CERIAS

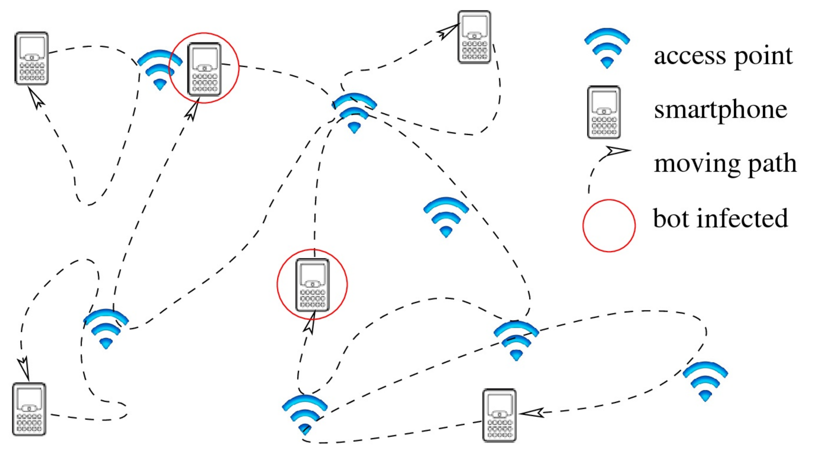
The Center for Education and Research in Information Assurance and Security

PURDUE UNIVERSITY

# $T$-dominance: A Stealthy Propagation Strategy for Mobile Botnet

Wei Peng, Feng Li, Jie Wu, and Xukai Zou
{pengw,fengli}@iupui.edu, jiewu@temple.edu, xkzou@cs.iupui.edu

## Forewarned is forearmed




Botnet, the chronic disease inflicting the Internet, is *going mobile* in the age of smartphones.

Forewarned is forearmed. We play the devil's advocate in this work by exploring techniques that could be used by mobile botnet to circumvent defense. Our work is motivated by the following observations.

1. Temporal/spatial proximity communication channels such as Wi-Fi and Bluetooth could be exploited by botnets for propagation.
2. Stealthiness is going to be a top priority for future botnets in light of the "remote kill switch" mechanism recently introduced by Google on the Android platform [1].
3. The rich information stored on smartphones could be abused by the (botnet) malware to derive users' mobility/social patterns, on which the stealthiness concept could be built.

## $T$-dominance and the distributed algorithm

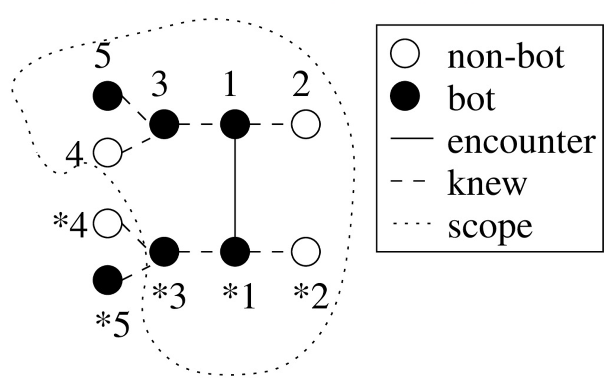
### § Reachability

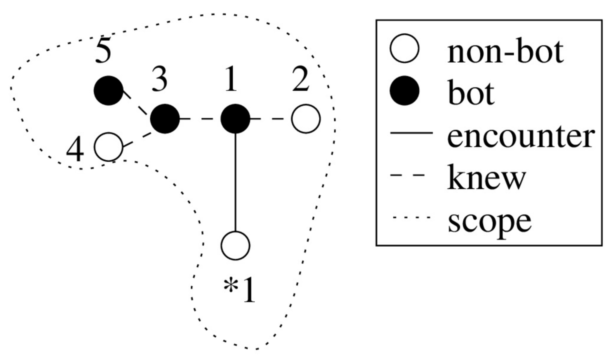
From the connectivity logs [2] of a pair of encountered smartphones, we can estimate their *reachability*, i.e., the average interval between consecutive encounters. Given a set of smartphones $\mathcal{P}$, let $\mathcal{G}(\mathcal{P})$ be the undirected weighted graph with $\mathcal{P}$ as vertices and reachability as the weight on the edges; $\mathcal{G}^T(\mathcal{P})$ be the subgraph of $\mathcal{G}(\mathcal{P})$ with weight-greater-than-$T$ edges removed.

**Definition 1** ($T$-dominance). *Let $\mathcal{P}$ be a set of smartphones and $\mathcal{B}$ be the set of bots in $\mathcal{P}$. $\mathcal{B}$ is* a botnet which $T$-dominates $\mathcal{P}$ at time $t$ *if 1) $\mathcal{B}$ is connected in $\mathcal{G}^T(\mathcal{P})$; 2) for any $p \in G^T(\mathcal{P})$, either $p \in \mathcal{B}$ or $p$ is a neighbor of a bot $b \in \mathcal{B}$ in $\mathcal{G}^T(\mathcal{P})$.*

### § Intelligence Exchange

Two kinds of encounters.




Nodes exchange intelligence at encounters. Two alternative approaches (*raw* and *processed*) are proposed.

### § Prune and Infect

Inspired by the Connected Dominating Set problem [3], we propose the *prune-and-infect* distributed algorithm for maintaining the $T$-dominance structural property.

**Prune** When a bot $u$ meets another bot $v$, $u$ decides whether to disinfect (*prune*) itself (for stealthiness). We propose two alternative prune algorithms (*individual* and *strong*) based on two alternative priority-comparison criteria (*strong* and *weak*).

**Infect** When a bot $u$ meets a clean node $v$, $u$ decides whether to infect $v$. The insight is that *$v$ should be infected unless it is likely to be pruned later*. To decide the likelihood of $v$ being pruned later, we check two criteria (*prunable* and *coverage*) consecutively.

### § Algorithm Properties

The prune-and-infect algorithm is *localized* and *delay-tolerant* in the following sense: if bot $b$ prunes itself in its local (and potentially outdated) view at time $t$, then, in the global (and updated) view, each of the smartphones $T$-dominated by $b$, including $b$ itself, is still $T$-dominated by some bot at $t$ in the global view.

Due to the speculative nature of the reachability metric, the $T$-dominance structural property provides no hard guarantee that a non-bot *will* be reached by $T$ after the attack time even if it is $T$-dominated by the botnet. However, our experiment shows that $T$-dominance provides a fairly good guarantee for reaching a majority in the smartphones pool.

## Contributions

1. We propose the concept of botnet-level stealthiness and a novel structural property, *$T$-dominance*, for a stealthy botnet. Instead of infecting all susceptible smartphones, a stealthy botnet malware with the $T$-dominance property only infects those smartphones which can reach each other smartphones within a time constraint of $T$ with a high probability.
2. We design a distributed algorithm which maintains the $T$-dominance structural property and prove that the algorithm is *localized* and *delay-tolerant* in the sense that the algorithm maintains the structural property despite relying solely on local and potentially outdated information.

## Results

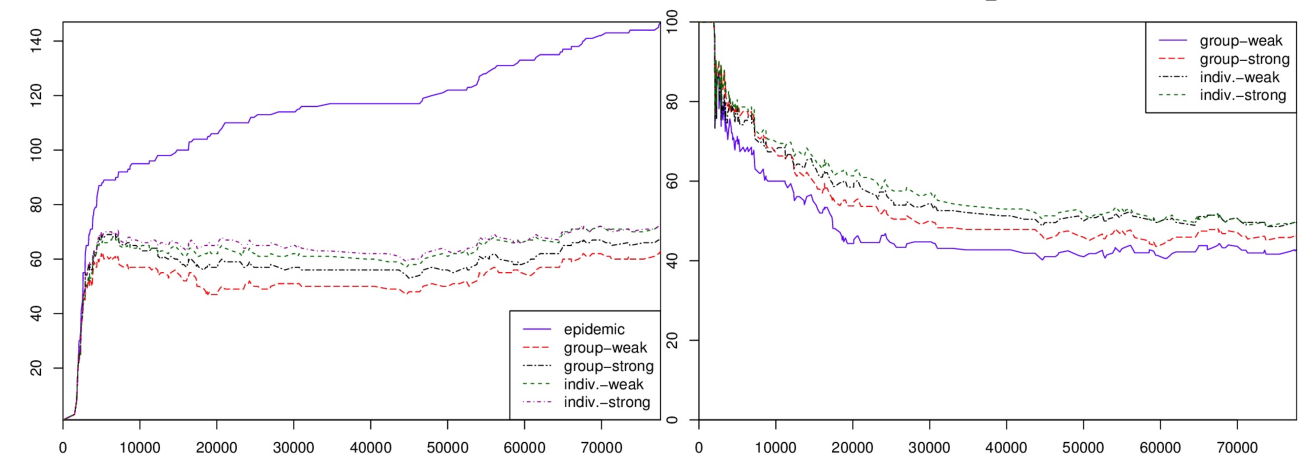
We use the dataset from the Wireless Topology Discovery (WTD) project [a] in our simulation.

Botnet's lifetime consists of two consecutive phases, *herding* and *attack*, with different goals. The goal of the herding phase is stealthiness as characterized by $T$-dominance; the goal of the attack phase is epidemic manifested in wide infection within $T$.

Different prune strategies under $T = 18,000$ (5 hrs).

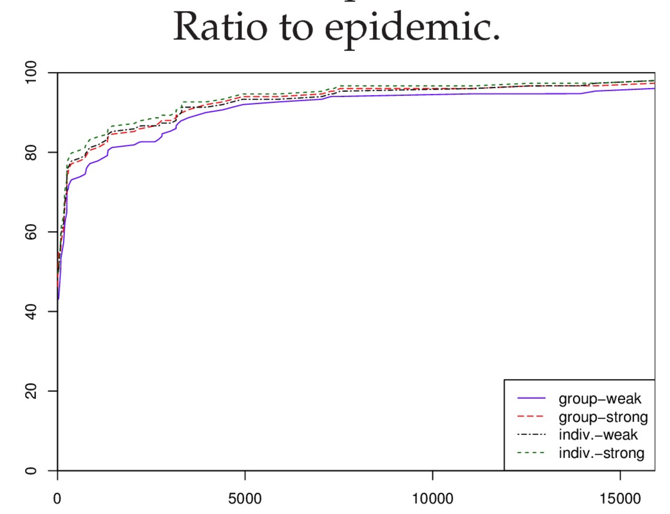
Herding phase.

Bot number. Ratio to epidemic.




Attack phase.

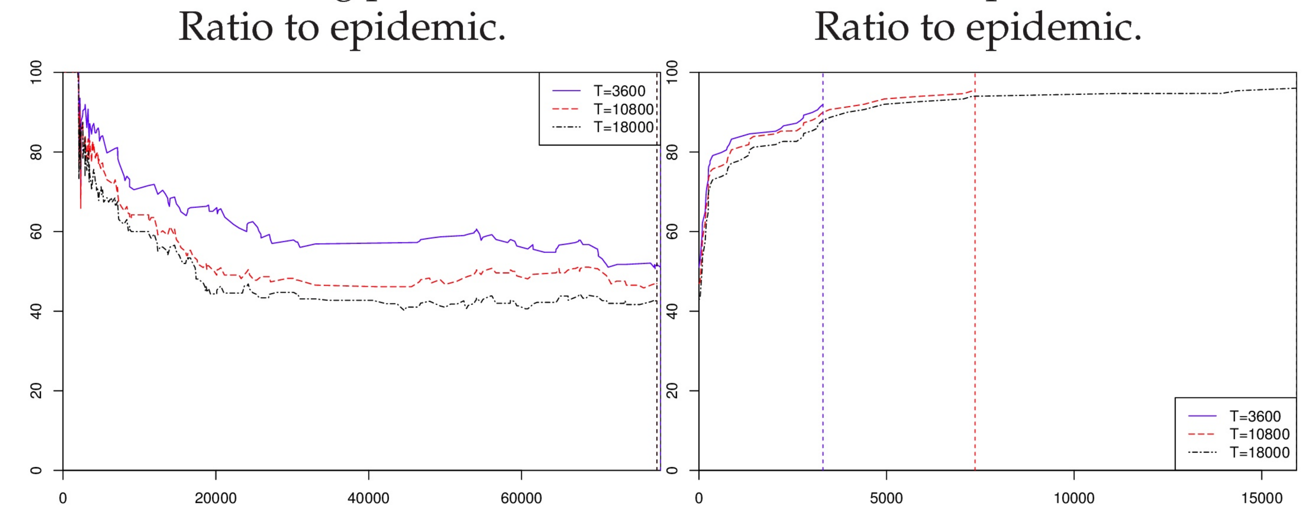
Ratio to epidemic.




The group-weak prune strategy under different $T$.

Herding phase. Attack phase.

Ratio to epidemic. Ratio to epidemic.




[a] http://sysnet.ucsd.edu/wtd/data_download/wtd_data_release.tgz

## References

[1] S. Hollister. Google flips android kill switch, destroys a batch of malicious apps. [Online]

[2] "Location histories for location aware devices," U.S. Patent 20 110 051 665, 2011.

[3] J. Wu, F. Dai, and S. Yang, "Iterative local solutions for connected dominating set in ad hoc wireless networks," *IEEE Transactions on Computers*, vol. 57, pp. 702–715, 2008.

## Thoughts on Defenses

**Cooperative defense.** Proximity malware propagation circumvents centralized defense; individual nodes lack the resource to defend against malware. A cooperative defense mechanism which autonomously coordinates nodes on the task of malware defense is the way out.

**Strategic sampling.** A *socially-aware* malware sampling based on the $T$-dominance property will choose a socially well-connected group for malware detection.

**Prioritized patching.** The $T$-dominance propagation can be applied in distributing malware patches. Instead of stealthiness, the balance between resource consumption and patch distribution efficiency are major concerns.

CERIAS

Discovery Park e-Enterprise Center