

On The Stochastic Modeling and Detection of RCS Worms



Kurt R. Rohloff, Tamer Başar

Coordinated Science Laboratory
The University of Illinois
Urbana, IL, USA
{krohloff,tbasar}@uiuc.edu

WORM

Nov. 11, 2005



Updated Contact Information

Kurt Rohloff

BBN Technologies

Cambridge, MA 02138

krohloff@bbn.com

www.bbn.com



RCS Worms

- Focus on **Random Constant Scanning** Worms.
Infected hosts randomly transmits infections.
Morris, CodeRed, Slammer, etc...
- Current RCS worm models deterministic.
Ignores stochastic nature of propagation.
- Analyze best-case anomaly detection limitations.
Trade-off detection time vs. false-alarm rate.



Stochastic RCS Worm Models

- **RCS Worms.**
- Markov Jump Model.
- Sequential Hypothesis Testing.
- Detection Properties.



Deterministic RCS Worm Model

- n : Population size.
- n_s : Susceptible population.
- $i(t)$: Number of infected hosts.
- $s(t)$: Number of susceptible uninfected hosts.

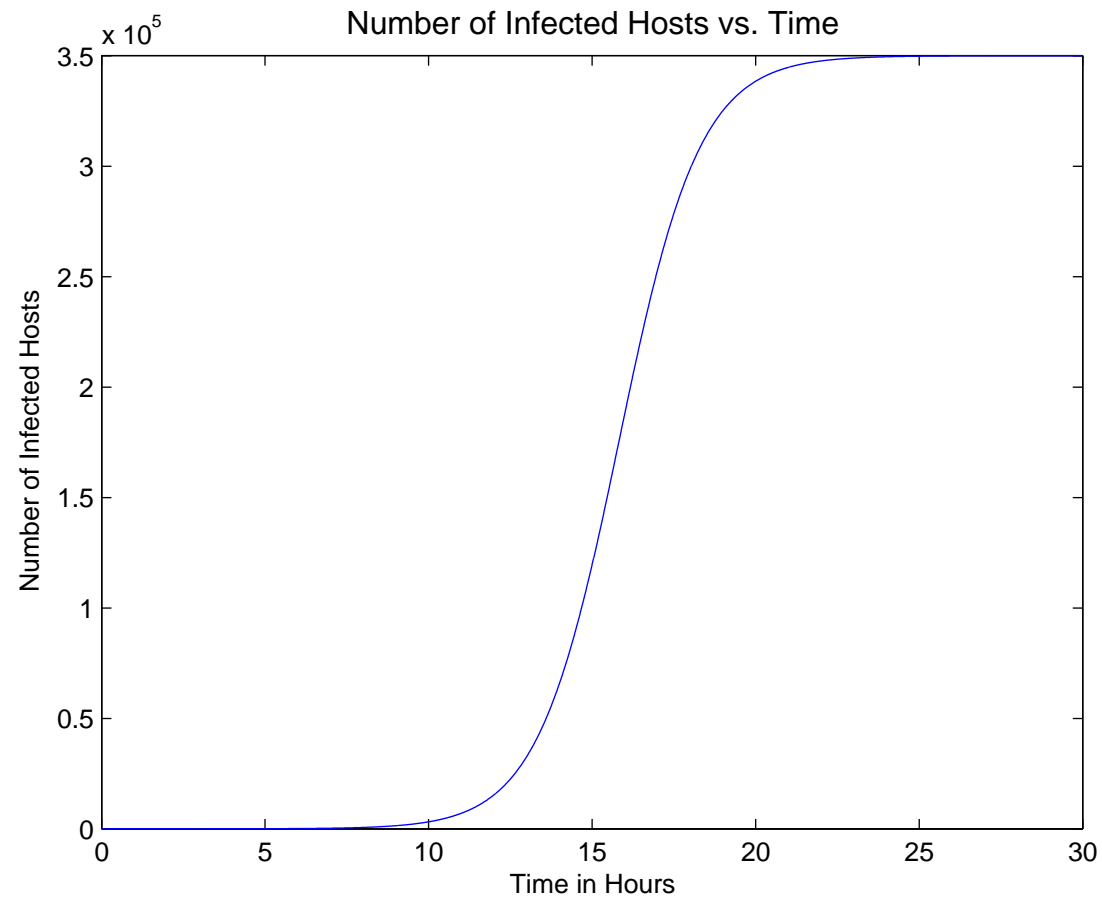
$$n_s = i(t) + s(t)$$

- B : Infection parameter.

$$\frac{di}{dt} = Bs(t)i(t)/n$$



CodeRed Propagation Simulation





Stochastic RCS Worm Models

- RCS Worms.
- **Markov Jump Model.**
- Sequential Hypothesis Testing.
- Detection Properties.



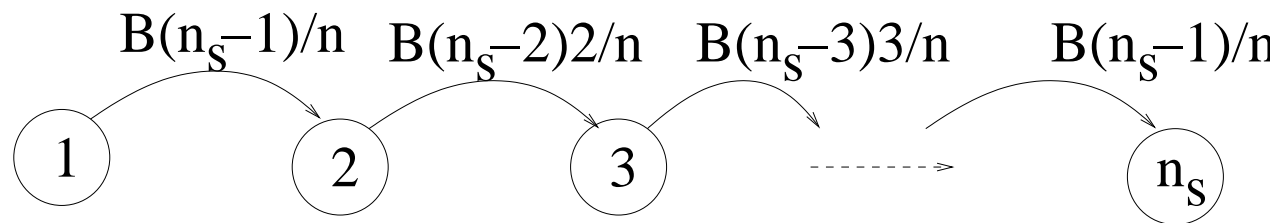
Markov Jump Model

- More accurate to model worm as **Markov Jump**

Process:

Exponential distr. on jump times.

Future behavior depends only on current state.





Markov Jump Model

- Jump intensity is a stochastic measure of jump rate.

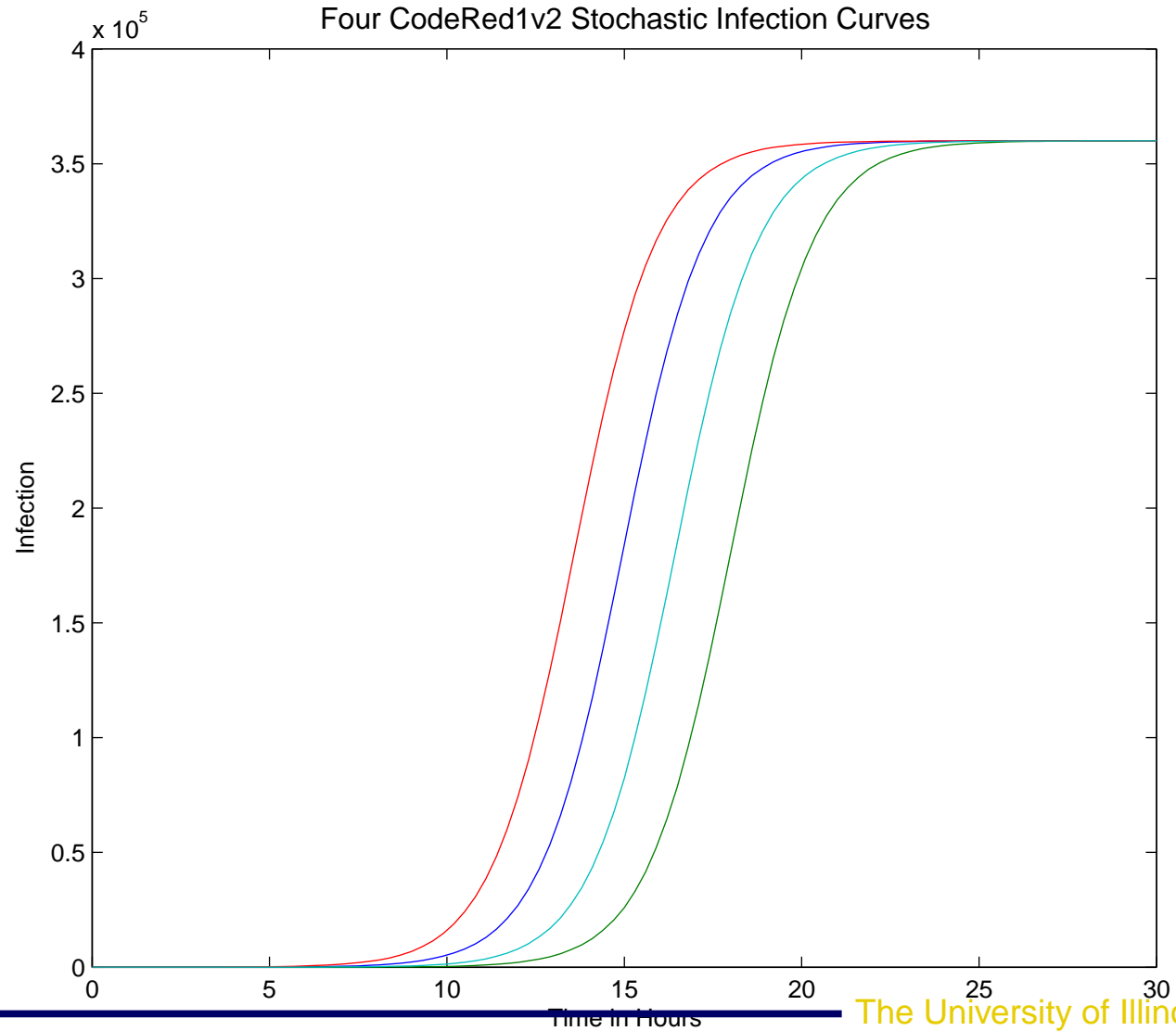
For RCS worms, jump intensity is density dependent.

Density dependent Markov jump model.

- K. Rohloff, T. Başar. Stochastic Behavior of Random Constant Scanning Worms. *In Proc. of 14th ICCCN, 2005.*



CodeRed Infection Propagation





Markov Jump Model

- Curiously, jump model simulations are uniform.
Same curve shifted in time.
- Effect on detection properties?
Use stochastic hypothesis testing methods?



Stochastic Model Analysis

- Stochastic propagation has Gaussian distribution about mean.

As n_s becomes large, deterministic model approaches stochastic model.

- Can develop better local scan observation model.
Model propagation with deterministic model.
Model local scanning with stochastic model.
- n_t : Size of local network.



Stochastic Model Analysis

- Scan interarrival distribution time complicated due to propagation growth.
- Can easily simulate interarrival times accurately.
- Any one scan has probability n_t/n of hitting local network.



Stochastic Model Analysis

- Consider global propagation scans in network.
First scan, second scan, n th scan...
Some scans hit local n_t hosts.
Each scan Bernoulli trial.
- P_m is index of m th scan that hits n_t local hosts.
- Model $P_1, P_2 - P_1, P_3 - P_2, \dots$ as exponentially distributed with mean n_t/n . (Bernoulli trials.)

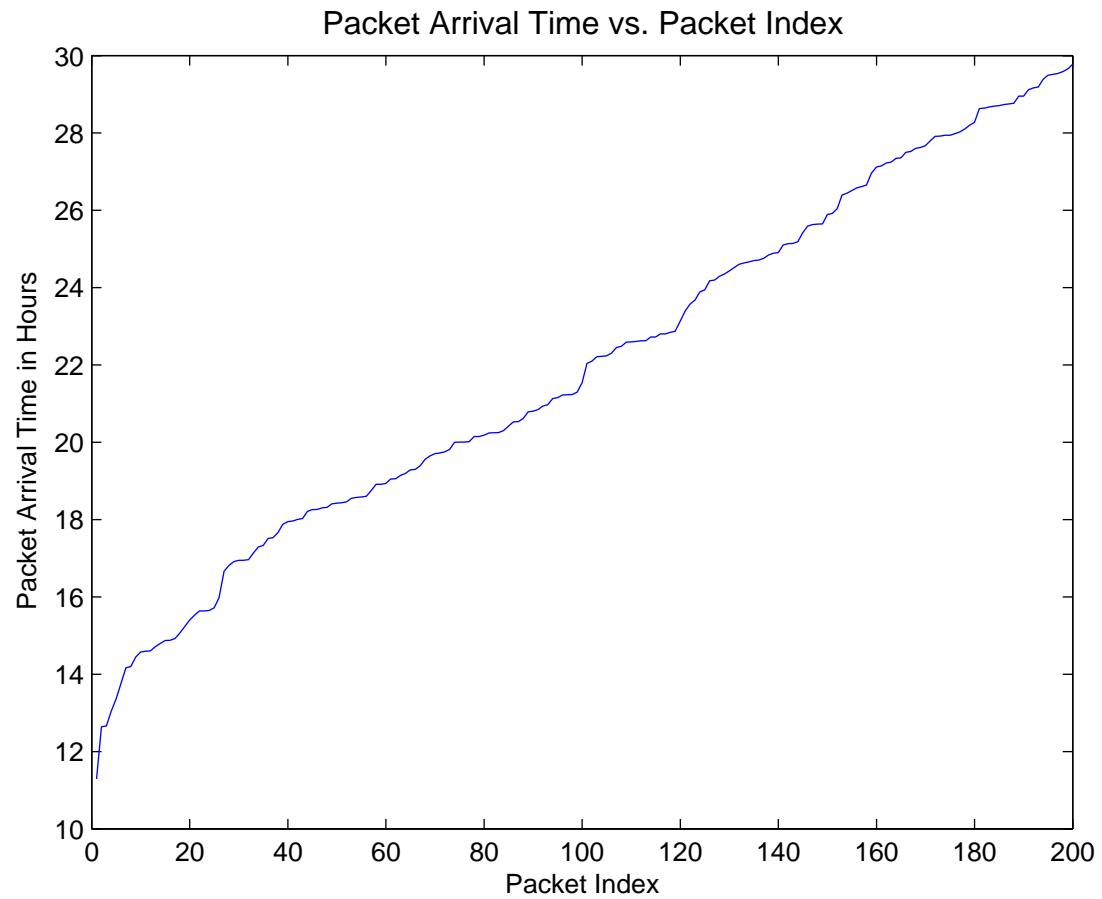


Stochastic Model Analysis

- Suppose $P_1, P_2 - P_1, P_3 - P_2, \dots$ are exponentially distributed with mean n_t/n . (Bernoulli scan trials.)
- Let $t(p) = \frac{n}{\beta n_s} \left[\frac{p}{n} + \log \left(\frac{n_s - (n_s - i_0) e^{-p/n}}{i_0} \right) \right]$.
- $t(P_1), t(P_2), t(P_3), \dots$ are scan interarrival times.
- Now have stochastic local scanning model.

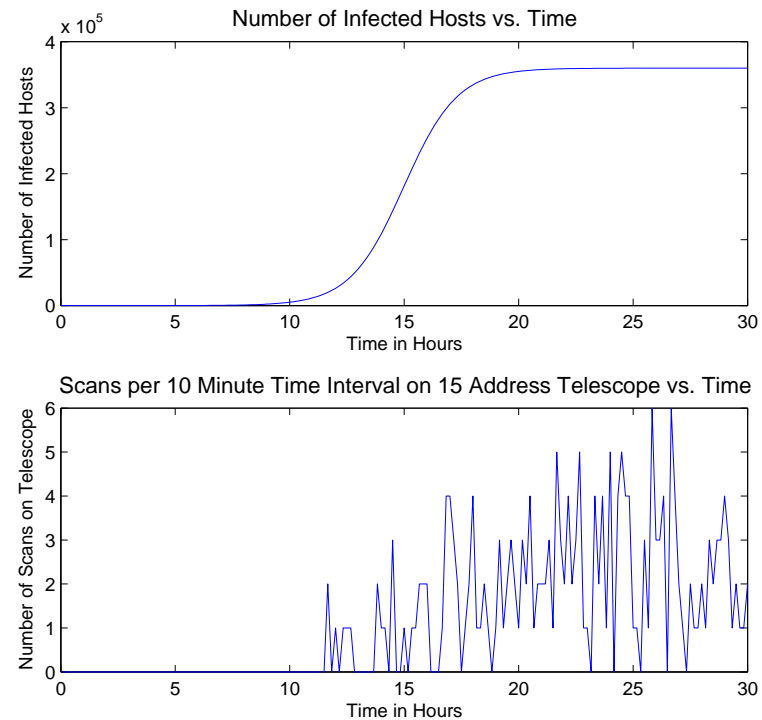


Scan Arrival Times





Hybrid Scanning Model



- Underlying global worm propagation.
- Top-level local observation model.



Stochastic RCS Worm Models

- RCS Worms.
- Markov Jump Model.
- **Sequential Hypothesis Testing.**
- Detection Properties.



Hypothesis Testing

- Suppose two possible hypothesis of world:

H_0 : No worm epidemic.

H_1 : Worm propagating on Internet.

- Suppose have set of observations of world:

$$Y^n = \{o_1, o_2, o_3, \dots\}$$

- Based on observations, want to determine which hypothesis is true.

Hypothesis testing.



Sequential Analysis

- Generally observations Y^n made online.
- Want to detect worm as observations being made.

Sequential hypothesis testing.

- Philosophy: Use all information available up to current time.



Probability Ratio

- Probability measures $p_0(Y^n)$, $p_1(Y^n)$.

$p_i(Y^n)$ is likelihood Y^n observed under H_i .

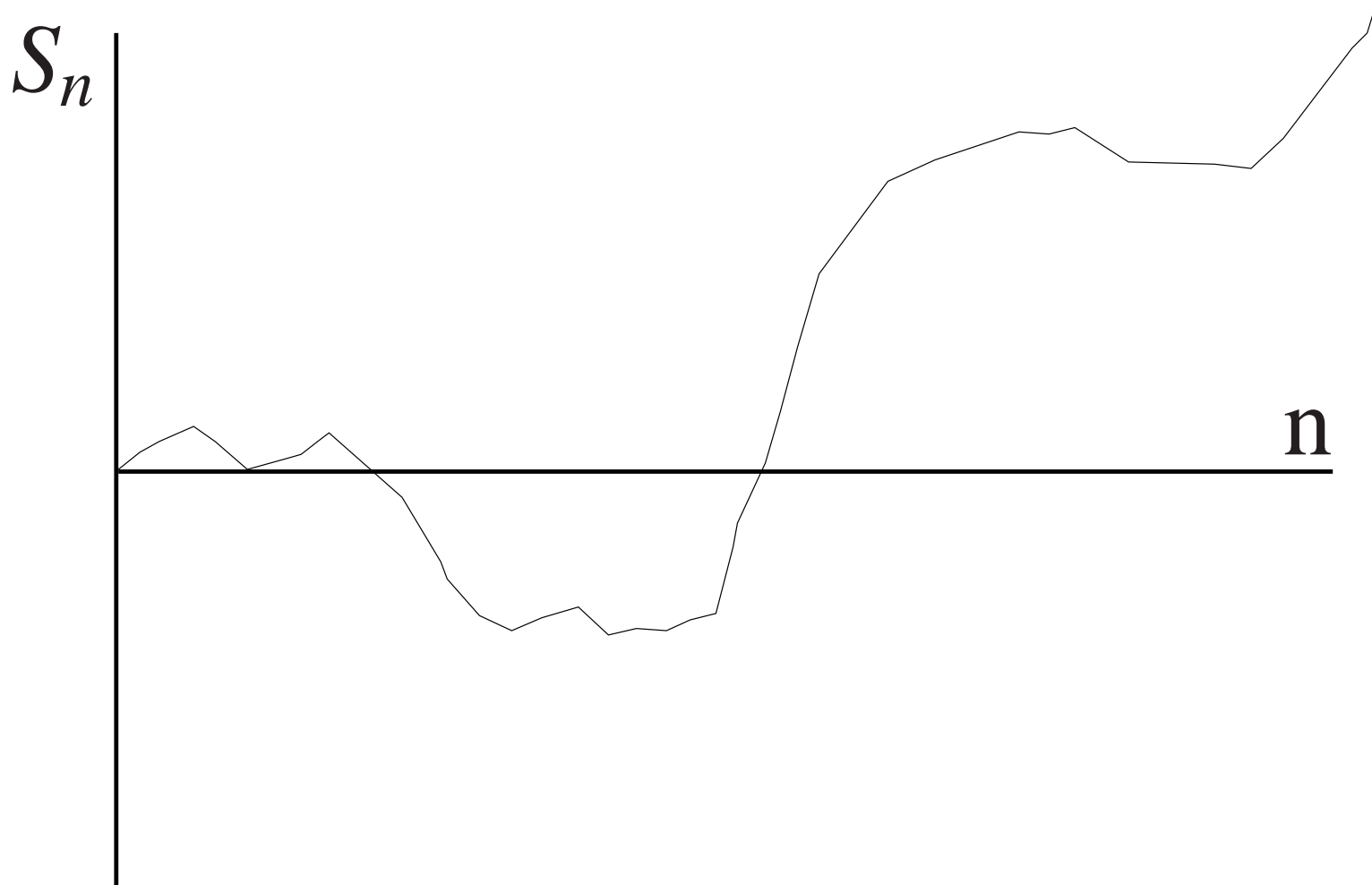
- Probability ratio test:

$$S_n = \ln \left(\frac{p_1(Y^n)}{p_0(Y^n)} \right).$$

- S_n is hypothesis confidence measure.



S_n vs. n Simulation





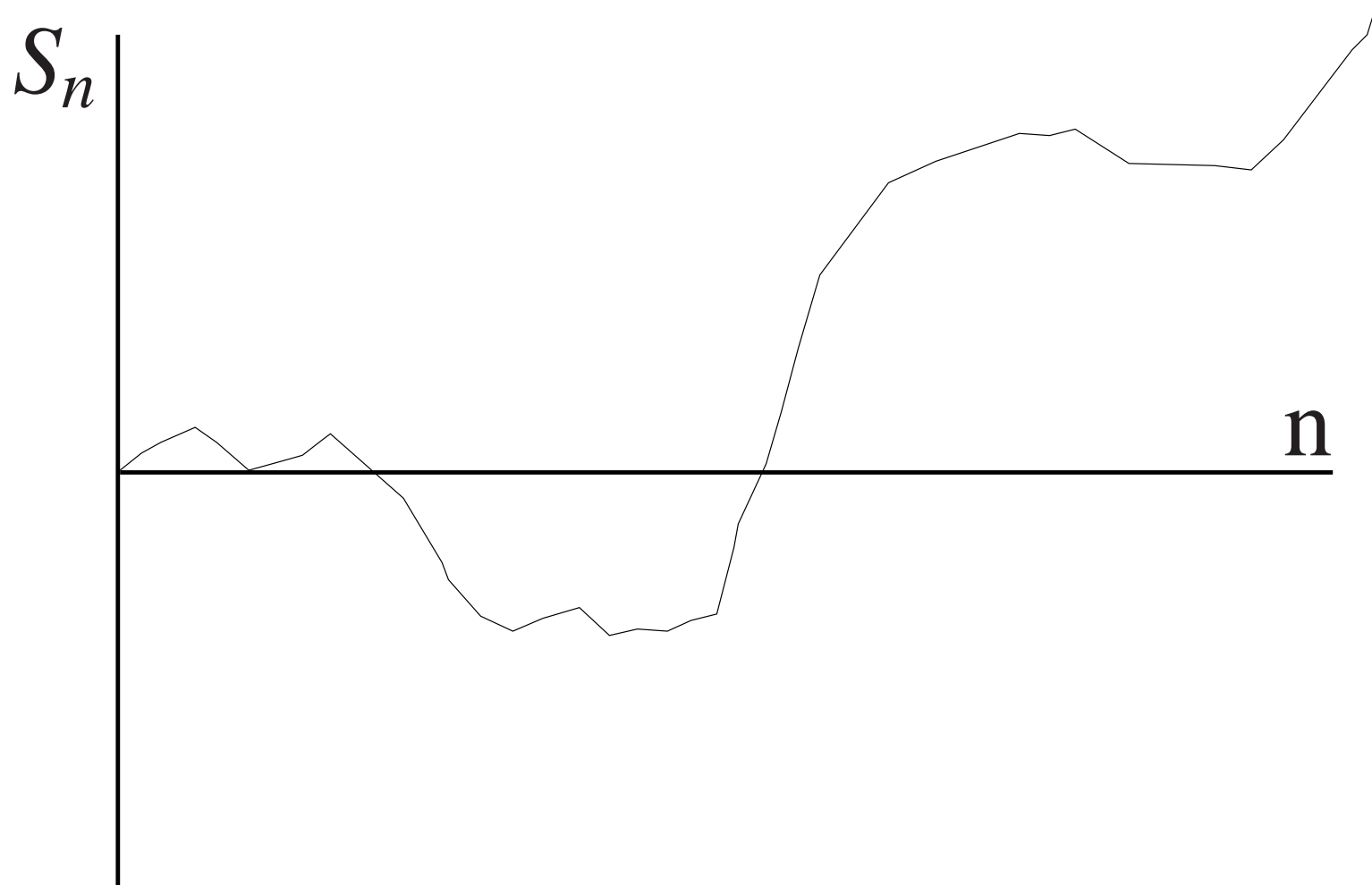
Threshold Passing

- $S_n = \ln \left(\frac{p_1(Y^n)}{p_0(Y^n)} \right)$.
If S_n large $\Rightarrow H_1$.
If S_n small $\Rightarrow H_0$.
- Threshold Passing Test:

$$g(Y^n) = \begin{cases} 1 & \text{if } S_n \geq h \\ 0 & \text{if } S_n \leq -a \end{cases} . \quad (1)$$

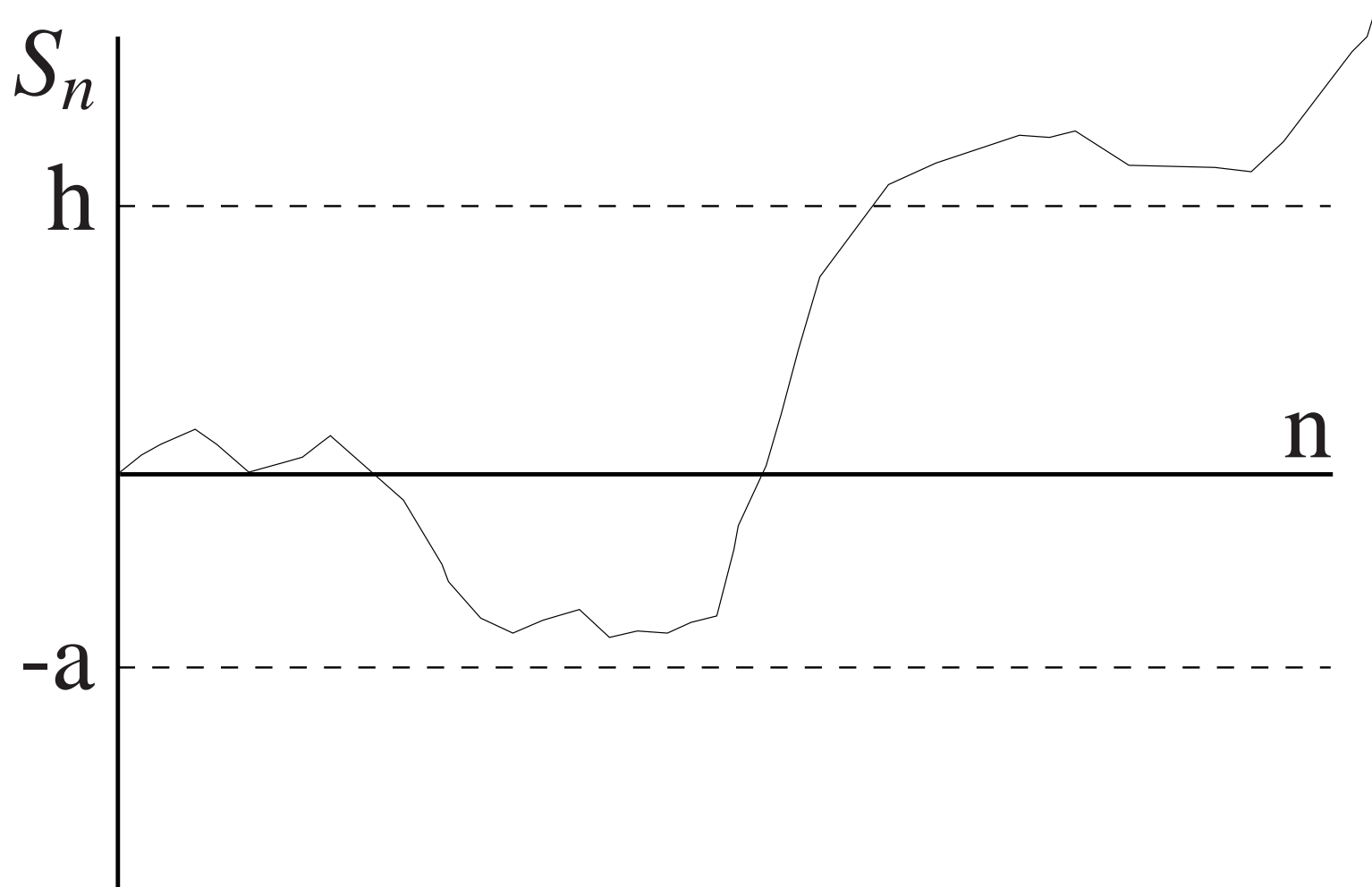


Threshold Passing



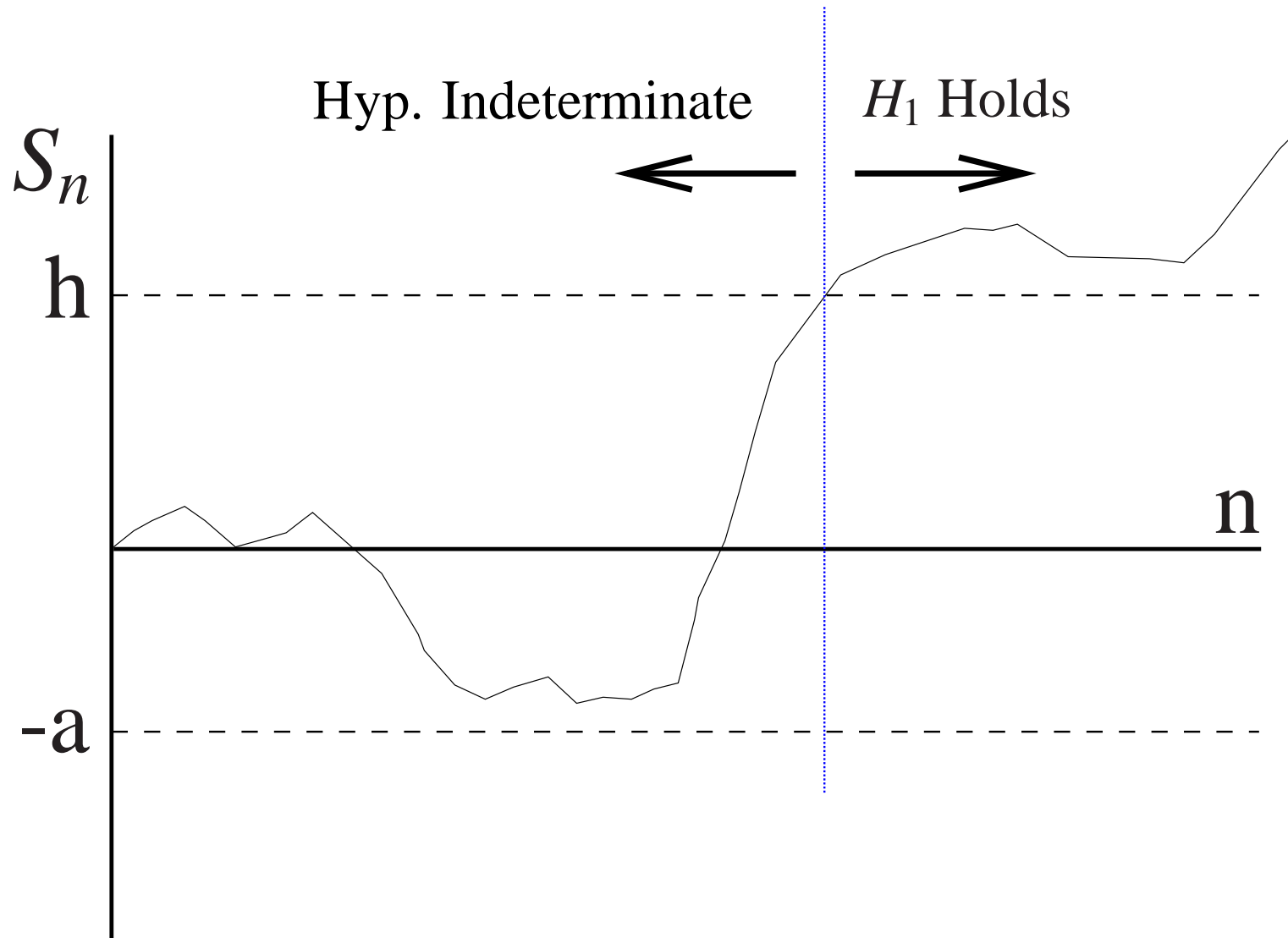


Threshold Passing





Threshold Passing





Sequential Analysis

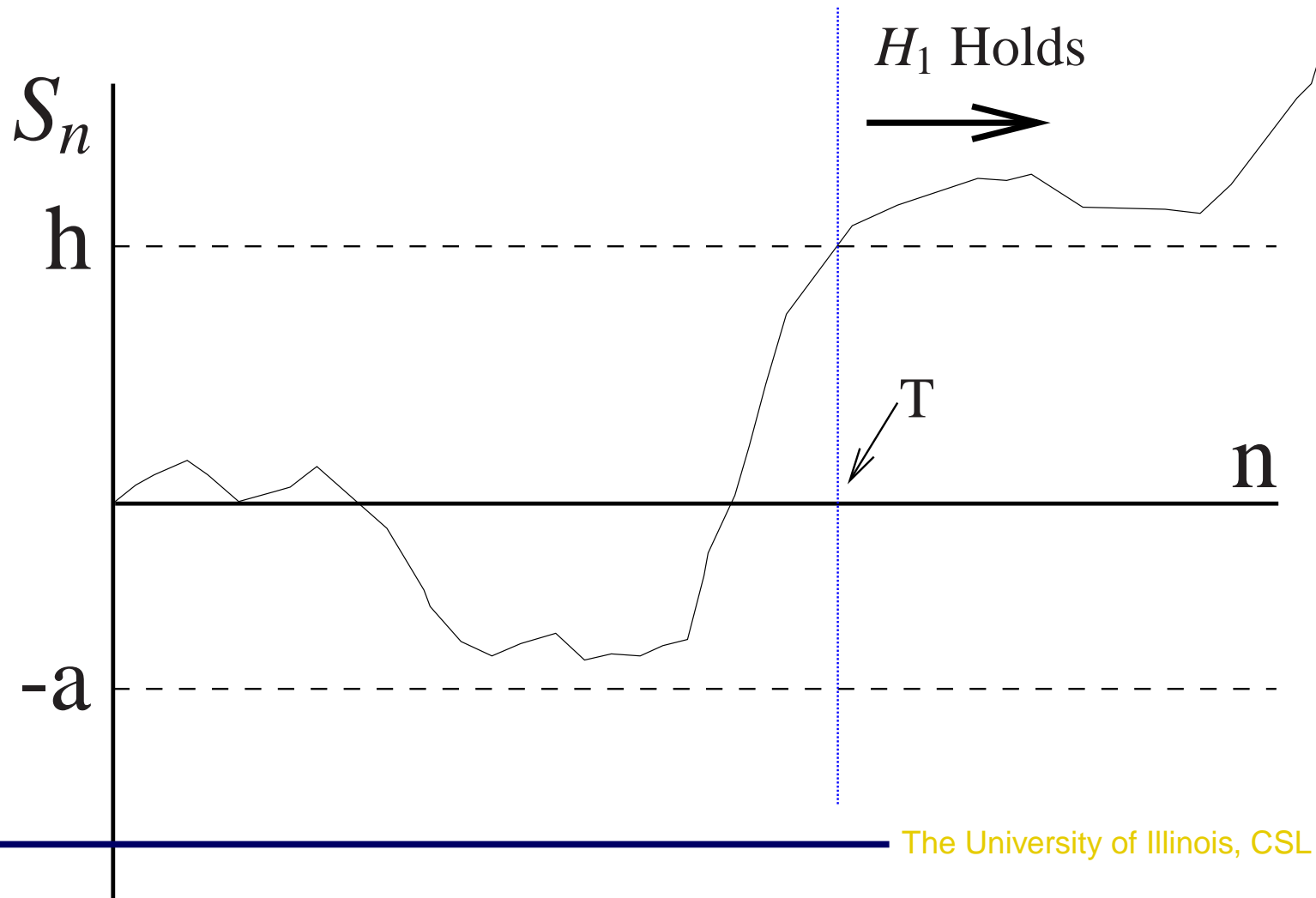
- $Y^n = \{o_1, o_2, o_3, \dots\}$.
- $T = \min\{n > 1 : g(Y^n) \in \{0, 1\}\}$.
- Sequential Probability Ratio Test (SPRT):
(Also called CUSUM test.)

$$g'(Y) = \begin{cases} 1 & \text{if } g(Y^T) = 1 \\ 0 & \text{if } g(Y^T) = 0 \end{cases} . \quad (2)$$



Threshold Passing

SPRT Result: $g'(Y^T) = 1$





Sequential Analysis

- $T = \min\{n > 1 : g(Y^n) \in \{0, 1\}\}$.

- T is a random variable.

$E[T]$ is **Average Run Length(ARL)**.

- α_0 is false alarm rate.

- SPRT is optimal sequential hyp. testing method.

No method with better false alarm rates, ARL's.



Stochastic RCS Worm Models

- RCS Worms.
- Markov Jump Model.
- Sequential Hypothesis Testing.
- **Detection Properties.**

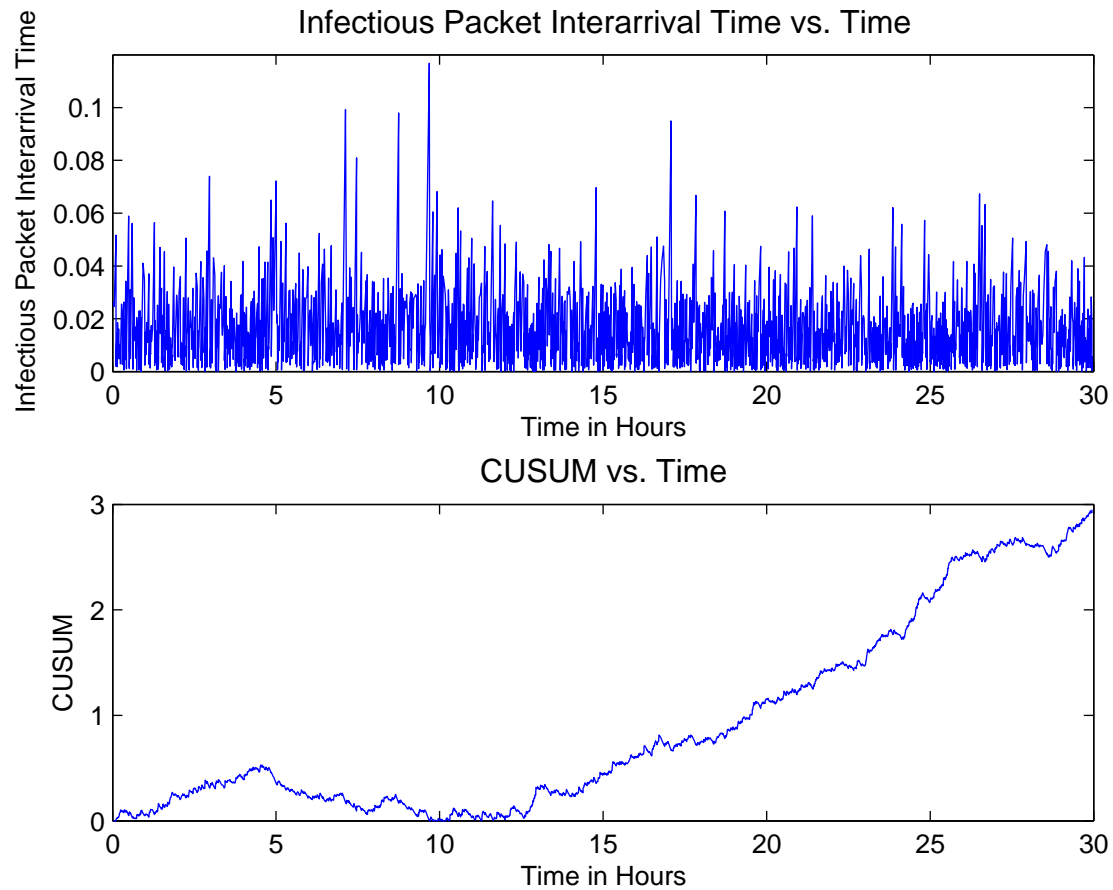


RCS Worm Sequential Hypothesis Testing

- H_0 : Background noise only.
 - Exponentially distributed background scanning on local network.
- H_1 : Background noise with worm scanning noise.
- Assume for now that have perfect information.
 - Worm parameters.
 - Worms start time.
 - Background noise parameters.

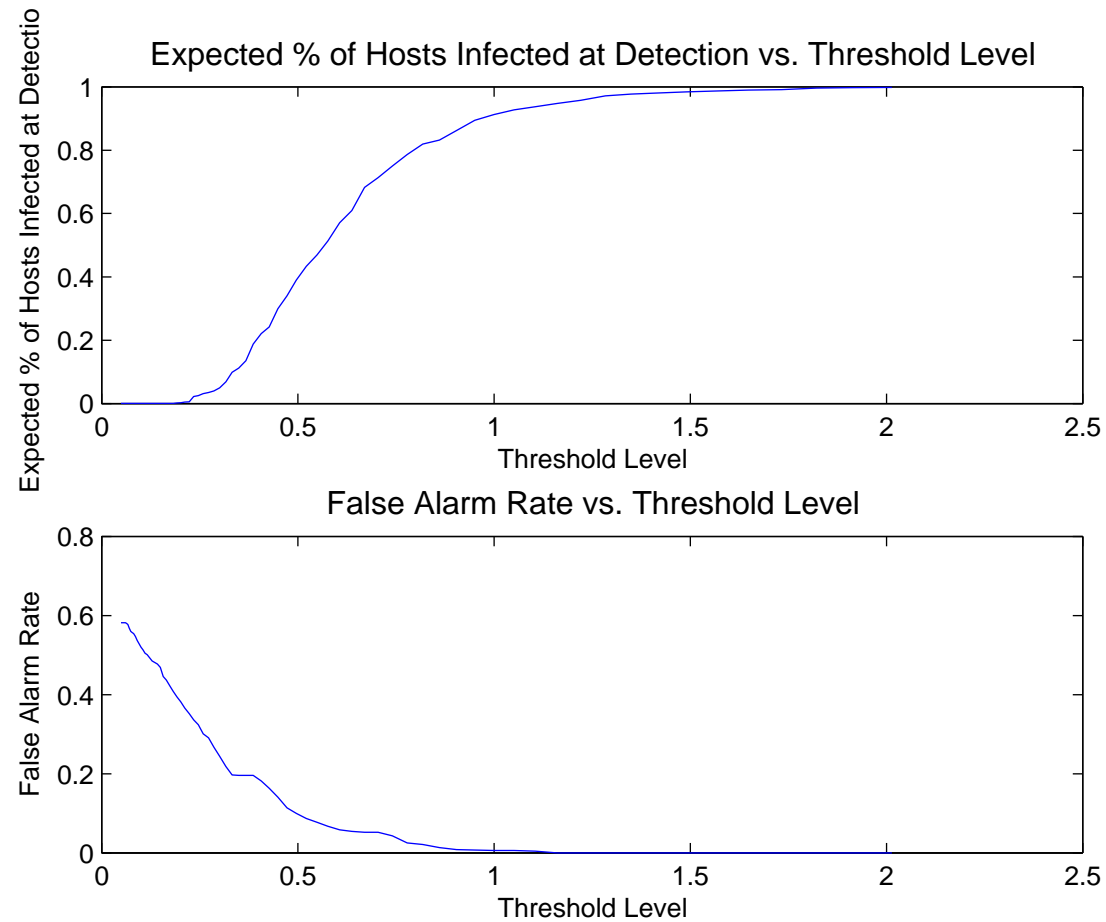


CodeRed Simulated SPRT Sample Run



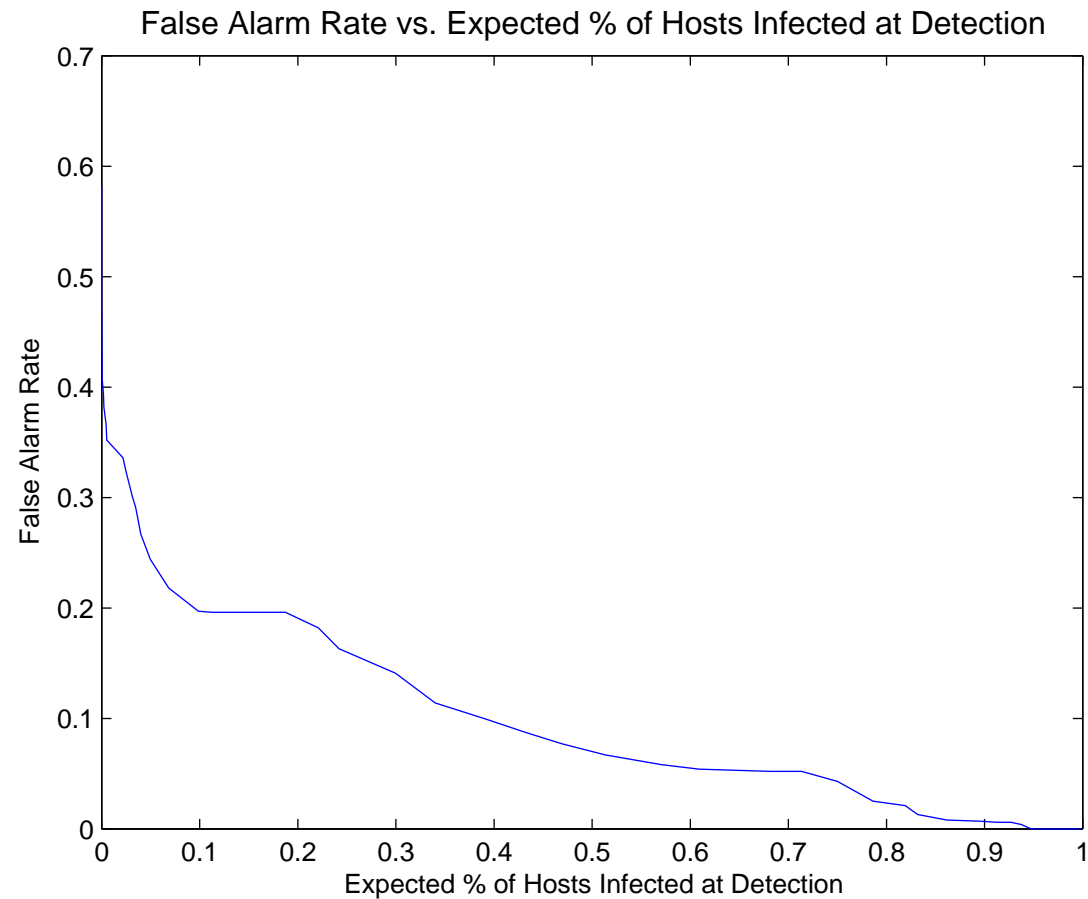


CodeRed ARL vs. α_0 Tradeoff



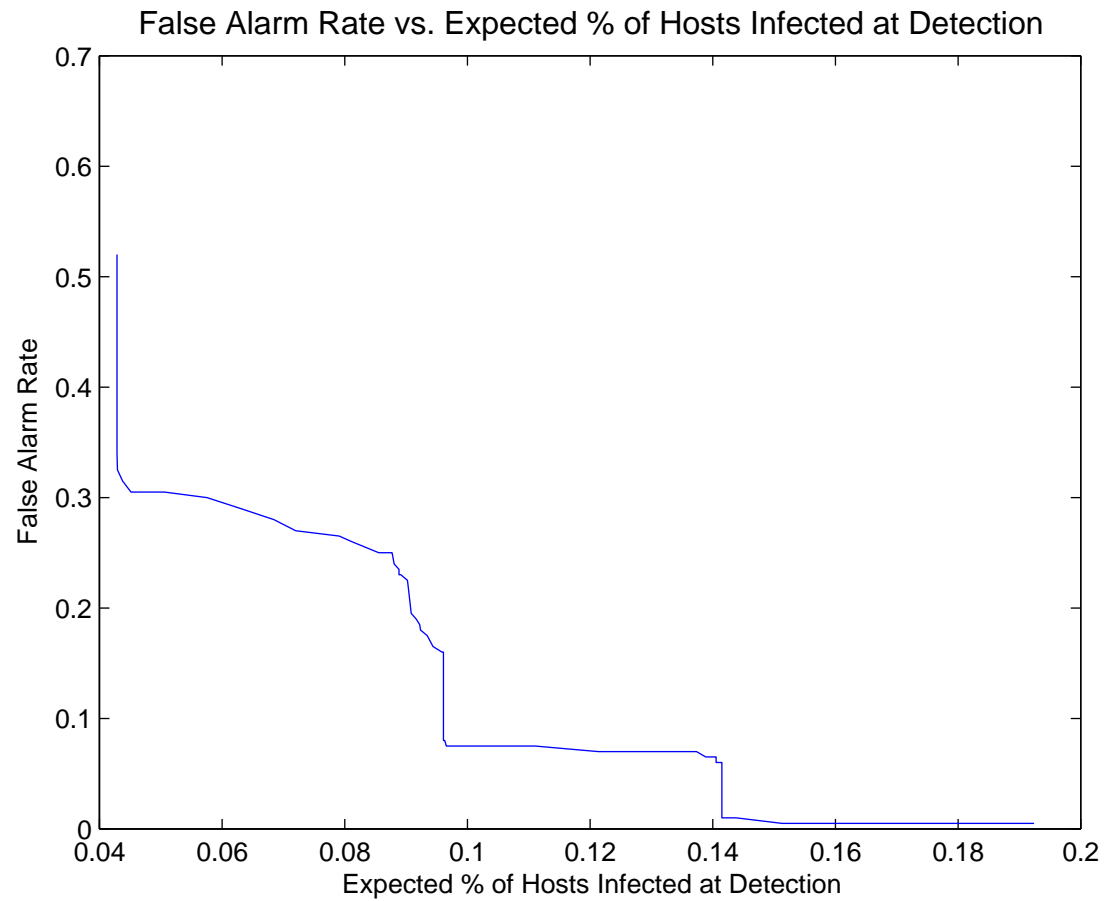


CodeRed ARL vs. α_0 Tradeoff





Slammer ARL vs. α_0 Tradeoff





Performance Limitations

- Graphs show some fundamental limits of detection methods.
- Aggressive worms like Slammer easier to detect faster.
- Analysis under idealized conditions.
- More "realism" makes detection more difficult.
 - Self-throttling of worm propagation?
 - Unknown parameters?



Conclusions

- Model of RCS worm.
 - Markov jump model.
- Analysis of detection limitations.
- Boundaries that cannot be crossed.



Conclusions

Thank you!

Publications are available.

krohloff@bbn.com

Any questions?