A Text-Mining Approach to Cyberrisk Management

Cyberrisk is one of the most pervasive threats facing the global community. The World Economic Forum (WEF) has listed cybersecurity failure as one of the top-five global risk factors since 2018.[1] In 2020, 39 percent of WEF survey respondents indicated that cyberattacks were highly likely and represented a high-impact risk for industries, governments and individuals alike.[2] During the coronavirus pandemic, many individuals shifted to working from home, making them lucrative targets because of the dilution of organizational cybersecurity practices.[3] Cyberattacks grew fivefold, according to a 2020 report by the World Health Organization (WHO).[4]

Distributed denial-of-service (DDoS) attacks are among the easiest to execute due to the lack of social engineering expertise or technical know-how needed to launch them. In the second half of 2020, DDoS attacks increased by 12 percent. The attack intensity peaked at 2.3 gigabits per second (Gbps) on Amazon Web Services (AWS) and 2.5 Gbps on the Google Cloud platform. Akamai also revealed that it blocked 809 million packets that targeted its Content Delivery Network (CDN) services.[5] In the first quarter of 2020, the number of DDoS attacks tripled compared with the same quarter in 2019 and accounted for 19 percent of the total number of incidents.[6] The attack duration increased by 25 percent over that one-year period. Educational institutions such as schools and colleges suffered disproportionately due to an increase in such attacks, which aggravated the digital divide. Governmental healthcare agencies were also targeted, leading to an increase in the chaos caused by the pandemic.[7] As masses of people indulged in online entertainment while sheltering in place, hackers also targeted game servers such as EVE Online, stranding gamers for nine days.[8]

In the face of unexpected and uncertain situations such as a worldwide pandemic and increasing cyberattacks, enterprises need to be prepared and resilient. Chief experience officers' (CXOs') initial preparedness may be challenged by the generation of enormous amounts of data with varied themes over time.[9] Decision makers must process this evolving information and determine whether the enterprise's cybersecurity protocol requires

Arunabha Mukhopadhyay
Is a professor in information technology and systems at the Indian Institute of Management (IIM) (Lucknow, India). He is also the academic advocate of the ISACA® Student Group (ISG) at IIM Lucknow. He can be reached at arunabha@iiml.ac.in.

Kalpit Sharma
Is a Ph.D. student in information technology and systems at the Indian Institute of Management (Lucknow, India). His research interests are privacy and risk issues in information systems, the economics of cybersecurity, healthcare IT, IT governance, and crowd-based digital business models. He can be reached at kalpit@iiml.ac.in.

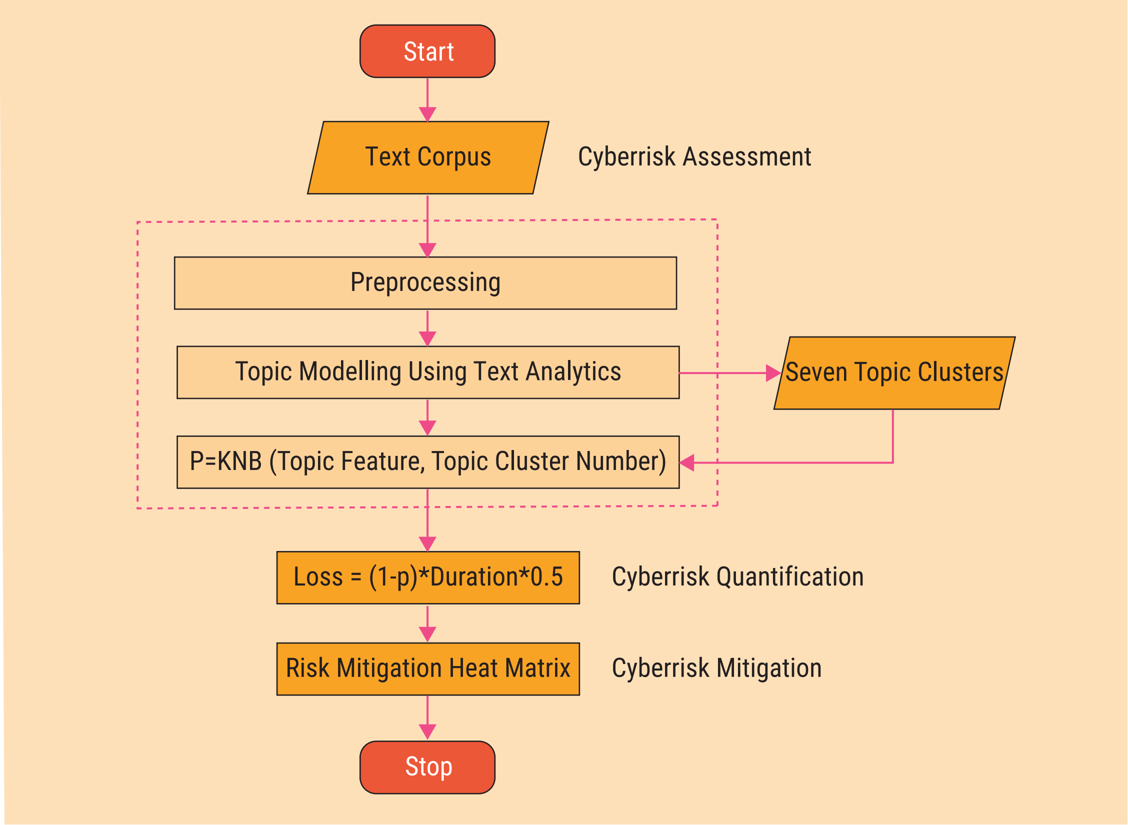
emergency revamping.[10] It is crucial to summarize and thematically analyze the various textual data generated around specific cyberattack incidents.

In this study, the text of web articles related to notable cyberattacks was input into the proposed model for cyberrisk management. The data were preprocessed into bigrams and trigrams using cybersecurity-related keywords. In terms of cyberrisk assessment, the existing web articles identified the routes and protocols exploited to launch attacks, highlighting the critical stages of the cyberrisk management process for similar cyberattacks in the future. The cost of an attack is a mix of tangible and intangible losses, and a robust response and top leadership communication are essential mitigation strategies. Successful cyberrisk management is contingent on chief technology officers (CTOs) following the critical themes extracted from the text of the articles used in this study. Failure to do so might delay proper loss prevention procedures or forgo the process altogether, culminating in extreme losses. This study extracts critical themes related to a particular cyberattack from existing web content and quantifies the potential losses if these themes are ignored. This sophisticated technical information can aid CXOs and CTOs as they tabulate the marketplace's published articles on cyberattacks to help them know what is current, what may happen and what the cost of future attacks could be.

Proposed Model

Figure 1 illustrates the proposed model, which comprises three modules: cyberrisk assessment, quantification and mitigation. The cyberrisk assessment module uses text analytic techniques to categorize the text data from web articles into three key themes related to attack route, attack cost and appropriate mitigation strategies. For each new web article, this module calculates the probability of

Figure 1—Flowchart of the Proposed Model


Start
Text Corpus
Cyberrisk Assessment
Preprocessing
Topic Modelling Using Text Analytics
Seven Topic Clusters
P=KNB (Topic Feature, Topic Cluster Number)
Loss = (1-p)*Duration*0.5
Cyberrisk Quantification
Risk Mitigation Heat Matrix
Cyberrisk Mitigation
Stop

correctly identifying the themes related to DDoS attacks using a Kernel Naive Bayes (KNB) classifier.[11] The cyberrisk quantification module calculates the expected losses by multiplying the probability calculated in the previous module by the loss incurred if the DDoS attack occurs. The cyberrisk mitigation module helps the CTO decide whether to transfer, accept or reduce the cyberrisk using technological intervention and cyberinsurance.

Data

The sample consists of eight web articles retrieved by using "DDoS" as the search term. On average, each of these documents is 25 lines long. The documents are described in terms of token types (e.g., letters, digits) and named-entity tags (e.g., person, location, organization). Tokens are predominantly letters and do not belong to any discernible entity. **Figure 2** illustrates the documents' composition.

Methodology

The methodologies used in the different modules include cyberrisk assessment, quantification and mitigation.

Cyberrisk Assessment

The cyberrisk assessment module first uses Latent Dirichlet Allocation (LDA) to divide the text data from the web articles into three key themes (and seven topic clusters) related to attack route, attack cost and appropriate mitigation strategies.[12] The seven topic clusters comprise four bigrams and three trigrams. The data sets in the bigrams and trigrams are divided in a ratio of 60:40. Next, inputs to the KNB classifier for bigrams and trigrams determine the probability of them belonging to the four topic clusters (topics 1, 2, 3 and 4) and three topic clusters (topics 5, 6 and 7), respectively. For each new web article, this module outputs the probability of correctly identifying the bigrams and trigrams related to the three key themes.[13, 14]

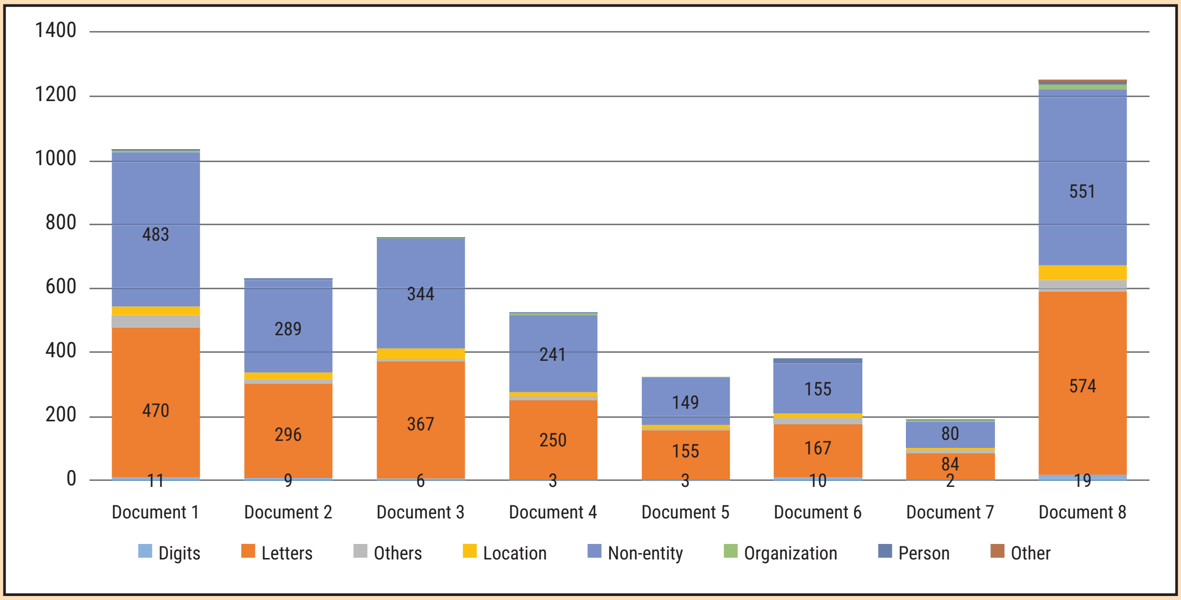
Cyberrisk Quantification

This module quantifies the expected loss based on the probability of wrongly identifying the topics, a loss of US$500,000 per hour from a DDoS attack and the hours of downtime.[15, 16, 17]

Cyberrisk Mitigation

The final module helps the CTO decide whether to reduce, accept or transfer the cyberrisk by using a

Figure 2—Text Composition



Enjoying this article?

- Read *Risk IT Framework, 2nd Edition. www.isaca.org/risk-it-f2*
- Learn more about, discuss and collaborate on risk management in ISACA's Online Forums. *https://engage.isaca.org/onlineforums*

combination of financial and technological interventions.

Figure 3 illustrates the steps in the three modules of the proposed model. MATLAB 2020b was used to analyze the data.

Results

It is helpful to understand the results related to each of the modules of cyberrisk assessment, quantification and mitigation.

Cyberrisk Assessment

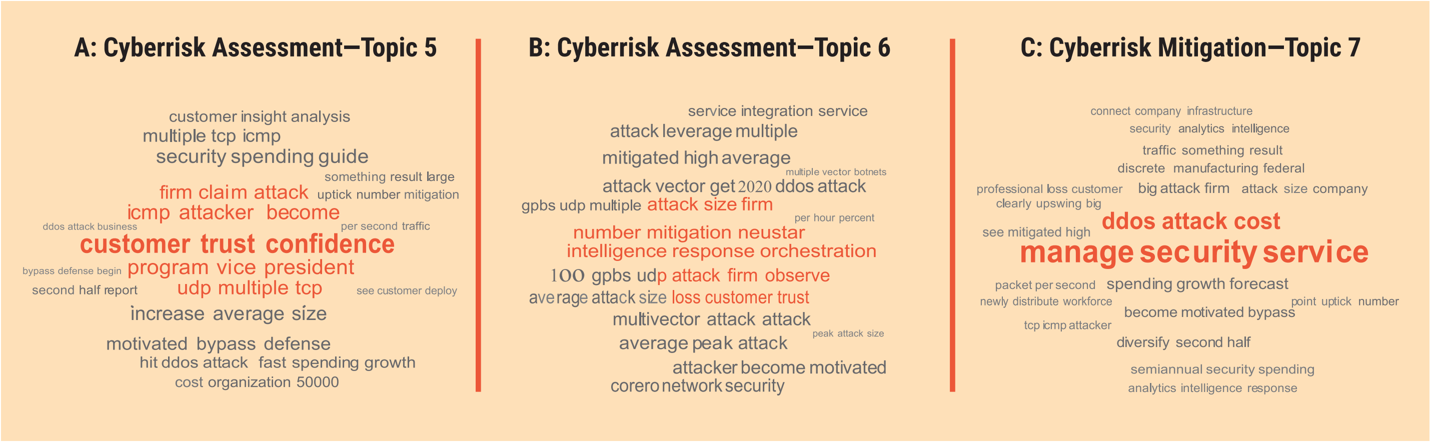
The topic modeling through LDA generates seven topic clusters: four clusters (topics 1, 2, 3 and 4) from the bigram model and three clusters (topics 5, 6 and 7) from the trigram model. **Figure 4A** depicts trigram-based topic clusters highlighting DDoS attacks' possible routes, such as exploiting Internet Control Message Protocol (ICMP), Transmission Control Protocol (TCP) and User Datagram Protocol (UDP). **Figure 4B** illustrates the trigram-based topic clusters related to losses in terms of attack cost, intensity and loss of customer trust and confidence in the enterprise's operations. **Figure 4C** shows that mitigation strategies, including the orchestration of a prompt response and top leadership communication, are necessary to allay customers' fears.

Next, a KNB classifier was applied to the data set, with different topic probabilities as the feature vector. **Figure 5** illustrates that the proposed model was able to classify the three critical themes (attack routes, attack cost and attack mitigation) using the bigram and trigram in 89 percent and 90 percent of cases, respectively. The model correctly classified attacks in 70 out of 78 cases in the bigram model and in 47 out

Figure 3—Steps in the Proposed Model

Cyberrisk Assessment	
Step 1	Preprocess the text
Step 2	Generate topic clusters using a topic modeling algorithm (LDA)
Step 3	Calculate the accuracy (p) of detecting the topic cluster correctly by using the classifier: p = KNB (topic_features, topic cluster label)
Cyberrisk Quantification	
Step 4	Calculate the expected loss E(L): $E(L) = (1 – p) * 0.5 * (\text{duration})$
Cyberrisk Mitigation	
Step 5	Propose mitigation strategies (e.g., technology, cyberinsurance): If (1 – p) > 0.18 AND E(L) > US$1.75 million, then implement technology + cyberinsurance

Figure 4—Trigrams Indicating Cyberrisk Assessment, Quantification and Mitigation



A: Cyberrisk Assessment—Topic 5

B: Cyberrisk Assessment—Topic 6

C: Cyberrisk Mitigation—Topic 7

of 52 cases in the trigram model. Topic 4 was classified most accurately (96 percent), and topic 2 was classified least accurately (67 percent).

Cyberrisk Quantification

Figure 6 tabulates the expected losses for each topic cluster. Misinterpretation of topic 2 incurs the highest expected loss, at US$3.17 million.

Cyberrisk Mitigation

Figure 7 depicts a heat matrix that situates the different attack classes in terms of risk × severity. Topic 2 is in the high-risk/high-expected-loss quadrant, while the other topics are in the low-risk/low-expected-loss quadrant. Enterprises at risk of misinterpreting or delaying information processing should implement mitigation strategies. The CTO should implement a highly accurate threat intelligence system with more comprehensive data sources and better text mining algorithms. Human tagging of topic clusters can also improve the accuracy of the classifier. A better understanding of the evolving cyberattack landscape can increase the probability of correctly detecting attacks and reduce losses due to delayed or wrong response orchestration. If an enterprise fails to identify topics in the low-low quadrant, it can subscribe to cyberinsurance, owing to the low-risk premium. Otherwise, enterprises can use a combination of technological intervention and cyberinsurance policies to move into the low-low quadrant.[18, 19, 20, 21]

Conclusion

This study discusses a programmatic algorithm for CTOs to fight cyberattacks by analyzing the text corpus related to cyberattacks in the industry. In 90 percent of cases, this study's proposed classifier

Figure 5—Confusion Matrix for Testing Data Set

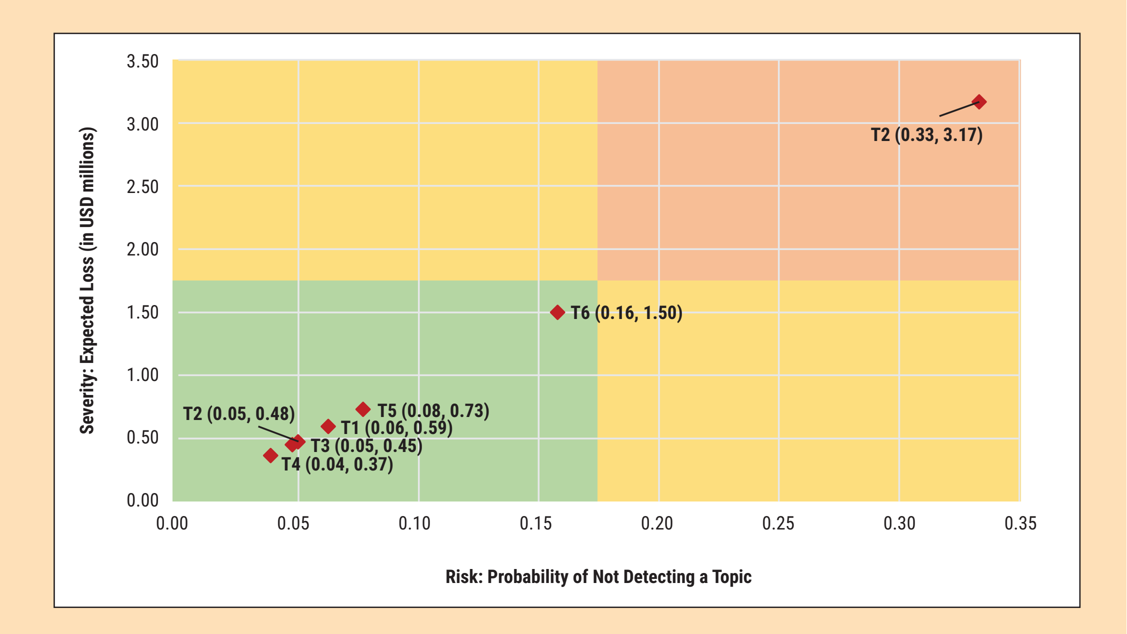
Topic Label	Bigram Topic Clusters				Total (N = 78)	Probability of Detecting Topic (p)
	T1	T2	T3	T4		
Topic 1	15	1	0	0	16	0.93
Topic 2	1	10	2	2	15	0.67
Topic 3	0	1	20	0	21	0.95
Topic 4	0	1	0	25	26	0.96
	Trigram Topic Clusters				Total (N = 52)	
	T5	T6	T7			
Topic 5	12	0	1		13	0.92
Topic 6	1	16	2		19	0.84
Topic 7	1	0	19		20	0.95

T^i = i^{th} topic cluster, where i = 1, 2, ..., 7

Figure 6—Expected Loss per Hour for Each Attack Class

Topic Label	Probability of Not Detecting a Topic (1 – p)	Duration of Attack (Hours)	Expected Loss per Hour (US$ millions) E(L) = (1 – p) * 0.5 * (duration)
T1	0.07	19	0.59
T2	0.33	19	3.17
T3	0.05	19	0.45
T4	0.04	19	0.37
T5	0.08	19	0.73
T6	0.16	19	1.50
T7	0.05	19	0.48

Figure 7—Risk-Severity Heat Matrix



correctly detected the topic from the text of selected web articles. Subsequently, it can help the CTO to estimate expected losses and determine mitigation strategies such as transferring, accepting or reducing the cyberrisk using technological and financial interventions.

Endnotes

1. World Economic Forum (WEF), *Global Risks Report 2021, 16th Edition*, 19 January 2021, *https://www.weforum.org/reports/the-global-risks-report-2021*
2. *Ibid.*
3. Interpol, "Interpol Report Shows Alarming Rate of Cyberattacks During COVID-19," 4 August 2020, *https://www.interpol.int/en/News-and-Events/News/2020/INTERPOL-report-shows-alarming-rate-of-cyberattacks-during-COVID-19*
4. World Health Organization (WHO), "WHO Reports Fivefold Increase in Cyber Attacks, Urges Vigilance," 23 April 2020, *https://www.who.int/news/item/23-04-2020-who-reports-fivefold-increase-in-cyber-attacks-urges-vigilance*
5. Hope, A.; "DDoS Attacks Increased Rapidly During the COVID-19 Pandemic as Hackers Exploited New Tools and Techniques," *CPO Magazine*, 29 January 2021, *https://www.cpomagazine.com/cyber-security/ddos-attacks-increased-rapidly-during-the-covid-19-pandemic-as-hackers-exploited-new-tools-and-techniques/*
6. Kaspersky, "DDoS During the COVID-19 Pandemic: Attacks on Educational and Municipal Websites Tripled in Q1 2020," 6 May 2020, *https://usa.kaspersky.com/about/press-releases/2020_ddos-during-the-covid-19-pandemic-attacks-on-educational-and-municipal-websites*
7. Nichols, S.; "US Health and Human Services Targeted by DDoS Scum at Just the Time It's Needed to Be Up and Running," *The Register*, 16 March 2020, *https://www.theregister.com/2020/03/16/hhs_reports_cyberattack/*
8. Fenlon, W.; S. Messner; "A DDoS Attack Has Kept Many EVE Online Players Offline for 9 Days With No End in Sight," *PCGamer*, 4 February 2020, *https://www.pcgamer.com/a-ddos-attack-has-kept-many-eve-online-players-offline-for-9-days-with-no-end-in-sight/*
9. Lohrmann, D.; "2020: The Year the COVID-19 Crisis Brought a Cyber Pandemic," Government Technology, 12 December 2020, *https://www.govtech.com/blogs/lohrmann-on-cybersecurity/2020-the-year-the-covid-19-crisis-brought-a-cyber-pandemic.html*

10 J. P. Morgan, "Developing a Culture of Cyber Preparedness," 29 October 2019, *https://www.jpmorgan.com/commercial-banking/insights/developing-culture-cyber-preparedness*

11 Hastie, T.; R. Tibshirani; J. Friedman; *The Elements of Statistical Learning*, Springer-Verlag, USA, 2009

12 Blei, D. M.; A. Y. Ng; M. I. Jordan; "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, vol. 3, 2003, p. 993–1022

13 *Ibid.*

14 *Op cit* Hastie

15 Sharma, K.; A. Mukhopadhyay; "Cyber Risk Assessment and Mitigation Using Logit and Probit Models for DDoS Attacks," 26th Americas Conference on Information Systems, 2020

16 Sharma, K.; A. Mukhopadhyay; "Assessing the Risk of Cyberattacks in the Online Gaming Industry: A Data Mining Approach," *ISACA® Journal*, vol. 2, 2020, *https://www.isaca.org/archives*

17 Tripathi, M.; A. Mukhopadhyay; "Financial Loss Due to a Data Privacy Breach: An Empirical Analysis," *Journal of Organizational Computing and Electronic Commerce*, vol. 30, iss. 4, 2020, p. 381–400

18 Mukhopadhyay, A.; S. Chatterjee; K. K. Bagchi; P. J. Kirs; G. K. Shukla; "Cyber Risk Assessment and Mitigation (CRAM) Framework Using Logit and Probit Models for Cyber Insurance," *Information Systems Frontiers*, vol. 21, iss. 5, 2019, p. 997–1018

19 Das, S.; A. Mukhopadhyay; G. K. Shukla; "I-HOPE Framework for Predicting Cyber Breaches: A Logit Approach," *Proceedings of the Annual Hawaii International Conference on System Sciences*, Institute of Electrical and Electronics Engineers (IEEE), 2013

20 Biswas, B.; A. Mukhopadhyay; G. Dhillon; "GARCH-Based Risk Assessment and Mean-Variance-Based Risk Mitigation Framework for Software Vulnerabilities," *AMCIS 2017: A Tradition of Innovation*, 23rd Americas Conference on Information Systems, 2017

21 Biswas, B.; A. Mukhopadhyay; "Phishing Detection and Loss Computation Hybrid Model: A Machine-Learning Approach," *ISACA Journal*, vol. 1, 2017, *https://www.isaca.org/archives*