Cloud Computing Assignment 2

Name : P Chandra Shekar
Roll No. : S20200010154

SUMMARY OF PAPER ONE :

TITLE : Optimization of Cryptography Algorithms in Cloud Computing

Introduction :

- Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.
- Cloud computing is becoming famous now a days and gaining high momentum
- Gaining more popularity attracts more attacks
- Shows how cryptography techniques can be incorporated while sending files on cloud.
- To encrypt the files RSA and DES are used.

Objective of the work :

- Nowadays cloud computing is gaining very good attention. So securing the data while in transit and at rest is very important.
- Cloud security is important because of the following reasons:
    - **Managing** : The data which is in cloud is not owned by the data owner because all of the work is handled by remote machines present somewhere else
    - **Access of Data** : As all the information is stored on the cloud, the cloud service provider(CSP) can misuse this information by giving it to some other competitors who are also in the same field.This will result in loss of the amount which the data owner might have received if the data was not lost.

      So, we can conclude that integrity, authentication and confidentiality of the data is a matter of concern.

- All the above mentioned concerns pose one major problem that is unwanted exposure of data to the cloud service provider which can be misused by the CSP which can result in software not working properly or cause some wanted bugs.

Major Contributions :

The solutions used to protect the privacy of the data is by :

- Using None of your business (NOYB)
- Privacy Manager
- Content Cloaking (CoClo)

Approach Referred :

- In this paper CoClo method of privacy is followed, which means, before any data is inserted into the cloud it should be encrypted, so that even the CSP will not have access to the underlying data.

- RSA and DES are used in this process to encrypt the data.

- It includes three steps as discussed in the paper :

    - **Preparation of data :** The data has to be encrypted should be preprocced.Here in this paper they accept only text input i.e, only files of type .java, .m, .txt are being allowed.No other type of data like pictures are not allowed.
    - **Encryption** : After the pre-processing of the data is complete they are then encrypted using RSA or DES. The encrypted file is now a binary file.
    - **Transfer of data :** Since we have encrypted the data using RSA and DES the file is now safe as no one can have the original data except the owner of the data.In this manner authentication, integrity and confidentiality is achieved.So Data Privacy in cloud is addressed by the above mentioned methods.

Experimental Details :

- Here RSA and DES are used to encrypt the data.
- Their performances using DES and RSA have been compared.

- A comparative analysis is given

Performance measuring metrics :

- **Encryption time :** The time taken by the algorithm to encrypt the input data.It means converting the plain text to cipher text.
- **Decryption time :** Time taken by the algorithm to decrypt the data.It means to convert the cipher text to plain text.
- **Total time in execution :** Total time taken to encrypt and decrypt.

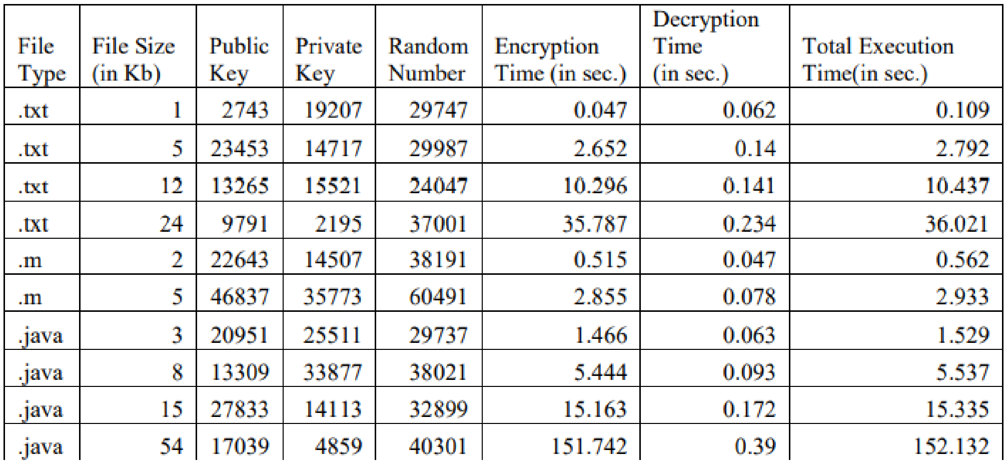
Total time = encryption time + decryption time

Results/Outcomes :

- Here RSA and DES algorithms are used to compare by giving the same input and judge them based on their total execution time.

- **Using RSA :**

Table 1 : RSA Encryption and Decryption time.



File Type	File Size (in Kb)	Public Key	Private Key	Random Number	Encryption Time (in sec.)	Decryption Time (in sec.)	Total Execution Time(in sec.)
.txt	1	2743	19207	29747	0.047	0.062	0.109
.txt	5	23453	14717	29987	2.652	0.14	2.792
.txt	12	13265	15521	24047	10.296	0.141	10.437
.txt	24	9791	2195	37001	35.787	0.234	36.021
.m	2	22643	14507	38191	0.515	0.047	0.562
.m	5	46837	35773	60491	2.855	0.078	2.933
.java	3	20951	25511	29737	1.466	0.063	1.529
.java	8	13309	33877	38021	5.444	0.093	5.537
.java	15	27833	14113	32899	15.163	0.172	15.335
.java	54	17039	4859	40301	151.742	0.39	152.132

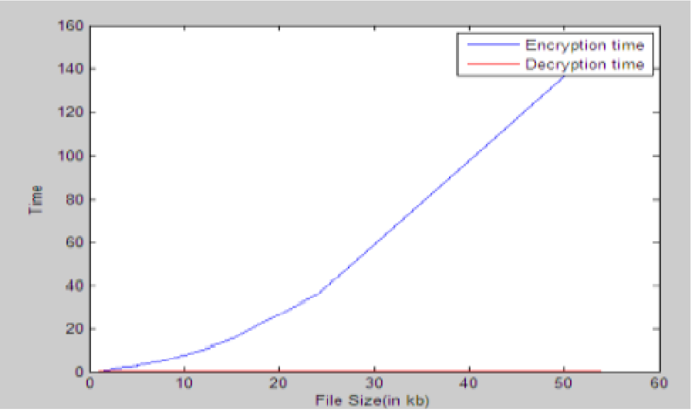


Figure 3: Graph of RSA

- Using DES :

Table 2: DES Encryption and Decryption time.

File Type	File Size (in Kb)	Encryption Time (in sec.)	Decryption Time (in sec.)	Total execution Time(in sec.)
txt	1	0.265	0.062	0.327
.txt	5	0.281	0.14	0.421
.txt	12	0.281	0.218	0.499
.txt	24	0.281	0.344	0.625
.m	2	0.265	0.11	0.375
.m	5	0.265	0.219	0.484
.java	3	0.265	0.125	0.39
.java	8	0.266	0.187	0.453
.java	15	0.266	0.327	0.593
.java	54	0.281	0.733	1.014

It can be seen from the table that encryption and decryption time even for the large file is in microseconds. This table is analyzed using the following graph:

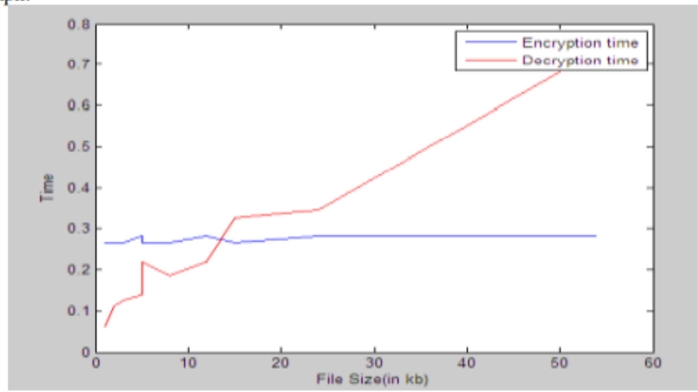


Figure 4: Graph of DES

- Now comparing both RSA and DES :

Table 3: Comparative analysis table of RSA and DES

File Type	File Size (in Kb)	RSA (Total execution Time)	DES (Total Execution Time)
.txt	1	0.109	0.327
.txt	5	2.792	0.421
.txt	12	10.437	0.499
.txt	24	36.021	0.625
.m	2	0.562	0.375
.m	5	2.933	0.484
.java	3	1.529	0.39
.java	8	5.537	0.453
.java	15	15.335	0.593
.java	54	152.132	1.014

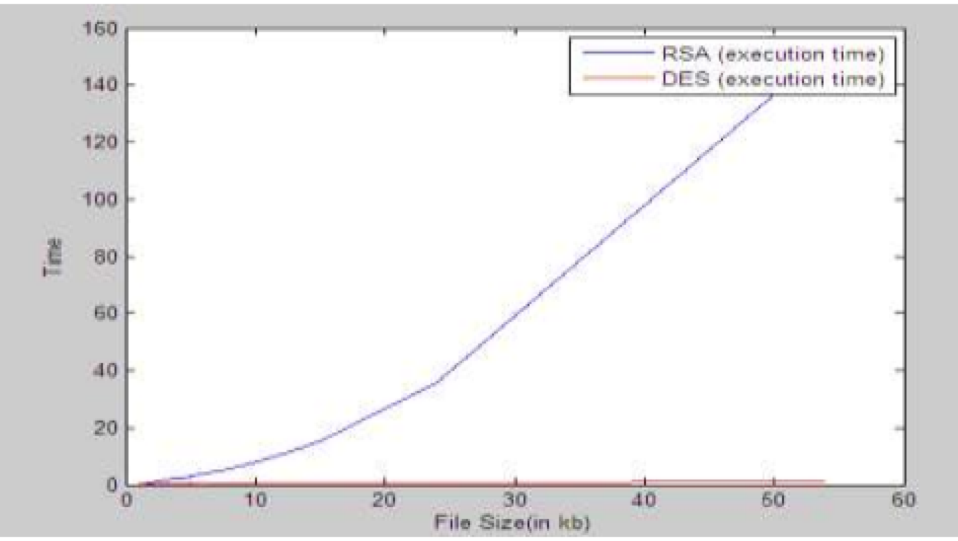


Figure 5: RSA and DES Comparative analysis

- With this they have concluded that DES is much faster than RSA and should be used if time is the main constraint.But from a higher security point of view RSA gives stronger protection and should be preferred more when security is of the highest concern.

Limitations of current work and future scope :

- The current work has only taken into consideration the text inputs but not proposed any way for other formats of data.
- The future works of this as the paper says include :
  - Using hybrid approaches like, we can instead of using a single named algorithm we can use multiple algorithm to encrypt the data.
  - The results brought from the RSA and DES can be used in integration.

Observations :

1. How does the design of the study address the research questions?

- The main concern of the study was to find a way to have integrity, confidentiality, authentication of the data stored in the cloud so that the data owner can feel safe while storing his data in the cloud.So they have used RSA and DES algorithm to do the same.Now since they have used this approach, they were able to achieve all the three objectives of data security prior to sending the data to the cloud.

2. How convincing are the results? Are any of the results surprising?

- The results were obtained by using RSA and DES for the same type of input.We can see that the time factor for encrypting the file using RSA was higher than DES.That is what is actually expected because we know that RSA is much bulkier and rigorous than DES.

3. What does this study contribute toward answering the original question?

- This study contributed towards how the client can be sure that his data is not being used or manipulate the data without his permission, or used to create bugs in the software.This has been achieved in the paper using cryptography using RSA and DES.

4. What aspects of the original question remain unanswered?

- As they have mentioned the title as Optimizations of cryptography in cloud computing, they have only shown how the cryptographic algorithms can be used to secure the data, but they have not optimized this process as a whole.This aspect of optimizations remain unanswered

SUMMARY OF PAPER TWO:

TITLE :A Novel Virtualization Enabled Cloud Infrastructural Framework for Enhancing Private Cloud Communication Security

Introduction :

- Approximately 2.5 quintillion bytes of data is generated everyday.
- To compute this mammoth amount of data for business intelligence contemporary methods are not sufficient.
- This introduces a new way of solving the problem, i.e, using non-conventional cloud computing technology.
- AWS kinesis and AWS Kafka are examples of high volume data analysis
- Virtualization, known as the backbone of cloud computing infrastructure, increases the computation power by using up the under-used computing resources.
- Since everything has a disadvantage, virtualization also brings up a mahot challenge with it, i.e, attackers have endless possibilities to attack the systems

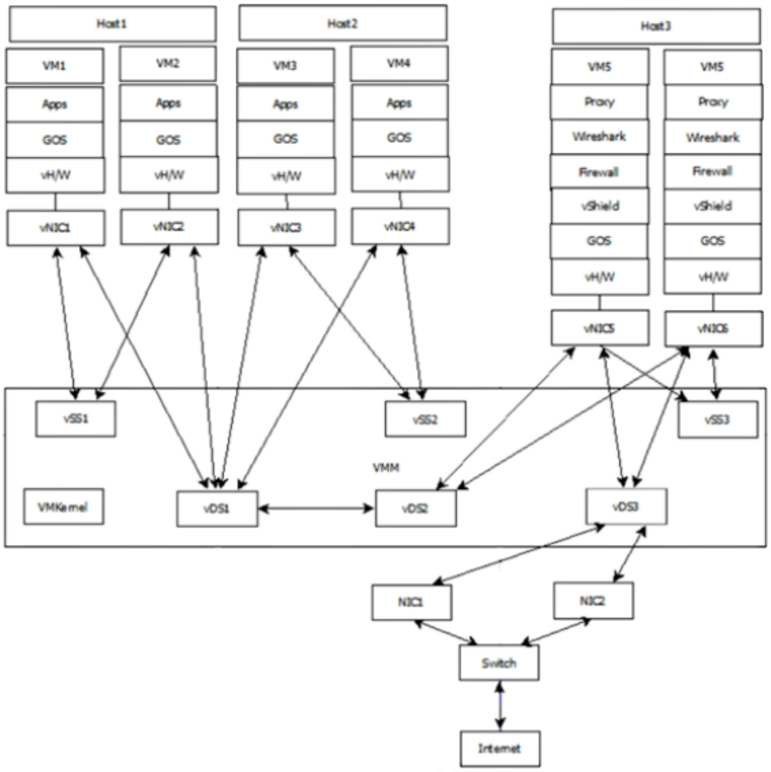
Objectives of the work :

- In this paper the majority concern was how to stop VMs from spreading the infection from an already compromised VM to a healthy VM so that remaining VMs can continue to work in a healthy state and also a methodology where the affected VM can be respawned.

Approach Referred :

- In this paper they have designed a framework where all the communication traffic must go on a predefined path and all other VMs will reveal nothing about them to the real world.
- This will still incorporate all the properties of virtualization ensuring isolation and other properties.
- In this paper they have taken a single physical machine and created three hosts from them.
- Each host will have 2 VMs in them which run either servers or clients.
- Now lets say these VMs have to connect to the internet they have no other way than connecting using Host 3

- Now since we are restricting the path of connecting to the internet, we are reducing the attack surfaces.
- Meanwhile, the VM3 is equipped with all the security measures.Host 3 is used as a filter for all the incoming and outgoing traffic.The filter here is just a basic firewall as mentioned in the paper.
- The following diagram shows the implementation :



- They have used a combination of standard switches and distributed switches to communicate between the VMs and the outer world, i.e, the Internet.
- **Communications Explained :**
  - If VM1 and VM2 want to communicate they are passes through VSS1
  - If VMs of different hosts want to communicate then they are passed through a distributed switch because the implementation of security features are more complex than VSS.This will ensure that the infection doesn't pass through to other VMs.
  - If any VM wants to communicate to the internet it should contact Host 3 VM, through the layered virtual distributed switches VDS1 and VDS2 and get redirected via another virtual distributed switch vDS3 to the physical switch through NIC1 (or NIC2, if required).

    - This way we can make sure that only, if infected, Host 3 will be infected.
- Here two VMs were used as filters so that a single point of failure does not occur.
- Following this method any infection of physical or logical resources falls down steeply.

Performance Metrics :

- The model is tested using the following equation :

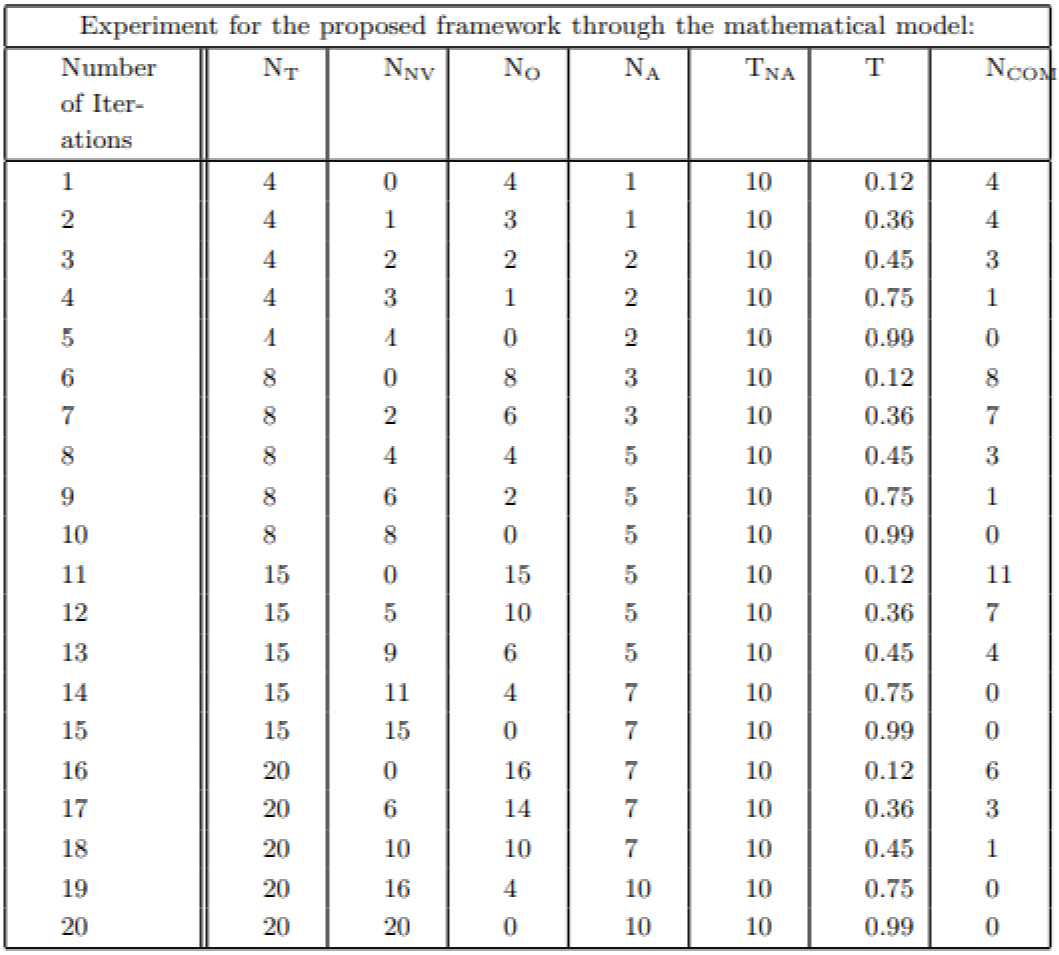
$$T = \alpha \left(N_A^{-C} \right)$$

Where T is Threat coefficient factor,
α is inverse coefficient
Na is list of attacking VMs
C is a constant

Experimental Details :

- The test-bed here prepared has been gone through a set of observations and based on the equation presented above
- NT represents the total number of VMs in the framework
- NNV represents invulnerable VMs
- NO represents the total number of Open VMs that are susceptible to attack and
- NA are the attacking VMs.
- TNA represents the total number of attacks in every iteration that has been taken as one of the parameters for the simulation.T is the threat factor that represents the probability of the NO VMs that can be compromised.
- NCOM represents the total number of newly compromised VMs after every iteration

<u>Results</u>:

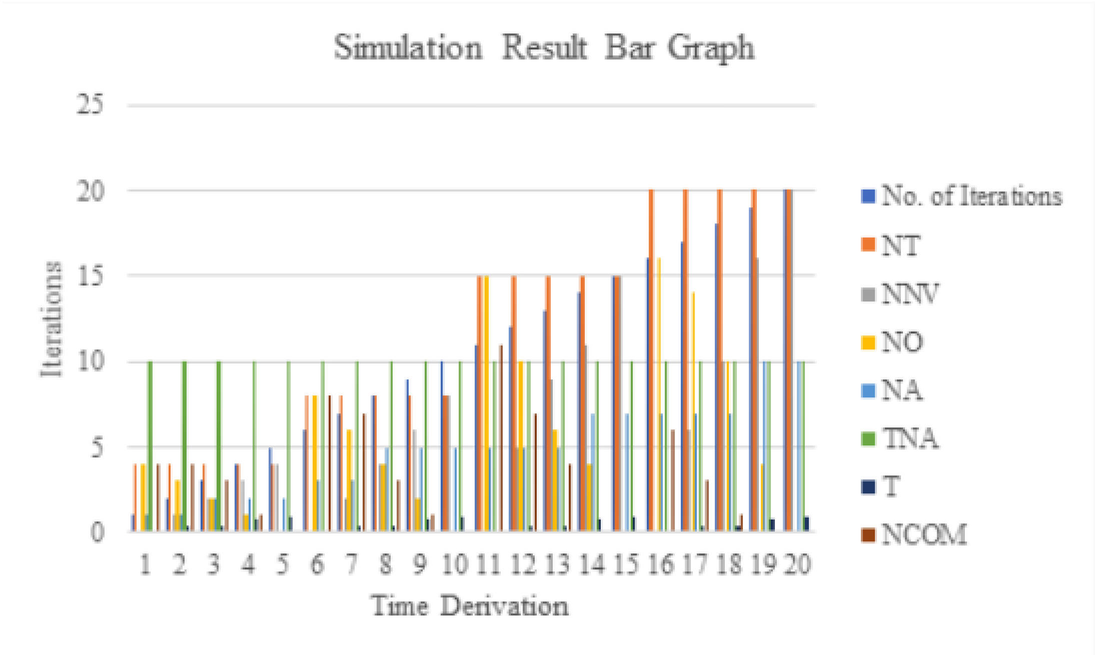


Experiment for the proposed framework through the mathematical model:

Number of Iterations	N_T	N_{NV}	N_O	N_A	T_{NA}	T	N_{COM}
1	4	0	4	1	10	0.12	4
2	4	1	3	1	10	0.36	4
3	4	2	2	2	10	0.45	3
4	4	3	1	2	10	0.75	1
5	4	4	0	2	10	0.99	0
6	8	0	8	3	10	0.12	8
7	8	2	6	3	10	0.36	7
8	8	4	4	5	10	0.45	3
9	8	6	2	5	10	0.75	1
10	8	8	0	5	10	0.99	0
11	15	0	15	5	10	0.12	11
12	15	5	10	5	10	0.36	7
13	15	9	6	5	10	0.45	4
14	15	11	4	7	10	0.75	0
15	15	15	0	7	10	0.99	0
16	20	0	16	7	10	0.12	6
17	20	6	14	7	10	0.36	3
18	20	10	10	7	10	0.45	1
19	20	16	4	10	10	0.75	0
20	20	20	0	10	10	0.99	0

- The above table shows the amount of infected VMs are decreasing as T is increasing which is in accordance with the equation above.
- So this, ensures that the private cloud infrastructure security is enhanced

TABLE 1
Simulation Results



The above bar graph points out that the difference in the inverse ratio increases with pro-portion to time derivation. It signifies that the number of newly compromised VMs are reducing. As a result, the risk factor for open VMs is reducing and hence the system is becoming safer

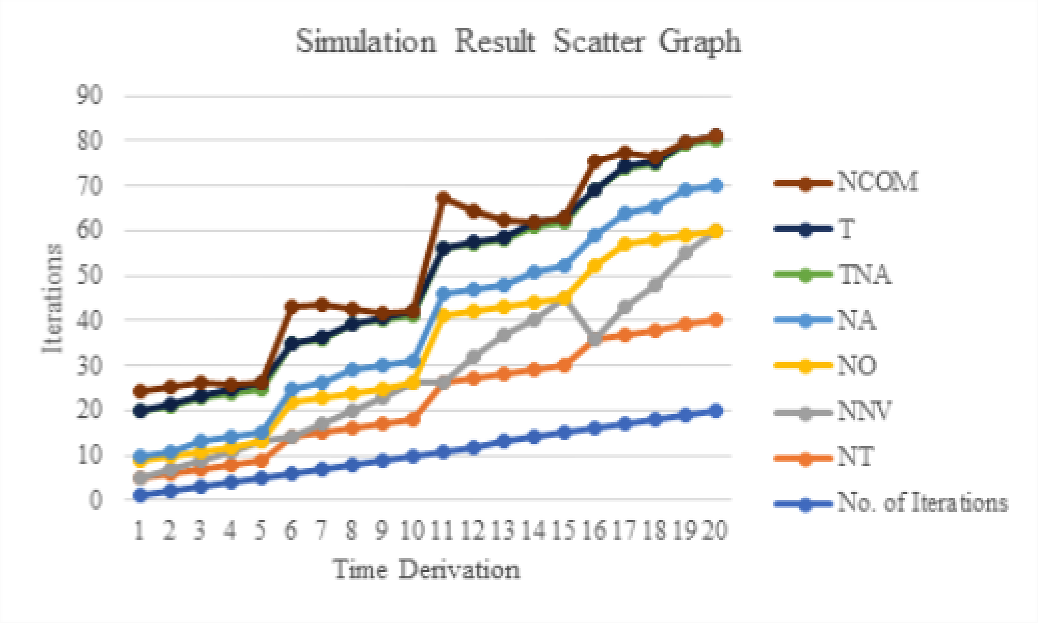


Fig. 4. Scatter graph of simulation results

- The scatter graph above depicts the peak values for each parameter along with time derivation.

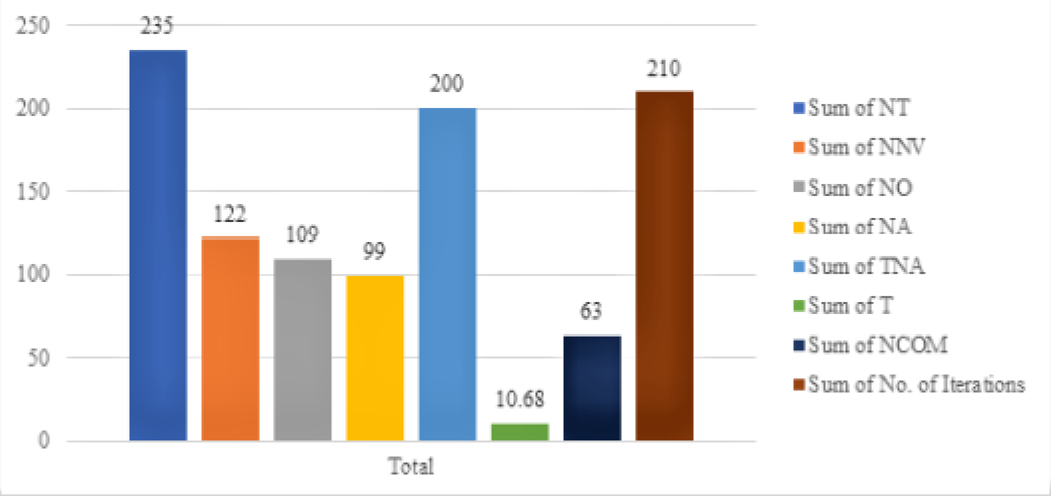


Fig. 5. Pivot Chart of simulation results

- The above pivot chart shows the aggregated values for all the parameters used in the simulations. Here we can observe that the total number of newly compromised VMs (NCOM) is only 63 while total number of VMs (NT) is 235 and total number of open VMs (NO) is 109. Hence, we can say that a lot of susceptible VMs have not been compromised by following the mathematical model and infrastructural framework designed.

Limitations of the work and future scope:

- The proposed design can protect a private cloud environment from different unknown attacks from the outer world. The proposed design can protect all the VMs under consideration through a master filter (VMs within a Host). This study has been carried out on the VMware platform.
- Further enhancement in security policy can be achieved through extensive research

Observations :

1. How does the design of the study address the research questions?

    - The process followed in this paper is as follows :
        - They have used a physical machine and created 3 hosts in which only one of them is responsible for communicating with the internet and all other VMs are connected using this VM.
    - The main concern here was how to spread the infection from one VM to another.
    - So by maintaining a single point of communication, the risk of spreading the infection decreases which is the main research question in this paper

2. How convincing are the results? Are any of the results surprising?

    - The results were quite satisfactory with regards to the question they had in hand.
    - They were able to come up with something that could prevent the spreading of infection.
    - We see that T is decreasing with the rise in the amount of VMs that are being infected which is in accordance with the equation that they have developed using their mathematical model

3. What does this study contribute toward answering the original question?

    - The study was able to answer one important question, that is how to protect the VMs.
    - Although VM have many advantages, its security remains a point of question when used in large scale, because of how an infection can easily spread between the VMs

- Now the method they have proposed for this question which has been discussed briefly shows how they were able to achieve the ability to decrease the spread of infections which is also reflected in the experiments they have conducted.

4. What aspects of the original question remain unanswered?

- As per the question they had in hand they were able to come up with a very good approach to solve it.
- The aspects which I think unanswered are how they are going to make this work on a large scale.How will they cope with single point of failure, because even though they have used 2 VMs to cope with SPOF, what if both of them are down.This brings up the question of how much time will a VM require if it is down.Now if both of them are not responding then connecting to the internet which passes through these VM, how will they handle connecting to the internet? Will the connection to the internet still be possible ? Will the other VMs still be as secure as they were if these controller VMs were present ?