Impact Analysis of Botnet Infection on Networked Systems using Timed Automata

SYSC 5500
Group 8

Alvi Jawad
Luke Newton

Outline

Preliminaries
- Botnets in the IoT Network
- Timed Automata & UPPAAL

Modeling
- Overview of the Mirai Botnet
- Modeling The Botnet Infrastructure
- Modeling devices in the network

Simulations
- The Concrete Simulator

Results
- Data Collection
- Device Infection Rate
- Network Traffic Generation
- Next steps!

What are Botnets?

Botnets

- A network of compromised devices
- Compromised devices are called bots
- Bots infect other vulnerable hosts in the network
- Infection rarely requires user interaction

The Internet of Things

- Vulnerable networked infrastructure
- IoT devices
  - have low computational resources
  - are poorly secured
  - have poor maintenance
- IoT devices can generate huge amounts of attack traffic in a botnet

Why use Timed Automata?

Timed Automata

- Finite State Machines extended with Clock variables
- A timed automaton is a 6-tuple (L, l_o, C, A, E, I)
  - L is a set of locations
  - $l_o \in L$ is the initial location
  - C is the set of clocks
  - A is a set of actions, co-actions, and the internal τ-action
  - $E \subseteq L \times A \times B(C) \times 2^C \times L$ is a set of edges
  - $I : L \rightarrow B(C)$ assigns invariants to locations

PROS

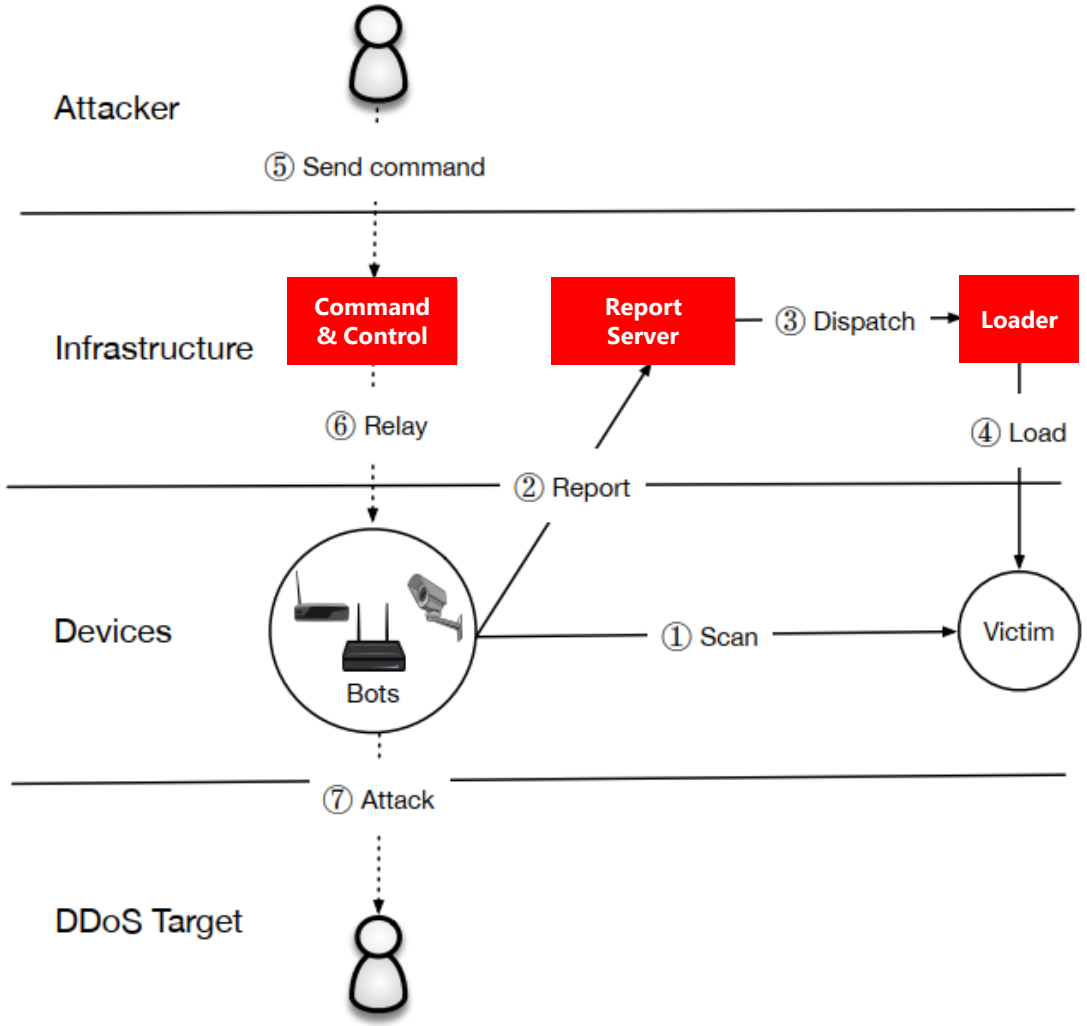
- Timed representation
- Hybrid-view of the system
- State Detection
- Behavior Prediction
- Formal verification

UPPAAL

- Verification of real-time systems modeled as networks of Timed Automata
- Query language: TCTL (Timed Computation Tree Logic)

Overview of the Mirai Botnet

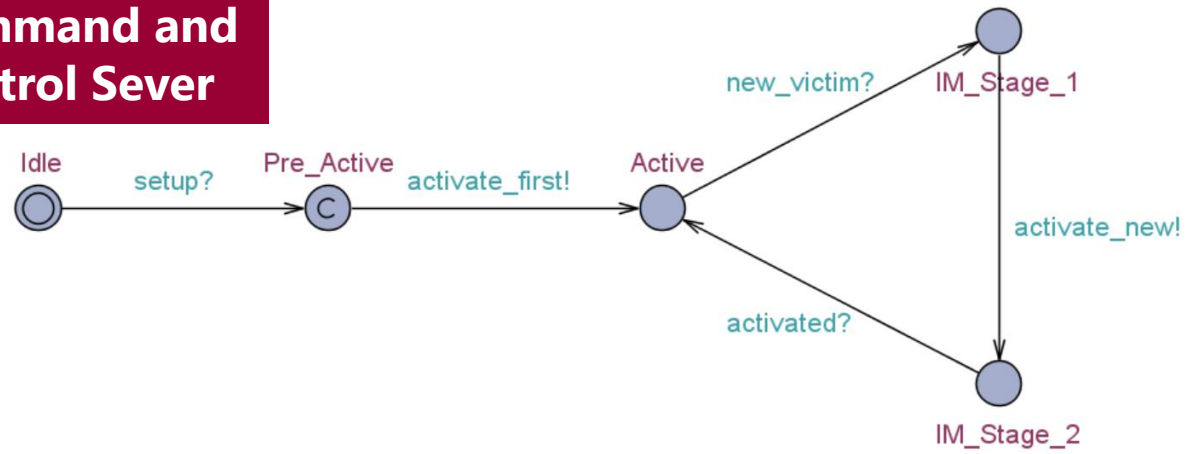
1. Scan
2. Report
3. Dispatch
4. Load
5. Send command
6. Relay
7. Attack



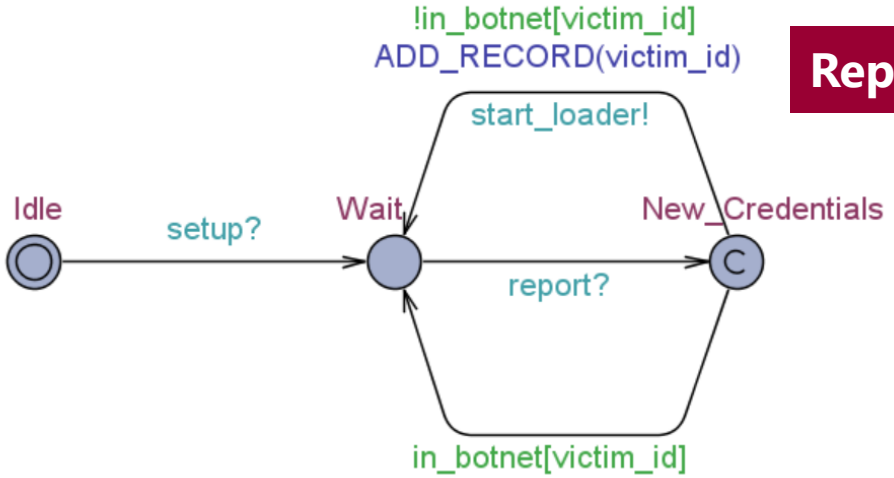
M. Antonakakis, T. April, M. Bailey, M. Bernhard, E. Bursztein, J. Cochran, Z. Durumeric, J. A. Halderman, L. Invernizzi, M. Kallitsis, D. Kumar, C. Lever, Z. Ma, J. Mason, D. Menscher, C. Seaman, N. Sullivan, K. Thomas, and Y. Zhou, "Understanding the mirai botnet," in *26th USENIX Security Symposium (USENIX Security 17)*, (Vancouver, BC), pp. 1093–1110, USENIX Association, Aug. 2017.

Modeling the Botnet Infrastructure

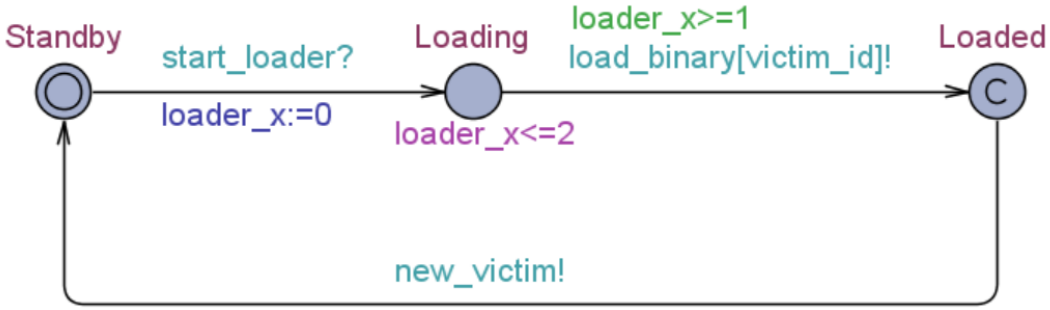
Command and Control Sever



Report Server

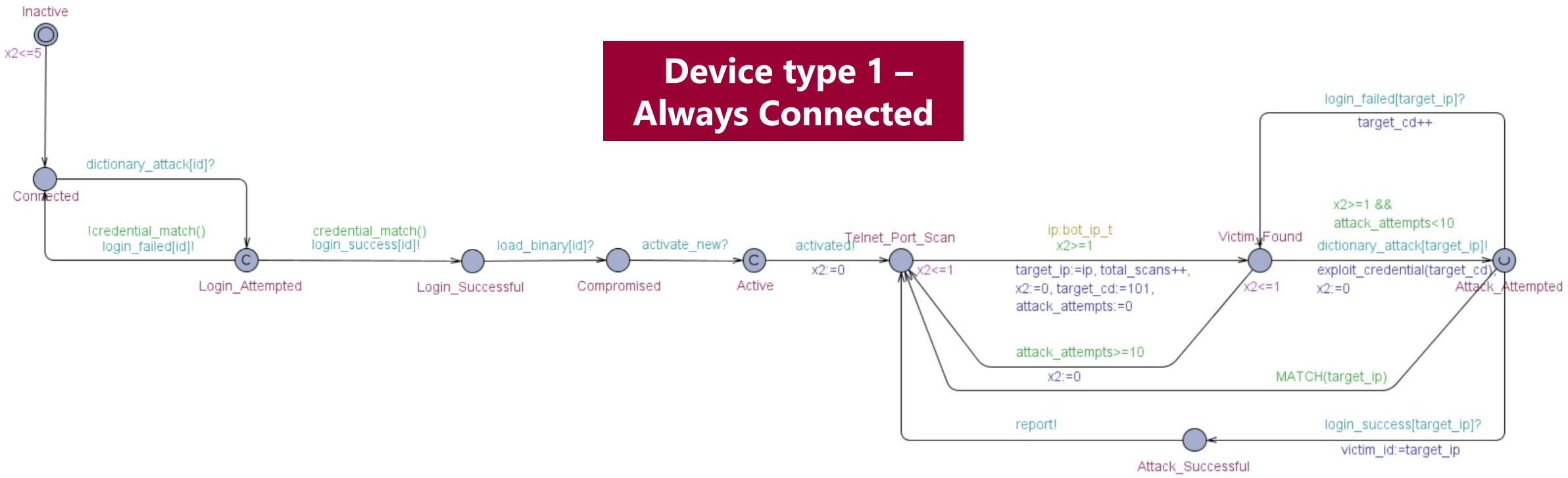


Loader



Modeling the Devices

Activation



Device type 1 – Always Connected

Scan and Report

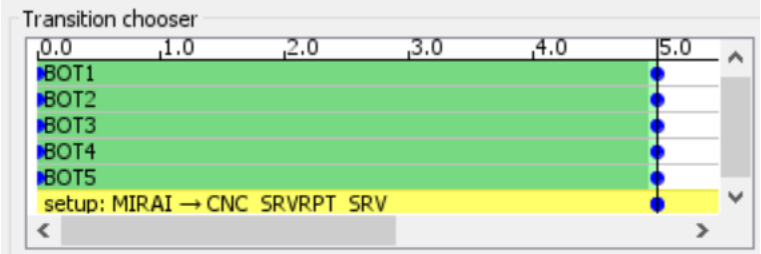
G. Kambourakis, C. Kolias and A. Stavrou, "The Mirai botnet and the IoT Zombie Armies," *MILCOM 2017 - 2017 IEEE Military Communications Conference (MILCOM)*, Baltimore, MD, 2017, pp. 267-272, doi: 10.1109/MILCOM.2017.8170867.

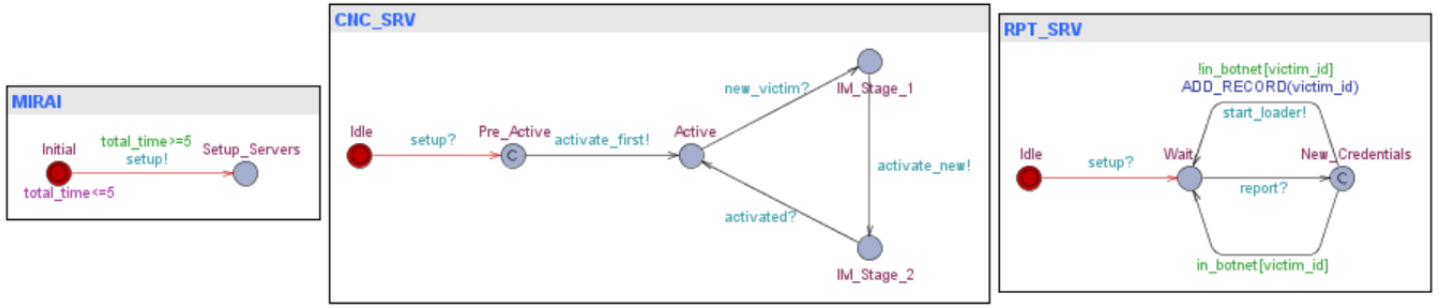
The Concrete Simulator – Initial State

State Variables

- <Global variables>
  - total_scans = 0
  - total_attempts = 0
  - current_number_bots = 0
  - victim_id = 1
  - in_botnet = {0,0,0,0,0,0}
    - [0] = 0
    - [1] = 0
    - [2] = 0
    - [3] = 0
    - [4] = 0
    - [5] = 0
  - credential_attempt = {101,101,101,101,101,101}
    - [0] = 101
    - [1] = 101
    - [2] = 101
    - [3] = 101
    - [4] = 101
    - [5] = 101
  - t(0) = 0
  - total_time = 0.000000
- BOT_DF
  - attack_attempts = 0
  - target_ip = 1
  - target_cd = 101
  - x1 = 0.000000
- BOT1
  - attack_attempts = 0
  - target_ip = 1
  - target_cd = 101
  - x2 = 0.000000
- BOT2
  - attack_attempts = 0
  - target_ip = 1
  - target_cd = 101
  - x2 = 0.000000
- BOT3
  - attack_attempts = 0
  - target_ip = 1
  - target_cd = 101
  - x2 = 0.000000
- BOT4
  - attack_attempts = 0
  - target_ip = 1
  - target_cd = 101
  - x2 = 0.000000
- BOT5
  - attack_attempts = 0
  - target_ip = 1
  - target_cd = 101
  - x2 = 0.000000

Choosing Transitions





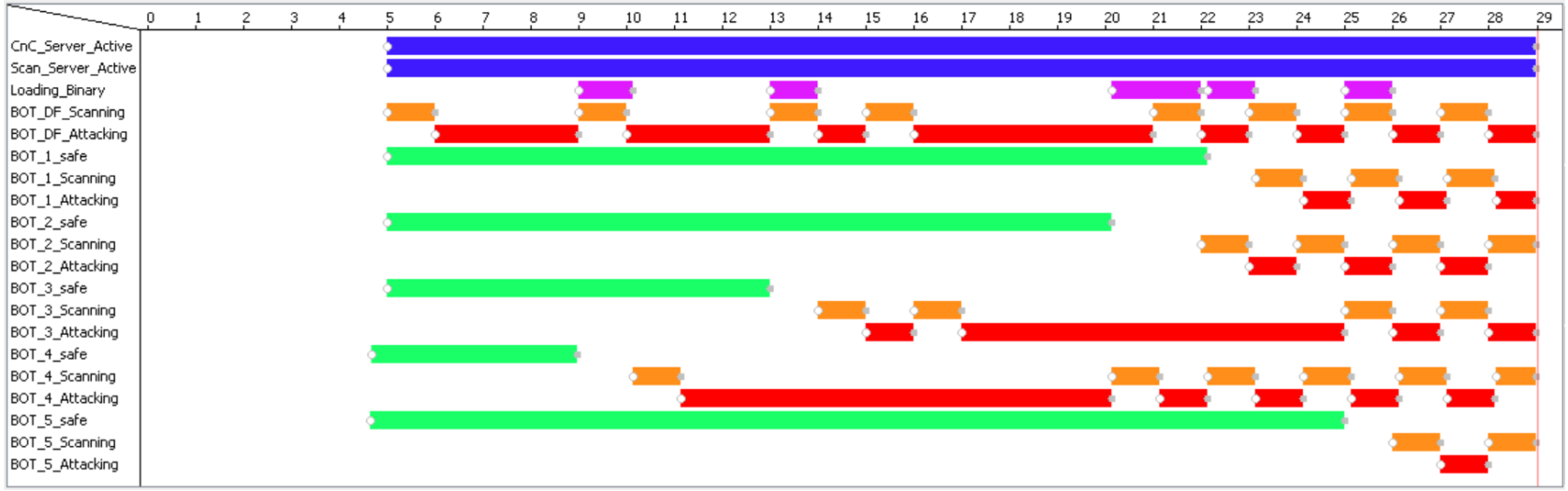
The Concrete Simulator – in Action

State at 18 minutes

- BOT_DF
  - attack_attempts = 1
  - target_ip = 2
  - target_cd = 102
  - x1 = 1.000000
- BOT1
  - attack_attempts = 0
  - target_ip = 1
  - target_cd = 101
  - x2 = 18.000000
- BOT2
  - attack_attempts = 0
  - target_ip = 1
  - target_cd = 101
  - x2 = 18.000000
- BOT3
  - attack_attempts = 1
  - target_ip = 5
  - target_cd = 101
  - x2 = 0.000000
- BOT4
  - attack_attempts = 6
  - target_ip = 2
  - target_cd = 107
  - x2 = 0.863843
- BOT5
  - attack_attempts = 0
  - target_ip = 1
  - target_cd = 101
  - x2 = 18.000000

- <Global variables>
  - total_scans = 7
  - total_attempts = 16
  - current_number_bots = 2
  - victim_id = 3
  - in_botnet = {1,0,0,1,1,0}
    - [0] = 1
    - [1] = 0
    - [2] = 0
    - [3] = 1
    - [4] = 1
    - [5] = 0
  - credential_attempt = {101,101,106,101,101,101}
    - [0] = 101
    - [1] = 101
    - [2] = 106
    - [3] = 101
    - [4] = 101
    - [5] = 101
  - t(0) = 0
  - total_time = 18.000000

Gantt Chart



- <Global variables>
  - total_scans = 26
  - total_attempts = 46
  - current_number_bots = 5
  - victim_id = 5
  - in_botnet = {1,1,1,1,1,1}
    - [0] = 1
    - [1] = 1
    - [2] = 1
    - [3] = 1
    - [4] = 1
    - [5] = 1
  - credential_attempt = {101,101,101,101,101,101}
    - [0] = 101
    - [1] = 101
    - [2] = 101
    - [3] = 101
    - [4] = 101
    - [5] = 101
  - t(0) = 0
  - total_time = 29.000000

State at 29 minutes

- BOT_DF
  - attack_attempts = 1
  - target_ip = 3
  - target_cd = 101
  - x1 = 0.000000
- BOT1
  - attack_attempts = 0
  - target_ip = 4
  - target_cd = 101
  - x2 = 0.863843
- BOT2
  - attack_attempts = 0
  - target_ip = 5
  - target_cd = 101
  - x2 = 0.000000
- BOT3
  - attack_attempts = 1
  - target_ip = 1
  - target_cd = 101
  - x2 = 0.000000
- BOT4
  - attack_attempts = 1
  - target_ip = 3
  - target_cd = 101
  - x2 = 0.863843
- BOT5
  - attack_attempts = 0
  - target_ip = 1
  - target_cd = 101
  - x2 = 0.000000

How do we collect data?

Concrete Simulator

- Run individual simulations and log data yourself
- See each run in great detail
- Slow, manual process

Verifier

- Formally check properties with temporal logic
- Exhaustive state space exploration
- Beware of state space explosion!

- Automate several simulation runs
- Graph frequencies and probability distributions
- Specify required confidence levels
- Export data for further analysis

Editor | Simulator | ConcreteSimulator | Verifier

Overview

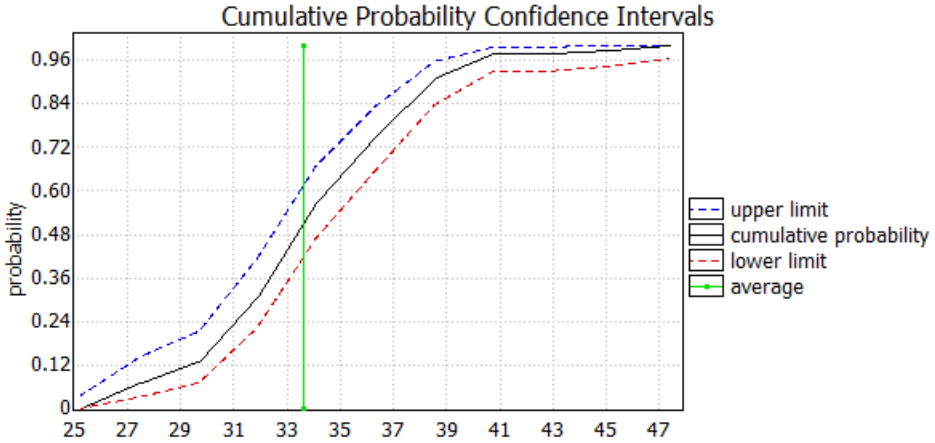
```
<----- MODEL VERIFICATION ----->
A[] !deadlock
A[] deadlock imply current_number_bots==total_devices-1

<----- INFECTION REACHABILITY ----->
A<> current_number_bots==total_devices-1
E<> current_number_bots==total_devices-1

<----- INFECTION TIMES ----->
A[] total_time==50 imply current_number_bots==total_devices-1
Pr[total_time<=100; 100] (<>current_number_bots==total_devices-1)
Pr[total_time<=50] (<>current_number_bots==total_devices-1)
Pr[total_time<=25] (<>current_number_bots==total_devices-1)
simulate [total_time<=100; 100] {current_number_bots}
simulate [total_time<=100; 100] {in_botnet[1], in_botnet[2], in_botnet[3], in_botnet[4], in_botnet[5]}
E [total_time<=25; 100] (max: current_number_bots)
E [total_time<=50; 100] (max: current_number_bots)
E [total_time<=100; 100] (max: current_number_bots)
Pr[<=50] (<>in_botnet[1])
```

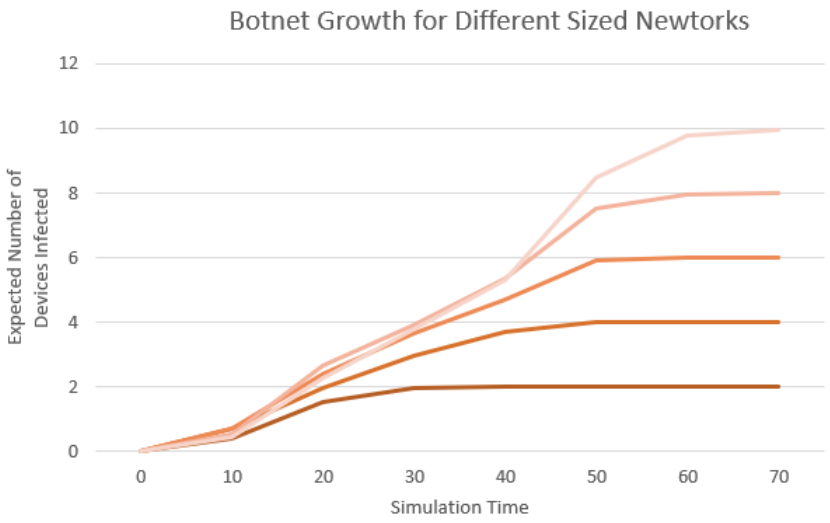
For now, we mainly produce results for a network of 5 devices

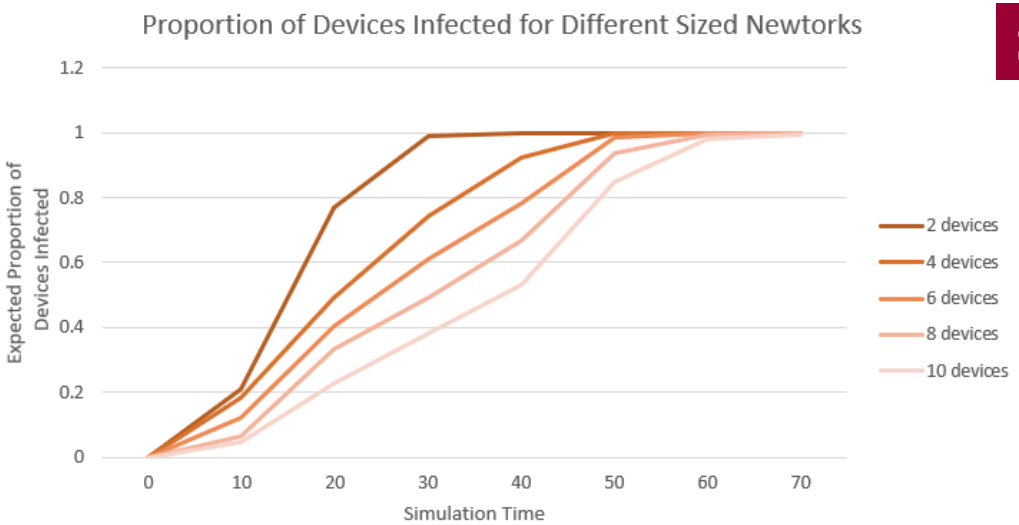
How quickly are devices infected?

Expected devices infected

- After 10 time units: 0.59 ± 0.0981
- After 25 time units: 2.93 ± 0.185
- After 50 time units: 4.99 ± 0.0198
- After 100 time units: 5 ± 0







Summary

- The model accurately depicts expected botnet growth
- Network size does not have a huge impact on expected time to infect all devices (so far)

Is network propagation "random"?

Devices are randomly targeted

- The order devices are infected should not matter
- Infection time distributions should be similar for each device

Kolmogorov-Smirnov Test

- Determine if two empirical distributions are sampled from the same distribution
- With confidence level 0.05 and 100 samples, threshold value for the test is 0.192
- Pairwise comparisons

What does this mean?

- **Location on the botnet password list significantly affects infection time**

Devices with different passwords

Device	2	3	4	5
1	0.37	0.23	0.23	0.29
2		0.29	0.32	0.17
3			0.10	0.18
4				0.28

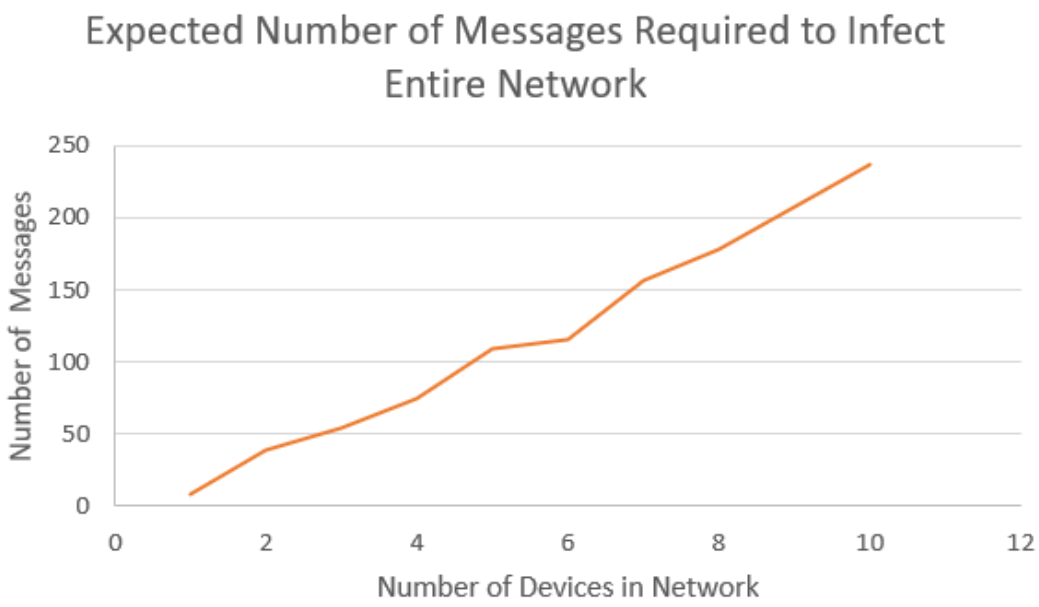
All devices with same password:

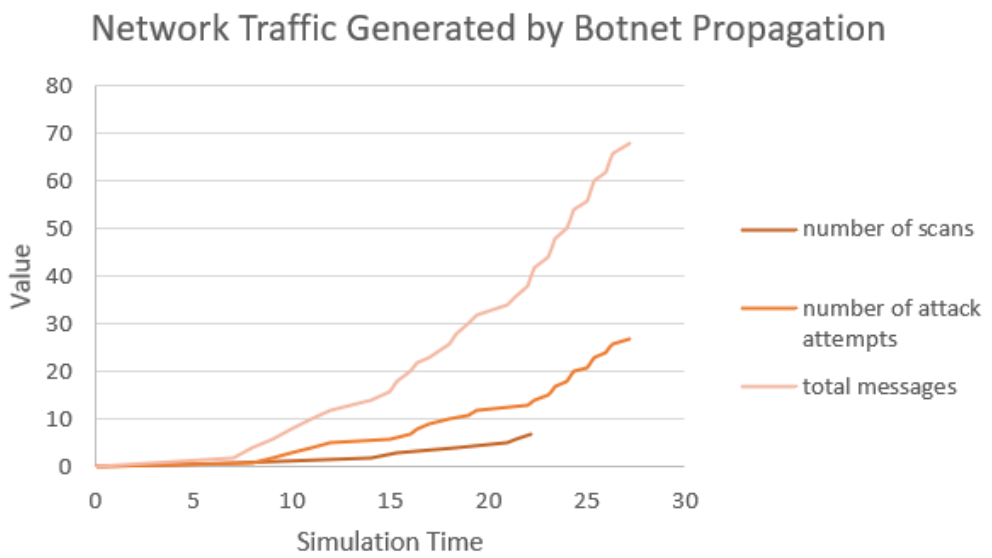
Device	2	3	4	5
1	0.06	0.11	0.09	0.12
2		0.11	0.09	0.11
3			0.07	0.08
4				0.11

How much network traffic is generated?

- Messages sent by a botnet use resources that would otherwise be used for legitimate network traffic

- Expected number of port scans: 16 ± 12.5
- Expected number of password attempts: 38.5 ± 15.4
- Expected total messages: 109 ± 54.4





Summary

- Total traffic generated by the botnet increases exponentially over time

- Required messages to infect whole network increases linearly with network size

What comes next?

Rebooting and Patching

- Rebooting a device clears any infection
- Consider device types that reboot periodically and following some distribution
- Model reinfectivity of devices after patching and subsequent reboots

Network Stability

- Adjust timing on transitions

Password Scheme

- Widen the range of possible credentials a device can have

New Questions

- With rebooting and patching, is it possible to infect all devices?
- With rebooting, does the botnet size reach a steady state?
- How many times will a device be infected before the whole network is infected?
- How do these new considerations affect the amount of network traffic produced?

Thank You!

Questions?