CAPCO

JOURNAL

THE CAPCO INSTITUTE JOURNAL OF FINANCIAL TRANSFORMATION

SECURITY

Cyber security ontologies supporting cyber-collisions to produce actionable information

MANUEL BENTO | LUIS VILARES DA SILVA
MARIANA SILVA

DIGITIZATION

#47
04.2018

JOURNAL

THE CAPCO INSTITUTE JOURNAL OF FINANCIAL TRANSFORMATION

RECIPIENT OF THE APEX AWARD FOR PUBLICATION EXCELLENCE

CONTENTS

ORGANIZATION

07 Implications of robotics and AI on organizational design
Patrick Hunger, CEO, Saxo Bank (Schweiz) AG
Rudolf Bergström, Principal Consultant, Capco
Gilles Ermont, Managing Principal, Capco

15 The car as a point of sale and the role of automotive banks in the future mobility
Zhe Hu, Associate Consultant, Capco
Grigory Stolyarov, Senior Consultant, Capco
Ludolf von Maltzan, Consultant, Capco

25 Fintech and the banking bandwagon
Sinziana Bunea, University of Pennsylvania
Benjamin Kogan, Development Manager, FinTxt Ltd.
Arndt-Gerrit Kund, Lecturer for Financial Institutions, University of Cologne
David Stolin, Professor of Finance, Toulouse Business School, University of Toulouse

35 Can blockchain make trade finance more inclusive?
Alisa DiCaprio, Head of Research, R3
Benjamin Jessel, Fintech Advisor to Capco

45 The aftermath of money market fund reform
Jakob Wilhelmus, Associate Director, International Finance and Macroeconomics team, Milken Institute
Jonathon Adams-Kane, Research Economist, International Finance and Macroeconomics team, Milken Institute

51 Costs and benefits of building faster payment systems: The U.K. experience
Claire Greene, Payments Risk Expert, Federal Reserve Bank of Atlanta
Marc Rysman, Professor of Economics, Boston University
Scott Schuh, Associate Professor of Economics, West Virginia University
Oz Shy, Author, How to price: a guide to pricing techniques and yield management

67 Household deformation trumps demand management policy in the 21st century
Iordanis Karagiannidis, Associate Professor of Finance, The Tommy and Victoria Baker School of Business, The Citadel
D. Sykes Wilford, Hipp Chair Professor of Business and Finance, The Tommy and Victoria Baker School of Business, The Citadel

CURRENCY

81 Security and identity challenges in cryptotechnologies
José Vicente, Chairman of the Euro Banking Association's Cryptotechnologies Working Group
Thomas Egner, Secretary General, Euro Banking Association (EBA), on behalf of the working group

89 Economic simulation of cryptocurrencies
Michael R. Mainelli, Chairman, Z/Yen Group, UK and Emeritus Professor of Commerce, Gresham College
Matthew Leitch, Z/Yen Group
Dionysios Demetis, Lecturer in Management Systems, Hull University Business School

101 Narrow banks and fiat-backed digital coins
Alexander Lipton, Connection Science Fellow, Massachusetts Institute of Technology (MIT), and CEO, Stronghold Labs
Alex P. Pentland, Toshiba Professor of Media Arts and Sciences, MIT
Thomas Hardjono, Technical Director, MIT Trust::Data Consortium, MIT

117 Quantitative investing and the limits of (deep) learning from financial data
J. B. Heaton, Managing Member, Conjecture LLC

SECURITY

125 Cyber security ontologies supporting cyber-collisions to produce actionable information
Manuel Bento, Euronext Group Chief Information Security Officer, Director, Euronext Technologies
Luis Vilares da Silva, Governance, Risk and Compliance Specialist, Euronext Technologies, CISSP
Mariana Silva, Information Security Specialist, Euronext Technologies

133 Digital ID and AML/CDD/KYC utilities for financial inclusion, integrity and competition
Dirk A. Zetzsche, Professor of Law, ADA Chair in Financial Law (Inclusive Finance), Faculty of Law, Economics and Finance, University of Luxembourg, and Director, Centre for Business and Corporate Law, Heinrich-Heine-University, Düsseldorf, Germany
Douglas W. Arner, Kerry Holdings Professor in Law, University of Hong Kong
Ross P. Buckley, King & Wood Mallesons Chair of International Financial Law, Scientia Professor, and Member, Centre for Law, Markets and Regulation, UNSW Sydney

143 Digital identity: The foundation for trusted transactions in financial services
Kaelyn Lowmaster, Principal Analyst, One World Identity
Neil Hughes, Vice President and Editor-in-Chief, One World Identity
Benjamin Jessel, Fintech Advisor to Capco

155 Setting a standard path forward for KYC
Robert Christie, Principal Consultant, Capco

165 E-residency: The next evolution of digital identity
Clare Sullivan, Visiting Professor, Law Center and Fellow, Center for National Security and the Law, Georgetown University, Washington D.C.

171 The future of regulatory management: From static compliance reporting to dynamic interface capabilities
Åke Freij, Managing Principal, Capco

Cyber security ontologies supporting cyber-collisions to produce actionable information

MANUEL BENTO | Euronext Group Chief Information Security Officer, Director, Euronext Technologies
LUIS VILARES DA SILVA | Governance, Risk and Compliance Specialist, Euronext Technologies, CISSP
MARIANA SILVA | Information Security Specialist, Euronext Technologies

ABSTRACT

In this article, we bring to the attention of key players in the financial services sector the continuous cyber security events affecting the industry globally. We also consider a possible solution for mitigation of such events through the introduction of new processes and technologies. Using a computational logic-based language by machine learning processes through artificial intelligence algorithms will improve prediction of unwanted cyber events via early warning alerts. A cyber-collision system concept is described by adjoining cyber security ontologies, security analyst experience, machine learning, and information sharing to protect the financial services sector.

1. INTRODUCTION

The current cyber security paradigm identifies well-defined activities, promoting multiple and increasing layers of defense in a reactive mode. We know from experience that "intrusion prevention systems" (IPS) do fail with an excessive number of false positives, which must be minimized through the tuning of the faulty detection signatures.

The current approach has failed in preventing ransomware attacks, phishing attacks, and the old social engineering attacks that continue to cause major problems for global corporations. Case studies of such attacks are well-known to most observers of the world of business. They include the Equifax debacle, the Yahoo bombshell, the WannaCry ransomware attack, the NotPetya malware outbreak at Maersk, Fedex, and Rosneft, among others.

The information security community expects 2018 to be not too different from the recent past. If anything, the proliferation of uncontrolled systems connected to the internet through the Internet of Things (IoT) could make matters worse.

To respond to this worsening situation, firms must commit resources to attain knowledge from beyond their firewalls, so that they can predict what attacks will become likely and decide where to invest. Consequently, facing the cyber enemies outside of the "comfort zone," and being able to prevent their attacks, or, even better, being able to avoid them, is paramount to cyber security. Battle-tested machine learning processes will, with the help of specialized security professionals, improve predictive analyses. Together with proactive financial business sector involvement, they could promote a cyber-collision system to handle positive cyber attack alerts from multiple sources using a centralized cyber attack index system employed for cyber defense support.

2. CURRENT REALITY – THE KNOWN WRONGS

Currently, enterprises, organizations, and governments have major difficulties in detecting information security attacks or even reacting to them when detected, especially when they affect multiple systems in many disperse geographical locations.

Knowledge of information security attack vectors is paramount to information security analysts, as cyber attackers have the capability to learn about any online business resource and evaluate its interconnectivity with systems within the same business sector. This is especially the case among financial services firms, who are not very open about sharing data to support peers (for example on failed access attempts). Nonetheless, willingness to start sharing and even creating a common approach to cybersecurity is helping the financial services sector with, at this stage, dealing and handling cyber response to known cyber attacks (e.g., FI-ISAC and FS-ISAC: Financial Services – Information Sharing and Analysis Center, mailing lists).

Recent cyber attacks have demonstrated that only after a process of public awareness of their real impact do companies call their security analysts to report on the cyber-resilience controls in place; typically with difficult to understand dashboards. Consequently, sharing security alerts at an early stage could improve the analysis process and minimize the impact of a cyber attack.

A data breach is the most disruptive cyber attack security incident. Consequently, firms should systematically identify and sanitize key lessons from cyber events in order to advance resilience capabilities.

Per Verizon's "2017 data breach investigations report," an incident is a security event that compromises the integrity, confidentiality, or availability of an information asset. On the other hand, a breach is an incident that results in the confirmed disclosure (not just potential exposure) of data to an unauthorized party.

> "Security specialists can perform deeper investigations to improve machine learning models and help transition from a reactive approach to a proactive one, and eventually to a predictive approach."

Data breaches frequently cause major reputational damage. They create uncertainty, reduce consumer trust, and can harm the firm's competitive edge in the markets.

Technology has become a vital part of people's daily lives and is crucial for societies to grow. As a result, security awareness must be among the major investments in information security by the financial services industry.

Technology may fail, however, early warnings through information sharing could create actionable information to mitigate cyber threats. It is common nowadays to see social engineering skills (e.g., through spear phishing) being used to explore weaknesses in order to obtain access to companies' valuable information.

On the other hand, the cyber threat landscape will be mostly about new applications leveraging new business models, based on new technologies, and making use of new infrastructure models (xaaS). Further, new rapid application development technologies, associated with complex algorithms to combine disparate data sources, including the merging of internal business specific data with external "big data" analysis to expand sector/oriented workflows and processes (such as using hybrid cloud systems), will make cyber security and data protection difficult disciplines to handle within one organization. A very real example can be found in poorly designed applications being dependent on built-in OS kernel libraries with obsolete algorithms for encryption/decription that cannot always be removed due to the loss of source code. Consequently, security measures need to be integrated at a later stage. A typical example is the undocumented use of "forked" Linux kernel libraries by Java applications.

Analysis of past data breaches alone, while helping us to understand common weaknesses, is not enough to stem the tide. What is needed is predictive analysis based on massive security event datasets to identify trends, predict impacts, and propose mitigating actions. Such analysis will be based on classification mechanisms that are underpinned by cyber security ontologies and feed AI algorithms allowing the identification of cyber-collisions, such as prediction of cyber attacks through clear alerts based upon experience or knowledge.

Through identification and study of past data breaches, we will be able to establish a well-defined baseline of behavioral and system activity against which we can apply machine learning techniques. Big data analysis, helped by cyber security ontologies and based on datasets of past events, enable algorithms to be trained to learn trends and impacts and propose mitigating solutions and consequently stop cyber attacks through learning collision mechanisms.

Table 1: Biggest data breaches of the 21st Century (U.S.£ million)

2017	Equifax	143
2016	Adult Friend Finder	412.2
2015	Anthem	78.8
2014	eBay	145
	JP Morgan Chase	76
	Home Depot	56
2013	Yahoo	3000
	Target Stores	110
	Adobe	38
2012	U.S. Office of Personnel Management (OPM)	22
2011	Sony's PlayStation Network	77
	RSA Security	40
2008	Heartland Payment Systems	134
2006	TJX Companies Inc.	94

Source: Armerding (2018)

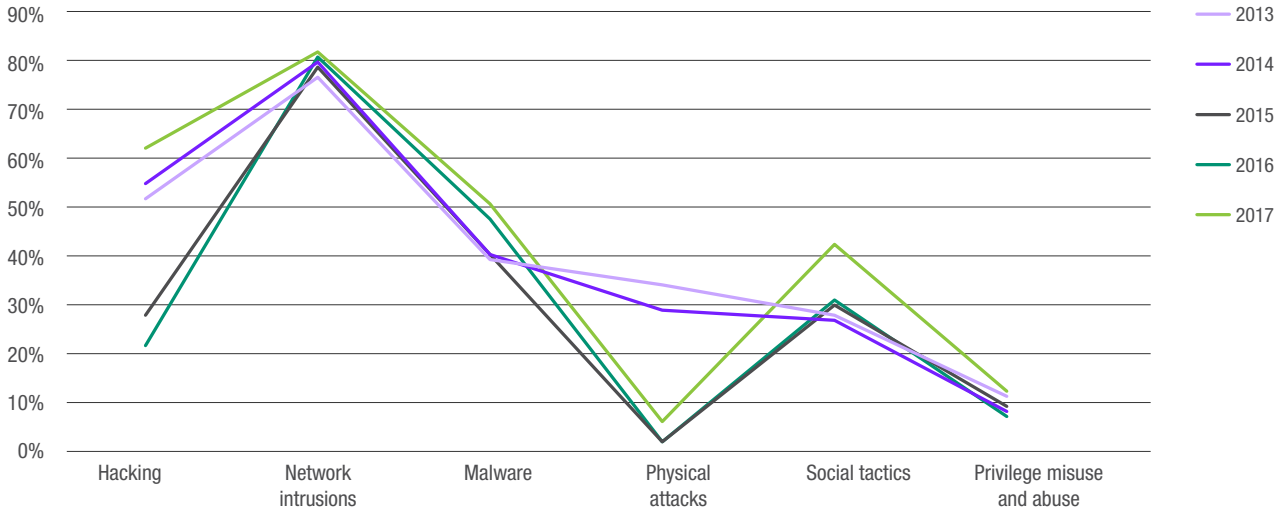
Table 1 depicts the biggest data breaches from this century.

In the financial services industry, the most prevalent types of attacks are: "denial of service attacks" (DDoS), web application attacks, and payment card skimming. Cyber attack methods are presented in Figure 1, showing clearly that the weakest link is still related to network intrusions and active hacking.

3. CHALLENGES OF DATA SCIENCE APPLIED TO CYBER SECURITY

At present, there are some challenges in applying data science to cyber security [Kolman (2014)]: data normalization, anomaly detection, high cost of errors, the data required is not public, data evaluation is difficult, semantic gap (with the difficulty to describe the information), lack of expertise, and an adversarial environment were permanently changing datasets imply a learning period for humans to adapt. Consequently, the use of data science in a cyber security context can be considered to require very specialized human skills and a large commitment of effort.

Figure 1: Verizon data breach investigations reports



As we can see from Figure 1, targeted attacks on computer networks are still the prevalent method of cyber attacks and the need for tools to support analysts to effectively hunt malicious activity within one's perimeter has increased dramatically. The existing security event information management (SIEM) systems help analysts through pre-defined schemas to identify logs or events that might be of interest within the aggregated logs. As described before, these systems are faced with some hard problems, namely the diverse schema of information sources that imply extra layers of technology – connectors – to properly incorporate information and security analyst expertise to help correlate the new source with existing log events.

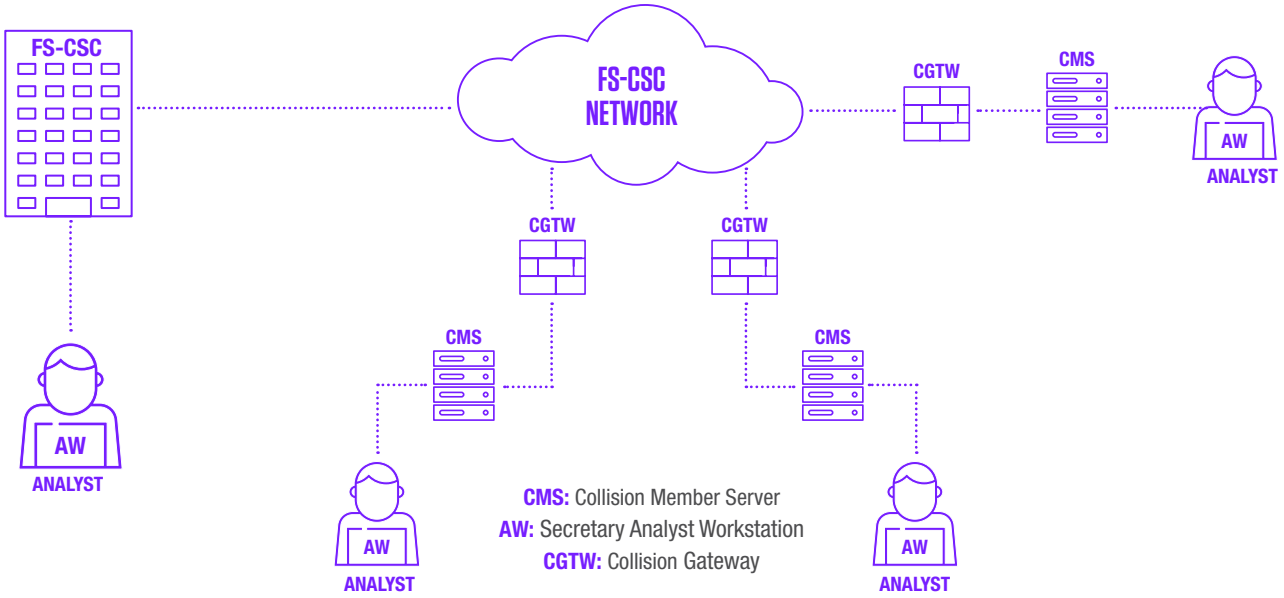
By adopting a unified way to support information integration and cyber situational awareness in cyber security systems, security analysts will be able to get better visibility on threats. As such, adopting a cyber security ontology will make available to security analysts, intrinsic properties, that with some "assumptions" (like "false positive multi location access" – for example a login at an office in London while the same login account is used within the same timeframe at HQ in Paris using mobile access to email), will overcome the workload limits, making it possible to analytically process the huge and constantly changing event datasets. The analysis process will, therefore, entail the adoption of the right learning and processing models. The process of further assigning specific metadata (i.e., required ontology labels) or attributes to identify appropriate analytical sources aiming to reduce the specialized security knowledge required for an analyst to be effective at understanding and evaluating a threat will also be incorporated.

4. MACHINE LEARNING HELPING CYBER SECURITY

Choosing ontologies as the unifying support to information integration, sharing generated cyber attack information through an event repository with heterogeneous data and knowledge schemas, can contribute to producing actionable information to feed a hypothetical cyber-collision mechanism. This proposed cyber-collision mechanism is fed by multivariate systems (systems that use incremental learning algorithms such as pattern recognition, data mining, or fuzzy logic), so attackers do not become familiar with a specific system. It is supported by security experts' cooperation to improve the models used by such systems. Machine learning systems are, therefore, the final element with highly integrated functions of high-performance analytics for predictive analysis and forecasting of cyber attacks.

Adopting the "unified cyber security ontology" (UCO) [Syed et al. (2016)], security specialists can perform deeper investigations to improve the machine learning models and help transition from a reactive approach to a proactive one, and eventually to a predictive approach. As we know from machine learning theory, experience increases task performance [p(t,e)>p(t)], which is the motivation for the

Figure 2: High level cyber-collision network alerting system



Source: Luis Vilares da Silva / Sofia Silva

cyber-collision mechanism to process the actionable information and predict cyber attacks. At the same time, if the cyber attack materializes, the model can improve situational awareness, which is of extreme importance when performing digital forensic investigations or defining a tactical cyber defense strategy.

The idea behind UCO is paramount for moving the cyber security paradigm from event correlation to an extensive cyber situational awareness in systems like the one proposed here as an example.

UCO ontology was mapped to many existing cyber security ontologies and concepts in the Linked Open Data cloud [Bizer et al. (2009)] and is an extension of the "intrusion detection system ontology." The UCO authors describe the ontology as the core for a cyber security "linked open data" (LOD) cloud as it represents the semantic version of the event exchange standard STIX, extended with other cyber security related standards, such as "common vulnerability and exposures" (CVE), "common attack pattern enumeration and classification" (CAPEC), etc.

The purpose of UCO is to serve as the core for the cyber security domain and its capacity to be extended serves the purpose of structuring event information to be shared, integrated, and reused within applications in the financial realm.

"Resource description framework" (RDF) and languages such as "ontology web language" (OWL), are used to represent entities through a set of abstract objects or concepts rather than only some strings of words. Both languages expose structures that represent information that is not only machine readable, but also machine understandable, and therefore facilitate the sharing of information from heterogeneous sources.

5. CYBER SECURITY ACTIONABLE COLLABORATION MODEL IN THE FINANCE SECTOR

Our hope is that the financial services industry can join efforts, through a consortium type of organization, to create a collaborative platform to properly predict cyber attacks through preemptive positive alerts producing actionable information enabling "early warnings" for the members of the consortium.

With such an ambition, the first set of questions arise: who will be in the consortium? Who should lead the process? What type of data should be shared? Where will the data be located? How is the communication processed in a synchronous way? Who is paying for it? Finally, which type of system are we proposing?

5.1 The consortium

To be broad but effective, the proposed consortium would include both frontline membership and consulting membership. The first type of membership in the finance sector would include trading companies and brokers, banks, credit card companies, and insurance companies. For the latter membership, international financial institutes and international law enforcement organizations are the interested parties to support the containment and undertaking further investigative assessments for proper incident handling.

The consortium should have a governing body, management structure, and governance model agreed by all stakeholders. This consortium could be described as the Financial System Cyber Security Center (FS-CSC).

5.2 Sharing data

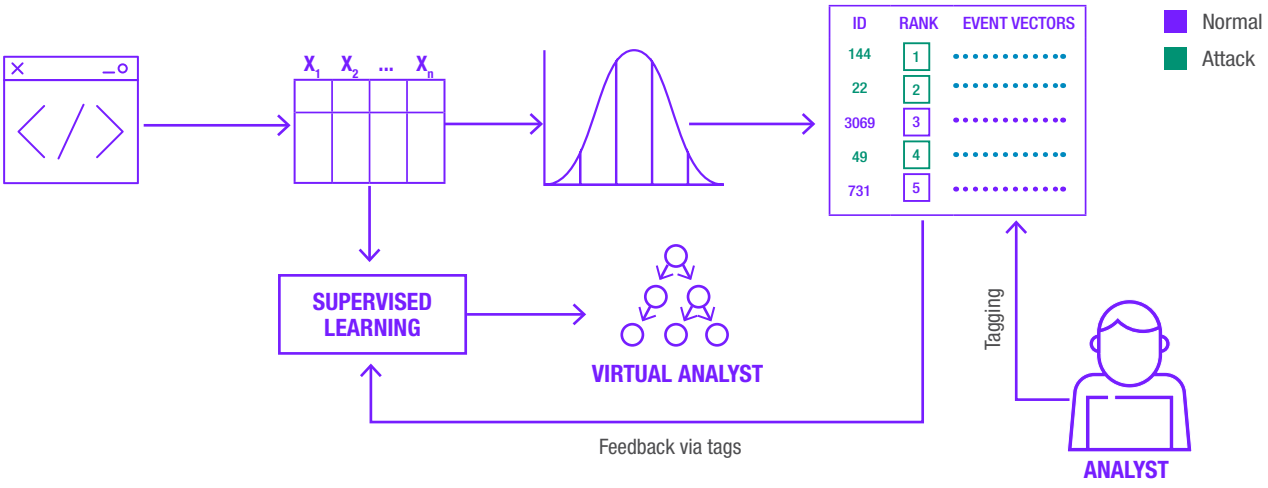
The most important element within collaborative platforms, as in any information system, is the data they process. For the proposed system and having in mind the reservations companies could have in sharing their data with competitors within the same sector, a specific and well-defined set of data would be considered. Using technology, such as intrusion prevention systems to analyze and convert scanned traffic into cyber security ontologies with features/attributes such as AttackPattern, BaseGroup, CCE, CVE, CVSSScoreType, ConfidenceType, Exploit, Malware, Origin, Attack, Attacker, Campaign, Consequence, etc... (i.e., an agreed subset of the UCO ontology), and centralize the output result in a large-scale logging system would, therefore, be a type of solution to share common cyber defense information.

5.3 Localization, network, and funding

The physical location of the event data is of major importance, with the financial services industry being heavily regulated. For a system that will share a security advantage with its members, an independent, stable, and regulated country should be chosen to host the infrastructure supporting it. Countries such as Switzerland, Luxembourg, Portugal, or Ireland would have the technology, stability, and regulations required to host an independent, advanced, data center such as this. They are also at a much lower risk of facing terrorist attacks.

For the communication to be secure, while effective, a private network should be created and configured to interconnect the security departments of the finance sector members who would also pay on a pro rata basis (whichever order of magnitude is used to classify the members).

Figure 3: Prediction system



Source: Kalyan Veeramachaneni/MIT CSAIL

5.4 System components – high level approach

From a topology perspective, the components are familiar to any InfoSec team within the industry. As depicted in Figure 2, the system has the normal business exposure to the internet, but there is a new routing mechanism for preemptive alerting of cyber attacks, which are described as Cyber-collision Gateways.

In the proposed system, each member's analyst will feed the system with tagged ontologies, allowing the learning process to improve the analysis of the traffic and, therefore, enhancing the detection of positive cyber attacks' alerts.

Ideally, the learning process will improve the analysis in a way that a virtual analyst will replace the member's analysts; a system similar to the virtual artificial intelligence analyst developed by the Computer Science and Artificial Intelligence Lab and the company PatternEx (Figure 3) that reduces false positives by factor of 5 [Connor-Simons (2016)].

Furthermore, the event database created is available to complement the alert mechanism with a search for actionable information through a hidden-hit mechanism.

The hidden-hit mechanism is a process that informs the owner of the information when his data was hit by a search and by whom. This process will allow the owner to decide if his information can be shared immediately or will trigger another process of peer communication between the searching actor and the information owner. This process is key for reporting purposes and to create security dashboards.

The key factor with this approach is that, on one hand, members do share information related to their internet traffic without sharing business information and, therefore, competitors will not take business advantage, and on the other hand, all members have access to more information with extra relevance.

6. CONCLUSION

Machine learning in cyber security will increasingly replace the current paradigm where reactive mechanisms protect our systems, but we continue to be vulnerable as the recent cyber attacks demonstrate.

Proper risk management practices will go beyond reactive controls and include proactive protection against unwanted future cyber events. The proposed approach for the financial services industry includes proper sharing of information related to internet traffic and, therefore, improve the defense perimeter. Having a common mechanism for improved alert on cyber attacks will accelerate cyber defense capabilities, which is also extremely important for advanced persistent threats. In the long run, this approach will save operational costs with a centralized virtualization of analysts.

References

Amerding, T., 2018, “The 17 biggest data breaches of the 21st century,” CSO Online, January 26, http://bit.ly/2ovBb24

Bizer, C., T. Heath, and T. Berners-Lee, 2009, “Linked data – the story so far,”
International Journal on Semantic Web and Information Systems 5:3, 1-22

Conner-Simons, A., 2016, “System predicts 85 percent of cyber attacks using input from human experts”,
MIT News, April 18, http://bit.ly/1SWb2OZ

CSO. “Biggest data breaches by year and accounts compromised”
https://goo.gl/KkNXLw – http://breachlevelindex.com

Kolman, Y., 2014, “Machine learning and big data in cyber security,”
http://bit.ly/2CMPzHE; http://bit.ly/2BVuoGs

Syed, Z., A. Padia, M. L. Mathews, T. Finin, and A. Joshi, 2016, “UCO: a unified cyber security ontology,”
Proceedings of the AAAI Workshop on Artificial Intelligence for Cyber Security

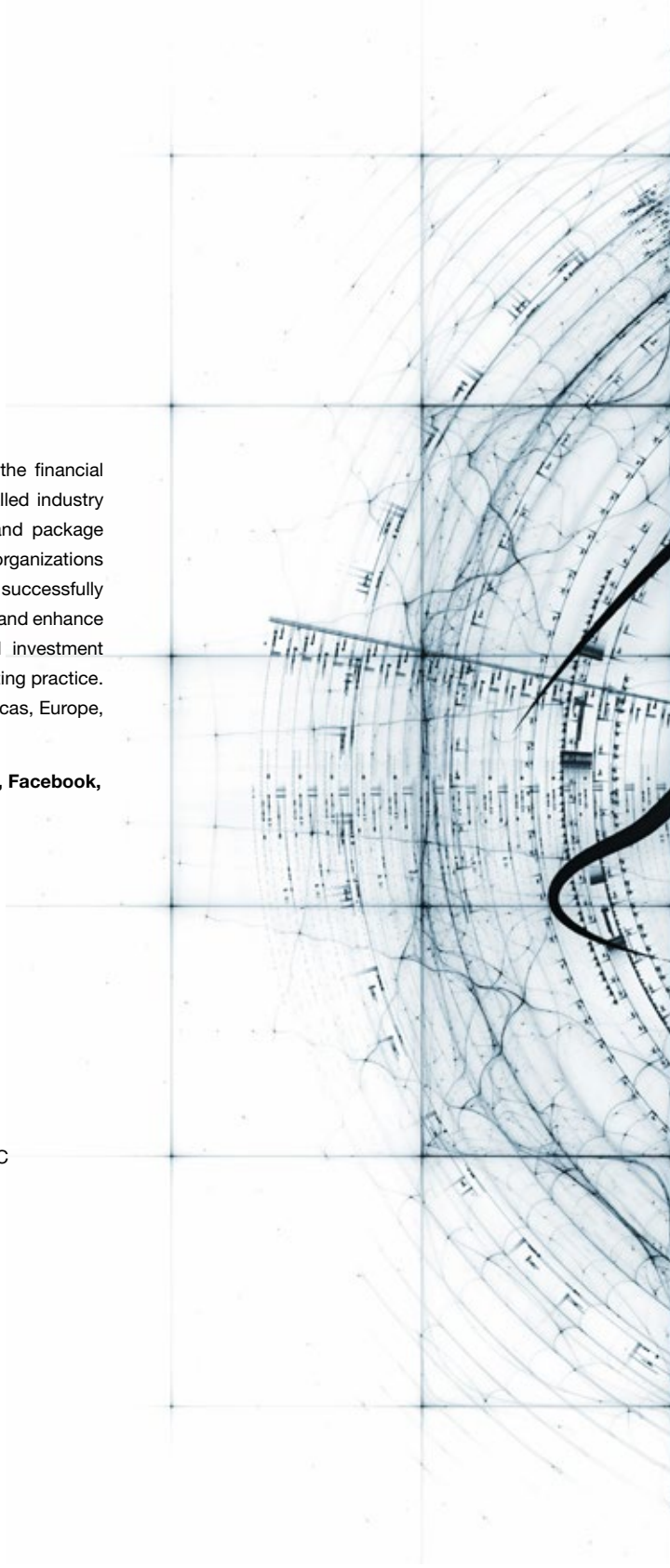
Verizon, 2017, Data breach investigations report

Verizon, 2016, Data breach investigations report

Verizon, 2015, Data breach investigations report

Verizon, 2014 Data breach investigations report

Verizon, 2013, Data breach investigations report



ABOUT CAPCO

Capco is a global technology and management consultancy dedicated to the financial services industry. Our professionals combine innovative thinking with unrivalled industry knowledge to offer our clients consulting expertise, complex technology and package integration, transformation delivery, and managed services, to move their organizations forward. Through our collaborative and efficient approach, we help our clients successfully innovate, increase revenue, manage risk and regulatory change, reduce costs, and enhance controls. We specialize primarily in banking, capital markets, wealth and investment management, and finance, risk & compliance. We also have an energy consulting practice. We serve our clients from offices in leading financial centers across the Americas, Europe, and Asia Pacific.

To learn more, visit our web site at **www.capco.com, or follow us on Twitter, Facebook, YouTube, LinkedIn and Xing.**

WORLDWIDE OFFICES

Bangalore
Bangkok
Bratislava
Brussels
Charlotte
Chicago
Dallas
Dusseldorf
Edinburgh
Frankfurt
Geneva
Hong Kong
Houston
Kuala Lumpur
London
New York
Orlando
Paris
Pune
São Paulo
Singapore
Stockholm
Toronto
Vienna
Warsaw
Washington, DC
Zurich

CAPCO.COM

CAPCO