Methods and Techniques of Intrusion Detection: A Review

Somya, Palak Bansal, and Tameem Ahmad(✉)

Department of Computer Engineering, Z. H. College of Engineering and Technology, Aligarh Muslim University, Aligarh, India
somya1595@gmail.com, palakb1995@gmail.com, tameemahmad@gmail.com

Abstract. Malware is an abbreviated term meaning "malicious software". This software has a capability to gain access or infect a system without the knowledge of the owner. In this paper, we have tried to provide brief information about different types of malwares known till date such as virus, rabbits, botnet, adware etc. Apart from those we have mentioned the cure to it i.e. intrusion detection. We have described various techniques of intrusion detection such as signature based, anomaly based, behavior based etc. Methods for implementing these techniques include neural networks, data mining etc. A brief description of Intrusion detection system is also provided which is a software application used to monitor the network and system activities and also to detect malicious actions. The objective of this paper is to provide complete study about the types of malware, techniques and methods of intrusion detection, challenges and applications.

Keywords: Intrusion detection · Techniques and methods · Malware · Signature based IDS · Anomaly based IDS

1 Introduction

Malwares are a big threat to modern computer world. These are mischievous programs crafted to prohibit the normal operations, gain unauthorized access to data and resources of the system that may lead to privacy violation and other abusive behavior. Poor separation between code and data is the foremost cause of malware. Intrusion detection is a technique that attempts to discover the unauthorized access to a computer by analyzing the malicious activities and negatively identifying all the non-attacks. This paper will begin with an introduction describing various types of malware. Section 2 will discuss about Intrusion detection. Various Techniques will be explained in Sect. 3. Methods to employ the two most widely used techniques (i.e. signature based and anomaly based) are described in Sect. 4.

At the end, we have concluded the paper with its future scope discussion continued with references list.

1.1 Types of Malware

Virus. A computer virus is a kind of a special program that loads onto computer without the knowledge of the user. Computer viruses are man-made and can also replicate and

A. Unal et al. (Eds.): SmartCom 2016, CCIS 628, pp. 518–529, 2016.
DOI: 10.1007/978-981-10-3433-6_62

therefore need some host for its propagation through some communication medium like email attachment or trade program [1]. It can easily corrupt a system, steal information or destroy the data (Fig. 1).

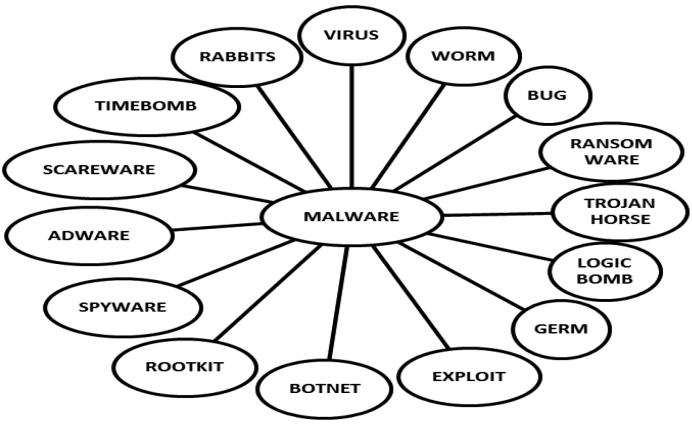

Fig. 1. Types of malware

Christmas tree, Melissa, CIH, Virus, Redlof, Autorun.abt, Peacomm, and NewHeur_PE is the most encountered examples.

Worms. Computer worms spread over computer networks by exploiting operating system vulnerabilities. It cause harm to their host networks by consuming bandwidth and overloading web servers. Furthermore, it can also contain payload to perform action on affected computers to steal or delete files or to create botnet. Worms have the ability to self-replicate and spread independently while viruses need human interference [14].

Blaster, Morris, Code red, Netsky, Stration, Sasser, Bagle, Skipi, no_virus are some of the examples.

Bugs. Bugs are nothing but human errors in the code or the program that compiles it. This flaw produces undesirable results. It can minutely affect programs behavior but when it persist for longer period, it can cause crashing or freezing.

JSONP is one of the most common bugs.

Ransomware. This malware holds a computer system captive while demanding a ransom. The malware limits user access to computer by displaying messages, encrypting files on hard disk or by form of denying the access on the system.

Examples are Reveton, CryptoLocker

Trojan Horse. It is a software program which trick users into downloading and installing malware. It could provide the access to the malicious party over the infected system and hence the attacker can easily bargain data, install more malwares, and modify files without the knowledge of legitimate user.

Some of the commonly encountered examples are Netbus, Nuker, Back Orifice, Limbo/NetHell, Pidief, ZeuS/PRG, Banker.bdn, PGPCoder, Torpig, and Gozi.

Logic Bomb. Logic Bomb are a group of line of code, kept associated with the program without knowledge, conditioned under some circumstance. As soon that condition met, it burst out with harmful effect. It does not replicate on other applications [13].

Examples are Jerusalem, SOBIG worm, Michelangelo.

Germs. Germs are considered as first generation virus which does not have a host program. It exists in its original form. When it infects files, it does not leave any mark of infection on that file leading to unawareness of the user. The problem continues in acquiring resources again and again to infect the unaffected file.

Examples are Germ was written for book sector.

Exploits. Exploits occur due to flaws which are exploited to create security hole to get into the system. Attackers execute a program to locate system's problem and used to make hidden paths to access confidential data.

White hat is an example.

Spyware. It's all about collecting information of the user or the organization without their knowledge that includes the spying the users' activity, keystrokes information collection, data harvesting in terms of account information and login credentials. These all are possible by putting in spyware program on someone's computer secretly and relay it to interested party.

Examples: Spyware Quake, Security toolbar, WhenUSave, PuritySCANVirtumonde.

Rootkits. It is malicious software program develop to control a computer without being detected by user. It is used to gain system administrator's access by compromising it. Because it continually hides its presence, typical security products are not effective in detecting rootkits [14].

Examples are Adore, Knark, LRK, AFX, SInAR, Rustock, and Mebroot.

Botnet. Bots are software program created to perform specific operation automatically. Bots can be used by botnets which is a collection of computers to be controlled by the third party for attacks. It is usually a zombie program [13].

Examples are LowSec, Rbot, Agobot, Slackbot, Mytob, SdBot, poebot, IRCBot, VanBot, MPack, and Storm.

Adware. It is advertising supported software. It includes the pop up ads on websites, free trial versions shown on an already infected system helping other malicious programs to get full access to the system including stealing information and tracking user activity [15].

Examples are DeskAd, AdBlaster.

Scareware. It usually comes from unreliable Internet sources like hacked websites, useful applications or exciting offers. When the user unknowingly downloads it, it produces vulnerable results leading to execution of malicious code and making system prone to attacks.

Smart Fortress, Android Defender are the most common examples.

Time Bomb. It is quite similar to logic bomb. As the name symbolize, there will be a particular time at which destructive effects will take place after the execution of the malware i.e. time is the parameter for its execution. Time bomb can be used by an internal user, who wishes to abolish the data of an organization due to some reason or his/her termination. Many incidents have occurred previously like an organization named Omega Engineering lost millions of dollars due to this malware [15].

Examples are The Christmas Day, Conficker.

Rabbits. Arabbit also known as "computer bacteria" is a computer program which lacks the logic bomb. It has a high rate of reproduction and replication in a shorter period of time. It slows down the computers and clog to the point of being nearly unusable. Cleaning a system is a long and complicated process in this case.

Fork Bombs, cmd.exe are few examples encountered.

2 Intrusion

An intrusion is an unauthorized attempt to break into a computer system. Such a break may force the system to move into an insecure state. It is a deliberate attempt made by the intruder to gain access of, manipulate or misuse valuable resources. If successful, it may result in rendering the resources as unreliable or unusable.

2.1 Intrusion Detection

The National Institute of Standards and technology classifies [4] Intrusion Detection as "The process of monitoring the events occurring in a computer system, or network and analyzing them for signs of intrusions, defined as attempts to compromise the confidentiality, Integrity, availability or to bypass the security mechanism of a computer or network" [5].

Intrusion detection is a type of security management system which is not only to identify an attack but also it includes the following [5]:

a. Monitoring the user and system activities
b. Analyzing system configurations and target vulnerabilities
c. Analysis of abnormal activity pattern
d. Accessing file integrity
e. Track of user policy violation
f. It provides the ability to recognize patterns typical of attacks by providing access to system and file integrity. If any target found vulnerable it would notify.

2.2 Intrusion Detection System

Intrusion Detection System is a security system that dynamically monitors and observes the target system for any misuse and handles the abnormal activity either by itself or by raising an alarm [2].

It can be configured to respond to predefined suspicious activities (e.g. when someone is trying to compromise the system's information through malicious activities) or it can even monitor the internet for latest attacks that could result into some future attack [10].

Primary criterions of measurement for IDS are as follows [3].

TRUE POSITIVE: legitimate attack (IDS gives alarm).
FALSE POSITIVE: no attack (IDS gives alarm).
FALSE NEGATIVE: legitimate attack (IDS gives no alarm).
TRUE NEGATIVE: no attack (IDS gives no alarm).

	T	F
T	TP	FN
F	FP	TN

Further IDS is classifies into three broad categories [8]:

Host Based IDS. This type of IDS collects information from an individual computer system.

It is the combination of signature based, rule based and heuristics based approach to detect intrusion. It monitors local host events, so can even detect attacks that a Network based IDS may miss. Events monitored include contents of operating systems, system and applications logs [9]. Its shortcoming is that it can easily be disabled by certain denial-of-service attacks.

E.g. Tripwire

Network Based IDS. This type of IDS collects information from the entire network. Sensors are placed that monitor all network traffic. Packets are collected and the captured data is then analyzed for predefined patterns and signatures to generate alerts [7]. There are 3 signatures that are very important [5]:

a. String signatures
b. Port signatures
c. Header signatures

This type of IDS does not actually impact the network performance, and are independent of the operating system. Its only when the networks are flooded, the packets are lost.

E.g. SNORT [11]

Vulnerability Assessment. This type is similar for scanning the attacks from the network and point out the drawbacks that are needed to be fixed. It is done on regular basis for the defensive works and ensures that the system or network is strengthened.

There are basically two models for the detection purpose: Anomaly detection and misuse detection which works simultaneously [16].

3 Techniques of Intrusion Detection System

3.1 Signature Based Intrusion Detection Technique

This is a primitive, simple and efficient technique of intrusion detection. This technique basically scans the malware program code and extracts its signature pattern [1]. Later it matches this signature pattern with the ones already fed in the database. During the extraction of signatures, all the system logs, executable files, records are taken into account. An alarm is raised immediately notifying the system user of the attack. The database is created by the antivirus developer, who analyses the new discovered malwares to find a specific pattern or a signature. Once the signature is extracted, it is then updated into the database. Since, detection rate and accuracy largely depends on the preexisting database, antiviruses need to be updated timely to provide better

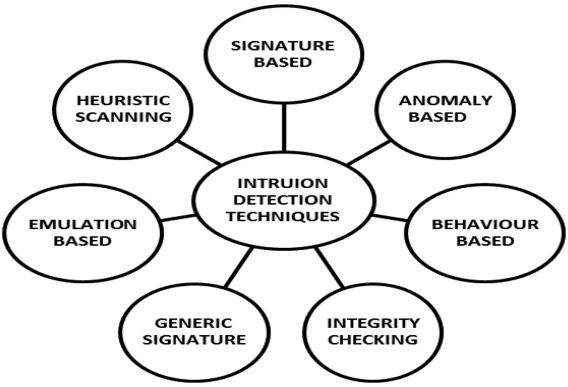

Fig. 2. Techniques of intrusion detection

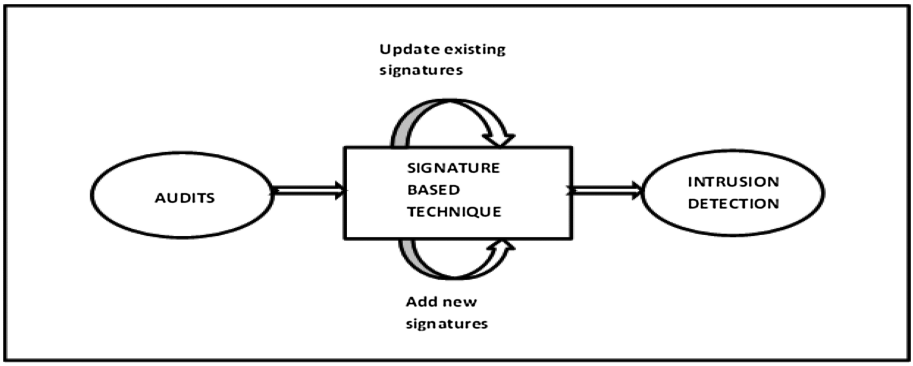

Fig. 3. Signature based intrusion detection

protection against the new upcoming latest malwares. This is in fact one of its major disadvantage [12] (Figs. 2 and 3).

3.2 Anomaly Based Intrusion Detection Technique

This technique overcomes the limitations of the signature based technique. The process of detection is divided into two phases namely, training phase and detection phase. In the training phase the system is trained about the normality whereas in the detection phase it compares the real data with the established profile to flag deviations and raise alerts [6]. It considers heuristics and artificial intelligence type techniques to differentiate between normal and abnormal activities.

Its main advantage is that it can easily detect unknown viruses as they also produce a different, anomalous behavior.

The only disadvantage with this technique is that there are a very large number of false positive attacks (Fig. 4).

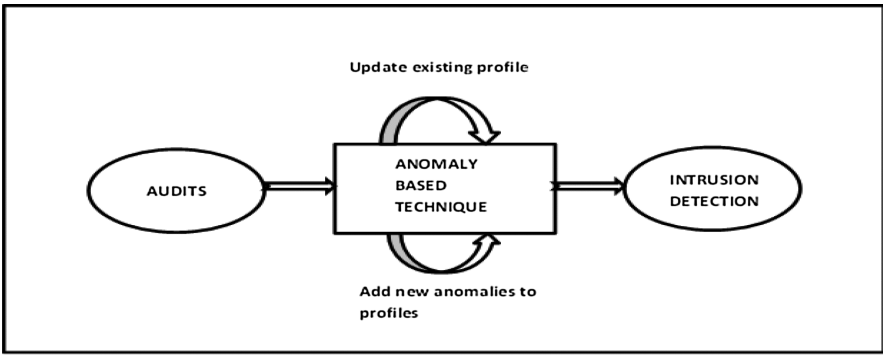

Fig. 4. Anomaly based intrusion detection

3.3 Emulation Based Intrusion Detection Technique

It is a trial-and-error detection technique that creates a virtual environment, and malware is executed by emulating its instructions. The recorded behavior is then used to propose a remedy. It is the best technique for the detection of metamorphic, polymorphic and encrypted viruses. The only disadvantage with this type of virus is that it does not detect all those malwares that do not show their abnormal behavior during the emulation.

3.4 Generic Signature Scanning Intrusion Detection Technique

This technique is used to detect malwares that belong to the same family i.e. they differ very little in their signatures.

This technique overcomes disadvantage of Signature based technique as it scans using patterns and wildcards to detect the various variants of a family. Generally new malwares are evolved by incorporating minor changes in the pre-existing malware

codes. That is that it cannot detect entirely different or new malwares and also they require a skilled researcher to extract variant patterns [1].

3.5 Integrity Checking Intrusion Detection Technique

It is simple to implement and gives efficient results. The first step includes computing and storing san uninfected file's hash value.

All other important files of the system like log files, boot files and executable files are scanned and their hash values calculated. This calculated hash value is compared with the ones already stored to detect intrusion. Integrity checker can also use other methods such as checksums or fingerprints of a file. The main disadvantage with this technique is that it may raise alerts even when minor changes are incorporated in any file.

3.6 Heuristic Scanning Based Intrusion Detection Technique

This is time consuming technique which requires extensive scanning of various boot record files, log files and executable files for code comparable to malicious code. Malicious program instruction include worm propagation routines, virus replication routine etc. Heuristic scanning can be of 2 types [1]:

Static Scanning: It includes direct scanning of a program to look for malicious code.
Dynamic Scanning: It includes emulation of the programs on emulators to detect for malware presence.

The main disadvantage of this technique is the generation of a lot of false positive alarms, and the chances of not detecting a known virus are also high in this case.

3.7 Behavior Based Intrusion Detection Technique

It is used for detection of viruses in terminals like PCs and mobile phones. This technique observes behavior of various malwares and fetches information like its source and destination, attachment size, type etc. Based on this behavior; it can flag a particular source as an intruder. It is capable of detecting known and unknown viruses with self-reference replication behavior. The only disadvantage is generation of false positives and false negatives [1].

4 Methods

4.1 Anomaly Based Intrusion Detection

Statistical Model. The parameters that are considered into account in the statistical methods are such as monitoring the CPU utilization by the user along with the user's session time and the Bandwidth (Fig. 5).

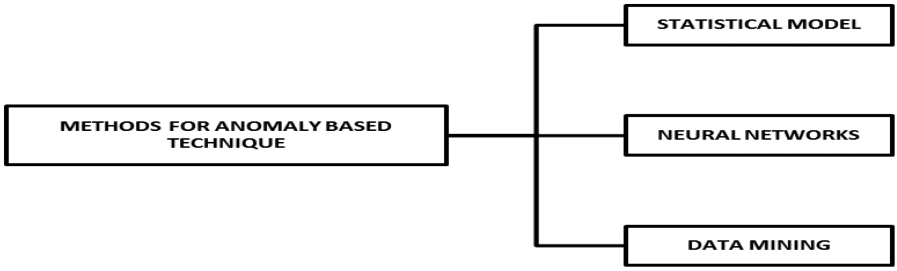

Fig. 5. Methods for anomaly based intrusion detection

These parameters are collected over time and are compared with the baseline criteria [7]. This technique works by looking for deviations from the ideal behavior and determining similarities of events with those which are typical of indicating an attack.

Neural Networks: This method takes into consideration typical characteristic of a system user and then establishing differences and variations from the user's established behavior.

It consists of three phases [9]

a. *Collection of training data*: This includes obtaining audits, log files for a user for a specified period..
b. *Training*: Neural networks are trained to identify users on the basis of these mapped audits.
c. *Performance*: The network is allowed to identify user according to its mapped audits. If the identified user and actual user differ then an anomaly is detected [3].

Data Mining: Data mining is the process of identifying valid, novel and useful patterns in data. The technique is the process of information extraction with an objective to trace the hidden facts in the repository [17]. The technique would verify this information and determine if it is an attack.

Data mining parameters include:

a. *Association*: Patterns are taken into consideration where there are dependencies and connection among them.
b. *Sequence analysis*: This parameter consider the patterns where one event takes to another event
c. *Classification*: looking for new patterns.
d. *Clustering:* finding facts not previously known.
e. *Forecasting*: discover patterns that can lead to sensible estimates.

4.2 Signature Based Intrusion Detection System

Signature Analysis: It is used as carving which is used as identifying the virus using predefined signatures and find the deleted files affected by the virus. It converts the

semantic format statement of an attack into an appropriate audit trial format. Low detection rate for zero day attacks is one of its disadvantages (Fig. 6).

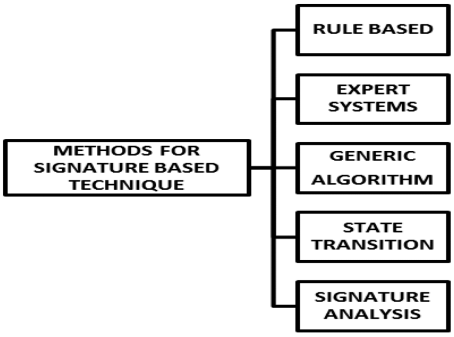

Fig. 6. Methods for signature based intrusion detection

State Transition Based: This model is collection of states, transitions and actions. An attack is described in a form of a transition that must be achieved by an intruder to compromise with a system [5]. It denotes various network states as states of a finite state machine. If a sequel state is identified from the network state of finite state machine, an intrusion is detected [3].

The graphic representation of intrusion is used for the successful completion.

Genetic Algorithms: It is a programming technique that mimics the biological evolutionary technique as a problem-solving strategy. The method possesses natural selection using a chromosome-like data structure and evolves that chromosome.. An evaluation function (also called "Fitness function") is used to compare the quality of each chromosome (survival for the fittest) with view of the desired goal [5].

It efficiently detects numerous types of network intrusion. This approach filters the traffic data using evolutionary theory, converges in less time and thus decreases the complexity.

Rule Based: Rule Based Approach observes deviations from the pre-defined rules; this information is provided to the expert system for detection of intrusion [7]. The demerit is that it could not able to detect the attacks that may occur over an extended period of time. The variations in an attack can affect the activity-rule comparison to a level that intrusion detection system fails to determine.

Expert Based. Previously, data audits and logs were manually analyzed. The disadvantage with this technique was extra time requirements and efforts for this analysis. The expert based System have been evolved to overcome this effort and mechanism have been developed with the creation of knowledge base [18] along with the set of rules, based on heuristics to trap the intrusion automatically without the human intervention.

5 Conclusion and Scope of Improvement

Security is an abstract concept, networks are prone to malware attacks. So, exposing the data or the resource, to an unauthorized person is risky. Proper steps must be taken care to prevent the network / system attacks and preserve the security and integrity. In this paper, a brief overview of different types of malwares, Intrusion detection techniques and methods is presented and it also highlights its advantages, disadvantages and applications.

For the sake of improvement in the existing intrusion detection systems we propose the idea of a modification in a form of approximate string matching which can increase the efficiency of present intrusion detection algorithms. According to this idea, we can compare the signature string and if at most four mismatches are detected, then this algorithm automatically puts some random values and detect the nearest possible match and then raises alarm. Along with alarm it can also update the database with the newly found malware. In case of Signature Based Intrusion Detection System, the new viruses are not detectable. So, with the help of the above proposed algorithm, this issue can be resolved to some extent. It can be considered more reliable and less resource consuming. Various malicious coders basically modify the signatures of existing malwares to develop new ones. We also know that this slight variation in signatures cannot be detected by signature based technique. The implementation of the proposed algorithm will reduce this failure rate as little changes in the signatures will still be detected.

References

1. Slade, R.: Guide to Computer Viruses: How to Avoid them, How to Get Rid of them, and How to Get Help. Springer, Heidelberg (2012)
2. Bashir, U., Chachoo, M.: Intrusion detection and prevention system: challenges and opportunities. In: International Conference on Computing for Sustainable Global Development (INDIAcom). IEEE (2014)
3. Asif, M.K., Khan, T.A., Taj, T.A., Naeem, U., Yakoob, S.: Network intrusion detection and its strategic importance. In: Business Engineering and Industrial Applications Colloquium (BEIAC) (2013)
4. Shah, B.: How to Choose Intrusion Detection Solution. SANS Institute Resources, 24 July 2001
5. Akbar, S., Rao, K.N., Chandulal, J.A.: Intrusion detection system methodologies based on data analysis. Int. J. Comput. Appl. 5(2), 14–18 (2010)
6. Nascimento, G., Correia, M.: Anomaly based Intrusion detection in software as service. In: 41st International Conference on Dependable Systems and Networks Workshops (DSN-W). IEEE (2011)
7. Mukherjee, B., Heberlein, L.T., Levitt, K.N.: Network intrusion detection. IEEE Netw. **8**(3), 26–41 (2002)
8. Sabahi, F., Movaghar, A.: Intrusion detection: a survey. In: 2008 Third International Conference on Systems and Networks Communications (2008)
9. Reddy, E.K.: Neural networks for intrusion detection and its applications. In: Proceedings of the World Congress on Engineering, vol II, London, UK (2013)
10. Kabiri, P., Ghorbani, A.A.: Research on intrusion detection and response: a survey. Int. J. Netw. Secur. **1**(2), 84–102 (2005)

11. Dharmapurikar, S., Lockwood, J.W.: Fast and scalable pattern matching for network intrusion detection systems. IEEE J. Sel. Areas Commun. **24**(10), 1781–1792 (2006)
12. Chou, T.-S.: Development of an intrusion detection and prevention course project using virtualization technology. Int. J. Educ. Develop. Using Inf. Commun. Technol. (IJEDICT) **7**(2), 46 (2011)
13. Hardikar, A.M.: Malware 101 – Viruses GSEC Gold Certification. SANS Institute (2008)
14. Szor, P.: The Art of Computer Virus and Defense. Addison Wesely Professional, Harlow (2005). ISBN 10: 0321304543, 13: 978-0321304544
15. Mathur, K., Hiranwal, S.: A survey on techniques in detection and analyzing malware executables. Int. J. Adv. Res. Comput. Sci. Softw. Eng. 3(4) (2013). ISSN: 2277 128X
16. Reddy, E.K.: Neural networks for intrusion detection and its applications. In: Proceedings of the World Congress on Engineering, vol. 2. no. 5 (2013)
17. Marakas, G.M.: Modern Data Warehousing, Mining, and Visualization: Core Concepts. Pearson Education, Upper Saddle River (2002). ISBN 0131014595
18. Ahmad, T., Ahmad, S., Jamshed, M.: A knowledge based Indian agriculture: With cloud ERP arrangement. In: 2015 International Conference on, Green Computing and Internet of Things (ICGCIoT), Noida, pp. 333–340 (2015). Doi:10.1109/ICGCIoT.2015.7380484