Enforcing Corporate Security Policies via Computational Intelligence Techniques

Antonio M. Mora,
Paloma De las Cuevas,
Juan Julián Merelo
University of Granada
Department of Computer Architecture and Technology, ETSIIT/CITIC
Granada, Spain
{amorag,paloma,jmerelo}@geneura.ugr.es

Sergio Zamarripa,
Anna I. Esparcia-Alcázar
S2 Grupo
Valencia, Spain
{szamarripa,aesparcia}@s2grupo.es

ABSTRACT

This paper presents an approach, based in a project in development, which combines Data Mining, Machine Learning and Computational Intelligence techniques, in order to create a user-centric and adaptable corporate security system. Thus, the system, named MUSES, will be able to analyse the user's behaviour (modelled as events) when interacting with the company's server, accessing to corporate assets, for instance. As a result of this analysis, and after the application of the aforementioned techniques, the Corporate Security Policies, and specifically, the Corporate Security Rules will be adapted to deal with new anomalous situations, or to better manage user's behaviour. The work reviews the current state of the art in security issues resolution by means of these kind of methods. Then it describes the MUSES features in this respect and compares them with the existing approaches.

Categories and Subject Descriptors

I.2.6 [**Artificial Intelligence**]: Learning - Induction; D.4.6 [**Operating Systems**]: Security and Protection - Access controls; I.2.1 [**Artificial Intelligence**]: Applications and Expert Systems

Keywords

Computational Intelligence; Evolutionary Computation; Corporate Security Policies; Security Rules

1. INTRODUCTION

Security in distributed systems has been a very profitable research area from the arising of the first client/server architectures [3]. Inside this, corporate security is one of the main topics. The landscape has changed dramatically in the last years, starting with the distribution of the information (instead of being centralised in corporate servers, it has been spread among multiple machines such as portable devices, external servers, or cloud storage systems); and continuing with the so-called Bring Your Own Device (BYOD) philosophy, in which the devices that access to the system are owned by the users (company's employees), and could contain both personal and professional information.

This scenario opens up new security issues [33], which should be dealt in a different way, taking into account both (company's) data security and (user's) privacy. In order to protect them, there are defined *Corporate Security Policies*.

To deal with this new situation, a novel system is being developed (inside an European Project). It is named *MUSES*, from *Multiplatform Usable Endpoint Security System* [31], which is a device-independent end-to-end user-centric tool. It considers a set of security rules defined as specifications of the Company Security Policies, and its main feature is the ability of 'learning' from the user's past behaviour and adapt, even inferring new ones, the set of rules in order to effectively manage potential future security incidents due to the user's behaviour. Then, the system will react, in a non-intrusive way, to the potentially dangerous sequence of actions (events) that he or she is conducting at any time.

To this end MUSES will analyse the users' behaviour by means of Data Mining (DM) techniques [24] and Machine Learning (ML) methods [6], extracting a set of patterns which will be later processed by means of Computational Intelligence (CI) algorithms, mainly Evolutionary Computation methods [5, 11, 23].

This is a step beyond the current state of the art in two senses: first regarding the current security systems for managing the new BYOD scenario inside the enterprises, as it can be read in [31]; and second concerning the application of Computational or Artificial Intelligence (AI) techniques to corporate security issues, focused on (and adapted to) the users' behaviour, as will be analysed in this work.

The paper is structured as follows. Next section gives a background in the current enterprise security issues. Section 3 reviews related work regarding the application of DM, ML, AI and CI techniques to a wide range of security problems inside the enterprise, but mainly focused on the user's behaviour and the consequent security policies adaptation, which are the main advantages of MUSES. The MUSES system's features regarding the application of those techniques are described in Section 4. Then, these features are com-

GECCO'14, July 12–16, 2014, Vancouver, BC, Canada.

http://dx.doi.org/10.1145/2598394.2605438.

pared with the existing works reaching some conclusions in Section 5.

2. ENTERPRISE SECURITY

Until these days, enterprises used to follow a static Security Policy devoted to control a certain structure [4], where the Information Assets and the devices were purchased and maintained by the company. Now that corporate networks are becoming dynamic for being adapted to the BYOD philosophy, there is an additional risk because the devices that the employees use are not always company-owned. A needed security policy, or in this case, an *Information Security Policy* (ISP from now on) should deal with the way of protecting a specific organisation information against a security breach. Though there are standards, such as the ISO27002 or the Security Forum's Standard of Good Practise[1], an ISP is defined depending on the characteristics of the community/organisation that they are built for.

Normally, the enterprise network architecture was being adapted to cope with external attackers [28]. However, with the consideration of BYOD, the threat is about corporate assets being compromised due to employees' devices with vulnerabilities [34], or leaked because they are being accessed from a device connected through an unsecured (public) network.

In Figure 1 there is a proposal which can be used for the beginning of the study of solutions that may make secure such a dynamic environment. It includes the possibility of having employee-owned mobile (smartphones and tablets) and portable (laptops) devices, and also the opportunity that the employees have of connecting these devices either from inside or outside the company premises. Moreover, company information assets are constantly accessed under these conditions, considering that an information asset means every *piece of information* that has a *value* (cost depending on the risk of being lost or leaked) for the company. It can be referred to files with sensitive information, to certain mails, or even to company applications.

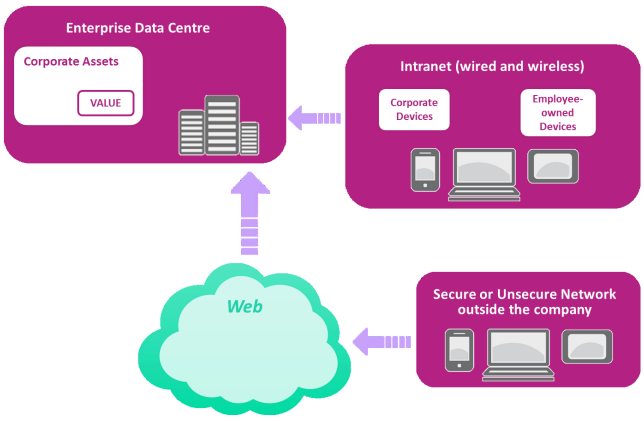

Figure 1: Architecture approach of an Enterprise Network assuming that the Company has adopted the BYOD philosophy.

The other issue to cope with is the elaboration of a good ISP, understandable for all the users of the company, and more importantly, non-intrusive for them. A lot of researchers have studied the natural tendency of employees whether to comply or not with the ISP [38, 7, 1], reaching conclusions such as the employees compliance with the security policies increases educating/training them in information security awareness [37], and decreases applying too much sanctions when a misuse or abuse occurs [17].

This situation leads to a need of protecting the organisation side, but also the users side, making non-interfering easy-to-follow ISPs, and leaving them to use their devices for personal purposes while working, without putting corporate information assets under risk. The compliance of these requirements would compose an End-to-End Security Solution (protecting both enterprise and employee), which is the aim of the MUSES project [31] (see Section 4).

3. STATE OF THE ART

Security is a wide area of research since the very beginning of the eighties [3]. Thousands of works have been published in a number of different issues in this topic. One of the most profiting fields is the application of Artificial Intelligence (AI) techniques to different security-based problems. This research line was started more than twenty years ago [10], and will be still open for several years further [32].

The topics addressed by the researchers are quite varied, including Data Mining (DM) [8, 22], and Machine Learning (ML) methods [12, 25], applied to many different problems.

Computational intelligence techniques have been also widely used in this area, being the most profiting methods the Evolutionary Computation (EC) metaheuristics: Genetic Algorithms (GAs) and Genetic Programming (GP).

There are several works using GAs for solving security issues, such as the intrusion detection (see [13] for a survey), the design and evaluation of security protocols [29, 43, 42], or the optimisation of different aspects related with security: IT security costs [21] and cryptographic protocols [44], to cite a few.

This work is focused on the application of different DM, ML and AI/CI techniques to a new set of security issues, which has arisen as a consequence of the new interactions between systems, and by the user's habits and behaviour (including the BYOD scenario), as it is described in Section 2. Then, the works that we are interested in are those related with the users' information and behaviour (in this scope), and the management (and adaptation) of Information or Corporate Security Policies (ISPs).

In this line, the paper by Greenstadt and Beal [14] combined biometrics signals with ML methods in order to get a reliable user authentication in a computer system. P.G. Kelley et al. [20] presented a method named *user-controllable policy learning* in which the user gives feedback to the system every time that a security policy is applied, so these policies can be refined according to that feedback to be more accurate with respect to the user's needs. This approach could be useful for a personal device, but our aim in MUSES is to have a global set of rules that could be adapted for all the users. On the other hand, policies could be created for enhancing user's privacy, as proposed by Danezis in [9], who defined a system able to infer privacy-related restrictions by means of a ML method applied in a social network environment. The idea of inferring policies will be also considered in MUSES, but in the scope of the company, and focused on ISPs.

[1]https://www.securityforum.org

A closer work to our approach is the one presented by Samak et al. [36], in which the authors use a clustering-based approach to infer new traffic security policies in a network. However our idea in MUSES, explained in Section 4, is to infer new Security Rules (as an specialisation of the ISPs) by means of GP.

Some other authors have applied this rule-based method for evolving (improving) a set of policies, such as Lim et al. [27, 26], who inferred new policies based in the decisions on a system, considering the user's feedback. A similar approach will be considered in MUSES, but there will be an automatic evaluation system for the new inferred rules, rather to the strict need of a user control. Moreover, these works have been tested on synthetic testbeds, but MUSES will run in real companies with real users and real data.

Finally, the work by Suarez-Tangil et al. [40], combines GP with the event correlation process which is also applied in MUSES [31]. However, their approach is designed to create the rules/engine for that process, instead of the security rules to be considered as the output of the event correlation, i.e. the decisions to be made according to the events produced and to the enterprise ISPs.

The next section briefly presents the MUSES system, and describes the techniques to be used inside it, as mechanisms for ISPs adaptation to the user's behaviour.

4. MUSES SYSTEM

As previously stated, MUSES will be a whole corporate security system aimed to deal with the new BYOD philosophy, i.e. it will manage user's accesses to the company servers from diverse own devices, which could be dangerous for several reasons, including the user's behaviour.

The defined MUSES architecture is shown in Figure 2. It is a *client/server* approach in which the *client* program will be installed in every user's mobile or portable device, independently of the platform (operating system and type of device). The *server* side would be installed in the corporate security operations centre. Both sides are connected through a secure channel (using HTTPS) over Internet.

One of the main features of this system will be the self-adaptation (to the user and context) of the set of Corporate Security Rules (specification of the ISPs). To this end, there is a component in the designed architecture (Figure 2, left side) named *MusKRS*, from MUSES *Knowledge Refinement System*. This will be run asynchronously in the server and will be in charge of analysing all the gathered information (events, context, user-related data), and adapting/refining the security rules to better deal with these events, also trying to predict future threats due to the user's behaviour.

This process will be composed by two steps: first, a Data Mining/Machine Learning procedure will be performed (in the *Data Miner* sub-component); second, a refinement and inference process will be done (in the *Knowledge Compiler* sub-component), considering the data 'extracted' in the first step, by means of Computational Intelligence techniques. It should be noticed that part of the refinement (or adaptation) of the security rules will be made using simpler methods, such as generalisation or specialisation of rules, for instance. Then, other parts of the process would be conducted using CI.

Another important fact is that MUSES will count with a human controller, normally the company Chief Security Officer (CSO), who will supervise the system activity by means of logs. Thus, adapted and inferred security rules will not be directly added to the current set of rules. Instead, they are proposed to this controller in order that he/she accepts them if they are interesting and correct. It is planned that the system will be able to 'learn' from this decisions so, after a so-called training or 'warm up' period, the rules would be directly accepted or rejected autonomously.

The following sections describe these processes: first focusing on DM techniques to be used both automatically by the KRS, and as a kind of decision-aid/monitoring tool for the CSO; second, the CI techniques are explained, mainly focusing in Evolutionary Computation approaches, since these methods perform very well, and have been widely used in security-based environments, as has been presented in Section 3.

4.1 Data Mining/Machine Learning

This task will be performed by the Data Miner module. It will take the 'raw' data from the database and will process the information, in order to yield a set of relevant data for the Knowledge Compiler sub-component or for the human controller. In the first case, this sub-component will take them as a reference in order to refine or adapt the current set of security rules (for instance, to deal with anomalous situations).

The process will be mainly non-supervised, and eventually the datasets can be huge (depending on the company's data flows), so Big Data processing methods [35] will be applied.

The DM/ML techniques will process the so-called patterns, which in this context correspond to events (and their related information) produced by the users' interactions with the system. The methods to be applied are:

- *Pattern Mining* [16]: This process will try to identify frequent or, on the contrary, anomalous patterns, in order to process them lately. The idea is that non-frequent patterns are potentially suspicious, and thus, could be of interest to be checked by the CSO or to serve as a reference for the rule-refinement process.

- *Classification* [30]: This technique tries to train a model (classifier) able to associate every pattern in the dataset to a class, so that the model could be used for assign a class for further incoming patterns with an unknown category. For instance, it could look for events (patterns) that had been marked as 'allowed' or 'denied' (according to the ISPs). When a new event arises, if it has not an assigned decision, the classifier should provide one based on the similarity with previous (and already labelled) patterns.

- *Clustering* [18]: The aim of this method is grouping the patterns considering some similarity criteria, in order to manage them as a set. This could be used for providing data visualisation mechanisms, in order to make it easier to interpret the data interaction and the distribution in clusters with respect to the different properties/features of the patterns.

- *Feature Selection* [15]: It consists on extract the most important features/variables from the data. This could be useful if we want to discard non-key features, which could be interesting in order to reduce the database weight, for improving the performance of other techniques (such as classification or clustering), and even

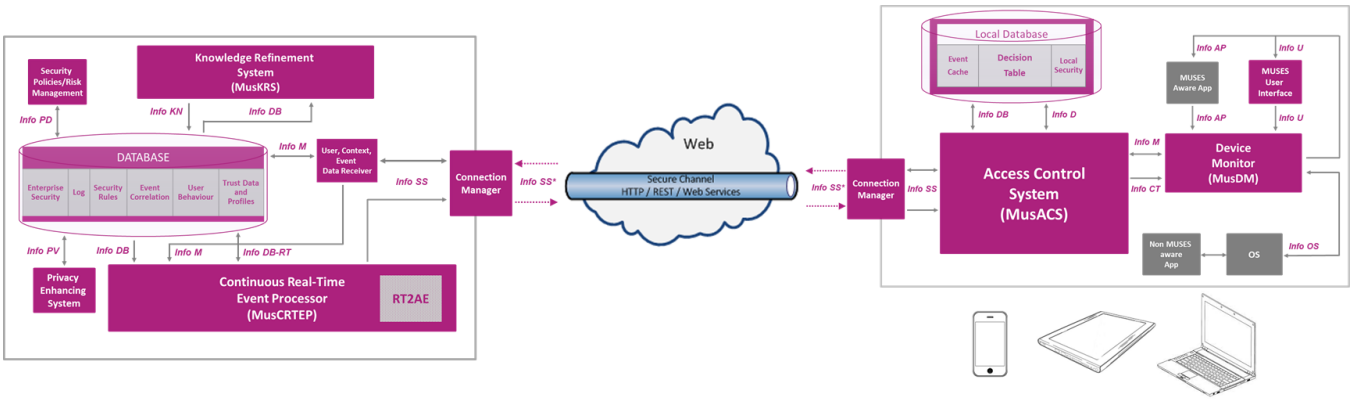

Figure 2: MUSES Architecture Overview (high-level components)

to improve the performance of whole the system, since less information would be gathered and transmitted.

- *Data Analysis*: This will provide the CSO with mechanisms to visualise interesting facts about the data, such as more frequent events, dangerous or suspicious users (according to their behaviour), more triggered rules, etc.

4.2 Computational Intelligence: Evolutionary Computation Methods

MUSES will use different EC approaches, initially, one in the DM/ML part of the process, and three in the rule-refinement/adaptation phase. They are based in GPs [23], and GAs [11].

The first evolutionary-based approach will be a *GP classification* method. This will be useful for two main reasons: first, in order to deal with the data class imbalance [19], very common in classification problems inside real systems (with real data); and second, to better manage categorical (non-numeric) data, since most of the features/variables and information gathered from the events take these kind of values.

Thus, this approach will be able to manage unbalanced datasets considering a fitness function in which a cost can be associated to the classifier accuracy at every epoch, having a penalty cost when the classifier makes a false negative (an element from the minority class which is classified as belonging to the majority class) [2]. Regarding the type of data, since GP algorithms can manage rule- or tree-based models, it will work perfectly with any categorical variable, yielding a good classifier as it has been made in other works (such as the aforementioned [2]).

The second set of EC approaches are, as stated, part of the rule-refinement (or adaptation) process. These techniques will be used for inference and optimisation, and will consider this data as a part of the process:

- The information extracted from the Data Miner subcomponent, mainly concerning the anomalous, unclassified or misclassified patterns. These are those patterns which did not match with any of the existing classes (they are quite different from the patterns belonging to those classes), so they cannot be included in any of the classes and thus they should be taken into account for a potential inference or update in the set of security rules, in order to 'cover' them.

- User-related information corresponding to those anomalous or unlabelled patterns (events). Thus, the user's ID, location and role, for instance, will be considered in order to select the applicable set of rules for that conditions.

- Context information for the same patterns, in order to also restrict the applicable set of rules considering this information.

Another useful information to consider in the refinement/adaptation process will be:

- Risk information extracted from the user profile (reputation), e.g. "Did the user received a lot of 'denies'/'allows' before?", i.e. "Is he/she trustworthy?". In case the user is not, more restrictive rules would be created for him, otherwise the corresponding rules could be 'eased' for that user.

- The information stored in logs along the system, which can, for instance, tell about how the user responds to system messages (either an action or if he gives feedback). This could result in the inference of new rules or in adaptation, in order to deal with, for instance, users that repeatedly ignore warning messages. Moreover, important log information regarding the parameters used or the decisions made in the different modules will be used for further tests of new inferred rules, as it is explained below.

So, the approaches that will be implemented are:

- *GP rule inference* method, which will generate/create new rules in order to 'cover' those situations non contemplated in the current set of rules. Thus, for instance, a new rule could be created in order to deal with the patterns to which the classifier could not assign a class. The generation will be done considering the so-called *dictionary*, i.e. a set of terms corresponding to all the possible inputs and actions in the system, which will be antecedents (conditions) and consequents in the security rules to be inferred. The evaluation of

these rules will be done considering the stored log information concerning the parameters along with the actions/decisions made in every component in the system. Thus, it will be possible to 'simulate' the whole system behaviour when the new rule is included and get a value of its performance.

- *GP rule refinement* approach, which will optimise the current set of rules, adjusting the values in the conditions (antecedents), for instance. Thus, some superfluous parts on the rules and even complete rules could be removed or improved, obtaining for instance specialisations or generalisations of existing rules which could mean a better performance. The evaluation (of the whole set of security rules) will be done considering the number of unlabelled patterns that will be 'covered' after the adjustments.

- *GA optimisation* algorithm for setting up and adapting the assets' values. These are numerical representations of the importance of the corporate assets, and are considered in the Real-Time Risk and Trust Analysis process, in order to assign a risk value to every potential decision that can be made by the system. If it is possible to evaluate the partial solutions proposed by the GA, this approach could be very useful for the CSO (who is in charge of assigning and adjusting these values over time). The adaptation or adjustment concerns the change in value that an asset could have due to a loss of importance, once an event has passed (a project presentation, for instance).

5. COMPARISON AND CONCLUSIONS

As it can be seen MUSES system will go quite far in the application of DM/ML techniques, with respect to other security-aimed systems. This was explained in the previous paper [31], where the system was compared with other existing (commercial) systems.

Regarding the scientific contribution, one of the main differences with respect to previous works is the consideration of security threats 'brought' by the user's behaviour inside the system, i.e. through interaction/events, rather than more general and external threats. Moreover, the techniques to be used here will work with real data (in a real system), as a difference to some research works.

Data Mining techniques have been used by the authors in some works, but usually aiming for a specific general objective, for instance the detection of threats (botnet) [8], or the recognition of anomalies [22], but they are not linked with a following process to improve the system (the refinement phase in MUSES).

There are some proposals in which security policies are inferred or refined [9, 36], but they do not affect the ISPs as in MUSES, and they are not based in the user's behaviour in order to do this.

Genetic Programming has been previously used by several authors [40, 27], even for creating new policies or rules in a security-aimed sense, but they do not affect the ISPs and moreover, our proposed evaluation functions (completely integrated in the system) for the refinement and inference approaches are novel.

With respect to Genetic Algorithms, they have been extensively used in the literature, mainly for the detection of anomalies and intrusions rather than for optimisation, as in our case. However, there are some examples that could be used as model for our approach, such as [21, 41].

Anyway, there is room for considering some of the proposed approaches that could be added as future features for MUSES such as the analysis of users via social networks [9, 25], the optimisation of security protocols [42], the implementation of intrusion detection mechanisms [13], or the application of novel privacy-related techniques [39], which is another feature also considered in MUSES.

6. ACKNOWLEDGMENTS

This work has been mainly supported by the MUSES European project (FP7-318508). In addition to projects SIPESCA (G-GI3000/IDIF, under Programa Operativo FEDER de Andalucía 2007-2013), EvOrq (TIC-3903), CANUBE (CEI2013-P-14), ANYSELF (TIN2011-28627-C04-02) and PYR-2014-17 included in GENIL - CEI BIOTIC (Granada).

7. REFERENCES

[1] A. Al-Omari, O. El-Gayar, A. Deokar, and J. Walters. Security policy compliance: User acceptance perspective. In *45th Hawaii International Conference on System Sciences*, pages 3317–3326. IEEE Press, 2012.

[2] E. Alfaro-Cid, K. Sharman, and A. Esparcia-AlcÃ¡zar. A genetic programming approach for bankruptcy prediction using a highly unbalanced database. In M. Giacobini, editor, *Applications of Evolutionary Computing*, volume 4448 of *Lecture Notes in Computer Science*, pages 169–178. Springer Berlin Heidelberg, 2007.

[3] A. J. P. Anderson. Computer security threat monitoring and surveillance. Technical report, James P. Anderson Co., Fort Washington, PA, 1980.

[4] S. Bacik. *Information Security Management Handbook*, volume 7, chapter Security Implications of Bring Your Own Device, IT Consumerization, and Managing User Choices, pages 133–142. Sixth edition, 2013.

[5] T. Back. *Evolutionary algorithms in theory and practice*. Oxford University Press, 1996.

[6] C. Bishop. *Pattern recognition and Machine Learning*. Springer, 2006.

[7] B. Bulgurcu, H. Cavusoglu, and I. Benbasat. Information security policy compliance: an empirical study of rationality-based beliefs and information security awareness. *MIS Quarterly*, 34(3):523–548, 2010.

[8] S. Chang and T. E. Daniels. P2p botnet detection using behavior clustering & statistical tests. In *Proceedings of the 2nd ACM Workshop on Security and Artificial Intelligence*, AISec '09, pages 23–30, New York, NY, USA, 2009. ACM.

[9] G. Danezis. Inferring privacy policies for social networking services. In *Proceedings of the 2Nd ACM Workshop on Security and Artificial Intelligence*, AISec '09, pages 5–10, New York, NY, USA, 2009. ACM.

[10] J. Frank and N. U. Mda-c. Artificial intelligence and intrusion detection: Current and future directions. In *In Proceedings of the 17th National Computer Security Conference*, 1994.

[11] D. E. Goldberg. *Genetic Algorithms in search, optimization and machine learning*. Addison Wesley, 1989.

[12] N. Görnitz, M. Kloft, K. Rieck, and U. Brefeld. Active learning for network intrusion detection. In *Proceedings of the 2Nd ACM Workshop on Security and Artificial Intelligence*, AISec '09, pages 47–54, New York, NY, USA, 2009. ACM.

[13] P. Gowher Majeed and S. Kumar. Genetic algorithms in intrusion detection systems: A survey. *International Journal of Innovation and Applied Studies*, 5(3):233–240, March 2014.

[14] R. Greenstadt and J. Beal. Cognitive security for personal devices. In *Proceedings of the 1st ACM Workshop on Workshop on AISec*, AISec '08, pages 27–30, New York, NY, USA, 2008. ACM.

[15] I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *J. Mach. Learn. Res.*, 3:1157–1182, 2003.

[16] J. Han, H. Cheng, D. Xin, and X. Yan. Frequent pattern mining: Current status and future directions. *Data Min. Knowl. Discov.*, 15(1):55–86, 2007.

[17] T. Herath and H. Rao. Protection motivation and deterrence: a framework for security policy compliance in organisations. *European Journal of Information Systems*, 18:106–125, 2009.

[18] A. K. Jain, M. N. Murty, and P. J. Flynn. Data clustering: A review. *ACM Comput. Surv.*, 31(3):264–323, Sept. 1999.

[19] N. Japkowicz and S. Stephen. The class imbalance problem: A systematic study. *Intell. Data Anal.*, 6(5):429–449, Oct. 2002.

[20] P. G. Kelley, P. Hankes Drielsma, N. Sadeh, and L. F. Cranor. User-controllable learning of security and privacy policies. In *Proceedings of the 1st ACM Workshop on Workshop on AISec*, AISec '08, pages 11–18, New York, NY, USA, 2008. ACM.

[21] T. Kirta and J. Kivimaab. Optimizing it security costs by evolutionary algorithms. In C. Czosseck and K. Podins, editors, *Conference on Cyber Conflict*, pages 145–160, Tallinn, Estonia, 2010. CCD COE Publications.

[22] M. Kloft, U. Brefeld, P. Düessel, C. Gehl, and P. Laskov. Automatic feature selection for anomaly detection. In *Proceedings of the 1st ACM Workshop on Workshop on AISec*, AISec '08, pages 71–76, New York, NY, USA, 2008. ACM.

[23] J. R. Koza. *Genetic Programming: On the programming of computers by means of natural selection*. MIT Press, Cambridge, MA, 1992.

[24] S. J. Lee and K. Siau. A review of data mining techniques. *Industrial Management & Data Systems*, 101(1):41–46, 2001.

[25] A. Leontjeva, M. Goldszmidt, Y. Xie, F. Yu, and M. Abadi. Early security classification of skype users via machine learning. In *Proceedings of the 2013 ACM Workshop on Artificial Intelligence and Security*, AISec '13, pages 35–44, New York, NY, USA, 2013. ACM.

[26] Y. T. Lim, P. C. Cheng, J. Clark, and P. Rohatgi. Policy evolution with genetic programming: A comparison of three approaches. In *Evolutionary Computation, 2008. CEC 2008. (IEEE World Congress on Computational Intelligence). IEEE Congress on*, pages 1792–1800, June 2008.

[27] Y. T. Lim, P. C. Cheng, P. Rohatgi, and J. A. Clark. Mls security policy evolution with genetic programming. In *Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation*, GECCO '08, pages 1571–1578, New York, NY, USA, 2008. ACM.

[28] R. Lippmann, K. Ingols, C. Scott, K. Piwowarski, K. Kratkiewicz, M. Artz, and R. Cunningham. Evaluating and strengthening enterprise network security using attack graphs. Project report ia-2, Massachusetts Institute of Technology, Lincoln Laboratory, October 2005.

[29] W. Lu and L. Traore. Detecting new forms of network intrusion using genetic programming. In *Proceedings of the 2003 Congress on Evolutionary Computation*, pages 2165–2172, 2003.

[30] J. MacQueen et al. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, page 14. California, USA, 1967.

[31] A. Mora, P. De las Cuevas, J. Merelo, S. Zamarripa, M. Juan, A. Esparcia-Alcázar, M. Burvall, H. Arfwedson, and Z. Hodaie. MUSES: A corporate user-centric system which applies computational intelligence methods. In D. S. et al., editor, *29th Symposium On Applied Computing*, pages 1719–1723, 2014.

[32] B. Morel. Artificial intelligence and the future of cybersecurity. In Y. Chen, A. A. Cárdenas, R. Greenstadt, and B. I. P. Rubinstein, editors, *AISec*, pages 93–98. ACM, 2011.

[33] R. Oppliger. Security and privacy in an online world. *IEEE Computer*, 44(9):21–22, September 2011.

[34] C. Orthacker, P. Teufl, S. Kraxberger, G. Lackner, M. Gissing, A. Marsalek, J. Leibetseder, and O. Prevenhueber. Android security permissions - can we trust them? In *MobiSec Session on Smartphone Security*, Aalborg, 2011.

[35] B. Ratner. *Statistical and Machine-Learning Data Mining: Techniques for Better Predictive Modeling and Analysis of Big Data, Second Edition*. CRC Press, Inc., Boca Raton, FL, USA, 2nd edition, 2011.

[36] T. Samak and E. Al-Shaer. Synthetic security policy generation via network traffic clustering. In *Proceedings of the 3rd ACM Workshop on Artificial Intelligence and Security*, AISec '10, pages 45–53, New York, NY, USA, 2010. ACM.

[37] R. Shaw, C. Chen, A. Harris, and H.-J. Huang. The impact of information richness on information security awareness training effectiveness. *Computers & Education*, 52:92–100, 2009.

[38] M. Siponen, S. Pahnila, and A. Mahmood. *New Approaches for Security, Privacy and Trust in Complex Environments*, volume 232, chapter Employees' adherence to information security policies: an empirical study, pages 133–144. IFIP International Federation for Information Processing, 2007.

[39] A. Solanas and A. Martínez-bal. *Advances in Artificial Intelligence for Privacy Protection and Security.* World Scientific Publishing Co., Inc., River Edge, NJ, USA, 2009.

[40] G. Suarez-Tangil, E. Palomar, J. Fuentes, J. Blasco, and A. Ribagorda. Automatic rule generation based on genetic programming for event correlation. In l. Herrero, P. Gastaldo, R. Zunino, and E. Corchado, editors, *Computational Intelligence in Security for Information Systems*, volume 63 of *Advances in Intelligent and Soft Computing*, pages 127–134. Springer Berlin Heidelberg, 2009.

[41] A. Tamjidyamcholo. Genetic algorithm approach for risk reduction of information security. *International Journal of Cyber-Security and Digital Forensics (IJCSDF)*, 1(1), 2012.

[42] L. Zarza, J. Forné Muñoz, J. R. Pegueroles Vallés, and M. Soriano Ibáñez. *Advances in artificial intelligence for privacy protection and security*, chapter Genetic algorithms for designing network security protocols, pages 325–358. World Scientific, 2010.

[43] L. Zarza, J. Pegueroles, and M. Soriano. Evaluation function for synthesizing security protocols by means of genetic algorithms. In *Proceedings of the The Second International Conference on Availability, Reliability and Security*, ARES '07, pages 1207–1213, Washington, DC, USA, 2007. IEEE Computer Society.

[44] L. Zarza, J. Pegueroles, M. Soriano, and R. Martínez. Design of cryptographic protocols by means of genetic algorithms techniques. In M. Malek, E. Fernández-Medina, and J. Hernando, editors, *SECRYPT*, pages 316–319. INSTICC Press, 2006.