# Cyber risk assessment and mitigation strategy for DDoS attacks in BFSI segment

*Kalpit Sharma*
*Doctoral Student,*
*Indian Institute of Management Lucknow*
*kalpit@iiml.ac.in*

*Arunabha Mukhopadhyay*
*Professor,*
*Indian Institute of Management Lucknow*
*arunabha@iiml.ac.in*

## Abstract

Malicious hackers launch Distributed Denial of Service (DDoS) attacks to disrupt business globally, by preventing legitimate users from accessing the original website. Last year alone the financial cost of a DDoS attack ranged from US$120K to US$2M, including direct and indirect losses. In 2018, DDoS attacks were 37% larger in size than the previous year. DDoS attacks flood the network with data packets as large as 300 Gbps to overwhelm the routers and servers. But in most cases, these attackers use any army of botnets, which carry out the DDoS attack. Thus, it is difficult to detect the attacker and prevent the attackers in future. In this study we wish to investigate the following: (a) what is the probability of correctly identifying a DDoS attack, (b) what is the expected loss due to a DDoS attack on an organization, and (c) how can an organization mitigate a DDoS attack. In this study, we use a dataset of DDoS attacks reported by a globally reputed Content Delivery Network (CDN). We employ data preprocessing steps such as hierarchical clustering and k-means clustering to divide our data into an optimal number of classes to increase class wise as well as the overall accuracy of the classifier. We use a Naive Bayes classifier that provides us with the probability of such an attack happening across different industries. In addition, we calculate the expected loss or opportunity cost that firms must bear when under such an attack. Subsequently, we suggest ways to mitigate the losses occurring due to restricted access to your firms' cyber-resources when under a DDoS attack.

## Introduction

Malicious hackers resort to Distributed Denial of Service (DDoS) attacks to disrupt business globally, by preventing legitimate users from accessing the original website. In 2018, the direct and indirect losses due to a DDoS attack ranged from US$ 120K to US$ 2M (Abrahams, 2018). In 2017, number of DDoS attacks increased by 29% from 2016. In 2017, the average attack size was recorded as high as 26 Gbps. [1]. But in most cases, these attackers use an army of botnets, which carry out the DDoS attack. Thus, it is difficult to detect the attacker and prevent the attackers in future.

DDoS attacks are executed in varying severity across different industries and networks. Industries like, BFSI, gaming, video streaming which rely heavily on real-time health of their networks suffered from losses amounting to US$ 4 million in 2015 (Zumberge, 2015). The losses have been increasing since. A hacktivist group attacked two banking giants in US, namely, Bank of America and JPMorgan Chase. They used DDoS attack in 28th January 2014 to bring some of the banks' cyber-resources to a halt. Twitter was inundated with tweets of banks' customers complaining of online outages during the same time. In 2018, two major Dutch banks, ABN AMRO, ING and Rabobank fell victim to DDoS attacks which resulted in

slowing down of their websites with no alleged security threat to customers' financial transactions. In October 2019, there was a wave of ransom-driven DDoS attacks on South African banks and city of Johannesburg. According to Akamai, India itself is 7th most targeted nation for DDoS attack specifically on BFSI web servers.

The BFSI industry have been second worst hit by DDoS attacks. Figure 1 shows that in 2018-19, the Financial & credit union industry has faced 6% of the total 35200 DDoS attacks that were recorded by Akamai's Prolexic platform. Hackers find it lucrative because of the amount of money involved in transactions as well as that can be claimed through ransom. Firms lose approximately US$ 50,000 per hour when under a DDoS attack (Bezsonoff, 2018). Thus, we intend to quantify the probability of detecting such a DDoS attack and expected loss associated with the same. Subsequently, we intend to map different classes and suggest ways to lower the risk as well as loss severity for ones with extremely high severity.

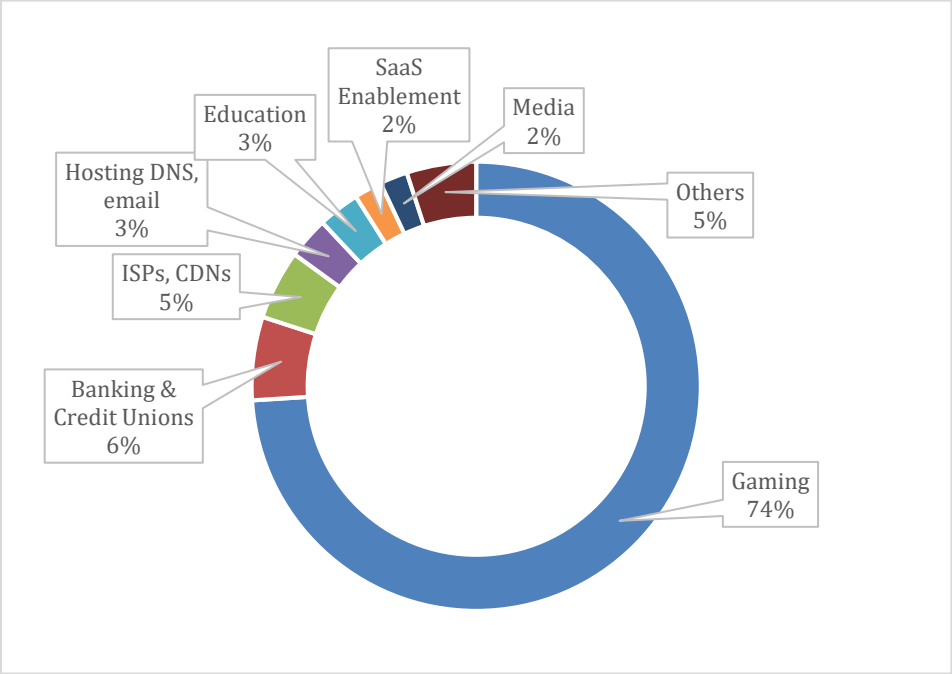


**Figure 1: Percentage of DDoS attacks across different industries (for years 2018-2019)**

## Research Question and our Proposed Model

Figure 2 illustrates the proposed Risk Assessment and Mitigation model for DDoS attacks for firms in the BFSI industry in detail. We intend to explore following issues related to DDoS attacks in BFSI industry:

- What is the probability (p) of not detecting a DDoS attack using our model for BFSI?
- What is the expected loss (*E(L)*), for a BFSI firm if our model fails to detect the attack vector?
- What are the ways to lower the risk (p) and expected loss (*E(L)*) for BFSI?

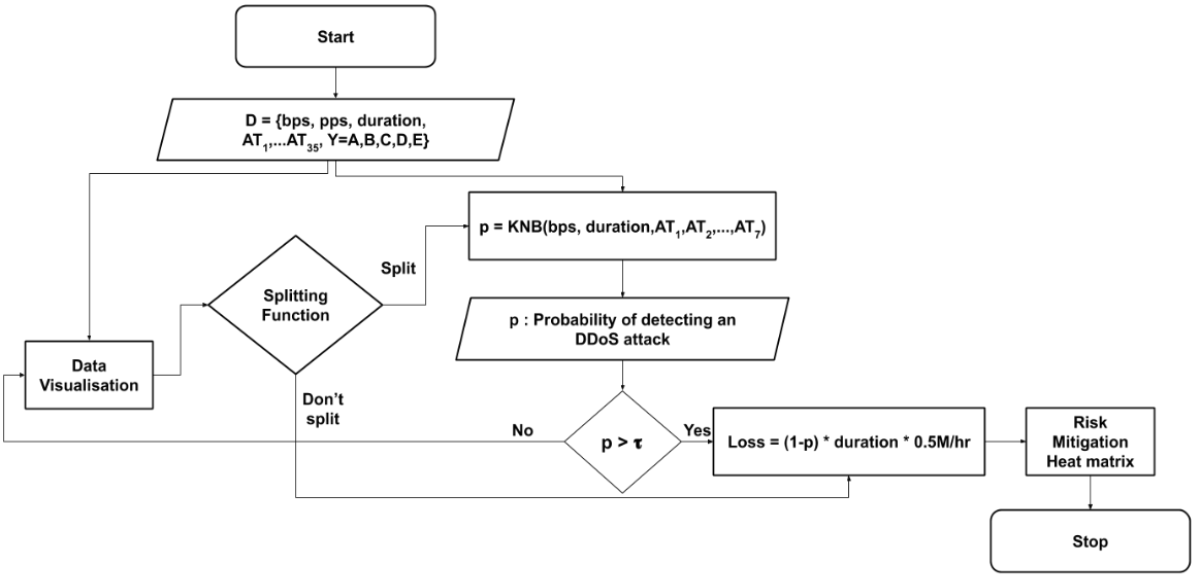


**Figure 2: Flowchart of the proposed model**

## Literature Review

Cyber risk assessment has been at the helm of cybersecurity research since the advent of Newer Information Technology for businesses (Gordon, Loeb, & Sohail, 2003). Assessment of risk helps identify and subsequently quantify the probability of a cybersecurity incident occurring provided the security protocols were in place. The cyber risk assessment also aids in evaluating the efficacy of IT risk management compliance structure already in place in organizations.

Cyber risk assessment methods intend to identify information assets (such as hardware, systems, laptops, customer data, and intellectual property) which can be under cyber-attack and their associated risks. Information assets are divided into multiple classes according to the perceived risk in order of their severity and broken into sub-parts to correctly identify the risky component of the asset and its type (tangible, intangible, etc.) (O'reilly et al., 2018)). The risk assessment stage is followed by quantification of identified risk with the help of diverse methods aiming at attaching a monetary value to it.

Cyber risk quantification methods rely on the probability of a risky incident occurring and meticulous calculation of loss amount for such incidents. Thus, the accuracy of such methods relies on the accuracy of risk identification as well as loss calculation. Loss estimation methods also evolve according to the unit of analysis and definition of loss for which we are undertaking the aforesaid exercise. Thus, the expected loss for entity resulting due to cyberattacks depends not only on the incident but also on the ability of an entity to accurately estimate its loss. These estimations also vary in their methodological rigor depending upon the type and granularity of data available to calculate them. Cyber risk quantification techniques range from mathematical risk modelling to data mining methods using empirical data available from security providers (Campbell & Stamp, 2004).

Many of the initial quantitative approaches, tried to model the cyber risk scenario as an uncertainty model where the probability of cyber risk occurrence is to be studied. Most of these classification model use traffic attributes such as TCP/IP layer used for the attack, quanta of bits used and packet structure to typify cyber-attacks' presence or otherwise. The uncertainty of classification can be modelled using various statistical methods which use some prior knowledge of the occurrence of a cyber breach and update it with current evidence through data. Mukhopadhyay et al (2017) used logit and probit models to calculate the probability of a cyber risk occurring using CSI-FBI survey data from 1997-2010. Biswas et al (2017) used machine learning techniques such as Bagger classifier and CART based hybrid classifier to assess phishing attacks. On the other hand, Balkanli et al (2015) augmented decision tree classifier by

using Chi-square and Symmetric uncertainty to analyse DDoS feature vectors from CAIDA dataset. Heratha et al (2007) used copula-based methods, that are quite popular with actuarial researchers, to quantify cyber risk and thus, propose insurance approaches. Alhazmi et al (2007) have previously shown that cyber risk attack vectors can be efficiently modelled using density estimation methods and thus, augment the accuracy of a method that relies upon distribution statistics to classify. Smith and Eloff (2002) used fuzzy logic-based RiMaHCoF method to analyse cyber-risks and quantify them in overlapping and conflicting risk classes.

Decision trees and their other variants like ensemble methods, hybrid classifiers, etc. are quite efficient with provision for decision rules for informing future decisions for classifying similar incident vectors. The use of a small number of independent features construct a very highly complex tree and pruning it becomes difficult given the trade-off with its accuracy.

Hubbard and Seiersen (2016) pointed out that the Bayesian statistical methods try to mimic human decision-making process under uncertainty where we update our decisions with every quantum of new information about the phenomenon. As independence of features is the underlying assumption in Bayesian methods, therefore, they perform well on datasets in which class boundary either overlap or contradict. The extant literature has not explored the Naïve Bayes classification method in the context of the multi-label classification problem. This study seeks to address the aforesaid research gap for estimating the probability of not detecting a DDoS attack.

## Data

In this study, we use a dataset from ***Akamai security platform***, that consists of 424 records of 7 DDoS attack types grouped into five overlapping attack classes in BFSI industry for the years 2012 to 2018 (McKeay, 2017; McKeay, 2019a; McKeay, 2019b) (Table 1).

**Table 1: Class wise frequency of attacks from 2012 to 2018 (n = 424)**

| *Attack Composition* | *Class* | *Count* |
|---|---|---|
| UDP Fragment, DNS Flood | A | 141 |
| NTP FLOOD | B | 101 |
| UDP Fragment, CharGEN Attack | C | 29 |
| SSDP Flood | D | 64 |
| UDP Flood | E* | 89 |

## Methodology

We have used Kernel based Naive Bayes (KNB) classifier, as our training data does not follow any parametric class distributions. Bayesian classifiers are statistical classifiers. They are used to predict class membership probabilities i.e. the probability that a given dataset record belongs to a class. The class which has the highest probability amongst the others is the dominant class for the record (Kamber and Pei, 2012). Naive Bayes classifier gives higher accuracy if the attributes are independent inherently. Our dataset has attack tuples with 2 independent classes overlapping and forming a hybrid attack vector. Thus, Kernel-based Naive Bayes (KNB) classifier serves the purpose of classification aptly.

Table 2 describes the steps that the proposed methodology follows to produce classification accuracy above some value τ (at least 70%) (Biswas and Mukhopadhyay, 2016). The threshold τ may be decided on the criticality of online mode of operation viz-a-viz their overall operations. Subsequently, it calculates the expected loss for the firm and suggests mitigation strategies to reduce the loss.

**Table 2: Steps of our Kernel-based Naive Bayes classifier model**

| | |
|---|---|
| Step 1 | Pass training dataset (TD) to a Kernel-based Naive Bayes (KNB) classifier. |
| Step 2 | KNB classifier assign all records to classes based on dominant posterior probability. |
| Step 3 | Calculate accuracy (p) of the detecting DDoS attack by the classifier.<br>$p = KNB(bps, duration, AT_1, AT_2 \ldots, AT_7)$ *(Eq.1)* |
| Step 4 | If accuracy (p) < threshold τ |

| | |
|---|---|
| | Then *Splitting_function (TD)*,<br>*Else*<br>Expected Loss, *E(L) = (1-p) * 0.5 * (duration) (Eq.2)* .<br>Propose mitigation strategies (technology pieces, cyber-insurance)<br>*If (1-p) > 0.5 AND E(L) > US$ 3 million, Then Implement Technology + Cyber-insurance* (Eq. 3)<br>*End* |
| Step 5 | TDnew = *Splitting_function (TD).* |
| Step 6 | TD = TDnew |
| Step 7 | Go to step 1. |
| | *Splitting_function(TD)* |
| Step A | k = Dendrogram(TD) |
| Step B | Generate new clusters using kmeans_algo(k). |

We used MATLAB 2019a to analyze the data through our model. Table 3 shows the confusion matrix of the initial run. The diagonal elements of Table 3 show the number of instances where attacks were correctly classified. For example, diagonal elements refer to instances correctly classified by our model (i.e., 53 for Class A) while 3 records have been misclassified as false positive and 2 records as false negative. Our model has correctly classified 96% of Class A attacks and 81 % of Class C attacks.

**Table 3: Confusion Matrix for Testing Dataset (n = 170)**

| | Predicted A | Predicted B | Predicted C | Predicted D | Predicted E | Total | Probability of detecting DDoS attack (p) |
|---|---|---|---|---|---|---|---|
| Actual A | 53 | 0 | 1 | 0 | 1 | 55 | 0.96 |
| Actual B | 0 | 43 | 0 | 0 | 0 | 43 | 1 |
| Actual C | 2 | 0 | 9 | 0 | 0 | 11 | 0.81 |
| Actual D | 0 | 0 | 0 | 18 | 2 | 20 | 0.9 |
| Actual E | 1 | 2 | 0 | 2 | 36 | 41 | 0.87 |
| Total | 56 | 45 | 10 | 20 | 39 | | |

We next split the class C with the highest number of misclassifications (i.e. 100-81 =19%). We, then visualize the dataset using hierarchal clustering (**Figure 3**) from which we note that there are 3 distinct branches (i.e., C1, C2, C3) based on dissimilarity index. So, we decide to split original Class C (n=29) into three sub-classes (C1(n=27), C2(n=1), C3(n=1)) using k-means clustering algorithm.

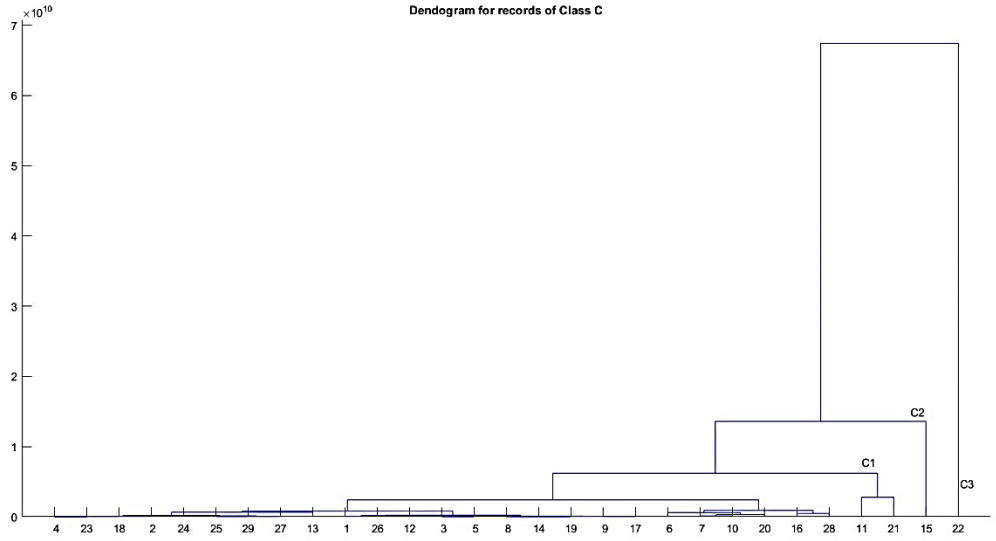


**Figure 3: Splitting of original Class C into three sub-classes C using Dendrogram**

## Results

The following sections discuss the results for computing the probability (p) of detecting a DDoS attack using the model, the expected loss (*E[LJ*) for the enterprise if the model fails to detect the attack vector and risk mitigation strategies.

### *Probability of Detecting a DDoS Attack*

**Table 4** shows that the overall accuracy of the classifier was 93.53 percent. Diagonal elements (shaded) represent instances in which the model correctly classified attacks (159 out of 170 records). The model correctly classified 100 percent of class A, C1, C2, C3 attacks and 71 percent of class E attacks.

**Table 4: Probability of Correct Classification for Testing Dataset (n=170)**

| | Predicted A | Predicted B | Predicted C | Predicted D | Predicted E1 | Predicted E2 | Predicted E3 | Probability of detecting an attack (p) |
|---|---|---|---|---|---|---|---|---|
| Actual A | 53 | 0 | 0 | 0 | 0 | 0 | 0 | 1.00 |
| Actual B | 1 | 46 | 0 | 0 | 0 | 0 | 0 | 0.98 |
| Actual C1 | 0 | 0 | 10 | 0 | 0 | 0 | 0 | 1.00 |
| Actual C2 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1.00 |
| Actual C3 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 1.00 |
| Actual D | 0 | 1 | 0 | 0 | 0 | 27 | 1 | 0.93 |
| Actual E | 3 | 2 | 1 | 0 | 0 | 2 | 20 | 0.71 |

### *Loss Computation After a DDoS Attack*

In **Table 5**, column 2 demonstrates the risk (i.e., probability of not detecting an attack [1-p]), and column 4 illustrates the severity of DDoS attack (i.e., of expected losses for each attack class based on Eq. 2).

**Table 5: Expected Loss Per Hour for Each Attack Class**

| Class | Risk : Probability of not detecting an attack (1-p) | Duration of attack (in hours) | Severity : Expected loss per hour (in millions US$) *E(L) = (1-p) * 0.5 * (duration)* |
|---|---|---|---|
| (1) | (2) | (3) | (4) |
| A | 0.00 | 55 | 0.00 |
| B | 0.02 | 49 | 0.49 |
| C1 | 0.00 | 42 | 0.00 |
| C2 | 0.00 | 60 | 0.00 |
| C3 | 0.00 | 43 | 0.00 |
| D | 0.07 | 62 | 2.17 |
| E | 0.29 | 66 | 9.57 |

## *Risk Mitigation Strategies*

**Figure 9** depicts a heat matrix calculated from the model that situates the different DDoS attacks in terms of Risk × Severity. This helps a chief technology officer (CTO) to prioritize the risk mitigation strategy, such as technological intervention to reduce the risk or transfer risk through cyberinsurance. For example, a DDoS attack of class E is in the high risk-high severity quadrant, while classes A, B, C1, C2 and C3 are in the low risk-low severity quadrant. For example, the CTO of an enterprise at risk of DDoS attacks of class E type should consider implementing the following risk mitigation strategies: First, add stringent firewalls or intrusion detection systems or divert excess or illegitimate traffic to backup servers or content delivery networks (CDNs) to reduce the risk and thus lower the severity of DDoS attack. Next, transfer the residual risk by subscribing to cyberinsurance policies, thus moving into the low risk-low severity quadrant. (Mukhopadhyay et al, 2017; Das and Mukhopadhyay, 2017; Biswas, Mukhopadhyay and Dhillon, 2017; Biswas and Mukhopadhyay, 2017)

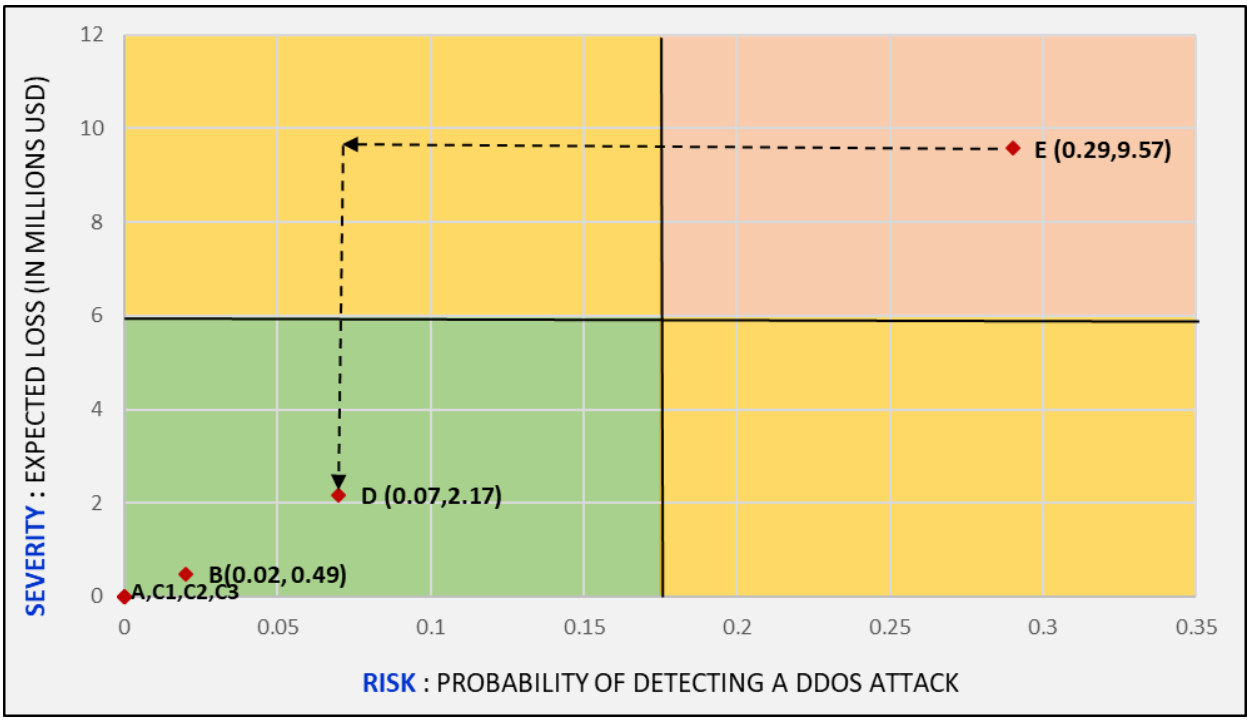


**Figure 4: Risk Mitigation Heat Matrix**

## Conclusion

Our two-step classification and loss computation model use Kernel based Naive Bayes classifier and clustering-based data pre-processing to calculate the probability of detecting a DDoS attack occurring on the BFSI industry firms. Our aforesaid classifier was able to detect a DDoS attack. Our model can also be scaled for large datasets (i.e. size and dimension). The proposed model also calculates class wise as well as overall expected loss amount. We also propose risk mitigation strategies for CTO/CSO of a firm, for each class of DDoS attack using a combination of technological intervention and cyber-insurance.

## References

Abrams, L.; "Dramatic Increase of DDoS Attack Sizes Attributed to IoT Devices," Bleeping Computer, 12 September 2018, https://www.bleepingcomputer.com/news/security/dramatic-increase-of-ddos-attack-sizes-attributed-to-iot-devices/

Zumberge, M.; "Cyber Attacks on the Rise in Media Biz Since Sony Hack: Survey (Exclusive)," Variety, 5 November 2015, https://variety.com/2015/digital/news/sony-hack-anniversary-cybersecurity-data-1201633671/

Shani, T.; "Updated: This DDoS Attack Unleashed the Most Packets per Second Ever. Here's Why That's Important," Imperva, 12 June 2019, https://www.imperva.com/blog/this-ddos-attack-unleashed-the-most-packets-per-second-ever-heres-why-thats-important/

Bezsonoff, N.; "The State of DDoS Attacks in 2017," Neustar, 11 October 2017, https://www.home.neustar/blog/neustar-global-attacks-and-cyber-security-insight-report

Chiel, E.; "Here Are the Sites You Can't Access Because Someone Took the Internet Down," Splinter News, 21 October 2016, https://splinternews.com/here-are-the-sites-you-cant-access-because-someone-took-1793863079

McKeay, M.; "Q4 2017 State of the Internet Security Report," Akamai, 28 December 2017, https://www.akamai.com/us/en/multimedia/documents/state-of-the-internet/q4-2017-state-of-the-internet-security-report.pdf

McKeay, M.; "2019 State of the Internet/Security: Web Attacks and Gaming Abuse," Akamai, June 2019, https://www.akamai.com/us/en/multimedia/documents/state-of-the-internet/soti-security-web-attacks-and-gaming-abuse-report-2019.pdf

McKeay, M.; "2019 State of the Internet/Security: DDoS and Application Attacks," Akamai, January 2019, https://www.akamai.com/us/en/multimedia/documents/state-of-the-internet/state-of-the-internet-security-ddos-and-application-attacks-2019.pdf

Han, J.; M. Kamber; J. Pei; Data Mining: Concepts and Techniques, Elsevier, the Netherlands, 2012

Biswas, B.; S. Pal; A. Mukhopadhyay; "AVICS-Eco Framework: An Approach to Attack Prediction and Vulnerability Assessment in a Cyber-Ecosystem," AMCIS 2016: Surfing the IT Innovation Wave—22nd Americas Conference on Information Systems

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, 21(5), 997–1018, 2017, https://doi.org/ 10.1007/s10796-017-9808-5

Das, S.; A. Mukhopadhyay; G. K. Shukla; "i-HOPE Framework for Predicting Cyber Breaches: A Logit Approach," 46th Hawaii International Conference on System Sciences, 2013, https://doi.org/10.1109/hicss.2013.256

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

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

Alhazmi, O. H., Malaiya, Y. K., and Ray, I. 2007. Measuring, analyzing and predicting security vulnerabilities in software systems. Computers and Security , 26(3): 219–228.

Balkanli, E., Nur Zincir-Heywood, A., and Heywood, M. I. 2015. Feature selection for robust backscatter DDoS detection. In Proceedings-Conference on Local Computer Networks, LCN, volume 2015-Decem, 611–618. IEEE.

Campbell, P. L. and Stamp, J. E. 2004. A classification scheme for risk assessment methods. Technical report, Sandia National Laboratories.

Gordon, L. A., Loeb, M. P., and Sohail, T. 2003. A framework for using insurance for cyber-risk Management. Communications of the ACM , 46(3): 81–85.

Hastie, T., Tibshirani, R., and Friedman, J. 2009. The Elements of Statistical Learning Springer-Verlag New York Inc.

Herath, H. S. and Herath, T. C. 2011. Copula-based actuarial model for pricing cyber-insurance Policies. Workshop on the Economics of Information Security , 2(1): 7–20.

Hubbard, D. W. and Seiersen, R. 2016. How to Measure Anything in Cybersecurity Risk . Wiley John + Sons.

OReilly, P., Rigopoulos, K., Feldman, L., and Witte, G. 2018. 2017 NIST/ITL Cybersecurity program annual report.

Silverman, B. W. 2018. Density estimation: For statistics and data analysis . Chapman and Hall.

Smith, E. and Eloff, J. 2002. A Prototype for Assessing Information Technology Risks in Health Care. Computers & Security , 21(3): 266–284