JUGAAD: Comprehensive Malware Behavior-as-a-Service

Sareena K P, Nikhilesh Singh, Chester Rebeiro, Kamakoti V.
Indian Institute of Technology Madras, India

15th Cyber Security Experimentation and Test (CSET) 2022.
Cyber Attacks – A Harsh Reality

Accenture Hit By Ransomware Attack, Latest Victim Of ‘Cyber-Pandemic’

‘If a $45 billion company like Accenture is vulnerable then everyone is vulnerable,’ says Michael Goldstein, CEO of Florida-based solution provider LAN Infotech.

By Joseph F. Kovar
September 20, 2021

A New Wave of Malware Attack Targeting Organizations in South America

Experts Uncover Spyware Attacks Against Catalan Politicians and Activists

Money & Banking

Beware of trojan malware attack, MeitY warns customers of 27 major banks

Debangana Ghosh | Mumbai | Updated on September 23, 2021
Cyber Attacks – A Harsh Reality

Despite having state-of-the-art Antivirus software!!

Accenture Hit By Ransomware Attack, Latest Victim Of ‘Cyber-Pandemic’

‘If a $45 billion company like Accenture is vulnerable then everyone is vulnerable,’ says Michael Goldstein, CEO of Florida-based solution provider L&H Infotech.

A New Wave of Malware Attack Targeting Organizations in South America

Experts Uncover Spyware Attacks Against Catalan Politicians and Activists

Money & Banking

MARIK

Beware of trojan malware attack, Mumbai: Contaminates 27 major banks

Debangana Ghosh | Mumbai | Updated on September

Malware
Confronting Cyber Attacks

Evolving and diverse threat landscape

Number of new malware reported in millions

Millions of new malware per year

Confronting Cyber Attacks

Evolving and diverse threat landscape

In-depth understanding of Ground-truth of malware behavior
Objectives, functionalities, and consequences

Run-time Behavioral Analysis

Run-time Behavioral Analysis

Access → Execute → Capture → Analyse

Malware
Access to Malware Samples

Access

1. Access to LIVE samples
   - Monopolized by Enterprises
   - High-Cost 12K samples @ 82K$

2. Older malware samples (Stale)

Malware Profiler  Command & Control Servers

LIVE
Executing Malware

Access

1. Access to LIVE samples
   Monopolized by Enterprises
   High-Cost 12K samples @ 82K$

2. Older malware samples (Stale)

Execute

3. Risks of spread; damage

4. Internet connectivity

5. Possibility of Evasion

6. Timely Execution
Capturing Behavior

**Access**

1. Access to LIVE samples
   - Monopolized by Enterprises
   - High-Cost 12K samples @ 82K$

2. Older malware samples (Stale)

![Diagram showing Access to LIVE samples and Command & Control Servers](image)

**Execute**

3. Risks of spread; damage

4. Internet connectivity

5. Possibility of Evasion

6. Timely Execution

![Diagram showing Execute process](image)

**Capture**

Different Perspectives

- Network
- Operating System
- Hardware
- Memory

![Diagram showing Capture process](image)
Access

1. Access to LIVE samples
   - Monopolized by Enterprises
   - High-Cost: 12K samples @ 82K$

2. Older malware samples (Stale)

Malware Profiler

Command & Control Servers

Execute

3. Risks of spread; damage

4. Internet connectivity

5. Possibility of Evasion

6. Timely Execution

Capture

Different Perspectives

- Network
- Operating System
- Hardware
- Memory

Analyse

Using ML/DL models
Multiple Pain Points

Access

1. Access to LIVE samples
   - Monopolized by Enterprises
   - High-Cost: 12K samples @ 82K$

2. Older malware samples (Stale)

Execute

3. Risks of spread; damage

4. Internet connectivity

5. Possibility of Evasion

6. Timely Execution

Capture

- Different Perspectives
  - Network
  - Operating System
  - Hardware
  - Memory

Analyse

- Using ML/DL models

Precise collection: Timely execution of live malware in real-world connected conditions
Malware Behavior-as-a-Service

Access
- Access to LIVE samples
- Monopolized by Enterprises
- High-Cost: 12K samples @ 82K$
- Older malware samples (Stale)

Execute
- Risks of spread; damage
- Internet connectivity
- Possibility of Evasion
- Timely Execution

Capture
- Different Perspectives
  - Network
  - Operating System
  - Hardware
  - Memory

Analyse
Using ML/DL models

Alternate Model: **Malware Behavior-as-a-Service**
JUGAAD: Malware Behavior-as-a-Service
For Malware Researchers

Users: Malware researchers from academia and industry
The Back-end for Precise Data Collection

1. Access to malware samples
   K7 SECURITY

2. LIVE malware samples
   VirusTotal
The Back-end for Precise Data Collection

1. Access to malware samples
2. LIVE malware samples
3. Containing risks
The Back-end for Precise Data Collection

1. Access to malware samples
2. LIVE malware samples
3. Containing risks
4. Internet connectivity
The Back-end for Precise Data Collection

1. Access to malware samples

2. LIVE malware samples

3. Containing risks

4. Internet connectivity

5. Evasion

Heterogeneous Real-world conditions
Desktops & Single board computers
The Back-end for Precise Data Collection

1. Access to malware samples
   - K7 Security

2. LIVE malware samples
   - VirusTotal

3. Containing risks

4. Internet connectivity

5. Evasion

6. Timely execution

Heterogeneous Real-world conditions
- Desktops & Single board computers

- Containing risks

- Timely execution
The Back-end for Precise Data Collection

1. Access to malware samples
2. LIVE malware samples
3. Containing risks
4. Internet connectivity
5. Evasion
6. Timely execution

Comprehensive view of runtime activity
Network, Operating System, Hardware

Heterogeneous Real-world conditions
Desktops & Single board computers

LIVE malware samples
Containing risks
Internet connectivity
Growing Data Corpus Till Date

2.7 TB of data, 22M behavioral snapshots
Conclusion

- Facilitates **precise, unbiased** view of diverse **perspectives** of malware activity
- **Offloads** time, efforts, and cost, while alleviating risk.
- Enables a **fair platform** for comparison of detection mechanisms
- Opens up the field for researchers in **non-security domains (data science)**
- Quickly explore and build novel solutions
- Future work: Include other **perspectives** (memory, instruction traces)
Backup Slides
The Back-end for Precise Data Collection

Timely execution

1. Download

Update Engine → Supply of Samples → Test Engine

Check for new samples → New Samples → Collect → Execute → Real-World Testbed

Dataset Corpus

Remote Power Control

Save Data
The Back-end for Precise Data Collection

1. **Timely execution**
   - Update Engine
   - Check for new samples
   - New Samples

2. **Real-world conditions**
   - Desktops & Single board computers

3. **Internet connectivity versus containment**
   - ERNET

4. **Comprehensive view of runtime activity**
   - Network, Operating System, Hardware

5. **Growing data corpus**
   - Up-to-date with evolving landscape

- **Growing data corpus**
- **Comprehensive view of runtime activity**
- **Real-world conditions**
- **Timely execution**