Scanning Activity Seen @ LBNL

Scan Activity Seen At LBL
Scans Per Scanner

Hosts Scanned Per Scanner
Ports Scanned Per Scanner

Scanning Speed
(a) Agobot Sources: UW I

(b) Agobot Sources: UW II
# Failed Conn’s Not Enough Info

![Graph showing CDF of # of remote hosts vs # of inactive local hosts for LBL.]

(a) LBL

---

Failure Ratio Much More Distinctive

![Graph showing CDF of # of remote hosts vs % of inactive local hosts.]

---

remainder

known_bad
Real-Time Detection

\[ \Lambda(Y) = \frac{\Pr[Y|H_1]}{\Pr[Y|H_0]} = \prod_{i=1}^{n} \frac{\Pr[Y_i|H_1]}{\Pr[Y_i|H_0]} \]

\[ \eta_1 \leftarrow \frac{\beta}{\alpha} \quad \eta_0 \leftarrow \frac{1-\beta}{1-\alpha} \]

- Event \( Y_n \)
- Update \( Y = (Y_1, \ldots, Y_n) \) and \( \Lambda(Y) \)
- \( \Lambda(Y) \geq \eta_n \):
  - Yes: Output \( H_1 \) (scanner)
  - No: \( \Lambda(Y) \leq \eta_0 \):
    - Yes: Output \( H_0 \) (benign)
    - No: Continue with more observations

Expected Time Until Decision

\[
E[N|H_0] = \frac{\alpha \ln \frac{\beta}{\alpha} + (1 - \alpha) \ln \frac{1-\beta}{1-\alpha}}{\theta_0 \ln \frac{\theta_1}{\theta_0} + (1 - \theta_0) \ln \frac{1-\theta_1}{1-\theta_0}}.
\]

\[
E[N|H_1] = \frac{\beta \ln \frac{\beta}{\alpha} + (1 - \beta) \ln \frac{1-\beta}{1-\alpha}}{\theta_1 \ln \frac{\theta_1}{\theta_0} + (1 - \theta_1) \ln \frac{1-\theta_1}{1-\theta_0}}.
\]
RB-SHT: Rate-Based Detection

- FCC’s interarrival times follow exponential dist. with mean \( \frac{1}{\lambda_1} \) (scanner) or \( \frac{1}{\lambda_0} \) (benign host). \( \frac{1}{\lambda_1} < \frac{1}{\lambda_0} \)

- \( T_n \): elapsed time until \( n \) FCC arrivals follows \( n \)-Erlang distribution

\[
\Lambda(n, T_n) = \frac{f_n(T_n | H_{scanning})}{f_n(T_n | H_{benign})} = \left( \frac{\lambda_1}{\lambda_0} \right)^n \exp\left(-\frac{\lambda_1}{\lambda_0} T_n \right)
\]