International Journal of Electronics Engineering (ISSN: 0973-7383)
Volume 11 • Issue 1 pp. 895-899 Jan 2019-June 2019 www.csjournals.com

Performance Analysis of Intrusion Detection System based Cloud using WEKA Tool

Amandeep Singh [1], Naresh Kumar [2]
[1,2] University Institute of Engineering & Technology, Department of Computer Science & Engineering, Kurukshetra, India

Abstract: The Proposed algorithm was decentralized to avoid the impact of the single point failure for the use of intrusion detection system (IDS) based classifiers, priority based and native based classifiers. The main objective was to combine some meta-heuristic properties of native bays and bays net to prepare a hybrid algorithm for effective scheduling and uniform distribution of workload. The parameters used were number of processing cores, million instruction per second searching, accuracy rate and time complexity. In the implemented algorithm, classifiers were considered in Waikato Environment for Knowledge Analysis (WEKA) tool. IDS based cloud uses WEKA tool for simulation and performance analysis.
Keywords: decentralized, intrusion detection system, classifiers, meta-heuristics, scheduling, WEKA.

Introduction

Cloud computing is an internet-oriented computing that provides shared processing elements, storage space and other devices whenever required. It is an architecture which provides services to users and helps to reduced IT overhead for end users, increases flexibility and reduces cost. By properly using resources, it helps to improve profit. One main problem is that there is scheduling of tasks for free resources. Many algorithms are used to solve the problem. It's an NP hard-problem (non-deterministic polynomial-time hardness). One of the drawbacks of cloud computing is if there is no internet connection then cloud computing is not possible hence you cannot access anything as it do not work with low connections. Some storage issues also arises like stored data may not be secured, it can be lost. Cloud computing resources are in cups, firewall, network forms, which are dynamically allocated to tasks according to the sequence. So, task scheduling is a dynamic problem in cloud. This means that no previous defined sequence may be useful for task processing. For dynamic scheduling the motive is uncertainty in task flow, uncertainty in execution path & uncertainty in available resources, because many number of tasks are sharing these resources simultaneously at that time. In scheduling of tasks, the best suitable resources are allocated to task in such a way that decreases the completion time. The priority is given to list of tasks in schedule. Then according to priority, tasks are allocated to resource virtual machine which satisfy an objective function.

Problem Statement

The main problem is to determine what type of application should be allocated to host that resources can efficiently use. In cloud there are numbers of resources which are differing from each other by meaning & cost. Task scheduling is different in cloud from the traditional methods of scheduling; attention must be needed on scheduling of tasks in cloud because cloud services depend upon them. Task scheduling is important for improving the flexibility & reliability of systems in cloud. In accordance with, time bound tasks are scheduled on resources, which help to find the best sequence in which many tasks are executed to give the suitable results to the users.

Cloud Service Scheduling

Cloud service is divided in two levels. First is 'User level' and second is 'System level'. User level scheduling deals with service troubles occuring among service providers and customers. For handling the resource management within data center, scheduling is used. Data center contains many virtual machines. Many tasks are received from users and then they assigned it to these virtual machines on the data center. This assignment of tasks affects the performance of data center. Many factors are considered like sharing of resource, QoS (Quality-of-Service), fault tolerance, reliability, flexibility etc. of utilization of system.

• Static & Dynamic Scheduling: In static scheduling pre-fetching is done for required data and pipelining is used for various stages of tasks execution. In dynamic scheduling there is no pre information about task. No information is known about execution time of tasks and during execution of application, allocation of tasks is done. The service request scheduling strategies in 3 tier cloud structure architecture consist of resource providers, service providers & customers/users that satisfy the aim. A preemptable scheduling enhance the performance of resources in clouds.

- Heuristics Scheduling

Optimization problems are included in NP-hard class. These types of problems are solved by evolution method heuristic approach. In evolution method, the most favorable solution can be selected if all probable solution is evaluates. Exhaustive evolution is not feasible for large number of instances problems of scheduling. In this case, heuristic approach is used to find optimal solutions, fastly by using suboptimal algorithm. This algorithm is used only for those problems which have exact polynomial time. For task completion in time at large scale, data processing system is required & needed to enhance the locality of task data. Many approaches are used to improve the locality of data which are either greedy & global optimization. Multi objective meta-heuristics scheduling algorithm is used to achieve high availability & fault tolerance, which helps to reduce the resource load.

- Real Time Scheduling

For increasing the throughput & minimizing the average response time, real time scheduling is used. Real time tasks are scheduled non preemptively for fully utilizing the service. A real time work load driven approach is proposed.

Booking of advancing errands on cloud is one of the examination issue, where the arrangement of machines and finish time of the assignments are considered. Reliable errand's arranging of machines issue is that, recognize number of dynamic hosts are p, number of VMs in each host are q. Most silly number of conceivable VMs to outline a solitary undertaking is (p*q). In the event that we have to outline r attempts, number of potential outcomes are (p*q)^r . So organizing of assignments is NP Hard issue. Satisfaction time need of steady undertaking is that if errand finish in dead line then just it is helpful else it isn't. In the event that it isn't beneficial then it is rejected. Most opportune Dead line First (EDF) estimation is excellent figuring for masterminding of incessant errands. Earliest Deadline First is Event Driven booking calculation which requires select as sensibly concerning their due dates. Endless undertakings can be intermittent, sporadic, aperiodic errands. Aperiodic and periodic model are utilized to assess execution of changes booking calculations. Everything considered EDF Scheduler outline the undertakings to such an extent, to the point that it assign the errand to the free open machine without considering the errand on that machine will meet the dead line or not.

- Work Scheduling

Work flow is used to structure the application in directed acyclic graph (DAG) form. Every node of DAG represents the constituent task and the applications inter task dependency is represented by edges. A work flow contains the set of tasks which are communicated with one another with in work. This kind of scheduling is most significant to manage the execution of work flow.

Cloud services and commercial products

There are various types of cloud computing services such as compute, storage, database, application, content delivery, analytics, deployment & management services and identity & access management services. Some are discussed below:

i) Storage Services: Cloud storage services allow storage and retrieval of any amount of data, at any time from anywhere on the web. These storage services organizes data into buckets or containers. Buckets or containers store objects which are individual pieces of data. Main features of cloud storage services are scalability, replication, access policies, encryption and consistency.

ii) Database Services: Cloud database services allow to set-up and operate relational or non-relational databases in the cloud. Main benefit of using these services is that it relieves the application developers from the time consuming database administration tasks. Some popular relational databases provided by various IT service providers are: MySQL, Oracle, SQL Server etc. The main features of these services are: scalability, reliability, performance and security. Some commercial products that are providing cloud database services are as follow:

- Amazon Relational Data Store: It is a relational database service from Amazon.
- Amazon Dynamo DB: It is a non-relational database service provided by Amazon.
- Google cloud SQL: It is a relational database service from cloud that allows hosting MySQL databases in the Google's cloud.
- Google cloud data store: It is fully managed non-relational database service from Google.
- Window Azure SQL database: It is a relational database service from Microsoft that is based on SQL server.
- Window Azure Table Service: It is a non-relational database service from Microsoft.

iii) Application Services: There are various cloud application services such as application runtimes and frameworks, queuing service, email service, notification and media services.

- Application runtime and frameworks: Cloud-based application runtimes and frameworks allow developers to develop and host applications in the cloud. Google App Engine and Windows Azure

Web Sites are the services from Google and Microsoft respectively that provide Platform as Service (PaaS) to cloud consumer.

- Queuing Services: It includes Amazon Simple Queue Service, Google Task Queue Service, and Windows Azure Queue Service.
- Email Services: Cloud-based email services allow applications hosted in the cloud to send emails. Amazon Simple Email service and Google Email Service are some popular email services from Amazon and Google respectively.
- Notification Services: Cloud-based notification services, also called push messaging services allow applications to push messages to internet connected smart devices such as smart phones, tablets etc. Some popular notification services are: Amazon Simple Notification Services, Google cloud messaging, Windows Azure Notification Hubs.
- Media services: These services include Amazon Elastic Transcoder, Google images Manipulation Services, Windows Azure Media Services and many more. These services are mainly used by applications for manipulating, transforming and transcoding media such as images, videos etc.

iv) Content Delivery Services: Content delivery services include Content Delivery Networks (CDNs) that are useful for serving static content such as text, images, scripts, etc. and streaming media. Two popular content delivery services are:

- Amazon Cloud front
- Windows Azure Content Delivery Network

WEKA Performance analysis

Step 1: WEKA tool is used for knowledge discovery in database (KDD) dataset. K-mean Clustering algorithm is implemented where K=2, Normal and Anomaly.

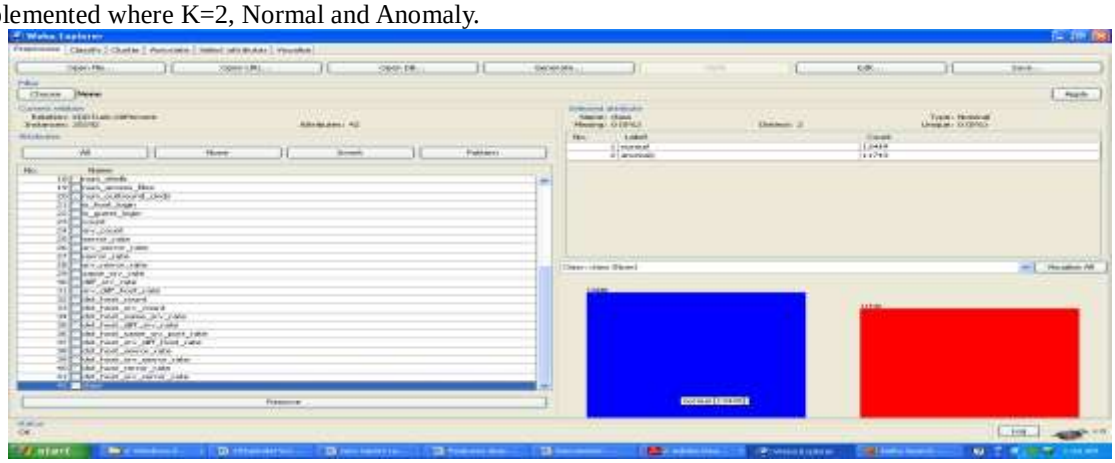

Fig. 1.1: Weka Explorer with dataset showing normal and anomaly

Step 2: Now centroid is found and weka 3.6.2 in cluster is selected

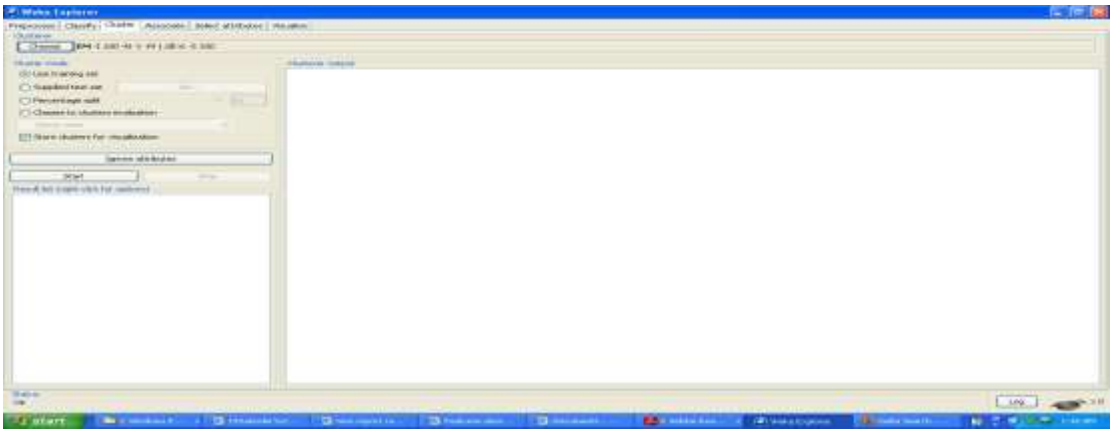

Fig. 1.2 : Select Cluster

Step 3: 42 Attribute in KDD is used.

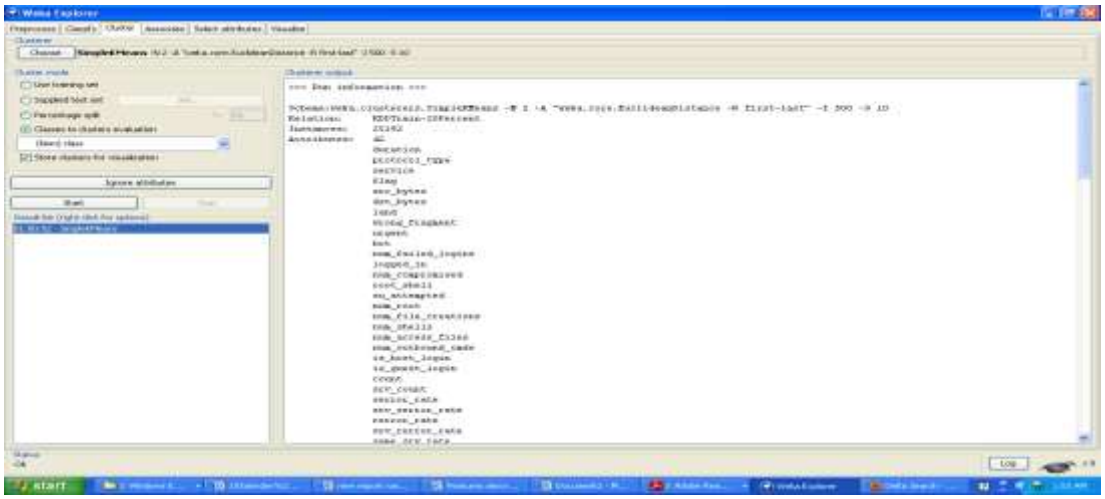

Fig. 1.3: Attribute in k-mean

Step 4: Cluster visual assignment is analyzed.

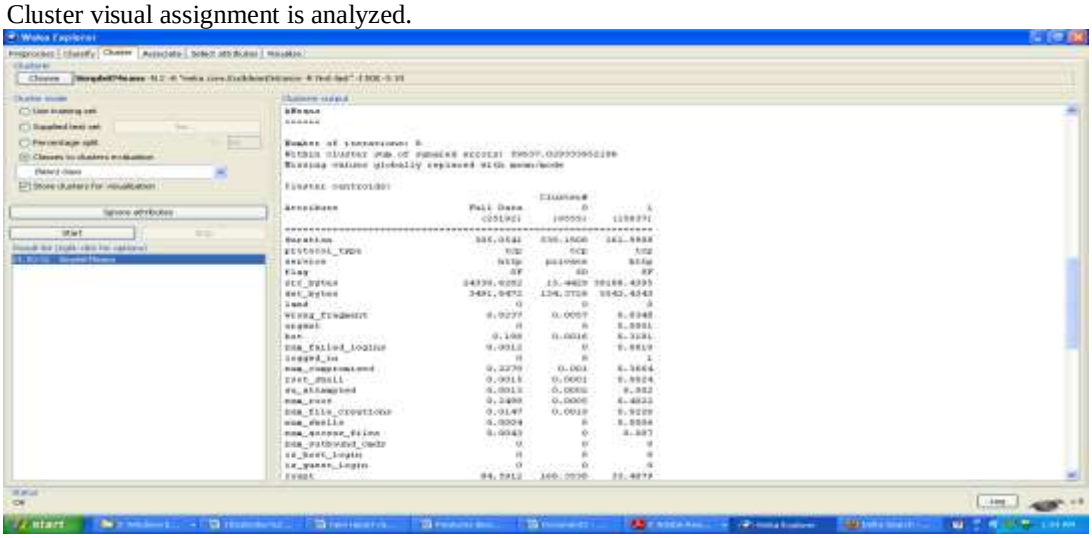

Fig. 1.4 :Anomaly and normal Detection Table

Conclusion

Implemented a set of classifiers for IDS among virtual machines. The main objective was to keep virtual machines busy all the time for their efficient utilization. Whenever a new task has to be submitted to the data center for processing that virtual machine had to be selected which had been assigned tasks minimum number of times. The implementation and testing of classifier was done using WEKA simulation tool. A comparative analysis was also done with the existing algorithms using same tool. The results depicted that new algorithm performed hews rather than existing one.

References

[1] Chirag N. Modi, Dhiren R. Patel, Avi Patel, Muttukrishnan RajarajaN, " Integrating Signature Apriori based Network Intrusion Detection System (NIDS) in Cloud Computing" Procedia Technology 6 (2012, 2nd International Conference on Communication, Computing & Security (ICCCS), 905-912.

[2] Mell, P. & T. Grance, "The nist definition of cloud computing (draft)" 2011, https://csrc.nist.gov/publications/drafts/800-145/draft-SP-800-145/final.

[3] Gens. F, "IT Cloud Services User Survey, pt.2: Top Benefits & Challenges", 2008, http://BLOGS.IDC.COM/IE/?=210.

[4] S. Beg, U. Naru1, M. Ashraf & S. Mohsin, "Feasibility of Intrusion Detection System with High Performance Computing: A Survey", International Journal for Advances in Computer Science, 2010, vol 1, pp 26-35.

[5] D. J. Brown, B. Suckow & T. Wang, "A Survey of Intrusion Detection Systems", Technical report Department of Computer Science, University of California, San Diego, 2002.

[6] D. Stiawan, A. H. Abdullah & M. Y. Idris, "The Trends of Intrusion Prevention System Network", 2nd International Conference on Education Technology and Computer (ICETC), 2010, Vol. 4 pp. 217–221.
[7] T. Dutkevyach, A. Piskozub & N. Tymoshyk, "Real-Time Intrusion Prevention and Anomaly Analyse System for Corporate Networks", in 4th IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, 2007, pp. 599–602.
[8] H. Zhengbing, S. Jun & V. P. Shirochin, "An Intelligent Lightweight Intrusion Detection System with Forensic Technique", in 4th IEEE Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, 2007, pp. 647–651.
[9] H. Zhengbing, L. Zhitang & W. Jumgi, "A Novel Intrusion Detection System (NIDS) Based on Signature Search of Data Mining", in WKDD First International Workshop on Knowledge discovery and Data Mining, 2008, pp. 10–16.
[10] Ahmed Patel, Mona Taghavi, KavehBakhtiyari, Joaquim Celestino Ju´nior, "An intrusion detection and prevention system in cloud computing: A systematic review", Journal of network & computer application 2012, vol 36, pp 25-41.
[11] Arshad J, Townend P, Xu J, "A novel intrusion severity analysis approach for Clouds", Future Generation Computer Systems Journal, 2013, vol. 29(1), pp. 416-428.
[12] Kizza J.M, "System intrusion detection and prevention", A Guide to Computer Network Security, Springer, 2009, pp. 273–298.
[13] Khorshed MT, Ali ABMS, Wasimi SA, "A survey on gaps, threat remediation challenges and some thoughts for proactive attack detection in cloud computing", Future Generation Computer Systems 2012, vol. 28, pp.833–851.
[14] Grobauer B, Walloschek T, Stocker E, "Understanding cloud computing vulnerabilities", Security & Privacy, IEEE 2011, vol. 9(2), pp. 50-57.
[15] Dastjerdi Amir Vahid, KA Bakar, and SGH Tabatabaei, "Distributed intrusion detection in clouds using mobile agents", in Third International Conference on Advanced Engineering Computing and Applications in Sciences, IEEE, 11-16 Oct 2009, pp 175-180.
[16] Viega J, McAfee, "Cloud computing and the common man", Computer 2009, vol. 42, pp 106-108.
[17] Wang C, Q Wang, K Ren, and W Lou, "Ensuring data storage security in cloud computing," in 17th International Workshop on Quality of Service, 2009, pp. 1–9.