



# International Journal for Innovative Engineering and Management Research

A Peer Reviewed Open Access International Journal

www.ijiemr.org

## COPY RIGHT



**ELSEVIER**  
**SSRN**

**2019IJIEMR.** Personal use of this material is permitted. Permission from IJIEMR must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works. No Reprint should be done to this paper, all copy right is authenticated to Paper Authors

IJIEMR Transactions, online available on 3rd Apr 2019. Link

[:http://www.ijiemr.org/downloads.php?vol=Volume-08&issue=ISSUE-04](http://www.ijiemr.org/downloads.php?vol=Volume-08&issue=ISSUE-04)

Title: **MODELING AND PREDICTING CYBER RISK ANALYSIS**

Volume 08, Issue 04, Pages: 64–71.

Paper Authors

**MR. V.SREEKANTH, MR. B.MASTHAN BABA**

Sree Vidyanikethan Institute Of Management from SV University.,(A.P),INDIA



USE THIS BARCODE TO ACCESS YOUR ONLINE PAPER

To Secure Your Paper As Per **UGC Guidelines** We Are Providing A Electronic Bar Code

## MODELING AND PREDICTING CYBER RISK ANALYSIS

<sup>1</sup>MR. V.SREEKANTH, <sup>2</sup>MR. B.MASTHAN BABA

<sup>1</sup>VI<sup>th</sup> semester, dept of MCA , Sree Vidyanikethan Institute Of Management from SV University.,(A.P),INDIA

<sup>2</sup>MCA,Assistant Professor Dept of MCA, Sree Vidyanikethan Institute Of Management from SV University, (A.P),INDIA

<sup>1</sup>sreekanthvathala@gmail.com

### ABSTRACT

Investigating digital episode informational indexes is a vital strategy for extending our comprehension of the development of the risk circumstance. This is a moderately new research point, and numerous investigations stay to be finished. In this paper, we report a measurable investigation of a rupture episode informational index comparing to 12 years (2005– 2017) of digital hacking exercises that incorporate malware assaults. We demonstrate that, as opposed to the discoveries announced in the writing, both hacking rupture episode between landing times and break sizes ought to be displayed by stochastic procedures, instead of by dispersions since they show autocorrelations. At that point, we propose specific stochastic procedure models to, separately, fit the between entry times and the break sizes. We additionally demonstrate that these models can anticipate the between entry times and the rupture sizes. So as to get further bits of knowledge into the development of hacking break occurrences, we lead both subjective and quantitative pattern examinations on the informational collection. We draw a lot of digital security experiences, including that the risk of digital hacks is to be sure deteriorating as far as their recurrence, yet not as far as the size of their harm.

**Keywords:** Hacking breach, data breach, cyber threats, cyber risk analysis, breach prediction, trend analysis, time series, Cyber security data analytics.

### I INTRODUCTION

While mechanical arrangements can solidify digital frameworks against assaults, information ruptures keep on being a major issue. This rouses us to describe the advancement of information rupture episodes. This not exclusively will profound our comprehension of information ruptures, yet in addition shed light on different methodologies for relieving the harm, for

example, protection. Many trust that protection will be valuable, however the improvement of precise digital hazard measurements to direct the task of protection rates is past the compass of the present comprehension of information breaks(e.g.,the absence of demonstrating approaches) [6].

As of late, analysts began demonstrating information rupture occurrences. Maillart and Sornette [7] examined the measurable properties of the individual personality misfortunes in the United States between year 2000 and 2008 [8]. They found that the quantity of rupture episodes drastically increments from 2000 to July 2006 however stays stable from that point. Edwards [9] broke down a dataset containing 2,253 break occurrences that length over 10 years (2005 to 2015) [1]. They found that neither the size nor the recurrence of information breaks has expanded throughout the years. Wheatley [10] broke down a dataset that is consolidated from [8] and [1] and compares to authoritative break episodes between year 2000 and 2015. They found that the recurrence of huge rupture occurrences (i.e., the ones that break in excess of 50,000 records) jumping out at US firms is autonomous of time, yet the recurrence of expansive break episodes striking non-US firms displays an expanding pattern.

## II RESEARCH AND DATASET DESCRIPTION

### SUPPORT VECTOR MACHINE ALGORITHM:

"Bolster Vector Machine" (SVM) is a directed AI calculation which can be utilized for both order and relapse difficulties. Be that as it may, it is generally utilized in arrangement issues. In this calculation, we plot every datum thing as a point in n-dimensional space (where n is number of highlights you have) with the estimation of each component being the estimation of a specific facilitate. At that point, we perform grouping by finding the hyper-plane that separate the two classes great (take a gander

at the underneath snapshot).Support Vectors are essentially the co-ordinates of individual perception. Bolster Vector Machine is an outskirts which best isolates the two classes (hyper-plane/line).More formally, a help vector machine builds a hyper plane or set of hyper planes in a high-or limitless dimensional space, which can be utilized for characterization, relapse, or different errands like anomalies recognition. Naturally, a great detachment is accomplished by the hyper plane that has the biggest separation to the closest preparing information purpose of any class (supposed practical edge), since as a rule the bigger the edge the lower the speculation blunder of the classifier. Though the first issue might be expressed in a limited dimensional space, it frequently happens that the sets to segregate are not directly distinct in that space. Therefore, it was suggested that the first limited dimensional space be mapped into an a lot higher-dimensional space, probably making the detachment less demanding in that space.

#### Algorithm 1 Algorithm for Predicting the VaR <sub>$\alpha$</sub> 's of the Hacking Incidents Inter-Arrival Times and the Breach Sizes Separately

Input: Historical incidents inter-arrival times and breach sizes, denoted by  $\{(d_{it}, y_{it})\}_{i=1, \dots, m+n}$ , where an in-sample  $\{(d_{it}, y_{it})\}_{i=1, \dots, m}$  as mentioned above was used for fitting and an out-of-sample  $\{(d_{it}, y_{it})\}_{i=m+1, \dots, n}$  is used for evaluation prediction accuracy;  $\alpha$  level.

- 1: for  $i = m + 1, \dots, n$  do
  - 2: Estimate the LACD<sub>1</sub> model of the incidents inter-arrival times based on  $\{d_{it}\}_{i=1, \dots, i-1}$ , and predict the conditional mean  $\Psi_i = \exp(\omega + a_1 \log(\epsilon_{i-1}) + b_1 \log(\Psi_{i-1}))$ ;
  - 3: Estimate the ARMA-GARCH of log-transformed size, and predict the next mean  $\hat{\mu}_i$  and standard error  $\hat{\sigma}_i$ ;
  - 4: Select a suitable Copula using the bivariate residuals from the previous models based on AIC;
  - 5: Based on the estimated copula, simulate 10000 2-dimensional copula samples  $(u_{1,i}^{(k)}, u_{2,i}^{(k)})$ ,  $k = 1, \dots, 10000$ ;
  - 6: For the incidents inter-arrival times, convert the simulated dependent samples  $u_{1,i}^{(k)}$ 's into the  $z_{1,i}^{(k)}$ 's by using the inverse of the estimated generalized gamma distribution,  $k = 1, \dots, 10000$ ;
  - 7: For the breach sizes, convert the simulated dependent samples  $u_{2,i}^{(k)}$ 's into the  $z_{2,i}^{(k)}$ 's by using the inverse of the estimated mixed extreme value distribution,  $k = 1, \dots, 10000$ ;
  - 8: Compute the predicted 10000 2-dimensional breach data  $(d_{it}^{(k)}, y_{it}^{(k)})$ ,  $k = 1, \dots, 10000$  based on Eq. (IV.1) and (IV.3), respectively;
  - 9: Compute the VaR <sub>$\alpha, d$</sub> ( $i$ ) for the incidents inter-arrival times and VaR <sub>$\alpha, y$</sub> ( $i$ ) for the log-transformed breach sizes based on the simulated breach data.
  - 10: if  $d_{it}^{(k)} > \text{VaR}_{\alpha, d}(i)$  then
  - 11: A violation to the incidents inter-arrival time occurs;
  - 12: end if
  - 13: if  $y_{it}^{(k)} > \text{VaR}_{\alpha, y}(i)$ ; then
  - 14: A violation to the breach size occurs;
  - 15: end if
  - 16: end for
- Output: Numbers of violations in inter-arrival times and breach sizes.

## **DATASET DESCRIPTION**

Given a dataset of digital hacking break occurrences, we need to utilize it to respond to the accompanying inquiries.

1) Should we utilize a circulation or stochastic procedure to portray the break occurrences between landing times, and which conveyance or procedure? This inquiry is imperative in light of the fact that noting it will specifically develop our comprehension of the dynamic digital hacking rupture circumstance from a fleeting point of view.

2) Should we utilize a dispersion or stochastic procedure to portray the rupture sizes, and which conveyance or procedure? This inquiry is critical on the grounds that noting it will straightforwardly extend our comprehension of the dynamic digital hacking rupture circumstance from a greatness point of view.

3) Are the break sizes and the occurrences between landing times autonomous of one another? If not, in what manner would it be advisable for us to portray the reliance between them? This inquiry is imperative on the grounds that noting it will straightforwardly develop our comprehension of the dynamic digital hacking break circumstance from a joint fleeting and greatness point of view.

4) Can we foresee when the following hacking occurrence will happen, and what the break size would be? This inquiry is essential on the grounds that noting it demonstrates our capacity to foresee the circumstance and conceivably direct proactive barrier at a little time scale (e.g., days or weeks early). For instance, when the

likelihood that a major break episode will happen amid the following week is high, the safeguard may powerfully alter the guard pose (e.g., implementing progressively confined approaches amid the following week). This is like what climate anticipating can do in the physical world.

5) What are the patterns that are displayed by hacking break occurrences? This inquiry is vital on the grounds that we can draw more elevated amount experiences into whether the circumstance is improving or more regrettable over a substantial time scale (e.g., 10 years), and to what degree.

## **III IMPLEMENTATION**

### **PROPOSED SYSTEM:**

In this paper, we make the accompanying three commitments. In the first place, we demonstrate that both the hacking rupture episode interarrival times (reflecting occurrence recurrence) and break sizes ought to be displayed by stochastic procedures, as opposed to by appropriations. We find that a specific point procedure can sufficiently depict the advancement of the hacking break episodes between entry times and that a specific ARMA-GARCH model can enough portray the development of the hacking rupture sizes, where ARMA is abbreviation for "AutoRegressive and Moving Average" and GARCH is abbreviation for "Summed up AutoRegressive Conditional Heteroskedasticity." We demonstrate that these stochastic procedure models can anticipate the between landing times and the break sizes. To the best of our insight, this is the main paper appearing stochastic procedures, instead of appropriations, ought to be utilized to display these digital risk

factors. Second, we find a positive reliance between the episodes between entry times and the rupture sizes, and demonstrate that this reliance can be satisfactorily depicted by a specific copula. We likewise demonstrate that while foreseeing between entry times and break sizes, it is important to think about the reliance; generally, the forecast outcomes are not precise. To the best of our insight, this is the principal work demonstrating the presence of this reliance and the result of disregarding it. Third, we lead both subjective and quantitative pattern examinations of the digital hacking break episodes. We find that the circumstance is to be sure deteriorating as far as the episodes between entry time in light of the fact that hacking rupture occurrences become increasingly visit, yet the circumstance is balancing out regarding the occurrence break estimate, demonstrating that the harm of individual hacking rupture episodes won't deteriorate. We trust the present examination will motivate more examinations, which can offer profound experiences into interchange chance alleviation approaches. Such bits of knowledge are helpful to insurance agencies, government offices, and controllers since they have to profoundly comprehend the idea of information break dangers.

## IV SYSTEM DESIGN

### SYSTEM ARCHITECTURE:

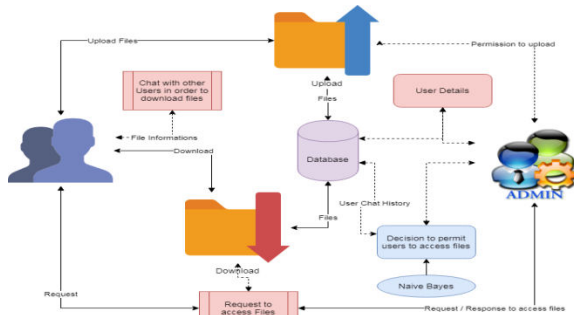
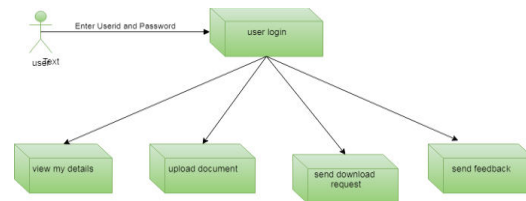


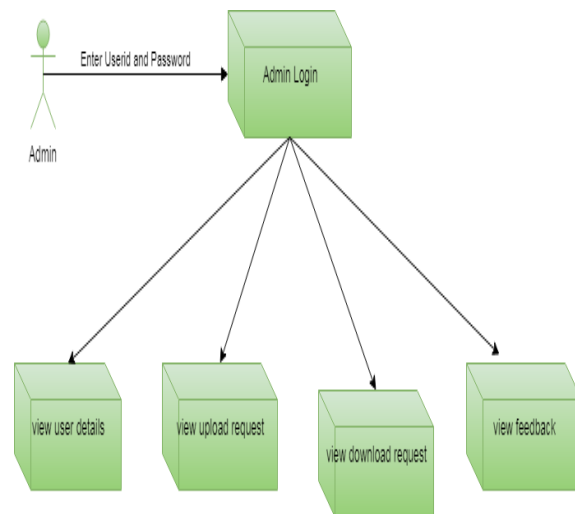
Figure 2: system architecture

### Component Diagram

#### a. User



#### b. Admin



### Basic Analysis of Hacking Breach Sizes

Table III condenses the fundamental measurements of the hacking break sizes. We see that three Business classifications have a lot bigger mean break sizes than others. We further see that there exists an expansive standard deviation for the rupture measure in every one of the injured individual classes, and that the standard deviation is in every case a lot bigger than the relating mean. plots the log-changed rupture sizes in light of the fact that, as we can see from Table III, the break sizes display extensive unpredictability and skewness (which is shown by the considerable distinction between the middle and the mean qualities), which make them

difficult to demonstrate without making changes.

**TABLE III**

STATISTICS OF HACKING BREACH SIZES, WHERE 'SD' STANDS FOR STANDARD DEVIATION

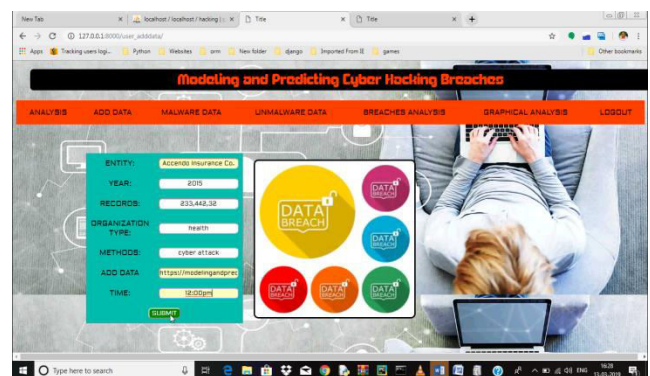
Size	Min	Median	Mean	SD	Max	Total
BSF	11	2000	2228000	10436141	76000000	69
BSO	11	6470	9677000	49488457	412000000	84
BSR	12	1464	2666000	14678814	100000000	60
EDU	20	10870	41940	95481.03	800000	165
GOV	24	14000	119400	293147.3	1700000	50
MED	180	4668	34140	96820.77	697600	163
NGO	444	15000	28190	34754.27	110000	9
Total	11	6324	1909000	19588938	412000000	600

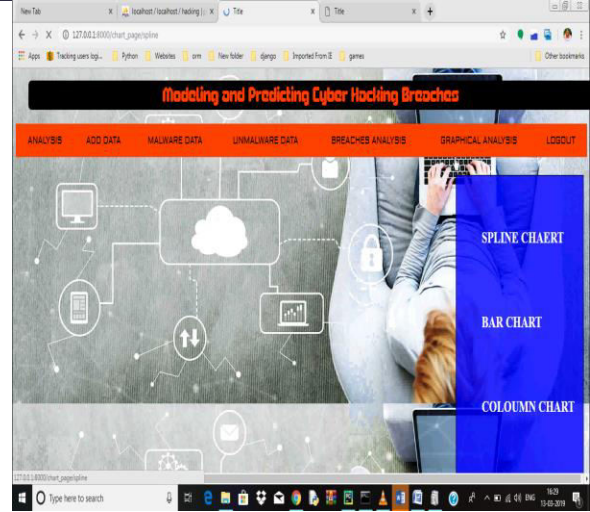
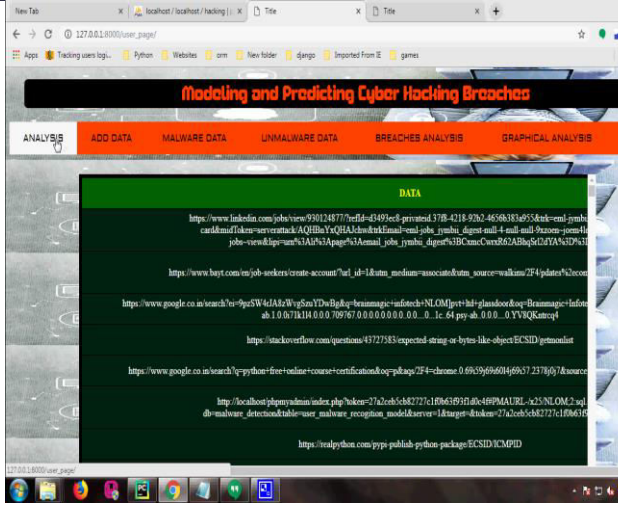
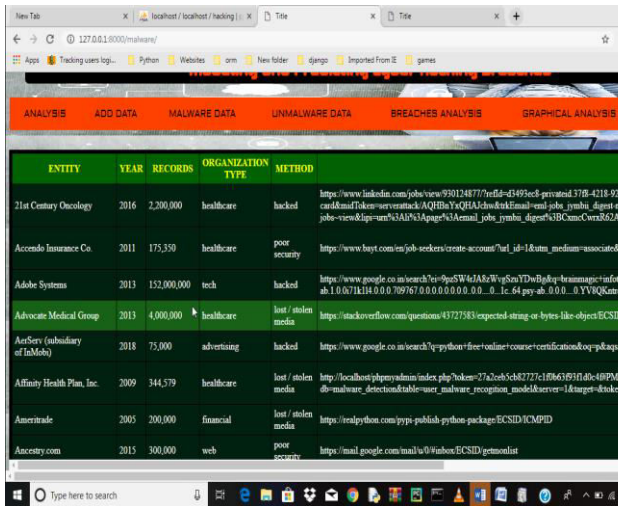
So as to respond to the inquiry whether the rupture sizes ought to be displayed by a dissemination or stochastic procedure, we plot the transient relationships between's the break sizes. plot the example ACF and PACF for the log-changed rupture sizes, individually. We watch connections between's the break sizes, implying that we should utilize a stochastic procedure, as opposed to a dispersion, to show the rupture sizes. This is rather than the understanding offered by past examinations [7], which recommends to utilize a skewed appropriation to show the break sizes. We characteristic the attracting of this understanding to the way that these examinations [7] did not investigate this due point of view of worldly connections. An imperative factor for deciding if to utilize a dispersion or a stochastic procedure to portray something, relies upon regardless of whether there is worldly autocorrelation between the individual examples. This is on the grounds that zero worldly

autocorrelation implies that the examples are autonomous of one another; something else, non-zero transient autocorrelation implies that they are not free of one another and ought not be displayed by a dissemination.

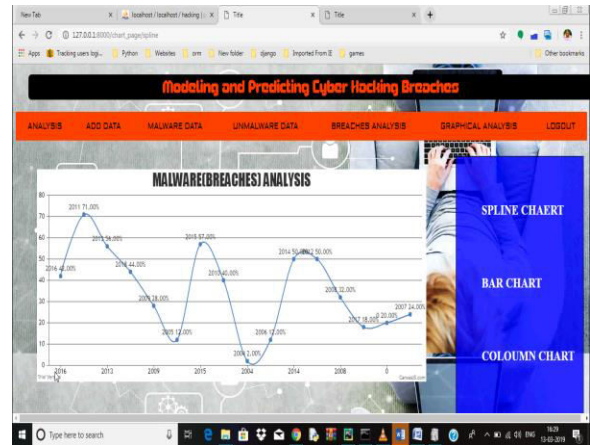
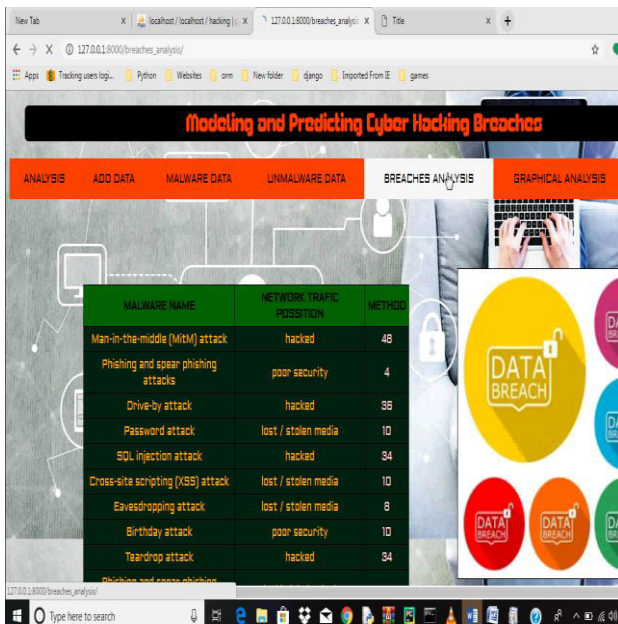
## V RESULTS

### Home Page:

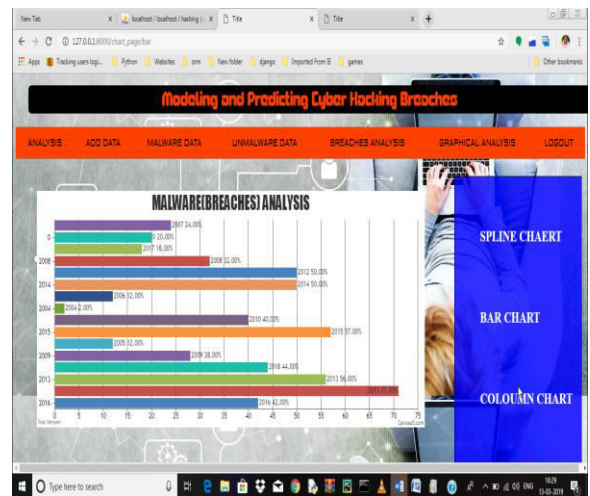


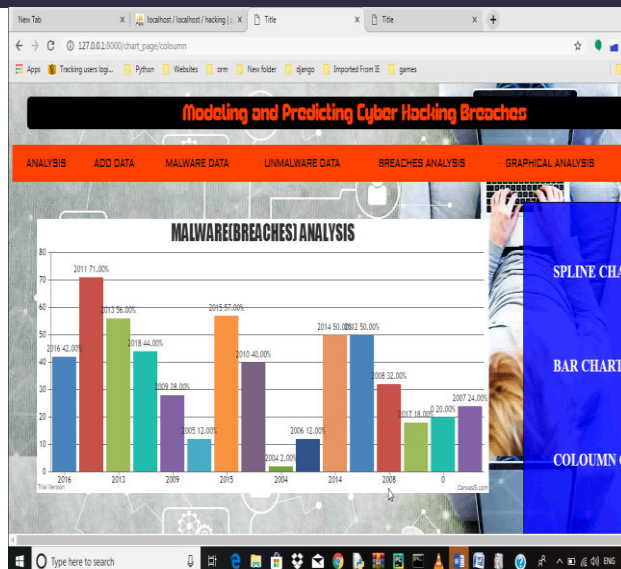



ENTITY	YEAR	RECORDS	ORGANIZATION TYPE	METHOD
21st Century Oncology	2016	2,200,000	healthcare	hacked
Accendo Insurance Co.	2011	175,350	healthcare	poor security
Adobe Systems	2013	152,000,000	tech	hacked
Advocate Medical Group	2013	4,000,000	healthcare	lost / stolen media
Airtel (subsidiary of InMobi)	2018	75,000	advertising	hacked
Affinity Health Plan, Inc.	2009	344,579	healthcare	lost / stolen media
Ameritrade	2005	200,000	financial	lost / stolen media
Ancestry.com	2015	300,000	web	poor security

MALWARE NAME	NETWORK / TRAFFIC POSITION	METHOD
Man-in-the-middle (MITM) attack	hacked	48
Phishing and spear phishing attacks	poor security	4
Drive-by attack	hacked	36
Password attack	lost / stolen media	10
SQL injection attack	hacked	34
Cross-site scripting (XSS) attack	lost / stolen media	10
Eavesdropping attack	lost / stolen media	8
Birthday attack	poor security	10
Teardrop attack	hacked	34





## VI CONCLUSION

We broke down a hacking break dataset from the perspectives of the occurrences between landing time and the rupture measure, and demonstrated that they both ought to be displayed by stochastic procedures as opposed to disseminations. The factual models created in this paper show tasteful fitting and expectation exactnesses. Specifically, we propose utilizing a copula-based way to deal with foresee the joint likelihood that an episode with a specific extent of break size will happen amid a future timeframe. Measurable tests demonstrate that the approaches proposed in this paper are superior to anything those which are exhibited in the writing, in light of the fact that the last disregarded both the fleeting connections and the reliance between the episodes between landing times and the break sizes. We led subjective and quantitative examinations to draw further bits of knowledge. We drew a lot of cybersecurity experiences, including that the danger of digital hacking rupture episodes is for sure

deteriorating as far as their recurrence, yet not the greatness of their harm. The system introduced in this paper can be embraced or adjusted to investigate datasets of a comparable sort.

## VII REFERENCES

- [1] P. R. Clearinghouse. Security Rights Clearinghouse's Chronology of Data Breaches. Gotten to: Nov. 2017. [Online]. Accessible: <https://www.privacyrights.org/information/breaks>
- [2] ITR Center. Information Breaches Increase 40 Percent in 2016, Finds New Report From Identity Theft Resource Center and CyberScout. Gotten to: Nov. 2017. [Online]. Accessible: <http://www.idtheftcenter.org/2016databreaches.html>
- [3] C. R. Focus. Cybersecurity Incidents. Gotten to: Nov. 2017. [Online]. Accessible: <https://www.opm.gov/cybersecurity/cybersecurity-episodes>
- [4] IBM Security. Gotten to: Nov. 2017. [Online]. Accessible: <https://www.ibm.com/security/information-rupture/index.html>
- [5] NetDiligence. The 2016 Cyber Claims Study. Gotten to: Nov. 2017. [Online]. Accessible: [https://netdiligence.com/wp-content/transfers/2016/10/P02\\_NetDiligence-2016-Cyber-Claims-Study-ONLINE.pdf](https://netdiligence.com/wp-content/transfers/2016/10/P02_NetDiligence-2016-Cyber-Claims-Study-ONLINE.pdf)
- [6] M. Eling and W. Schnell, "What do we think about digital hazard and digital hazard protection?" J. Hazard Finance, vol. 17, no. 5, pp. 474– 491, 2016.
- [7] T. Maillart and D. Sornette, "Overwhelming followed circulation of digital dangers," Eur. Phys. J. B, vol. 75, no. 3, pp. 357– 364, 2010.



[8] R. B. Security. Datalosddb. Gotten to: Nov. 2017. [Online]. Accessible: <https://blog.datalosddb.org>

[9] B. Edwards, S. Hofmeyr, and S. Forrest, "Promotion and substantial tails: A more intensive see information ruptures," J. Cybersecur., vol. 2, no. 1, pp. 3– 14, 2016.

[10] S. Wheatley, T. Maillart, and D. Sornette, "The extraordinary danger of individual information breaks and the disintegration of protection," Eur. Phys. J. B, vol. 89, no. 1, p. 7, 2016.

[11] P. Embrechts, C. Klüppelberg, and T. Mikosch, Modeling Extremal Events: For Insurance and Finance, vol. 33. Berlin, Germany: Springer-Verlag, 2013.

[12] R. Böhme and G. Kataria, "Models and measures for relationship in digital protection," in Proc. Workshop Econ. Inf. Secur. (WEIS), 2006, pp. 1– 26.

[13] H. Herath and T. Herath, "Copula-based actuarial model for evaluating digital protection approaches," Insurance Markets Companies: Anal. Actuarial Comput., vol. 2, no. 1, pp. 7– 20, 2011.

## AUTHORS



**Mr. V.SREEKANTH**, VI<sup>th</sup> semester, dept of MCA , Sree Vidyanikethan Institute Of Management from SV University.,(A.P),INDIA .Email ID: sreekanthvathala@gmail.com



**Mr. B.Masthan Baba**, having 11 years of teaching experience. Guided 90 students' projects. Membership in the Professional bodies like CSTA, ACM and ISCA. Attended good number of conferences and workshops. Published 02 Papers in National Conferences and attended 1 Faculty Development programs.



# International Journal for Innovative Engineering and Management Research

*A Peer Reviewed Open Access International Journal*

[www.ijiemr.org](http://www.ijiemr.org)