International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-6S, August 2019

Malicious Traffic Detection System using Publicly Available Blacklist's

Sudarshan N, P.Dass

Abstract— In this fastest growing technology with the increase in internet usage, the communication became much faster and easier which resulted in the massive growth in digitalization. With this the cyber crimes were increasing day-by-day . They employ every possible technique and trick to make the users as zombies for their malicious activities or Crypto mining. In recent years we are facing issues with ransomware' which result in the loss of data integrity and confidentiality along with our privacy and anonymity. The malware' can spread all over the network within no time. Using anti virus programs alone for safeguarding our network is a bad practice because they filter the traffic on signature based. Here problem is if the user is not up to date with the definitions from the AV provider, then he will be prone to the attack. In this model a system to track malicious trails in a network is done. This employs online malware detection system (Virus Total) and open source dynamic black lists which contain malware or suspicious programs along with some static pre compiled blacklists from different antivirus providers and our own definitions of block to filter the traffic which gives the detailed log report on the suspicious trails, this is from domain name or IP address or malicious scripts in the webpage.

Keywords—Cyber crimes, malware, VirusTotal, black lists, Filtering, log report

I. INTRODUCTION

Malicious code, unremarkably named "malware," persistently displays one in all the highest security worries for associations. Worldwide cash connected misfortunes thanks to malware is quite several bucks per annum. Average malware incorporates infections, worms, Trojan steeds, spyware, adware, and others. Since the primary computer infection surfaced within the mid-1980s, malware has fashioned into thousands of variations that change in contamination system, unfold element, dangerous payload, and completely different highlights. Among them, infections and worms, the 2 most frequently ascertained forms of malware, have drawn a lot of business and analysis thought than different types of malware due to their self-imitating nature, sensational proliferation speed, and conceivably extreme dangerous results.

Infections can occur in many forms through many hierarchical methods in this generation. Many advancements in the programming languages and inheritance is being changed from the program to program. This makes the detection hard to the anti virus programs including with the wireless devices in the network and the local area network devices this makes the data loss in an organization in highest form with the infection. The infection includes the switches, routers and firewalls as well. Some malware might unfold through innovative systems.

It should be in any kind damages the infrastructure's information confidentiality. This project aims to observe malicious traffic in any forms like malware, DOS, mass scans and different suspicious tries, that inward to the network, supported publically offered blacklists from completely different resources. It aims to dam the traffic, which can hurt the infrastructure in either of the forms mentioned higher than.

The sections in the paper are like this the section two offers the summary on literature survey, section three offers the look and implementation of the project details, section four offers the small print on differing types of outputs supported the log report, section five concludes the paper.

II. LITERATURE REVIEW

Cova et al [1] executed (JSAND) is an instrument used for identifying, breaking down pernicious site pages. The apparatus utilized ten highlights to portray considerable occasions of a download-when-browsing. These are chosen dependent by means that typically follows doing a assault. What's more, the highlights were classified into essential highlights that are required for a fruitful endeavor, for example, the quantity of instant segments (modules) and valuable highlights that are not entirely required to dispatch an effective assault, for example, how many times we are being redirected to the different URLs. All together in a framework to choose if the situation given happened component esteem happened, it doled out a likelihood score to each element esteem dependent on a few models.

Rima Masri [2] proposed and actualizes a framework for naturally recognizing pernicious ads. It has three various web malware domain name discoveries frameworks for malignant notices recognition its activities and reports the quantity of distinguished noxious notices utilizing every framework.

Manuscript published on 30 August 2019.
***** Correspondence Author (s)
Sudarshan N, Bachelor's student,Electronics and Communication Engineering, Saveetha School of Engineering, Chennai, Tamilnadu, India (email: 1nasum.sudharshan@gmail.com)
P.Dass, Assistant Professor, Electronics and Communication Engineering, Saveetha School of Engineering, Chennai, Tamilnadu, India (email: pdass@saveetha.com)

What's more, we contemplate the efficiency of every framework by ascertaining the perplexity network and exactness. We find that URLVoid is the best as far as exactness (73%) on the grounds that it utilizes a blend of surely understood site scanners and space boycotts.

Xing et al. [3] contemplated the likelihood of encouraging Malvertisements arrangement through program augmentations. They created Expector, which is anestimation system used to recognize program augmentations that infuse promotions in website pages. Expector was planned dependent on various things, for example, distinguishing the sites that may prompt the promotion infusion functionalities and activating occasions that expansion may be keen on for advertisement infusion. They executed Exclusion on right around Eighteen thousand Google's Chrome browser augmentations and identified two hundred and ninety two expansions that offers promotion infusion. They concluded that fifty six expansions resulted to malicious destinations. At long last, the paper reasoned that utilizing augmentations with advertisement infusing property may prompt Malvertising dangers.

Portage et al. [4] planned and actualized an apparatus called OdoSwiff to fragment down Flash substance and identify malevolent practices dependent on specific qualities, for example, powerful internet browser cross site scripting's. He utilized non dynamic investigation module to extract the contents of Flash format files and dynamic examination module to run the Flash file application and noted down the running flow. The device was allowed to get a major gathering of Flash format of files that display pernicious practices. The outcomes demonstrated that author could recognize pernicious Flash ads bitterly contrasted and existing frameworks that examine Flash applications.

HONG GUO [5] This paper receives hypothetical ideas and techniques from the field of informal community examination, to be specific, area measures and sub divided the investigation, to identify the basic attributes of the interpersonal organization along with mechanical system. Specifically, in light of the one of a kind results of the malware engendering procedure, author distinguish irregular way between the suitable main area measure to assess the basic post of one of the stand-alone hubs. Measurement quality augmenting deterioration is then connected to break down the installed subgroup structure. In light of the determined main area execution data and they found one of the group structure, we figure the basic hazard demonstration to look over the effect on one of the, aggregate , the system based attributes on malicious script engendering elements. Subtleties on auxiliary hazard demonstrate gave in exploration show segment.

III. DESIGN AND IMPLEMENTATION

3.1 Architecture of the proposed project

Maltrail will depends on the architecture with the flow like Traffic which is the real-time and then the sensor which monitorzs the traffic in the device and then the server which then logs the results. Sensor is an stand-alone script which is running in the end devices in moniter mode to inspect the packets which are moving in and out. It mainly works with an principle of PCapy an python script to inspect the web traffic. This then controls the low of logs to the server where the black listed traffic have to be necessarily farwarded to the server which is the centralized one accurately for example an honeypots. If there should be an occurrence of a match that triggers, it forwards the occasion results to centralized local Server. In the event that Sensor is being kept running on an un identified machine from Server (default setup), logs are put away specifically into the neighborhood logging index. Else, they will be forwarded to the server in the method od user data gram messages to server.

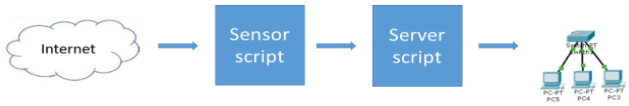


Figure 1 Architecture

The various steps involved in project for compiling all packets for malicious scripts or other suspicious activities. This is shown in figure 2 below. This includes the steps which will the packet inspection undergoes. The first step will be the packet capturing based on the standalone script that is running in the device, generally this will be running in the core router or after the firewall for mirroring the traffic. The second step will be the defragmentation of the packet to separate the packet's source, destination IP, source, destination port. This will give us enough understanding about the packet whether it can be trusted or not.

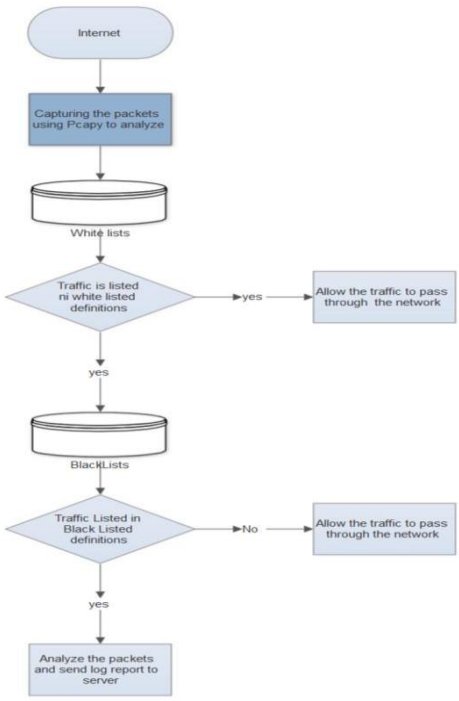


Figure 2 Flow chart

3.2 Black lists

The black lists contain the VirusTotal and open source available (black) lists which contain harmful scripts or actions or suspicious triggers, and for this in addition we use the blocked websites data and IP addresses from various anti-virus providers specific definitions in the text files which we are going to interpret with our sensor. This also contains the white lists those can be sent into the network without delaying it with filtering process. This contains the list of websites that are reported as malicious and the websites which provide the anonymizing services using proxy chains, this can be blocked with the help of most commonly used TOR exit nodes. The websites which provide dynamic IP while requesting services from the internet are also kept under suspicious category to improve our security over the network. This also includes the websites which provide direct downloads while visiting the webpage are also blocked, this excludes some well known websites that allow direct downloads like google, amazon, skype, firefox etc., This definitions will gets updated within the interval of time so that we will be getting better results. Suspicious domains which have slightly change in their domain name can also be blocked for example g00gle, amazn, flipkrat, etc., This can be availed from various AV vendors and publicly available databases which are open source so it contains better results all over the world.

3.3 Sensor

This is a standalone script written in python programming language which uses the python framework known as Pcapy for capturing the data packets in the network core layer which are inbounded into the network. This uses the set of algorithms for compiling the data packets

Sensor algorithm
import pcapy
capture packets
for each packet
do
if domain_name then
split url (.) into array[]
if array[] in whitelist[]
retuen true
else send packet to array1[]
if array1[] in black-list[]
check for (query,sec,source_ip.source_port,source,destination,dstination_ip,dstination_port,proto)
send to server
if packet has multiple requests
log to mass scans
if packet has multiple protocols
log to mass scans
if packet has dynamic ip
log to suspicious
if packet has multiple dst_ports
log to mass scans
if packet has executable
log to direct_downloads
else
allow packet to pass
end

3.4 Server

The logs which we obtained from the sensor will gets stored in the location which we are specifically included in our script. The obtained logs must be presented infront of the admin so that he can be able to get the prior understanding on what is happening in the network. So that we use web based client access for the results. The server can be the another host were the SIEM exists or it can be the host in which it is running. This will be different in different cases. Where the organization will have tens of hosts don't require SIEM. They can deploy the server in the host itself and access it via web client. It gives the results based on the appearance time, host, and source and destination IP's along with port numbers. This even provide us the severity of the traffic. This also gives us the information about the IP using the search engines like DuckDuckGo which is a privacy focused search engine.

Server algorithm
enable authentication
if auth_details==true
import logs
plot_graphs(severity)
do query(domain)
update(details)
if port_scans>usual
return (scan_details)
if ip=dynamic
scan (src_ip) with query
if dst_ports are dynamic with in the session
return suspicious_ports
else
end

IV. RESULTS AND DISCUSSION

4.1 Client authentication

This server reports must be monitored via web interface by the admin. So there must be some authentication to secure the log reports being erased by the intruders or malware. This can be done by using the authentication given to the admin for analysis and alerting purpose

Figure 3 Client login

4.2 Dashboard for analysis

The client is given the detailed report on the analysis done to the traffic by filtering the severity and the number of threats are given in front of the user with the login done. The graph showing all the types of attacks and the threats were given in detail.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Figure 4 Dashboard

4.3 Threat details

Threat details including its source and destination details along with the destination and source ports will be given. This will also be having the number of packets its filtered and marked as the threats.

Figure 5 Threat details

4.4 Reverse DNS Lookup

The DNS lookup is done with the IP to check its severity and its details like when and where it is hosted in what place and what time, this will be helpful to block the whole content from the specific web server.

Figure 6 Reverse DNS lookup

4.5.Mass scans detection

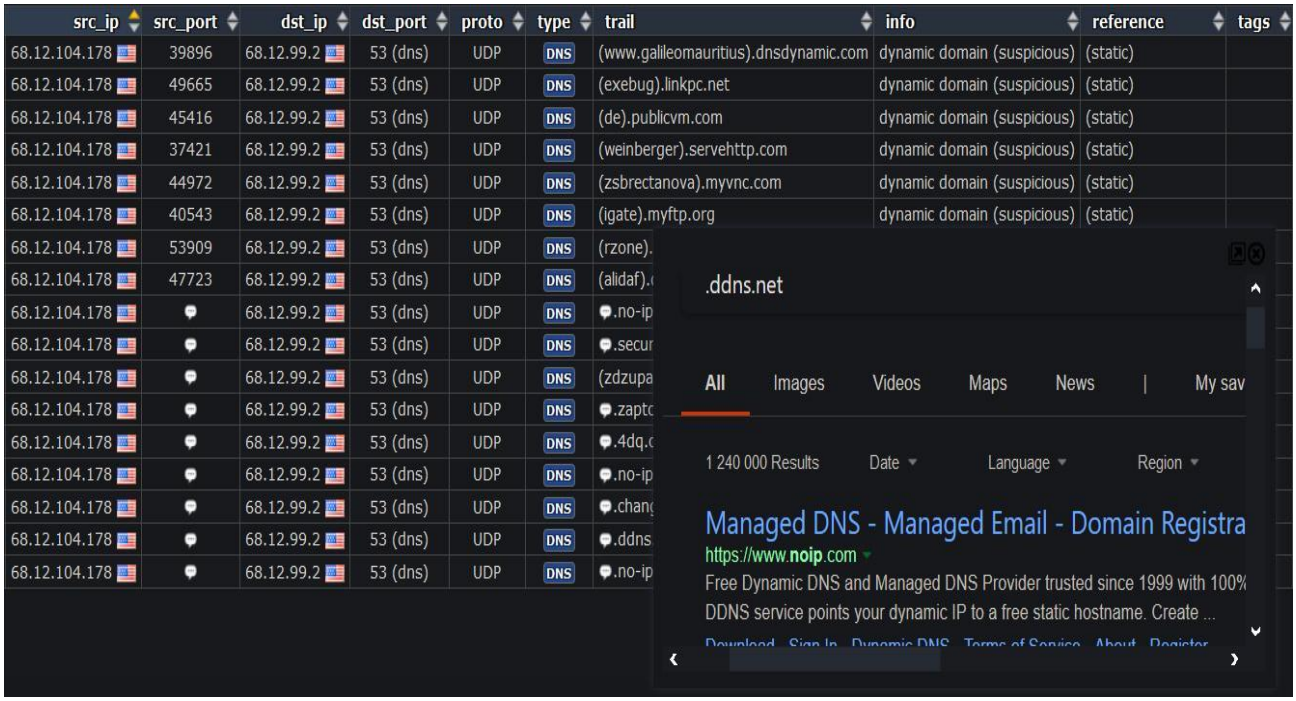
threat	sensor	events	severity	first_seen	last_seen	sparkline	src_ip	src_port	dst_ip	dst_port	proto	type	trail	info	reference
35efdb6d	r2d2	8	medium	10th 02:49:36	10th 08:40:00		61.240.144.64	60000	68.12.104.178		TCP	IP	61.240.144.64	incoming masscan detected	autoshun.org
7df13688	r2d2	8	medium	10th 00:57:09	10th 09:49:21		61.240.144.65	60000	68.12.104.178		TCP	IP	61.240.144.65	incoming masscan detected	autoshun.org
75f739bf	r2d2	10	medium	10th 00:17:33	10th 10:19:11		61.240.144.66	60000	68.12.104.178			IP	61.240.144.66	incoming masscan detected	autoshun.org
1ce3dc3a	r2d2	2	medium	10th 07:33:47	10th 07:33:47		62.210.248.159	5114	68.12.104.178	5060 (sip)	UDP	IP	62.210.248.159	sipvicious scan	autoshun.org
f1f65e35	r2d2	2	medium	10th 01:31:01	10th 01:31:01		178.162.201.166	27926	68.12.104.178	5060 (sip)	UDP	IP	178.162.201.166	sipvicious scan	autoshun.org
8ba8f62a	r2d2	2	medium	10th 09:28:33	10th 09:28:33		178.162.211.207	5277	68.12.104.178	5060 (sip)	UDP	IP	178.162.211.207	sipvicious scan	autoshun.org
0932a55a	r2d2	2	medium	10th 08:15:19	10th 08:15:19		192.198.115.139	5081	68.12.104.178	5060 (sip)	UDP	IP	192.198.115.139	sipvicious scan	autoshun.org
bbaaca06	r2d2	2	medium	10th 06:59:01	10th 06:59:01		199.189.87.8	5071	68.12.104.178	5060 (sip)	UDP	IP	199.189.87.8	sipvicious scan	autoshun.org
7dc60924	r2d2	4	medium	10th 00:17:55	10th 07:38:47		199.217.116.159		68.12.104.178	5060 (sip)	UDP	IP	199.217.116.159	sipvicious scan	autoshun.org
704d1461	r2d2	2	medium	10th 07:46:49	10th 07:46:49		212.83.132.65	5119	68.12.104.178	5060 (sip)	UDP	IP	212.83.132.65	sipvicious scan	autoshun.org
7cb0a54f	r2d2	2	medium	10th 08:44:00	10th 08:44:00		212.83.171.94	5109	68.12.104.178	5060 (sip)	UDP	IP	212.83.171.94	sipvicious scan	autoshun.org

Figure 7 Mass scans

Mass scans will be detected on the network based on the known web based IPs. This will save us from attackers who were doing the reconnaissance and information gathering.

4.6 Dynamic domain scanning

The websites which will under go with all the dynamic IPs for accessing the exit nodes of the network will be blocked with the help of TOR exit nodes.

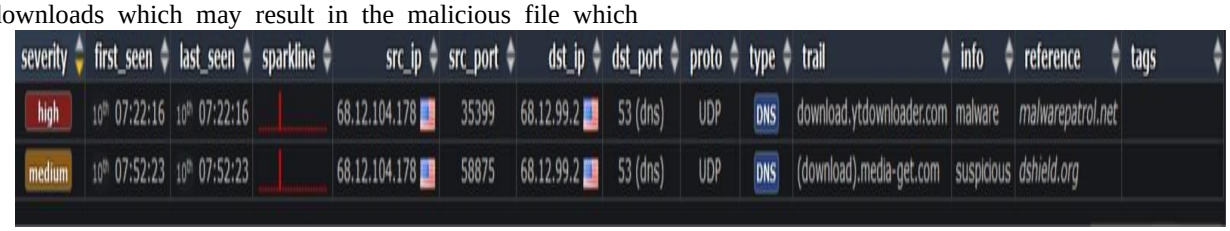


src_ip	src_port	dst_ip	dst_port	proto	type	trail	info	reference	tags
68.12.104.178	39896	68.12.99.2	53 (dns)	UDP	DNS	(www.galileomauritius).dnsdynamic.com	dynamic domain (suspicious)	(static)	
68.12.104.178	49665	68.12.99.2	53 (dns)	UDP	DNS	(exebug).linkpc.net	dynamic domain (suspicious)	(static)	
68.12.104.178	45416	68.12.99.2	53 (dns)	UDP	DNS	(de).publicvm.com	dynamic domain (suspicious)	(static)	
68.12.104.178	37421	68.12.99.2	53 (dns)	UDP	DNS	(weinberger).servehttp.com	dynamic domain (suspicious)	(static)	
68.12.104.178	44972	68.12.99.2	53 (dns)	UDP	DNS	(zsbrectanova).myvnc.com	dynamic domain (suspicious)	(static)	
68.12.104.178	40543	68.12.99.2	53 (dns)	UDP	DNS	(igate).myftp.org	dynamic domain (suspicious)	(static)	
68.12.104.178	53909	68.12.99.2	53 (dns)	UDP	DNS	(rzone).			
68.12.104.178	47723	68.12.99.2	53 (dns)	UDP	DNS	(alidaf).			
68.12.104.178		68.12.99.2	53 (dns)	UDP	DNS	.no-ip			
68.12.104.178		68.12.99.2	53 (dns)	UDP	DNS	.secur			
68.12.104.178		68.12.99.2	53 (dns)	UDP	DNS	(zdzupa			
68.12.104.178		68.12.99.2	53 (dns)	UDP	DNS	.zaptc			
68.12.104.178		68.12.99.2	53 (dns)	UDP	DNS	.4dq.(			
68.12.104.178		68.12.99.2	53 (dns)	UDP	DNS	.no-ip			
68.12.104.178		68.12.99.2	53 (dns)	UDP	DNS	.chan(			
68.12.104.178		68.12.99.2	53 (dns)	UDP	DNS	.ddns			
68.12.104.178		68.12.99.2	53 (dns)	UDP	DNS	.no-ip			

Figure 8 Dynamic domains filtering

4.7 Direct File downloads filtering

The sites will acquire and option known as one click downloads which may result in the malicious file which results in compromising the target within no time this must be avoided by using this filter



severity	first_seen	last_seen	sparkline	src_ip	src_port	dst_ip	dst_port	proto	type	trail	info	reference	tags
high	10th 07:22:16	10th 07:22:16		68.12.104.178	35399	68.12.99.2	53 (dns)	UDP	DNS	download.ytdownloader.com	malware	malwarepatrol.net	
medium	10th 07:52:23	10th 07:52:23		68.12.104.178	58875	68.12.99.2	53 (dns)	UDP	DNS	(download).media-get.com	suspicious	dshield.org	

Figure 9 Direct file download filter

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

V. CONCLUSIONS

In this paper we conclude that the design of malicious traffic control system will reduce the infrastructure damage the threats that come from the internet or intranet by utilizing publicly available black lists. In this way of approach we are not relying on the signatures that are being given to the end user by the AV provider. Because we are updating our definitions with various AV reports and our custom definitions to block the content.Network Security is a field of study which offers infinite possibilities for research and development. There has been a constant and remarkable progress in this field due to these possibilities. The current study also provides some aspects that can be improved.

REFERENCES

1. M. Cova, C. Kruegel, and G. Vigna, "Detection and analysis of driveby-download attacks and malicious javascript code," in Proceedings of the 19th international conference on World wide web. ACM, 2010, pp. 281-290
2. Automated Malicious Advertisement Detection using VirusTotal, URLVoid, and TrendMicro Rima Masri? and Monther Aldwairi?† ?College of Technological Innovation
3. X. Xing, W. Meng, B. Lee, U. Weinsberg, A. Sheth, R. Perdisci, and W. Lee, "Understanding malvertising through ad-injecting browser extensions," in Proceedings of the 24th International Conference on World Wide Web, ser. WWW '15. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee, 2015, pp. 1286-1295. [Online]. Available: https://doi.org/10.1145/2736277.2741630
4. S. Ford, M. Cova, C. Kruegel, and G. Vigna, "Analyzing and detecting malicious ?ash advertisements." in ACSAC, 2009, pp. 363-372.
   Impact of Network Structure on Malware Propagation: A Growth Curve Perspective HONG GUO, HSING KENNETH CHENG, AND KEN KELLEY Intrusion Prevention/ Intrusion Detection System (IPS/IDS) for Wi-Fi networks Michal, Jaroslav, Frantisek International Journal of Computrt Networks & Communications (IJCNC) Vol.6 , No.4,July 2014
5. J. Segura. (2013, Dec) Malvertising and the joys of online advertising. [Online]. Available: https://blog.malwarebytes.com/threatanalysis/2013/12/malvertising-and-the-joys-of-online-advertising/
6. Virustotal scanner. [Online]. Available: http://securityxploded.com/virus-total-scanner.php
7. M. Aldwairi, "Hardware ef?cient pattern matching algorithms and architectures for fast intrusion detection." Ph.D. dissertation, North Carolina State University, 2006.
8. M. Kharbutli, M. Aldwairi, and A. Mughrabi, "Function and data parallelization of wu-manber pattern matching for intrusion detection systems," Network Protocols and Algorithms, vol. 4, no. 3, pp. 46-61, 2012.
9. Vikram. Urlvoid online scanner to ?nd if a website is safe to visit. [Online]. Available: http://www.technorms.com/1284/urlvoidonline-scanner-to-?nd-if-a-website-is-safe-to-visit
10. What is urlvoid? [Online]. Available: http://www.urlvoid.com/about-us/ [20] Site safety center. [Online]. Available: https://global.sitesafety.trendmicro.com
11. K. Markham. (2014, March) Simple guide to confusion matrix terminology. [Online]. Available: http://www.dataschool.io/simpleguide-to-confusion-matrix-terminology/
12. Chen, A.; Lu, Y.; Chau, P.Y.K.; and Gupta, S. Classifying, measuring, and predicting users' overall active behavior on social networking sites. Journal of Management Information Systems, 31, 3 (2014), 213-253. .
13. Chen, P.; Cheng, S.; and Chen, K. Optimal control of epidemic information dissemination over networks. IEEE Transactions on Cybernetics, 44, 12 (2014), 2316-2328.
14. Matook, S.; Cummings, J.; and Bala, H. Are you feeling lonely? The impact of relationship characteristics and online social network features on loneliness. Journal of Management Information Systems, 31, 4 (2015), 278-310.
15. EC-Council, Ethical Hacking and Countermeasures Version 6 Module XVII Web Application Vulnerabilities: International Council of E-commerce Consultants.
16. R. E. Overill, "ISMS insider intrusion prevention and detection", Published in Elsevier, Information security technical report 13, pp. 216-219, 2008.
17. N. Godbole, Information systems security: Security Management, Metrics, Frameworks and Best Practices: JOHN WIELY, All Chapters, 2009.
18. N. F. Mir, Computer and Communication Networks: Prentice Hall, Chapter 5, 9, 10, 2006.
19. B. Forouzan, Data Communications and Networking: McGraw Hill, Fourth Edition, Chapter 13, 14, 15, 31, 32, 2006.
20. A. S. Ashoor and S. Gore, "Importance of Intrusion Detection system (IDS)", International Journal of Scientific and Engineering Research, vol 2, issue 1, January 2011.
21. U. A. Sandhu, S. Haider, S. Naseer, O. U. Ateeb, "A Survey of Intrusion Detection & Prevention Techniques", International Conference on Information Communication and Management IPCSIT: IACSIT Press, Singapore 2011.
22. K. Scarfone and P. Mell, Guide to Intrusion Detection and Prevention Systems (IDPS), Recommendations of the National Institute of Standards and Technology: NIST Special Publication, February 2007.
23. B. Menezes, Network Security and Cryptography: CENGAGE Learning, Chapter 14, 18, 19, 21, 22, 24, 2010.
24. J. R. Vacca, Computer and information security handbook: Morgan Kaufmann Series in Computer Security, First edition, May 4, 2009.