International Journal of Current Trends in Engineering & Research (IJCTER)
e-ISSN 2455–1392 Volume 2 Issue 3, Month Year pp. 163 - 165
Scientific Journal Impact Factor : 3.468
http://www.ijcter.com

Applying Data Mining Techniques on Fog Computing

Mahajan Megha[1], Laware Pallavi[2], Jadhav Swapnali[3], Shivam Singh[4], Yashpal Singh[5]

[1,2,3,4,5] *Government College of Engineering and Research Awasari, Pune*

Abstract— Now days the era is changing so fast as the people are dependent only on the internet. Whole task requires internet. To connect to the social media we require that internet facility. Also the overall data of the social media or any web service is stored virtually and that is nothing but cloud. So now a days Cloud computing is prominent. But many times this cloud service may go in trouble by bypassing the security. The other person may know the password of owner and may damage the data so to avoid this we have to provide extra security to the social account. We had provided the facility that the attacker will get decoy data and the original data will be secured. So the Fog computing means providing cover to the cloud by giving decoy data to the intruder and original only to the owner.

I. INTRODUCTION

The internet is an enabler, and one of the most important things it's enabled is cloud services. Computer Science technologies have been sprouting up like wild weed in a rainforest. We now have better computers and faster systems. Along with system performance scientists have managed to boost data storage capacities which have opened up a whole new window of avenues for the computing community.

All medium and large scale industries use cloud. This obviously supports better operational efficiency, but comes with risks, perhaps the most serious of which are data theft attacks, insider. This is considered as one of the top threats to cloud computing by the Cloud Security Alliance. While most Cloud computing. If the insider knows the password of owner then he can easily enter the system. So to avoid this security more than the password should be provided. In our system we are finding and comparing the behavior and can easily identify the original user from all intruders. If the user is not authorized then decoy file will be provided to it.

The overall project is dependent upon original data, decoy data, naive bayes algorithm and the database of user behavior.

1. Original data

Original data is only provided to the owner of the system. Also it will be protected from attacker.

2. Decoys

This data will be provided to the intruder. This data contains decoy information such as, catalog files, excel sheet, decoy documents, hone files, honey pots, and various other bogus information can be generated on demand and serve as a means of detecting unauthorized access to information and to poison the ex-filtrated information. This data is only to confuse the intruder. This technology may be integrated with user behavior profiling technology to secure a user's information in the Cloud. Whenever abnormal access to a cloud service is noticed, decoy information may be returned by the Cloud and delivered in such a way as to appear completely legitimate and normal. Combining User Behavior Profiling and Decoy Technology for Masquerade Detection User Behavior Profiling Legitimate users of a computer system are familiar with the files on that system and where they are located. Any search for specific files is likely to be targeted and limited. The decoy files are downloaded by the legitimate user and placed in highly-conspicuous locations that are not likely to cause any interference with the normal user activities on the system. Combining the Two Techniques

.We hypothesize that detecting abnormal search operations performed prior to an unsuspecting user opening a decoy file will corroborate the suspicion that the user is indeed impersonating another victim user. This scenario covers the threat model of illegitimate access to Cloud data. Furthermore, an accidental opening of a decoy file by a legitimate user might be recognized as an accident if the search behavior is not deemed abnormal. In other words, detecting abnormal search and decoy traps together may make a very effective masquerade detection system. Combining the two techniques improves detection.

3. Hadoop Implementation

Big Data is a collection of large datasets that cannot be processed using traditional computing techniques. It is not a single technique or a tool, rather it involves many areas of business and technology. And that Big data is managed efficiently by Hadoop technology. In project the map reduces technique of hardtop is used to map behaviors of user. , MapReduce is a parallel programming model for writing distributed applications devised at Google for efficient processing of large amounts of data (multiterabyte data-sets), on large clusters (thousands of nodes) of commodityHadoop is an Apache open source framework written in java that allows distributed processing of large datasets across clusters of computers using simple programming models .

4. Naive Bayes

This algorithm is used in this project for calculating the probability of the user behavior occurrence. It will compare the current user's behavior with history data and then it will make decisions by finding the probability. It will be in the percentage probability and will give confirm decision that user is valid or not and which file will be provided to it. Before that the support vector machine algorithm was used but the identification was on classification so it was unable to give proper decisions. So this drawback overcame by naive bayes algorithm.

II. IMPLEMENTATION

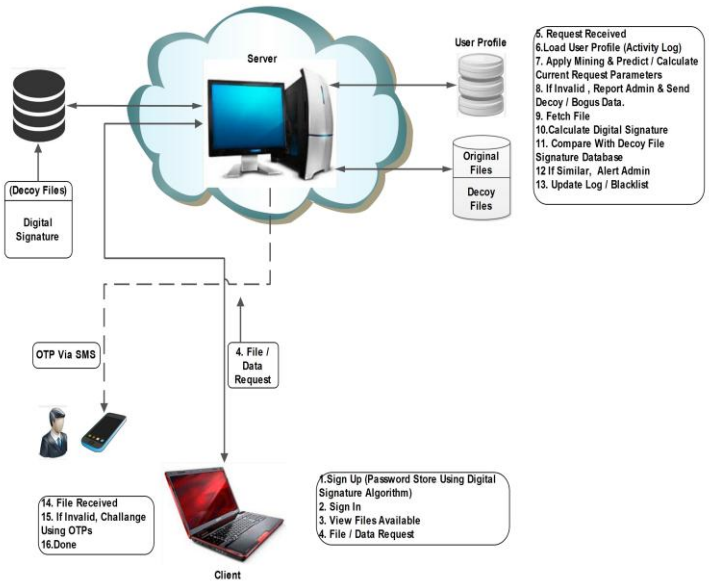

Fig 1, Implementation Diagram

III. ADVANTAGES

- Though intruder knows the password of owner, it can't affect the system.
- And also can't get original data files so data becomes secure.
- For the owner OTP service is provided.

IV. DISADVANTAGES

- If intruder's behaviors completely match with owner then identification becomes difficult.

V. FUTURE SCOPE

- We can add more attributes to Naïve Bayes to make the system more strong.
- Future work will expand on the Fog computing paradigm in Smart Grid.

VI. CONCLUSION

In this position paper, we present a novel approach to securing personal and business data in the Cloud. We propose monitoring data access patterns by profiling user behavior to determine if and when a malicious insider illegitimately accesses someone's documents in a Cloud service. Decoy documents stored in the Cloud alongside the users real data also serve as sensors to detect illegitimate access. Once unauthorized data access or exposure is suspected, and later verified, with challenge questions for instance, we inundate the malicious insider with bogus information in order to dilute the users real data. Such preventive attacks that rely on disinformation technology, could provide unprecedented levels of security in the Cloud and in social networks.

REFERENCES

[1] D. Takahashi, French hacker who leaked Twitter doc-aments to TechCrunch is busted, March 2010 https://cloudsecurityalliance.org/topthreats/csathreats.v1.0.pdf

[2] P. Allen Obamas Twitter password revealed after French hacker are-rested for breaking into U.S. presidents account, March 2010. http://www.dailymail.co.uk/news/article-1260488/Barack-Obamas-Twitterpassword-revealed-French-hacker-arrested.html

[3] B. M. Bowen and S. Hershkop, Decoy Document Distributor: http://sneakers.cs.columbia.edu/ids/fog/, 2009.

[4] M. Ben-Salem and S. J. Stolfo Combining a baiting and a user search profiling techniques for masquerade detection, in Columbia University Computer Science Department, Technical Report cucs-018-11, 2011. https://mice.cs.columbia.edu/getTechreport.php?techreportID=1468

[5] Van Dijk and A. Juels, on the impossibility of cryptography alone for privacy-preserving cloud computing, in Proceedings of the 5th USENIX conference on hot topics in security, ser. HotSec10. Berkeley, CA, USA: USENIX Association, 2010, pp. 18. [Online]. Available: http://dl.acm.org/citation.cfm?id=1924931.1924934

[6] J. Peptone, Dropboxs password nightmare highlights cloud risks, June 2011.

[7] M. Ben-Salem and S. J. Stolfo, Modeling user search-behavior formasquerade detection, in Proceedings of the 14th International Symposium on Recent Advances In Intrusion Detection. Heidelberg: Springer, September 2011, pp. 120.] https://mice.cs.columbia.edu/getTechreport.php?techreportID=1468