Cloud Based Malware Detection Technique

Authors

Manish Kumar Gupta
Sagar Shaw
Prof. Sanjay Chakraborty

Institute Of Engineering & Management,
Kolkata

Overview

- Abstract
- Introduction
- Signature Based Detection Is Insufficient
- Proposed Work
- Conclusion & Future Work
- References

Abstract

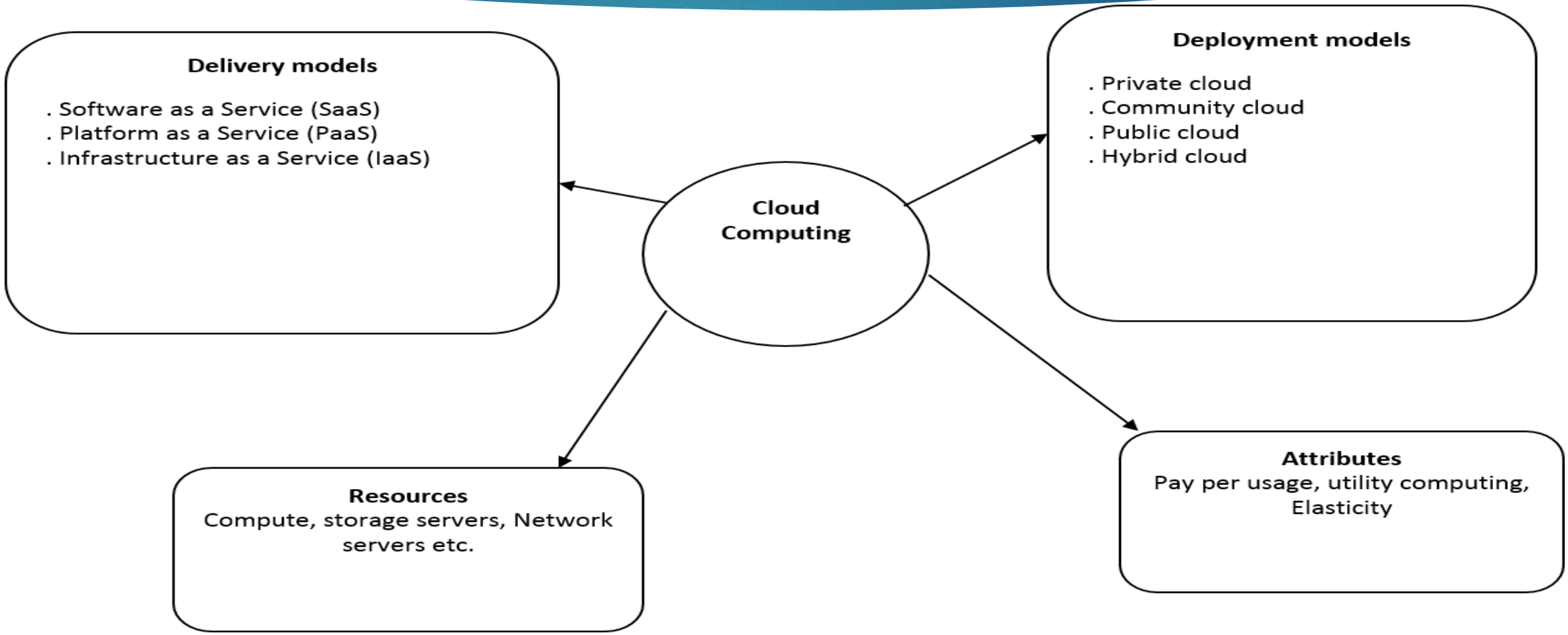
- Security is one of the major concerns in cloud computing.
- Many antivirus software unable to detect many modern malware threats.
- This work counsel a new model for malware detection on cloud architecture.
- In this work we use combined detection techniques, DNA Sequence Detection Process, Symbolic Detection Process and Behavioural Detection Process.
- The Proposed approach (PMDM) can be deployed on a VMM which remains fully transparent to guest VM and to cloud users. PMDM prevents the malicious code running in one VM (infected VM) to spread into another non-infected VM with help of hosted VMM.
- A prototype of PMDM is partially implemented on one popular open source cloud architecture – Eucalyptus.

Introduction

Antivirus software is one of the most widely used tools for detecting and stopping malicious and unwanted files. However, traditional host based antivirus is now questionable.

Antivirus software fails to detect many modern threats and its increasing complexity has resulted in vulnerabilities that are being exploited by malware.

Cloud Architecture



What Is Malware

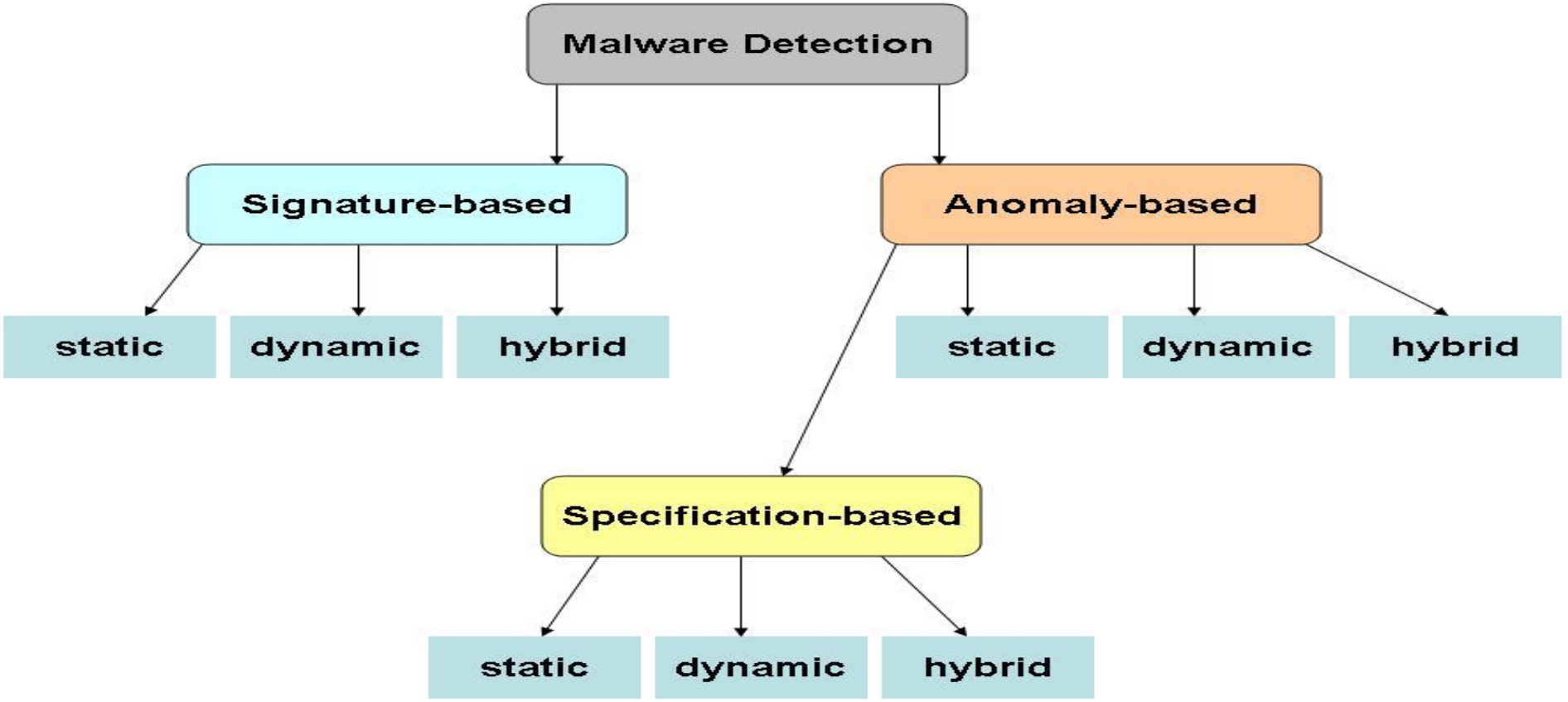
- A Malware is a set of instructions that run on your computer and make your system do something that an attacker wants it to do.
- For Example :

  Crash, Hang, Compromised privacy.

Type Of Malware

- Virus
- Trojan horse
- Scare ware
- Adware
- Worm

Malware Detection Techniques



Anomaly Based Detection

- An anomaly-based detection technique uses its knowledge of what constitutes normal behaviour to decide the maliciousness of a program under inspection.
- Anomaly-based detection usually occurs in two phases–a training (learning) phase and a detection (monitoring) phase.
- In training phase the detector attempts to learn the normal behavior. The detector could be learning the behavior of the host.

Specification Based Detection

- Specification-based detection is a special type of anomaly-based detection.
- Specification-based detection specifies a set of valid behavior to decide the maliciousness of a program.
- The main limitation of specification-based detection is that it is often difficult to specify completely and accurately the entire set of valid behaviors a system should exhibit.

Signature Based Detection

- Signature-based detection attempts to model the malicious behavior of malware. This model of malicious behavior is often referred to as the signature.
- Ideally, a signature should be able to identify any malware exhibiting the malicious behavior specified by the signature. Like any data that exists in large quantities which requires storage, signatures require a repository. This repository represents all of the knowledge the signature-based method has, as it pertains to malware detection.

Survey Work (1)

- Signature Optimizing Pattern Matching :

This method is used depends on the signature, which storage already in the database.

For this purpose, they used a string matching algorithm, comparison variants of which arise in finding similar DNA or protein sequences. [1] .

Survey Work (2)

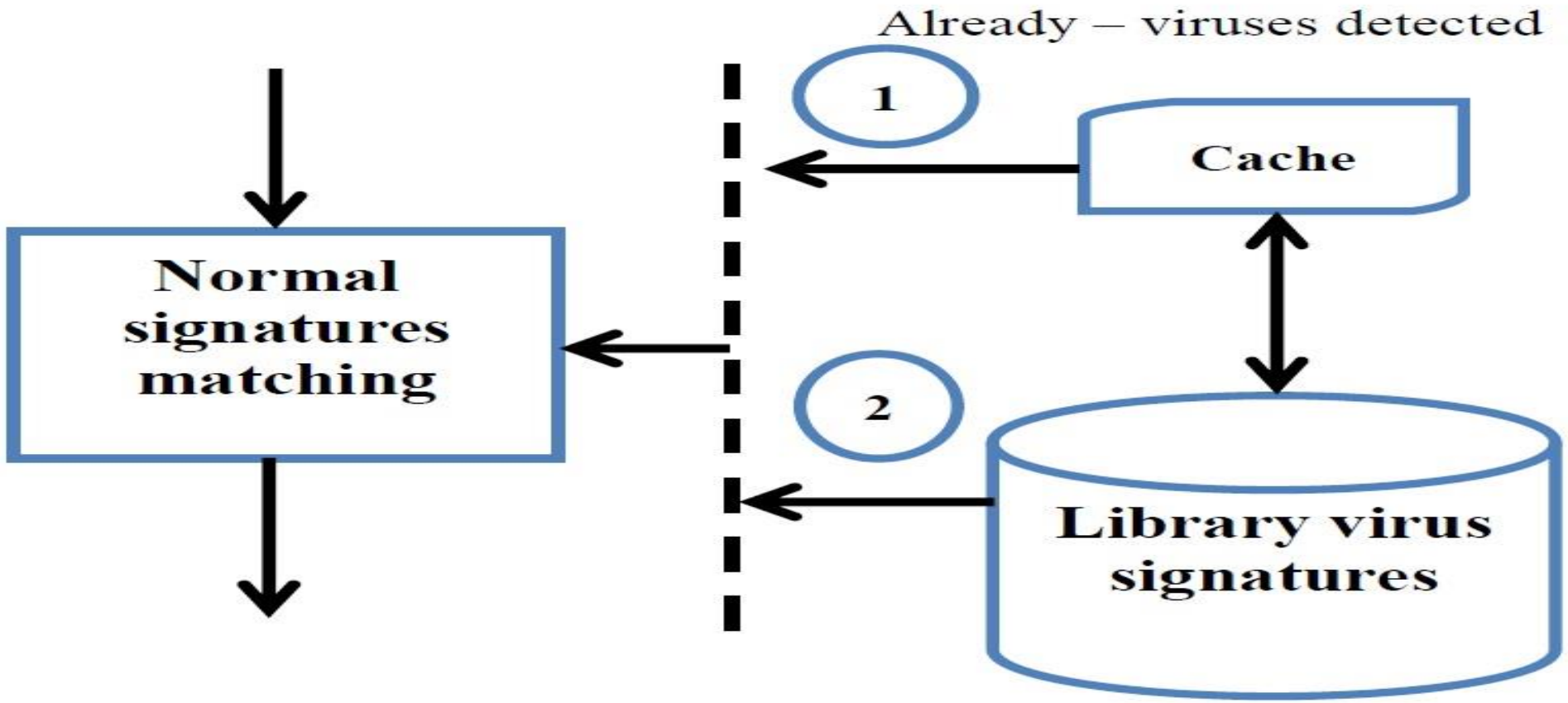
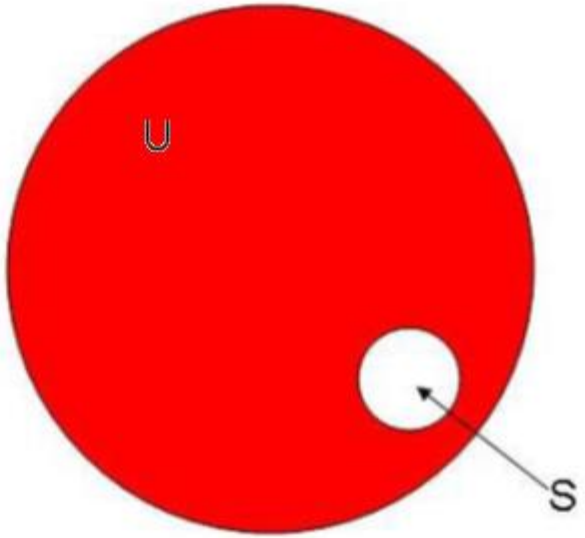


Fig. 2. Shows the process for Optimizing Pattern Matching of Library

Signature Based Detection Is Insufficient

U = set of all malicious behavior
S = set of all known signatures



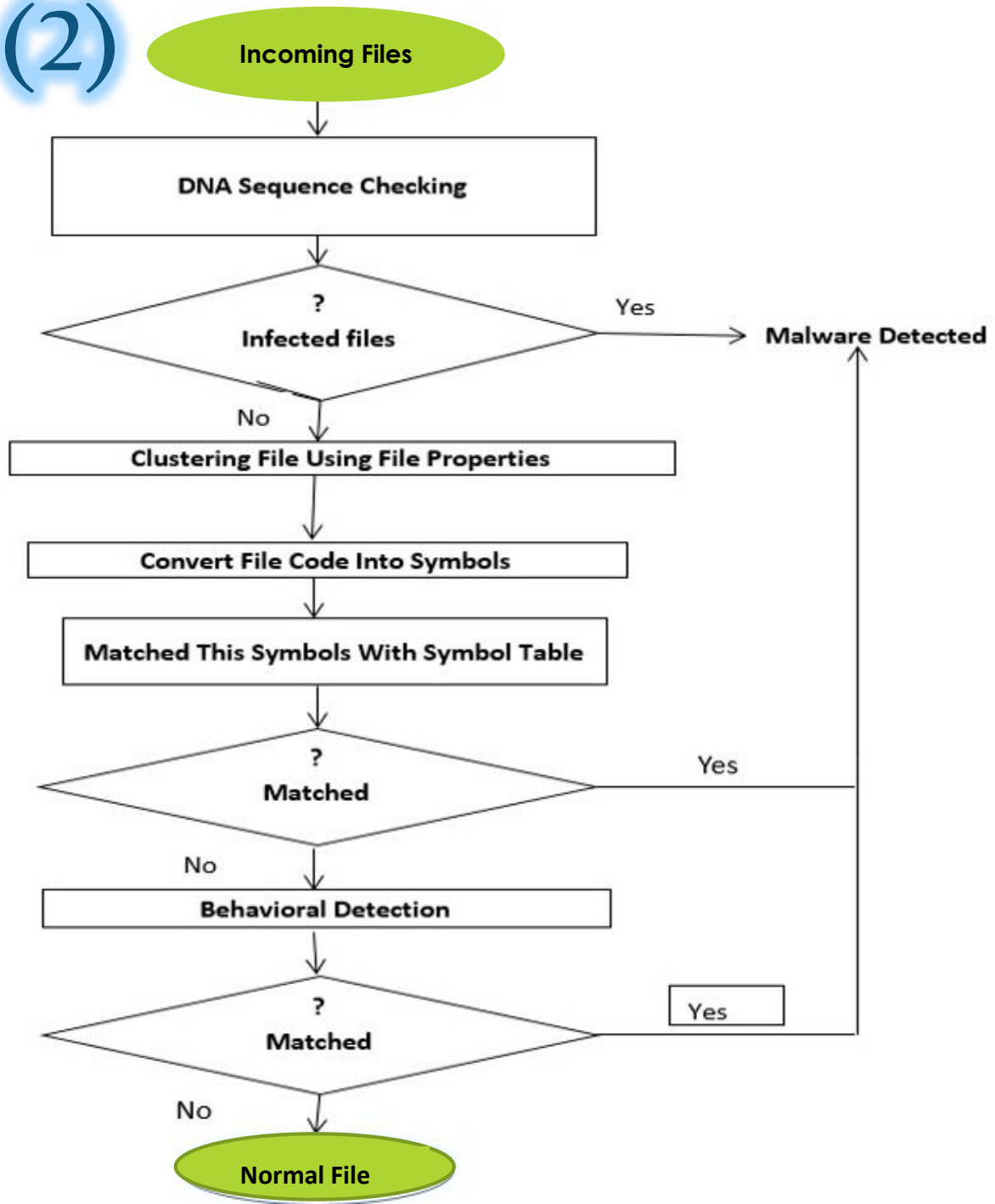
Pitfalls Of Signature Based Detection

- Most Conservative
- No zero day protection
- Post infection protection
- Cannot cope with malware variants
- False positives

Proposed Work (1)

- Step 1 : Start
- Step 2 : DNA Sequence Detection Process.
- Step 3 : Symbolic Detection Process.
- Step 4 : Behavioural Detection Process.
- Step 5 : Cloud Deployment Model.
- Step 6 : End.

Proposed Work (2)



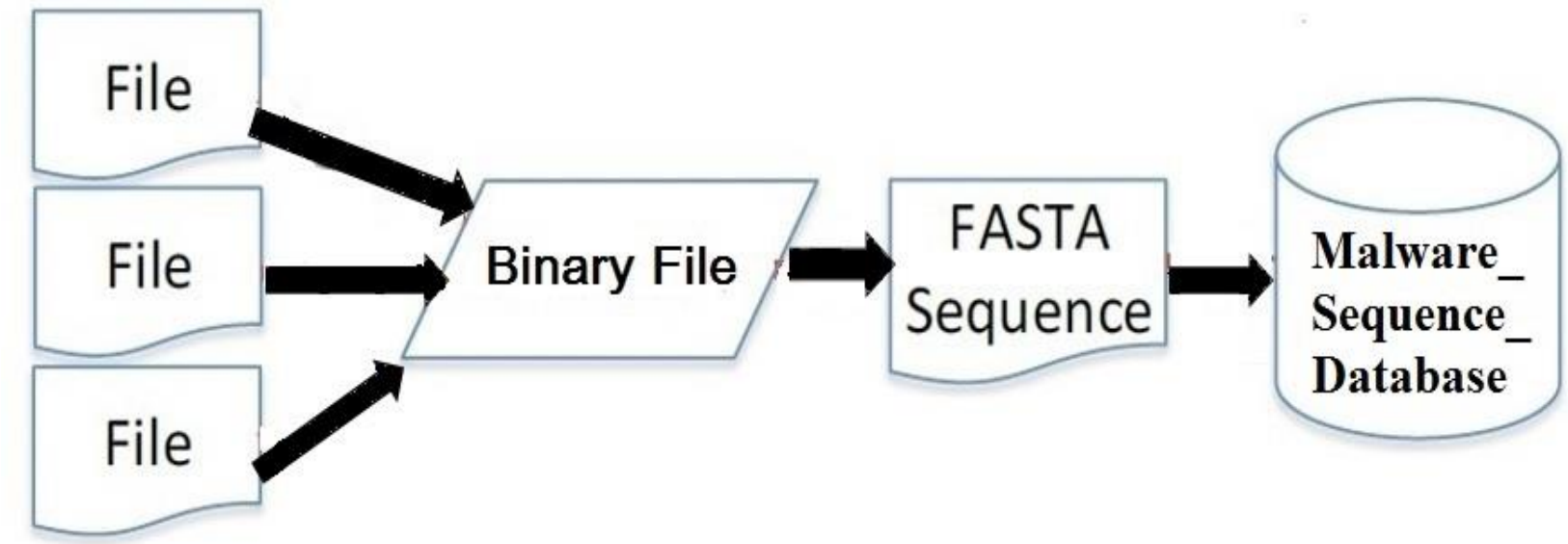
Proposed Work (3)

- Extraction of DNA sequence from a file.

Binary Bits	DNA Character
00	T
01	G
10	C
11	A

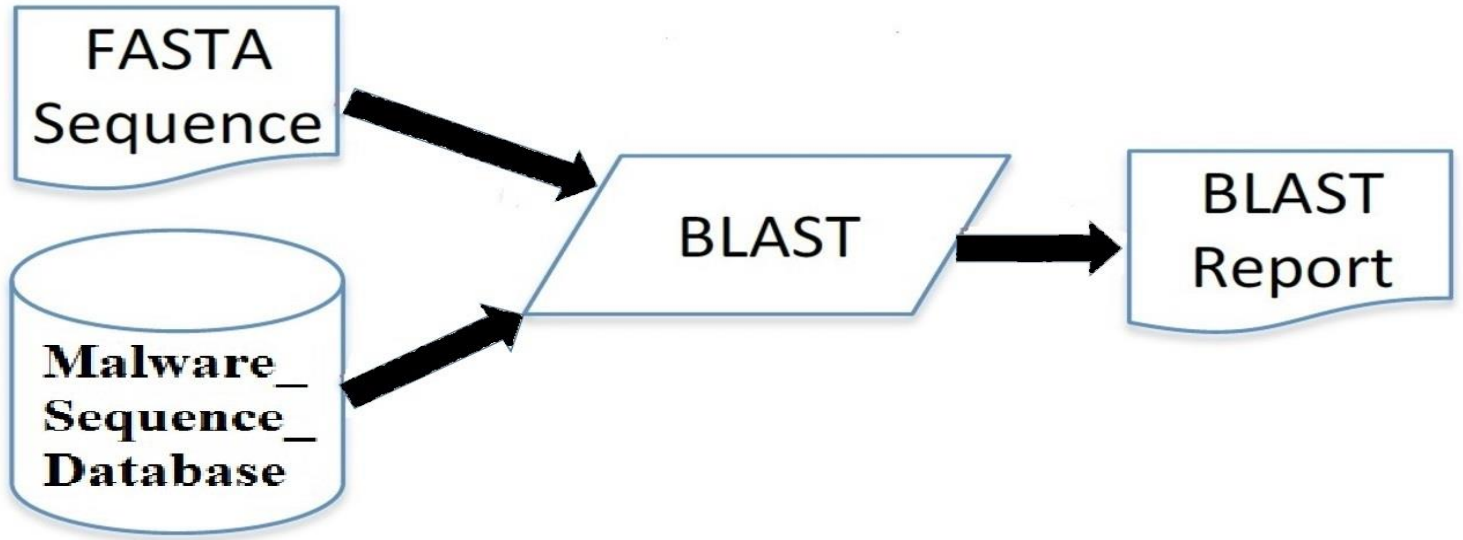
Proposed Work (4)

- Creating a Malware_Sequence_Database.



Proposed Work (5)

- Comparing FASTA Sequence with Malware_Sequence_Database.



- Using Blast Software [1].

Proposed Work (6)

- The result of this comparison is a BLAST report.
- And, malware detected file will be blocked.
- Otherwise proceed for next process i.e. Symbolic detection.

- In Process 2 the files which are pass through the first process are only go for the second process.

Proposed Work (7)

Cluster Using File Properties :

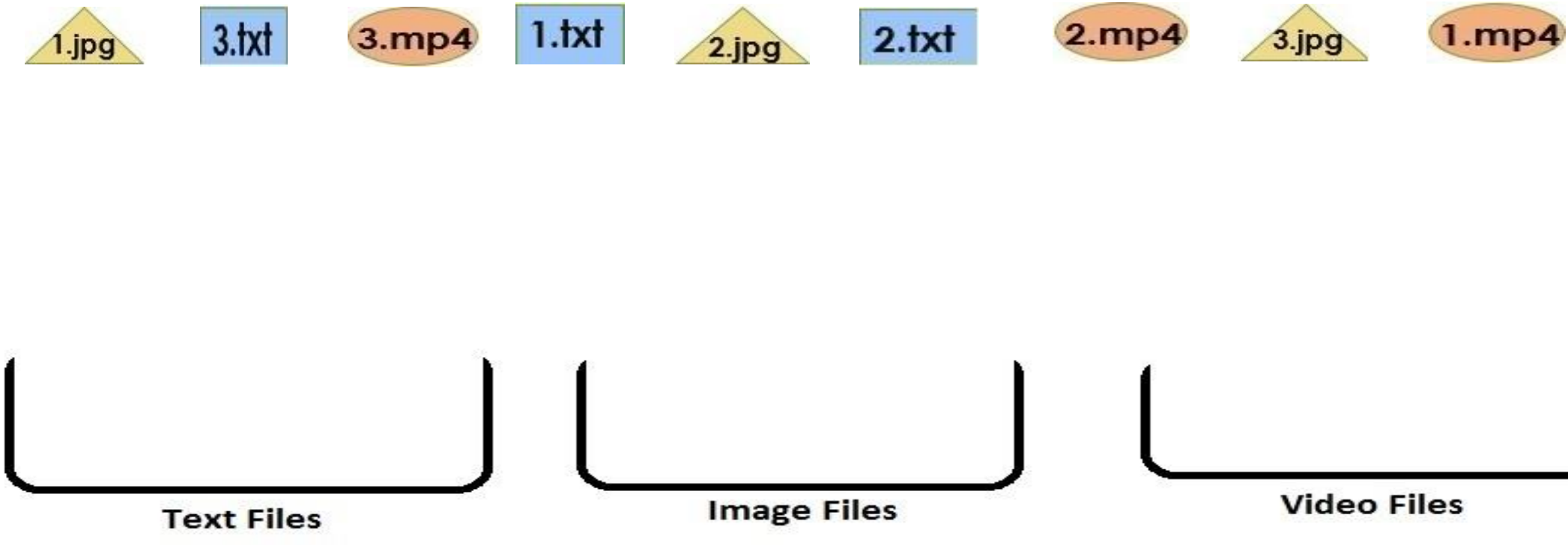


Fig :- Clustering Using File Format

Proposed Work (8)

- Malware execute its code, to work.
- Detect malware by executing its 1st line.
- Example of conventional malware signature code.

```
1. define Imain(){ //1st line of conventional malware signature
2. infect();
3. }
```

Proposed Work (9)

- 1. define Imain(){ ==========➔Converted into Symbol $1
- Store the symbol into database.

SL No.	String	Symbol
1.	:A	◄
2.	Start	↕
3.	**Imain()**	**$1**
4.	explorer	!!
5.	shutdown	ż

Proposed Work (10)

- Take Infected File code to check
- Infected File code :

1. main(){ ==========➔Converted into Symbol N1
2. int n1, n2, r; ==========➔Converted into Symbol N2
3. r=n1+n2; ==========➔Converted into Symbol N3

Proposed Work (11)

4. printf("Sum= ", r); ===➔Converted into Symbol N4
5. define Imain(){ ========➔Converted into Symbol $1

(compare with existing database)

6. infect(); ========➔Stop Converting.
7. }
8. }

Proposed Work (12)

- Malware detected file will be blocked.
- Otherwise proceed for next process i.e. behavioral detection.

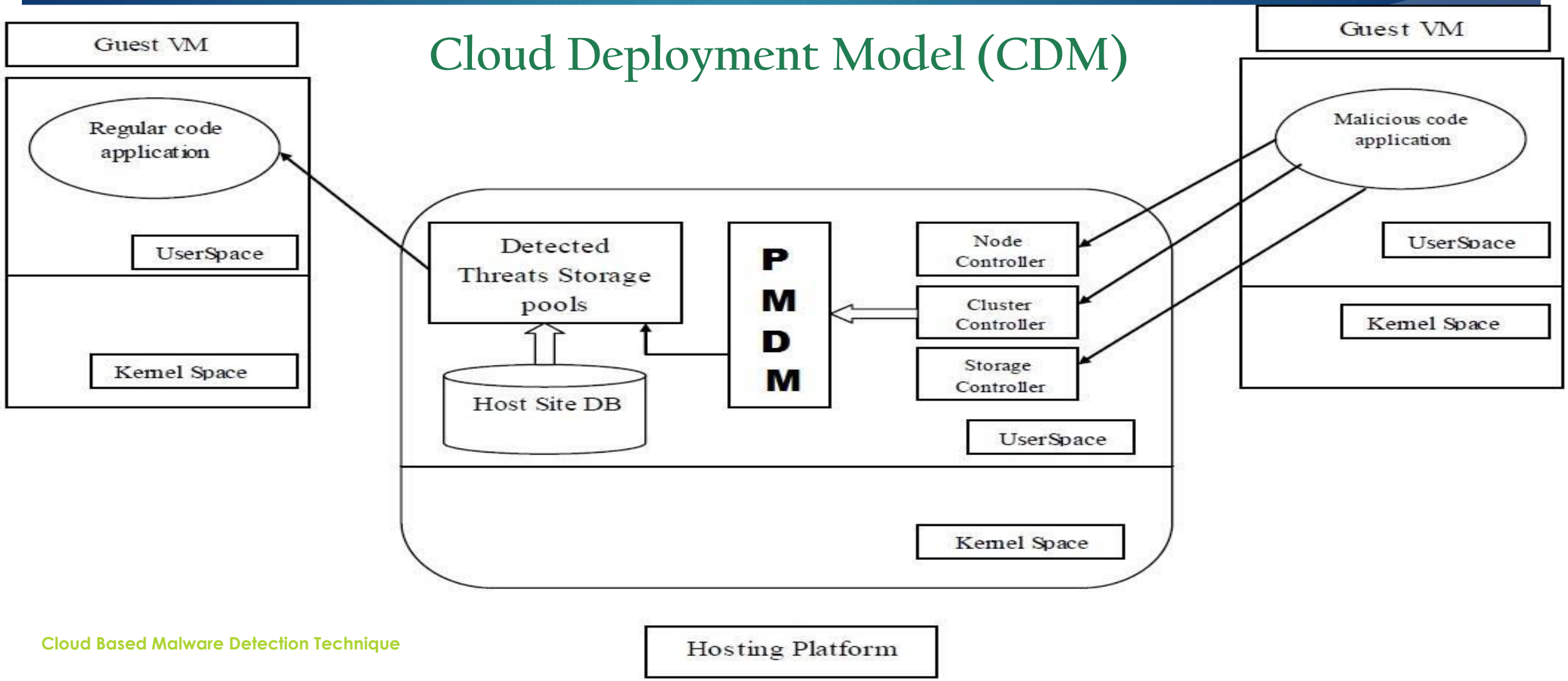
- In Process 3 the files which are pass through the second process are only go for the third process.

Proposed Work (13)

- Detecting malicious files using a virtual machine.
- Testing and running the file into a sandbox.
- We use Anubis sandbox which is free available [4].
- Combining above three process we get Proposed Malware Detection Model (PMDM)
- This Proposed Malware Detection Model (PMDM) discuss above is deploy into cloud architecture i.e. Cloud Deployment Model (CDM).

Proposed Work (14)

Cloud Deployment Model (CDM)



Result & Analysis (1)

- Result of DNA Sequence Process
  - Document Gathering

186 Text files	220 Executable Files
169 Java Files	99 Binary Files
152 Image Files	194 HTML Files

Result & Analysis (2)

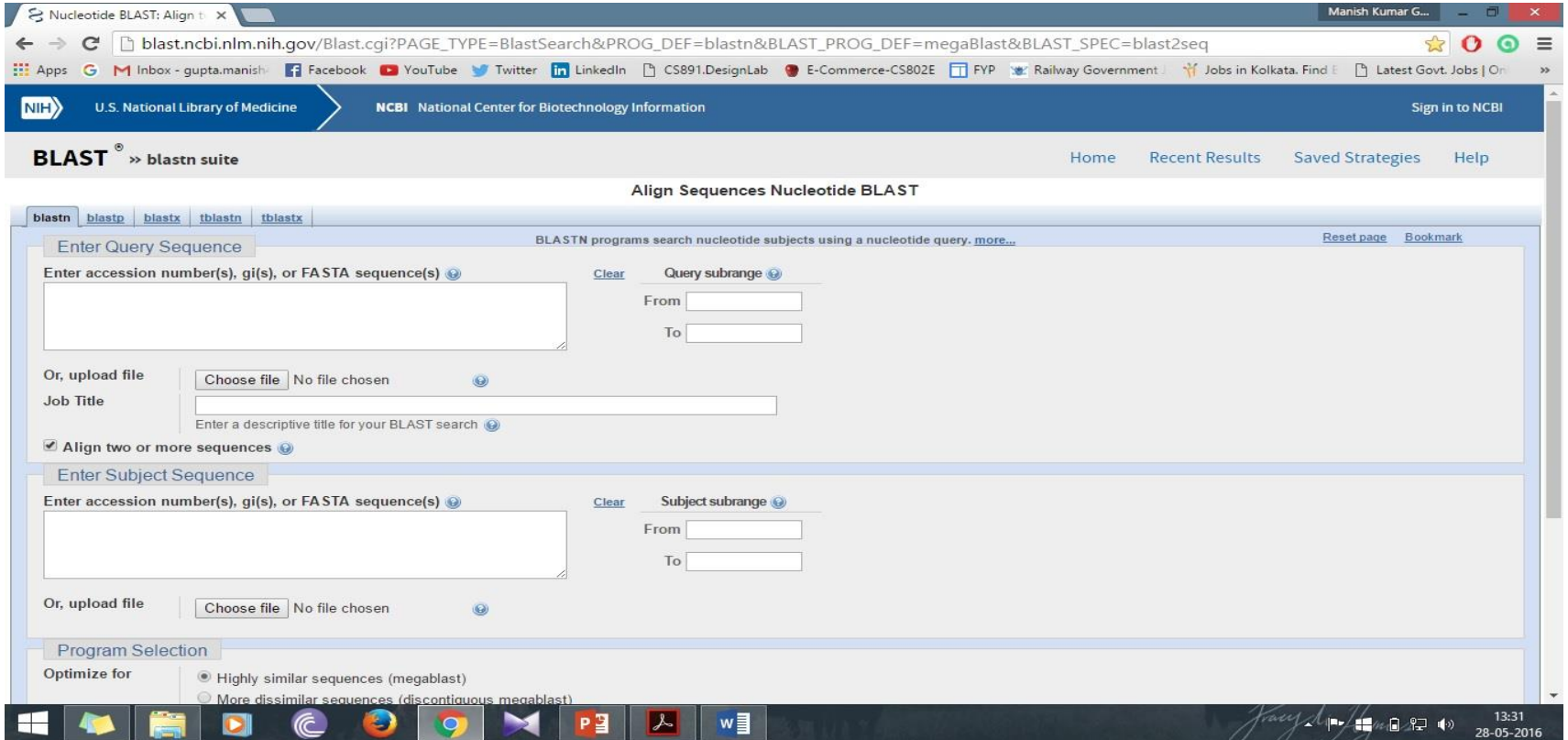
- Modify DNA Sequence

  - c: /user/name.txt

GTAGGGCCCGTTTGGCCAAAAATTTTTTTT.

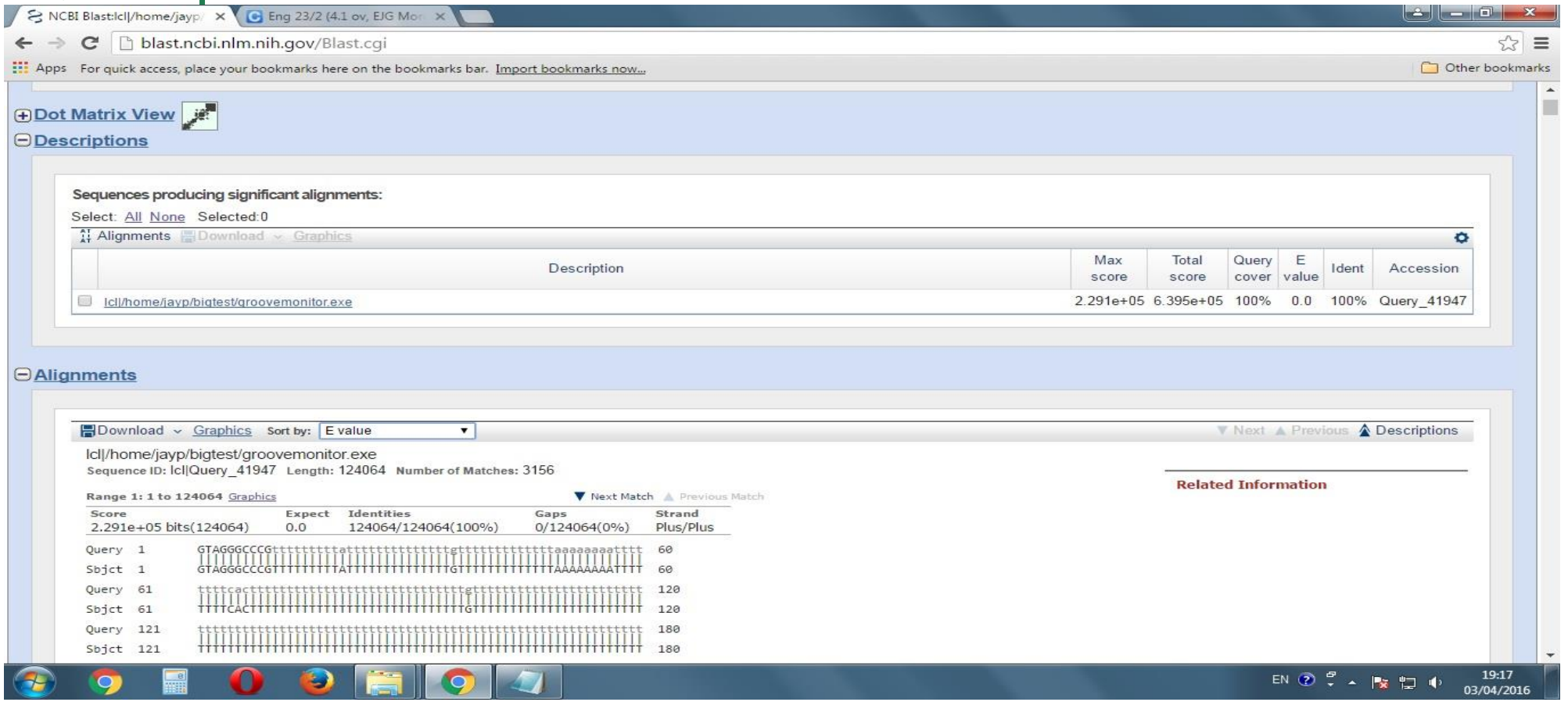
Result & Analysis (3)

- Database and software



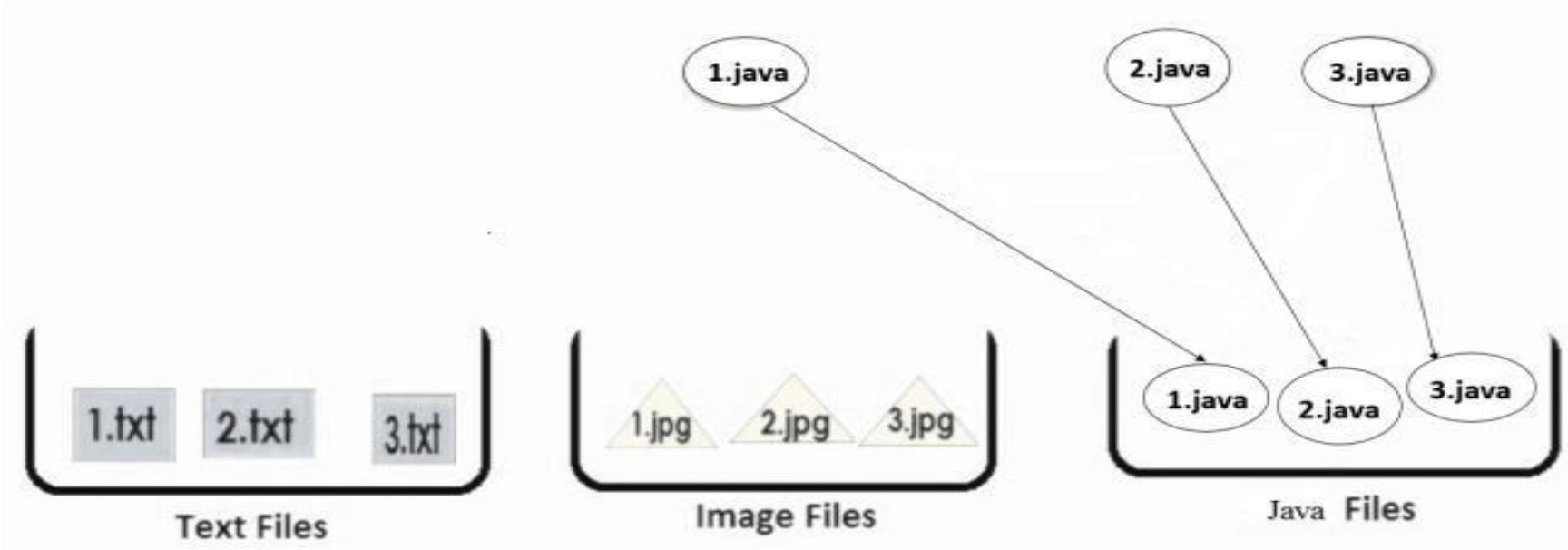
Result & Analysis (4)

- Blast Report



Result & Analysis (5)

- Result of Symbolic Detection Process
  - File Clustering.



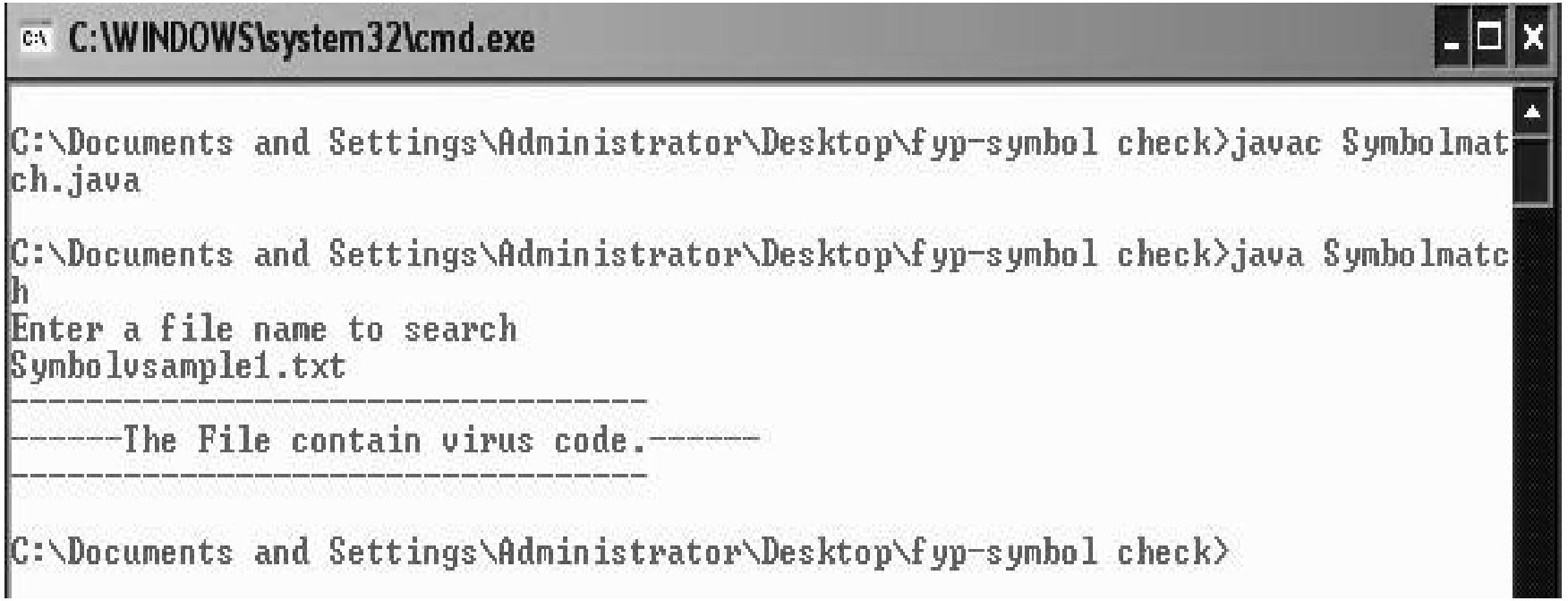
Result & Analysis (6)

- Converting File Characters into Symbols.

vsample1.txt(input file)	Symbolvsample1.txt(output file)
@echo off :A start explorer goto :A	ech◘Ö◄↕‼§◄

Result & Analysis (7)

- Matching Symbol with Symbol Table Database.



Benefits of Proposed Work

- File attribute checking, detect malicious file without open the file .
- Post infection protection is overcome by clustering we know which portion of file content we have to see exactly for malware.
- Symbolic Detection technique will increase time efficiency as we not see for whole malware signature matching.
- Cannot cope with malware variants or Zero day protection is overcome by behavioral detection.

Conclusion & Future Work

- The proposal of this work is to find the best solutions to the problems of anti-viruses and improve performance and find possible alternatives for a better working environment without problems with high efficiency and flexibility.
- Future work on this field will focus on the dependence of cloud computing. Cloud technologies have become possible because of shearing physical server resources between multiple virtual machines (VMs).

References

- [1] G. E. Dahl, J. W. Stokes et al., "Large-scale malware classification using random projections and neural networks", 2013 IEEE International Conference, pp. 3422 - 3426 , 31 May 2013.
- [2] Jay Pedersen, Dhundy Bastola, et al., "BLAST Your Way through Malware Malware Analysis Assisted by Bioinformatics Tools", International Conference on Security and Management 2012, 2011.
- [3] Safaa Salam Hatem, Dr. Maged H. wafy, et al., "Malware Detection in Cloud Computing", International Journal of Advanced Computer Science and Applications (IJACSA), vol 5, Science and Information, 2014.
- [4] Dan C. Marinescu, "Cloud Computing: Theory and Practice", MK Publication, 2013.
- [5] Mark Graham, "Behaviour of Botnets and Other Malware in Virtual Environments", The Open Web Application Security Project 2014.

Questions & Answers

THANK YOU

STAY PROTECTED