An Intuitive Way to Unmask In-Browser Cryptojacking in Network Level using Support Vector Machine (SVM) in Machine Learning

Pruthvi Raj Kantamani, Geetha Manoj Potru, Yovan Felix A

ABSTRACT--- cryptocurrency mining is skyrocketed in the recent times. The hackers are turning their heads towards this type of attacks mainly through browser-based mining which makes them undetected. This leads to the increase in Electricity bills and heavy load on the computer processing. This can be avoided by disabling the java script in the device. The disabling Java Script leads to poor user experience of the websites as the websites in recent times are powered with Java Script frameworks. This paper discusses about a live network filtering tool which detects the illicit In-browser miningwithin the range of a network of devices.

Keywords—cryptojacking, SVM, In-browser mining, wireshark, cryptocurrency, illegal cryptomining.

I. INTRODUCTION

Almost a decade ago a cryptocurrency named as Bitcoin [1] has emerged. From that point It has grown to a very popular cryptocurrency by the introduction of transparency in the transactions and security to the block chain. At the beginning of the Bitcoin many people didn't about how huge it will grow. So they used their personal computers for mining. As Technology broadens more people are being introduced to cryptocurrencies, searching for the easy alternatives for mining cryptocurrencies by combining the CPU power of their computers with GPUs. Some other deploy the JavaScript code snippets on to their websites. So, the visitors who visit their websites which contain the code snippets of particular mining JavaScript will get their computing resources used for computing the hashes without the visitor's consent by the owner of the website. The emergence of the application-specific integrated circuits and collective mining pools resulted in growth of mining bitcoins to over 43 Peta hashes. The Pirate Bay [5] which is a torrent index and magnet links provider has been caught using the cryptocurrency mining JavaScript in their website to add a revenue stream. This mining of cryptocurrency by stealing the computing power of CPUs and GPUs of the users is often termed as *cryptojacking* which is also known as coinjacking. Crypto mining is not illegal. Many people tend to mine the cryptocurrencies in their own computers using CPUs and high-end GPUs. It is mining the cryptocurrency by using others computing cycles that makes illegal. This illicit mining is not only confined to the personal computers but also to large enterprises. A recent report has analyzed that over 1 million malicious miners are deployed over a period of 12 years from 2007-2018 [2].

II. LITERATURE SURVEY

A paper on first look at browser-based cryptojacking [7] has explained in a good manner about how the In-browser mining is rampant in the recent days and also educated that there is an appreciable growth in cryptojacking. It has also mentioned about the mining pools by which the attackers will start the cryptojacking campaigns.

Another paper named "A First Look at the Crypto-Mining Malware Ecosystem: A Decade of Unrestricted" [2] state that their analysis over 4.4 million malware samples reveals that campaigns with multi-million earnings, associating over 4.3% of Monero with Illicit mining.The observations from these two papers show that the In-browser cryptomining has taken a steep curve in recent days.

III. GROWTH OF NETIZENS

As the world is moving towards the digitization more individuals are making themselves familiar with the cryptocurrencies or digital currencies as they are more secure than the conventional bills or paper currency. By having the money in the form of digital currency makes not only their money secure but also, they can be used irrespective of location and nation. In addition to that the payments are more transparent so that the change in the transactions is limited to none. By the privileges that digital currency had, more people are turning their heads towards them. So are the hackers. As there is rise in the price and demand of cryptocurrencies, the greedy nature of the malicious people who are making their way to mine the cryptocurrencies increases. In the recent days many national governments are have made the digital currencies legal. By legalizing the digital currencies in their countries, the demand increases. This triggered the rise in the illicit mining of cryptocurrencies.

Revised Manuscript Received on July 10, 2019.

Pruthvi Raj Kantamani, School of Computing Sathyabama Institute of Science and Technology, Chennai, T.N, India. (E-mail: prudhvi.kantamani@gmail.com)

Geetha Manoj Potru, School of Computing Sathyabama Institute of Science and Technology, Chennai, T.N, India. (E-mail: manojpotru9@gmail.com)

Yovan Felix A, School of Computing Sathyabama Institute of Science and Technology, Chennai, T.N, India. (E-mail: yovanfelix@sathyabama.ac.in)

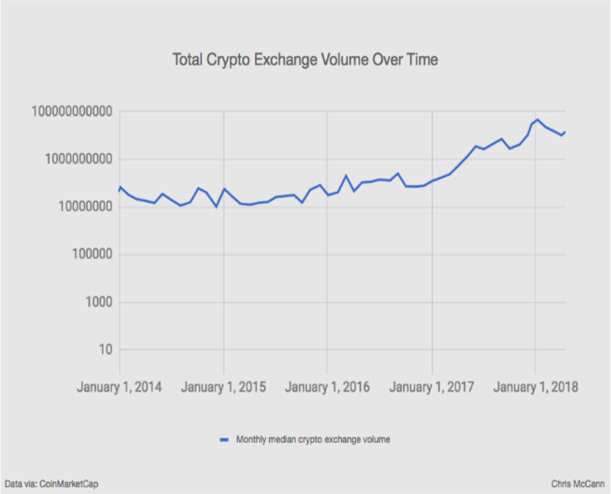


Figure 1.Total Crypto Exchange Volume Over Time. (Source: https://medium.com)

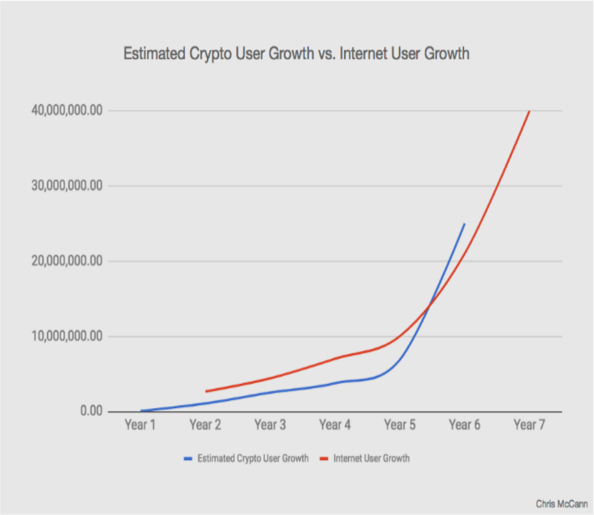


Figure 2.Estimated Crypto User Growth vs. Internet User Growth. (Source: https://medium.com)

IV. MINING MANNER

Mining is termed as a process of record keeping, done through the use of computer processing power [11]. In view of Bitcoin, when a transaction is done, a block is created. In order to process the transaction, the block which was created should be solved. This is the work of miner. Once miner has solved the block the miner will be rewarded with some portion of bitcoins and a transaction fee in terms of Bitcoins.

A. Mining pools

At the beginning of the cryptocurrency mining era, the mining is done with personal computers CPU power. Later as the hashes become more and more the blockchain size increases and made it difficult to mine the cryptocurrencies with personal computers. So, this made to miners turn towards the GPUs which is more reliable than the CPUs. The first ever Bitcoin coin block mined on July 18th, 2010 by a user named ArtForz [3], by using his own created algorithm by himself. Then at the mid of 2011 others started creating and released as open source GPU-based mining tools. These tools are in great help with increased in efficiency for mining. As the competition in submitting hashes increases, the usage of GPUs in the parallel manner also increased. This type of using GPUs in parallel is also known as mining rigs. On reaching mid-2012, companies started selling ASICs (Application Specific Integrated Circuits) which are specifically designed for the cryptocurrency mining purpose. With the introduction of ASICs to the mining, the hash rate is increased drastically. These are very expensive and consume a lot of electricity and generates a lot of heat. Thus the persons who own these high end ASICs join together to form a group often called Mining pools. In these mining pools the blocks are mined together. So, there is a higher chance of obtaining Bitcoins rather than mining individually.

B. Browser-based mining

The browser-based mining is the most common type of illicit mining that takes place remotely in the user's personal computer. Bitcoin plus is one example that can be given for bitcoin browser miners. The browser-based mining provides a decent way to make money from the visitors computing cycles without the user's consent. The rise in the browser-based mining took place with introduction of Bitcoin miners such as JSminer (2011) and Mine Crunch (2014). Although In-browser mining makes a decent money, it is not as much reliable as GPU and ASIC-based mining. The example of Browser-based mining is Monero [4] which is a cryptocurrency alternative to Bitcoin. Monero uses the specifically an algorithm called Crypto Night [6].

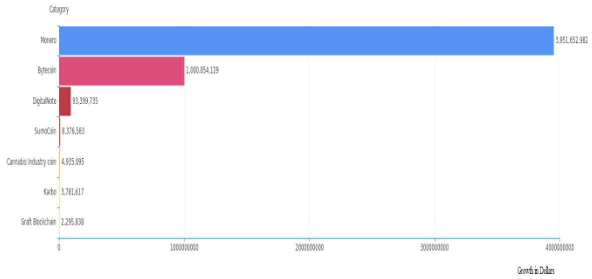


Figure 3.Crypto Night based cryptocurrencies market capitalization, June 2018.

V. RANSOMWARE VS. CRYPTOJACKING & RESULTS

In Ransomware the attacker encrypts the user's data and asks for ransom. If the user pays the ransom, the data may be decrypted. If the user does not have any confidential data or data that is not important, there is a chance that user would not pay the ransom to the attacker. Hence, it might not be profitable for the attacker. On the other hand, cryptojacking makes use of every device. The attacker uses the device's computing power and use it to mine cryptocurrencies. There is very less chance that the mining is being detected in the user's system. Unlike Ransomware, the cryptojacking is hard to detect. So, the hackers are making their way towards the cryptojacking.

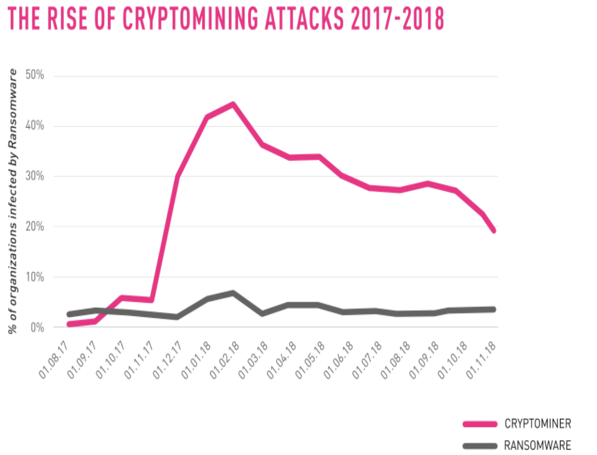
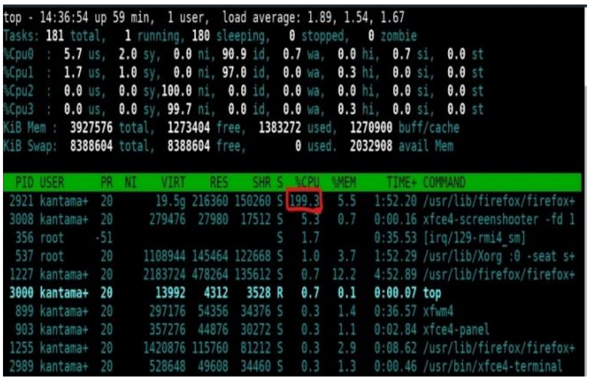


Figure 4.The rise of cryptomining attacks between 2017 and 2018 (Source: https://www.bleepingcomputer.com)

OBSERVATIONS

The cryptojacking is a simple passive attack yet a serious considerable issue that make use of other user's computing cycles without their concern. In order to stop cryptojacking in the user's computer with in browser cryptojacking, we have to disable the JavaScript in the browser settings. Most of the websites now a days are being developed by using Java Script. This Java Script is most dominant in the MEAN stack, which uses the Java Script frameworks as both Front end and back end. Disabling Java Script makes the user Interface look bad. When a particular website that include an In-browser miner is accessed in the web browser. The miner runs in background and starts using the user's computing resources to perform complex mathematical calculations. These complex mathematical calculations need a large computing resources. So, there is a sudden hike in the CPU usage which results in over heating issues and makes the CPU overloaded. This is a serious issue when coming to the enterprise level which has a large network of devices.



```
top - 14:36:54 up 59 min,  1 user,  load average: 1.89, 1.54, 1.67
Tasks: 181 total,   1 running, 180 sleeping,   0 stopped,   0 zombie
%Cpu0  :  5.7 us,  2.0 sy,  0.0 ni, 90.9 id,  0.7 wa,  0.0 hi,  0.7 si,  0.0 st
%Cpu1  :  1.7 us,  1.0 sy,  0.0 ni, 97.0 id,  0.0 wa,  0.3 hi,  0.0 si,  0.0 st
%Cpu2  :  0.0 us,  0.0 sy,100.0 ni,  0.0 id,  0.0 wa,  0.0 hi,  0.0 si,  0.0 st
%Cpu3  :  0.0 us,  0.0 sy, 99.7 ni,  0.0 id,  0.0 wa,  0.3 hi,  0.0 si,  0.0 st
KiB Mem :  3927576 total,  1273404 free,  1383272 used,  1270900 buff/cache
KiB Swap:  8388604 total,  8388604 free,        0 used.  2032908 avail Mem

  PID USER      PR  NI    VIRT    RES    SHR S  %CPU %MEM     TIME+ COMMAND
 2921 kantama+  20   0   19.5g 216360 150260 S 199.3  5.5   1:52.20 /usr/lib/firefox/firefox+
 3008 kantama+  20   0  279476  27980  17512 S   5.3  0.7   0:00.16 xfce4-screenshooter -fd 1
  356 root     -51                           S   1.7          0:35.53 [irq/129-rmi4_sm]
  537 root      20     1108944 145464 122668 S   1.0  3.7   1:52.29 /usr/lib/Xorg :0 -seat s+
 1227 kantama+  20     2183724 478264 135612 S   0.7 12.2   4:52.89 /usr/lib/firefox/firefox+
 3000 kantama+  20       13992   4312   3528 R   0.7  0.1   0:00.07 top
  899 kantama+  20      297176  54356  34376 S   0.3  1.4   0:36.57 xfwm4
  903 kantama+  20      357276  44876  30272 S   0.3  1.1   0:02.84 xfce4-panel
 1255 kantama+  20     1420876 115760  81212 S   0.3  2.9   0:08.62 /usr/lib/firefox/firefox+
 2989 kantama+  20      528648  49608  34460 S   0.3  1.3   0:00.46 /usr/bin/xfce4-terminal
```

Figure 5.Hike in CPU usage (seen as 199%) when accessing the particular website that has miner script.

SYSTEM ARCHITECTURE

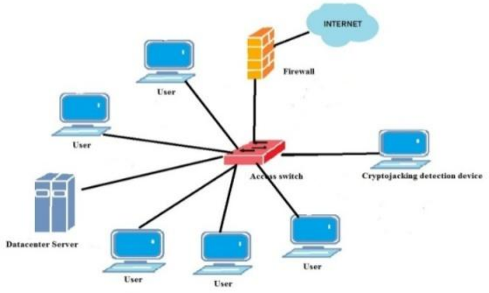


Figure 6.picture of Architecture detecting cryptojacking in the network of devices

PROPOSED WORK

This work is done using Machine Learning approach in which we train the model with a certain predefined data sets and also with the packets that are captured from the Wireshark [8]. As a start, the TCP (Transmission Control Protocol) packets are to be captured from the wireless interface or wire connection as well. The model is trained with datasets that are from different sources like general browsing, GPU mining and CPU mining. The packets captured from the Wireshark are used to detect whether the mining is being done or not in the particular device. There is also a live filtering model which is used to detect the in-browser mining on the go.

A. SVM (Support Vector Machine):

Support Vector Machine (SVM) is one of the most popular and effective classification algorithms and has attracted much attention in recent years [9]. The SVM (support vector machine) algorithm is used for classification. This classification is primarily used to group data. The classification is important to distinguish between the TCP packets that are transferred from the infected web sites. It is used to find the optimal hyperplane between two classes. This gives the more generalization ability to distinguish. The separating hyperplane is determined by using an important sample in the training datasets. Instead of training SVM model on the whole training data set, a small subset of dataset was selected and this subset will act as a final training set. The K-means clustering helps in selecting the most informative samples and the SVM classifier gets trained on the dataset that we provide.

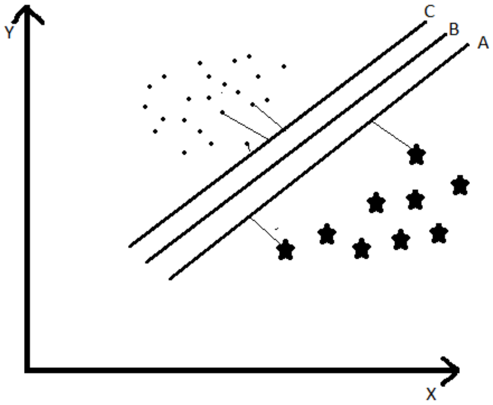


Figure 7.depicting the concept of SVM Algorithm

B. Capturing TCP packets with Wireshark:

By using Wireshark, we can capture the packets of information when the user is browsing. By deploying the TCP filter the TCP packets alone get filtered into a file with extension of '.pcapng'. The TCP packets alone get captured from the given interface that may be wired or wireless interface.

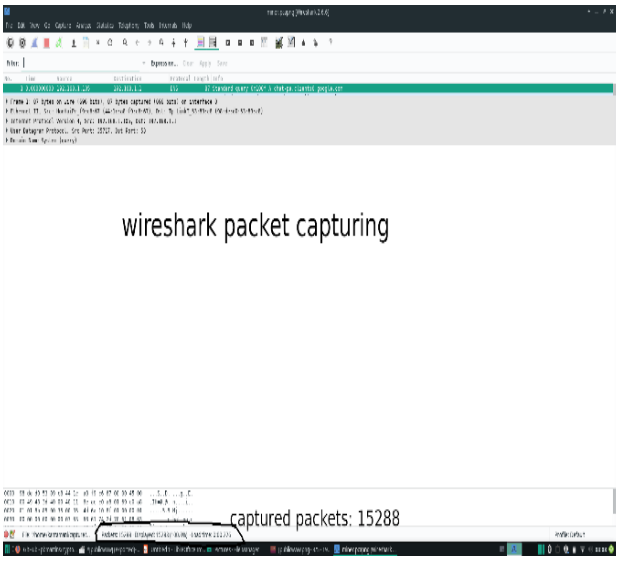
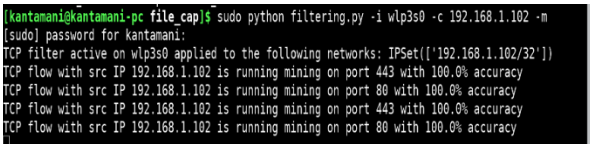


Figure 8. depicting the capturing of packets using Wireshark.

This file has to be parsed for the TCP packets. After capturing the TCP packets, it is to be converted to a readable format (.dat) file which is used to train the model. The model is to be trained by using either SVM or Neural Networks. By using the SVM algorithm the classification of the data gets much easier. By this model we can check that computing resources are being mined or not. As in Figure 6 explains the filtering can be applied on the network of devices. So, the device in which the mining is being done can be known.

C. Live filtering Model:

In the live filtering model, the classification is done using already trained model, In the live filtering model the network with a set of devices is analyzed for mining in the browser. This filter is applied on the set of IP addresses in the network. If there is any illicit mining is being done it will notify on what IP address and port the mining is being done. This is very helpful in detection of mining with a range of one device to a large number of devices that are in the same network. In figure 8 the TCP filtering is enabled on the wireless Interface and IP address is given. After detection of the cryptojacking on the device, the network administrator or the user can terminate the process that is illegally using the computing resources to mine cryptocurrencies.



```
[kantamani@kantamani-pc file_cap]$ sudo python filtering.py -i wlp3s0 -c 192.168.1.102 -m
[sudo] password for kantamani:
TCP filter active on wlp3s0 applied to the following networks: IPSet(['192.168.1.102/32'])
TCP flow with src IP 192.168.1.102 is running mining on port 443 with 100.0% accuracy
TCP flow with src IP 192.168.1.102 is running mining on port 80 with 100.0% accuracy
TCP flow with src IP 192.168.1.102 is running mining on port 443 with 100.0% accuracy
TCP flow with src IP 192.168.1.102 is running mining on port 80 with 100.0% accuracy
```

Figure 9. shows the working of filtering model.

In this live filtering model, a set of IP addresses of the devices in the network which are to be detected are to be given as arguments and also to which interface (here wlp3s0 is wireless interface) the filter is to be activated should also be given. Then based on the TCP packets transferring through the port which uses the HTTP and HTTPS requests, the model will return what are the devices that are undergoing illegal crypto mining can be detected.

CONCLUSION

The machine learning concept has a lot of scope when coming to the intrusion detection in the network. The attackers would not stop the illegal usage of others computing cycles but they may change the way in the stealing them for the mining purposes as the usage of cryptocurrencies will raise in the future. There is an urgent need for which the devices are to be maintained secured from the attackers. There are a lot of people out there that do not know that their precious computing cycles are being stolen. This paper shows a way in which the devices in the network can be made free from illegal browser mining. The virtual space "Cloud" is on the other side of the coin that makes the most of the advancements in the recent times. There is no wonder that the attackers might steal the computing cycles of the heavy processors that are situated in some virtual location and connected to the Internet. The Internet is not Immune to the attackers. It is what the users aware of makes the Internet Immune.

REFERENCES

1. Satoshi Nakamoto. Bitcoin: A Peer-to-Peer Electronic Cash System, 2009.
2. Sergio Pastrana, Guillermo Suarez-TangilKing, Title: A First Look at the Crypto-Mining Malware Ecosystem: A Decade of Unrestricted Wealth, dated: 3 Jan 2019.
3. Important milestones of the bitcoin project. https://en.bitcoin.it/wiki/Category:History, 2009.
4. MalteMöser, Kyle Soska, Ethan Heilman, Kevin Lee, Henry Heffan, Shashvat Srivastava,Kyle Hogan, Jason Hennessey, Andrew Miller, Arvind Narayanan, and Nicolas Christin. An Empirical Analysis of Traceability in the Monero Blockchain, published in Proceedings on Privacy Enhancing Technologies in 2018.

5. In-browser Cryptojacking at full throttle – A report by Quick Heal Security Labs written by Prashanth Kadam. (https://blogs.quickheal.com/browser-cryptojacking-full-throttle-report-quick-heal-security-labs/) on January 23, 2018
6. Nicolas van Saberhagen, CryptoNote v 2.0, October 17, 2013.
7. ShayanEskandari, Andreas Leoutsarakos, Troy Mursch, Jeremy Clark, Bad Packets Report. A first look at browser-based cryptojacking, published in IEEE SECURITY & PRIVACY ON THE BLOCKCHAIN (IEEE S&B) on 7 Mar 2018.
8. Usha Banerjee, AshutoshVashishtha and MukulSaxena, Evaluation of the Capabilities of Wireshark as a tool for Intrusion Detection, International Journal of Computer Applications in 2010
9. Durgesh K. Srivastava, Lekha Bhambhu, Data Classification using Support Vector Machine, Journal of Theoretical and Applied Information Technology, 2009
10. W. Wang, B. Ferrell, X. Xu, K. W. Hamlen, and S. Hao, “SEISMIC:SEcure In-lined Script Monitors for Interrupting Cryptojacks,” in European Symposium on Research in Computer Security. Springer, 2018, pp. 122–142.
11. What is Bitcoin Mining, https://cointelegraph.com/bitcoin-for-beginners/what-is-mining
12. D. Y. Huang, H. Dharmdasani, S. Meiklejohn, V. Dave, C. Grier, D. McCoy, S. Savage, N. Weaver, A. C. Snoeren, and K. Levchenko, “Botcoin: Monetizing stolen cycles.” in NDSS. Citeseer, 2014.
13. M. Musch, C. Wressnegger, M. Johns, and K. Rieck, “Web-based cryptojacking in the wild,” arXivpreprint arXiv:1808.09474, 2018.
14. A. Vamshi, “Netskope: Technical analysis of Xbooster parasitic Monero miner,” May 2018. [Online]. Available: https://perma.cc/8RZG-5QBS
15. J. Grunzweig, “Palo Alto Networks: The Rise of the Cryptocurrency Miners,” June 2018. [Online]. Available: https://perma.cc/4VZL-J45Q
16. Gladence, L. Mary, M. Karthi, and V. Maria Anu. "A statistical comparison of logistic regression and different Bayes classification methods for machine learning." ARPN Journal of Engineering and Applied Sciences 10, no. 14 (2015): 5947-5953.
17. Y. Bevish Jinila , K. Komathy (2013),” A privacy preserving authentication framework for safety messages in vanet”, 4th International Conference on Sustainable Energy and Intelligent System (SEISCON 2013), December 12-14, 2013, pp. 456-461, IET
18. Karthika, J. K., V. Maria Anu, And A. Veeramuthu. "An Efficient Attribute Based Cryptographic Algorithm For Securing Trustworthy Storage And Auditing For Healthcare Big Data In Cloud." (2006).
19. Rajalakshmi, V., M. Lakshmi, and V. Maria Anu. "A Complete Privacy Preservation System for Data Mining Using Function Approximation." J. Web Eng. 16, no. 3&4 (2017): 278-293