

IDRBT JOURNAL OF BANKING TECHNOLOGY

Volume 2 | Number 1 | Jan - Jun 2018



**Institute for Development and Research
in Banking Technology**
(Established by Reserve Bank of India)

Editor-in-Chief

Dr. A. S. Ramasastr

Institute for Development and Research
in Banking Technology (IDRBT), India.

Editorial Board

Prof. Chin-Teng Lin, University of
Technology, Australia.

Prof. Constantin Zopounidis, Technical
University of Crete, Greece.

Prof. Dirk Van den Poel, Ghent
University, Belgium.

Prof. D. B. Phatak, Indian Institute of
Technology, Bombay, India.

Prof. Kalyanmoy Deb, Michigan State
University, USA.

Mr. Marc Hollanders, Bank for
International Settlements, Switzerland.

Mr. Massimo Cirasino, The World
Bank, USA.

Prof. Paolo Bellavista, Università di
Bologna, Italy.

Prof. Rajkumar Buyya, University of
Melbourne, Australia.

Prof. R. K. Shyamasundar, Indian
Institute of Technology, Bombay, India.

Prof. Sushil Jajodia, George Mason
University, USA.

Mr. Thomas Lammer, European Central
Bank, Germany.

Prof. Venu Govindaraju, University at
Buffalo, USA.

Aims and Scope

The aim of the IDRBT Journal of Banking Technology (IJBT) is to promote new thinking, conceptual frameworks and research/practice-oriented innovative ideas on topics of current interest and broad relevance to application of Technology in Banking and Financial Services.

The scope of the journal covers all aspects of technology, which directly or indirectly contributes to the technological growth of the banking and financial sector, both from researchers' as well as practitioners' perspectives. It publishes original research/practice articles on all aspects of computing and communication technologies, which are/can be used in banking and finance, including case studies, experimental and survey articles.

IDRBT Journal of Banking Technology (IJBT)

Copyright Information

For Authors

Authors retain copyright of their work and sign an exclusive licensing agreement, which grants IDRBT the right but not the obligation to defend the work against improper use by third parties.

Authors can post the accepted, peer-reviewed version – known as the “pre-print” – to the following sites, with a Link to the Definitive Version of Record on the Journal Website:

- On Author’s own Home Page and
- On Author’s Institutional Repository and
- In any repository legally mandated by the agency funding the research on which the work is based.

Authors can reuse any portion of their own work in a new work of their own as long as a citation and link to the Version of Record on the Journal Website are included.

Authors can include partial or complete papers of their own in a dissertation as long as citations and link to the Versions of Record in the journal website are included. Authors can use any portion of their own work in presentations and in the classroom.

For Readers

While the advice and information in this journal is believed to be true and accurate at the date of its publication, neither the authors, the editors, nor the publisher can accept any legal responsibility for any errors or omissions that may have been made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

All articles published in this journal are protected by copyright, which covers the exclusive rights to reproduce and distribute the article (e.g., as offprints), as well as all translation rights. No material published in this journal may be reproduced photographically or stored on

microfilm, in electronic data bases, on video disks, etc., without first obtaining written permission from the publisher. The use of general descriptive names, trade names, trademarks, etc., in this publication, even if not specifically identified, does not imply that these names are not protected by the relevant laws and regulations.

For permission to reuse any content, please write to ijbtqueries@idrbt.ac.in.

Journal Website

<http://www.idrbt.ac.in/ijbt.html>

Subscription Information

The IDRBT Journal of Banking Technology is published twice a year in January and July. For information on subscription rates, please contact: publisher@idrbt.ac.in

Advertisements

Please contact publisher@idrbt.ac.in for advertisement rates.

Disclaimer

IDRBT publishes advertisements in this journal in reliance upon the responsibility of the advertiser to comply with all legal requirements relating to the marketing and sale of products or services advertised. IDRBT and the editors are not responsible for claims made in the advertisements published in the journal. The appearance of the advertisement in IDRBT publications does not constitute endorsement, implied or intended, of product advertised or for the claims made for it by the advertiser.

Office of Publication

Publications Office,
Institute for Development and Research in
Banking Technology,
Castle Hills, Road No.1, Masab Tank,
Hyderabad-500 057, India.
e-mail: publisher@idrbt.ac.in

© IDRBT 2018

Editorial

Dr. A. S. Ramasastrri

Research Articles

1. From prediction to anticipation of cyber attacks 01
Michael Weiss
2. Decision-making under uncertainty 12
Monika, Hao Quan, Dipti Srinivasan
3. Machine learning in computer security 23
V. Rao Vemuri

Practitioners' Perspective

4. Innovating a seamless customer experience 41
Bipin Sahni
5. National payment system – overview of regulatory mandates 47
Gynedi Srinivas, Harish Natarajan

Dr. A. S. Ramasastrri

One of the primary objectives of banks is to deliver what customer needs. Technology has been assisting banks in achieving this objective by providing appropriate solutions. Quite often than not, the technology solutions do leave a few vulnerabilities. It is exactly here that the researchers and practitioners play an important role – to help banks in building secure, robust and convenient products and services.

Traditionally, banks have been using technology solutions that can do routine day-to-day activities more accurately and efficiently. Over the past few years, the scope of technology solutions has changed. There have been useful proven techniques to assist banks in prediction and decision-making due to recent developments in fuzzy systems and neural networks. In-store robots, software robotics, chatbots, smart cameras and self-learning risk models are being used by banks.

The keywords in today's banking technology are artificial intelligence, cyber defence and digital payments. We are happy that the papers in the current (second) issue of our Journal on Banking Technology focusses on these important areas.

There are three stages in building good cyber defence systems. During the first stage, detection and prevention systems are put in place; such systems are mostly based on past knowledge. During the second stage, systems are built to predict and prevent cyber attacks; such systems are based on past knowledge and future predictive power. It is only in the final stage that systems are built to anticipate attacks and prevent them; they are in the realm of the unknown. The first paper in the journal by Michael Weiss titled "From Prediction to Anticipation of Cyber Attacks" is about the developments in this area.

More uncertain the environment, more difficult the decision-making of an individual or an institution. But, it is only the uncertainty that warrants effective decision-making. The second paper by Monika, Hao Quan and Dipti Srinivasan titled "Decision-Making under Uncertainty" discusses the issues of decision-making in detail.

Security is of utmost concern to banks. More secure the systems of banks, more trust the customers will have in banks. And banks function on the principle of trust.

Dr. A. S. Ramasastrri (✉)

Director, Institute for Development and Research in Banking Technology, Hyderabad, India
e-mail: asramasastrri@idrbt.ac.in

Security is at the core of trust. The third paper by Rao Vemuri titled “Machine Learning in Computer Security” dwells on the use of machine learning for security.

While Bipin Sahni presents from his experience “Innovating a Seamless Customer Experience” in the fourth paper, Gynedi Srinivas and Harish Natarajan, the researchers from World Bank give a detailed account of “National Payment System – Overview of Regulatory Mandates” in the fifth paper.

We trust the papers prove useful to both academia and industry.

Dr. A. S. Ramasastry
Editor-in-Chief

From prediction to anticipation of cyber attacks

Michael Weiss¹

Abstract With the rising volume and variety of cyber attacks, it has become increasingly harder for businesses and organizations to defend against attacks. The paper makes the case that to respond to this challenge, we need to anticipate new threats, not merely react to known threats. It reviews reactive approaches to cyber attacks where current actions are based on past behavior, and proactive approaches guided by predictions about the future.

Keywords: Cybersecurity · Prediction · Anticipation · Threats · Cyber attacks

1 Introduction

In this paper, we discuss that to address the mounting threat of cyber attacks we need to anticipate new threats, not merely react to known threats. With the rising volume and variety of cyber attacks, it has become increasingly harder for businesses and organizations to defend against attacks. The paper makes the case that to respond to this challenge, we need to switch from a reactive to a more proactive approach to cybersecurity.

The existing approach to cybersecurity has been mostly reactive. For example, traditional mechanisms to defend against malware are based on matching attacks against known signatures. As new strains of malware are discovered, signatures are added to the list of known attacks. This approach works only as long as the volume and variety of attacks is low. With the increase in the number of attacks, however, by the time a new attack has been identified, significant damage may already have been done [1].

The literature on the cognitive basis of prediction [2–4] provides an overarching perspective for this paper. As suggested by [2], prediction comprises two types of activities: on one hand, forecasting or prediction in the narrow sense, and anticipation on the other. The key distinction between both is that in the former, current actions are based on past behavior and that in the latter, predictions about the future guide current actions.

In the remainder of the paper, after describing the method used to conduct the review of existing approaches, we first describe approaches to predicting cyber attacks from past behavior. We then argue that to address the challenges imposed by the

✉ Michael Weiss:
michael_weiss@carleton.ca

¹ Department of Systems and Computer Engineering, Carleton University, Ottawa, Canada

rapidly changing cybersecurity environment, approaches inspired by the study of anticipatory thinking are required.

2 Method

The review of techniques reported in this paper was conducted by first identifying the candidate papers through databases like Google Scholar. In a second step, by following references, for applying topic modeling to extract the underlying, latent themes from those papers. Topic modeling provides an alternative to manual clustering of articles and allows us to identify non-obvious connections between the ideas expressed in the selected papers.

2.1 Selection of papers on prediction and anticipation

For this paper, we collected a set of 108 documents related to reactive and proactive approaches to prediction. The papers included both highly cited articles matching either the keywords “machine learning” or “anticipation” and “cybersecurity”. Also included are key papers cited by those papers, e.g., papers on the cognitive foundation of anticipatory thinking. To allow for emerging topics to be represented in the collection, we also handpicked recent conference papers and theses with a lower number of citations. A more thorough systematic review of the literature is the subject of future work.

2.2 Creating the topic model

Topic modeling is a probabilistic technique for extracting latent topics from a set of documents [5, 6]. It does not require a human to label the documents, and, thus, belongs to the class of unsupervised learning algorithms. Topic modeling has been applied to areas such as analyzing emerging trends of security vulnerabilities [7] and the evolution of scientific fields [8].

A common topic modeling technique is Latent Dirichlet Allocation (LDA) [5]. In LDA, documents are represented as a bag of words, i.e., the order of the words does not matter. Given a sample of documents and the number of topics, LDA produces a distribution $P(z/d)$ that a document d is about a given topic z and a distribution $P(w/z)$ that a topic z is associated with a word w .

To construct the topic model, we used only paper abstracts as documents. There is a balance between the length of the documents provided to a topic model, the number of potential topics each document contains, and the ease with which we can understand the created topic model. By focusing on the abstracts, we emphasize highlights of the articles as summarized by its authors. Abstracts can also be scanned more quickly than the full articles when examining the articles associated with a given topic.

Table 1 The major latent topics in prediction and anticipation research

ID	Weight	Keywords	Name
9	0.583	security, information, detection, systems, network, system, attacks	Intrusion detection
8	0.243	cybersecurity, game, failures, domain, afd, failure	Adversarial thinking
6	0.238	data, prediction, learning, science, processing, model, mining	Predictive analytics
4	0.230	software, vulnerabilities, vulnerability, metrics, code, security, development	Vulnerability prediction
5	0.225	attack, risk, engineering, scenarios, cyber, set, graphs	Scenario modeling
3	0.158	attacks, anticipation, hijacking, types, emails, phishing, mechanisms	Complex attack detection
7	0.103	identity, theft, botnet, breach, news, company, software	Text and network mining
1	0.098	anomalies, events, performance, feature, ability, hros, extrapolation	Anticipation of failure
2	0.097	malware, patterns, threat, behavior, classification, hotspot, file	Behavioral analysis

We then created a topic model, and iterated the model with different numbers of topics, until a set of mostly independent clusters of documents emerged. The literature also suggests that 10-12 topics are a good heuristic value for the number of topics [9]. Table 1 shows the output of this step. For each topic, the keywords associated with the topics produced by the topic model, a name assigned by the researcher, and the topic weight are listed.

As apparent from Table 1, the topic “intrusion detection” is the topic with the highest weight and the topic “behavioral analysis” is the topic with the lowest weight.² The clusters obtained by topic modeling provide the basis for our review of prediction and anticipation techniques. The application of topic modeling enabled the discovery of subtle connections between articles via common topics that may have easily been missed in a manual expert review.

3 Prediction

All techniques reviewed in this section assume that to predict the future you are restricted to examining the past. This premise covers the most current machine learning and predictive analytics techniques. The techniques correspond to six of nine topics in Table 1 and comprise the majority of the articles found. They include predictive analytics, intrusion detection, behavioral analysis, text and network mining, complex attack detection, and vulnerability prediction.

² A high topic weight indicates an area that much of the existing research has focused on, whereas a low weight often suggests an emerging research area.

3.1 Predictive analytics and machine learning

Predictive analytics is the art of building and using models to make predictions. It uses machine learning to build those models [10]. There is no fixed set of methods, yet, for applying machine learning to cybersecurity. Some peculiarities of cybersecurity also make it more challenging to apply machine learning: the evolution of attacks that requires learning to be incremental; a high data volume; a high cost of errors; the need to label training data which requires substantial effort preparing the data; and a lack of data sets [11].

A survey of machine learning techniques for cybersecurity is provided in [12]. Machine learning techniques can be grouped into supervised, unsupervised, and hybrid techniques. Supervised methods require data instances to be assigned categories or labels (e.g., “spam” or “ham”). These include methods like decision trees, Naive Bayes and Support Vector Machines. Unsupervised methods do not require labels. They include k-means clustering and association mining. Hybrid methods can be used in supervised or unsupervised modes, and include neural networks, genetic algorithms, and Bayesian networks.

One particular challenge in the cybersecurity context is that machine learning models can themselves be attacked [13]. Through carefully crafted attacks, attackers can gain an understanding of the internal state of a machine learning model, which allows them to attack more effectively in the future. Attacks against machine learning models fall into two categories: integrity attacks (an attacker tries to get the algorithm to accept a harmful attack as benign) and availability attacks (attackers train the model to classify benign instances as harmful) [13]. Recent research [14] calls for algorithms that can unlearn what they had incorrectly “learned” from attacks against them.

3.2 Intrusion detection

Intrusion detection is the process of monitoring a network or system for intrusions or attacks [15]. Common types of attacks include scanning attacks which are used to gather information about a network or system, penetration attacks in which an attacker tries to gain unauthorized access to a system, and Denial of Service (DoS) attacks that aim to exhaust the resources of a network. Systems built to detect intrusion attempts can be grouped into misuse and anomaly detection [15]. Misuse detection can detect known attacks (i.e., abnormal behavior) with predefined characteristics. Conversely, anomaly detection considers intrusions to be deviations from normal behavior.

Manual analysis of intrusions is limited to mitigating known attacks. As attacks evolve, human-created rules used to detect them become ineffective [1, 16]. Due to the volume, variety, intensity, and velocity of data that they have to process analysts are also missing many malicious security events [16]. Machine learning helps automate the analysis of attacks. Machine learning methods for misuse detection are typically based on a classification of attacks against signatures [12]. Clustering algorithms can support anomaly detection. They can also be used to extract new signatures for misuse detection [12].

In recent work, hybrid analyst-in-the-loop approaches have been proposed that combine the best of human analysts and machine learning algorithms [17, 18]. They respond to the problems of how to present analysts with the right information and the lack of current datasets for training machine learning classifiers. Analysts can become

overwhelmed with monitoring real-time security events given their limited time “budget” to investigate alerts. One solution is to combine automation (rule-based detection of potential threats) and exploration (visualization of the associated information) [17]. Exploration allows analysts to discover new anomalies that are not yet covered by rules, select features useful for detection, and validate existing rules.

The lack of labeled data makes it difficult to use supervised machine learning models. The system proposed in [18] combines unsupervised and supervised learners. It uses outlier detection to identify suspicious activities. Suspicious activities are presented to the analyst who labels activities as actual attacks or normal behavior. These labels can then be used to train a supervised learner. The analyst helps the system identify new and evolving attacks, while machine learning can predict known attacks without input from the analyst.

3.3 Behavioral analysis

The traditional approach to detecting malware is to define signatures and match incoming malware samples against them. Manually defining signatures is a laborious process and inadequate to keep up with changing strains of malware. There are over 6 million new strains of malware each year [19]. It is also easy for an attacker to circumvent signature-based approaches. Malware can be modified without altering its behavior but changing its signature.

More recent approaches to malware detection are based on behavioral analysis. Behavioral analysis profiles malware by creating a trace of the instructions it calls and uses a combination of clustering and classification techniques [20]. It does not require a manually labeled set of malware samples. Instead, malware samples are initially clustered by the similarity of their behavioural profiles. These clusters can then be used as labels to train a classifier. Behavioral analysis is also robust against the evolution of malware. Even when changes are made to the malware, its behavior will be similar.

3.4 Text and network mining

Text and network mining algorithms have shown much promise for predicting cyber attacks. Methods discussed elsewhere in this paper heavily rely on text analysis. For example, behavioral profiles used for malware analysis (section 3.3) contain sequences of instructions and their arguments that the malware invokes on the host system. These can be translated into features by extracting n-grams from those sequences and creating a bag-of-words representation on which the classification and clustering algorithms can operate [20].

Similarly, methods for vulnerability prediction (section 3.6) often rely on documents (e.g., CVE reports) or sources that can be interpreted as documents (e.g., source code). In [7], topic modeling is used to examine trends in CVE (Common Vulnerability and Exposure) reports. The authors first identify latent topics in the reports to categorize vulnerabilities. They then compute the weight of each topic across different years to understand vulnerability trends. In [21], source code files are treated as documents and symbols in the code as words. A bag-of-words

representation of the code provides input to a classifier that predicts the likelihood of a software system containing vulnerabilities.

Network mining applies techniques from social network analysis to help identify, e.g., the central nodes in a network, clusters of nodes, or the roles nodes play in a network. Nodes represent actors or systems and the ties that connect them stand for their relationships. Network mining has been applied to analyze botnets [22]. The authors create a network model of the communication patterns between the hosts of a botnet using NetFlow data and analyze the network using a PageRank algorithm to identify the central hosts.

3.5 Complex attack detection

Increasingly, attacks are executed in multiple steps, making them harder to detect. Such complex attacks require that defenders recognize the separate stages of an attack, possibly carried out over a longer period, as belonging to the same attack. Complex attacks can be divided into exploration and exploitation phases [23]. Exploration involves identifying vulnerabilities and scanning and testing a system. It is how an attacker gathers information about the system. Exploitation involves gaining and maintaining access. At this stage, the attacker applies the know-how gathered during the exploration stage.

An example of a complex attack that combines exploration and exploitation is a sequence of a phishing attack, followed by an exfiltration attack. First, attackers will attempt to collect information on the organization they intend to attack, e.g., names of key employees. Then, they will craft a targeted phishing attack. The phishing attack allows the attackers to gain access to the user's system and install malware. The purpose of the malware could be to extract files from the user's machine or to use the user's machine as an attack vector to attack other machines in the organization's network.

A phishing attack is usually carried out by sending an email purporting to come from a trusted source and tricking its receiver to click on a URL that results in installing malware on the user's system. This malware then creates a backdoor into the user's system for staging a more complex attack. Phishing attacks can be recognized both by the types of keywords used in the email (as with a spam email), as well as by the characteristics of URLs included in the message [24]. Features that have been used successfully to detect phishing attacks include URLs that include IP addresses, the age of a linked-to domain, and a mismatch between anchor and text of a link.

3.6 Vulnerability prediction

Vulnerabilities are weaknesses, flaws or deficiencies that can be exploited by threats to cause harm to an asset. This section focuses on software vulnerabilities. Given that not all vulnerabilities are of equal impact and that resources are limited, authors of software need to prioritize on which patches to create and system administrators on which of these patches to deploy. Vulnerability prediction can assist in this task by

predicting the kinds of vulnerabilities that exist in a system and the risk of them being exploited.

There are two ways to predict vulnerabilities: based on metadata (i.e., information about the vulnerabilities) or from inherent properties of a system. A natural question to ask is whether we can predict the timing and impact of an exploit from the information in a CVE report. A classifier for this task is described in [25]. The features it uses include text fields of the report (e.g., description), timestamps (e.g., the time between the first time the vulnerability was reported and the time of its exploit), and cross-references to other reports. As new information about vulnerabilities that have been exploited becomes available, the classifier can be retrained to incorporate this information.

A machine learning model to predict the likelihood that a software system contains vulnerabilities from the system's source code (an inherent property) has been described in [21]. This approach trains a classifier directly on the source code rather than on quality metrics derived from the code. It was even found to be capable of predicting vulnerabilities in future releases of the same software. It has been demonstrated that architectural flaws (another inherent property) are also good indicators of security issues [26]. Changes to the architecture (including patches to fix security issues) can result in new vulnerabilities. The authors calculate structural (dependencies between files) and evolutionary metrics (co-changes among files) from source code and its revision history. Certain patterns in those metrics (e.g., frequent changes) indicate architectural flaws known to correlate with security issues.

4 Anticipation

The techniques in this section allow us to select actions based on their anticipated consequences. They furthermore enable us to operate in a continually evolving environment. This ability sets anticipatory techniques apart from predictive techniques. Techniques in this section include adversarial thinking, scenario modeling, and anticipation of failure. They are embodied in some of the more recent approaches to dealing with cyber attacks.

4.1 Adversarial thinking

Game theory studies the strategic interaction between players. Training in game theory has been shown to help sensitize students to the role of human adversaries in cybersecurity [27]. It can also be used to model multiple levels of reasoning like level- k reasoning (e.g., a level-2 strategy would be for operators to expect attackers to try to anticipate their moves and to act accordingly). Game theoretic models have also been implemented in algorithms to protect critical infrastructures and for mechanism design [28].

Coherence networks provide a way of representing competing hypotheses and the evidence they explain [29]. It has been used to anticipate, understand and respond to actions of an opponent in adversarial problem-solving situations (e.g., military decisions) [29]. In this approach, a hypothesis and its evidence is considered 'the more coherent, the less supporting evidence' the hypothesis requires. Evidence and

hypotheses have associated activation levels and these activation levels are propagated through the links among them. As evidence is observed, it propagates through the coherence network and updates the activation level of associated hypotheses and evidence. This technique can capture the dynamics of the evolution of hypotheses and evidence.

4.2 Scenario modeling

Scenario modeling was first developed as a management approach to support strategic decision-making [30]. It is a method to examine possible alternative futures given a projection of trends. The goal of scenario modeling is not to predict the future, but to prepare for an uncertain, unfolding future. Often only external scenarios are modeled. However, modeling internal scenarios (resources and capabilities) provides insights into an organization's capability to execute. In the context of cybersecurity, external refers to attacks and attackers and internal to defenders and their capabilities.

A scenario can be thought of as a sequence of observable indicators or signals. In the context of scenarios, we often focus on "weak signals" as signals that we need to pay close attention to because of their far-reaching impact [30]. Multiple scenarios can be combined into a tree in which internal nodes indicate indicators and branches represent possible alternatives [31]. This construct provides the basis for scenario generation and failure analysis.

Applications of scenario modeling to cybersecurity include: modeling attacks, generating attack scenarios, and assessing the impact of attacks. Different approaches have been proposed to generate potential attack scenarios: merging existing attacks [32] and applying attack patterns [33]. An AI planning approach for generating attack scenarios has been described in [35]. Recent work on intrusion prevention systems also suggests that we can use scenarios to assess the impact of ongoing complex attacks [34].

For instance, common attack patterns can be extracted from a public collection of attack types (CAPEC [36]) and codified in the form of patterns [33].

These patterns capture knowledge about attacks from the perspective of an attacker: each captures an attacker's goal and the steps to carry out the attack. The authors then show how this collection of attack patterns can be used to generate possible attack scenarios given an attacker's goals. This approach was able to replicate an expert vulnerability analysis.

4.3 Anticipation of failure

Mindfulness is the capability to discover and manage unexpected behavior [37]. It combines the concept of anticipation with that of resilience. Anticipation comprises preoccupation with failure, reluctance to simplify, and sensitivity to operation [37]. When we apply those processes to cybersecurity, they imply that we must pay close attention to signs of abnormal behavior in our networks and systems, question what we take for granted (i.e., expect attacks to evolve), and always search for a coherent explanation of our observations (i.e., maintain multiple competing hypotheses about the state of the world).

Anticipatory Failure Determination (AFD) [31] is an approach for envisioning failure scenarios. Its focus is not on learning from what failures have occurred in the past, but on discovering what failures may occur and how they can be brought about. AFD has recently been applied to model failure scenarios in cybersecurity [38]. The goal of the approach is to build an inventory of resources (indicators, tools, people, vulnerabilities, and information) that have enabled failures in the past. Failure scenarios start from failure indicators and work their way back through a causally linked chain of resources.

5 Discussion

Approaches to cybersecurity based on prediction (in the narrow sense, in which we have been using it in this paper) are limited in the extent to which they can cope with the evolution of cyber attacks. With the authors of [4], we made a distinction between prediction – basing actions on the past – and anticipation – basing current actions on future consequences of those actions.

Anticipation is a “future-oriented action, decision, or behavior based on a (implicit or explicit) prediction” [4]. Anticipatory systems include a forward model [4] that allows them to form hypotheses about the next response from the environment. Predictions from the forward model enable a system to compare predicted and observed responses and adjust its behavior [3]. This agrees with psychological experiments that show that current behavior is a function of both the context and the expected consequences of the behavior [3].

The forward model of biological anticipatory systems has its equivalent in the representation of possible futures in anticipatory techniques. In the case of adversarial thinking, possible futures are explored by the level-k reasoning of game theory or the competing hypotheses of coherence networks. In scenario modeling, the forward model consists of the generation and subsequent monitoring of attack scenarios. In anticipation of failure, mindful practices and reasoning backward from failures provide the anticipatory capability.

6 Conclusion

From this paper, we understand the need for a shift in the mindset on how we deal with cyber attacks, from a reactive to a more proactive approach founded in the emerging techniques of anticipatory thinking. This shift is required to manage an environment characterized by a significant increase in the volume and variety of cyber attacks. The paper also calls for more research on the proactive approach to cybersecurity.

References

1. Chen, HM, Kazman, R, Monarch, I, Wang, P: Can cybersecurity be proactive? A big data approach and challenges. *IEEE Hawaii International Conference on System Sciences*, 5978–5987 (2017)
2. Bubic A, Von Cramon DY, Schubotz RI: Prediction, cognition and the brain. *Frontiers in Human Neuroscience*, 4, 25:1-25:15 (2010)
3. Butz, MV, and Pezzulo, G: Benefits of anticipations in cognitive agents. *The Challenge of Anticipation*, 45–62, Springer (2008)

4. Pezzulo, G, Butz, MV, and Castelfranchi, C: The anticipatory approach: definitions and taxonomies. *The Challenge of Anticipation*, 23–43, Springer (2008)
5. Blei, DM, Ng, AY, and Jordan, MI: Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993–1022 (2003)
6. Blei, DM: Probabilistic topic models. *Communications of the ACM*, 55(4), 77–84 (2012)
7. Neuhaus, S, and Zimmermann, T: Security trend analysis with CVE topic models. *IEEE International Symposium on Software Reliability Engineering*, 111–120 (2010)
8. Hall, D, Jurafsky, D, and Manning, CD: Studying the history of ideas using topic models. *Conference on Empirical Methods in Natural Language Processing*, 363–371, Association for Computational Linguistics (2008)
9. Mathew, G, Agarwal, A, and Menzies, T: Trends in topics at SE conferences (1993-2013). *arXiv preprint arXiv:1608.08100*, 1–19 (2017)
10. Kelleher, JD, Mac Namee, B, and D’Arcy, A: *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies*. MIT Press (2015)
11. Sommer, R, and Paxson, V, *Outside the closed world: On using machine learning for network intrusion detection*. *IEEE Symposium on Security and Privacy (SP)*, 305–316 (2010)
12. Buczak, AL, and Guven, E: A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials*, 18(2), 1153–1176 (2016)
13. Barreno, M, Nelson, B, Joseph, AD, and Tygar, JD: The security of machine learning. *Machine Learning*, 81(2), 121–148 (2010)
14. Ballelli, T, Gad, M, and Shah, A: *Intrusion Learning: An Overview of an Emergent Discipline*. *Technology Innovation Management Review*, 6(2), 15–20 (2016)
15. Deka, RK, Kalita, KP, Bhattacharya, DK, and Kalita, JK: Network defense: Approaches, methods and techniques. *Journal of Network and Computer Applications*, 57(C), 71-84 (2015)
16. Ferguson, AJ, and Harris, NE: Moving big-data analysis from a forensic sport to a contact sport using machine learning and thought diversity. *Journal of Information Warfare*, 14(2), 53–70 (2015)
17. Shah, A, Abualhaol, I, Gad, M, and Weiss, M: Combining exploratory analysis and automated analysis for anomaly detection in real-time data streams. *Technology Innovation Management Review*, 7(4), 25–31 (2017)
18. Veeramachaneni, K, Arnaldo, I, Korrapati, V, Bassias, C, and Li, K: AI2: training a big data machine to defend. *IEEE International Conference on Big Data Security on Cloud (Big Data Security)*, *International Conference on High Performance and Smart Computing (HPSC)*, and *International Conference on Intelligent Data and Security (IDS)*, 49–54 (2016)
19. Benzmu’ller, R: *Malware trends 2017*. *G DATA Security Blog*, <https://www.gdatasoftware.com/blog/2017/04/29666-malware-trends-2017> (2017)
20. Rieck, K, Trinius, P, Willems, C, and Holz, T: Automatic analysis of malware behavior using machine learning. *Journal of Computer Security*, 19(4), 639–668 (2011)
21. Scandariato, R, Walden, J, Hovsepian, A, and Joosen, W: Predicting vulnerable software components via text mining. *IEEE Transactions on Software Engineering*, 40(10), 993–1006 (2014)
22. Francois, J, Wang, S, and Engel, T: *BotTrack: tracking botnets using NetFlow and PageRank*. *International Conference on Research in Networking*, 1–14, Springer (2011)
23. Esteves, J, Ramalho, E, and De Haro, G: To improve cybersecurity, think like a hacker. *MIT Sloan Management Review*, 58(3), 71–77 (2017)
24. Fette, I, Sadeh, N, and Tomic, A: Learning to detect phishing emails. *ACM International Conference on World Wide Web*, 649–656 (2007)
25. Bozorgi, M, Saul, LK, Savage, S, and Voelker, GM: Beyond heuristics: learning to classify vulnerabilities and predict exploits. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 105–114 (2010)
26. Feng, Q, Kazman, R, Cai, Y, Mo, R, and Xiao, L: Towards an architecture-centric approach to security analysis. *Working IEEE/IFIP Conference on Software Architecture (WICSA)*, 221-9230 (2015)
27. Hamman, ST, Hopkinson, KM, Markham, RL, Chaplik, AM, and Metzler, GE, *Teaching game theory to improve adversarial thinking in cybersecurity students*, *IEEE Transactions on Education*, 2017 (in press)
28. Alpcan, T: Game theory for security. *Encyclopedia of Systems and Control*, 495–499 (2015)
29. Thagard, P: Adversarial problem solving: Modeling an opponent using explanatory coherence. *Cognitive Science*, 16(1), 123–149 (1992)

30. Schoemaker, PJ: Scenario planning: a tool for strategic thinking. *Sloan Management Review*, **36**(2), 25–40
31. Kaplan, S, Haimes, YY, and Garrick, BJ: Fitting hierarchical holographic modeling into the theory of scenario structuring and a resulting refinement to the quantitative definition of risk. *Risk Analysis*, **21**(5), 807–807 (2001)
32. Ahn, W, Chung, M, Min, BG, and Seo, J: Development of cyber-attack scenarios for nuclear power plants using scenario graphs. *International Journal of Distributed Sensor Networks*, **11**(9), 836258 (2015)
33. Li, T, Paja, E, Mylopoulos, J, Horkoff, J, and Beckers, K: Security attack analysis using attack patterns. *IEEE International Conference on Research Challenges in Information Science (RCIS)*, 1–13 (2016)
34. Albanese, M, and Jajodia, S: A Graphical Model to Assess the Impact of Multi-Step Attacks. *The Journal of Defense Modeling and Simulation*, 1548512917706043, 1–15 (2017)
35. Pasquale, L, Hanvey, S, Mcgloin, M, and Nuseibeh, B: Adaptive evidence collection in the cloud using attack scenarios. *Computers & Security*, **59**, 236–254 (2016)
36. MITRE, Common Attack Pattern Enumeration and Classification. <https://capec.mitre.org/about/index.html>, last accessed in Nov. 2017
37. Weick, KE, Sutcliffe, KM, and Obstfeld, D: Organizing for high reliability: Processes of collective mindfulness. *Crisis management*, **3**(1), 81–123 (2008)
38. Badalkhani, P: Using publicly available information to predict cyber failures. Master's Thesis, TIM Program, Carleton University (2016).

Decision-making under uncertainty

Monika¹ · Hao Quan² · Dipti Srinivasan¹

Abstract Choosing actions on the basis of imperfect observations with unknown outcomes is called decision-making under uncertainty, and exists in many important problems. It unifies researchers across various disciplines to develop and model tools and methodologies to solve real-world decision-making problems under uncertainty. This paper proposes an efficient uncertainty modelling method using neural network-based prediction intervals (NN-based PIs) based on [1]. PIs are excellent tools to model uncertainties. Particle swarm optimisation (PSO)-based lower upper bound estimation (LUBE) method is used to construct NN-based PIs. Thereafter, a scenario generation method is used to generate scenarios from a list of PIs. These scenarios are further incorporated into the stochastic model for decision-making.

Keywords: Decision-making · Uncertainty · Neural networks · Particle swarm optimization · Lower Upper Bound Estimation · Prediction

1 Introduction

Every decision-making process holds a level of uncertainty. Uncertainty can never be eliminated, but can undeniably be condensed by using various methodologies and following some ideas like reducing time horizons for decisions, determining the worst case scenario, managing decisions adaptively, clarifying the uncertainty, increasing knowledge, and so on [2]. Uncertainty can be defined as the scenario of not having complete knowledge due to intrinsic deficiencies in acquired knowledge [3]. On the basis of its sources, we can classify uncertainty as ambiguity, approximations and likelihood. The ambiguity comes from the prospect of having multiple outcomes for processes or systems. The process of approximation can involve the use of imprecise semantics in language, approximate reasoning, and dealing with complexity, by emphasising relevance. Approximations can be viewed to include imprecision, abrasiveness and simplification. The likelihood can be defined in the context of chance, odds and gambling. Likelihood possesses the primary components of randomness and sampling [4]. Uncertainty can also be employed for characterizing the state of a system such as uncertainty of the result. The sources of uncertainty, while using in the field of modelling and computation of various aspects of the world

Monika:
sermo@nus.edu.sg

Hao Quan:
quanh@epgc.a-star.edu.sg

✉ Dipti Srinivasan:
dipti@nus.edu.sg

¹ Department of Electrical and Computer Engineering, National University of Singapore, Singapore

² Experimental Power Grid Centre, Agency for Science, Technology and Research (A*STAR), Singapore

may include parameter uncertainty, model inadequacy, parametric variability, observation error, numerical uncertainty, interpolation uncertainty, etc. [1].

Decision-making under uncertainty (DMUU) modelling faces numerous challenges like what is uncertainty, level and extent of information available, approaches for analysing uncertainty, how to incorporate the uncertainty modelling methods into practices for decision-making and henceforth. There are two types of problems in quantification of uncertainty, forward uncertainty propagation and inverse uncertainty quantification. The first one focuses on the response of system for outputs when inputs are uncertain. The later one attempts to estimate bias correction and parameter calibration. The forward uncertainty propagation methods include probabilistic methodologies [5][6][7][8][9][10][11][12][13] such as simulation-based methods (Monte-Carlo simulations, importance sampling, etc.), local expansion-based methods (such as Taylor series, perturbation method, etc.), functional expansion-based methods (like Neumann expansion, polynomial chaos expansion, etc.), most probable point models, numerical integration based models; and non-probabilistic approaches [14] like interval analysis, fuzzy theory, possibility theory and evidence theory.

The essential categorization of probability is the probability distribution function (PDF), which can be defined as the relative likelihood for an arbitrary variable to take on a given value. The PDF is non-negative ubiquitously and the integral of PDF over the entire space is equal to one. The density of a random variable, X is f , where f is a non-negative Lebesgue-integrable function, if [15]:

$$P [A \leq X \leq B] = \int_A^B f(x)dx$$

Let us say that F is the cumulative distribution function (CDF) of X , then F can be represented as:

$$F(x) = \int_{A-\infty}^x f(u)du$$

If f is continuous at x then,

$$f(x) = \frac{d}{dx} F(x)$$

Instinctively, one can think that $f(x)dx$ is probability of X falling within the infinitesimal interval $[x, x + dx]$. Figure 1 shows the uniform, Gaussian and binomial PDF and CDF.

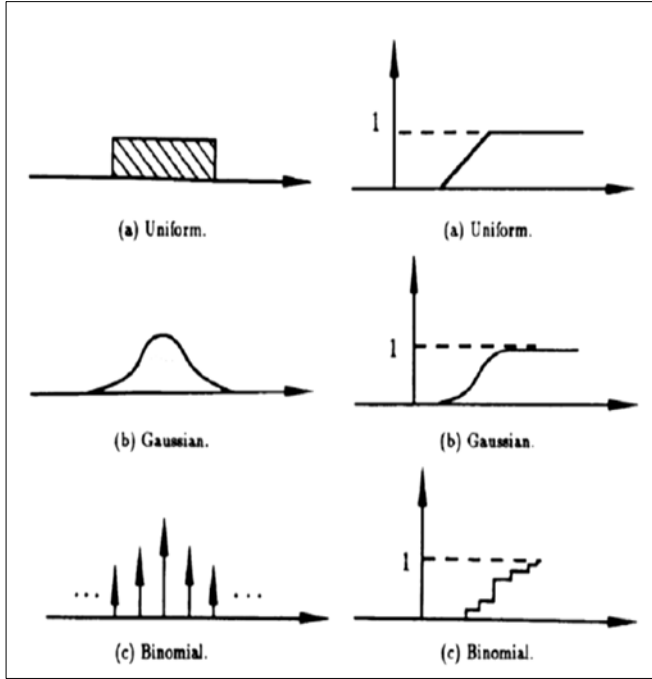


Fig. 1 The uniform, Gaussian and binomial PDF (left) and CDF (right)

Let us say that X is a random variable whose value is given and its distribution states a probability density function f , then the expected value of X (if it exists) can be calculated as,

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

In regression applications of neural networks, point forecasts have been mainly used. Nevertheless, owing to certain disadvantages of point forecasts like decrease in reliability of point forecasts with increase in level of certainty and lack in information about prediction accuracy, prediction intervals (PIs) are becoming increasingly popular.

The rest of the paper is organized as follows. PIs and their evaluation indices are introduced in Section 2. PSO-based LUBE method to construct PIs and stochastic modelling for decision-making are introduced in Sections 3 and 4 respectively. Finally, Section 5 draws the conclusion.

2 Prediction intervals

Prediction interval is a terminology used in statistics, which infers to an estimate of interval for future observations, based on the available observations. A PI consists of lower and upper bounds that brace a future unknown target value with a certain probability $((1-\alpha) \%)$ which is known as the confidence level. A typical PI comprises of three units: the upper bound, lower bound and the coverage probability. With the use of high quality PIs, it becomes convenient for the decision makers to confidently draw up future plans, better manage risks, and maximize their benefits. Like point forecasts are evaluated using the indices MSE (Mean Square Error) and MAPE (Mean Absolute Percentage Errors), on the similar grounds PIs are evaluated using the indices PICP (Prediction Interval Coverage Probability), PINAW (prediction interval normalized average width) and PINRW (prediction interval normalized root-mean-square width) whose mathematical representations are given below:

$$PICP = \frac{1}{n} \sum_{i=1}^n c_i$$

$$PINAW = \frac{1}{nR} \sum_{i=1}^n (U_i - L_i)$$

$$PINRW = \frac{1}{R} \sqrt{\frac{1}{n} \sum_{i=1}^n (U_i - L_i)^2}$$

$$CWC = PINAW + \phi(PICP)e^{-\eta(PICP - \mu)}$$

where n is the number of samples, U_i and L_i are the upper and lower bounds of PIs, c_i is an indicator of the PI coverage, R is the range of the underlying targets. PINAW and PINRW correspond to the 1-norm and 2-norm of the width of PIs. The arrangement of PINAW is analogous to MAPE when point forecasting is concerned, and it gives equal weights to all widths of PIs. The other index PINRW is functionally analogous to MSE, and amplifies wider intervals. For higher quality of prediction intervals, larger PICP and narrower PINRW form the essential properties. These properties of PICP and PINRW are contradictory, as demanding a larger PICSP will always result in wider PINRW, and tapering PINRW may lead to an unsatisfactorily low PICP. When considered from the optimization point of view, this is a two-objective problem. For the simplification and comprehensive comparisons for different PIs, it is necessary to transform this primary multi-objective problem into a single-objective one. An example of cost function for complete assessment of PIs is coverage width-based criterion (CWC). In CWC, a step-function, $\Phi(PICP)$, is defined whose value is majorly dependent on the fulfilment of PICP, with the parameters μ and η , controlling the position and amount of jump in CWC.

3 PSO-based LUBE method

In this paper, a method called LUBE (Lower Upper Bound Estimation) is proposed for the construction of prediction intervals (PIs). This method is preferred over the traditional methods for reasons such as:

- It is simpler and paradigms PIs with higher quality in a single step. On the contrary, traditional methods first perform point forecasting followed by construction of PIs.

- As a matter of fact, the real-world data with uncertainty such as wind power, solar energy, market response, etc. is unstable and intermittent in nature. Thus, assumptions about distribution of data seems problematic and doubtful. The LUBE method is a nonparametric method and hence no assumption on data distribution is made. Whereas, traditional methods constantly consider a parametric distribution (e.g., Normal) of data and then attempt to find its parameters for construction of intervals.

- LUBE method has significantly lower computational burden for PI construction when compared to alternative methods, which owes to the fact that the developed NN directly generates PIs. Alternatively, other approaches first need to linearize NN models or calculate complex matrices such as the Jacobian and Hessian matrices.

PIs with a high coverage probability and narrow width are expected for decision-making. However, these two aspects of PIs contradict with each other. For example, increasing the coverage probability will also widen the PIs while squeezing the PIs may lead to a lower coverage probability. This multi-objective problem can be transformed into a constrained single-objective problem by reassigning the coverage probability and then treating it as a hard constraint and the only objective is to minimize the width of PIs.

The basic concept of LUBE method is to adopt a NN with two outputs to directly generate the upper and lower bounds of PIs. The first and second outputs correspond to the upper and lower bounds of PIs separately. The symbolic NN with two outputs for the LUBE method is shown in Figure 2.

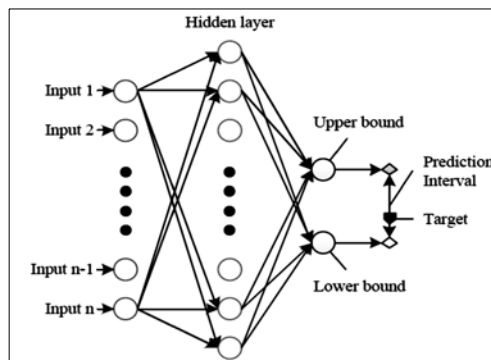


Fig. 2 NN model for LUBE method to generate upper and lower bounds of PIs

The plotting between the inputs and outputs of the system is shown below [16]:

$$y_i = f_i \left(\sum_{j=1}^{N_h} w_{ij} f_2 \left(\sum_{k=1}^{N_i} v_{jk} x_k + b_{vj} \right) + b_{wi} \right)$$

where y_i is the output of the i^{th} node on the output layer; x_k is the input of the k^{th} node in the input layer; w_{ij} represents the connection weight between nodes in the hidden and output layers; v_{jk} is the connection weight between nodes in the input and hidden layers; and b_{wi} and b_{vj} are bias terms that represent the threshold of the transfer function f_1 and f_2 . The number of nodes in the input, hidden and output layers are N_i , N_h and N_o respectively. In LUBE method, $N_o = 2$, and y_1, y_2 correspond to the upper and lower bounds of PIs. The LUBE method generates the upper and lower bounds directly, making PI construction as simple as point forecasts.

Neural Networks make the quality of PIs sensitive to their own structure. The question is how to determine layers and number of neurons in each layer. Too small or too large NNs have a low simplification power and often suffer under-fitting or over-fitting problems. The determination of the NN structure is key to the successful construction of high quality PIs. In this paper, fully connected feed-forward three-layered NNs are chosen and the number of neurons in the hidden layer is varied from 1 to 20. In order to determine an optimal structure of NNs, a k-fold cross validation method is employed. This method relies on a k-fold cross correlation. In order to maintain independence of the training and test, the k-fold cross validation is implemented on the training set. The entire training set is divided equally into k complementary folds, out of which, usually k-1 are used for training the candidate NNs, and the remaining fold is used for the purpose of validation [17]. For avoiding biased sampling, every structure of NN is trained and validated for k times using k different sub-training and validation datasets. A 5-fold cross validation is implemented here. Traditionally error-based measures such as MAPE and MSE are used for selecting the optimal structure of NNs. The focus of this paper is on construction of high quality PIs. Therefore, it is more reasonable to determine the optimal structure of NNs directly using PI-based evaluation indices, such as PICP, PINAW and CWC. Determination of the optimal NN structure needs to balance between the network complexity, generalization and learning capacity of NNs.

Meta-heuristic approaches are computational intelligence methods inspired by natural evolution or social behaviour. These methods include expert systems [18], fuzzy logic [19], GA [20][21], evolutionary programming, simulated annealing, tabu search, PSO [22], ant colony optimization and differential evolution [23]. PSO (Particle Swarm Optimization) is a parameter optimization technique which is employed to solve the problem of optimizing the PIs.

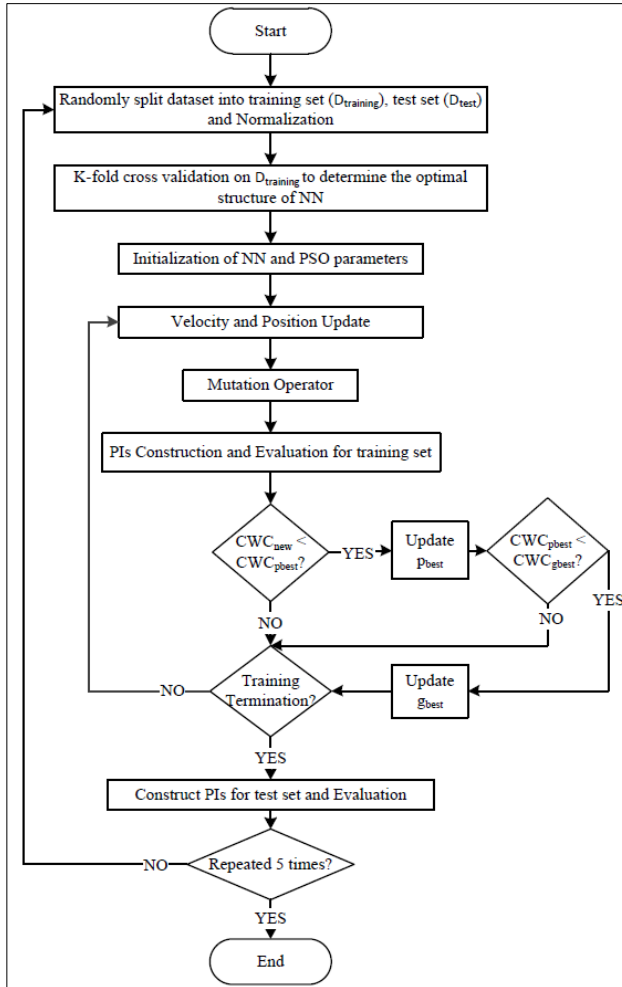


Fig. 3 PSO-based LUBE method for construction and evaluation of PIs

The objective of using PSO has mainly two aspects. Firstly, PSO is used to solve the single-objective problem which is to handle the constraints and optimize the objective. Secondly, PSO with mutation operator is used as the training algorithm through optimizing the connection weights [24] of NN models.

The flow chart of the proposed PSO-based LUBE method is shown in Figure 3.

The data is first split into data set and test set randomly, whereafter the training dataset is normalised, following the same procedure for test dataset. The training set ($D_{training}$) is further split into sub-training sets and validation sets using the 5-fold cross validation method. Median value of CWC is used to determine the optimal structure of NNs. Then the parameters of NNs and PSO are initialised. A bad initialisation may not converge to a very good result, hence initialisation of parameters is very important for this algorithm. After investigation [25], it was found that Nguyen-Widrow (NW) method obtains the best and most stable results. Hence NW method is used in this algorithm for NN weight initialisation. PSO parameter initialization consists of

particle position and velocity initialization. Since the weights of NN connection are represented as the position of particles, hence position initialization gets completed in weight initialization. Zero symmetric numbers are used to initialize the particle velocity which is a process of random initialization. Velocity and position update are the core of the PSO algorithm. The particles will exchange their “findings” with each other in the update process. In this way, the information gets exchanged efficiently throughout the whole swarm. The classic formulas for velocity and position update [26] are shown below:

$$v_n(t+1) = W \cdot v_n(t) + C_1 \text{rand}() (p_{best,n} - x_n(t)) + C_2 \text{rand}() (g_{best,n} - x_n(t))$$

$$x_n(t+1) = x_n(t) + v_n(t+1)$$

where v_n is the particle velocity in the n^{th} dimension, $\text{rand}()$ is a random number between 0 and 1, W is a scaling factor, and C_1 and C_2 are scaling factors that determine the relative “pull” of p_{best} and g_{best} .

Besides the two updates, the ranges for velocity and position are limited to V_{max} and X_{max} separately. Selection, crossover and mutation form the three main operators in GA. Mutation operator, which helps achieve diversity in GA, is integrated into PSO. This integration strongly enhances the searching capacity and avoids being trapped into local optima. In flow chart shown in Figure 3, Gaussian mutation is added to each connection weight after the position update. The mean and standard deviation of Gaussian distribution are the weight value and 10% of that weight value respectively. The mutation rate exponentially decreases as the optimization continues. In the validation step, the training and validation sets are combined together to train the NN. LUBE method is applied to construct new PIs after updating the NN connection weights. PI assessment indices (PICP and PINRW) are calculated. p_{best} is the personal best value of each particle and g_{best} denotes the best value of the whole swarm. These values are updated during the process of training. Once the training process terminates, the g_{best} value is chosen to generate PIs for the test set. PICP and PINAW instead of PINRW are calculated and recorded. For the comprehensive evaluation purpose, CWC is also calculated. The process is repeated a couple of times. This PSO-based LUBE method is convergent.

4 Stochastic modeling for decision-making

In case of point forecasts, there is only one forecast value which can be used directly for decision-making. But in case of PIs, even a single level PI consists of three components namely upper bound, lower bound and corresponding confidence level. As decision-making becomes difficult in the presence of two bounds, linkage between the PIs and decision-making becomes important. In this paper, a computational framework is proposed for building up this important linkage by introducing a scenario generation method and stochastic modelling.

In a decision-making process, for a given lead time, a single level PI does not suffice making an optimal decision. Theoretically, a sufficient number of PIs or quantiles are required to represent any type of probabilistic distributions. Using the PSO-based LUBE method the various levels of PIs are constructed, which are applied to estimate the empirical cumulative distribution function (ECDF) of unknown

probabilistic distributions of the parameters under consideration. Now the challenge is to apply these PIs to mathematical models for decision-making. As multilevel PIs are presented and PIs suffer from multi-valued problem for decision-making, a scenario generation technique is proposed to properly represent the uncertainties and these scenarios are mathematically involved into the stochastic model for computational purpose for decision-making. The steps involved in implementation of the proposed generation scenario include:

1. Apply the PSO-based LUBE method to make forecasts for a list of PIs with multiple confidence levels;
2. If the constructed PIs are assumed to be central PIs, then each PI can be uniquely decomposed into two quantiles. The $(1 - \alpha) \%$ PIs consists of two bounds, i.e. the lower and upper bounds. They correspond to the $(\alpha/2) \%$ and $(1 - \alpha/2) \%$ quantiles respectively;
3. Obtain discrete points on ECDF and generate the ECDF curve fitting;
4. Monte Carlo simulation is applied to generate scenarios from the fitted ECDF.

This method can be easily implemented, avoiding the complex computation like the covariance matrix. It also avoids strong assumptions on probabilistic distributions, the only assumption being that the constructed PIs are central PIs. This method builds an important bridge between the PIs and the scenarios used in stochastic model. These generated scenarios are further computationally involved for decision-making.

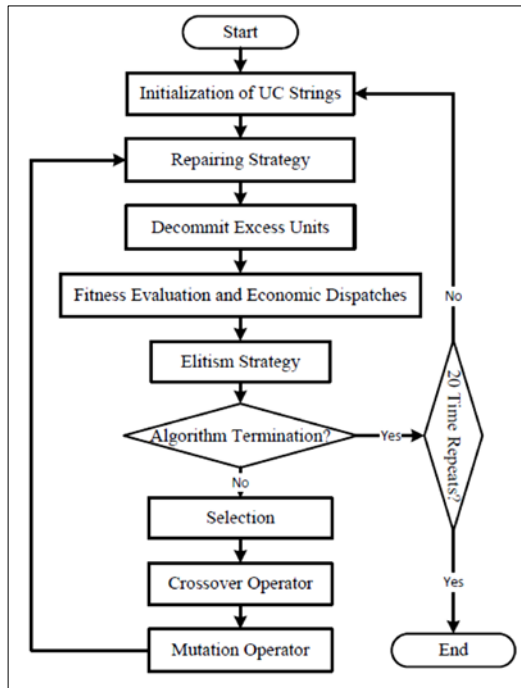


Fig. 4 Flowchart of the GA-based solution method

To solve the stochastic problem, a GA-based solution is proposed whose flowchart is presented in Figure 4, which shows an example of this GA-based solution method applied to unit commitment problems in electrical power systems. But this framework can be fitted to other optimization problems as well.

Initialization is crucial to the repeatability and success rate of the GA-based solution method for the stochastic problem. During initialization, all the constraints are not considered, hence the chromosomes need repairing periodically. The fitness function is defined in the following equation and is positively oriented, i.e., the greater, the better. A is a system-dependent constant used to avoid getting too small fitness values. $E(X, P)$ is the objective of the stochastic problem representing expected cost.

$$Fitness = \frac{A}{E(X,P)}$$

Elitism is a useful strategy frequently used in GA. The idea is to reserve the best one or more populations in previous generation and directly copy them to the next generation without any modifications. In this way, the best fitness values will not decrease with increasing number of iterations. In this paper, the two fittest solutions are reserved first, and the two worst solutions after the selection are replaced by the two elitism members. q -tournament selection is implemented in the GA. The most used value for q is 2. The larger value the q is, the higher the selection pressure becomes in the population. A hefty value of q means the entire population is dominated by the members with high fitness values. There is a need to maintain a balance between the selection pressure and the diversity of the population.

The crossover operator is a two-point crossover. Two-point crossover calls for two points to be selected on the parent chromosome strings, under a certain crossover probability. Everything between the two points is swapped amongst the parent binary strings, rendering two child chromosomes.

Adaptive mutation is chosen here, whose mutation rate exponentially decreases with the increase of the number of generations. For each mutation, a random number is generated. If the random number is smaller than the mutation rate, the value on this bit is flipped; otherwise, the value remains the same as before.

The termination criterion can be set as the reach of the maximum number of iterations or few improvements are made in a certain number of generations.

5 Conclusion

In this paper, PIs are proposed to quantify uncertainties from forecasting. We propose a new method called PSO-based LUBE method for creating multilevel PIs, which involves neural networks. These PIs are then used for generation of scenarios using the ECDF and Monte Carlo Simulation. These generated scenarios are thereafter employed into a stochastic model for decision-making.

References

1. H. Quan: Uncertainty modeling in distributed power systems. Ph. D diss. National University of Singapore (2014)
2. Making effective decisions in high uncertainty, <http://www.decision-making-solutions.com/decision-making-in-uncertainty.html>
3. B. M. Ayyub, G. J. Klir: Uncertainty modelling and analysis in engineering and the sciences. Chapman & Hall/CRC Press (2006)
4. N. Attoh-Okine, B. M. Ayyub: Applied research in uncertainty modeling and analysis, ser. International Series in Intelligent Technologies. Springer (2005)
5. S. H. Lee, W. Chen: A comparative study of uncertainty propagation methods for black-box type functions. ASME 2007 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference. American Society of Mechanical Engineers. 1275–1284 (2007)
6. S. Raychaudhuri: Introduction to Monte Carlo simulation. Simulation Conference, WSC 2008. Winter. IEEE. 91–100 (2008)
7. P. J. Smith, M. Shafi, H. Gao: Quick simulation: A review of importance sampling techniques in communications systems. IEEE Journal on Selected Areas in Communications. **15**(4), 597–613 (1997)
8. C. G. Bucher: Adaptive sampling an iterative fast Monte Carlo procedure. Structural Safety. **5**(2), 119–126 (1988)
9. B. Lallemand, G. Plessis, T. Tison, and P. Level: Neumann expansion for fuzzy finite element analysis. Engineering computations. **16**(5), 572–583 (1999)
10. K. Sepahvand, S. Marburg, and H.-J. Hardtke: Uncertainty quantification in stochastic systems using polynomial chaos expansion. International Journal of Applied Mechanics. **2**(2), 305–353 (2010)
11. X. Du and W. Chen: A most probable point-based method for efficient uncertainty analysis. Journal of Design and Manufacturing Automation. **4**(1), 47–66 (2001)
12. F. Xiong, K. Guo, and W. Zhou: Uncertainty propagation techniques in probabilistic design of multilevel systems. International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering (ICQR2MSE). IEEE, 874–878 (2011)
13. H. Xu and S. Rahman: A generalized dimension-reduction method for multidimensional integration in stochastic mechanics. International Journal for Numerical Methods in Engineering. **61**(12), 1992–2019 (2004)
14. C. Chatfield: Calculating interval forecasts. Journal of Business and Economic Statistics. **11**(2), 121–135 (1993)
15. H. Tijms: Understanding probability. Cambridge University Press (2012)
16. C.-H. Chen, J.-C. Wu, and J.-H. Chen: Prediction of flutter derivatives by artificial neural networks. Journal of Wind Engineering and Industrial Aerodynamics. **96**(10-11), 1925 – 1937(2008)
17. J. Wright and M. Manic: Neural network architecture selection analysis with application to cryptography location. International Joint Conference on Neural Networks (IJCNN), 1–6 (2010)
18. P.-H.Chen: Two-level hierarchical approach to unit commitment using expert system and elite PSO. IEEE Transactions on Power Systems. **27**(2), 780–789 (2012)
19. B. Venkatesh, P. Yu, H. Gooi, and D. Choling: Fuzzy MILP unit commitment incorporating wind generators,” IEEE Transactions on Power Systems, **23**(4), 1738–1746 (2008)
20. S. Kazarlis, A. Bakirtzis, and V. Petridis: A genetic algorithm solution to the unit commitment problem. IEEE Transactions on Power Systems. **11**(1), 83–92(1996)
21. K. Swarup and S. Yamashiro: Unit commitment solution methodology using genetic algorithm. IEEE Transactions on Power Systems. **17**(1), 87–91 (2002)
22. T. Logenthiran and D. Srinivasan: Particle swarm optimization for unit commitment problem. IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS). 642–647 (2010)
23. X. Yuan, A. Su, H. Nie, Y. Yuan, and L. Wang: Application of enhanced discrete differential evolution approach to unit commitment problem. Energy Conversion and Management. **50**(9), 2449–2456 (2009)
24. Q. Sun, Y. Yu, Y. Luo, and X. Liu: Application of BFNN in power flow calculation in smart distribution grid. Neurocomputing. **125**, 148–152 (2014)
25. A. Pavelka and A. Proch: Algorithms for initialization of neural network weights random numbers in matlab. Control Engineering. **2**, 453–459(2004)
26. G. Pulido and C. Coello: A constraint-handling mechanism for particle swarm optimization. IEEE Congress on Evolutionary Computation, CEC 2004, 2, 1396–1403 (2004).

Machine learning in computer security

V. Rao Vemuri¹

Abstract The past few years have witnessed a rise in the use of AI and Machine Learning techniques to a variety of application areas, such as image understanding and autonomous vehicle driving. Wireless and cloud technologies have also made it possible for millions of people to access and use services available via the internet. During the same period, the world has also witnessed a rise in cyber-crime, with criminals continually expanding their methods of attack. Weapons like ransomware, botnets, and attack vectors became popular forms of malware attacks. This paper examines the state-of-the-art in computer security and the use of machine learning techniques therein. True, machine learning did make an impact on some narrow application areas such as spam filtering and fraud detection. However – in spite of extensive academic research – it did not seem to make a visible impact on the problem of intrusion detection in real operational settings. A possible reason for this apparent failure is that computer security is inherently a difficult problem. Difficult because it is not just one problem; it is a group of problems characterized by a diversity of operational settings and a multitude of attack scenarios. This is one reason why machine learning has not yet found its niche in the cyber warfare armory. This paper first summarizes the state-of-the-art in computer security and then examines the process of applying machine learning to solve a sample problem.

Keywords: Machine Learning · Computer Security · Intrusion Detection

1 The status of cyber warfare

Cyber security, a subset of information security, is the practice of defending an organization's networks, computers and data from unauthorized digital access, attack or damage by implementing various processes, technologies and practices.

Network security, a subset of cyber security, aims to protect any data that is being sent through devices in a network to ensure that the information is not changed or intercepted. The role of network security is to provide protection from all types of cyber threats including viruses, worms, Trojan horses, zero-day attacks, hacker attacks, denial of service (DoS) attacks, and attacks by spyware, adware, ransomware, and so on.

✉ V. Rao Vemuri
rvemuri@ucdavis.edu

¹ Professor, Department of Computer Science, University of California, Davis, USA

With the advent of wireless technology, cloud computing and Internet of Things (IoT), the world is becoming one huge integrated network. This means that an increasing amount of personal and corporate information exists in the cloud where it potentially interchanges with a multitude of devices. A compromised network does not only mean access to private banking details, but also access to public infrastructures such as traffic lights, GPS tracking systems, water services, and power plants. Under these circumstances, faster and more reliable intrusion detection techniques become necessary both to diagnose and prevent attacks. To meet this challenge, some are suggesting the use of the power of parallel computing platforms like MapReduce and Hadoop, an open-source software framework for distributed storage and processing of big data [7, 11].

1.1 Hacking has gone pro

At the turn of the century, almost all threats to computer systems were malware programs (viruses, worms, and Trojans) written by pranksters. Although some malware did harm, most simply annoyed people. Professional and state-sponsored hackers were there, but they were not the norm.

Nowadays almost all malware is created to steal money or corporate secrets [14]. Professional hackers make millions of dollars, victimizing individuals and corporations with almost no fear of being prosecuted. Malware has morphed from innocuous, funny viruses and worms to identity-stealing programs and ransomware. Advanced persistent threats (APTs), such as mobile surveillanceware like JadeRAT, officially or unofficially working on behalf of a foreign government, are the new normal. According to the FBI, cyberwar will turn black in the coming years with sinister activities like taking control of a moving vehicle from a distance or remotely turning off a heart pacemaker [1].

1.2 Breach detection tools have improved

Once antivirus scanners were the main tools for breach detection. Now, products have been developed to detect when someone is doing something malicious, even if that someone is a “legitimate” user.

Event monitoring systems are improving. Many companies are now storing and analyzing billions of events a day, using huge disk storage arrays. Intrusion detection has moved beyond detecting simple malicious activity to detecting anomalous events that are out of character for a company and its employees. Connections to known, questionable networks are tracked and reported like the antivirus detections of yesteryear. Data leak protection (DLP)¹ has become big business.

¹ Roger A. Grimes, “Make stolen data worthless”, <http://www.infoworld.com/article/2969372/security/how-to-make-stolen-data-worthless.html>, August 11, 2015

1.3 Multifactor authentication and encryption are becoming the new default

Many still use passwords, but most sites now offer two-factor authentication. Mobile phones and popular operating systems come with biometric identification by default. Identity authentication startup Trusona is making a world without passwords a reality.

Default encryption is on the rise despite nearly all governments protesting it². Today, most popular operating systems, computers, and mobile devices come with built-in, default-enabled disk encryption. More and more websites are using SSL³ encryption by default although police and government agencies are trying to get rid of default encryption or enabling backdoors in the name of stopping criminals.

2. Computer security landscape

Security is a hard problem. Hard because it is a complex problem with many facets. There are many types of attacks, each targeting a different layer of the ISO/OSI model as shown in Table 1.

Table 1 Some Popular Attack Modes

Layer Name	Popular Protocols	Popular Attack Modes
Layer 7: Application	DNS, DHCP, HTTP, FTP, IMAP, SSH, NTP, SMTP, SNMP, Telnet, TFTP	DNS poisoning, Phishing, SQL injection, spam.
Layer 6: Presentation		
Layer 5: Session	SMB, NFS, Socks	
Layer 4: Transport	TCP (connection-based), UDP (connection-less)	TCP attack, Routing attack, SYN Flooding, Sniffing
Layer 3: Network	IP-v4, IP-v6, ICMP, IPSec	Ping, ICMP Flooding
Layer 2: Data Link	PPTP, Token Ring	ARP spoofing, Mac flooding
Layer 1: Physical		

Ping sweeps and port scans are used for reconnaissance. Sniffing captures packets as they travel through man-in-the-middle attacks intercepting messages intended for a third party. In spoofing, one sets up a fake device and tricks people to send messages

² Paul Venezia, "The Deep End", <http://www.infoworld.com/article/2946064/encryption/encryption-with-forced-backdoors-is-worse-than-useless-its-dangerous.html>, July 13, 2015

³ Roger A. Grimes, "10 security technologies destined for the dustbin", <http://www.infoworld.com/article/2970447/security/10-security-technologies-destined-for-the-dustbin.html>

to it. Hijacking means taking control of a session. Advanced persistent threats (APT) and multi-stage attacks are other challenging problems.

There are many components to a network security system that work together to improve the security posture. The most common network security components include firewalls⁴, anti-virus software, intrusion detection and prevention systems⁵, and virtual private networks⁶.

2.1 Attack scenarios

Attackers can take advantage of vulnerabilities in hardware, software, and protocols.

Active attacks are based on an alteration of the original message, or the creation of a false message. An active attack can be an interruption (e.g., masquerading) or modification (e.g., denial of service). In a masquerade, an unauthorized entity pretends to be authorized. (e.g. Due to a lack of authentication, Bob doesn't know if Tom is masquerading as Alice.) Modification results in loss of integrity. A modification attack, in turn, can be a replay or alteration attack. In a replay, if Alice wants to send \$100 to Tom's account, Tom captures that message and makes a second transfer of \$100. In an alteration, Tom captures Alice's message and alters it to read \$200. Denial of service attacks prevent legitimate users from using some services.

In a passive attack, the attacker does not intend to modify but indulges in monitoring the transmission to find out what is happening. There are two types of passive attacks. In Release of Message, the goal is to capture confidential data and put it publicly on the network. In Traffic Analysis, the attacker tries to find similarities between encrypted messages and deduces the original content.

The Speculative Execution attack or Spectre and Meltdown are two newly discovered attack scenarios that exploit critical vulnerabilities in modern processors. These hardware vulnerabilities allow programs to steal data that are currently being processed. While programs are typically not permitted to read data from other programs, these exploits can get hold of secrets stored in the memory of other running programs. This might include passwords stored in a password manager or browser, personal data such as photos, emails, instant messages and even business-critical documents. Unlike usual malware, Meltdown and Spectre are hard to distinguish from regular benign applications. These exploits do not leave any trace in a log file. Meltdown and Spectre work on personal computers, mobile devices, and in the cloud. Depending on the cloud provider's infrastructure, it might be possible to steal data from other customers. Intel has responded to this disclosure in terms of both software patches and firmware updates.

⁴ "Managed Firewall Simplify and Streamline the Management and Monitoring of Your Firewall Device", <https://www.secureworks.com/capabilities/managed-security/network-security/managed-firewall>

⁵ "Managed IDS/IPS Two Devices You Shouldn't Be Without",

<https://www.secureworks.com/capabilities/managed-security/network-security/managed-ids-ips>

⁶ Chey Cobb, "Ensuring Network Security with a VPN (Virtual Private Network)",

<http://www.dummies.com/programming/networking/ensuring-network-security-with-a-vpn-virtual-private-network/>

Meltdown breaks the most fundamental isolation between user applications and the operating system. If a computer has a vulnerable processor and runs an unpatched operating system, it is not safe to work with sensitive information without the chance of leaking the information. Luckily, there are software patches against Meltdown. Spectre, on the other hand, breaks the isolation between different applications. It allows an attacker to trick error-free programs, which follow best practices, into leaking their secrets. In fact, the safety checks of said best practices actually increase the attack surface and may make applications more susceptible to Spectre.

Speculative Execution is a legitimate procedure that may inadvertently create an opportunity for an attack. In order to improve performance, many CPUs may choose to speculatively execute instructions based on assumptions that are considered likely to be true. During speculative execution, the processor would be verifying these assumptions. If they are valid, then the execution continues. If they are not valid, then the execution is unwound, and the correct execution path can be started based on the actual conditions. It is possible for this speculative execution to have side effects which, if not restored when the CPU state is unwound, can lead to information disclosure.

The “bounds check bypass attack” (Variant 1) allows malicious code to circumvent bounds checking features built into most binaries. Even though the bounds check fails, the CPU will speculatively execute instructions after the bounds checks, and can access memory that the code could not normally access. When the CPU determines the bounds check has failed, it discards any work that was done speculatively; however, some changes to the system can be still observed (in particular, changes to the state of the CPU caches). Malicious code can detect these changes and read the data that was speculatively accessed.^{16, 17, 18}

“The branch target injection attack” (Variant 2) uses the ability of one process to influence the speculative execution behavior of code in another security context running on the same physical CPU core. Modern processors predict the destination for indirect jumps and calls that a program may take and start speculatively executing code at the predicted location. The tables used to drive prediction are shared between processes running on a physical CPU core, and it is possible for one process to influence (pollute) the branch prediction tables of another process or kernel code. In this way, an attacker can cause speculative execution of any mapped code in another process, in the hypervisor, or in the kernel, and potentially read data from the other protection domain using techniques like Variant 1.

This vulnerability can be fixed either by a CPU microcode update from the CPU vendor, or by applying a software mitigation technique. This mitigation may be applied to the operating system kernel, system programs and libraries, and individual software programs, as needed.^{16, 19, 20}

¹⁶ “Hacker News”, <http://hn.premii.com/#/article/16073874>

¹⁷ “CVE-2017-5753”, <http://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2017-5753>

¹⁸ “CVE-2017-5715”, <http://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2017-5715>

¹⁹ “CVE-2017-5754”, <https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2017-5754>

²⁰ “Meltdown and Spectre”, <https://meltdownattack.com/>

2.2 Intrusion detection

An intrusion detection system (IDS) is the modern-day equivalent to the burglar alarm; it constantly monitors the network to look for suspicious activity and can be configured to notify security administrators of any suspected intrusion. An Intrusion Protection System (IPS) goes a step further by sending an alert and preventing the attempted intrusion, say by dropping traffic if a malicious act is detected. Based on the location in a network, an IDS can be host-based, network-based or application-based [2, 8, 13, 18].

Host-based IDSs, typically software, are installed on host computers and are used to analyze traffic received by the host. For example, an antivirus device may detect unwelcome traffic and log it for further analysis. Host-based systems might also monitor the OS, system calls, audit logs, and error messages on the host system. While network probes can detect an attack, only host-based systems can determine whether the attack was successful. Additionally, host-based systems can record what the attacker performed on the compromised host.

Network-based IDSs (NIDS) use strategically positioned probes to monitor and analyze all traffic on the target network. While host-based detection cannot detect a ping sweep or a port scan across multiple hosts, network-based IDSs can easily detect such reconnaissance attacks. Network-based sensors generate an alert when these reconnaissance attacks are discovered. As network speeds increase, so must the capabilities of the intrusion detection probes. As the network grows, more probes can be added to ensure proper coverage and security. None of these systems were effective in pinpointing attacks quickly; they were generally used as forensic tools to examine security incidents *ex post facto*.

Depending on how they function, NIDS can be divided into two types: (a) Behavior-based (or, anomaly-based or statistical) IDS, and (b) Signature-based (or, pattern-based) IDS, also known as misuse detection.

In behavior-based systems, the IDSs try to detect intrusions via deviations from normal or expected behavior. Here, IDSs make a profile of every user during normal operation. When a deviation of this normal behavior is detected, the IDS triggers its alarm. This type of IDS can detect the type of intrusion that has no record of its previous occurrence. In that sense, behavior-based systems can detect new type of attack patterns. A large number of false alarms are a problem with this system.

In signature-based systems, the IDSs maintain a database of known exploits and their attack patterns (also called signatures). While analyzing network packets, if the IDS finds any pattern match to one of those known attack patterns, then it triggers an alarm. This type of IDS needs to analyze every packet in the network as it looks for known attack patterns. This type of IDS produces less number of false positives. Since there are many network-based exploits coming on each month, these need to be updated frequently.

While both techniques can be effective in real-time, both suffer from significant limitations: behavior-based techniques break down under heavy traffic or sudden traffic bursts and signature-based techniques cannot guard against unknown

intrusions. State-of-the art anomaly detection methods can identify deviations from normal behavior, but they do not offer guidance on what to do next.

Application-based IDS appears to be the latest novelty in the intrusion detection field. These systems fall into three sub-categories:

- Application self-protection (with OWASP⁷ being the most well-known Web application security project)
- Web application firewall (WAF⁸)
- Dynamic application security testing (DAST⁹)

When it comes to web applications, the majority of companies do not have resources to fix vulnerabilities. A good option is to apply an automated process combining WAF and DAST. Being a passive security system, firewall itself is not capable of recognizing bad content in the traffic. Redirecting the flow to DAST and receiving an “alert” about malicious traces, firewall memorizes the rules and collects them. In its turn, DAST produces reports on application security vulnerabilities.

There are a number of commercial intrusion detection systems in the market: Juniper, McAfee, Cisco, Symantec, etc. These IDSs generally do not provide an ideal performance as advertised. Some of the effective host-based open source IDSs are: OSSEC¹⁰, and Tripwire¹¹. Some of the popular open source NIDS are: Snort¹², Real SecureNet, Suricata¹³, and Bro.

Snort generates thousands of alerts in a small time. A sample is shown in the Table 2 below. A table like this is usually the starting point for the application of machine learning. Each row of such a table describes an observation or alert. The fields (columns) in each observation may include source IP, destination IP, start and end times, protocol type, number of failed login attempts, number of file operations, number of failed connections from the same host, and so on. These fields are also called features (or attributes). A typical data set may contain millions of alerts (rows) and well over fifty features. An alert is nothing but a capture of an event (or events) that occurred at a given time; it does not mean an intrusion has occurred. Our goal is to look for a meaning in these patterns. Certain collections of alerts may indicate some nefarious activity. Collecting alerts from an IDS is just one way of taking the vital signs of a computer system. There are other ways, such as looking at the system calls generated by the operating system [15].

⁷ “Welcome to OWASP”, https://www.owasp.org/index.php/Main_Page

⁸ Chandan Kumar, “5 Open Source Web Application Firewall for Better Security”, <https://geekflare.com/open-source-web-application-firewall/>, September 4, 2016

⁹ Ian Muscat, “DAST vs SAST: A Case for Dynamic Application Security Testing”, <https://www.acunetix.com/blog/articles/dast-dynamic-application-security-testing/>, September 6, 2017

¹⁰ “OSSEC: Open Source HIDS SECurity”, <http://ossec.github.io/index.html>

¹¹ “Tripwire”, <https://github.com/Tripwire/tripwire-open-source>

¹² “Snort”, <https://www.snort.org/>

¹³ “Suricata”, <https://suricata-ids.org/>

Table 2 A Sample Data Set

T	Time	src_port	Dest port	src_ip	dest_ip	Severity
Adobe Products Violation	Sep 01 '14 00:17:50	Medium	Large4	46.178.180.36	33.85.222.155	High
Aggressive Aging	Sep 01 '14 00:23:00	Medium	Large4	15.226.10.38	25.46.150.139	Low
Apache Server Protection Violation	Sep 01 '14 04:17:11	Medium	Large4	53.197.36.240	20.213.101.107	High
Application Servers Protection Violation	Sep 01 '14 04:23:00	Medium	Large4	21.69.151.82	3.138.44.87	High
BACKDOOR: SSH Server Running on Non-Standard Port	Sep 01 '14 08:34:55	Medium	Large5	2.93.222.122	27.131.204.222	Low
BitTorrent: DHT Tracker Communications (UDP)	Sep 01 '14 08:41:36	Medium	Large5	51.105.11.21	15.40.86.223	High
Content Protection Violation	Sep 01 '14 08:47:00	Medium	Large1	19.39.134.253	19.177.132.149	High

3. Modern face of machine learning

Banks and others in financial industry use Machine Learning (ML) to gain insights from data, and to prevent fraud. The insights can also identify investment opportunities, or help investors know when to trade. ML can also identify clients with high-risk profiles.

Machine Learning posits that computers can learn from data without being explicitly programmed to perform specific tasks. In fact, data is key when it comes to building systems using ML. The adage that more data means better models is true when it comes to ML. The input to a ML model could be structured or unstructured data, network data, click-stream data or behavior data and the output is a score or a class label.

Data sets are growing larger. As the volume, velocity, and variability of data streams increase, so does the challenges. Even within a single network, the most basic characteristics – such as bandwidth, duration of connections, and application mix – can exhibit immense variability, rendering them unpredictable. While tools often

work well with thousands of records and a few megabytes of data, they do not scale up well while handling real-world problems, measured in gigabytes or even terabytes of data.

As models are exposed to new data, they incrementally adapt. They learn from previous model stages to produce reliable, repeatable decisions and results. It is a science that is not new – but one that has gained fresh momentum [5, 21, 23, 25].

3.1 Data collection

Table 2 is an example of an unlabeled data set. Traditionally, a security expert examines each row of this table and assigns a label YES for a breach and NO, otherwise. Such labelled data, created laboriously, constitutes the training data for a supervised machine learning algorithm. Then the algorithm learns to classify the rows into two categories: intrusive and benign. Thus, a well-crafted anomaly detection algorithm may be able to generalize (from what it has learned from the training data) and detect an intrusion when presented with data it has not seen earlier [2, 24, 25].

Broadly speaking, there are three stages in the application of ML to solve any problem: preprocessing, training and validation of a model, and post-processing. The training and validation stage has been automated with excellent algorithms and opensource codes (e.g. Scikit-learn¹⁴, Tensor Flow¹⁵) to implement these algorithms. The bulk of the time and effort is normally spent during the pre- and post-processing (evaluation) stages where the purpose is to gain insight as such human intervention is still required.

A significant challenge in devising an evaluation plan is the lack of publicly available datasets, such as those available in, say, image processing. The two publicly available datasets that – the DARPA/Lincoln Labs packet traces [16, 17], and the KDD Cup dataset derived from them [12] – are almost two decades old, and no longer adequate for any current study.

3.2 Pre-processing

Raw data is often unusable for ML purposes. Data wrangling is the process of cleaning and unifying messy and complex data sets for easy access and analysis. This process typically includes manually converting/mapping data from one raw form into another format to allow for more convenient consumption and organization of the data. There may be missing data, inconsistent data use (e.g., a cardinality feature may contain “North”, “north”, and “N”, all identical in meaning), and numeric data with non-numeric characters, among many other possible problems. This step also involves combining multiple data sources to a single usable source and normalizing data so that all feature values are within a standard range, say in [0,1].

¹⁴ “scikit-learn”, <http://scikit-learn.org/stable/>

¹⁵ “About TensorFlow”, <https://www.tensorflow.org/>

For illustration, assume that we have a labelled data set, comprised of 41 columns and millions of rows. To detect an intrusion, we might use a classification algorithm and make a prediction of the class label. A first step of the procedure would be to download the data and examine it. To read the data and the labels we can use Python's `pandas.read_csv` function and process the resulting data frame using the following `process_data` function.

```
def process_data(X, y):
    X = X.drop(41, 1)
    X[1], uniques = pandas.factorize(X[1])
    X[2], uniques = pandas.factorize(X[2])

    num_examples = 10**6
    X = X[0:num_examples]
    y = y[0:num_examples]

    X = numpy.array(X)
    y = numpy.array(y).ravel()

    return X, y
```

All this function does is drop the label field (column 41) from `X` and turn categorical features in columns 1 and 2 into integers, then picks the first million rows and returns the resulting numpy arrays. This idea can be used if one is trying to map source IP's to some numerical values, thus:

```
df['ID'] = pd.factorize(df.SrcIP)[0]
```

Result

SrcIP	ID
192.168.1.112	0
192.168.4.118	1
192.168.1.112	0
192.168.4.118	1
192.168.5.122	2

The next step is feature engineering [3, 19]. Feature engineering is the process of using domain knowledge to create features that make machine learning algorithms work. Feature engineering is fundamental to the application of machine learning, and is both difficult and expensive. The table above shows only six features: time, src port, dest port, src ip, dest ip and severity. (What the table did not show is whether or not the alert is a result of a breach of security.) A security expert may feel that the features

selected may not suit her objective. To develop a user profile, subjective features like host ID, time of log-in, commands used, data entry modalities (speed of typing, keyboard Vs mouse usage) may be more relevant. To profile a program, objective features like system calls generated, resources used (CPU time, Memory, Buffers, etc.) may be more appropriate. In the context of intrusion detection, anomalous actions often happen in bursts rather than as isolated events. Due to this reason, time-based features like (a) the number of flows to unique source/destination IP addresses inside the network in the last T seconds to/from the same destination/source and (b) the number of flows from the source IP to the same destination port in the last T seconds, [4] occur.

Another rule of machine learning is to use a training set with instances drawn from all classes in equal proportions. Also, these algorithms perform better when trained with large data sets from each class. One finds only a few anomalies in a large data set. A machine trained with large chunks of normal data (positive examples) performs better in recognizing normal data and it would be foolhardy to expect it to perform well on abnormal data (negative examples).

Dealing with unbalanced datasets entails strategies such as improving classification algorithms or balancing classes in the training data during preprocessing. The later technique is preferred as it has wider application. The main objective of balancing classes is to either increase the frequency of the minority class or decrease the frequency of the majority class using a variety of techniques like under-sampling the majority class, over-sampling the minority class, synthetic minority over-sampling (SMOTE) – where a subset of data is taken from the minority class as an exemplar and new instances are created [9]. This author had used an Earth model to synthetically create a data set of minority class (underground nuclear explosions) while discriminating explosions from earthquakes [16].

3.3 Model building and validation

To make a prediction, one can use any of the classifiers available in the library. For example, logistic regression is used in the next section.

3.4 Evaluation of classification models

Evaluation of the classification model is an important post-processing step. In this connection, there are several important issues that need to be addressed:

- (a) What is the purpose of model evaluation? We use a model evaluation procedure to estimate how well a model will generalize to out-of-sample data. We also need a model evaluation metric to quantify model performance.
- (b) What are some of the evaluation procedures?
 - i. A simple way to evaluate a model is to split the data set into a training set and a test set. This split is best done with packaged s/w tools built

into Scikit or TensorFlow. While fast and simple, this step also gives a better estimate of out-of-sample performance, although it is still a high-variance estimate.

```
# split X and y into training and testing sets
from sklearn.cross_validation import train_test_split
X_train, X_test, y_train, y_test=train_test_split(X,y,
random_state=0)
```

- ii. K-fold Cross Validation: This systematically creates “K” training splits and averages the results. This runs K-times slower, but gives even a better estimate of out-of-sample performance, although it is still a high-variance estimate.
- iii. The model itself can be evaluated depending on the type of problem the model is solving. In regression problems, the usual metrics are mean absolute error, mean square error and the root mean square (RMS) error. If it is a classification problem, accuracy is the traditional metric. The question to be answered is: Can we predict an intrusion status given some measurements on system health?

```
# train a Logistic regression model on the training set
from sklearn.linear_model import LogisticRegression
logreg = LogisticRegression()
```

```
logreg.fit(X_train, y_train)
```

```
# make class predictions for the testing set
```

```
y_pred_class = logreg.predict(X_test)
```

- (c) Now calculate the classification accuracy which is defined as the percentage of correct predictions

```
# calculate accuracy
from sklearn import metrics
print(metrics.accuracy_score(y_test, y_pred_class))
0.692708333333
```

- (d) Null accuracy is the accuracy that could be achieved by predicting the most frequent class (here, normal with no intrusion). For brevity, this calculation is not shown here.
- (e) For the binary classification (normal Vs intrusive) model being built here from 192 (rows) alerts, only the first 28 results are shown below:

```
# print the first 25 true and predicted responses
from __future__ import print_function
print('True:', y_test.values[0:25])
print('Pred:', y_pred_class[0:25])
```

```
True: [1 0 0 1 0 0 1 1 0 0 1 1 0 0 0 0 1 0 0 0 1 1 0 0 0]
Pred: [0 0 0 0 0 0 0 1 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0]
```

(f) The Confusion Matrix gives a more complete picture of how the classifier is performing. Also, it allows one to compute various classification metrics, and these metrics can provide valuable guidance in model selection. Here, this is a 2x2 matrix because there are 2 response classes. Every observation in the testing set is represented in exactly one box. The basic terminology used is as follows:

- **True Positives (TP):** The model *correctly* predicted that there IS an intrusion
- **True Negatives (TN):** The model *correctly* predicted that there is NO intrusion
- **False Positives (FP):** The model *incorrectly* predicted that there IS intrusion (a “Type I error”)
- **False Negatives (FN):** The model *incorrectly* predicted that there is NO intrusion (a “Type II error”).

#IMPORTANT: first argument is true values, second argument is predicted values

```
print(metrics.confusion_matrix(y_test, y_pred_class))
```

```
[[118  12]
 [ 47  15]]
```

```
# save confusion matrix and slice into four pieces
confusion = metrics.confusion_matrix(y_test, y_pred_class)
TP = confusion[1, 1]
TN = confusion[0, 0]
FP = confusion[0, 1]
FN = confusion[1, 0]
```

n=192	Predicted: 0	Predicted: 1	
	Actual: 0	TN = 118	FP = 12
	Actual: 1	FN = 47	TP = 15
		165	27

The confusion matrix can be used to calculate several performance metrics of the classifier, as discussed below:

Classification accuracy, defined as $(TP + TN)/(TP + TN + FP + FN)$, is the easiest metric to understand.

```
print((TP + TN) / float(TP + TN + FP + FN))
print(metrics.accuracy_score(y_test, y_pred_class))
0.692708333333
```

Other metrics can be calculated likewise by replacing the first line of the above code snippet by the appropriate formula, as shown below:

Sensitivity = $(TP)/(TP + FN) = 0.241935483871$ tells how sensitive the classifier is in detecting positive (intrusive) instances. That is, how often the prediction is correct when the actual value is positive?

Specificity = $(TN)/(TN + FP) = 0.907692307692$ tells how often the prediction is correct when the actual value is negative?

False Positive Rate = $(FP)/(TN + FP) = 0.0923076923077$ tells how often is the prediction incorrect when the actual value is negative.

Precision = $(TP)/(TP + FP) = 0.555555555556$ tells how often is the prediction correct when a positive value is predicted. That is, how “precise” is the classifier while predicting positive instances?

3.5 Which metrics to use?

The choice of metric depends on the objective in modeling. For spam filtering (positive class is “spam”), one optimizes for precision or specificity because false negatives (spam goes to the inbox) are more acceptable than false positives (non-spam is caught by the spam filter). For a fraud detector (positive class is “fraud”), one optimizes for sensitivity because false positives (normal transactions that are flagged as possible fraud) are more acceptable than false negatives (fraudulent transactions that are not detected). In intrusion detection applications, a false positive, requires spending expensive analyst time examining the reported incident only to eventually realize that it reflects benign activity and can quickly render the NIDS ineffective.

False negatives, on the other hand, have the potential to cause serious damage to an organization: even a single compromised system can seriously undermine the integrity of the IT infrastructure. Systems that aim at minimizing FNs also tend to increase FPs. To guarantee that no terrorist passes through an airport security, you may have to risk frisking a grandmother!

3.6 Adjusting classification threshold

A threshold of 0.5 is normally used in binary classification problems to relate class predictions to probabilities. It is straightforward to print, say the first 10, predicted responses using the command `logreg.predict(X_test)[0:10]` and the corresponding predicted probabilities by the command `logreg.predict_proba(X_test)[0:10,:]` and plot a histogram of the predicted probabilities.

3.7 ROC and AUC

Sensitivity and specificity have inverse relationship. Although both are affected by the threshold, it is possible to study the effect of the threshold by plotting the ROC (Receiver Operating Characteristic) curve, which is a plot of the true-positive rate on the y-axis (i.e., sensitivity) and false-positive rate (i.e., $1 - \text{specificity}$) on the x-axis, for all possible classification thresholds. For an ideal classifier (high sensitivity and high specificity), the ROC curve hugs the upper left corner of the graph [Fawsett, '06]. This curve can be obtained by running the ROC Curve function from the Scikit-Learn's Metrics module with the true values of the testing set stored in `y_test`, as the first argument and the predicted probabilities stored in `y_pred_prob`, (NOT `y_pred_class`) as the second argument.

The AUC is literally the area under the ROC curve and represents the percentage of the total area under the ROC. A higher AUC value is indicative of a better overall classifier as such AUC is often used as a single-number indicator of the performance of the classifier as an alternative to classification accuracy.

4. Summary

The first part of this paper examined the state-of-the-art of computer security in the light of evolving advances in mobile and cloud computing with special attention to intrusion detection. The second part identified spam filtering and fraud detection where ML scored significant successes. Intrusion detection did not score as well in terms of its large-scale adaptation in commercial products probably because it is not a single problem but a syndrome. It seems that ML is good at identifying known patterns of attack (is this pattern there?) and not so good at identifying evolving attack patterns, especially if the attacker can inspect the models being used. As Richard Hamming famously said, "The purpose of computation is insight, not numbers," ML can be profitably used to gain insight into the operation and interpretation of intrusion detection systems because it can do a lot more, a lot faster.

References

1. F. Abagnale: Talks at Google, <https://www.youtube.com/watch?v=vsMydMDi3rI>, November 27, 2017
2. E. Aroms: NIST Special Publication 800-94 Guide to Intrusion Detection and Prevention Systems (Idps). CreateSpace, Paramount, CA (2012)
3. Arvin L. Bluma and Pat Langley: Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 97, 245-271, 1997
4. Varun Chandola, Eric Elertson, Levent Ert'oz, Gy'orgy Simon and Vipin Kumar: Data Mining for Cyber Security. *Data Warehousing and Data Mining Techniques for Computer Security*, <http://minds.cs.umn.edu/publications/chapter.pdf>, Springer (2006)
5. Vu Dao and V. Rao Vemuri. *Computer Network Intrusion Detection: A Comparison of Neural Networks Methods*. *Differential Equations and Dynamical Systems (Special Issue on Neural Networks)*, 2002
6. Tom Fawcett: An Introduction to ROC Analysis. *Pattern Recognition Letters*, 27, 861-874 (2006)
7. R. Fontugne, J. Mazel, K. Fukuda: Hashdoop: a mapreduce framework for network anomaly detection. *IEEE conference on computer communications workshops (INFOCOM WKSHP)*. 494-499 (2014)
8. P. Garcia-Teodoro, J. Diaz-Verdejo, G. Macia-Fernandez, and E. Vazquez: Anomaly-based network intrusion detection: techniques, systems and challenges. *Computers & Security*, 28 (1-2), 18-28 (2009)
9. Fithria Siti Hanifah, Hari Wijayanto, Anang Kurnia: SMOTE Bagging Algorithm for Imbalanced Data Set in Logistic Regression Analysis. *Applied Mathematical Sciences*, 9, 2015
10. G.S. Jang, F. U. Dowla, and V. Vemuri: Application of neural networks for seismic phase identification. *Proc. IJCNN 91*, Singapore, 899-904 (1991)
11. S. Kamaruddin and V. Ravi: Credit card fraud detection using big data analytics: use of PSOANN based one-class classification. *International Conference on Informatics and Analytics*, Pondicherry, August 2016
12. KDD cup data. <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html>
13. Nathan Keegan, Soo Yeon Ji, Aastha Chaudhary, Claude Concolato, Byunggu Yu and Dong Hyun Jeong: A survey of cloud-based network intrusion detection analysis. *Human-centric Computing and Information Sciences*, 6:19 (2016)
14. Raghu Krishnapuram and Anirban Mondal: Upcoming research challenges in the financial services industry: a technology perspective. http://www.idrbt.ac.in/assets/publications/Journals/Volume_01/Chapter_04.pdf, *IJBT*, 1, 66-84 (2017)
15. Yihua Liao and V. Rao Vemuri: Using text categorization techniques for intrusion detection. *Proc. Usenix*, San Francisco, August 2002
16. R. Lippmann, R. K. Cunningham, D. J. Fried, I. Graf, K. R. Kendall, S. E. Webster, and M. A. Zissman: Results of the 1998 DARPA offline intrusion detection evaluation. *Proc. Recent Advances in Intrusion Detection*, 1999
17. R. Lippmann, J. W. Haines, D. J. Fried, J. Korba, and K. Das: The 1999 DARPA off-line intrusion detection evaluation. *Computer Networks*, 34(4), 579-595, October 2000
18. Chilukuri K. Mohan and Kishan G. Mehrotra: Anomaly detection in banking operations. http://www.idrbt.ac.in/assets/publications/Journals/Volume_01/Chapter_02.pdf, *IJBT*, 1, 16-48, 2017
19. Nguyen H. T., Petrović S., Franke K.: A comparison of feature-selection methods for intrusion detection. In: Kotenko I., Skormin V. (eds.) *Computer Network Security, MMM-ACNS 2010*, *Lecture Notes in Computer Science*, 6258, Springer, Berlin, Heidelberg (2010)
20. Sanjay Rawat, Arun K. Pujari, V. P. Gulati, V. Rao Vemuri: Intrusion detection using text processing techniques with a binary-weighted cosine metric. *Journal of Information Assurance and Security*, 2005

21. S.A.R. Shah, B. Issac: Performance comparison of intrusion detection systems and application of machine learning to snort system. *Future Generation Computer Systems*, Elsevier, ISSN 0167-739X (2017)
22. Robin Somer and Vern Paxson: Outside the closed world: On using machine learning for network intrusion detection. *IEEE Xplore*, July 8, 2010
23. Bayu Adhi Tama and Kyung-Hyune Rhee: A Detailed Analysis of Classifier Ensembles for Intrusion Detection in Wireless Network. *J Inf Process Syst*, 13(5), 1203~1212, October 2017
24. C. F. Tsai, Y. F. Hsu, C. Y. Lin, and W. Y. Lin: Intrusion detection by machine learning: a review. *Expert Systems with Applications*, 36(10), 11994-12000 (2009)
25. V Rao Vemuri, (Ed.): *Enhancing computer security with smart technology*, CRC Press, Auerbach Publications, November 21, 2005.

Innovating a seamless customer experience

Bipin Sahni¹

Abstract The pace of innovation presents multiple challenges for companies, but perhaps the most critical is how to meet and exceed customers' ever-increasing user experience expectations. However, wherever there is a challenge, there is also an opportunity; innovating faster and better than your competitors can increase market share. In this article, we examine four innovations – Internet of Everything (IoE), Artificial Intelligence, Biometrics, and Mixed Reality – and discuss how they can deliver frictionless transactions to improve the user experience for our customers.

Keywords: Financial Services · Financial Technology · Innovations · Artificial Intelligence · Biometrics · Mixed Reality · User Experience

1 Introduction

Innovation is constant. New consumer tech products and services have elevated customer expectations when it comes to the way products are designed and more importantly, how they are experienced. In many cases, a technology's transition – smartphones, for example – from shiny new thing to status quo has become not just astonishingly quick, but normal. There are constant demands to continuously improve and integrate across platforms. On top of building a rewarding experience, in the case of the finance industry, building a secure experience is of utmost importance.

For large organizations, keeping up with innovation and delivering new products and services at the necessary pace is an opportunity. Fall behind the fast moving curve of customer expectations and you can lose business.

Any innovation for new and emerging platforms must have two key principles at their core: security and user experience. Innovation around any new platform must have strong security technology incorporated from the outset. Equally, the objective of an innovation must be to enhance the experience of the customer or client. Innovation for the sake of innovation detracts from your ability to focus on delivering the seamless user experience we have all come to expect and demand.

As we are confronted by so many shiny objects, we apply our overarching principles to some of the areas we anticipate having the greatest impact for our

✉ Bipin Sahni
InnovationGroup@wellsfargo.com

¹ Head – Innovation, R&D, Innovation Group, Wells Fargo

customers. These include the Internet of Everything, Artificial Intelligence, Biometrics, and Mixed Reality.

2 Internet of everything

2.1 Connecting data in one ecosystem

According to Analytics Week [1], by the year 2020, there will be about 1.7 megabytes of new information created every second for every human being on the planet. Part of what will drive the seemingly ever-increasing explosion of data is the Internet of Everything.

The Internet of Everything enables networking and data-sharing with any connected device, and the number of connected devices is only going to grow: from iPhones and wearables to refrigerators and cars. IoE, and the smart networks that support and analyze all the data these objects generate, will mean people will be able to interact with devices in a way that has not been possible before.

For example, a smart fridge could alert you when you need milk. By saying “yes” when prompted by your smartphone, you will trigger approval for the refrigerator to place an order for more milk, which could then be delivered by your local grocery store.

Another example could be streaming the latest Hollywood or Bollywood movie to your smart TV. The fee for the film could depend on the number of people watching in your living room, which could be calculated by a smart LED lightbulb detecting the presence of people in its vicinity.

With both of these examples, people are doing ordinary things – buying milk and watching movies – but in a radically different way. And what is that radical difference? A financial transaction that is frictionless within the overall user experience.

2.2 Endless applications

Other use cases are already proving themselves. For example, applications such as auto insurance telematics that capture driver performance, or commercial real estate building-management systems that use built-in sensors to manage energy usage, environmental comfort, and security, are already becoming prevalent.

Within finance, examples include IoE-sensitive lease pricing, which uses data collected from connected devices to better understand collateral values for leasing deals, or using IoE data to assess creditworthiness. As devices and data become more connected, banks also have the opportunity to begin functioning as platforms for micropayments.

The financial applications of IoE are endless, and will become integrated into the facets of daily life. After all, there is a financial component to almost everything we do.

2.3 Overcoming fraud and avoiding risk

Of course, smart technology never comes risk-free. The ability to clone IoE devices or take control of them could lead to disastrous effects for our connected world. Enacting encryption, privacy policies and compliance programs is therefore critical. While financial institutions are building applications with security in mind, customers should review what types of authentication controls are available within their connected devices. Technology owners must continuously review their applications for vulnerabilities. As we see an increased number of connected devices being built, it's more important than ever for vendors to build in strong authentication controls during the development stage.

3 Artificial intelligence

3.1 Existing data, new experiences

Artificial Intelligence (AI) enables us to use data in a whole new way. We are exploring ways to use AI to deliver insightful and personalized experiences for customers and team members. In a basic example, using a virtual personal assistant, customers could ask, "What is my checking account balance?"

But if a virtual assistant can also unlock your account after a fraud alert, or answer questions about fees and schedule upcoming payments, what else can it do? As AI evolves, expect to see increased enhancements to automated investment portfolios beyond traditional "robo-advisors." These computer-generated experts will evolve beyond simple rules-based models to understand complex financial issues and can help you set and achieve personalized short and long-term goals.

3.2 Facebook Messenger AI Assistant

We piloted an AI assistant within the Facebook Messenger platform with 5,000 customers and team members. AI can help augment the customer experience by bringing a "best of both worlds" approach: high-performance AI technology enhancing the personal touch of a human banker. Our chatbot [2] for Messenger pilot has provided an excellent opportunity to innovate on platforms customers use on a daily basis giving them options on when, where and how they want to be served.

Customers who are looking for information can direct message to the virtual banking assistant and engage in an interactive chat. The assistant can aid with queries related to account balances and most recent transactions, or even more intricate information, like how much the customer spent on food last week or the location of the nearest ATM. With every interaction, the chatbot will become more conversational and ultimately transition more intricate conversations to a banker for a seamless customer experience.

In the future, how a bank interacts with its customers will depend on where and when the customer demands and expects. Fail to make yourself available on the customer's platform of choice? He or she will find a bank that will.

3.3 Overcoming concerns

Sceptics usually have some reservations around AI. One is that the advancement of AI – particularly in the form of process automation – will replace people in their jobs. In reality, AI will primarily eliminate the tasks that are monotonous and don't require a human touch. In this way, it's actually more likely AI will improve the workforce, helping people to be more efficient and freeing them up to deliver higher value work.

AI technology has the potential to become “creepy,” acting like Big Brother and peering in on our lives. To quell these concerns, it's key that institutions use AI to address customer experience pain points, offering solutions that are timely, useful, and relevant.

But most important of all, for AI to deliver on its promise, customers must opt in to the service, knowing that there is a tradeoff between providing personal details and a seamless experience. If AI experiences are delivered to the customer without their buy-in, that is when the creepiness factor can negate the whole objective.

3.4 There will always be risks

These concerns aside, AI does pose risks that need to be addressed by the industry. Like any technology, AI is sure to meet unforeseen security vulnerabilities. It's crucial that policy and rules around customer interaction, data usage and privacy are in place, as well as a plan to mitigate reputation and regulatory risk.

Even with these safeguards, it remains to be seen how people will interact with AI. As with any technology, generations will take to advancement in varying degrees. The type of product and its applications will also likely determine the consistency of adoption, depending on how natural it feels for customers to integrate the new service into their lives. Financial services firms hoping to integrate AI should focus not just on making AI work, but making it work in a way that's meaningful and comfortable for customers.

4 Biometrics

4.1 What you know, what you have, what you are

The death of the password is almost upon us. As major companies have experienced breach after breach, they are looking to secure their systems by ensuring people can gain access not by what they know (passwords) but what they have (cell phones or tokens) and increasingly, what they are (fingerprints or eye scans).

Biometric authentication aims to make a user's body the new password, and it is already becoming a common part of our lives, from airports to smartphones. Some companies are ahead of the curve and have already implemented biometrics in financial services applications. A common example of this is Apple Pay, where a customer is required to use his or her fingerprint, or more recently, face recognition, to complete a transaction.

In the very near future, we are likely to see more types of biometric solutions in the market. These include voice authentication, either as a standalone verification method or in conjunction with another method such as voice and face recognition at the same time. Other methods could include using your heartbeat – a unique personal identifier – to provide access to your financial accounts. Or, in the category of “what you are” behavioral authentication, the gait of your walk or how you normally handle your phone. It's too soon to say when these new technologies may permanently replace the password.

4.2 Biometric challenges

What if you're in a loud room attempting to use voice authentication? Not an unlikely scenario. To be entirely fool-proof, technologists should offer multiple forms of authentication for situational use, and account for environmental settings like lighting and noise. These and other challenges in biometric adoption will be common, but while there are challenges, and they may slow down adoption for some people, the potential of biometrics makes it inevitable that everyone will use “what they are” to login one day very soon.

As with IoE and AI, data must be encrypted to protect customers. Privacy and compliance programs involving biometrics should be thorough and will need to work securely with existing applications. People need to feel comfortable that everything is secure before offering biometric information as a form of authentication, because it's not something that can be changed like a password.

4.3 Better security, better experience

These challenges aside, no matter what particular kind of authentication method a company uses, biometrics breaks the security trend of the past 10 to 20 years – that increasing security results in a worse user experience. First, there were passwords, next, there were more complicated passwords, and then, there were tokens. Each step increased security but decreased convenience. Biometrics means, we can deliver on two key goals – increase security for our customers and deliver a more seamless and convenient user experience.

5 Mixed reality

5.1 The future might not be right now, but it's coming soon

Banks are paying attention to advances in technologies and customer experiences across emerging platforms. Over the past decade, the banking industry has made significant strides in making the day-to-day experience compatible with mobile.

Hardware supporting mixed reality is still limited, and existing applications mostly support gaming and entertainment. For financial institutions, what is useful is, to think about the experience their organization would want to create on that platform, and experiment by building prototypes and proofs of concept. By doing so, companies will be prepared to react and deliver their services as new platforms gain traction.

5.2 Understanding the *why* of successful innovation

But the user experience is always about more than just the technology – it is how people use it. We bring our customers into the process, through pilots and interviews, to identify what they are looking for and explore concepts collectively. We also learn from popular experiences (in this case often gaming and entertainment) to understand why are they successful, and where can we emulate their strengths.

The hardware and devices for this technology will help determine the timing of adoption, as will the experiences created by organizations across industries. Mixed reality platforms will not only serve as a new medium to introduce financial information, but will almost assuredly incorporate the principles of IoE, AI and biometrics, creating new sources of data, and new ways to authenticate.

6 Conclusion

Customers gravitate toward products and services that reduce hassle and save time. Investment in innovation is therefore pivotal to improving the end user experience. We of course cannot predict what the future will hold, but we're working tirelessly to create a framework that can deliver a seamless, frictionless experience for every financial transaction, regardless of what platforms emerge in the future.

We must look beyond our own organization to ensure we are exposed to the latest innovations and trends that will impact our business.

References

1. Vishal Kumar: Big data facts, Analytics Week. <https://analyticsweek.com/content/big-data-facts/>, March 26, 2017
2. Wells Fargo: Startup Accelerator. www.wellsfargo.com/accelerator

National payment system – overview of regulatory mandates*

Gynedi Srinivas¹ · Harish Natarajan¹

Abstract This article discusses the concept of the national payment system (NPS) and the different regulatory models being applied in select countries for conducting oversight over different components of the NPS. We describe the core components of the NPS and their use in the World Bank in its payment system technical assistance projects. Further, we discuss the need to define the regulatory mandates and parameters when different regulators are involved for the oversight of the different components of the NPS, drawing upon international examples comprising the European Union (Germany and Luxembourg) and Turkey.

Keywords: National payment system · Regulatory models · Nine pillars methodology · Financial stability · Payment systems and operations

1 Introduction

The oversight of payment and settlement systems is recognized as a “central bank function whereby the objectives of safety and efficiency are promoted by monitoring existing and planned systems, assessing them against these objectives and, where necessary, inducing change”.² The oversight of payment and settlement systems over the years has often been as a critical function of the central bank contributing to the overall financial stability, along with prudential supervision (which function could be with an independent banking supervisory authority).

Worldwide, the trend has been to recognise and strengthen the role of the central banks in ensuring financial stability. As payment and settlement systems is the major transmission channel of risk, the need to regulate and oversee payment and settlement systems as part of the wider mandate of ensuring financial stability of the central banks is being underscored in all the economies world over. While this is the case, the advent of technology and the entry of non-banks into the payments area, especially in the area

* The views expressed in this article represent the personal views of the authors and do not represent the views of the World Bank.

✉ Gynedi Srinivas
gstrinivas1@worldbank.org

Harish Natarajan
hatarajan@worldbank.org

¹ Senior Financial Sector Specialist, Payment Systems Development Group, Finance and Markets, World Bank

¹ Lead Financial Sector Specialist, Finance and Markets Global Practice, World Bank

² See Bank for International Settlements, Committee on Payment and Settlements Systems Central bank oversight of payment and settlement systems, May 2005.

of retail payments, and to promote innovation and competition, the oversight over certain components of the national payments system (NPS) are being reassessed in several countries. For instance, in the case of UK, a new Payment Systems Regulator³ has been set up to regulate and oversee the retail payment systems, with the Bank of England⁴ being responsible for the oversight and supervision of financial market infrastructures (FMIs), as part of its mandate for ensuring financial stability. In contrast, all aspects of the payments system in Indonesia are regulated and overseen by Bank of Indonesia⁵, as part of its mandate of achieving and maintaining the stability of the value of the Rupiah and ensuring financial system stability. The Financial Services Authority (OJK)⁶, a separate regulator in Indonesia is tasked with regulating and supervising financial services activities in banking, capital markets, non-banking financial sector and consumer protection.

Based on the legal mandate and the division of responsibilities, it is quite often the case that different authorities including the central bank are involved in discharging the oversight function over the national payments system (NPS). For instance, the licensing and regulation of non-bank payment service providers and operators could be vested in a different regulatory authority. Other public policy objectives such as anti-money laundering and combating the financing of terrorism (AML/CFT), consumer protection, anti-competitive practices, could all be under the aegis of different authorities.

This article while taking into account these broad trends, throws light on the concept of the NPS and the different regulatory models being applied in select countries for conducting oversight over different components of the NPS. In section 2, a definition and outline of the NPS is presented. The core components of the NPS as outlined in the “General guidance of national payment system development⁷”, the “nine pillars” methodological approach derived from the above and used by the World Bank in its payment system technical assistance projects, are discussed in section 3. Section 4 talks about the need to define the regulatory mandates and parameters when different regulators are involved for the oversight of the different components of the NPS, drawing upon international examples comprising the European Union (Germany and Luxembourg) and Turkey. These examples are used to present the two stylized models of oversight and supervision over payment and settlement systems that are followed in these countries. Section 5 provides the conclusion, highlighting the need for an effective cooperative framework between the regulators for the safe and efficient functioning of the NPS.

³ The Payment Systems Regulator (PSR) is a subsidiary of the Financial Conduct Authority of the UK, and is independent, with its own board and managing director. (Source: <https://www.psr.org.uk/>)

⁴ The Bank of England supervises three main types of FMIs: (i) recognised payment systems; (ii) central securities depositories; and (iii) central counterparties. It works with the Financial Conduct Authority (FCA) and overseas regulators to supervise FMIs. (Source: <https://www.bankofengland.co.uk/financial-stability/financial-market-infrastructure-supervision>)

⁵ <http://www.bi.go.id/en/tentang-bi/fungsi-bi/tujuan/Contents/Default.aspx>

⁶ <http://www.ojk.go.id/en/tentang-ojk/Pages/Tugas-dan-Fungsi.aspx>

⁷ See Bank for International Settlements, Committee on Payments and Market Infrastructures (formerly CPPS), General guidance for national payment system development, 2006. The guidance framework is widely used in all World Bank payments system technical assistance projects.

2 The national payment system (NPS)

A country's National Payment System (NPS) consists of a defined group of institutions and a set of instruments and procedures, used to facilitate the circulation of money within the country and internationally. The main elements of a modern national payments system include:⁸

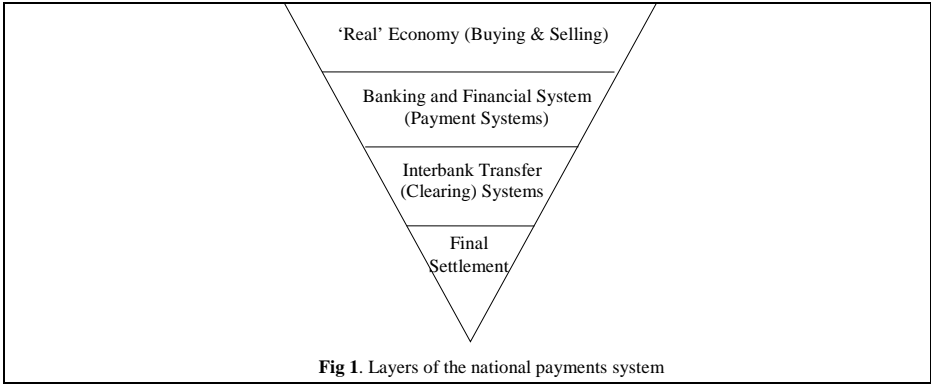
- Payment instruments used to initiate and direct the transfer of funds between the accounts of payers and payees;
- Payment infrastructures for transacting and clearing payment instruments, processing and communicating payment information, and transferring the funds between the paying and receiving institutions;
- Electronic book-entry securities system(s) to register and record changes in ownership of both private and government securities;
- Financial institutions that provide payment accounts, instruments and services to consumers, and businesses and organizations that operate payment transaction, clearing and settlement service networks for those financial institutions;
- Non-financial institutions that provide payment and access to payment-related services and offer various products to satisfy market needs;
- Market arrangements such as conventions, regulations and contracts for producing, pricing, delivering, and acquiring the various payment instruments and services;
- Laws, standards, rules, and procedures set by legislators, courts, and regulators that define and govern the mechanics of the payment transfer process and the conduct of payment service markets.

As can be seen from the above, the NPS is a broad concept and encompasses a country's entire matrix of institutional and infrastructure arrangements and processes for initiating and transferring monetary claims in the form of commercial bank and central bank liabilities. Seen from this perspective, any country's economy can be viewed as a series of layers in an inverted pyramid, as depicted in Figure 1, in which each layer is supported by the layers beneath it. The broadest layer of the pyramid represents the real economy and the financial markets, i.e. the buying and selling of goods and services throughout the nation. It is supported by the country's banking system – the next level of the pyramid – which provides payment services to all sectors of the economy.⁹ The third level consists of a limited number of interbank value transfer systems through which payment and other financial transactions are processed. The final settlement of funds transfers takes place across the accounts

⁸ See Bank for International Settlements, Committee on Payments and Market Infrastructures (formerly CPPS), General guidance for national payment system development, 2006.

⁹ Comprising the individual, retail, industrial and commercial, financial, government, and international sectors.

which approved institutions hold with the central bank, whose pivotal role is vital to the functioning of the economy as a whole.



The NPS is therefore a core component of the broader financial system and is the infrastructure that provides the economy with the channels or circuits for processing the payments resulting from the many different types of economic transactions that take place on a daily basis. It can be viewed as the ‘lubrication system’ for the engine of the economy.

A well-functioning NPS requires a delicate balance between market-driven competition, cooperation and public good considerations. In a mature environment, banks and other payment service providers should compete for payments business and customers, while achieving the benefits and efficiencies that stem from the sharing of non-competitive infrastructures. In the national interest, it is imperative that economies of scale are achieved and that the national payment service infrastructure allows as many participants as possible to offer their services to the public.

2.1 Scope and components of a NPS

Taking account of the previous points, it is clear that a comprehensive NPS comprises, a country’s entire matrix of institutional and infrastructure arrangements and processes for initiating and transferring monetary claims in the form of commercial bank and central bank liabilities. The main elements of a NPS, therefore, include the following:

- Institutions providing financial intermediation;
- Non-financial institutions and financial institutions that provide payment services;
- Businesses and organizations that operate payment transaction, clearing and settlement service networks (payment systems);
- Institutions providing services to payment systems (payment system operators – PSOs) and payment service providers (PSPs);
- A legal and statutory framework;
- Rules, regulations and agreements;
- Appropriate payment instruments;

- Processing systems and procedures;
- A cost-effective technological infrastructure for telecommunications;
- Clearing and settlement mechanisms that adequately balance risk and efficiency requirements;
- An appropriate oversight framework including inter-regulatory cooperation.

The trends in NPS development focusses on initiatives such as: (i) broadening the range of payment instruments and services including a wider pool of providers inclusive of non-banks; (ii) contributing to financial inclusion; (iii) encouraging innovation; (iv) improving cost efficiency; (v) augmenting consumer protection measures; (vi) improving anti-money laundering and countering the financing of terrorism measures; (viii) enhancing interoperability and resilience; (ix) better containing legal, operational, financial and systemic risks in payment infrastructures; (x) improving market contestability by limiting anti-competitive practices; (xi) creating more suitable supervisory, oversight¹⁰ and regulatory regimes for the national payment system; and (xi) enhancing the overall safety and efficiency of the NPS.

In some countries, all of these responsibilities vest with the central bank which is charged with the overall development of the NPS, while in other countries some of the aspects of these responsibilities and activities vest with other regulators. Notwithstanding the presence of different regulators, it is particularly important to recognize the need for cooperation among the regulators to achieve the above objectives in totality for the overall safety, efficiency and development of the NPS. This is important and relevant as the infrastructures for retail payment, securities settlement and large-value payment systems are interconnected and interdependent. As the systems develop, the country's principal financial institutions generally become participants in all of them. Settlement in one system could affect the safety and efficiency of settlement in the others. Therefore, in order to mitigate the potential cross-system risks, the legal, operational, financial and systemic risks need to be monitored and well-managed through a cooperative framework among the regulators. The safety and efficiency of the NPS in terms of the various payment systems and instruments also has a bearing on the public confidence in money and in the use of more electronic payments.

3 The “nine pillars” methodology

The above framework in conjunction with the recommendations made in various Committee on Payments and Market Infrastructures (CPMI, formerly CPSS) publications and other World Bank publications¹¹, notably the World Bank's Retail payments package¹² and the experience and previous work of the World Bank on payment systems development in several countries around the world, the World Bank has developed the “nine pillars” methodological approach to the development of the

¹⁰ By convention, the term oversight is reserved to designate the specific responsibilities and tools central banks have with regard to payment and settlement systems due to their unique character of being both a public authority and a bank. See Bank for International Settlements, Committee on Payment and Settlements Systems Central bank oversight of payment and settlement systems, May 2005.

¹¹ See Bank for International Settlements, Committee on Payments and Market Infrastructures (formerly CPSS) (i) “General guidance for national payment systems development”; (ii) “General principles for international remittance services”

¹² “Developing a comprehensive national retail payments strategy”; “A practical guide for retail payments stocktaking”; “From remittances to m-payments: Understanding ‘alternative’ means of payment within the common framework of retail payments system regulation”; “Innovations in retail payments worldwide: A snapshot: Outcomes of the global survey on innovations in retail payments instruments and methods 2010” <http://www.worldbank.org/en/topic/paymentssystemremittances/brief/retail-package>

NPS strategy. The “nine pillars” strategy, aims to provide public authorities and market participants with detailed guidance on how to develop and implement a comprehensive, strategic payments reform process; identifies a methodology for undertaking a detailed stock-taking of a country’s national payments landscape; explores the development of a normative framework that underpins the development of a safe and efficient payment system including retail payments involving innovative payment mechanisms.

The nine pillars take into account all the parameters necessary for the successful development of a country’s national payment system. These comprise: Pillar 1 – the legal framework; Pillar 2 – the settlement mechanisms for large-value and time-critical payments; Pillar 3 – the securities settlement systems; Pillar 4 – the inter-bank money markets; Pillar 5 – the retail payment systems; Pillar 6 – government payments; Pillar 7 – remittances; Pillar 8 – oversight of the NPS; and Pillar 9 – co-operative framework for the development of the NPS.

To provide a better understanding, a brief description of each of the “nine pillars” is given below along with their respective roles in the country’s national payment system.

Legal framework: The legal framework comprises the general laws in the country as well as statutes specific to the safe and efficient functioning of the NPS and includes the relevant rules and regulations framed thereunder. All elements of the NPS should be based on a well-founded, clear, transparent, and enforceable legal basis. The legal basis provides the foundation for relevant parties in the NPS to define respective rights and obligations, provides the rules and procedures which should be enforceable, and provides the overall framework for a sound risk management. In addition, regulatory and supervisory powers of the regulator including licensing, sanctions etc., are derived from the legal framework.

Large-value payments system: A payment system is generally categorized as either a large-value payment system (LVPS) or as a retail payment system. An LVPS is a funds transfer system that typically handles large-value and high-priority payments. In contrast to retail systems, many LVPSs are operated by central banks, using an RTGS or equivalent mechanism. LVPS are generally classified as systemically important payment systems. In general, a LVPS is considered to be systemically important – if it has the potential to trigger or transmit systemic disruptions; is the sole payment system in a country or the principal system in terms of the aggregate value of payments; and could be systems that mainly handle time-critical, high-value payments; and systems that settle payments used to effect settlement in other systemically important financial market infrastructures.

Securities Settlement Systems (SSS) including Central Securities Depositories (CSD): A securities settlement system enables securities to be transferred and settled by book entry according to a set of predetermined multilateral rules. Such systems allow transfer of securities either free of payment or against payment. When transfer is against payment, many systems provide delivery versus payment (DvP), where delivery of the security occurs if and only if payment occurs. A central securities depository provides securities accounts, central safekeeping services, and asset services, which may include the administration of corporate actions and redemptions, and plays an important role in helping to ensure the integrity of securities issues (that is, ensure that securities are not accidentally or fraudulently created or destroyed or

their details changed). A CSD can hold securities either in physical form (but immobilized) or in dematerialized form (that is, they exist only as electronic records). In many countries, a CSD also operates a securities settlement system. SSS in the broadest sense of the term are also critical for the functioning of the LVPS for facilitating the provision of collateralized credit facilities to the participants in the LVPS; and in the implementation of monetary policy operations.

Inter-bank money markets: Inter-bank money markets serve an important purpose in serving and fulfilling the liquidity management needs of the banks. They serve as channels of liquidity distribution between banks in conjunction with any liquidity support provided by the central bank to the banks. The inter-bank borrowing could either be on a collateralized basis or on an uncollateralized basis. The lending and borrowing of funds is effected through the LVPS and where the borrowing is collateralized, the SSS along with the LVPS is also involved.

Retail payment systems: A retail payment system is a funds transfer system that typically handles a large volume of relatively low-value payments in such forms as cheques, credit transfers, direct debits, card payment and increasingly mobile payment transactions. Retail payment systems may be operated either by the private sector or the public sector, and are usually settled on a multilateral deferred net settlement (DNS) basis. Retail payment systems are used by individuals, households, businesses with one of the biggest users of the retail payments being the Government both for its payments as well as its receipts.

While banks were previously largely involved in retail payments (non-banks were involved in card schemes), with the advent and greater use of technology a variety of non-banks have come to be involved in every stage of the retail payment transaction life cycle including in the clearing and settlement process. A stylized retail payments model is provided in Figure 2 below to capture this in greater detail.

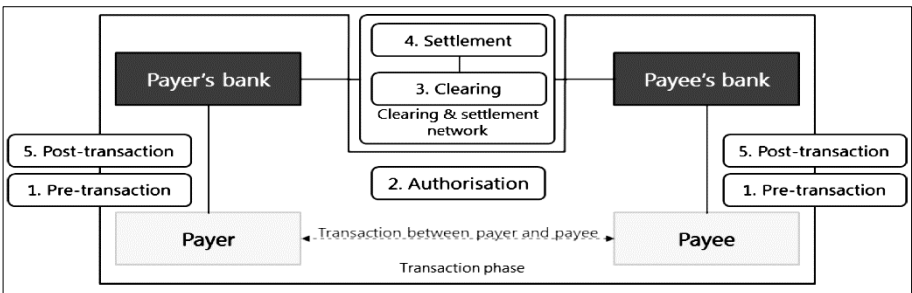


Fig 2. Retail payments, “Non-banks in retail payments”, <http://www.bis.org/cpmi/publ/d118.pdf>, September 2014

In the above retail payments model, non-banks can be involved at all stages of the payments process. Typically, the role of non-banks can be classified under four distinct categories, based on the roles played by them at the different stages of the

payment chain; the type of service provided; and the predominant type of relationship with banks. Non-banks can accordingly be classified as¹³:

- (i) **Front-end providers:** Non-banks as front-end providers usually comprise mobile wallets, internet payment gateway providers, credit card acquirers or, payment institutions and e-money institutions in the EU. They are not usually involved in clearing and settlement, but operate in the pre-transaction, initiation, and post-transaction stages of a payment transaction. They may often compete with banks for providing these services, while also usually having a cooperative framework with banks for the clearing and settlement of their transactions.
- (ii) **Back-end providers:** Back-end non-bank providers do not have a direct relationship with the payer or payee as is the case with front-end providers. They mostly provide specialized back-end services relating to several payment instruments to banks, either through an outsourcing or a cooperative arrangement, focusing on or two stages of the payments process. Typical examples include information technology (IT) services, providers of data centre services, trusted service managers, data security firms, or entities that provide back office operations, anti-money laundering, audit or compliance.
- (iii) **Operators of retail payment infrastructure:** As indicated by the name, these entities operate the retail payment infrastructure by providing clearing and processing services with respect to retail payment instruments. The card networks are a typical example of this kind of activity.
- (iv) **End-to-end providers:** Payers and payees have a direct relationship with end-to-end providers. These are closed-loop systems, where the movement of funds from a payer's account to a payee's account does not necessarily involve the use of a bank. Examples of such closed-loop systems are the three-party card schemes, some types of remittance services and e-money products. Banks could still be used to redeem or fund end-user accounts with such providers and could also act as agents for such providers. Other non-banks could also act as agents for such end-to-end providers.

Government payments: As indicated above, the Government is one of the biggest users of the payments infrastructure in the country. The Government uses the payments system to affect salary payments, vendor payments, and for the disbursement of social security benefits. On the receipts side, the Government uses the payment systems extensively for collection of taxes, customs, revenues, etc.

Remittances: Remittances are one of the sources of external financing in developing countries and contribute significantly to the GDP in several countries. These remittances are largely constituted by the flow of funds from migrant workers back to their families in their home country. The payment system aspects of remittances, deal with the remittance services market to be contestable, transparent, accessible and sound.

¹³ See Bank for International Settlements, Committee on Payments and Market Infrastructures (formerly CPPS) "Non-banks in retail payments" September, 2014 <http://www.bis.org/cpmi/publ/d118.pdf>

Oversight and cooperation: The last pillar is oversight and cooperation. By convention, the term oversight is reserved to designate the specific responsibilities and tools central banks have with regard to payment and settlement systems due to their unique character of being both a public authority and a bank. The broad definition of oversight used here, also subsumes the catalyst function within the oversight framework, insofar as the same is part of the broader oversight objectives of the central bank. The oversight function as indicated in the definition consists of three activities: (i) monitoring; (ii) assessing; and (iii) inducing change.

In addition to central banks, in certain jurisdictions, other regulatory authorities could have regulatory and supervisory powers over particular aspects of payment and settlement systems. These include other authorities such as banking supervisors, securities regulator, competition authority, consumer protection authority and financial intelligence unit authorities. The range of regulatory and supervisory activities could include licensing, periodic submission of data, on-site inspections and off-site monitoring. In all such cases, it is imperative that a cooperative framework is established between and amongst the relevant authorities to avoid duplication of efforts amongst them, ease regulatory burden of the overseen entities and avoid and minimise any potential regulatory arbitrage.

It is usually the case that in these situations, a Memorandum of Understanding is entered into by the regulators involved clearly outlining the roles and responsibilities of each regulator. It also identifies areas of mutual cooperation in terms of sharing of information, data and reports and coordinating any crisis management situations. Thus, each authority should have well-defined responsibilities and specific tools to carry out the responsibilities, assigned to it under its regulatory mandate.

In addition, it is also critical to note that the regulators should establish a forum for interaction with all the payment system stakeholders, as a consultative and collaborative approach for successful payment system modernization. In several countries, central banks as payment system overseers have been entrusted with the responsibility to set up and lead the NPC, with a view to creating a safe, sound and efficient national payment system that meets the evolving needs of the various stakeholders while preserving the safety of their financial transactions. Implementation of industry-wide initiatives require enlisting the support of all payment system stakeholders, to ensure a successful buy-in. Accordingly, in many countries a permanent National Payment Council (NPC) has been established in which all the relevant sectoral regulators (central bank, securities regulator, banking supervisor, telecom authority), the Government (as a major user of the payment system) are represented along with all other payment system stakeholders (banks and non-banks as participants and payment system operators and providers), and consumer organisations. The NPC serves as the forum for such cooperation and collaboration.

The NPC in many countries is tasked with preparing a strategic and holistic action plan with defined timelines for the modernization and development of payment and settlement systems in line with international standards. To this end, specific taskforces of the members are constituted to work on issues related to payment systems, such as

legal and regulatory matters; IT issues; devising plans to leverage existing payment infrastructure to promote retail electronic payment products; introducing interoperable e-money products; promoting online payments through the internet and mobile with adequate security measures; deepening the PoS infrastructure; rationalizing the fee structure; consumer protection issues, etc. It should however be noted that the role of NPC is to act only as a consultative body, and not as a decision-making body.

4 Defining and establishing regulatory mandates – two models

Two international models of regulatory mandates covering the NPS are presented below. In both the models, more than one regulator is involved for different aspects related to the NPS. In Model 1, the examples of Germany and Luxembourg are presented under the overarching EU framework; and in Model 2, the example of the regulatory structure in Turkey is outlined. In both the Models, it should be noted that the authorities have taken into account the strides in technology and entry of non-banks into the payments arena and have framed their regulatory structures outlining the roles of each individual regulator based on their respective legal mandates, in the light of these developments.

4.1 Model 1 – Germany

In Germany, the Deutsche Bundesbank is responsible for the oversight of all payment and settlement system inclusive of payment instruments and it carries out this oversight in consonance with the ECB oversight framework and relevant national laws. The legal basis for oversight activities is enshrined in Article 127 (2) of the Treaty on the Functioning of the European Union and Article 3.1 of the Protocol on the Statute of the European System of Central Banks and of the European Central Bank (Statute of the ESCB), under which the ESCB is inter alia mandated “to promote the smooth operation of payment systems”¹⁴. Article 22 of the ESCB Statute provides an additional legal basis. According to this, the ECB and national central banks may provide facilities, and the ECB may make regulations, to ensure efficient and sound payment and settlement systems within the currency union and with other countries.

The Bundesbank not only acts as an operator of payment systems, but is also entrusted with the task of oversight. This is reflected in section 3 of the Bundesbank Act (Gesetz über die Deutsche Bundesbank) which states that the Bundesbank “shall arrange for the execution of domestic and cross-border payments and shall contribute to the stability of payment and clearing systems.”¹⁵ Its oversight activities are focused on financial market infrastructures, payment instruments and critical service providers for infrastructures and banks. Amongst other aspects related to payment instruments, the security features of payment instruments, such as payment cards, credit transfers, direct debits and e-money, also forms a part of Bundesbank’s oversight activities.

¹⁴ See https://www.bundesbank.de/Navigation/EN/Tasks/Payment_systems/Oversight/oversight.html

¹⁵ See https://www.bundesbank.de/Navigation/EN/Tasks/Payment_systems/Oversight/oversight.html

In terms of section 8 (1) of the Payment Services Supervision Act (PSSA)¹⁶, any institution wishing to provide payment services as a payment institution in Germany, needs written authorization from the Federal Financial Supervisory Authority (BaFin). The same requirement is applicable to e-money institutions pursuant to section 8a of the PSOA. The provision of payment services is included in the license to conduct e-money business.

Further, the PSSA in section 1 defines payment service providers and payment services. Sub-section 2 of section 1 defines payment services and sub-section 10 (7) excludes “payment transactions carried out within a payment or securities settlement system between settlement agents, central counterparties, clearing houses or central banks and other participants of the system, and payment service providers” from the ambit of payment services”. Read with the relevant statutes (Bundesbank Act, ESCB statute and the ECB statute), it is evident that the Bundesbank is charged with the oversight of all the activities excluded from the ambit of BaFin under sub section 10 (7) of the PSSA. The provisions in the respective statutes as indicated above clearly establish the respective regulatory mandates of the Bundesbank and the BaFin.

Since the PSSA coming into force, these institutions have been legally defined as two categories of payment service providers. As defined in section 1 (2) of the PSSA, payment services comprise: (i) deposit and withdrawal transactions; (ii) payment transactions in the form of direct debits, credit transfers and card payments (excluding the granting of credit payment transactions involving the granting of credit); (iii) payment authentication transactions; (iv) digital payment transactions; and (v) money transmission services.

As part of the licensing procedure, both groups of institutions are required to submit, a business model, a business plan with a budget plan for the first three financial years and a description of the measures required to fulfil the security requirements of section 13 of the PSSA. Initial capital requirements are different for payment institutions and e-money institutions¹⁷.

4.1.1 Cooperation between the Bundesbank and BaFin

The cooperation between Bundesbank and BaFin is defined by the PSSA Act and the Banking Act¹⁸. Section 3 (3) of the PSSA enjoins upon BaFin and the Deutsche Bundesbank to work together in accordance with PSSA and states that section 7 of the Banking Act shall apply accordingly. Section 7, sub-section 3 of the Banking Act states that “BaFin and the Deutsche Bundesbank shall exchange observations and findings that are necessary for the performance of their respective functions”; while section 4 authorizes sharing of data between both the regulators including access to each other’s data bases for discharging their regulatory functions.

¹⁶ https://www.BaFin.de/SharedDocs/Downloads/EN/Aufsichtsrecht/dl_zag_en.pdf?__blob=publicationFile&v=2

¹⁷ https://www.bundesbank.de/Navigation/EN/Tasks/Banking_supervision/Payment_institutions/payment_institutions.html

¹⁸ https://www.BaFin.de/SharedDocs/Downloads/EN/Aufsichtsrecht/dl_kwg_en.pdf?__blob=publicationFile&v=3

4.2 Model 1 – Luxembourg

The Banque centrale du Luxembourg (BCL), as a member of the Eurosystem, has the mission to promote the smooth functioning of payment and settlement infrastructures. The BCL's mission is stipulated in the national law of 23 December 1998 as amended and in particular articles 2 (5), 27-3 and 34 (1). The BCL's mission is in conformity with the article 127 (2) and (5) of the Treaty of the Functioning of the European Union. The BCL is charged with ensuring the efficiency and safety of payment systems and securities settlement systems, as well as the safety of payment instruments¹⁹. In this context, the BCL adopted a regulation²⁰, which specifies the scope of its oversight mission, the general framework and the means by which it carries out oversight activities.

The second regulator in Luxembourg, the Commission de Surveillance du Secteur Financier (CSSF) supervises, regulates, authorizes, informs, and, where appropriate, carries out on-site inspections and issues sanctions, with regard to payment service providers (which include payment and electronic money institutions). Payment institutions and electronic money institutions are governed by the law of 10 November 2009 on payment services (“the PSL”)²¹. Payment services are defined in Article 1(38) of the PSL as “any business activity performed on a professional basis listed in the Annex”. Electronic money is defined in Article 1(29) of the PSL. Payment services are listed in the Annex to the PSL.

No payment institution or electronic money institution may be established in Luxembourg without holding a written authorization from CSSF (Articles 6 and 24-2 of the PSL). The entity submits a formal application to the CSSF, including all the documents and information required using the relevant application forms. The entity makes a presentation on the project to the CSSF. The CSSF then carries out a detailed analysis of the application and having been satisfied, issues a letter authorizing the applicant to transmit the application file to the competent Minister. After the authorization has been granted, the authorized company is published on the official register maintained by the CSSF²².

4.2.1 Cooperation between BCL and CSSF

The cooperation between BCL and the CSSF is governed by the relevant provisions of the Organic law of the Central Bank of Luxembourg, Law of 23 December 1998 (as amended –CBL law)²³ and the Law of 10 November 2009 on payment services, on the activity of electronic money institution and settlement finality in payment and securities settlement systems (PS law).²⁴

Article 2 (5) of the CBL law, states that CBL and CSSF shall cooperate and coordinate with each other to ensure the safety and efficiency of payment systems and

¹⁹ See http://www.bcl.lu/en/payment-systems/surv_sys/cadre_juridique/index.html

²⁰ http://www.bcl.lu/en/Legal-Framework/documents_national/regulations/_reglements_de_la_bcl/2016_21/index.html

²¹ <http://www.cssf.lu/en/supervision/payment-institutionselectronic-money-institutions/>

²² <https://www.cssf.lu/en/supervision/payment-institutionselectronic-money-institutions/authorisation/>

²³ http://www.bcl.lu/en/Legal-Framework/documents_national/loi_organique/loi_list/organic_law_1-April-2015.pdf

²⁴ http://www.cssf.lu/fileadmin/files/Lois_reglements/Legislation/Lois/L_101109_psd_eng_upd_211212.pdf

securities settlement systems, as well as the safety of payment instruments. Such cooperation will be based on agreements between the two regulators based on their individual mandates as specified in the relevant statutes. Article 27 (3) authorizes the CBL to undertake on-site visits as part of its oversight activities in coordination with CSSF; while Article 33 (2) permits sharing of data between the two regulators without any prejudice to professional secrecy obligations.

In terms of Article 33 (1) of the PS law, the CSSF is mandated to cooperate with CBL for enabling it to discharge its functions as an oversight authority. Article 33 (2) permits the CSSF to exchange information with CBL as the overseer for payment and settlement systems in Luxembourg.

4.3 Model 2 – Turkey

In Turkey, the Türkiye Cumhuriyet Merkez Bankası, (the Central Bank of the Republic of Turkey – CBRT), has regulatory and oversight responsibility for payment and securities settlement systems. The banking supervisor, the Banking Regulation and Supervisory Agency (BRSA), regulates payment service providers which include commercial banks, payment institutions and electronic money institutions, and all payment instruments including cards.

The Law on the Central Bank of the Republic of Turkey²⁵ No. 1211, and the Law on Payment and Securities Settlement Systems, Payment Services and Electronic Money Institutions²⁶ No. 6493, designate and grant powers to CBRT to regulate, oversee and act as the licensing authority for the payment and securities settlement systems in Turkey. Law No. 6493, provides CBRT with the powers to grant licenses (Article 5); and grants oversight powers over such licensed systems (Article 8).

Accordingly, the main objective of the CBRT is to ensure “safe, uninterrupted, efficient and effective functioning of the systems.”²⁷ In pursuance of this objective, the CBRT:

- (i) Issues regulations, which are in compliance with international standards;
- (ii) Performs oversight of the systems in order to ensure their compliance with national and international regulations and standards such as the “Principles for Financial Market Infrastructures”;
- (iii) Cooperate with the authorities, which are responsible for supervision of the financial system, the system operators and the system participants.

The CBRT is thus responsible for leading all initiatives related to payment and securities settlement systems, operating the core systems, and overseeing all payment and securities settlement systems in Turkey.

The Banking Regulation and Supervisory Agency (BRSA), was established in June 1999 according to Banks Act No. 4389 and began to operate in August 2000.

²⁵[http://www.tcmb.gov.tr/wps/wcm/connect/94d352d1-0d45-45a9-99c7-](http://www.tcmb.gov.tr/wps/wcm/connect/94d352d1-0d45-45a9-99c7-b64eab088e09/Law.pdf?MOD=AJPERES&CACHEID=ROOTWORKSPACE94d352d1-0d45-45a9-99c7-b64eab088e09)

[b64eab088e09/Law.pdf?MOD=AJPERES&CACHEID=ROOTWORKSPACE94d352d1-0d45-45a9-99c7-b64eab088e09](http://www.tcmb.gov.tr/wps/wcm/connect/94d352d1-0d45-45a9-99c7-b64eab088e09/Law.pdf?MOD=AJPERES&CACHEID=ROOTWORKSPACE94d352d1-0d45-45a9-99c7-b64eab088e09)

²⁶ https://www.bddk.org.tr/WebSitesi/english/Legislation/129166493kanun_ing.pdf

²⁷ <http://www.tcmb.gov.tr/wps/wcm/connect/TCMB+EN/TCMB+EN/Main+Menu/PAYMENT+SYSTEMS/Objectives+and+Policies>

According to Banking Law²⁸ No. 5411, the role of BRSA is to establish confidence and stability in financial markets, the sound operation of the credit system, foster the development of the financial sector, and protect the rights and interests of depositors. In this regard, the responsibilities of BRSA are three-fold: (i) regulate, monitor and supervise banks and financial holding companies, as well as leasing, factoring and financing companies, and ensure the enforcement of their regulations; (ii) become member of international financial, economic and professional organizations in which domestic and foreign equivalent agencies participate, to sign memorandum of understanding with the authorized bodies of foreign countries regarding the issues that fall under the Agency's duty field; and (iii) fulfil other duties assigned by the law.

The Law on Payment and Securities Settlement Systems, Payment Services and Electronic Money Institutions No. 6493, and the bank cards and credit cards law²⁹, recognize and grant powers of regulation and supervision to BRSA over payment institutions, electronic money institutions and payment instruments including cards. Law no. 6493, provides powers to BRSA to grant permission to a payment institution to offer payment services (Article 14); to an electronic money institution (Article 18); conduct supervision over payment and electronic money institutions (Article 21); power to enforce sanctions, undertake investigations and launch legal proceedings (Article 27).

In addition, all entities intending to establish a card system, issue cards, enter into agreements with merchants, exchange information, and engage in clearing and settlement activities for card based transactions in Turkey, are required to obtain a license from the BRSA in terms of Article 4 of the bank cards and credit cards law. The law provides BRSA with the necessary powers to regulate and supervise the cards market and also has provisions for consumer protection.

The Capital Markets Board (CMB), is the regulatory and supervisory authority for the capital markets. The Borsa Istanbul (BIST) is the only securities exchange and Merkezi Kayit Kurulusu (MKK) functions as Turkey's central securities depository (CSD) and registrar of dematerialized corporate securities. Capital market trades conducted over BIST are settled and cleared by the Istanbul Settlement and Custody Bank (Takasbank). Takasbank is licensed as a non-deposit-taking sector-specific investment bank. Takasbank is regulated and supervised by the BRSA. However, the CMB and CBRT also oversee Takasbank: as CCP and TR by the CMB; and as SSS by the CBRT.

4.3.1 Cooperation between regulators in Turkey

Law 6493, outlines the cooperative framework between CBRT and BRSA. Article 15, states that the BRSA should consult the CBRT while issuing a license to a payment institution or an electronic money institution, and inform the CBRT while revoking and terminating a license in terms of Articles 16 and 17. Article 26, provides for

²⁸ Banking Law No. 5411, https://www.bddk.org.tr/WebSitesi/english/Legislation/14905banking_law_december_2013.pdf. The Banks Act No.4389 has since been abolished.

²⁹ <https://www.bddk.org.tr/WebSitesi/english/Legislation/8917bankcardsandcreditcardslaw.pdf>

cooperation and information exchange between BRSA and CBRT with regard to the enforcement of the provisions of the law in so far as they are related with the payment and electronic money institutions. Further, Article 29 allows for joint supervision and audit of card companies by the BRSA and CBRT. In addition, Article 30 allows for information exchange between the BRSA and CBRT for implementing and enforcing the provisions of the law.

The legal basis for CMB's cooperation with the CBRT is derived from Article 128 (1-ç) of the CML that gives the duty to exchange information and cooperate in any manner with other financial regulatory and supervisory institutions in order to ensure financial stability or fulfil the requirements of national or international legislation to the CMB.

The CBRT, BRSA and CMB discuss and cooperate with other authorities regarding specific issues under the Financial Stability Committee (FSC). This committee is primarily responsible for financial stability and systemic risk. The FSC provides the coordination for the cooperation among relevant authorities with respect to FMI's in Turkey. The committee comprises of the heads of the BRSA, the CMB, the CBRT, Saving Deposits Insurance Fund, and the Undersecretariat of Treasury. The chair of the FSC is the Deputy Prime Minister of Turkey Responsible for Economy.

4.4 Comparison of the two models

In both the models, the respective regulatory domains are clearly defined with overlaps if any which are recognized and addressed in the relevant applicable statutes.

In the case of Model 1 (Germany and Luxembourg), the central bank is responsible for the oversight of payment and settlement systems including payment instruments, while a different regulator is responsible for authorizing, licensing and supervising payment service providers. Model 2 in Turkey is a variation of Model 1 in so far as regulating payment instruments is concerned with the banking supervisor being entrusted with this task and not the CBRT. However, in both the models, the respective central banks are the overseers for payment and settlement systems.

In case of both the models, the aspect of cooperation between the regulators is clearly emphasized and laid down in the respective laws. The case studies above provide examples of cooperation ranging from: (i) consultations at the time of licensing a payment service provider; (ii) conducting joint on-site inspections; (iii) sharing of information and data; and (iv) coordination of crisis management.

5 Conclusion

The NPS being a critical pillar of financial stability, the safe and efficient functioning of the NPS is very critical for continued financial stability. The complexity of the NPS has evolved over the years, especially with the advent of technology and the entry of non-bank providers of payment services and operations. This has created newer

challenges in carrying out oversight over the NPS by the relevant authorities including the central banks.

These new developments have brought an urgency to review, revisit, and reorient existing policy objectives. The large-scale shifts being witnessed in the area of retail payments whether in terms of newer non-bank providers, payment methods or innovative technologies, outline the need for increasing competition while ensuring consumer protection and financial stability. Where certain aspects of the NPS are regulated and overseen by a regulator other than the central bank, the issue gains more criticality and underscores the need to quickly evolve and operationalise a sound framework of cooperation between the authorities for the safe, sound and efficient functioning of the NPS in a country.

It is evident from the above, that there is no one size fits all approach while carrying out oversight over the NPS in a country. The powers and responsibilities of the relevant authorities over different facets of the NPS are a function of their respective mandates. While, this being so it is essential that the mandates of the individual regulators involved are fully captured and reflected in a framework of cooperation and collaboration in developing a sound and risk-based supervisory, regulatory and oversight framework for the NPS as a whole, in line with international best practices and standards. Going forward, where a multiplicity of regulators is involved, it is a best practice to have a formal arrangement in the form of a Memorandum of Understanding between them to avoid regulatory overburden and avoid and minimise any regulatory arbitrage. Adopting such an approach, would enable the NPS to flourish by embracing innovations and in providing faster, efficient, safer and cheaper payment services to the end-users in the economy.

IDRBT Journal of Banking Technology

Volume 2 | Number 1 | Jan-Jun 2018

Editorial

Dr. A. S. Ramasastrri

Research Articles

1. From prediction to anticipation of cyber attacks 01
Michael Weiss
2. Decision-making under uncertainty 12
Monika, Hao Quan, Dipti Srinivasan
3. Machine learning in computer security 23
V. Rao Vemuri

Practitioners' Perspective

4. Innovating a seamless customer experience 41
Bipin Sahni
5. National payment system – overview of regulatory mandates 47
Gynedi Srinivas, Harish Natarajan