COMPUTER SCIENCE & ENGINEERING
TEXAS A&M UNIVERSITY

PeerPress: Utilizing Enemies' P2P Strength against Them

Zhaoyan Xu[1], Lingfeng Chen[1], Guofei Gu[1] , Christopher Kruegel[2]
[1]Texas A&M University, College Station
[2]University of California, Santa Barbara

Wed, Octo 17th, 2012

Agenda

- Introduction
- Approach Overview
- PeerPress: Port Extraction
- PeerPress: Informed enforCed Execution
- Evaluation
- Conclusion

Introduction: P2P Malware

Botnet's Evolution

- Early botnets use centralized C&C architecture

  - Centralized C&C is Fragile and Easy to be detected

- More advanced robust peer-to-peer architectures for C&C

Status

Kaspersky Security Reports:

" ... More than 2.5 million P2P malware incidents per month ... "

Examples of P2P Malware

Conficker (10,500,000+ bots), Sality (1,000,000+ bots, Waldec(80,000+ bots), Storm (1,000,000+ bots) ....

Current Research

Network-level Detection

- Perform Clustering and Correlation to identify suspicious traffic
  - Apply multiple statistics techniques.
  - Fail in front of encryption, pattern manipulation.
- Structure/Graph Analysis
  - Only P2P structure regardless of whether the traffic is malicious.
  - Requires tremendous resources, such as global ISP-level cooperation

Host-level Detection

- Signatures Matching
  - Suffer from obfuscation/polymorphism.
- Runtime Behavior Matching
  - Typically Expensive
- Both Require Client-side Installation
  - Not Scalable for Large-Scale Deployment

Agenda

- Introduction
- Approach Overview
- PeerPress: Port Extraction
- PeerPress: Informed enforCed Execution
- Evaluation
- Conclusion

Our Approach

Is it possible to combine the strength of both approaches?

- Host-level Dynamic Analysis
  - Insight: P2P malware has build-in remotely-accessible logic for peer communication/control.
  - Target: Extract the access/control conversation logic as detection evidence.
- Network-level Active, Informed Probing
  - Insight: P2P malware has to open some port
  - Target: Actively probe machine in the network to detect malware-infected machines

Our Approach

Overview of PeerPress System

Dynamic Analysis

Analyze two logic components:

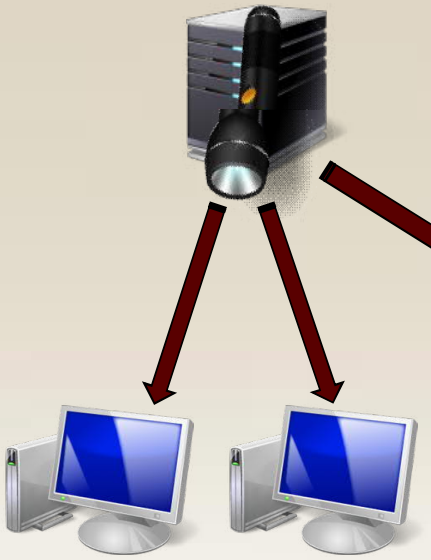
P_1: Open Service Port

P_2: Parse request and generate response

Malware Control Birthmark:

Portprint: Port that malware listens on

MCB probing: Request-and-Response Pair



MCB-assisted network probing

Advantage of Our Approach

Fast and Proactive

- Apply probing technique to make the detection as fast as network scanning.
- Able to detect malware even before the start of malicious communication/activity.

Reliable

- Probing content is extracted directly from malware binary.
- Control logic is usually unique to each malware family.

Scalable

- Easy for large-scale deployment.

Why PeerPress?

Dynamic Analysis

- Analyze Malware Peer's Logic to find MCB against themselves.

Informed Active Probing

- Scan the Peers' Machine to press them expose the malware-infected machine.

Not only Applicable to P2P Malware

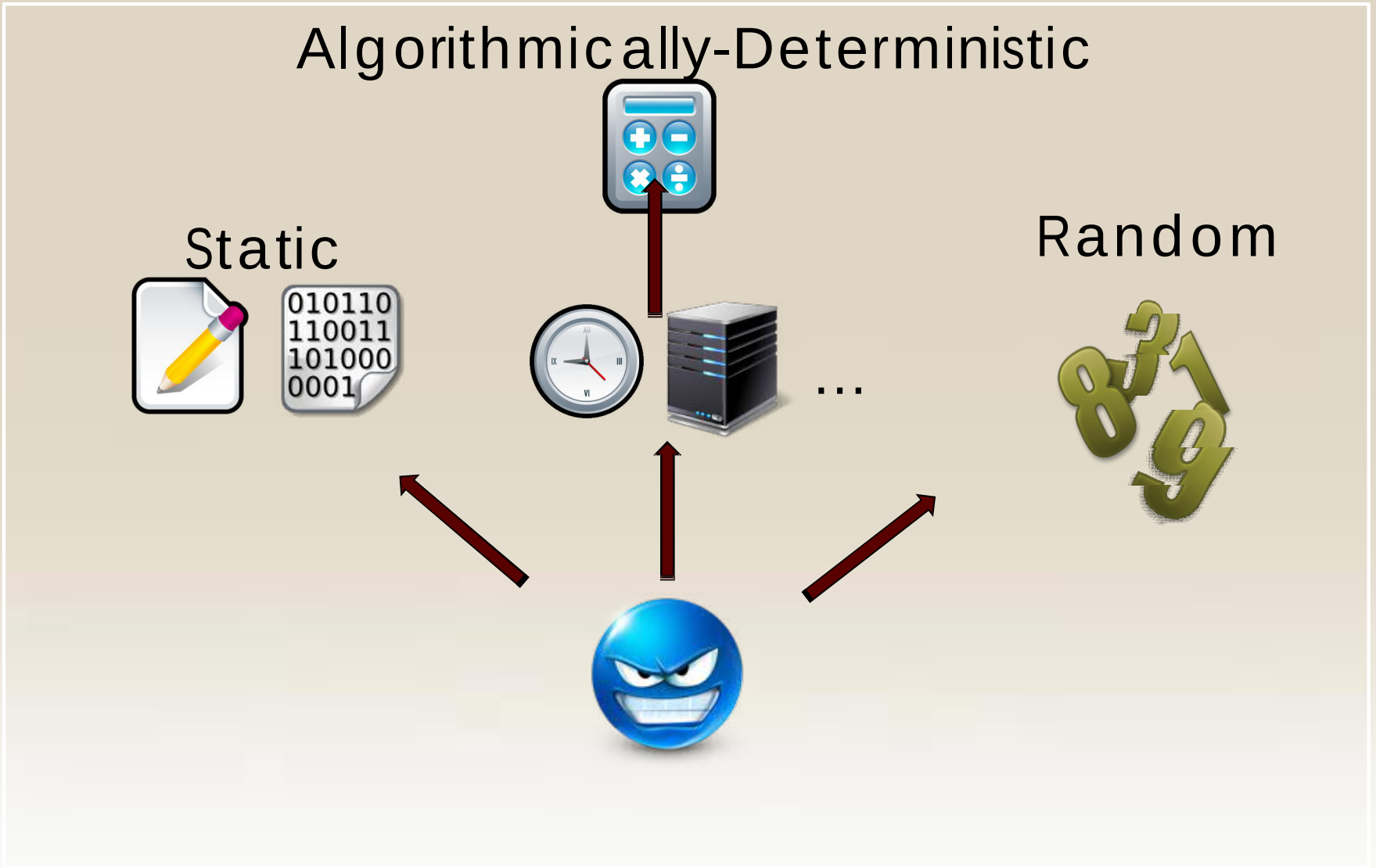
- Trojan Horses or any malware that contains Malware Control Birthmark

Agenda

- Introduction
- Approach Overview
- PeerPress: Port Extraction
- PeerPress: Informed enforCed Execution
- Evaluation
- Conclusion

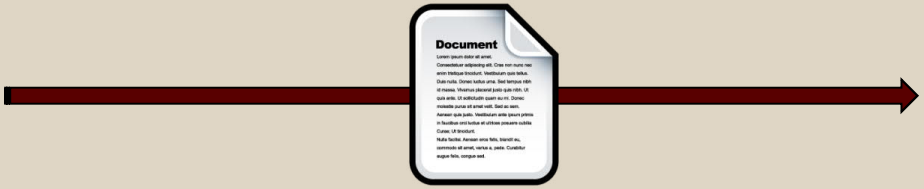
System : Portprint Extraction

Challenges: Malware binds to different port



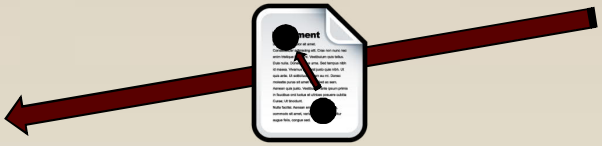
System : Portprint Extraction

Solution: Backward Taint Analysis + Program Slicing



bind()

1. Execution Trace Collection from Malware Booting to Port Bind



2. Backward Taint Analysis

Many-unknown-sources-to-one-known Sink

System/Library Calls whose

Parameter has Semantic Meaning

Constant Value

Read-only Segment

.....

3. Derive Portprint Type and Source of Data Dependence

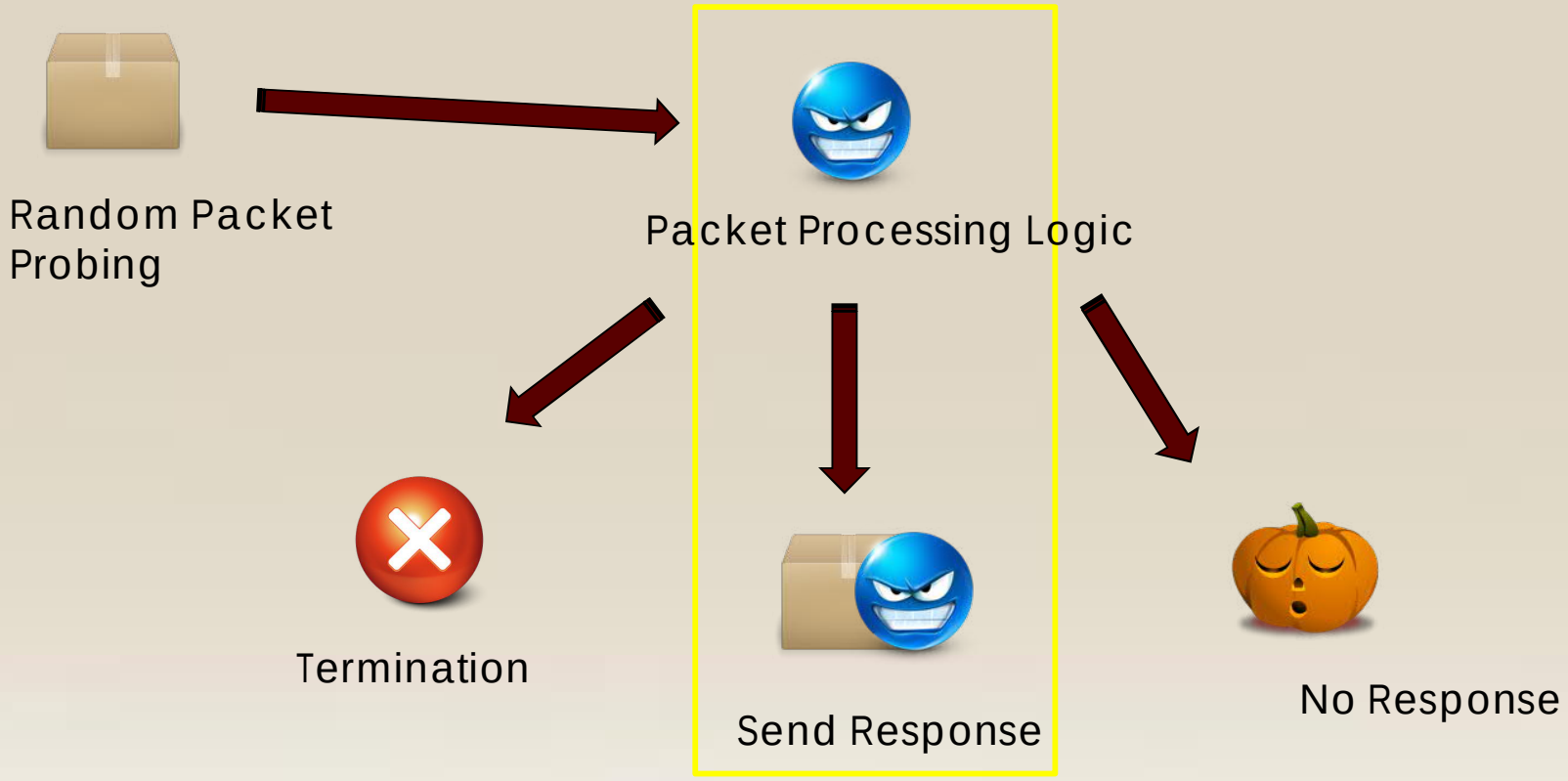
4. Program Slicing and Port Generation Logics Extraction

Agenda

- Introduction
- Approach Overview
- PeerPress: Port Extraction
- PeerPress: Informed enforCed Execution
- Evaluation
- Conclusion

System : MCB Probing Extraction

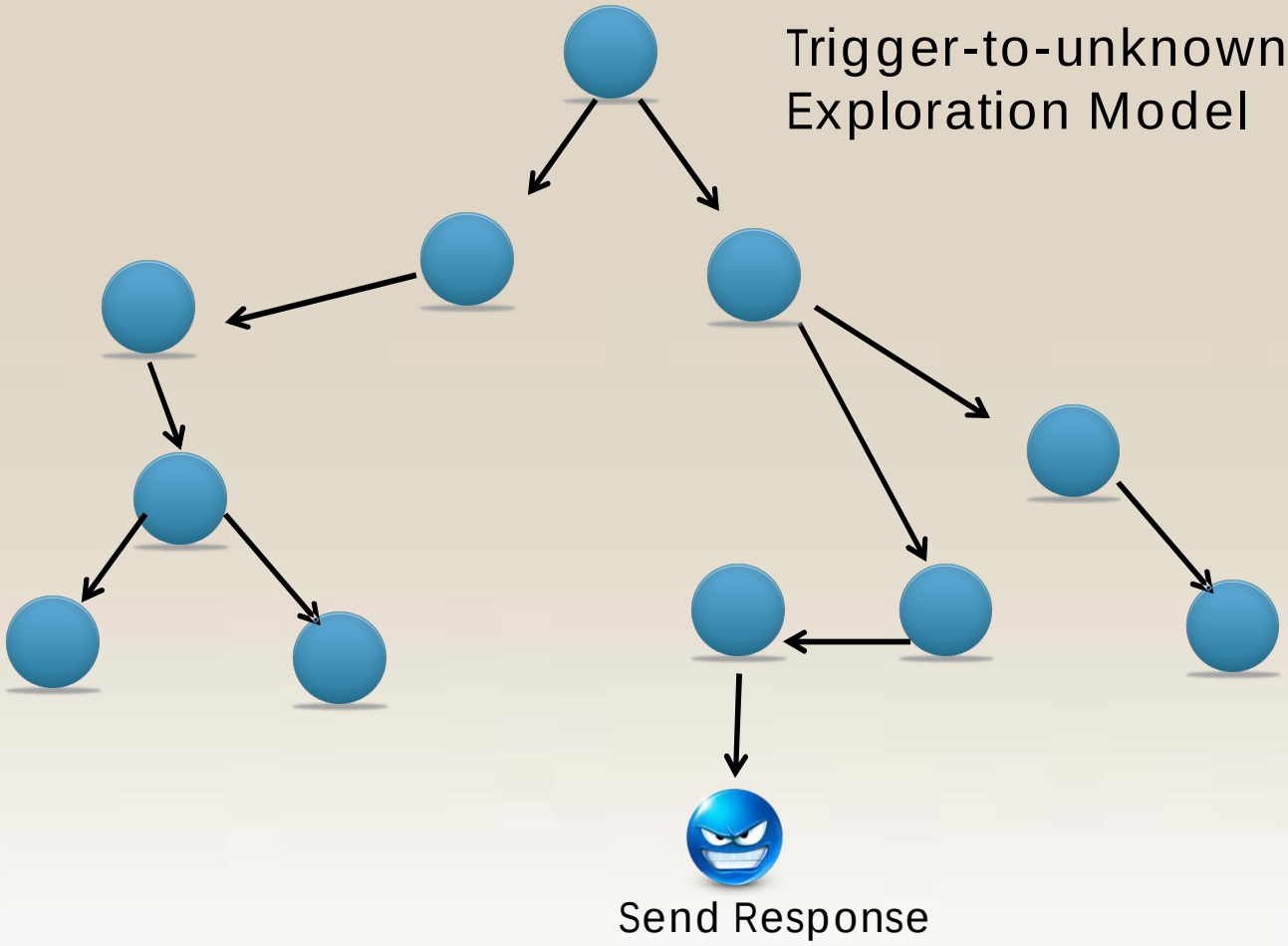
Challenges



MCB Paths: All possible execution paths from packet receiving to packet transmitting

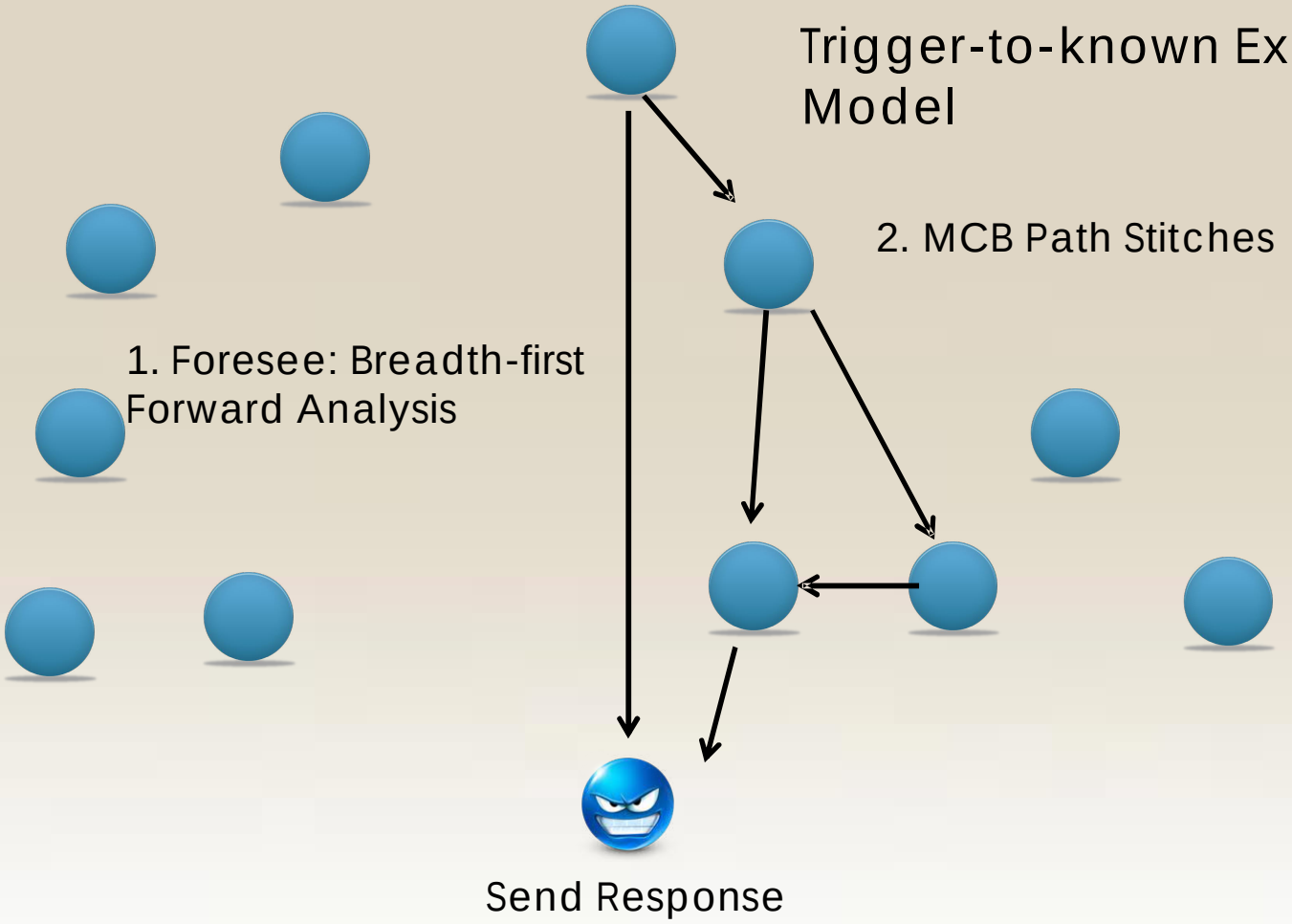
System : ICE

Traditional Multipath Exploration



System : ICE

Our Informed enforCed Execution(ICE) Scheme



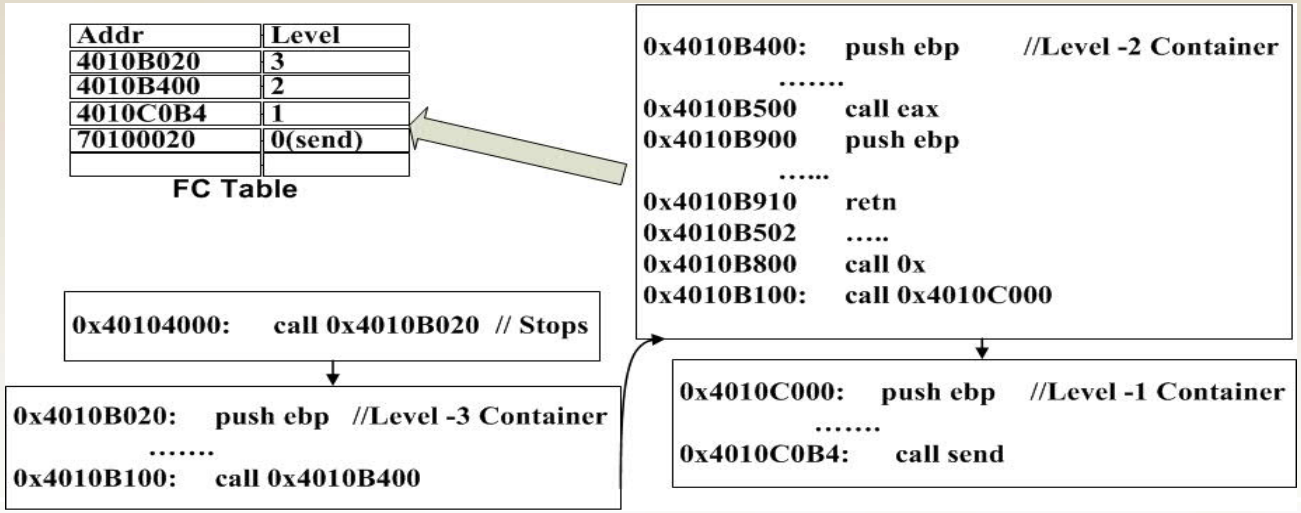
System : ICE

How to *Quickly* Find the Send out Routine

Insight: From the booting of P2P Malware, it starts sending out packets for peer communication. Such observable sending out routine may be reused in its server logic.

We define Function Container:

- Any desired or undesired sinkholing system/library calls are function containers, such as send() or closesocket()
- The function directly or indirectly contains an existing function container.



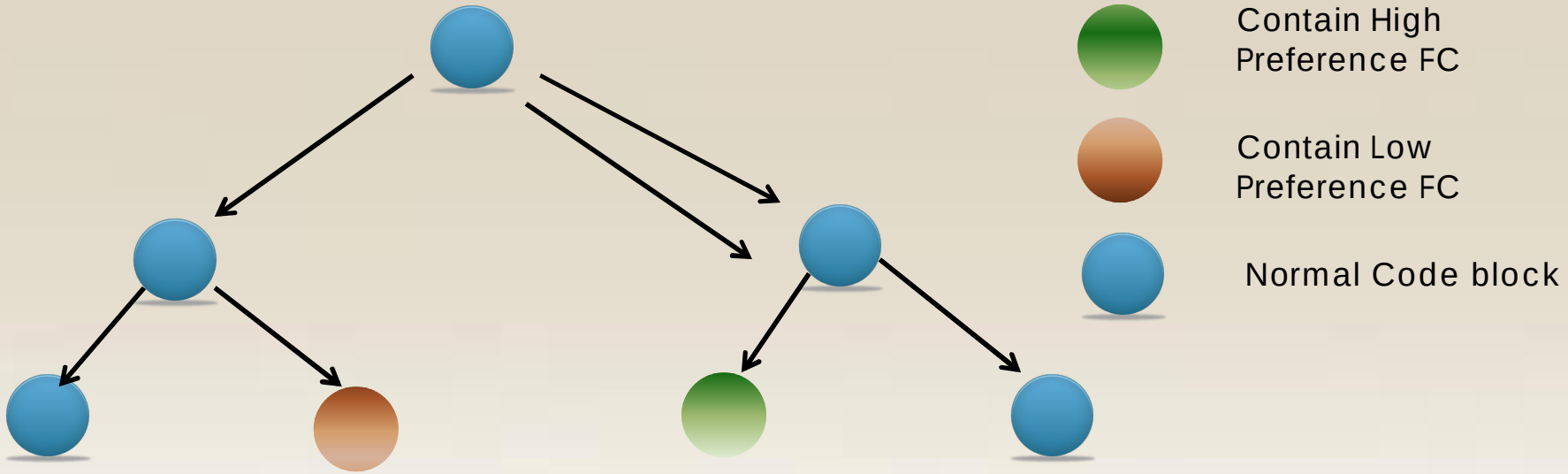
System : ICE

Path Foreseeing

Online Enforced Execution to explore MCB paths

Foreseeing, look forward, *k* code blocks

to search for the calls to any recorded function container.



System : ICE

Stitching Dynamic Symbolic Execution

- Expand all possible paths that are sensitive to tainted packets bytes (related to network packets)
- Apply combination of concrete and symbolic execution to filter out Invalid and Unreachable paths.
- Reconstruct MCB probing based on symbolic equations.

System : MCB Probing Extraction

Verifier: Filtering False Positive Cases

- First round:

  Verify whether probing packets can trigger the malware to execute the MCB path.

- Second round:

  Verify whether the reply is unique or not.

  + Probe benign software and make sure their replies are different.

Agenda

- Introduction
- Approach Overview
- PeerPress: Port Extraction
- PeerPress: Informed enforCed Execution
- Evaluation
- Conclusion

Evaluation

Real world P2P Malware and Trojan Horse Families			
Name	Type	Name	Type
Conficker	P2P Bot	Nugache	P2P Bot
Phabot	P2P Bot	Sality	P2P Bot
NuclearRAT	Trojan Horse	BackOrfice	Trojan Horse
Penumbra	Trojan Hose	Storm/Peacomm	P2P Bot
NuCrypt	Trojan Horse/Worm	Wopla	Trojan Horse
WinCrash	Trojan Horse	WinEggDrop	Spyware

Evaluation

Effectiveness of Portprint Extraction

Malware	Type	Port Number
Conficker	algorithms	46523/TCP, 18849/UDP
Nugache	static, random	8/TCP, 3722/TCP
Sality	algorithms	6162/UDP
Phabot	random	1999/TCP
Storm/Peacomm	static	7871, 11217/UDP
BackOrfice	static	31337/TCP

Evaluation

Effectiveness of ICE

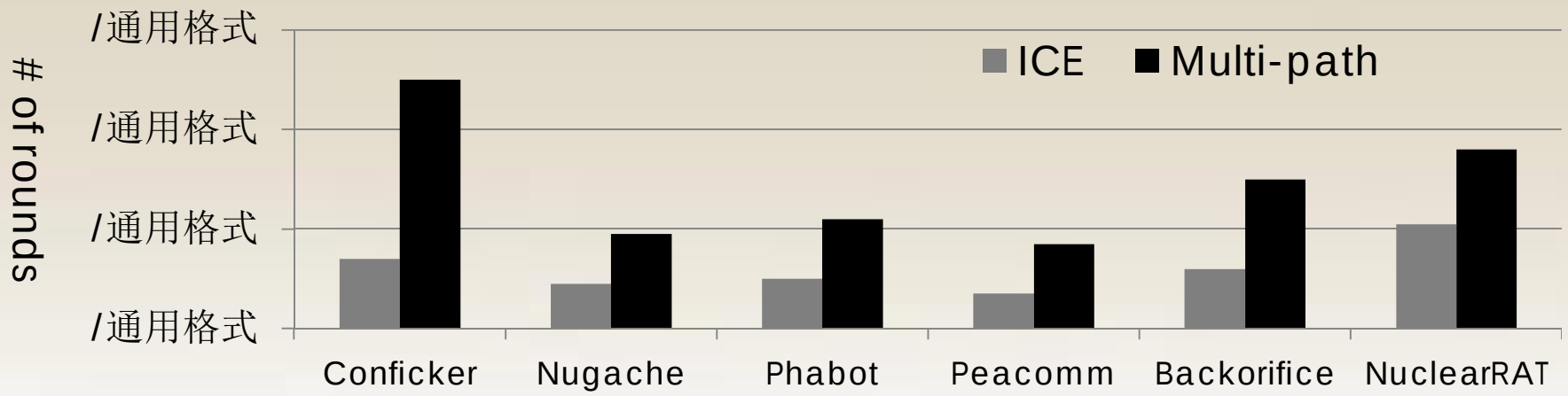
- We set the maximum call depth for function containers as 4:

  Locate average 28 function containers per malware sample

- Overhead:

  Compare ICE with Multipath Explorations

  Measure the number of rounds to generate one MCB path



Evaluation

Outcome of MCB

Malware	# of MCB	Malware	# of MCB
Conficker	3	Peacomm	3
Sality	1	BackOrifice	14
Phabot	9	NuclearRAT	12
WinEggDrop	8	Penumbra	13
Nugache	7	WinCrash	1
NuCrypt	2	Wopla	2

Evaluation

Detection Results through Active Probing

In Virtual Networks

- Install samples for each family on our virtual environment
- Install well-known benign server software, such as Apache, eMule.

Detection Results:

- PeerPress correctly detects all the exsiting malware
- Average 1.103 seconds to detect each malware

Evaluation

False Positive Test

In Real Networks of our Campus

- Scan 3 /24 networks using extracted MCBs
- Scan common ports for HTTP, P2P, FTP services

Results:

- No false positives

Evaluation

Comparison with State-of-the-art Detection System

Deploy State-of-the-art network based system, BotHunter, in the virtual network

Results:

No malware detected.

Discussion:

Reasonable result, because Bothunter needs collecting enough network traffic for evidence.

PeerPress is more proactive.

Agenda

- Introduction
- Approach Overview
- PeerPress: Port Extraction
- PeerPress: Informed enforCed Execution
- Evaluation
- Conclusion

Conclusion

- We propose a novel two-phase detection framework for P2P Malware.
- PeerPress combines the merits of both dynamic binary analysis and network-level informed active probing.
- We develop techniques such as ICE to improve the analysis performance.

Q&A