J Inf Process Syst, Vol.13, No.5, pp.1203~1212, October 2017
https://doi.org/10.3745/JIPS.03.0080

ISSN 1976-913X (Print)
ISSN 2092-805X (Electronic)

JOURNAL OF INFORMATION PROCESSING SYSTEMS JIPS

A Detailed Analysis of Classifier Ensembles for Intrusion Detection in Wireless Network

Bayu Adhi Tama*,** and Kyung-Hyune Rhee**

Abstract

Intrusion detection systems (IDSs) are crucial in this overwhelming increase of attacks on the computing infrastructure. It intelligently detects malicious and predicts future attack patterns based on the classification analysis using machine learning and data mining techniques. This paper is devoted to thoroughly evaluate classifier ensembles for IDSs in IEEE 802.11 wireless network. Two ensemble techniques, i.e. voting and stacking are employed to combine the three base classifiers, i.e. decision tree (DT), random forest (RF), and support vector machine (SVM). We use area under ROC curve (AUC) value as a performance metric. Finally, we conduct two statistical significance tests to evaluate the performance differences among classifiers.

Keywords

Classifier Ensembles, Classifier's Significance, Intrusion Detection Systems (IDSs), Wireless Network

1. Introduction

Intrusion detection systems (IDSs) play very prominent roles in the modern security system. They are placed at the foremost position to obstruct the attacks that might be happened in computer network. Attacks are widely known as the most severe issues in the security systems. There exist countless number of attacks that have been identified, yet more novel attacks are continuously mushrooming. Moreover, they take advantage of the vulnerability of a system and try to make different kind of damages. They could make a particular resource unavailable at a certain time, modify communication data or contents in a system, or leak sensitive data and information of a system, for instance.

Machine learning and data mining techniques have been fascinated by researchers worldwide due to a great performance result in various application domains. In the realm of IDSs, these techniques show the promising result by predicting future attack patterns using learning paradigm [1]. Learning is the process of constructing a predictive model using data set. It lies in several categories, i.e., supervised, unsupervised, and reinforced learning. In particular, supervised learning uses the labelled samples to create a model and the future unknown samples would be labelled using the model. Furthermore, the objective of supervised learners is to obtain high classification accuracy and to reduce false positive rate (FPR) [2,3].

Manuscript received February 26, 2016; accepted July 5, 2016.
Corresponding Author: Kyung-Hyune Rhee (khrhee@pknu.ac.kr)
* Laboratory of Information Security and Internet Applications, Pukyong National University, Busan, Korea (khrhee@pknu.ac.kr, bayuat@gmail.com)
** Faculty of Computer Science, Sriwijaya University, Inderalaya, Ogan Ilir, Sumatera Selatan, Indonesia (bayuat@gmail.com)

However, with a large number of features in the data set, obtaining a good predictive accuracy is computationally expensive. In the context of modern intrusion detection and prevention, fast detection capability with higher accuracy and lower FPR are quite necessary. To do so, fast detection approach might be achieved using lightweight classifier, whilst higher detection accuracy could be obtained using classifier ensembles approach.

In addition, classifier ensembles or multiple classifier systems (MCSs) are employed to improve the classification accuracy. MCSs are deployed by incorporating a number of base classifiers to predict the final class. The performance of MCSs rely on several elements, i.e. the selection of base classifiers, combination methods, and the MCSs architecture [4]. In this study, we focus on the performance evaluation of two combination schemes, i.e., voting [5] and stacking [6] using three base classifiers, i.e., decision tree (DT) [7], random forest (RF) [8], and support vector machine (SVM) [9]. We consider the wide diversity of those base classifiers as one of the underlying principle on designing MCSs.

The major contribution of this work is a thorough performance evaluation of the aforementioned MCSs schemes against the base classifier models for intrusion detection system in a wireless ecosystem. We declared a hypothesis that, given the outstanding performance behavior in other application domains, MCSs model will outperform the base classifier models. To evaluate our proposed method, we conduct a number of experiments using publicly available data set for intrusion detection research in IEEE 802.11 environment, so-called GPRS [10]. It is a custom-made data set for evaluating IDSs in wireless network.

There rest of the paper is structured as follows. Section 2 describes the prior works related to classifier ensemble in intrusion detection. Section 3 provides the details of our proposed method for evaluating the classifier ensembles against the base classifiers. Section 4 discusses the experimental results, and finally Section 5 draws some concluding remarks.

2. Related Work

Prior researches of IDSs have been summarized and discussed in several works [11-13]. Table 1 shows the existing methods of intrusion detection using classifier ensembles available in the literature. As shown Table 1, most prior researches have been focused on an old data set, so called KDD Cup 1999. However, the data set has been received a lot of criticisms since it does not represent current attack patterns. Therefore, in this study we use data set which is specifically intended for IDSs research in wireless network. To the best of our knowledge, this is the first attempt of applying several classifier ensembles for IDSs in IEEE 802.11 wireless network.

Moreover, other classifier fusion techniques such as stacking has been underexplored in the existing literature. Most works have been emphasized on the combination approaches, i.e., majority vote, weighted majority vote, and product rule while we address more complex integration models, such as stacking. Rokach [14] classifies stacking, along with combiner tree and the grading approaches, as meta-learning based integrating method. Notwithstanding voting schemes have been applied in the previous works, in order to distinguish between our approach and the previous ones, we also consider to employ the three voting strategies, e.g. product of probabilities, minimum probability, and maximum probability [15] as the classifier combination techniques.

Concerning base classifiers, we consider an ensemble composed of heterogeneous classifiers in order

to generate diverse methods of each classifier in representing the knowledge from different perspectives. According to the common knowledge, it is straightforward to understand that to gain for combination, the individual classifiers must be different, and otherwise the improvement might not be obtained if identical classifiers were combined. Therefore, by combining the outputs of diverse classifiers, the final resultant model is supposed to have better prediction accuracy than each individual classifier.

Table 1. Prior works of intrusion detection using classifier ensemble

	Ensemble scheme	Base classifiers	Dataset	Performance metrics	Statistical test
[28]	Weighted ensemble	Classification and regression trees (CART), Bayesian networks	KDD Cup 1999	Accuracy	No
[29]	Majority voting	Neural network, support vector machine, and multivariate regression splines	KDD Cup 1999	Accuracy	No
[30]	Weighted ensemble	Decision tree, support vector machine	KDD Cup 1999	Accuracy	No
[31]	Boosting	Decisions stumps	KDD Cup 1999	Precision, false alarm rate	No
[32]	Product rule	NA	Private	AUC	No
[33]	Min, Max, and product rule	k-means, v-SVC	KDD Cup 1999	Precision, false alarm rate	No
[34]	Voting	Neural network, decision tree	KDD Cup 1999	TP rate, FP rate, Precision, Recall, and F1 measure	No
[35]	Bagging	Multilayer perceptron, radial basis function	Private	Accuracy	No
This study	Voting, stacking	Decision tree, random forest, and support vector machine	GPRS	AUC	Yes

3. Methodology

3.1 Dataset

For this study we employed GPRS dataset which possesses two distinct wireless network topologies, i.e., WEP/WPA and WPA2. WEP/WPA dataset is composed by 4 type attacks class (37.5%) and normal class (62.5%) with 15 variables and 1 class label variable. We consider 9600 instances for training set. WPA2 dataset comprises 7500 instances with 4 type attacks class (40%) and normal class (60%). It has 16 variables and 1 class label attribute.

3.2 Experimental Techniques

In this section we describe the two fusion techniques (e.g., voting and stacking) and the three single classifiers (e.g., DT, RF, and SVM).

3.2.1 Ensemble techniques

We hypothesize that combining weak classifiers might outperform the performance of the best

classifier members in ensemble. The combiner generates the best decision boundary based on error boundaries offered by each classifier in ensemble. In this study we consider the two ensemble techniques (e.g., voting and stacking).

3.2.1.1 Voting

Voting is the most fundamental combination policy. Supposed a set of N classifiers $\{h_1, \dots, h_N\}$ is given and we have task to combine h_i's to predict the class label from a set of l possible class labels $\{c_1, \dots, c_l\}$. For every instance x, the outputs of classifier h_i are given as an l-dimensional vector $(h_i^1(x), \dots, h_i^l(x))^T$, where $h_i^j(x)$ is the output of h_i for the class label c_j. Different output of $h_i^j(x)$ can be as follows [13].

- Crisp label: $h_i^j(x) \in \{0,1\}$ which takes value one if h_i predicts c_j as the class label and zero otherwise.
- Class probability: $h_i^j(x) \in [0,1]$, which can be considered as an estimate of the prior probability $P(c_j|x)$.

In this paper, we use the three different voting strategies, i.e., product of probabilities, minimum probability, and maximum probability.

3.2.1.2 Stacking

Stacking trains a learner to combine the base classifiers. Here, the base classifiers are called the level-0 classifiers, and the combiner is called meta-classifier [6,16]. One of the issue in stacking is obtaining the suitable base classifiers and the meta-classifier, especially in relation to each specific dataset [16]. Therefore, in this study we consider to evaluate stacking for intrusion detection data set. Fig. 1 shows the structure of stacking ensemble. To predict a new instance, the level-0 classifiers produce a vector of prediction that is the input to the meta-classifier, which in turn predict the class [16]. In this study, we use multi-response linear regression (MLR) as a meta-classifier [6,17].

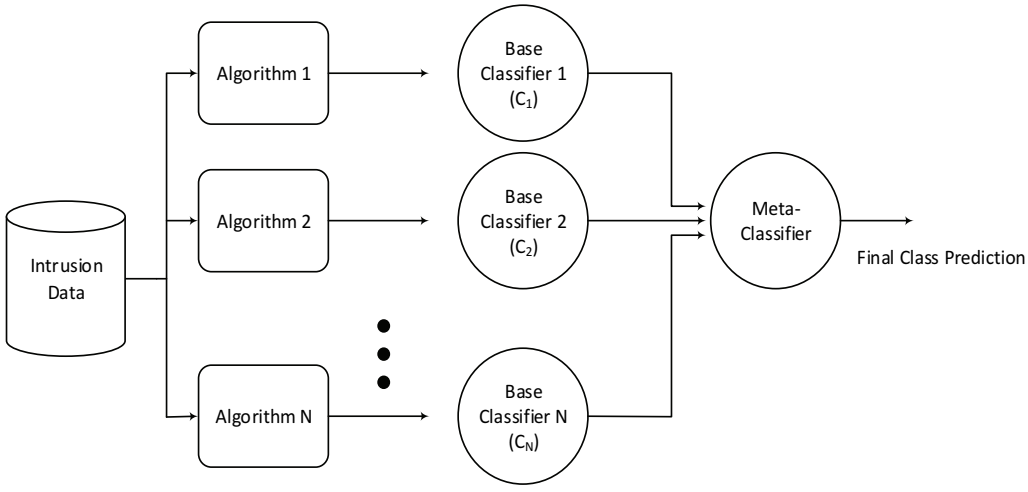

Fig. 1. Ensemble of classifiers using stacking [16].

3.2.2 Single base classifiers

3.2.2.1 Decision tree

DT (C4.5) is the most applicable and popular tree construction algorithm among the machine learning researches [18]. It is the descendent of the ID3 approach to inducing decision trees [19]. Tree is composed by a non-leaf node denoting a test on an attribute, each branch depicts an outcome of the test, and each leaf node represents a class label. Tree pruning is typically required to avoid *over-fitting* the data and to improve accuracy of the classifier on unseen instances.

Tree is fully generated in a bottom-up manner, and finally pruning is then performed. The C4.5 uses the concept of information gain to select the optimal splitting criterion that 'best' separates a given data. In addition to simple and fast tree construction, DT generally has good classification accuracy. Since there is only one parameter that can be fine-tuned, in this experiment we conduct several experiments to choose the best confidence C parameter by implementing a *grid* search on $C = \{0.05, 0.10, \dots, 0.50\}$. We report only the result with the best parameter.

3.2.2.2 Random forest

RF generates a number of trees and chooses the variables to put into each model by random selection [8]. The tree is generated to maximum size but it is not pruned. The strategy on incorporating of various trees resulting good predictive accuracy and avoiding *over-fitting*. There are two tuning parameters in RF: the number of q of variables to be selected in each node, which is generally kept constant on all nodes, and the number of trees, that make up the forest.

Compared to other classifiers, RF has several advantages such as lower computational burden since every single tree is based on fewer variables and easier implementation in parallel computing manner that can further accelerate the algorithm [20]. For this experiment, we use the number of trees is 500 and set the number of q is the square root of the total number of predictors.

3.2.2.3 Support vector machine

SVM was firstly proposed by Cortes and Vapnik [21] and the earlier version was developed for two-class classification problem. Hitherto, it has been extended for multi-class classification problem and regression. Training vectors are mapped into a higher dimensional space and SVM tries to find the linear separating *hyperplane* with the maximal margin using kernel functions. There are four basic kernel functions, i.e. linear, polynomial, radial basis function (RBF), and sigmoid. In this paper, we use radial basis function RBF kernel. It requires a width parameter of Gaussian function γ and cost parameter $\mathcal{C}$. To choose the optimal parameters, we follow [22] by implementing a *grid* search on $\mathcal{C} = [2^{-5}, 2^{-4}, \dots, 2^{13}]$ and $\gamma = [2^{-15}, 2^{-13}, \dots, 2^{3}]$ to obtain the best parameter combination. We report only the result obtained with the best parameters.

3.2.3 Model selection and evaluation metric

For model selection, as recommended by [23], we use five times two cross-validation (*5x2f cv*). It splits the data randomly into two equal parts, one part is for training set and the other part is for testing set or validation set. This process is then repeated five times. The result of this process is 10 performance values. This method is found to be more robust than usual k-cross-validation since it

overcomes the problem of underestimated variance. We use area under ROC curve (AUC) as the most adequate classifier performance measure [24,25]. It has advantages over other performance measures since simple classification accuracy is often a poor metric for measuring performance. Furthermore, it includes all cut-off values compared to other performance measures. AUC is a portion of the area of the unit square, its value will always be between 0 and 1.0 [24]. AUC is defined as follows:

$$AUC = \int_0^1 \frac{TP}{TP+FN} d\frac{FP}{FP+FN} = \int_0^1 \frac{TP}{P} d\frac{FP}{N} \quad (1)$$

where TP is true positives, TN is true negatives, P is positive, and N is negative. We present the average value of five times two cross-validation (*5x2f cv*).

4. Performance Analysis

We compare and report the performance result of classifier ensemble applied on WEP/WPA dataset and WPA2 dataset. We show that by incorporating multiple classifiers, the final ensemble performance is supposed to increase significantly. Fig. 2 presents the median AUC value of five times two folds cross-validation per classifier for WEP/WPA and WPA2 dataset. Pro-V, Min-V, Max-V, and STA denote product of probabilities, minimum probability, maximum probability voting, and stacking, respectively.

To determine whether the differences between the classifiers in term of AUC values are significant we use the Friedman test [26]. It is the nonparametric counterpart of the repeated-measures one-way ANOVA test [23,27]. Each classifier is ranked for each fold separately, according to the AUC value, in ascending order, from the best performer to the worst performer. Consider n folds (10 in our case) and k classifiers (7 in our case) to evaluate. The Friedman statistic is defined as:

$$\chi_F^2 = \left[\frac{12}{n\times k\times(k+1)} \times \sum_{j=1}^{k}\left(R_{.j}\right)^2\right] - 3 \times n \times (k+1) \quad (2)$$

where $R_{.j}$ is the mean rank of the jth of k algorithms. The mean rank is defined as $\bar{R}_{.j} = \frac{1}{n}\sum_1^n R_{ij}$ where R_{ij} is the rank of jth of k classifiers on the ith of n folds.

It is suggested that only if the null hypothesis (all algorithms have the same performance value) of the Friedman test is rejected we conduct post hoc test using Nemenyi test. It computes a q statistic over the difference in average mean ranks of the classifiers. For any two classifiers, q statistic is computed as:

$$q = \frac{\bar{R}_{.j1} - \bar{R}_{.j2}}{\sqrt{\frac{k(k+1)}{6n}}} \quad (3)$$

The null hypothesis is rejected if the q value exceeds the critical values q_α. The critical values q_α are the studentized range statistic scale divided by $\sqrt{2}$, n is the number of folds, and k is the number of classifiers [23,27].

Firstly we are going to discuss the results of WEP/WPA dataset. From the Fig. 2, it is clear STA is the best performer followed by DT, Min-V, Pro-V, Max-V, RF, and SVM. Table 2 confirms the results of the average ranking of 10-folds AUC values for WEP/WPA dataset. Among the performance outputs,

STA is the top performer. Furthermore, STA has brought significant improvement over the three single classifiers, e.g., DT, RF, and SVM with the percentage improvement is 0.41%, 1.29%, and 13.35%, respectively. However, the other classifier ensembles (e.g., Min-V, Max-V, and Pro-V) have not brought significant improvement over the DT classifier. We also found that across the entire results of single classifiers performance, SVM has performed worst. The Friedman test indicates that the differences of classifier performance in term of AUC values are significant with $\chi^2(7) = 38.9571$, p-value = 7.29739E-07 ($p < 0.05$), and 6 degrees of freedom. We can conclude that there is a significant difference among the seven classifiers on the WEP/WPA dataset.

Table 2. Mean ranking of the folds for AUC

Dataset	Classifier						
	DT	RF	SVM	Pro-V	Min-V	Max-V	STA
WEP/WPA	2.7	5.3	7.0	3.6	3.2	4.4	1.8
WPA2	6.0	3.6	7.0	2.6	3.5	3.0	2.3

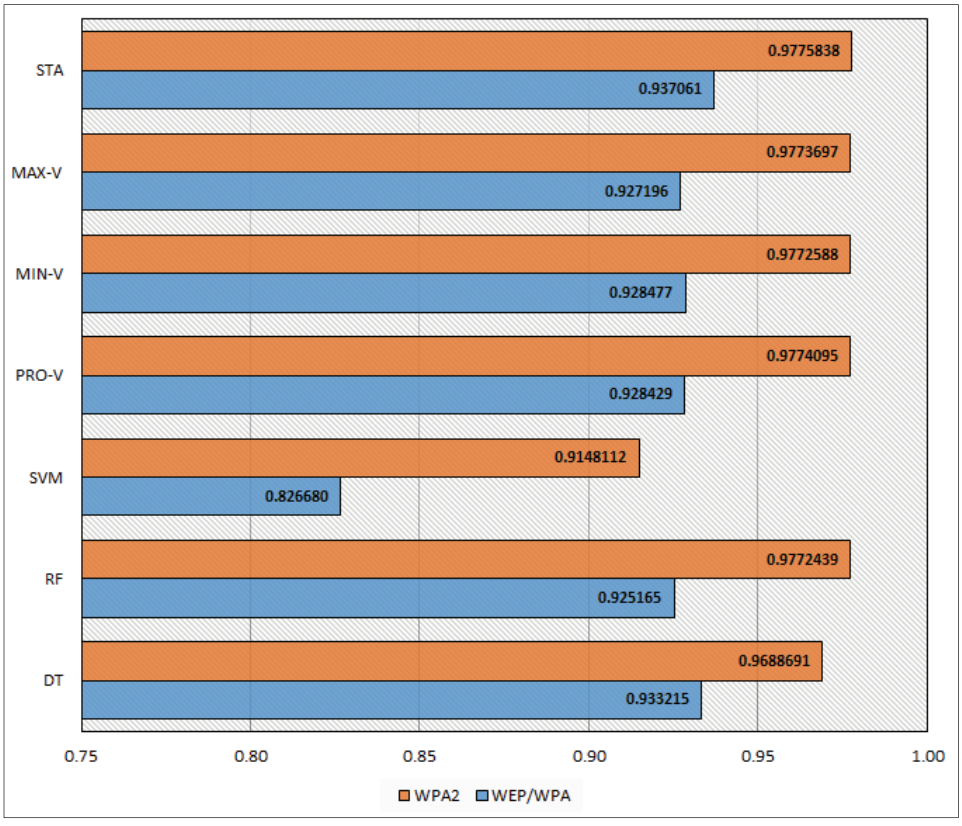

Fig. 2. *5x2f cv* average AUC value of two datasets per classifier.

For WPA2 dataset, among the four classifier ensemble schemes, STA has performed best followed by Pro-V, Max-V and Min-V. STA has brought significant improvement over the three single classifiers, e.g., DT, RF, and SVM with the percentage improvement is 0.90%, 0.03%, and 6.86%, respectively. In

addition, the other classifier ensemble schemes (e.g., Min-V, Max-V, and Pro-V) outperform the three single classifiers significantly. Surprisingly, among the three single classifiers, RF outperforms DT and SVM still performs worst considerably. The Friedman significant test yields significant difference of classifier performance in term of AUC metric with $\chi^2(7) = 38.2857$, p-value = 2.24256E-06 ($p < 0.05$), and 6 degrees of freedom. Hence, we can confidently conclude that we cannot reject the null hypothesis on the WPA2 dataset.

5. Conclusions

This paper studies the classifier ensembles applied on IDSs for IEEE 802.11 wireless network. The performance of several classifier ensembles, e.g., product of probabilities voting, minimum probability voting, maximum probability voting and stacking were thoroughly evaluated and compared with the three different single classifiers (e.g., DT, RF, and SVM). To the best of our knowledge, this is the first contribution to knowledge by considering a set number of classifier ensembles in IDSs for wireless network. We found that among the classifier ensembles, stacking is the top performer followed by voting schemes. Hence we suggest to include stacking or voting in the domain of IDSs in wireless network since their performance significantly outperform single classifiers.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2014R1A2A1A11052981).

References

[1] B. A. Tama and K. H. Rhee, "Performance analysis of multiple classifier system in DoS attack detection," in *Information Security Applications, LNCS, vol. 9503*. Cham, Switzerland: Springer International Publishing, 2016, pp. 339-347.

[2] B. A. Tama and K. H. Rhee, "A combination of PSO-based feature selection and tree-based classifiers ensemble for intrusion detection systems," in *Advances in Computer Science and Ubiquitous Computing*. Singapore: Springer Singapore, 2015, pp. 489-495.

[3] B. A. Tama and K. H. Rhee, "Data mining techniques in DoS/DDoS attack detection: a literature review," *Information*, vol. 18, no. 8, pp. 3739-3747, 2015.

[4] M. P. Ponti, "Combining classifiers: from the creation of ensembles to the decision fusion," in *Proceedings of the 24th SIBGRAPI Conference on Graphics, Patterns and Images Tutorials (SIBGRAPI-T)*, Alagoas, Brazil, 2011, pp. 1-10.

[5] D. Ruta and B. Gabrys, "Classifier selection for majority voting," *Information Fusion*, vol. 6, no. 1, pp. 63-81, 2005.

[6] A. K. Seewald, "How to make stacking better and faster while also taking care of an unknown weakness," in *Proceedings* of *the 19th International Conference on Machine Learning*, Nevada, LA, 2002, pp. 554-561.

[7] J. R. Quinlan, *C4.5: Programs for Machine Learning*. San Mateo, CA: Morgan Kaufmann, 1993.

[8] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5-32, 2001.
[9] V. N. Vapnik, *Statistical Learning Theory*. New York, NY: Wiley, 1998.
[10] D. W. F. V. Vilela, E. T. Ferreira, A. A. Shinoda, N. V. de Souza Araujo, R. de Oliveira, and V. E. Nascimento, "A dataset for evaluating intrusion detection systems in IEEE 802.11 wireless networks," in *Proceedings of the IEEE Colombian Conference on Communications and Computing (COLCOM)*, Bogota, Colombia, 2014, pp. 1-5.
[11] P. Garcia-Teodoro, J. Diaz-Verdejo, G. Macia-Fernandez, and E. Vazquez, "Anomaly-based network intrusion detection: techniques, systems and challenges," *Computers & Security*, vol. 28, no. 1-2, pp. 18-28, 2009.
[12] C. F. Tsai, Y. F. Hsu, C. Y. Lin, and W. Y. Lin, "Intrusion detection by machine learning: a review," *Expert Systems with Applications*, vol. 36, no. 10, pp. 11994-12000, 2009.
[13] N. C. Oza and K. Tumer, "Classifier ensembles: select real-world applications," *Information Fusion*, vol. 9, no. 1, pp. 4-20, 2008.
[14] L. Rokach, "Ensemble-based classifiers," *Artificial Intelligence Review*, vol. 33, no. 1, pp. 1-39, 2010.
[15] L. I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithm*, 2nd ed. Hoboken, NJ: John Wiley & Sons, 2014.
[16] M. P. Sesmero, A. I. Ledezma, and A. Sanchis, "Generating ensembles of heterogeneous classifiers using stacked generalization," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 5, no. 1, pp. 21-34, 2015.
[17] K. M. Ting and I. H. Witten, "Issues in stacked generalization," *Journal of Artificial Intelligence Research*, vol. 10, pp. 271-289, 1999.
[18] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York, NY: John Wiley & Sons, 2001.
[19] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, no. 1, pp. 81-106, 1986.
[20] P. Cichosz, *Data Mining Algorithms: Explained Using R*. Chichester, UK: John Wiley & Sons, 2015.
[21] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273-297, 1995.
[22] C. W. Hsu, C. C. Chang, and C. J. Lin, *A Practical Guide to Support Vector Classification*. Taipei City, Taiwan: National Taiwan University, 2010.
[23] J. Demsar, "Statistical comparisons of classifiers over multiple data sets," *Journal of Machine Learning Research*, vol. 7, pp. 1-30, 2006.
[24] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861-874, 2006.
[25] F. Provost, T. Fawcett, and R. Kohavi, "The case against accuracy estimation for comparing induction algorithms," in *Proceedings of the 15th International Conference on Machine Learning (ICML-98)*, Madison, WI, 1998, pp. 445-453.
[26] M. Friedman, "A comparison of alternative tests of significance for the problem of m rankings," *The Annals of Mathematical Statistics*, vol. 11, no. 1, pp. 86-92, 1940.
[27] N. Japkowicz and M. Shah, *Evaluating Learning Algorithms: A Classification Perspective*. New York, NY: Cambridge University Press, 2011.
[28] S. Chebrolu, A. Abraham, and J. P. Thomas, "Feature deduction and ensemble design of intrusion detection systems," *Computers & Security*, vol. 24, no. 4, pp. 295-307, 2005.
[29] S. Mukkamala, A. H. Sung, and A. Abraham, "Intrusion detection using an ensemble of intelligent paradigms," *Journal of Network and Computer Applications*, vol. 28, no. 2, pp. 167-182, 2005.
[30] S. Peddabachigari, A. Abraham, C. Grosan, and J. Thomas, "Modeling intrusion detection system using hybrid intelligent systems," *Journal of Network and Computer Applications*, vol. 30, no. 1, pp. 114-132, 2007.

[31] W. Hu, W. Hu, and S. Maybank, "AdaBoost-based algorithm for network intrusion detection," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 38, no. 2, pp. 577-583, 2008.
[32] J. B. D. Cabrera, C. Gutierrez, and R. K. Mehra, "Ensemble methods for anomaly detection and distributed intrusion detection in mobile ad-hoc networks," *Information Fusion*, vol. 9, no. 1, pp. 96-119, 2008.
[33] G. Giacinto, R. Perdisci, M. Del Rio, and F. Roli, "Intrusion detection in computer networks by a modular ensemble of one-class classifiers," *Information Fusion*, vol. 9, no. 1, pp. 69-82, 2008.
[34] S. S. S. Sindhu, S. Geetha, and A. Kannan, "Decision tree based light weight intrusion detection using a wrapper approach," *Expert Systems with Applications*, vol. 39, no. 1, pp. 129-141, 2012.
[35] M. Govindarajan and R. M. Chandrasekaran, "Intrusion detection using neural based hybrid classification methods," *Computer Networks*, vol. 55, no. 8, pp. 1662-1671, 2011.

Bayu Adhi Tama http://orcid.org/0000-0002-1821-6438

He received his bachelor degree in electrical engineering from Universitas Sriwijaya and master degree in information technology from Universitas Indonesia in 2004 and 2008, respectively. Currently, he is working toward a Ph.D. degree in information systems at the Laboratory of Information Security and Internet Applications (LISIA), Department of IT Convergence and Application Engineering, Pukyong National University, Korea. His research interests include data mining and machine learning techniques in cyber-security applications.

Kyung-Hyune Rhee

He received his M.S. and Ph.D. degrees from Korea Advanced Institute of Science and Technology (KAIST), Korea in 1985 and 1992, respectively. He worked as a senior researcher in Electronic and Telecommunications Research Institute (ETRI), Korea from 1985 to 1993. He also worked as a visiting scholar in the University of Adelaide, the University of Tokyo, and the University of California, Irvine. He has served as a Chairman of Division of Information and Communication Technology, Colombo Plan Staff College for Technician Education in Manila, the Philippines. He is currently a professor in the Department of IT Convergence and Application Engineering, Pukyong National University, Korea. His research interests center on key management and its applications, mobile communication security and security evaluation of cryptographic algorithms.