Received: 14 July 2018 | Revised: 3 December 2018 | Accepted: 11 January 2019

DOI: 10.1002/widm.1306

WILEY WIREs DATA MINING AND KNOWLEDGE DISCOVERY

OVERVIEW

Machine learning in cybersecurity: A review

Anand Handa | Ashu Sharma | Sandeep K. Shukla

Interdisciplinary Centre for Cyber Security and Cyber Defense of Critical Infrastructures, Department of Computer Science & Engineering, Indian Institute of Technology, Kanpur, Uttar Pradesh, India

Correspondence
Sandeep K. Shukla, Interdisciplinary Centre for Cyber Security and Cyber Defense of Critical Infrastructures, Department of Computer Science & Engineering, Indian Institute of Technology, Kanpur, Uttar Pradesh, India.
Email: sandeeps@cse.iitk.ac.in

Machine learning technology has become mainstream in a large number of domains, and cybersecurity applications of machine learning techniques are plenty. Examples include malware analysis, especially for zero-day malware detection, threat analysis, anomaly based intrusion detection of prevalent attacks on critical infrastructures, and many others. Due to the ineffectiveness of signature-based methods in detecting zero day attacks or even slight variants of known attacks, machine learning-based detection is being used by researchers in many cybersecurity products. In this review, we discuss several areas of cybersecurity where machine learning is used as a tool. We also provide a few glimpses of adversarial attacks on machine learning algorithms to manipulate training and test data of classifiers, to render such tools ineffective.

This article is categorized under:
Application Areas > Science and Technology
Technologies > Machine Learning
Technologies > Classification
Application Areas > Data Mining Software Tools

KEYWORDS
adversarial learning, intrusion detection, machine learning, malware analysis

1 | INTRODUCTION

Machine Learning is based on the idea of automated learning from examples and experience, without being explicitly programmed. There are several kinds of machine learning algorithms such as supervised learning, unsupervised learning, reinforcement learning, etc. Machine learning is widely used in areas of computer vision (Harandi, Taheri, & Lovell, 2012), speech recognition (Braho, Pike, & Pike, 2018), object-recognition and also for content-based retrieval from multimedia (Lew, Sebe, Djeraba, & Jain, 2006). It is also used in building prediction systems like stock market predictions (Choudhry & Garg, 2008). It has its application to recommendation systems that help in recommending web pages, news article, television programs, and items for sale (Pazzani & Billsus, 2007). As expected these days, a lot of cybersecurity applications of machine learning techniques are seeing widespread usage. For example, machine learning is used to detect zero day attacks (He, Raghavan, Chai, & Lee, 2018) such as STUXNET (Farwell & Rohozinski, 2011), Sony Zero day attack, etc. or variants of known attacks. However, hackers or malware authors are also moving at the same pace to combat the cyber defense. The problem arises from the fact that machine learning techniques (Kotsiantis, Zaharakis, & Pintelas, 2007) were originally designed for stationary environments in which the training and test data are assumed to be generated from the same distribution. In the presence of intelligent and adaptive adversaries (Huang, Joseph, Nelson, Rubinstein, & Tygar, 2011), this working hypothesis is likely to be violated to some extent. In fact, a malicious adversary can manipulate the input data, exploiting specific vulnerabilities of learning algorithms (Barreno, Nelson, Joseph, & Tygar, 2010) to compromise the security of an entire system. In the following section, we discuss some applications of machine learning in cybersecurity and also the adversarial

threats to such methods. This paper is a review of some of the application of machine learning to cybersecurity and definitely is not exhaustive.

2 | APPLICATIONS OF MACHINE LEARNING IN CYBERSECURITY

In this section, we discuss a few applications of the machine learning in cybersecurity. Various domains such as power system security, industrial control systems, intrusion detection in supervisory control and data acquisition (SCADA) systems, intrusion detection for vehicular ad-hoc network (VANET), malware analysis, etc., have witnessed the usage of the machine learning techniques in recent days (Mitchell & Chen, 2014). In this short review, we do not aim to be comprehensive. We sample a few examples to provide the readers with a taste of the applicability of machine learning in cybersecurity.

2.1 | Power system security

Electric power systems are crucial to any country's economy and security. Blackout (Liang, Weller, Zhao, Luo, & Dong, 2017), the most severe form of power loss to a relatively wide area, that is, posting tremendous societal consequences, could result from different causes such as a fault at a power transmission line, or intentional attacks. Blackouts become widespread by initial failures propagating in a diverse and intricate cascade of rare events. However, when the power system is in a *stressed* state, unnecessary line trips can greatly exacerbate the severity of the outage, contribute to the geographical propagation of the disturbance, and may even lead to a cascading blackout. Therefore, while the power system is in a stressed state, a limited cyber attack can prove to be disastrous by cascading failures.

Many new digital technologies such as digital protection relays in power systems (Sortomme, Venkata, & Mitra, 2010), adaptive protection methods, phasor measurement units (PMUs), etc. are enhancing the cyber attack surfaces in power systems. In (Bernabeu, Thorp, & Centeno, 2012), the authors used a machine learning technique to detect the stressed condition of the power system. They used decision tree based approach to define a discrimination function to classify the system state to either as *stressed* or *safe*. The Potential predictors, that is, voltage magnitudes, angle differences, mega volt-amps (reactive) (MVAr) flows, current magnitudes, etc. obtained by PMU are used for training the classifier. For the classification, the supervised machine learning algorithm namely, decision trees (DT) (Safavian & Landgrebe, 1991) is used in predicting the appropriate reliability balance of the adaptive protective scheme based on wide area measurements. Furthermore, machine learning has been used to optimize the number of PMU's to be deployed as well as the most critical location of those PMUs. The decision tree helps in the splitting of attributes that determine the locations where PMUs are to be deployed. Despite the growing levels of complexity and uncertainty in power systems, the machine learning helps power-system engineers to respond to the challenge of planning and operating future power systems with an acceptable level of security.

2.2 | Detection of cyber attacks on industrial control systems by zone divisions

Recently cyber-attacks have become a serious threat even for control systems. Hence, we need to secure them. Machine learning algorithms can be used to have an automatic intrusion detection system that can secure control systems and critical infrastructures (CI). CI majorly rely on interconnected Information and Communication Technology (ICT) (Wihersaari, 2015) which are vulnerable. STUXNET (Farwell & Rohozinski, 2011) worm infection is an example of such a cyber-attack. STUXNET is a computer worm, which attacked the Windows-based industrial control and took over the control of Programmable Logic Controllers (PLCs) (Bolton, 2015) for causing substantial damage to Iran's nuclear program. Hence, there is an urgent need for industrial control systems (ICS) to develop reliable security and safety features. A PLC is a small computer with a built-in operating system. They are used in industry to control machine operations. If they got attacked or compromised then it may result in a huge loss. Machine learning algorithms are used to design such a detection system to detect any concealed cyber-attacks. Morita et al. (2013) have proposed an automatic detection system based on machine learning techniques. In order to carry out the experiment, a simple plant was considered and calibrated in which the hot water is circulated between two tanks that vary in heights. A SCADA system (Figueiredo & da Costa, 2012) and operators are used specifically for both the tanks. Principal component analysis (PCA) has been used to detect any abnormality. It basically detects the patterns of data being collected from the plant and SCADA systems. For example, depending on the laws of physics, if the level of the water column is high in one tank then it must be low in the other tank and vice versa. In case of abnormal patterns, PCA will report about the abnormal behavior of the two plants. This is done by training the algorithm using the data obtained from the SCADA systems. PCA helps in finding patterns to reduce the dimensions of the dataset with minimal loss of information.

There exists an adversarial attack (Huang et al., 2011) (Rubinstein et al., 2009) on the above discussed machine learning algorithm. If somehow, an intruder to the system can observe and detect the patterns of the training data then he can change

the results of the machine learning algorithm. For example, in the experiment performed by Morita, if an intruder into the system can modify the height of the column of one tank by intercepting sensor reporting the height of the column of another tank and inject false information to mislead the controller, then the algorithm will not detect the abnormality and may give false results. This adversarial effect can be detected by limiting the intrusion to a zone. According to Morita, they must place the two tanks in two different networks or zones. This reduces the probability that an intruder controls both the zones and thus it helps the defense system to detect the intrusion pattern. Such approaches combined with machine learning algorithm can be helpful to build the defense system for CI which are more robust.

2.3 | Intrusion detection in SCADA systems

SCADA systems are essential for managing and monitoring CI such as electric power generation, transmission and distribution or a water/sewage treatment plant. But the security of SCADA systems has now become more complicated because they have now been linked to the IT networks. This has been done in order to allow better linking of the field operations to the business network. It has increased the threats and risks posed by hackers on the SCADA system (Maglaras et al., 2018). Hence, there is an urge to design and develop some alert systems that can provide CI operators with a tool that can support them in detecting ongoing intrusions. These systems must provide them with a better protection from cyber-attacks such as SLAMMER worm (Moore et al., 2003). The worm affected the two United States utilities and a nuclear power plant eventually resulting in a blackout in the North-Eastern US. Hence, there is a need to develop intrusion detection systems (IDS), which are used to identify an attack and initiate proper alerts that may help take appropriate actions. IDS may fail to deal with all kinds of attacks and it may generate false alarms. These false alarms may lead to high economic risks. An analyst can apply machine learning algorithms like support vector machine (SVM) (Scholkopf & Smola, 2001) to differentiate between normal and malicious traffic. SVM is a method for classifying data. One class support vector machine (OCSVM) (Jiang & Yasakethu, 2013) which is a supervised machine learning algorithm that learns a decision function for detecting new or unknown data.

Maglaras and Jiang (2014) have proposed two distributed IDS which are able to detect attacks that occur in a SCADA system. Both of them are developed and evaluated for Cockpit CI project (Cruz et al., 2015). It is based on real-time perimeter IDS, which provides the core cyber-analysis that helps in assessing and protecting the security perimeter of each CI. Two OCSVM modules were developed and tested using the datasets from a small scale test bed. The first method is K-OCSVM which helps in distinguishing real alarms from the false alarms. It is a combination of the OCSVM method and K- means clustering method (Kanungo et al., 2002). K-OCSVM is a different method from all existing similar methods that require preselection of parameters with the use of cross-validation. Cross-validation is a validation technique for assessing how the results of a statistical analysis will generalize to an independent data set. K-OCSVM is trained offline by the help of network traces and the output of the detection module is communicated to the system by intrusion detection message detection exchange format (IDMEF) files that contain the information about the source, time and severity of the intrusion. Though the performance of this method was good as it has low overhead and has good accuracy but it achieves this by ignoring small fluctuations in the communication network that may obscure attacks.

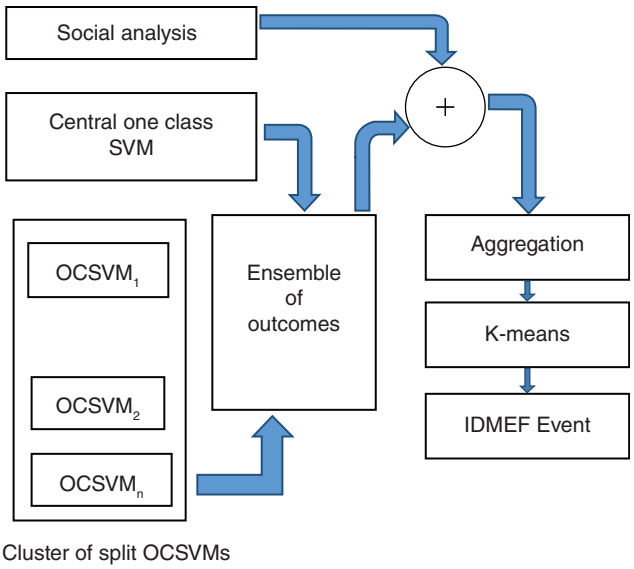

FIGURE 1 IT-OCSVM machine implementation (Maglaras & Jiang, 2014)

For another method IT-OCSVM was developed that has both high accuracy and low overhead. Figure 1 shows the implementation of IT-OCSVM. The main purpose of this is to perform anomaly detection in a time efficient way and with high accuracy and low overhead. It has the following seven stages. First, it performs the preprocessing of raw data obtained from the testbed. Second, out of time based and content based features, time based features are selected. Third, the dataset is then split on the basis of source and an OCSVM module is developed and trained for each split dataset. The cluster of split OCSVMs run in parallel with a central OCSVM. It basically produces errors targeted to a specific source. Fourth, based on the models created from the training phase the new dataset is tested. Fifth, an ensemble based mechanism is used to combine the outcomes of different OCSVM modules. In order to add weight to alerts produced from different sources Spearman's rank correlation coefficient (Sedgwick, 2014) is used. Finally, the outcomes of the different models are gathered and communication is done through IDMEF. Hence, machine learning techniques have helped the operators to build some containment strategies in case of attacks like SLAMMER worm.

2.4 | Intrusion detection systems for VANETs

VANET (Youse, Mousavi, & Fathy, 2006) is an emerging technology in modern transportation systems for providing safety and valuable information. It provides the benefits of road safety and travelers comfort while protecting driver's privacy from different types of attack. They are usually vulnerable to a number of active and passive attacks such as eavesdropping and interference respectively. IDS can be used to mitigate the threats such as control violations and unauthorized intrusions (Kumar, Srivastava, & Lazarevic, 2006) by detecting the abnormal or malicious behaviors. This detection can be done more accurately if there is a collaboration among vehicles in VANET. Distributed machine learning is a methodology for designing a collaborative detection algorithm over VANET. But the major problem to collaborative learning is that nodes exchange data among them and privacy can be compromised. A malicious node can interfere and can obtain sensitive information about other participating nodes. Zhang and Zhu (2018) have proposed a machine-learning-based collaborative IDS (CIDS) architecture. It basically trains a classifier to detect intrusions in VANET. The CIDS helps the vehicles to use the knowledge of the labeled training data of other vehicles. It basically increases the size of the training data size for each vehicle. Hence, it reduces the burden of each vehicle as this work is distributed to all the vehicles in the network. CIDS also helps the vehicles to share the information without exchanging the training data. A distributed machine learning approach called the alternating direction method of multipliers (ADMM) (Boyd, 2011) has been used. This approach helps in decentralizing the machine learning technique over a network that allows the node to share their classification results. The main concern is privacy as any malicious outsider can observe their classification results. For this, a privacy-preserving mechanism (Ev mievski, 2002) is used. It is a well-defined concept that can provide a strong privacy guarantee by which a change of any single entry of the dataset can only change the distribution of the responses of the dataset with minimal effect. Hence, this approach assures road safety and secures the driver's information.

2.5 | Malware analysis

"Malware refers to a program that is inserted into a system, usually covertly, with the intent of compromising the confidentiality, integrity, or availability of the victim's data, applications, or operating system or of otherwise annoying or disrupting the victim" (Ranveer & Hiray, 2015). Attackers exploit vulnerabilities of web applications, operating system, network services, etc. or use social engineering on victims, to execute the malware on the victim's device in order to infect it and propagate it to other devices. From the last few decades, malware are continuously growing and evolving with high pace. Now, these malware become resilient to the traditional detection system and are smart enough to evade them. Combating malware is very important for software/systems security, but to prevent it from the advanced malware, viz. polymorphic and metamorphic malware is a challenging task, as it changes the structure/code after each infection (Sharma & Sahay, 2014). Earlier, malware was detected and prevented by signature-based techniques but these techniques suffer from difficulties to detect zero-day malware or advanced malware. To overcome these limitations, machine learning techniques are gaining popularity in malware detection. The advantage of machine learning techniques is that it will not only detect known malware but also deliver knowledge for the detection of new or obfuscated malware (Rieck, Trinius, Willems, & Holz, 2011).

In malware analysis, the machine learning work in two phases as shown in Figure 2. First is the training phase, where a machine learning algorithm mathematically formalizes the set of features extracted from the known malicious and benign (nonmalicious) data to build a predictive model. These features can be extracted by static analysis, without executing the executables such as opcode or bytecode n-grams, PE header, etc. or dynamic analysis, during the execution of executable such as API or system calls, network traffic, etc. (Ranveer & Hiray, 2015). Second, in the testing phase, obtained predictive model process the properties of unknown data and predict whether it is malware or benign. The popular machine learning techniques among the researchers for the detection of second generation malware are Naive Bayes (Sharma, Sahay, & Kumar, 2016)

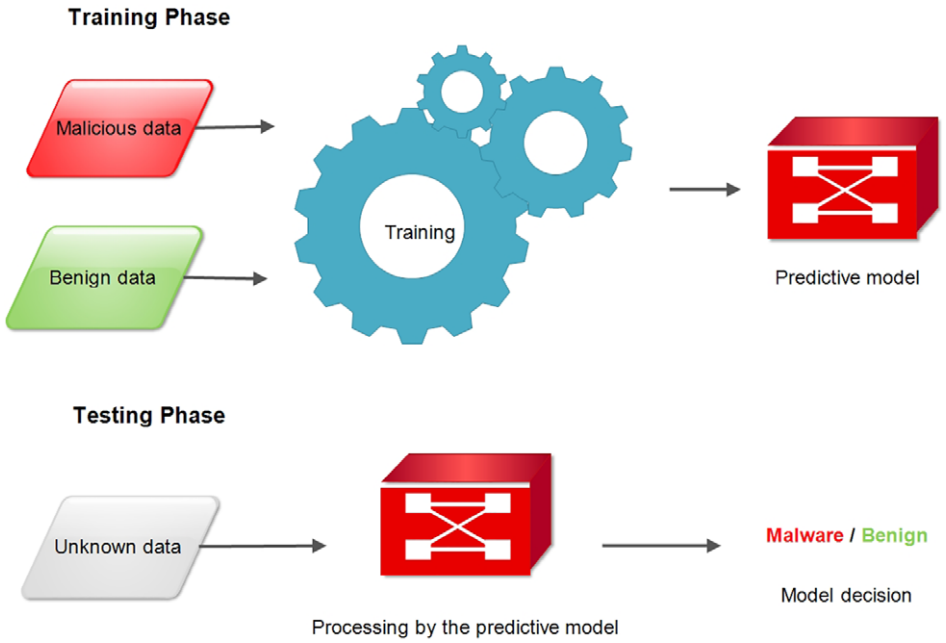

FIGURE 2 Machine learning in malware analysis

(Shabtai, Kanonov, Elovici, Glezer, & Weiss, 2012), Decision Tree (Sharma & Sahay, 2016), Data Mining (Santos, Brezo, Ugarte-Pedrero, & Bringas, 2013), Neural Networks (Dahl, Stokes, Deng, & Yu, 2013), Hidden Markov Models (Raghavan, 2018), Deep Learning (Tobiyama, Yamaguchi, Shimada, Ikuse, & Yagi, 2016), etc.

3 | ADVERSARIAL ATTACKS

In the present scenario, attackers are continually seeking for a new attack strategy to breach targets. Machine learning is an emerging field but unfortunately, like many advanced and innovative technological processes, it attracts hackers as well. Machine learning can be leveraged for both beneficial enterprise purposes as well as for malicious activity (Friedrichs, Huger, & O'donnell, 2014). As many applications such as malware detection or anomaly detection are relying on machine learning for automated decision making, there is a scope of exploiting potential vulnerabilities in machine learning algorithms to render such detection useless. An attacker can manipulate the output of a machine learning modeling during the training phase. An example of adversarial attack can be a malware that has been classified as benign due to such manipulation (Biggio, Nelson, & Laskov, 2011) (Biggio, Fumera, & Roli, 2014) (Biggio et al., 2014) (Biggio, Nelson, & Laskov, 2012).

Machine learning adversarial attacks (Huang et al., 2011) are categorized into two types of high-level attacks. One is the causative/ poisoning attack, where the attacker alters the training process by influencing the training data and harms the classifier performance whereas the second one is the exploratory or evasion attack which does not alter the training process but use other strategies, such as probing the learner or offline analysis, to discover information to manipulate the predictions of an already-trained classifier (Biggio et al., 2013). Most of the attackers focus on exploratory attacks because it may be difficult to access the training data. (Barreno, Nelson, Sears, Joseph, & Tygar, 2006). These Machine learning adversarial attacks are further classified into black-box and white-box evasion attacks (i.e., with knowledge of the model parameters) (Tramer et al., 2017).

The advancement of computation capability and network communication technology has led to the development of cyber-physical systems (CPS) (Baheti & Gill, 2011). But this advancement led to an increase in threats which are caused by adversarial attacks. If enough security is not provided to the hardware and as well as to the software assets, an attacker can manipulate the system dynamics by inducing perturbations.

Deep neural networks (DNNs) has been used in various applications of Artificial Intelligence (AI), computer vision and recognition models (Liu et al., 2017). Nowadays, these are extremely effective and flexible for extracting actionable, high-level information from the raw data produced by a wide variety of sensors in CPSs (Sonntag, Zillner, Smagt, & Lorincz, 2017). They are also used in computer security applications such as malware detection. One of the challenges in developing such models are intelligent adversaries who are actively trying to evade them by perturbing the trained model (Grosse, Papernot, Manoharan, Backes, & McDaniel, 2017).

Kolosnjaji et al. (2018) have investigated the vulnerability of malware detection methods. They have used deep neural networks to learn from raw bytes of binaries. They proposed a gradient-based attack and claimed that the attack is capable to evade a detection system (which is built on deep neural network model) by manipulating few bytes of each malware binaries and predict it as a benign binary. The results being achieved are promising and it shows that adversarial malware binaries evade the targeted defense system with high probability, even though the modification in malware binaries are less than 1%.

The various attacks described in this section depend on specific models using a particular machine learning algorithm. However, a protective system for adversarial attack for one machine learning model may not work for other models. Therefore, a regular study is required to put more sophisticated defense systems to protect machine learning systems from adversarial attacks.

4 | CONCLUSION

Machine learning provides a solution to various cyber-attack detection problems like malware detection, detection of intrusions, and most importantly in security issues related to CI like power system security, industrial control systems, intrusion detection in SCADA systems, intrusion detection for VANET, etc. These problems involve efficient and effective training and classification of data in huge volumes. The existence of adversarial attacker who can evade such tools by manipulating the classifiers is a major growing concern. This review provides a sample of some areas of cybersecurity where machine learning is being used. Also, threats of adversary attacks which can manipulate training and test data for classifiers has been discussed. Attacks are used to manipulate the model based predictions with a malicious objective. The goal of this review is to create awareness on machine learning applications in cybersecurity. A sample of approaches that have been used by adversaries to threat current machine learning based defense against cyber attacks is presented.

CONFLICT OF INTEREST

The authors have declared no conflicts of interest for this article.

REFERENCES

Baheti, R., & Gill, H. (2011). Cyber-physical systems. *The Impact of Control Technology*, *12*(1), 161–166.

Barreno, M., Nelson, B., Joseph, A. D., & Tygar, J. (2010). The security of machine learning. *Machine Learning*, *81*(2), 121–148.

Barreno, M., Nelson, B., Sears, R., Joseph, A. D., & Tygar, J. D. (2006). Can machine learning be secure? In *Proceedings of the 2006 ACM symposium on information, computer and communications security* (pp. 16–25). New York, NY: ACM.

Bernabeu, E. E., Thorp, J. S., & Centeno, V. (2012). Methodology for a security/dependability adaptive protection scheme based on data mining. *IEEE Transactions on Power Delivery*, *27*(1), 104–111.

Biggio, B., Corona, I., Maiorca, D., Nelson, B., Srndic, N., Laskov, P., … Roli, F. (2013). Evasion attacks against machine learning at test time. In *Joint European conference on machine learning and knowledge discovery in databases* (pp. 387–402). Berlin, Heidelberg: Springer.

Biggio, B., Fumera, G., & Roli, F. (2014). Security evaluation of pattern classifiers under attack. *IEEE Transactions on Knowledge and Data Engineering*, *26*(4), 984–996.

Biggio, B., Nelson, B., & Laskov, P. (2011). Support vector machines under adversarial label noise. In *Asian conference on machine learning* (pp. 97–112). Cambridge, MA: Microtome Publishing.

Biggio, B., Nelson, B., & Laskov, P. (2012). Poisoning attacks against support vector machines. *arXiv*, *1206*, 6389.

Bolton, W. (2015). *Programmable logic controllers*. Oxford, England: Elsevier Science & Technology.

Boyd, S. (2011). Alternating direction method of multipliers. In *Talk at nips workshop on optimization and machine learning*. Boston: Now publishers.

Braho, K. P., Pike, J. P., & Pike, L. A. (2018, March 27). *Methods and systems for identifying errors in a speech recognition system.* Google Patents. (US Patent 9,928,829).

Choudhry, R., & Garg, K. (2008). A hybrid machine learning system for stock market forecasting. *World Academy of Science, Engineering and Technology*, *39*(3), 315–318.

Cruz, T., Barrigas, J., Proenca, J., Graziano, A., Panzieri, S., Lev, L., & Simões, P. (2015). Improving network security monitoring for industrial control systems. In *IFIP/IEEE International Symposium on Integrated Network Management (IM) IM 2015*. (pp. 878–881).

Dahl, G. E., Stokes, J. W., Deng, L., & Yu, D. (2013). Large-scale malware classification using random projections and neural networks. In *2013 I.E. international conference on acoustics, speech and signal processing (ICASSP)* (pp. 3422–3426). New York, NY: IEEE.

Ev mievski, A. (2002). Randomization in privacy preserving data mining. *ACM Sigkdd Explorations Newsletter*, *4*(2), 43–48.

Farwell, J. P., & Rohozinski, R. (2011). Stuxnet and the future of cyber war. *Survival*, *53*(1), 23–40.

Figueiredo, J., & da Costa, J. S. (2012). A SCADA system for energy management in intelligent buildings. *Energy and Buildings*, *49*, 85–98.

Friedrichs, O., Huger, A., & O'donnell, A. J. (2014, October 28). *Method and apparatus for detecting malicious software using machine learning techniques.* Google Patents. (US Patent 8,875,286).

Grosse, K., Papernot, N., Manoharan, P., Backes, M., & McDaniel, P. (2017). Adversarial examples for malware detection. In *European symposium on research in computer security* (pp. 62–79). Cham, Switzerland: Springer.

Harandi, M., Taheri, J., & Lovell, B. C. (2012). Machine learning applications in computer vision. In *Machine learning algorithms for problem solving in computational applications: Intelligent techniques* (pp. 99–132). Hershey, Pennsylvania: IGI Global.

He, Z., Raghavan, A., Chai, S., & Lee, R. (2018). Detecting zero-day controller hijacking attacks on the power-grid with enhanced deep learning. *arXiv preprint arXiv: 1806.06496.*

Huang, L., Joseph, A. D., Nelson, B., Rubinstein, B. I., & Tygar, J. (2011). Adversarial machine learning. In *Proceedings of the 4th ACM workshop on security and artificial intelligence* (pp. 43–58). New York, NY: ACM.

Jiang, J., & Yasakethu, L. (2013). Anomaly detection via one class SVM for protection of SCADA systems. In *2013 International conference on cyber-enabled distributed computing and knowledge discovery (CYBERC)* (pp. 82–88). New York, NY: IEEE.

Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., & Wu, A. Y. (2002). An efficient k-means clustering algorithm: Analysis and implementation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, *7*, 881–892.

Kolosnjaji, B., Demontis, A., Biggio, B., Maiorca, D., Giacinto, G., Eckert, C., & Roli, F. (2018). Adversarial malware binaries: Evading deep learning for malware detection in executables. *arXiv*, *1803*, 04173.

Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. *Emerging Artificial Intelligence Applications in Computer Engineering*, *160*, 3–24.

Kumar, V., Srivastava, J., & Lazarevic, A. (2006). *Managing cyber threats: Issues, approaches, and challenges* (Vol. 5). New York, NY: Springer Science & Business Media.

Lew, M. S., Sebe, N., Djeraba, C., & Jain, R. (2006). Content-based multimedia information retrieval: State of the art and challenges. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, *2*(1), 1–19.

Liang, G., Weller, S. R., Zhao, J., Luo, F., & Dong, Z. Y. (2017). The 2015 Ukraine blackout: Implications for false data injection attacks. *IEEE Transactions on Power Systems*, *32*(4), 3317–3318.

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, *234*, 11–26.

Maglaras, L. A., & Jiang, J. (2014). Intrusion detection in SCADA systems using machine learning techniques. *Science and information conference (SAI)*, 626–631.

Maglaras, L. A., Kim, K.-H., Janicke, H., Ferrag, M. A., Rallis, S., Fragkou, P., … Cruz, T. J. (2018). Cyber security of critical infrastructures. *ICT Express*, *4*, 42–45.

Mitchell, R., & Chen, I.-R. (2014). A survey of intrusion detection techniques for cyber-physical systems. *ACM Computing Surveys (CSUR)*, *46*(4), 55.

Moore, D., Paxson, V., Savage, S., Shannon, C., Staniford, S., & Weaver, N. (2003). Inside the slammer worm. *IEEE Security & Privacy*, *99*(4), 33–39.

Morita, T., Yogo, S., Koike, M., Hamaguchi, T., Jung, S., Koshijima, I., & Hashimoto, Y. (2013). Detection of cyber-attacks with zone dividing and PCA. *Procedia Computer Science*, *22*, 727–736.

Pazzani, M. J., & Billsus, D. (2007). Content-based recommendation systems. In *The adaptive web* (pp. 325–341). Berlin, Heidelberg: Springer.

Raghavan, A. (2018). *Boosted hidden markov models for malware detection. Master's Projects, 623*. https://scholarworks.sjsu.edu/etd_projects/623

Ranveer, S., & Hiray, S. (2015). Comparative analysis of feature extraction methods of malware detection. *International Journal of Computer Applications*, *120*(5), 1–7.

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

Rubinstein, B. I., Nelson, B., Huang, L., Joseph, A. D., Lau, S.-h., Rao, S., … Tygar, J. (2009). Stealthy poisoning attacks on PCA-based anomaly detectors. *ACM SIGMETRICS Performance Evaluation Review*, *37*(2), 73–74.

Safavian, S. R., & Landgrebe, D. (1991). A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man, and Cybernetics*, *21*(3), 660–674.

Santos, I., Brezo, F., Ugarte-Pedrero, X., & Bringas, P. G. (2013). Opcode sequences as representation of executables for data-mining-based unknown malware detection. *Information Sciences*, *231*, 64–82.

Scholkopf, B., & Smola, A. J. (2001). *Learning with kernels: Support vector machines, regularization, optimization, and beyond*. Cambridge, MA: MIT Press.

Sedgwick, P. (2014). Spearmans rank correlation coefficient. *BMJ*, *349*, g7327.

Shabtai, A., Kanonov, U., Elovici, Y., Glezer, C., & Weiss, Y. (2012). Andromaly: A behavioral malware detection framework for android devices. *Journal of Intelligent Information Systems*, *38*(1), 161–190.

Sharma, A., & Sahay, S. K. (2014, March). Article: Evolution and detection of polymorphic and metamorphic malwares: A survey. *International Journal of Computer Applications*, *90*(2), 7–11.

Sharma, A., & Sahay, S. K. (2016). An effective approach for classification of advanced malware with high accuracy. *International Journal of Security and Its Applications*, *10*(4), 249–266.

Sharma, A., Sahay, S. K., & Kumar, A. (2016). Improving the detection accuracy of unknown malware by partitioning the executables in groups. In *Advanced computing and communication technologies* (pp. 421–431). South Korea: Science and Engineering Research Support Society.

Sonntag, D., Zillner, S., van der Smagt, P., & Lorincz, A. (2017). Overview of the CPS for smart factories project: deep learning, knowledge acquisition, anomaly detection and intelligent user interfaces. In *Industrial internet of things* (pp. 487–504). Cham, Switzerland: Springer.

Sortomme, E., Venkata, S., & Mitra, J. (2010). Microgrid protection using communication-assisted digital relays. *IEEE Transactions on Power Delivery*, *25*(4), 2789–2796.

Tobiyama, S., Yamaguchi, Y., Shimada, H., Ikuse, T., & Yagi, T. (2016). Malware detection with deep neural network using process behavior. In *2016 I.E. 40th annual computer software and applications conference (COMPSAC)* (Vol. 2, pp. 577–582). New York, NY: IEEE.

Tramer, F., Kurakin, A., Papernot, N., Goodfellow, I., Boneh, D., & McDaniel, P. (2017). Ensemble adversarial training: Attacks and defenses. *arXiv*, *1705*, 07204.

Wihersaari, K. (2015). Intelligence acquisition methods in cyber domain: Examining the circumstantial applicability of cyber intelligence acquisition methods using a hierarchical model.

Youse, S., Mousavi, M. S., & Fathy, M. (2006). Vehicular ad hoc networks (vanets): challenges and perspectives. In *2006 6th international conference on its telecommunications proceedings* (pp. 761–766). New York, NY: IEEE.

Zhang, T., & Zhu, Q. (2018). Distributed privacy-preserving collaborative intrusion detection systems for vanets. *IEEE Transactions on Signal and Information Processing over Networks*, *4*(1), 148–161.

How to cite this article: Handa A, Sharma A, Shukla SK. Machine learning in cybersecurity: A review. *WIREs Data Mining Knowl Discov*. 2019;e1306. https://doi.org/10.1002/widm.1306