treatment script 3) component served as the end-line. The base ANN with only the self-correction component was also assessed (treatment script 2). Each system was tested on 14 user pair datasets, where each pair consisted of training data from a legitimate user, test data from a legitimate user and test data from an intrusive user. Data for these 14 datasets was produced from the MIT reality mining dataset (Eagle & Pentland, 2006).The mean overall accuracies, false positive rates and true positives rates for each system applied to each dataset were examined.

The overall accuracy of a behaviour-based authorisation method depends on the rate that it can determine intrusive activity, as well as the rate that it produces false positives from legitimate activity. Overall accuracy did not always increase, but did for four of the 14 datasets demonstrating its capability to do so. Examining the mean result for all datasets together found that the overall accuracy did increase but not significantly.

False positives are legitimate activities mistaken for intrusive activities. Overall, the mean false positive rate was reduced when the intrusion detection components were applied. With the exception of two datasets, the behaviour-based authorisation method with intrusion detection concepts applied was able decrease false positives to below 5%.

True positives are intrusive activities correctly identified as intrusive. Overall the mean true positives rate was found to be less when both intrusion detection concepts had been applied. The self-correction component significantly increased the true positive rate, but this was always reduced with the application context-awareness. It was also found that if the base ANN had a true positive below 10%, the behaviour-based authorisation method with both intrusion detection concepts would perform at a rate below 5%, in most cases 0%.

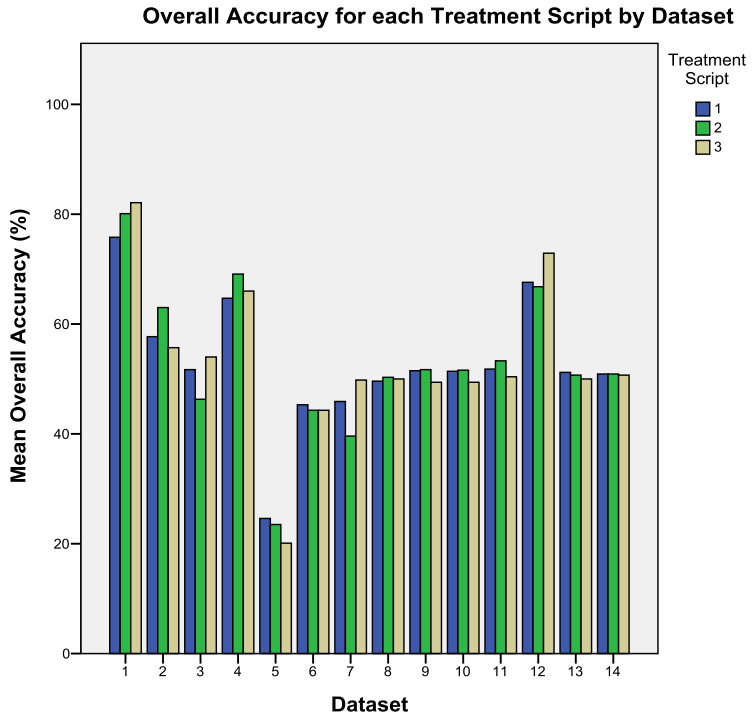

Figure 5 Bar chart of different mean overall accuracies separated by Dataset and Treatment Script, with datasets 1 through to 14 being plotted across treatment scripts 1, 2 and 3.

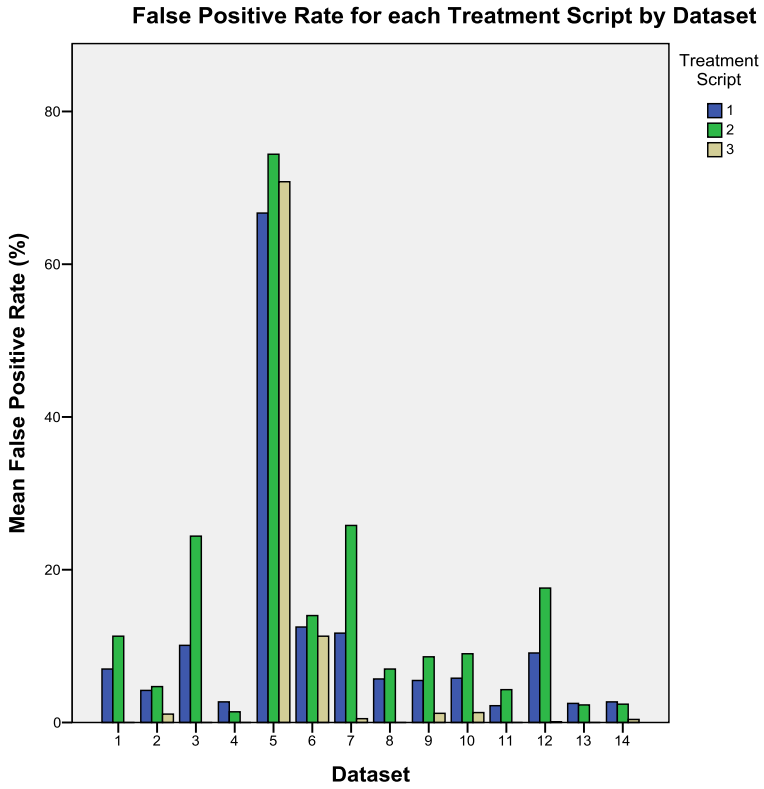

Figure 6 Bar chart of different mean false positive rates separated by Dataset and Treatment Script, with datasets 1 through to 14 being plotted across treatment scripts 1, 2 and 3.

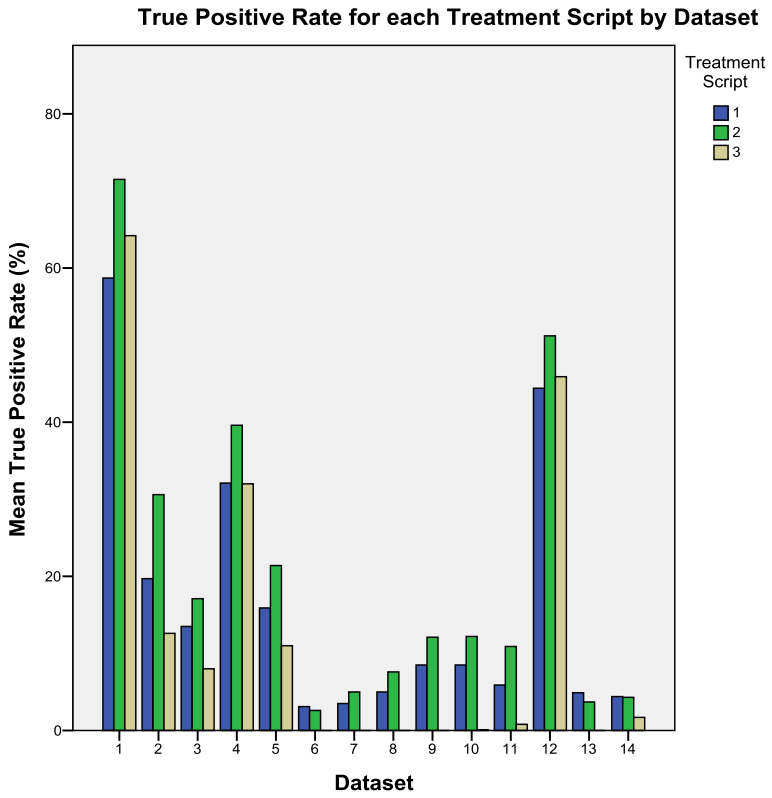

Figure 7 Bar chart of different mean true positive rates separated by Dataset and Treatment Script, with datasets 1 through to 14 being plotted across treatment scripts 1, 2 and 3

CONCLUSION

This research involved the development of modular components based on intrusion detection methods to find out whether or not said intrusion detection methods can improve the performance of behaviour-based authorisation on mobile devices. The intrusion detection methods were based around:

- Self-correction
- Context awareness
- Hybridisation

The results of experimentation found that the performance of a behaviour-based authorisation method for mobile devices can be improved by applying a hybrid self-correction and context-aware intrusion detection component, but only if the initial base behaviour-based authorisation method can detect intrusions at a rate of at least 10%. The self-correction intrusion detection component, when added, improved intrusion detection rates but increased the rate of false alarms. The hybrid self-correction and context-aware components, when applied, reduced false alarms to an average rate of below 5%, many producing a rate of 0%. However, the hybrid self-correction and context-aware components also reduced true positive (intrusion detection) rates for most datasets. Although the false positive reduction indicates an increase in performance, if the intrusion detection rate falls to 0, the performance of a behaviour-based authorisation method is also effectively 0. The experiments found that if the base behaviour-based authorisation method's intrusion detection rate was at least 10%, the intrusion detection rate would not fall to 0 when the hybrid self-correction and context-aware components are applied.

A reduction in false positives indicates an improvement of performance for a behaviour-based authorisation method, as it maintains the accessibility of a device while potentially preventing breaches of privacy and integrity. However if the ability to detect intrusions does not exist, the performance of a behaviour-based authorisation method also does not exist, regardless of lower false positive rates. As an improvement to the performance of the behaviour-based authorisation method only appears to occur when the initial method has an

intrusion detection rate of 10% or higher, future research should focus on experimenting with other behaviour-based authorisation methods that are capable of fulfilling the requirements for improvement.

REFERENCES

Anandarajan, M. (2002). Profiling Web Usage in the Workplace: A Behavior-Based Artificial Intelligence Approach. *Journal of Management Information Systems, 19*(1), 243-266. doi: 10.2307/40398573

Chang, G. W., Cheng, I. C., & Yu-Feng, T. (2010). Radial-Basis-Function-Based Neural Network for Harmonic Detection. *Industrial Electronics, IEEE Transactions on, 57*(6), 2171-2179. doi: 10.1109/TIE.2009.2034681

Eagle, N., & Pentland, A. (2006). Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing 10*(4), 255-268. doi: 10.1007/s00779-005-0046-3

Federal-Communications-Commission. (2014). *Report of Technological Advisory Council (TAC) Subcommittee on Mobile Device Theft Prevention (MDTP)*. US: Technological Advisory Council.Retrieved from https://transition.fcc.gov/bureaus/oet/tac/tacdocs/meeting12414/TAC-MDTP-Report-v1.0-FINAL-TAC-version.pdf

Home-Office. (2014). Reducing Mobile Phone Theft and Improving Security. *The Behavioural Insights Team.* Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/390901/HO_Mobile_theft_paper_Dec_14_WEB.PDF

Krenker, A., Kos, A., & Bešter, J. (2011). *Introduction to the artificial neural networks*: INTECH Open Access Publisher.

Li, F., Clarke, N., Papadaki, M., & Dowland, P. (2013). Active authentication for mobile devices utilising behaviour profiling. *International journal of information security, 13*(3), 229-244. doi: 10.1007/s10207-013-0209-6

Li, F., Wheeler, R., & Clarke, N. (2014). An Evaluation of Behavioural Profiling on Mobile Devices *Human Aspects of Information Security, Privacy, and Trust* (pp. 330-339): Springer. doi: 10.1007/978-3-319-07620-1_29

Majeed, K., Jing, Y., Novakovic, D., & Ouazzane, K. (2014). Behaviour Based Anomaly Detection for Smartphones Using Machine Learning Algorithm*International conference on Computer Science and Information Systems*, *held in Dubai (UAE),* 17th-18th October 2014: International Institute of Engineers.

Patel, A., Qassim, Q., Shukor, Z., Nogueira, J., Júnior, J., Wills, C., & Federal, P. (2011). *Autonomic agent-based self-managed intrusion detection and prevention system.* Paper presented at the Proceedings of the South African Information Security Multi-Conference (SAISMC 2010), Port Elizabeth, South Africa

Piotrowski, A. P., & Napiorkowski, J. J. (2013). A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling. *Journal of Hydrology, 476*, 97-111.

Schmeck, H., Müller-Schloer, C., Çakar, E., Mnif, M., & Richter, U. (2010). Adaptivity and self-organization in organic computing systems. *ACM Transactions on Autonomous and Adaptive Systems (TAAS), 5*(3), 10. doi: 10.1145/1837909.1837911

Shi, E., Niu, Y., Jakobsson, M., & Chow, R. (2011, 25th October 2010). *Implicit authentication through learning user behavior.* Paper presented at the Proceedings of the 13th international conference on Information security, Boca Raton, FL, USA. doi: 10.1007/978-3-642-18178-8_9

Sood, S., Mehmi, S., & Dogra, S. (2015, 19th-20th March 2015). *Artificial intelligence for designing user profiling system for cloud computing security: Experiment.* Paper presented at the Computer Engineering and Applications (ICACEA), 2015 International Conference on Advances in, Ghaziabad, India. doi: 10.1109/ICACEA.2015.7164645

Spathoulas, G. P., & Katsikas, S. K. (2010). Reducing false positives in intrusion detection systems. *Computers & Security, 29*(1), 35-44. doi: http://dx.doi.org/10.1016/j.cose.2009.07.008

Tapellini, D. (2014). Smart phone thefts rose to 3.1 million in 2013. *Consumer Reports.* Retrieved from http://www.consumerreports.org/cro/news/2014/04/smart-phone-thefts-rose-to-3-1-million-last-year/index.htm

Tetko, I. V., Livingstone, D. J., & Luik, A. I. (1995). Neural network studies. 1. Comparison of overfitting and overtraining. *Journal of Chemical Information and Computer Sciences, 35*(5), 826-833. doi: 10.1021/ci00027a006

Wasserman, P. (1989a). The Backpropagation Training Algorithm *Neural Computing: Theory and Practice* (pp. 44-54). New York: Van Rostrand Reinhold.

Wasserman, P. (1989b). Fundamentals of Artificial Neural Networks *Neural Computing: Theory and Practice* (pp. 11-26). New York: Van Rostrand Reinhold.

Yang, Q.-L., Lv, J., Tao, X.-P., Ma, X.-X., Xing, J.-C., & Song, W. (2013). Fuzzy self-adaptation of mission-critical software under uncertainty. *Journal of Computer Science and Technology, 28*(1), 165-187. doi: 10.1007/s11390-013-1321-9

Yassin, W., Udzir, N. I., Muda, Z., & Sulaiman, M. N. (2013). *Anomaly-Based Intrusion Detection through K-Means Clustering and Naives Bayes Classification.* Paper presented at the Proceedings of the 4th International Conference on Computing and Informatics (ICOCI), Sarawak, Malaysia. doi: 10.1109/CITA.2011.5999520