Mobility-Pattern-Based Anomaly Detection Algorithm in Mobile Networks

Chaoli Cai and Ajay Gupta

Wireless Sensornet (WiSe) Laboratory
Department of Computer Science
Western Michigan University
Kalamazoo, MI, 49008, USA

Outline – MPB Anomaly Detection

- Motivation
- Problem Description
- Previous Work
- Proposed Algorithm
- Simulation Results
- Contributions
- Future Work

Motivation

- Characteristics of mobile networks
  - Mobile nodes have no fixed infrastructure
  - Arbitrary node movement
  - Lack of centralized control
- Prevention vs detection
- Misuse detection vs anomaly detection

Anomaly Detection
Problem Description

Case 1: Suppose R1 is the normal route, are other 3 routes normal or anomaly?

Case 2: Suppose R3 is the normal route. Based on the obvious deviations, it's still easy to classify R1 and R2 as anomalies, but what about R4?

2.1 During the training process, should we treat R3 and R4 as 2 distinct routes or integrate them as 1 route?

2.2 During the testing process, how can we classify R4 as the normal route or as an anomaly?

Previous Work

• Zhang et al, MobiCom 2000, wireless IDS framework.
• Cai et al, GlobeCom 2006, Mahalanobis distance is used for similarity of mobility patterns

Proposed Algorithm

- Mine data smoothly during the training process and distinguish the subtle difference during the testing process.

Start

Data Collection

Data Mining & Fuzzy Logic

Testing & Training

Normal Mobility Profile

Generate Alert

Distance > Threshold?

NO

YES

Stop

Comprehensive Algorithm Pattern Mined

Testing or Training?
Proposed Algorithm (Cont’d)

Training process:
- If (MinACD>C_k), add (k+1)^th cluster associated with its mobility pattern string P_1, P_2, ..., P_n.
- else, select the cluster which has the minimum MinACD, say C.
  - compute a total of m distances of P from the m pattern strings of C
  - compute m average pattern strings distance of P from the n hops of C
  - determine the minimum average pattern strings distance (MinAPSD)
  - If (MinAPSD>P_k), add (m+1)^th pattern string, that is, P_{m+1} to C.

Overall overhead of the MPB algorithm is O(N+2k+3m).

M: # of points, k: # of destination clusters, m: # of pattern strings, n: # of hops

Simulation Results

- Achieve good performance in terms of detection rate (≥90%) and false alarm rate (≤6%) for all the ranges of velocity.
- Threshold (P_{thr} = 0.6) can be efficiently determined for a given C_{thr}.
Contributions

- An efficient MPB anomaly detection Algorithm in mobile networks
- Easily generate and maintain the normal profile
- Handle close points and approximate mobility patterns
- Achieve good performance without strong assumptions such as restricted velocity ranges
- Efficiently determine the design parameter - threshold

Future Work

- Generalize mobility profile with \( n \) levels (sequences)
- Compare performance using other Data Mining techniques
- Maximum fusion vs Average fusion
- Detect individual significant deviation along with the sequence increases
- Distributed MPB
- Assign probability value to represent for the frequency of the pattern string collected from the different node

Thank you

Questions?

http://www.cs.wmich.edu/wise