Securing Internet Coordinate Embedding Systems

Mohamed Ali Kaafar, INRIA, Fr
Laurent Mathy, Lancaster University, U.K
Chadi Barakat, INRIA, Fr
Kavé Salamatian, LIP6, Fr
Thierry Turletti & Walid Dabbous, INRIA, Fr

Introduction

- ICS (Internet Coordinates Systems):
  - Embed RTTs into geometric spaces

Decentralized Approaches
- Landmark-based Approaches (e.g. NPS)
- P2P Approaches (e.g. Vivaldi)

8/28/2007
ICS: Applications

Accurate, Scalable, robust
Long convergence time
Deployed as an “Always-On and Large scale Service”

Insider Attacks

- Passively: not cooperating, falsifying coordinates
- Actively: delaying probes
- Or Both!

Securing ICS is crucial for their deployment
Rationale

- ICS are dynamics
  - Coordinates keep changing

- Model of Normal behavior
  - Need a “Clean” system
    - Allow Abnormal behavior Detection
      - Comparing Model Predictions with observations (claims)

Clean System?
The Surveyors infrastructure

- Small Subset Trusted of nodes
  - EXCLUSIVELY positioned using mutual measurements
    - No malicious Activity
  - Involved in the ICS
    - USED by other nodes for positioning in the usual way
- Idea: Local behavior of a Surveyor is similar to that of other nearby nodes
Nodes Behavior Model

- Measured Relative Error
  \[ D_n = \left| \frac{predicted - measured}{measured} \right| \]

- Model
  \[ D_n = \Delta_n + U_n \]
  \[ \Delta_{n+1} = \beta \Delta_n + W_n \]

Many sources contribute to errors
- Assumption: \( U_n \) and \( W_n \) follow Gaussian Distribution
- Empirically Validated

Kalman Filter

- Separating nominal signal from noisy measurement
- Predicts the relative error
  \[ \hat{\Delta}_{n|n-1} \]
- Characterizes the innovation process
  \[ \eta_i = D_i - \hat{\Delta}_{i|i-1} \]

Abnormality = Significant Deviation of the values of \( \eta_i \)
Strategy

Step 1) Calibrate Model

Step 2) Provide Nodes with Model parameters

Step 3) Node uses the Model parameters to run its own filter

Abnormal Behavior Detection

- Simple Hypothesis Testing
  - $H_0$: the node has a Normal behavior
  - FIND the threshold value $t_n$ such that
    \[
    P\left(|D_n - \hat{\Delta}_n|_{n-1} \geq t_n \mid II_0\right) = \alpha.
    \]
  - $\alpha$ “significance-level”
Abnormal Behavior Detection(2)

- If Observed deviation exceeds $t_n$:
  - Hypothesis is rejected
    - Node is flagged as abnormal (potentially suspicious),
    - Embedding step is aborted, and
    - Measured relative error is discarded

Validation

- Simulations
  - King dataset
- PlanetLab
  - 280 nodes
- Vivaldi and NPS
  - Similar Results
Validation (2): Model Accuracy

- Self-calibration of the Kalman filter
  - at every node
  - in a clean system

Representativeness of Surveyors

- Number of Surveyors Needed?
- Randomly-chosen surveyors: 8% of population
  - This is a conservative upper bound
- 1% with a simple K-means deployment
- Optimal is still open research issue
Which Surveyor is representative?

- Closer is better
- Close enough is good enough: No need for THE closest.

Evaluation

- ROC Curves
- Tradeoff FPR/TPR
- Higher Curves are better!
- Detection: Excellent up to 20% of malicious nodes
- Still performs well, under heavy attack, up to 30% of malicious Nodes

\[ \alpha = 5\% \]
Embedding System Performance

- Practically immune to the attack $\alpha = 5\%$

Conclusions

- General Detection Method
  - Decentralized
  - Independent from the dimensions and the embedding protocol
- Very efficient although
  - No trust propagation among normal nodes
Thanks!

Questions?

Using Surveyors for positioning other nodes
Future Works

- Don’t Forget
  - Still need to Secure the distance estimate phase
  - Blatantly lie about coordinates when requested
  - Certified Coordinates
Kalman Filter equations

- 2 Steps: Prediction and Update
- Prediction Step: \( \hat{\Delta}_{i-1}^\beta = \beta \hat{\Delta}_{i-1} + W \)

Its a posteriori error variance is:
\[
P_{i|i-1} = \beta^2 P_{i-1|i-1} + W_N.
\]

- Update Step, integrates the observed \( D_i \):
\[
\hat{\Delta}_{i|i} = \hat{\Delta}_{i|i-1} + K_i (D_i - \hat{\Delta}_{i|i-1})
\]
\[
K_i = \frac{P_{i|i-1}}{P_{i|i-1} + W_U}
\]
\[
P_{i|i} = \frac{P_{i|i-1}}{P_{i|i-1} + W_U} + W_U P_{i|i-1}
\]

The vast majority of estimations are excellent
MALICIOUS BEHAVIOR DETECTION

- $H_0$: The hypothesis that the peer node has a normal behavior (honest)
- PB: FIND the threshold value $t_n$ such that
  \[ P\left( \left| D_n - \hat{\Delta}_n \right| / \tau \geq t_n \mid I(0) \right) = \alpha \]

- We can demonstrate that:
  \[ t_n - \sqrt{\nu_n} Q^{-1}(\alpha/2) \]
  where $Q(x) = 1 - \Phi(x)$

  $\Phi(x)$ CDF of $N(0,1)$, and $\alpha$ “significance-level”
State ($t_0$) \rightarrow \text{Prediction} (t_1) \rightarrow \text{Difference} \rightarrow \text{State} (t_1)

Large Difference \rightarrow \text{Anomaly!!}

\textbf{Surveyor nodes}
Strategy

Surveyors          Nodes

- Inter-Surveyors measurements
- Calibrate Filters using Clean measurements
- Provide nodes with filter parameters

- Select the closest Surveyor
- Run their filter with the provided parameters
- Filter-out abnormality

Kalman Filter – KF

- System state: unknown system parameters
- Measurement
- KF
- System: our knowledge of the system
Attacks exploiting cooperation

- Means
  - Passively: not cooperating, falsifying coordinates
  - Actively: delaying probes

- Attacks
  - Disorder (DoS)
  - Isolation
  - Repulsion (Free Riding)
  - Collusion

Vivaldi: Main Algorithm

\[ x_i = x_i + \delta \cdot (\text{rtt} - \| x_i - x_j \|) \cdot u (x_i - x_j) \]
Which Surveyor?

Isolation attack
Linear state space model

\[ \Delta_{n+1} = \beta \Delta_n + W_n \]
\[ D_n = \Delta_n + U_n \]

- Obtain relative error predictions from this model.

EM (Expectation Maximization) Method

- [http://www.gatsby.ucl.ac.uk/~zoubin/software.html](http://www.gatsby.ucl.ac.uk/~zoubin/software.html)
Using the closest surveyor’s parameters