Using Feature Analysis to Guide Risk Calculations of Cyber Incidents

Benjamin Aziz[a] and Alaa Mohasseb[b]
School of Computing, University of Portsmouth, Portsmouth, U.K.

Keywords: Cyber Security, Datasets, Risk Analysis, Text Mining, Machine Learning.

Abstract: The prediction of incident features, for example through the use of text analysis and mining techniques, is one method by which the risk underlying Cyber security incidents can be managed and contained. In this paper, we define risk as the product of the probability of misjudging incident features and the impact such misjudgment could have on incident responses. We apply our idea to a simple case study involving a dataset of Cyber intrusion incidents in South Korean enterprises. We investigate a few problems. First, the prediction of response actions to future incidents involving malware and second, the utilisation of the knowledge of the response actions in guiding analysis to determine the type of malware or the name of the malicious code.

1 INTRODUCTION

The Internet has become the backbone for both private and public sectors due to its importance in providing the main infrastructure of communication, data transformation and services across every domain of life. However, the frequent occurrences of Cyber incidents, such as viruses, spyware, spam and other malware programs coupled with their increasing complexity over the years have caused financial losses for worldwide organisations. In a recent report published by the UK government and pwc (HM Government and PWC,), it was indicated that the cost of Cyber security incidents is on average £1.46M-£3.14M to large organisations and £75K-£311K to small organisations, per year.

According to the same report, organisations are increasingly spending more on information security purposes in order to decrease the risk of Cyber incidents. Risk, informally defined as anything that adversely impacts an organisation's business, cannot be avoided completely, but can rather be managed (Kaplan and Garrick, 1981), and the prediction of incident features, based on data mining and machine learning techniques, can play a crucial role in managing risk and reducing its impact.

For example, data mining and text analysis has widely been used in literature to detect and classify malware (e.g. (Suh-Lee et al., 2016; Kakavand et al., 2015; Norouzi et al., 2016; Fan et al., 2015; Helal and Romdhane, 2016; Lu et al., 2010; Fan et al., 2016; Rieck et al., 2011; Ding et al., 2013)) and malicious code analysis (e.g. (Bahraminikoo et al., 2012; Schultz et al., 2001; Shabtai et al., 2012)).

This paper introduces the idea that risk probability can be derived from the accuracy measure of data classification tools (Chinchor, 1992). Risk probability is seen as the complement of accuracy, and therefore, it can be combined with meaningful impact to derive risk values in a classical manner. We show how this idea can be used to evaluate the risk for a simple case study of real data representing Cyber intrusion incidents collected from a number of small and medium Korean companies, where text classification tools are trained using the current dataset to predict the values of certain features.

The rest of the paper is structured as follows. In Section 2, we give an overview of related work. In Section 3, we give an overview of the Cyber intrusion incidents dataset used in the case study. In Section 4, we discuss the experimental study and the results obtained. In Section 5, we introduce our idea that risk can be defined based on the accuracy of the classification algorithms for the class of problems being predicted. In Section 6, we apply our idea of calculating risk based on prediction accuracy to the case study dataset. Finally, in Section 7, we conclude the paper and give directions for future work.

[a] https://orcid.org/0000-0001-5089-2025
[b] https://orcid.org/0000-0003-2671-2199

Aziz, B. and Mohasseb, A.
Using Feature Analysis to Guide Risk Calculations of Cyber Incidents.
DOI: 10.5220/0011561200003318
In *Proceedings of the 18th International Conference on Web Information Systems and Technologies (WEBIST 2022)*, pages 190-195
ISBN: 978-989-758-613-2; ISSN: 2184-3252

2 RELATED WORK

The current internet technologies are plagued by Cyber security challenges and threats. Hence, it became a key element of every enterprise to protect business and secure the underlay systems. As Cyber security threats are growing to cause a venue of vulnerability for each organisation, a considerable amount of research has been conducted to consider the Cyber security challenges from different perspectives. Numerous amount of probabilistic and statistical methods for risk assessment have been proposed such as (Sommestad et al., 2010; Shin et al., 2013; Cherdantseva et al., 2016; Ruan, 2017; Paté-Cornell et al., 2018; Santini et al., 2019). However, recently, machine learning started to be applied widely in Cyber security and risk applications; this is due to the efficacy of machine learning techniques against the statistical risk models as demonstrated in (Kakushadze and Yu, 2019).

Naïve Bayes (NB), k-Nearest Neighbor and neural networks to filter spam. In addition, authors in (Lu et al., 2019) showed that a Cyber security prediction model with fewer prediction errors can be achieved by applying the Grey Wolf Optimisation algorithm (Mirjalili et al., 2014) to optimise the SVM parameters.

Furthermore, authors in (Oprea et al., 2018) proposed MADE (Malicious Activity Detection in Enterprises) to detect the malicious activities in the enterprise networks and score the risk of the external connections based on the predicted probabilities. While the combination of both supervised and unsupervised learning has been investigated in (Sarkar et al., 2019) in order to capture the Cyber security incidents such as malware and malicious emails. The study showed how the network structure of dark-web forums data can be used to predict Cyber security incidents. Moreover, the way of highlighting risk factors of network security incidents using data mining was presented in (Gounder and Nahar, 2018). The authors showed that rule mining could play a role in detecting anomaly patterns and preventing their risk. In (Gai et al., 2016) the authors' proposed Decision Tree-based Risk Prediction (DTRP) algorithm to reduce the risk of data sharing among financial firms. The proposed approach aims to predict the hazardous conditions that firms can be incurred due to information sharing. In addition, in (Huang et al., 2017) a unified risk assessment framework for SCADA networks was proposed. The proposed framework adjusts the risk parameters by learning from historical data and also incrementally from online observations. While in (Feng et al., 2017) a user-centric machine learning approach was presented to classify the Cyber security incidents and categorise them based on different risk levels. Authors in (Cheong et al., 2019) presented a new approach to quantifying a company's cyber security risk. The newly proposed method is based on text analytics and the advanced autoencoder machine learning technique. In (Figueira et al., 2020) a new predictive model for risk analysis was proposed in which the risk has been calculated based on the future threat probabilities rather than historical frequencies. Recent surveys in (Rawat et al., 2019; Torres et al., 2019) highlight more detailed works related to applications of machine learning techniques to Cyber security.

Although these studies provide important insights into the area of risk assessment and cyber security, such studies remain narrow in focus dealing only with the correctly predicted incidents without taking into account the risk impact of the wrongly predicted incidents. Therefore, in this paper, we focus on defining risk as the product of the probability of misjudging incident features and the impact such misjudgment could have on incident responses.

3 THE KAITS CYBER INTRUSION DATASET

The dataset used in our case study represents Cyber security intrusion incidents in five Small and Medium Enterprises (SMEs) in South Korea, collected over a period of ten months from 1 January 2017 until 31 October 2017 by the KAITS Industrial Technology Security Hub (KAITS,). As a public-private partnership, the Hub aims to encourage the sharing of knowledge, experience and expertise across Korean SMEs. The data for each SME is stored in a separate file. 4643 entries (as a row) and the following six features (i.e. metadata, labels) are included in the data :

- `Date and Time of Occurrence`: this is a value representing the date and time of the incident's occurrence.
- `End Device`: this is a value representing the name of the end device affected in the incident.
- `Malicious Code`: this is a value representing the name of the malicious code detected in the incident.
- `Response`: this is a value representing the response action that was applied to the malicious code.
- `Type of Malware`: this is a value representing the type of malware (malicious code) detected in the incident.
- `Detail`: this is a free text value to describe any other detail about the incident.

We focus in our case study next on two of the above features, namely `malicious code` and `response`. In addition to the above metadata, the dataset also contains statistics on the technical responses to incidents carried out by each of the five SMEs.

4 EXPERIMENTAL STUDY AND RESULTS

The objective of the experimental study is to assess the risk calculation of cyber incidents using feature analysis. Four machine learning algorithms were used for the classification process; J48 Decision tree (J48), RandomForests (RF), Naïve Bayes (NB) and Support Vector Machine (SVM). The data distribution in the KAITS dataset is shown in Table 1.

Table 1: The KAITS dataset data distribution.

Company Name	**Total Number of Incidents**
Company 1(DF)	932
Company 2(MT)	633
Company 3(SE)	923
Company 4(EP)	448
Company 5(MS)	1707

The experiments were set up using 10-fold cross-validation, and typical performance indicators were used, such as accuracy, precision, recall, and F-measure (Chinchor, 1992) and are calculated as shown in the following formulæ:

$$Accuracy = \frac{\textit{\# of correct predictions (TP+TN)}}{\textit{\# of predictions (TP+TN+FP+FN)}}$$

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

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

$$F = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Where, True Positive (TP) is a positive instance classified correctly as positive, True Negative (TN) is a negative instance classified correctly as negative, False Positive (FP) is a negative instance classified wrongly as positive and False Negative (FN) is positive instance classified wrongly as negative.

4.1 Results

In this section, we present the results of the accuracy of the machine learning algorithms. These results are summarised in Table 2

Table 2: Performance of the Classifiers for identifying the types of response based on malicious code- Best results are highlighted in bold.

Company Name	J48	SVM	RF	NB
Company 1(DF)	83%	**87%**	82%	84%
Company 2(MT)	86%	**87%**	**87%**	85%
Company 3(SE)	**89%**	**89%**	**89%**	85%
Company 4(EP)	86%	**91%**	84%	87%
Company 5(MS)	**93%**	**93%**	**93%**	89%

The overall results for the identification of the different types of responses based on the given malicious code indicated that SVM was the best classifier for all five companies in terms of performance and accuracy. In addition, most classifiers could not identify response categories such as “none”, “blocked” and “deleted”.

5 A FEATURE PREDICTION-BASED FORMULA FOR RISK

We start by reiterating the classical formula for risk, first suggested by IBM’s Robert Courtney, Jr. (Robert H. Courtney, 1977):

$$risk = probability \times impact$$

which states that risk is the product of probability and impact. Based on this, we define $\mathcal{M} = \{m_1, \ldots, m_k\}$ as the set of impact levels that an organisation would utilise. It is possible to assume further that $\mathcal{M}$ is ordered by some ordering relation $\sqsubseteq_{\mathcal{M}}$, which specifies how the values $m_1, \ldots, m_k$ compare to one another, either partially or totally. For example, $\mathcal{M}$ could refer to some monetary values or some computational values such as the increase/decrease in available processing power or time.

We assume that a Cyber incident is described by a set of features (labels), which represent the metadata for that incident. For example, in the dataset we consider here, described in the next section, there are six such features. We refer to such features by the variables $\ell_1, \ldots, \ell_k$. The impact of not predicting a particular feature of an incident, ℓ, given that all the other features are known, is defined using the following function:

$$impact(\ell) = m_\ell \in \mathcal{M}$$

In other words, *impact*(ℓ) defines the impact on the organisation in case the value of ℓ is predicted incorrectly. For example, if ℓ represents the type of response required, say from knowing the malicious code in the incident, then m_ℓ is the impact on the IT infrastructure or the organisation of misjudging this response.

The probability of making such misjudgment on a feature ℓ is referred to by the value P_ℓ defined as the complement of accuracy:

$$P_\ell = 1 - Accuracy_\ell$$

Where $Accuracy_\ell$ is the accuracy value of the classification algorithm used in predicting ℓ. Accuracy, itself, is defined by the following general formula (Chinchor, 1992):

$$Accuracy = (number\ of\ correct\ predictions/ \\ number\ of\ all\ predictions)$$

An example of P_ℓ would be the probability of predicting wrongly the type of response given the malicious code involved in an incident.

We can now define feature prediction-based risk, resulting from the incorrect prediction of some incident feature ℓ, in terms of the following equation:

$$risk_\ell = P_\ell \times m_\ell$$

We demonstrate next the application of this definition on a real case of a dataset representing Cyber intrusion incidents in Korean enterprises.

6 A RISK ANALYSIS OF THE KAITS DATASET

We explain in the following sections, through the use of a simple example from the KAITS dataset (KAITS,), our approach to the calculation of risk within the context of feature prediction in Cyber incidents.

6.1 Risk Probability

As we mentioned earlier, our main hypothesis rests on the assumption that the incorrect prediction of an incident's feature represents a risk, e.g. due to all the consequences (impact) that will result from such misjudgment. Therefore, the prediction accuracy measure can be used as a measure of risk probability.

We give here one example of measuring the accuracy of predicting the type of response to an incident given the malicious code detected in that incident.

Based on the values of Table 2, which define the $Accuracy_{response}$ variable, Table 3 presents the risk probability values for each of these and hence defining the value of $P_{response}$.

6.2 Impact

The KAITS dataset does not include any explicit information about the impact incurred as a result of the incidents, other than statistics related to numbers and types of responses. For our purposes, we shall assume a simple model based on these to illustrate how impact can be combined with the risk probabilities of the previous section.

We assume that for each company, technical response to an incident costs, on average, a single monetary unit for that company, which we term c_i (in other words, company i's single unit of currency). Table 4 represents one example of an impact factor resulting from Cyber intrusion incidents, which is the average cost per response to a ticket issued for servicing an incident. The table contains the number of tickets issued for each of the five companies based on the statistics reported in the dataset.

6.3 Risk Calculation

Based on the probability of risk and the example impact assumed, we can calculate a value for risk. Table 5 shows the risk values for each of the five companies associated with the incorrect prediction of the type of response from the malicious code based on the example impact given. The table thus represents a calculation of $risk_{response}$.

The rationale behind the data in this table is that the incorrect prediction of the type of response to an incident will lead to a misjudgment of the kind or level of service required and therefore will lead to no value in return for the cost in the worst-case scenario. Hence the numbers in the table represent the worst possible costs of incorrect predictions per algorithm parameterised by each company's currency. These numbers can be interpreted as the limit of the acceptable level when making a cybersecurity decision in the wrong way. However, the real value underlying these data will be determined by the value of the currencies themselves.

7 CONCLUSION

We demonstrated in this paper how risk can be defined based on the probability of inaccurate predictions of Cyber incident features, e.g. the kind of responses given the malicious code used in the incident, and the impact those predictions can have in terms of the number of responses served. Hence the quality of prediction determines the risk probability. We used a sample Cyber incidents dataset to demonstrate this

Table 3: Risk probability of the classifiers for identifying the types of response based on the malicious code.

Company Name/Algorithm	J48	SVM	RF	NB
Company 1(DF)	17%	13%	18%	16%
Company 2(MT)	14%	13%	13%	15%
Company 3(SE)	11%	11%	11%	15%
Company 4(EP)	14%	9%	16%	13%
Company 5(MS)	7%	7%	7%	11%

Table 4: Example impact resulting from the incidents.

Company Name	Number of response tickets served (KAITS,)	Assumed average monetary cost
Company 1(DF)	3925	$3925 \times c_1$
Company 2(MT)	13	$13 \times c_2$
Company 3(SE)	27	$27 \times c_3$
Company 4(EP)	88	$88 \times c_4$
Company 5(MS)	19	$19 \times c_5$

Table 5: Risk associated with the incorrect identification of the types of response based on malicious code.

Company Name/Algorithm	J48	SVM	RF	NB
Company 1(DF)	$667.25c_1$	$510.25c_1$	$706.5c_1$	$628c_1$
Company 2(MT)	$1.82c_2$	$1.69c_2$	$1.69c_2$	$1.95c_2$
Company 3(SE)	$2.97c_3$	$2.97c_3$	$2.97c_3$	$4.05c_3$
Company 4(EP)	$12.32c_4$	$7.92c_4$	$14.08c_4$	$11.44c_4$
Company 5(MS)	$1.33c_5$	$1.33c_5$	$1.33c_5$	$2.09c_5$

concept by applying text analysis and classification algorithms.

This approach is meaningful in that data prediction is developed to further risk analysis. Considering that risk analysis is gaining momentum in companies, a proactive approach taken in this study will work as a positive impetus for the development of the risk analysis domain. In the future, we plan to generalise this idea to other domains that carry a notion of risk, e.g. safety and reliability. This would then lead to more comprehensive definitions of risk. We also plan to extend the analysis to larger Cyber security datasets, particularly those available on open platforms such as VCDB (VERIZON,), SecRepo (Mike Sconzo,) and CAIDA (Center for Applied Internet Data Analysis,).

REFERENCES

Bahraminikoo, P., Yeganeh, M., and Babu, G. (2012). Utilization data mining to detect spyware. *IOSR Journal of Computer Engineering (IOSRJCE)*, 4(3):01–04.

Center for Applied Internet Data Analysis. CAIDA Data.

Cheong, A., Cho, S., No, W. G., and Vasarhelyi, M. A. (2019). If you cannot measure it, you cannot manage it: Assessing the quality of cybersecurity risk disclosure through textual imagification. *SSRN*.

Cherdantseva, Y., Burnap, P., Blyth, A., Eden, P., Jones, K., Soulsby, H., and Stoddart, K. (2016). A review of cyber security risk assessment methods for scada systems. *Computers & security*, 56:1–27.

Chinchor, N. (1992). Muc-4 evaluation metrics. In *Proceedings of the 4th Conference on Message Understanding*, MUC4 '92, pages 22–29, Stroudsburg, PA, USA. Association for Computational Linguistics.

Ding, Y., Yuan, X., Tang, K., Xiao, X., and Zhang, Y. (2013). A fast malware detection algorithm based on objective-oriented association mining. *computers & security*, 39:315–324.

Fan, C.-I., Hsiao, H.-W., Chou, C.-H., and Tseng, Y.-F. (2015). Malware detection systems based on api log data mining. In *Computer Software and Applications Conference (COMPSAC), 2015 IEEE 39th Annual*, volume 3, pages 255–260. IEEE.

Fan, Y., Ye, Y., and Chen, L. (2016). Malicious sequential pattern mining for automatic malware detection. *Expert Systems with Applications*, 52:16–25.

Feng, C., Wu, S., and Liu, N. (2017). A user-centric machine learning framework for cyber security operations center. In *2017 IEEE International Conference on Intelligence and Security Informatics (ISI)*, pages 173–175. IEEE.

Figueira, P. T., Bravo, C. L., and López, J. L. R. (2020). Improving information security risk analysis by including threat-occurrence predictive models. *Computers & Security*, 88:101609.

Gai, K., Qiu, M., and Elnagdy, S. A. (2016). Security-aware information classifications using supervised learning for cloud-based cyber risk management in financial big data. In *2016 IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecu-*

rity), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS), pages 197–202. IEEE.

Gounder, M. P. and Nahar, J. (2018). Practicality of data mining for proficient network security management. In *2018 5th Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE)*, pages 149–155. IEEE.

Hellal, A. and Romdhane, L. B. (2016). Minimal contrast frequent pattern mining for malware detection. *Computers & Security*, 62:19–32.

HM Government and PWC. 2015 Information Security Breaches Survey. https://www.pwc.co.uk/assets/pdf/2015-isbs-executive-summary-02.pdf.

Huang, K., Zhou, C., Tian, Y.-C., Tu, W., and Peng, Y. (2017). Application of bayesian network to data-driven cyber-security risk assessment in scada networks. In *2017 27th International Telecommunication Networks and Applications Conference (ITNAC)*, pages 1–6. IEEE.

KAITS. Industrial Technology Security Hub. http://www.kaits.or.kr/index.do.

Kakavand, M., Mustapha, N., Mustapha, A., and Abdullah, M. T. (2015). A text mining-based anomaly detection model in network security. *Global Journal of Computer Science and Technology*.

Kakushadze, Z. and Yu, W. (2019). Machine learning risk models. *Journal of Risk & Control*, 6(1):37–64.

Kaplan, S. and Garrick, B. J. (1981). On the quantitative definition of risk. *Risk analysis*, 1(1):11–27.

Lu, H., Zhang, G., and Shen, Y. (2019). Cyber security situation prediction model based on gwo-svm. In *International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, pages 162–171. Springer.

Lu, Y.-B., Din, S.-C., Zheng, C.-F., and Gao, B.-J. (2010). Using multi-feature and classifier ensembles to improve malware detection. *Journal of CCIT*, 39(2):57–72.

Mike Sconzo. SecRepo.com - Samples of Security Related Data.

Mirjalili, S., Mirjalili, S. M., and Lewis, A. (2014). Grey wolf optimizer. *Advances in Engineering Software*, 69:46 – 61.

Norouzi, M., Souri, A., and Samad Zamini, M. (2016). A data mining classification approach for behavioral malware detection. *Journal of Computer Networks and Communications*, 2016:1.

Oprea, A., Li, Z., Norris, R., and Bowers, K. (2018). Made: Security analytics for enterprise threat detection. In *Proceedings of the 34th Annual Computer Security Applications Conference*, pages 124–136. ACM.

Paté-Cornell, M.-E., Kuypers, M., Smith, M., and Keller, P. (2018). Cyber risk management for critical infrastructure: a risk analysis model and three case studies. *Risk Analysis*, 38(2):226–241.

Rawat, D. B., Doku, R., and Garuba, M. (2019). Cybersecurity in big data era: From securing big data to data-driven security. *IEEE Transactions on Services Computing*.

Rieck, K., Trinius, P., Willems, C., and Holz, T. (2011). Automatic analysis of malware behavior using machine learning. *Journal of Computer Security*, 19(4):639–668.

Robert H. Courtney, J. (1977). Security Risk Assessment in Electronic Data Processing Systems. In *Proceedings of the June 13-16, 1977, National Computer Conference*, AFIPS '77, pages 97–104, New York, NY, USA. ACM.

Ruan, K. (2017). Introducing cybernomics: A unifying economic framework for measuring cyber risk. *Computers & Security*, 65:77–89.

Santini, P., Gottardi, G., Baldi, M., and Chiaraluce, F. (2019). A data-driven approach to cyber risk assessment. *Security and Communication Networks*, 2019.

Sarkar, S., Almukaynizi, M., Shakarian, J., and Shakarian, P. (2019). Mining user interaction patterns in the darkweb to predict enterprise cyber incidents. *Social Network Analysis and Mining*, 9(1):57.

Schultz, M. G., Eskin, E., Zadok, F., and Stolfo, S. J. (2001). Data mining methods for detection of new malicious executables. In *Security and Privacy, 2001. S&P 2001. Proceedings. 2001 IEEE Symposium on*, pages 38–49. IEEE.

Shabtai, A., Moskovitch, R., Feher, C., Dolev, S., and Elovici, Y. (2012). Detecting unknown malicious code by applying classification techniques on opcode patterns. *Security Informatics*, 1(1):1.

Shin, J., Son, H., and Heo, G. (2013). Cyber security risk analysis model composed with activity-quality and architecture model. In *International Conference on Computer, Networks and Communication Engineering (ICCNCE 2013)*. Atlantis Press.

Sommestad, T., Ekstedt, M., and Johnson, P. (2010). A probabilistic relational model for security risk analysis. *Computers & security*, 29(6):659–679.

Suh-Lee, C., Jo, J.-Y., and Kim, Y. (2016). Text mining for security threat detection discovering hidden information in unstructured log messages. In *Communications and Network Security (CNS), 2016 IEEE Conference on*, pages 252–260. IEEE.

Torres, J. M., Comesaña, C. I., and García-Nieto, P. J. (2019). Machine learning techniques applied to cybersecurity. *International Journal of Machine Learning and Cybernetics*, pages 1–14.

VERIZON. VERIS Community Database.