IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-3, Issue-5 E-ISSN: 2347-2693

Securing Data from Data Manipulation Using NAIVE String Search Algorithm and NTRU Algo in Cloud Computing

Indu Sharma[1*] and Mandeep kaur[2]

[1*,2] Dept. of Computer Science Engg., *Rayat and Bahra Institute of Engg., Sahuaran, PTU Punjab , India*

www.ijcseonline.org

Received: May /12/2015 Revised: May/16/2015 Accepted: May/27/2015 Published: 30/May/ 2015

***Abstract*— Cloud computing nowadays is dealing with most of security issues and privacy issues, no matter how strong security we use but achieving the 100% security will always a tough notch. However this research deals with area of cloud computing in building the trust level of cloud brokers among the general public or its distributers and buyers. The trust level is achieved by following a novel idea of checking the data manipulation in uploaded content. The SAAS cloud network in general is always secured with encryption so in this part for security NTRU is been implemented. Further this method describes the naive bayes string pattern matching algorithm. The data which is uploaded in server is auto synchronizing for back up onto another server which is a hidden database server. The user doesn't know that his/her data is kept in this server also whenever the query for download is been generated to download the server the existing data in server is cross verified with hidden database to check the manipulation if any which is accommodated using Naïve Bayes String Search Algorithm. If no data manipulation is been occurred it defines the superiority and trustworthiness of cloud storage among the public and cloud buyers.**

Keywords*— *Data Security, Cloud, Computing, Privacy, NTRU, NAIVE

I INTRODUCTION

Cloud is essentially a bunch of commodity computers networked together in same or different geographical locations, operating together to serve a number of customers with different need and workload on demand basis with the help of virtualization. Cloud computing is a system if tasks and data are held on the internet rather than on individual devices, giving on-demand capability. Resources are provided and administered by a service provider. Programs are started on a remote server and then sent to the user. To maintain the data integrity and data availability many people proposed several algorithms and methods that enable on demand data correctness and verification. So Cloud servers are not only used to store data like a ware house , it also provides frequent updates on data by the users with different operations like insert, delete , update and append.

1.2 Security and Privacy in Cloud

One of the core themes of cloud computing and cloud storage in general is that service should be independent of the location. Some of these aspects affect the way the cloud service provider creates his service and might lead to security and privacy issues for the consumer of the services.

Some of the characteristics of the infrastructure are detailed as follows.

Location Independent Services: The very characteristics of the cloud computing services is the ability to provide services to their clients irrespective of the location of the provider, the physical hardware below could be moved anywhere but the services should still be available. This feature also applies to the consumers of the services, with the advent of computing platform the consumption of the services cannot be restricted to a particular location but may be requested from any location as per the choices of the customer.

Communications: Due to the vary nature of the cloud computing infrastructure communications is a major component in every design. These communication lines could exist from few seconds to hours based on the services being consumed. So the security of this communication lines should be persistent as long as the connection between the provider and consumer exists at minimum and cover some buffer period too.

Infrastructure: The infrastructure that is used for these services should be secured appropriately to avoid any potential security threats and should cover the life time of component. This lifetime can be estimated to be about 10 years.

Storage Security: The data that is stored on the cloud services often would last longer than the security that should be ensured of the components which are used to store or compute these data. This would entail the storage services should be robust enough to achieve component and hardware changes easily and transparently. This applies to the algorithms and encryptions schemes that are used to secure this data; they could become obsolete and might become easy targets to brute force attacks as the processing powers of the various devices keep increasing.

Backup Storage: In this aspect the security should outlast general storage security and the life span could be assumed

to be greater than thirty years, and as with normal storage services the technologies should be resistant to component and hardware changes as well as the algorithms used to store the data.

II Need and significance of proposed research Work

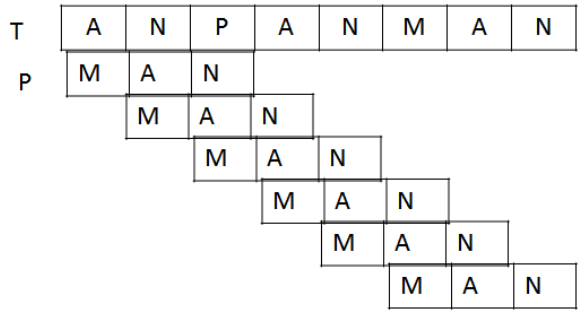
In this research of thesis, various issues of cloud computing are surveyed and analyzed by looking over current scenario. The main problem and issues are defined on data privacy/Security. The data uploaded by user can be manipulated by hackers or security breakers or can be stolen and uploaded right back away after manipulating. In our approach we are going to preserve the privacy of data manipulation to avoid it and detect it. For privacy, algorithm called, **Naive string search algorithm** is used in which each character of the pattern is compared to a substring of the text which is the length of the pattern, until there is a mismatch or a match. For prevention of intrusion using a unique key generation provision to user, i.e., every time user logins the system will generate the unique key which will be entered by user and he/she will be redirected to the content upload and download page automatically. For the data security NTRU Algorithm for encryption and Decryption is used. However our main area of research is to predict the data manipulation on cloud computing.

III Objectives:

- Implementing Naive String Search Algorithm for analyzing any type of manipulation in data
- Implement NTRU Algorithm for encryption and Decryption of data
- To compare it with current state of art techniques.

IV RESEARCH METHODOLOGY/PLANNING OF WORK

Naive string matching algorithm:
It is also known as Brute Force algorithm**.** It has no pre-processing phase, needs constant extra space. It always shifts the window by exactly one position to the right. It requires 2n expected text characters comparisons. It finds all valid shifts using a loop that checks the condition P[1....m]=T[s+1.....s +m] for each of the n-m+1 possible values of s . Consider the following example.
T=ANPANMAN
P=MAN
ANPANMAN A brute force method for string matching algorithm shown

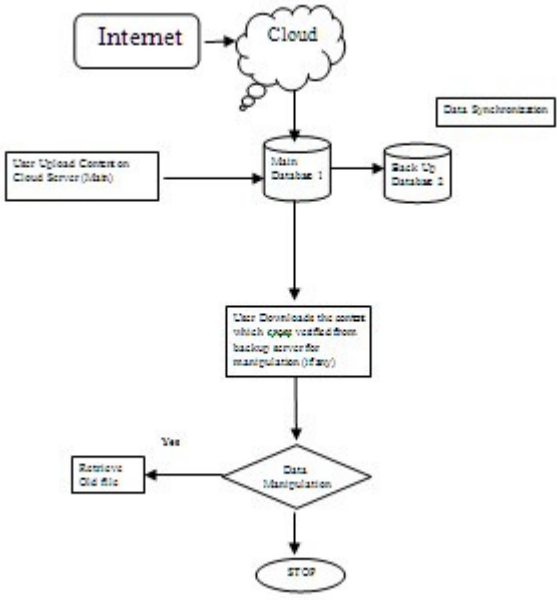


Figures: working of Naïve String Search Algoritm

Naive string matching algorithm takes time O((nm+ 1)m), and this bound is tight in the worst case. The worst case running time is thus O((n-m+1)m). The running time of Naive String Matching algorithm is equal to its matching time, since there is no preprocessing.

V Design and implementation

System flow design In this design we mentioned the entire steps we are going to perform in our entire work like uploading the content and then encrypting the data using NTRU algorithm with its cipher text formation and public and private key is enabled by NTRU and then the naive search algorithm will check any type of data manipulation and leakage of information and will respond according to the situation and the user will be notified immediately and if the results are positive after analysis then the user can download the content and after that decryption is performed.



System design of the proposed work

Algorithm Level Design

1. First the user will upload the data on the cloud servers.
2. Encryption of the data will happen with help of NTRU Algoritm.
3. The uploaded data will get store in 2 servers.
4. When user wants to download the data, the uploaded data will get decrypt and comes in original form.
5. Then the content of the data will be compared with the backup server and respond accordingly.
6. If the content matched then the user can download the data file
7. If the content doesn't match then the error of manipulation appears and the client can contact the service provider about this.

VI RESULTS AND DISCUSSION

In this research of thesis, various issues of cloud computing are surveyed and analyzed looking over current scenario. The main problem and issues are defined on data privacy/Security. The data uploaded by user can be manipulated by hackers or security breakers or can be stolen and uploaded right back away after manipulating. In this approach the manipulation in data file is detected with an algorithm called, **Naive string search algorithm** where each character of the file is matched with backup file until there is a mismatch or a match. For prevention of intrusion using a unique key generation provision to user, i.e., every time user logins the system will generate the unique key which shall be entered by user and he/she will be redirected to the content upload and download page automatically, however our main area of research is to predict the data manipulation on cloud computing and preventing using **NTRU encryption algorithm**.

In this research the encryption and decryption timing is calculated with regards to NTRU as to how they behave at different data packets and what their results are accordingly. So in this we will mainly focuses on the encryption and decryption timings and to analyze them as to get the answer of our problem via NTRU's better position than RSA and also to calculate and compare the the throughput of RSA and NTRU together. Moreover in this research the timing analysis of the naive String Search algorithm is also calculated which helps in detecting any type of data manipulation.

6.1 Performance analysis of NTRU and RSA with respect to encryption and decryption.

Encryption analysis of NTRU

Encryption Timings

Input size	RSA Timing	NTRU Timing
14 kb	10 ms	3 ms
24 kb	13 ms	6 ms
30 kb	16 ms	9 ms
39 kb	17 ms	9 ms
62 kb	17 ms	11 ms

The table constructed above shows the variations in the timings for the encryption in NTRU and RSA.

Decryption analysis of NTRU and RSA

Decryption Timings

Input size (Kb)	RSA Timing (ms)	NTRU Timing (ms)
14 kb	52 ms	32 ms
24 kb	61 ms	43ms
30 kb	66 ms	44 ms
39 kb	101 ms	51 ms
62 kb	142 ms	77 ms

The table constructed above shows the variations in the timings for the decryption in NTRU and RSA. Eg. when includes 62 kb input size for RSA, it will take 142 ms for decrypting the data, whereas in NTRU only 77 ms required for decryption.Therefore NTRU is speedier while decryption

6.2. Throughput Evaluation of NTRU and RSA.

Throughput is inversely proportional to number of recourses used, more the throughput less will be the resource utilization. Here the observations shows that throughput of NTR is more than that of RSA making the NTRU a much beneficial algorithm to use as there is less resource utilization than RSA.

Throughput evaluation of NTRU and RSA with regards to encryption

The throughput of NTRU is better than RSA it means that it uses less resources that is the reason why it is so efficient.

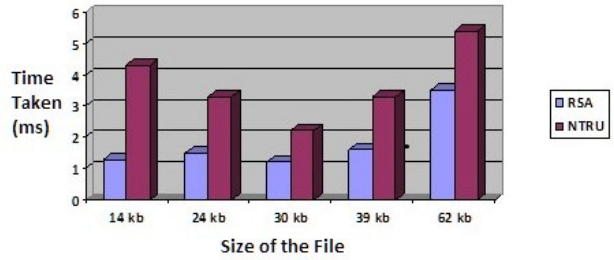


Figure : Throughput Evaluation of Encryption

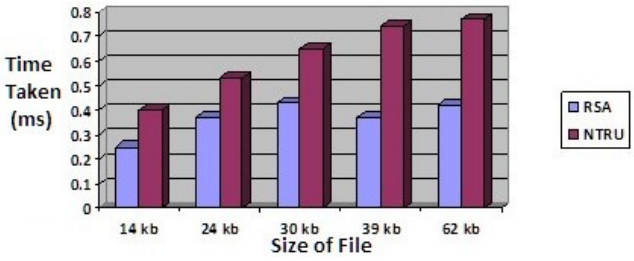
Throughput for encryption

Input size (Kb)	RSA throughput (kb/sec)	NTRU throughput (kb/sec)
14 kb	1.3	4.3
24 kb	1.5	3.3
30 kb	1.2	2.2
39 kb	1.6	3.3
62 kb	3.5	5.4

The table constructed above shows the throughput for the encryption in NTRU and RSA. When includes 62 kb input size for RSA, it will give 3.5 kb/sec throughput during encryption, whereas in NTRU generates 5.5 kb/sec throughput. Therefore more the throughput less will be the consumption of resources, so NTRU is better than RSA.

Throughput evaluation of NTRU and RSA with regards to Decryption

Here shown that the throughput of NTRU is better than RSA it means that it uses less resources that is the reason why it is so efficient.



Throughput for decryption

Input size (Kb)	RSA throughput (kb/sec)	NTRU throughput (kb/sec)
14 kb	0.25	0.40
24 kb	0.37	0.53
30 kb	0.43	0.65
39 kb	0.37	0.74
62 kb	0.42	0.77

The table constructed above shows the throughput for the decryption in NTRU and RSA. When includes 62 kb input size for RSA, it will give 0.42 kb/sec throughput during decryption, whereas in NTRU generates 0.77 kb/sec throughput. Therefore more the throughput less will be the consumption of resources, so NTRU is better than RSA.

Analysis of time taken by Naive String Search Algorithm

Time Computation of Naive String Search Algorithm

Input size (Kb)	Naive String Timing (ms)
14 kb	11ms
24 kb	18 ms
30 kb	29 ms
39 kb	37 ms
62 kb	66 ms

The table above describes the time taken by naive string search algorithm while performing the detection of any manipulation in the data files. The file which is downloadable by the client will compare with the backup server. Here the data files of different sizes are analyzed and time taken by the algorithm is calculated which is shown below in the table

VII CONCLUSION AND FUTURE SCOPE

The research concluded that the implementation is been carried out is successful which opened the many number of advancement towards this particular area of field in cloud computing. The naive bayes string search algorithm was successful in finding the manipulation and detection which gives a better idea of trust level towards the cloud storage providers. The application for its fastness was implemented along with the NTRU. It was observed that NTRU is been a fastest encryption algorithm gave a better outcome of security trust towards the cloud service purchasers. The manipulation is nearly impossible in this case but however looking over future perspectives we cannot guarantee the 100% achievement of security. This research can help in monitoring the passive attacks on particular algorithm since in passive attack the attacker manipulate the file and send it back to destination after stealing the content. Moreover majority number of traffic is shifting from PC's to mobile. Users don't like to sit for long over pc's to access the internet as its easily available on smart phone devices. The era of 4G now is shifting most of its users to mobile devices hence this research is also needs to be enabled on mobile cloud area network, which helps in monitoring this area also among the masses. By this approach it can cover the whole mass of network.

REFERENCES

[1] Bhadauria, R., & Sanyal, S. (2012). Survey on security issues in cloud computing and associated mitigation techniques. *arXiv preprint arXiv:1204.0764*.

[2] Bhadauria, R., Chaki, R., Chaki, N., & Sanyal, S. (2011). A survey on security issues in cloud computing. *arXiv preprint arXiv:1109.5388*.

[3] Behl, A. (2011, December). Emerging security challenges in cloud computing: An insight to cloud security challenges and their mitigation. In *Information and Communication Technologies (WICT), 2011 World Congress on* (pp. 217-222). IEEE.

[4] Dan Boneh, Eu-Jin Goh, and KobbiNissim. Evaluating 2-DNF formulas on ciphertexts. In Theory of Cryptography Conference, TCC'2005, volume 3378 of Lecture Notes in Computer Science, pages 325-341. Springer, 2005.

[5] Dhage, S. N., & Meshram, B. B. (2012). Intrusion detection system in cloud computing environment. *International Journal of Cloud Computing*, *1*(2), 261-282

[6] Jansen, W. A. (2011, January).Cloud hooks: Security and privacy issues in cloud computing. In *System Sciences (HICSS), 2011 44th Hawaii International Conference on* (pp. 1-10). IEEE.

[7] Jansen, W., & Grance, T. (2011). Guidelines on security and privacy in public cloud computing. *NIST special publication*, *800*, 144.

[8] Jaeger, P. T., Lin, J., & Grimes, J. M. (2008). Cloud computing and information policy: Computing in a policy cloud?. *Journal of Information Technology & Politics*, *5*(3), 269-283.

[9] Julien Bringe and al. An Application of the Goldwasser-MicaliCryptosystem to Biometric Authentication, Springer-Verlag, 2007.

[10] Ko, R. K., Jagadpramana, P., Mowbray, M., Pearson, S., Kirchberg, M., Liang, Q., & Lee, B. S. (2011, July). TrustCloud: A framework for accountability and trust in cloud computing. In *Services (SERVICES), 2011 IEEE World Congress on* (pp. 584-588). IEEE

[11] Kolodner, E. K., Tal, S., Kyriazis, D., Naor, D., Allalouf, M., Bonelli, L & Wolfsthal, Y. (2011, November). A cloud environment for data-intensive storage services. In *Cloud computing technology and science (CloudCom), 2011 IEEE third international conference on* (pp. 357-366). IEEE.

[12] Mather, T., Kumaraswamy, S., & Latif, S. (2009). *Cloud security and privacy: an enterprise perspective on risks and compliance*. " O'Reilly Media, Inc.".

[13] Michener, W. K., & Jones, M. B. (2012). Ecoinformatics: supporting ecology as a data-intensive science. *Trends in ecology & evolution*, *27*(2), 85-93.

[14] Mishra, R., Dash, S. K., Mishra, D. P., & Tripathy, A. (2011, April). A privacy preserving repository for securing data across the cloud. In *Electronics Computer Technology (ICECT), 2011 3rd International Conference on* (Vol. 5, pp. 6-10). IEEE.

AUTHORS PROFILE

Indu Sharma.is a M.tech student in Rayat and Bahra college Sahuaran. Her areas of interest are Operating System, Cloud Computing.

Mandeep Kaur Kang is a Assistant Professor in Rayat and Bahra college Sahuaran. Her area of interest is Digital Image processing, Cloud computing.