Security (Models): Prospects, Perspectives, Directions

By
Evangelos Kranakis
School of Computer Science
Carleton University

A Fundamental Question

- *Privacy:*
  Preventing the unauthorized extraction of information from communications over an insecure channel.
- Can two people that have never met before create and share a secret key?

1975: A Defining Moment

A private conversation between two people with no prior acquaintance is a common occurrence in business, however, and it is unrealistic to expect initial business contacts to be postponed long enough for keys to be transmitted by some physical means.

W. Diffie and M. E. Hellman
New Directions in Cryptography, 1976
IEEE Information Theory Workshop, Lenox MA, 1975

How about Security?

- What is security?
- What is the fundamental question?
- Why is it such a complex subject?

What is the Fundamental Question?

- There isn't any fundamental question!
- Better yet, there are too many questions!
- Why?

What is Security

IT	Physical	Political	Other	...
Computer	Airport	State	Banking	...
Data	Home	International	Financial	...
Information	Food	National	Passports	...
Network	Power Plant	Military	Health	...
⋮	⋮	⋮	⋮	⋮

A Complex Dynamic System

Security studies often require deep analysis using a complex dynamic system whose behavior cannot be described with just a few parameters!

Security Models

- A security model is a theoretical construct that represents *a situation*, with a set of variables and a set of logical and quantitative relationships between them in order to facilitate the study of security.
- Limitations of Models
  - You curb the number of parameters to solve the problem, and your solution is likely irrelevant!
  - You enlarge the pool to include most relevant parameters, and a solution becomes infeasible!

Every Model has a Story to Tell

- BGP Routing
- ($\star$) Zero Day Worms
- Device Authentication
- Active Worm Containment
- ($\star$) Barrier Coverage in Sensor Networks
- ($\star$) Best Effort
- ($\star$) Conclusion

A bit philosophical...before we start!

...Philosophy is written in the grand book–I mean the universe–which stands continuously open to our gaze, but it cannot be understood unless one first learns to comprehend the language and intercept the characters in which it is written, it is written in the language of Mathematics, and its characters are...

Gallileo Gallilei, *Il Saggiatore*, 1623

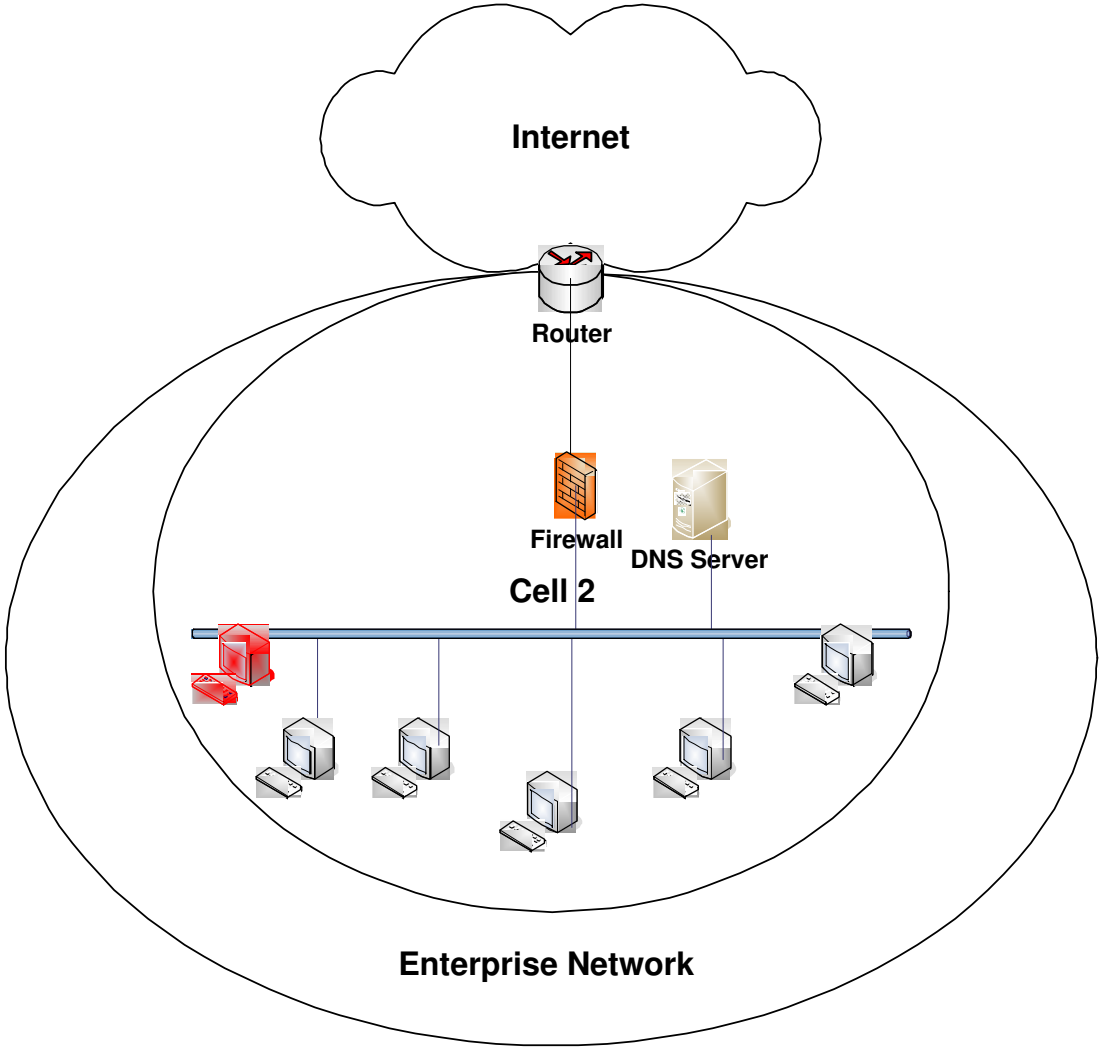
Models for Zero Day Worms

(with D. Whyte, P. Van Oorschot, NDSS 2005)

The Problem

- A *Zero day* worm is a previously-unknown computer malware for which antivirus software signatures are not yet available.
- Scanning worm propagation can occur extremely fast: Slammer infected 90% of vulnerable Internet hosts in less than 10 mins.
- Automated countermeasures are required for worm containment and suppression
- Worm propagation detection methods are usually limited by
  - Speed of detection
  - Inability to detect zero-day worms
  - Inability to detect slow scanning worms
  - High false positive rate

Affecting Enterprise Networks



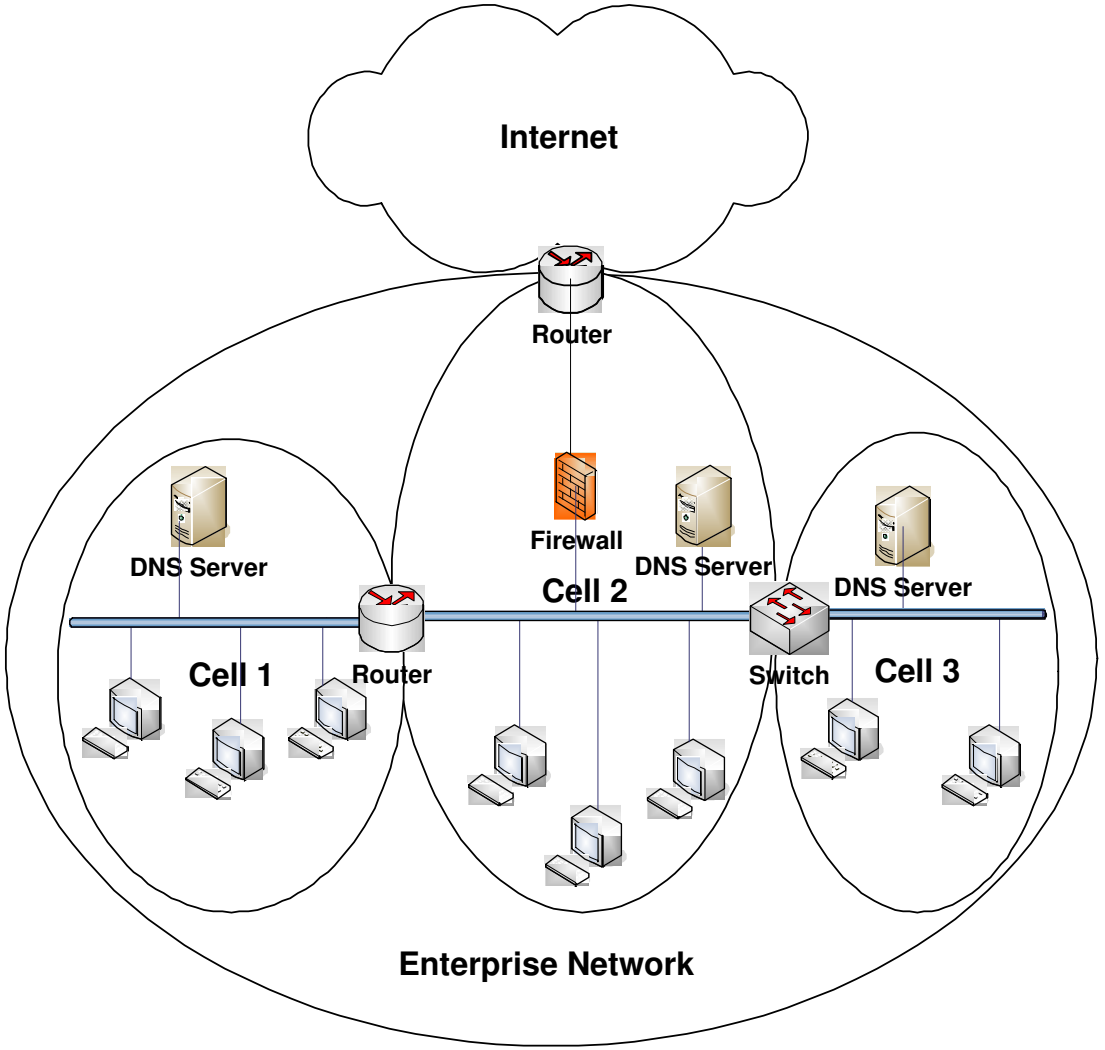
How to discover Worms

- Identifying large flows in real time with small amounts of memory
- Counting the number of sources and destinations
- Determining scanning by counting the number of connections attempts to unused portions of the IP address

Scanning Worm Characteristics

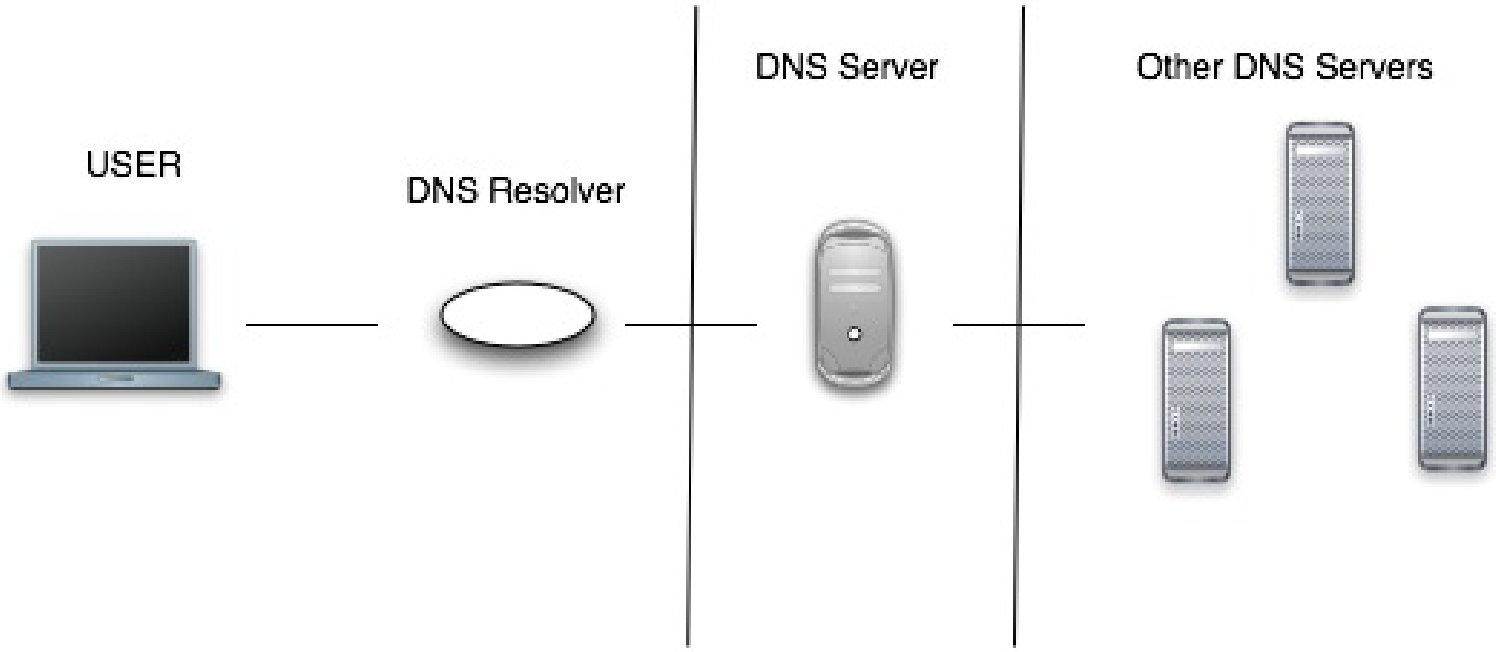
- Scanning worms can employ a variety of strategies to infect systems
  - Topological scanning
  - Slow scanning
  - Fast scanning
- So far, all make use of pseudo random generated 32-bit numbers to determine their targets
- The use of numeric IP addresses does not require a DNS lookup (Violation of typical network behavior (i.e. DNS).)

Cells in Enterpirse Networks



Queries/Answers in a DNS

The DNS is used on the Internet to correlate between IP address and readable names.



Resolver (part of system sending queries) is on the client side of the configuration. The name server answers the queries.

DNS-based Scanning Worm Detection Approach

- Inline network device
- Divide network into cells
- Gather DNS requests, embedded IPs in HTTP requests
- Construct a candidate connection list (CCL) respecting Time to Live (TTLs)
- Observe outgoing connections
- Those outgoing connections not matching an entry in the CCL generate an alert

Training Period and Rapid Response

- Technique can be used to rapidly determine if a host within an enterprise network is trying to infect external systems
- Anomaly-based "training period" required to generate whitelists
- Whitelists are valid non-DNS using protocols/activities
- Detects local to remote (L2R) and local to local (L2L) inter-cell propagation

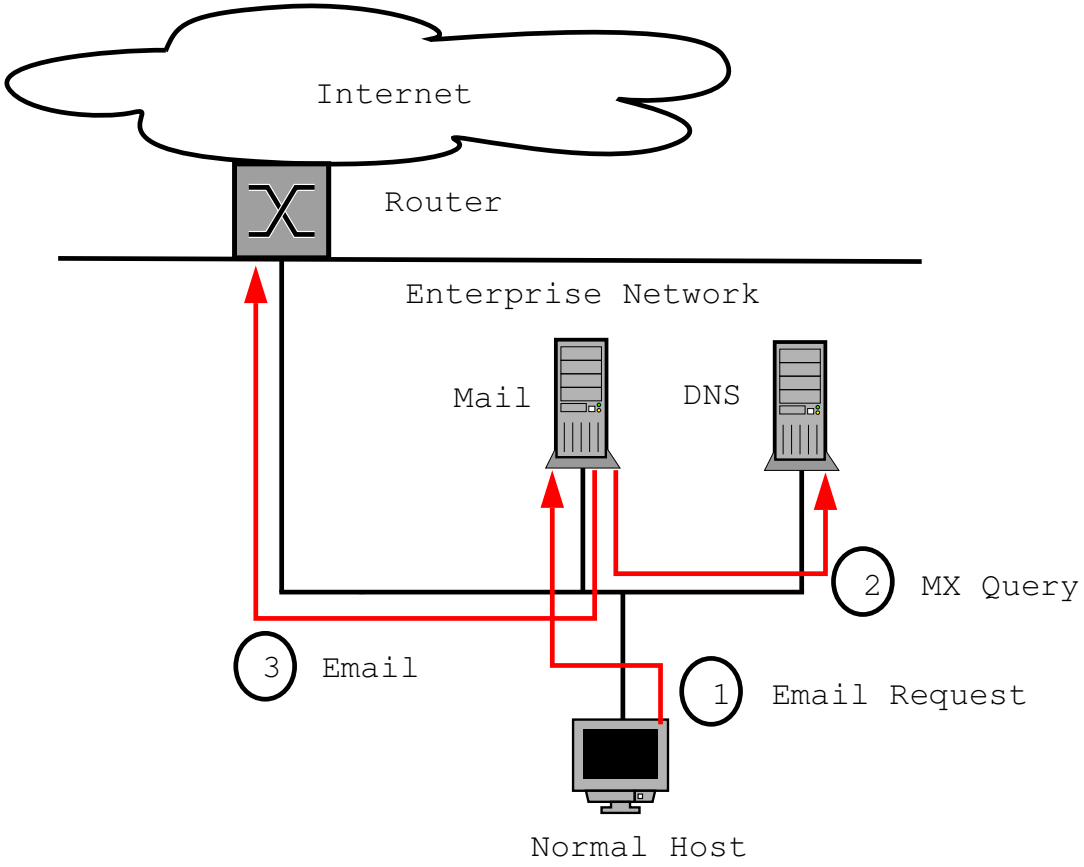
Anomaly-based Worm Detection Advantages

- Detection of zero-day worms / attack tools
- Detection of low and slow attacks - no threshold
- Low maintenance
- Relies on observation of a protocol found in all networks

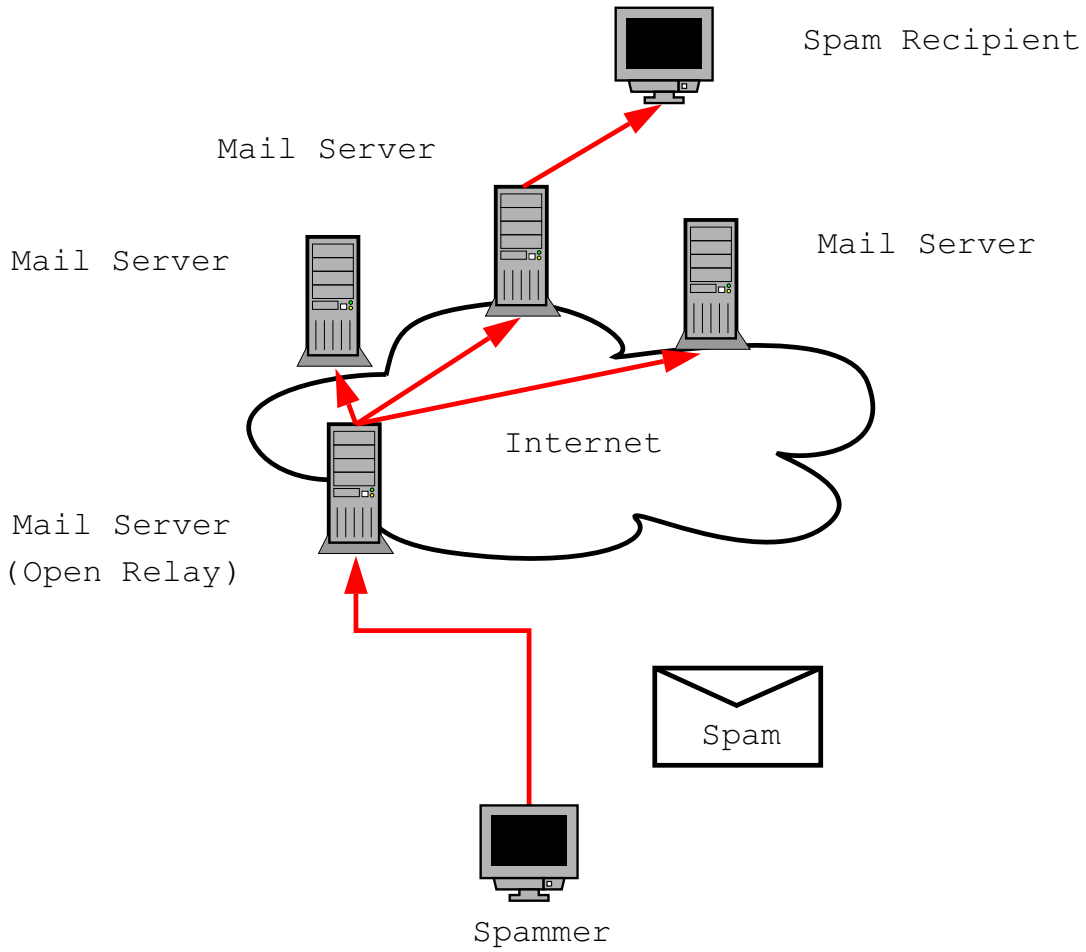
Limitations

- Will not detect intra-cell propagation
- Open networks cause large whitelists and the potential for false negatives
- Will not detect network share traversal propagation or mass mailing worms
- Automated scanning/attack tool false positives

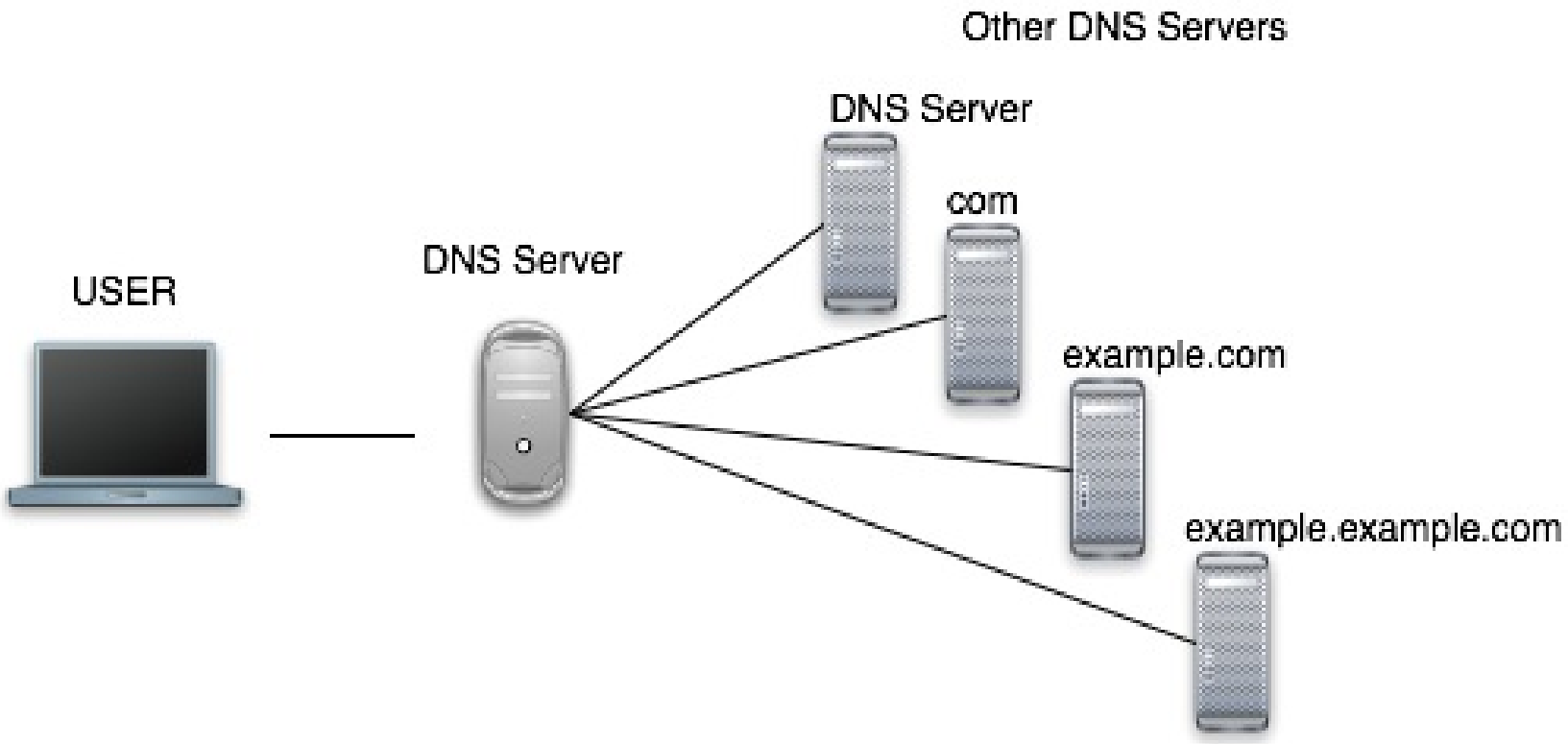
Extending the idea: Normal Hosts...



...and Spammers



...and Multiple DNS



The additional DNS servers could themselves form a network

Extensions

- ARP-based detection of scanning worms in an enterprise network
- Detection of Mass-mailing worm detection
- DNS-based selection and filtering
- Automated attack tool detection
- Covert communication detection

Optimal Movement of Mobile Sensors for Barrier Coverage of a Planar Region

(with Bhattacharya, Burmester, Hu, Shi, Wiese, COCOA 2008)

The Problem: Intrusion Detection

- Intrusion detection and border surveillance are important applications for wireless sensor networks.
- A major goal in these applications is to detect intruders.
- This is accomplished by *appropriate coverage*.
- Washington Post, Thursday, February 28, 2008
  - Project 28
  - U.S. Retooling High-Tech Barrier After 28-Mile Pilot Project Fails
  - Virtual Fence Along Border To Be Delayed

Types of Coverage

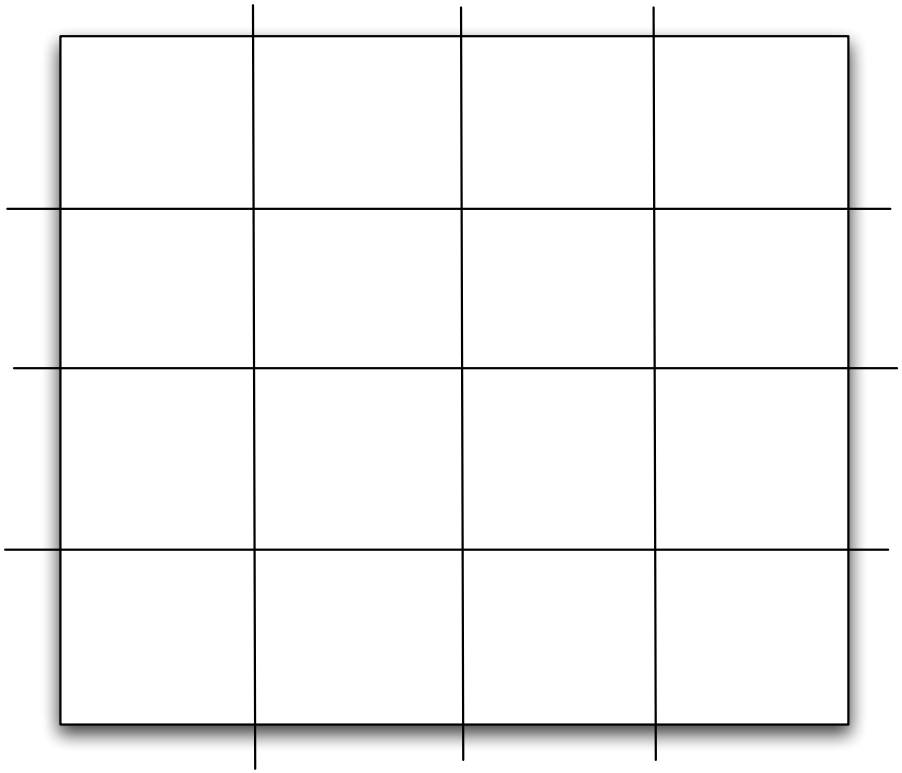
- There are two types of Coverage:
  - *Full* Coverage
  - *Barrier* Coverage

On Full Coverage (1/3)

- A major goal of full coverage is to detect the presence of intruders in a given protected area.
- This type of coverage is referred to as *(full)* coverage, where the sensors cover fully a region.
- A given region is said to be *k-fully covered* by a sensor network if every point inside the region is covered by (i.e., is within the range of) at least k sensors.
- k is a safety parameter that assumes an appropriate value.

On Full Coverage (2/3)

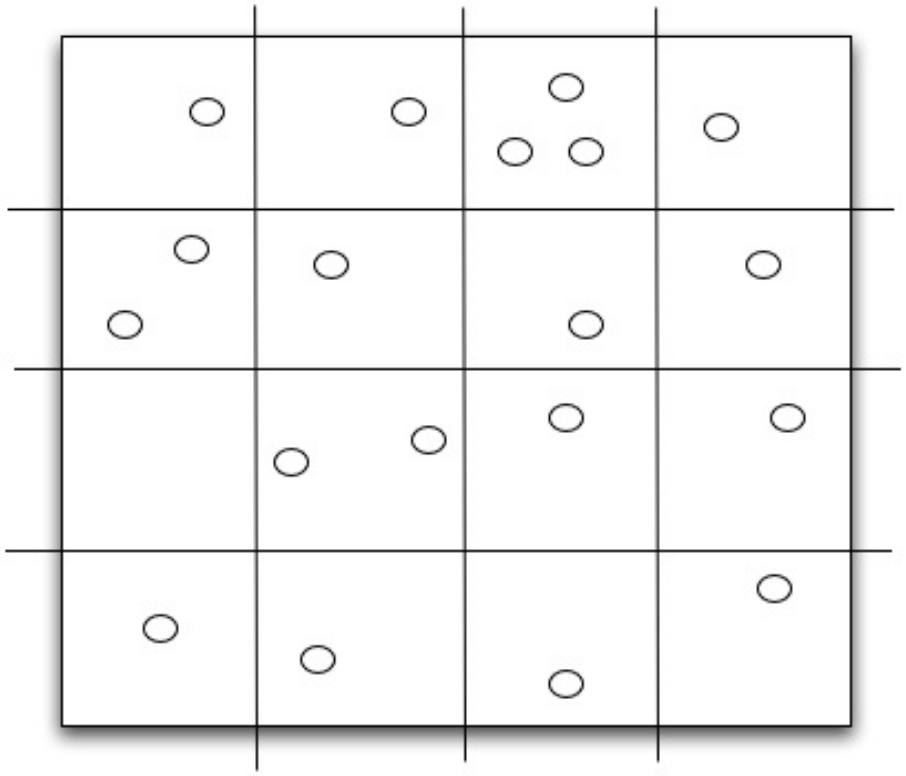
Assume the region is a unit square.



Partition it into subsquares.

On Full Coverage (3/3)

Throw sensors randomly and independently.



Questions about Full Coverage

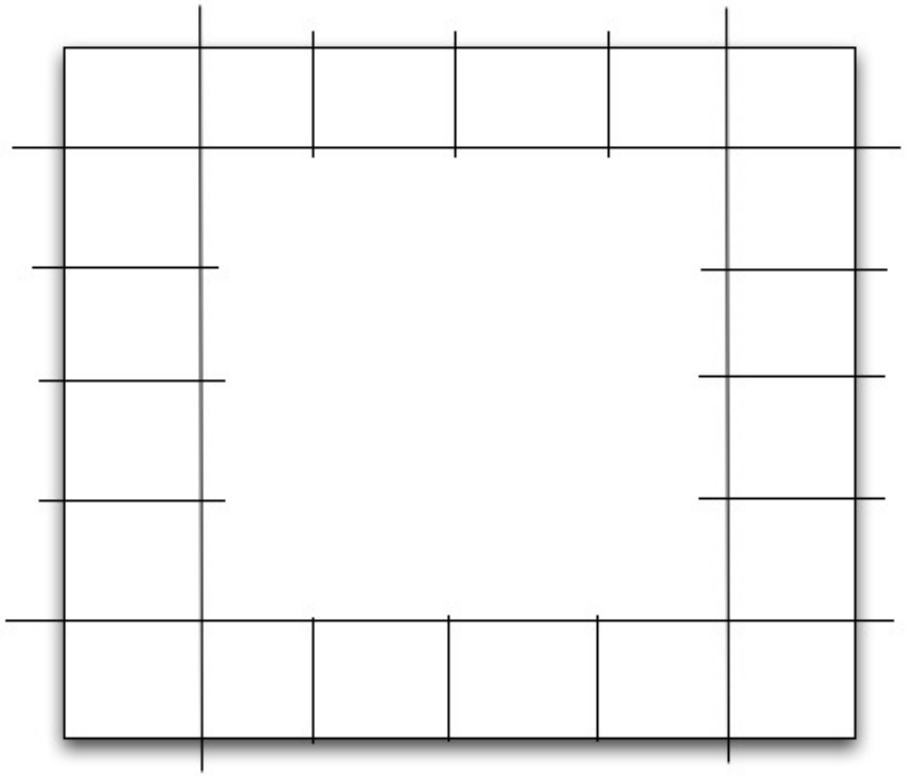
- Have you accomplished full coverage?
- Have you accomplished k-full coverage?
- For a given range $r > 0$ of the sensors:
  Are n sensors enough? With high probability?

On Barrier Coverage (1/3)

- A major goal in *barrier coverage* is to detect intruders as they cross a border or as they penetrate a protected area.
- This type of coverage is referred to as *barrier coverage*, since the sensors form a barrier for the intruders.
- A given belt region is said to be *k-barrier covered* by a sensor network if all crossing paths through the region are k-covered, where a crossing path is any path that crosses the width of the region completely.
- k is a safety parameter that assumes an appropriate value.

On Barrier Coverage (2/3)

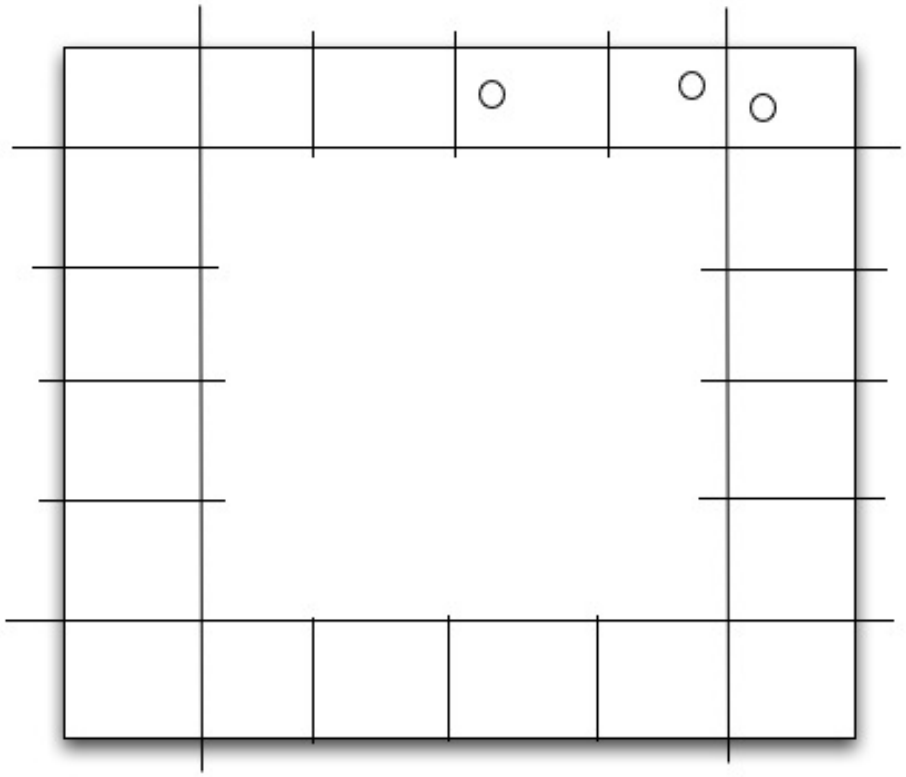
Assume the region is a unit square.



Partition "the perimeter" into subsquares.

On Barrier Coverage (3/3)

Throw sensors randomly and independently.

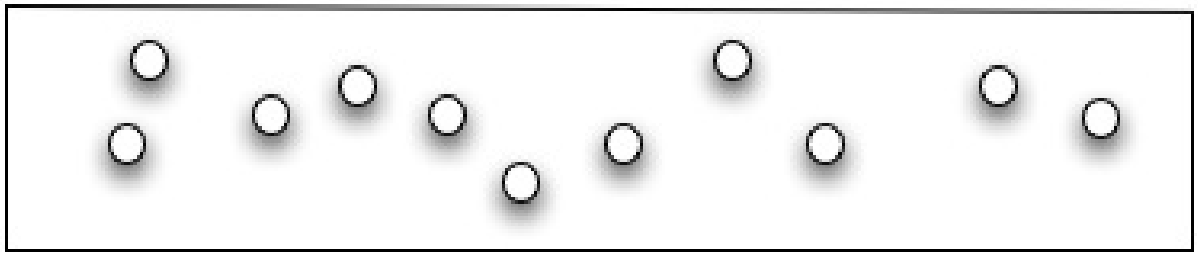


Questions about Barrier Coverage

- Have you accomplished barrier coverage?
- Have you accomplished k-barrier coverage?
- For a given range $r > 0$ of the sensors:
  Are n sensors enough? With high probability?

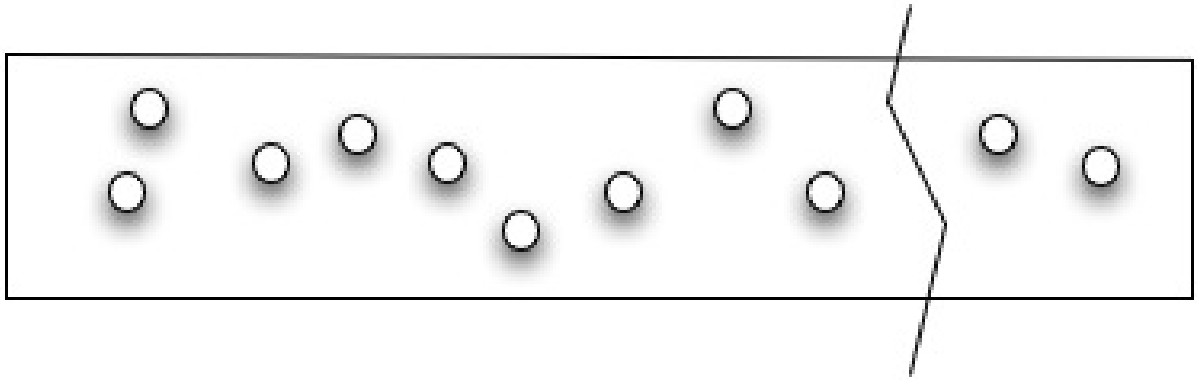
Intruder Detection

Suppose that sensors are already placed in a region.



Intruder Detection

Can an intruder penetrate the region undetected?



- A natural question then is how does one determine the minimum number of sensors to deploy to have k-barrier coverage in a given belt region?
- And, how does one determine, after deploying sensors in a region, whether the region is indeed k-barrier covered?

Intruder Detection

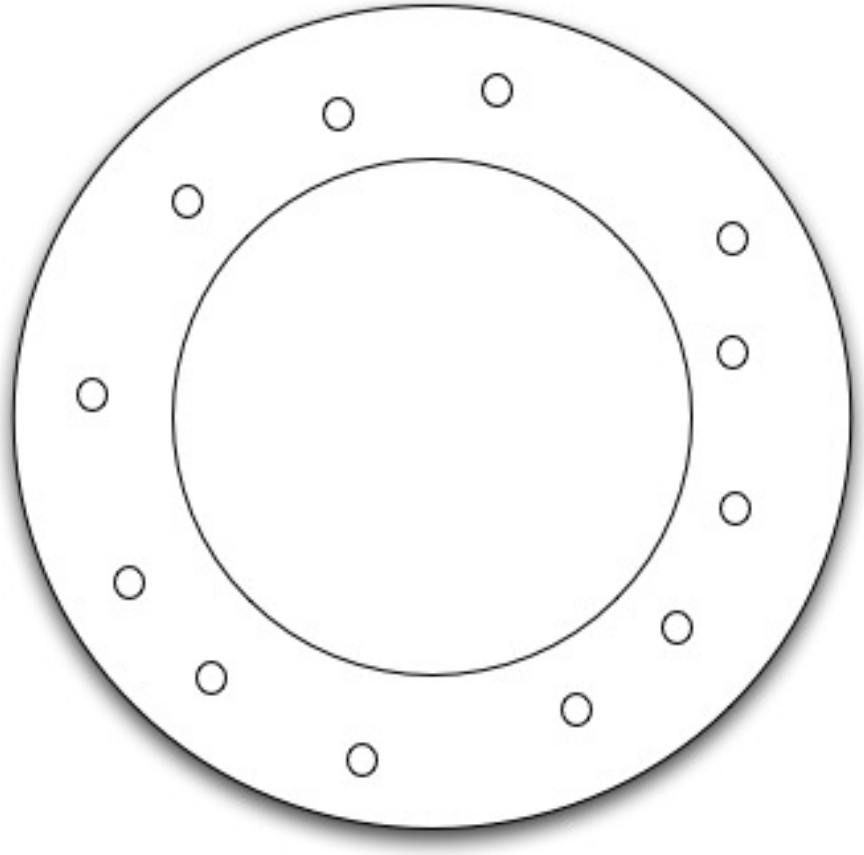
S. Kumar,T. H. Lai, A. Arora in MobiCom 2005:

- establish equivalence conditions between k-barrier coverage and existence of node-disjoint paths between two vertices in a graph
- prove that when deploying sensors deterministically, the optimal deployment pattern to achieve k-barrier coverage is to deploy k-rows of sensors on the shortest path across the length of the belt region such that consecutive sensors sensing disks abut each other.

Related to **Menger's theorem** (1927): Let G be an undirected graph and x, y two nonadjacent vertices. Then the size of the minimum vertex cut for x and y (the minimum number of vertices whose removal disconnects x and y) is equal to the maximum number of pairwise vertex-independent paths from x to y.

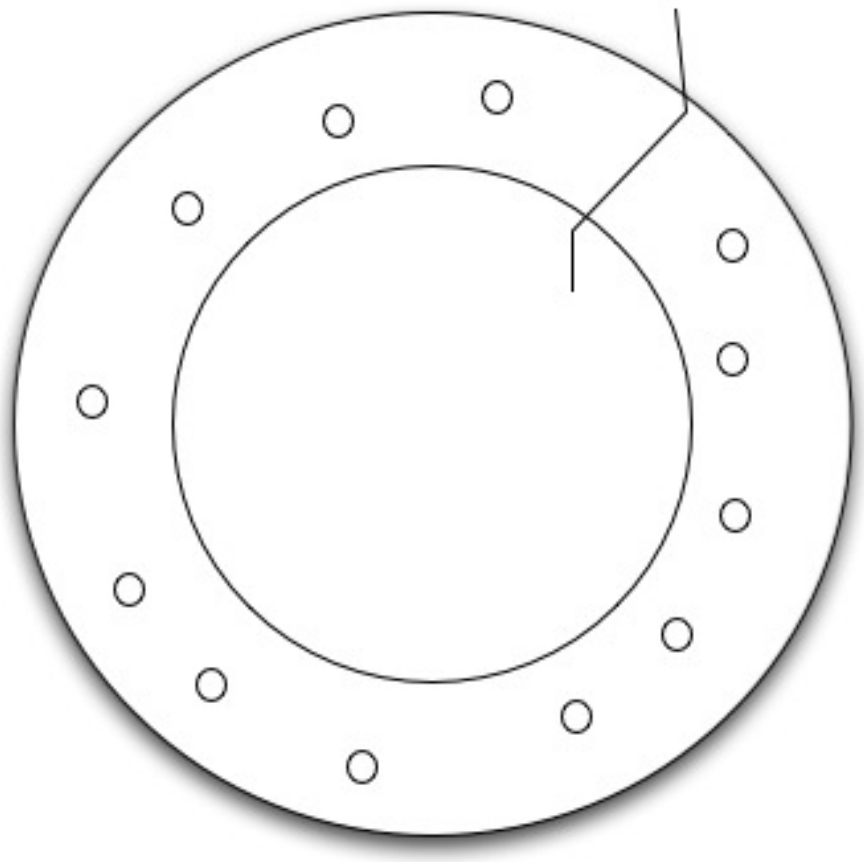
Intruder Detection

Suppose that sensors are already placed in a circular region.



Intruder Detection

Is intrusion possible?



Similar Questions

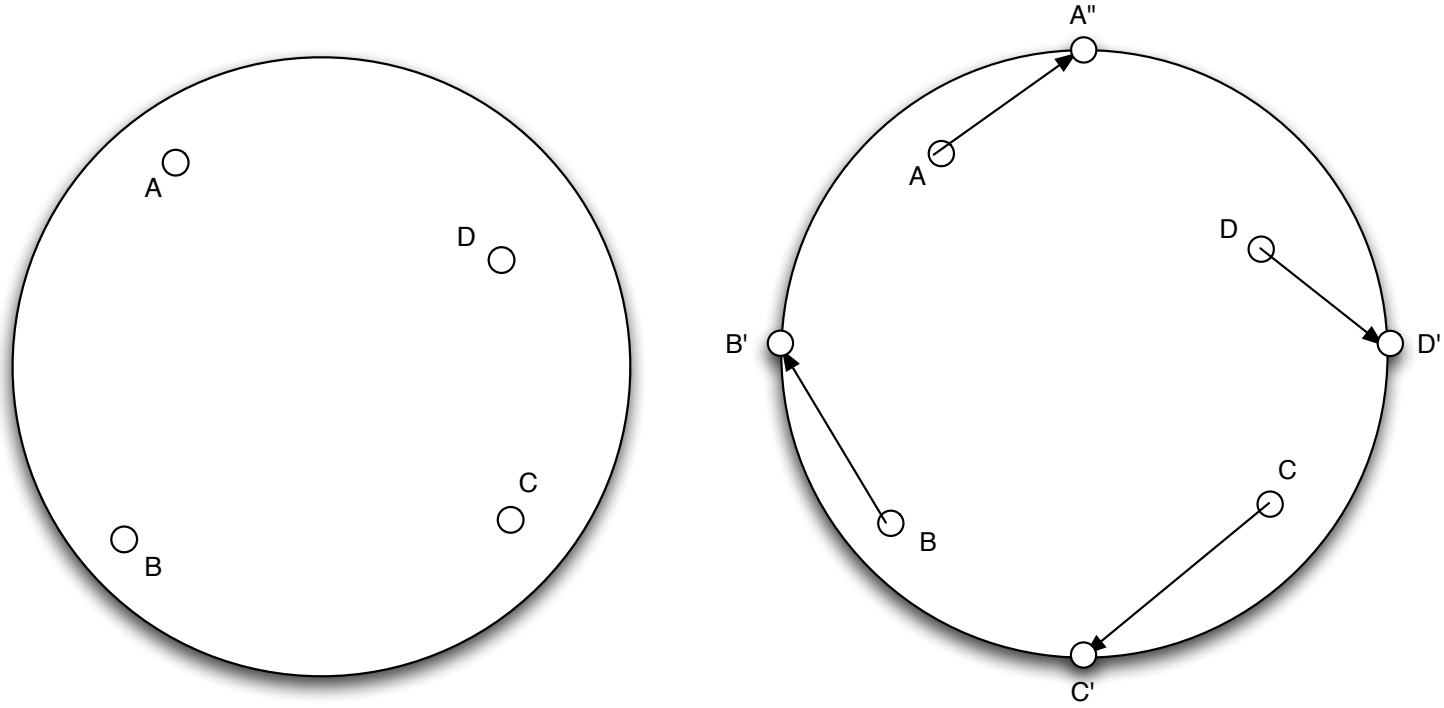
- Can an intruder penetrate the region undetected?
  - A natural question then is how does one determine the minimum number of sensors to deploy to have k-barrier coverage in a given belt region?
  - And, how does one determine, after deploying sensors in a region, whether the region is indeed k-barrier covered?
- Lex Schrijver (1991) proved necessary and sufficient conditions for the existence of pairwise vertex disjoint simple closed curves "homotopic" to given closed curves.
- Resulting complexity bounds are hard to determine: this indicates the problem is hardly closed!

Mobile Sensors: Changing the Model

- Previous models are static.
- Assuming
  - the sensors are mobile, and
  - the set of sensors does not *barrier cover* a region
- How do we move the sensors optimally so as to provide the best possible barrier coverage?

  What do we mean and why do we care about optimality?

Placing Robots in the Perimeter



Four sensors A, B, C, D located in the interior of a disk move to new positions A', B', C', D' towards the perimeter of the disk so that $A'B'C'D'$ forms a regular 4-gon.

Mobile Sensor Barrier Coverage

- n given sensors located inside a circle and starting in positions

  $$A_1, A_2, \ldots, A_n$$

  move to new positions

  $$A'_1, A'_2, \ldots, A'_n,$$

  respectively.

- To optimize barrier coverage every sensor moves from its current position A_i to a new position A'_i so that the new positions $A'_1, A'_2, \ldots, A'_n$ form the corners of a regular n-gon.

- The distance from A_i to A'_i is $d(A_i, A'_i)$.

Two Objectives

- **Min-Sum:** Minimize the sum

$$\sum_{i=1}^{n} d(A_i, A_i'). \tag{1}$$

- **Min-Max:** Minimize the maximum:

$$\max_{i=1}^{n} d(A_i, A_i'). \tag{2}$$

It is clear that in each case the function is minimized when each sensor moves to its new position in a straight line.

Formalities

- Let the n sensors have coordinates $A_i = (a_i, b_i)$, for $i = 1, 2, \ldots, n$.
- Let us parametrize the regular polygon with respect to the angle of rotation, say θ.
- The n vertices of the regular n-gon that lie on the perimeter of the disk can be described by

$$(a_i(\theta), b_i(\theta)) = (\cos(\theta + (i-1)2\pi/n), \sin(\theta + (i-1)2\pi/n)),$$

for $i = 1, 2, \ldots, n$), respectively.

Minimizing the Sum

- We are interested in minimizing the sum

$$S_n(\sigma,\theta) := \sum_{i=1}^{n} \sqrt{\left(a_i - a_{\sigma(i)}(\theta)\right)^2 + \left(b_i - b_{\sigma(i)}(\theta)\right)^2} \qquad (3)$$

  as a function of the angle θ and permutation σ.

- The optimization problem is

$$\min_{\sigma,\theta} S_n(\sigma,\theta). \qquad (4)$$

Minimizing the Max

- We are interested in minimizing a maximum

$$M_n(\sigma,\theta) := \max_{1\leq i\leq n} \sqrt{\left(a_i - a_{\sigma(i)}(\theta)\right)^2 + \left(b_i - b_{\sigma(i)}(\theta)\right)^2} \qquad (5)$$

  as a function of the angle θ and permutation σ.

- The optimization problem is

$$\min_{\sigma,\theta} M_n(\sigma,\theta). \qquad (6)$$

Summary of Results

- Min-Max optimization problem can be solved
  - 1D: optimally in $O(n)$ time
  - 1.5D: optimally in $O(n)$ time, and
  - 2D: optimally in $O(n^4 \log n)$ time.
- Min-Sum optimization problem can be solved
  - 1D: optimally in $O(n)$ time,
  - 1.5D: optimally in $O(n)$ time, and
  - 2D: $1+\epsilon$ approximation algorithm in time $O(\frac{1}{\epsilon}n^4)$.

Open Problems

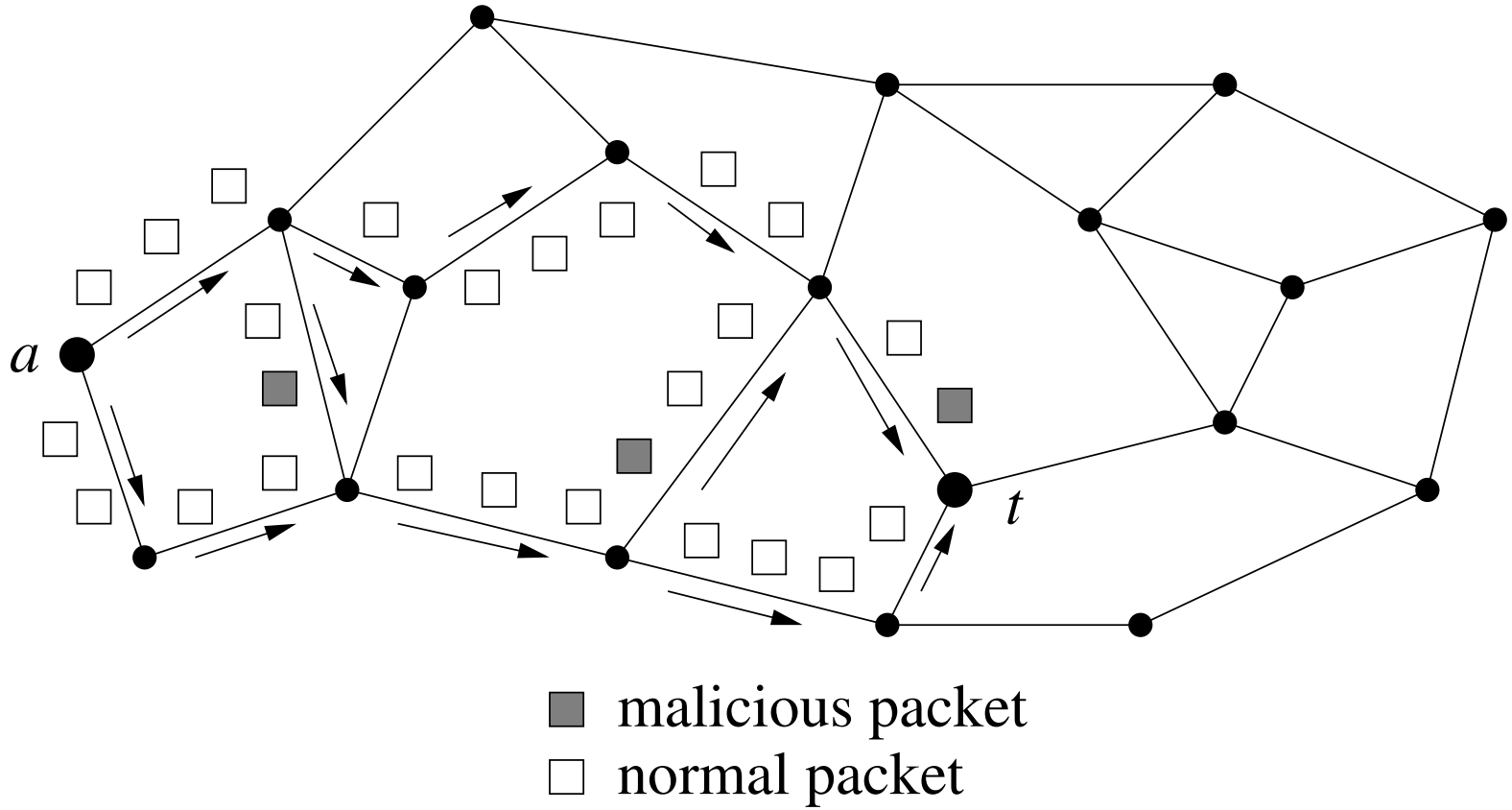
1. Can we solve the min-sum problem in polynomial time?
2. Can we improve existing results?
3. Different variants of the problem on simple polygons and regions:
   - regions with holes,
   - sensor placements,
   - Are some of them NP-hard?
4. Can we Improve
   - Motion model?
   - Network model?
   - Communication model?

Best Effort Models

The Problem

- For a given security scenario what is the best you can do so as to optimize the security performance?

Detect Presence of Malicious Packets in a Network



a is the attack node (from which an attack is initiated).

t is the targeted node (affected by the attack).

Mathematized Intrusion Detection

Problem:

- Intruder sends malicious packets to a given node.
- How do you detect malicious packets with packet-sampling?

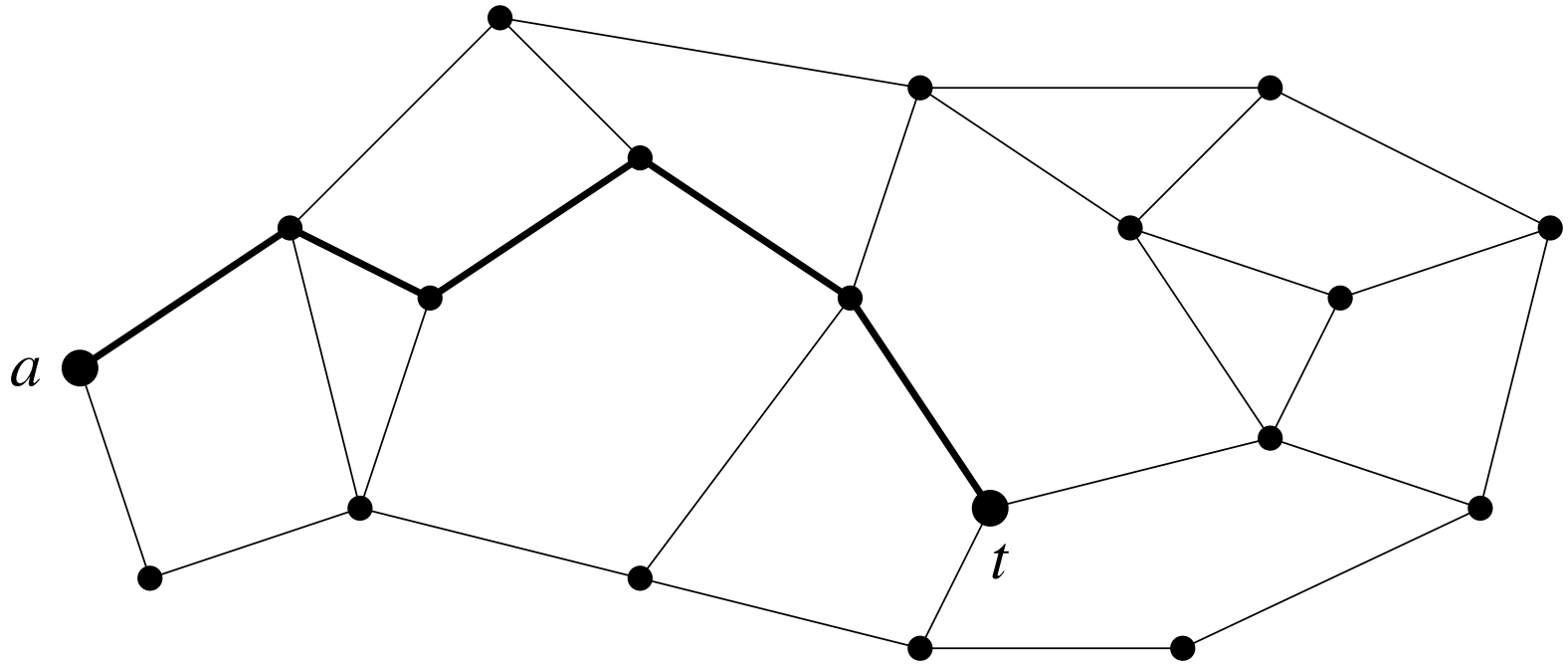
What does detection entail?

- Only a portion of "flowing" packets in designated links can be sampled.
- Examination may involve
  - specific packet header fields,
  - other data content details.

And the sampling costs?

- Must be kept low!

Intruder's Problem



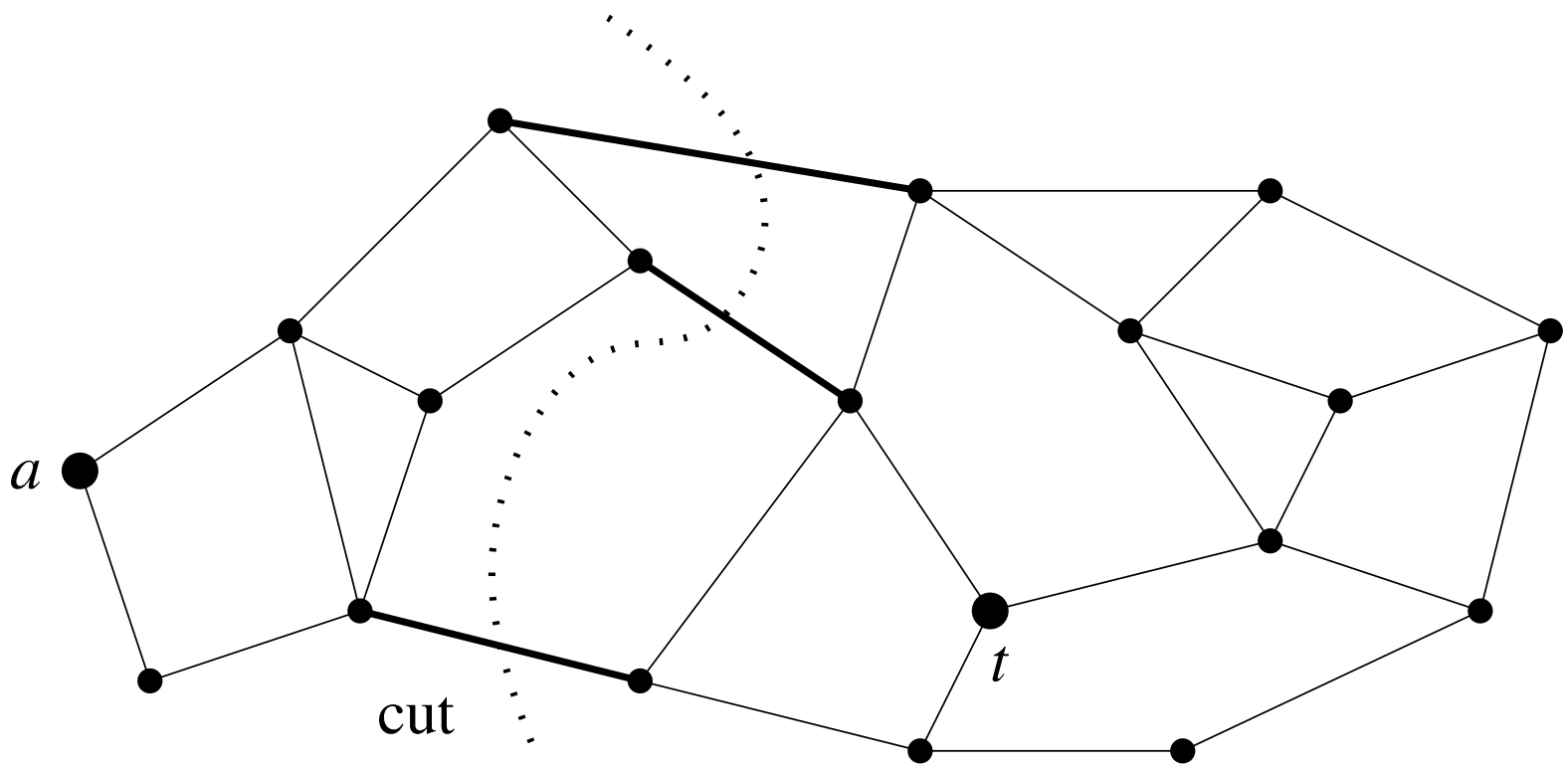
Intruder's Strategy:

pick a path on which the malicious packet will be sent!

Defender's Strategy:

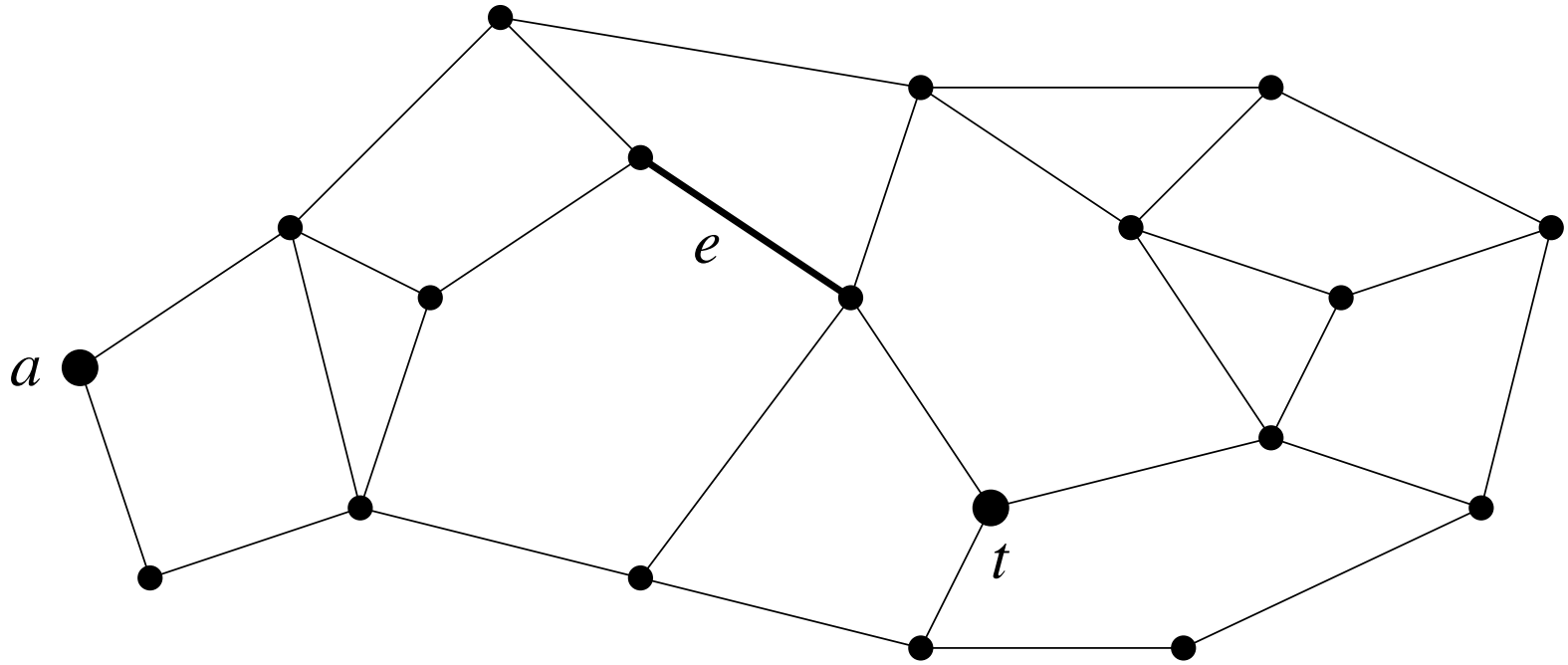
pick a sampling rate on the network links!

Defender's (i.e., Internet Provider) Problem



Sample links that belong to an $a - t$ mincut!

Capacity, Flow, and Sampling



A given link e is associated with:

- c_e: capacity,
- f_e: flow,
- s_e: sampling rate.

Players and Available Information

- Defender:
  - The defender has complete knowledge of the network.
  - The defender is an internet provider, and can get this information from the Link State Routing Protocol.
- Intruder:
  - The intruder may or may not be able to have complete knowledge of the network.
  - However, the most powerful intruder is one that also has complete knowledge of the network.

What are the Costs?

Probability of Detection in a Link e:

$$p_e := \Pr[\text{detecting a malicious packet at link } e] = s_e/f_e$$

Probability of Detection in a Path P from a to t:

$$1 - \prod_{e \in P} (1 - p_e)$$

Within Budget B:

$$\sum_{e \in E} s_e \leq B.$$

If you think of p_e as the sampling probabilities then the budget constraint becomes

$$\sum_{e \in E} f_e \cdot p_e \leq B.$$

Determining the Payoffs

Let $\mathcal{P}(a,t)$ be the set of paths from a to t. The intruder has a probability distribution $q(P)$ over paths $P \in \mathcal{P}(a,t)$ such that

$$\sum_{P \in \mathcal{P}(a,t)} q(P) = 1,$$

i.e., $q(P)$ is the probability that path $P \in \mathcal{P}(a,t)$ is selected.

For a given path $P \in \mathcal{P}(a,t)$, the expected number of times a packet is detected is equal to

$$\sum_{e \in P} p_e.$$

Expected number of times packet is detected as it moves from s to t

$$\sum_{P \in \mathcal{P}(a,t)} q(P) \sum_{e \in P} p_e$$

Objectives and Payoffs

Intruder: chooses a distribution $q :=< q(P) : P \in \mathcal{P}(a,t) >$ over the paths so as to minimize the maximum possible detection

$$\min_q \max_p \sum_{P \in \mathcal{P}(a,t)} q(P) \sum_{e \in P} p_e$$

Defender: chooses sampling distribution $p :=< p_e : e \in E >$ over the links so as to maximize the minimum possible detection

$$\max_p \min_q \sum_{P \in \mathcal{P}(a,t)} q(P) \sum_{e \in P} p_e$$

Min-Max: There exists an optimal solution that is achieved when

$$\min_q \max_p \sum_{P \in \mathcal{P}(a,t)} q(P) \sum_{e \in P} p_e = \max_p \min_q \sum_{P \in \mathcal{P}(a,t)} q(P) \sum_{e \in P} p_e$$

We do not know how to find such "equilibrium" strategies!

Intruder's Optimal Strategy

Given the network flow $f :=< f_e : e \in E >$

1. Find the maximum flow $M_{a,t}(f)$ from a to t.
2. Decompose the maximum flow into flows on paths

   $$P_1, P_2, \ldots, P_l,$$

   with flows $m_1, m_2, \ldots, m_l$, respectively. Note that

   $$\sum_{i=1}^{l} m_i = M_{a,t}(f).$$

3. **The intruder sends malicicious packet along path P_i which maximizes**

   $$\frac{m_i}{M_{a,t}(f)}.$$

Defender's Optimal Strategy

Given the network flow $f :=< f_e : e \in E >$

1. Find the links $e_1, e_2, \ldots, e_r$ of a minimum cut flow, say

$$f_{e_1}, f_{e_2}, \ldots, f_{e_r}.$$

   Note that

$$\sum_{i=1}^{r} f_{e_i} = M_{a,t}(f).$$

2. **The defender samples links $e_1, e_2, \ldots, e_r$ at the rate**

$$B \cdot \frac{f_{e_i}}{M_{a,t}(f)}$$

Conclusion

Metaphysics and Empiricism

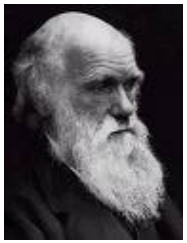
- *Metaphysics* is the branch of philosophy concerned with understanding existence and knowledge.
- *Empiricism* ($E\mu\pi\epsilon\iota\rho\iota\sigma\mu os$) is a theory of knowledge emphasizing the role of experience in the formation of ideas, while discounting the notion of *innate* ideas.

John Locke and State of Mind

- *John Locke* (Aug 29, 1632 to Oct 28, 1704) considered the first of the British Empiricists, important to social contract theory.
- He postulated that the mind was a "blank slate" and contrary to Cartesian or Christian philosophy, Locke maintained that people are born without "inborn" ideas.

Locke, Darwin, Metaphysics,...,and Baboons

...He who understands baboon will do more toward understanding metaphysics than Locke.



Darwin, August 16, 1838

Security Metaphysics

Pessimist:

...He who understands human behavior will do more toward understanding security than any mathematical model...

Optimist:

...He who understands mathematical models will do more toward understanding security than any security analyst...

Related Literature

F. Akujobi, I. Lambadaris, E. Kranakis, Endpoint-Driven Intrusion Detection and Containment of Fast Spreading Worms in Enterprise Networks, In Proceedings of MILCOM 2007, Oct 29-31.

M. Barbeau, J. Hall, E. Kranakis, Detecting Impersonation Attacks in Future Wireless and Mobile Networks, In proceedings of MADNES 2005

M. Kodialam, T. V. Lakshman, Detecting Network Intrusions via Sampling: A Game Theoretic Approach, INFOCOMM 2003.

D. Whyte, E. Kranakis, P. Van Oorschot, DNS-based Detection of Scanning Worms in an Enterprise Network, NDSS 05.

B. Bhattacharya, M. Burmester, Y. Hu, E. Kranakis, Q. Shi, A. Wiese, Optimal Movement of Mobile Sensors for Barrier Coverage of a Planar Region. COCOA'08