Visualizing Anomalies in Sensor Networks

Qi Liao*, Lei Shi†, Yuan He‡, Rui Li$, Zhong Su†, Aaron Striegel^Δ, Yunhao Liu^x

* Central Michigan University (qi.liao@cmich.edu), † IBM Research China ({shllsh,suzhong}@cn.ibm.com), ‡ Hong Kong University of Science and Technology (heyuan@cse.ust.hk)
$ Xi'an Jiao Tong University (rli@mail.xjtu.edu.cn), Δ University of Notre Dame (striegel@nd.edu), X Tsinghua University & HKUST (liu@cse.ust.hk)

ACM SIGCOMM, August 15 – 19, 2011, Toronto, Canada.

Problem

Sensor networks are important components of future internet, i.e., Internet of Things (IoT). However, diagnosing a large-scale sensor network is tremendously challenging due to

- Resource and bandwidth constraints on sensor nodes;
- Spatiotemporally dynamic network behaviors in hostile environments;
- Lack of accurate models to understand silent faults.

Traditional network management, monitoring and analytic methods are less effective. The unpredictable, spatiotemporally dynamic, silent failures and anomalies cannot be readily detected quickly, and most importantly the underlying causes for these anomalies are nearly impossible to be found.

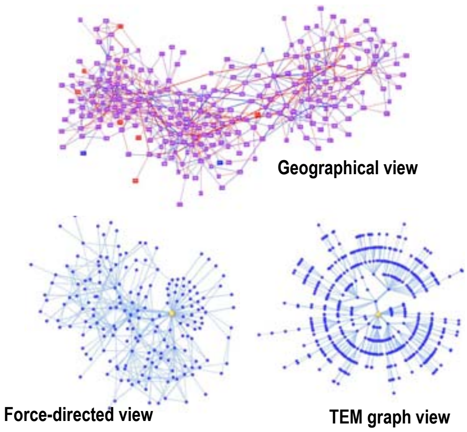
Solution

Rather than deriving hypothesis directly with the inference model, visual analytics solution focuses on collecting better evidence for the human to improve the decision-making process.

We developed **Sensor Anomaly Visualization Engine** (**SAVE**), which combines both anomaly detection algorithms and visual exploration process that can bring in experts' domain knowledge.

Several novel visualizations have been proposed to present and correlate time-varying high-dimensional sensor data facets – routing topology, sensor networking status and physical sensor readings.

Topologies Visualization



GreenOrbs

The latest deployment in Zhejiang Forestry University includes 500 physical sensor nodes that have been in continuous operation since year 2010.

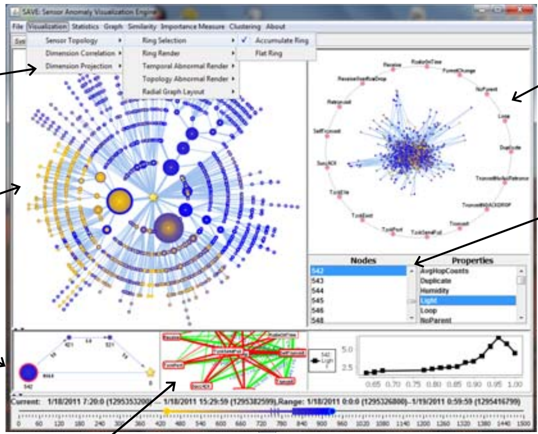
Data Collection

Every 10 minutes, each sensor node reports the following:

- Sensor readings
- Routing paths
- Wireless link status/quality
- Network statistics (counters)

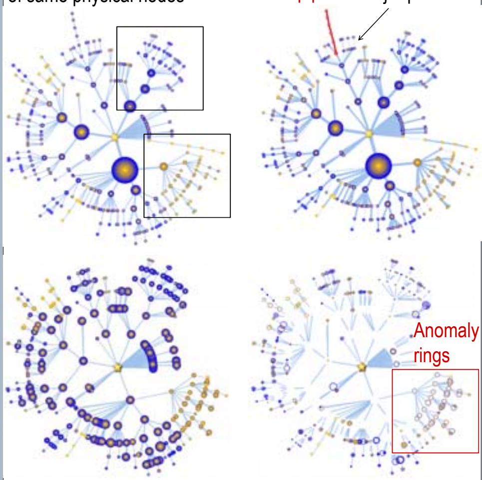
Sensor Anomaly Visualization Engine

SAVE: Analyzing spatiotemporal anomalies for wireless sensor networks.



Temporal Expansion Model (TEM)

- Key feature of sensor networks: all nodes send packets to the sink node.
- Split each physical sensor node into multiple logical nodes according to its separate routing paths to the sink.
- Routing graphs → directed trees (better for visualization and navigation)
- *Temporal rings*: edge weighted → node weighted
  - Size = number of sent / relayed packets
  - Color = time (orange = earliest, blue = latest)
  - Line: solid = ascending time, dash = anomalies



Evaluations

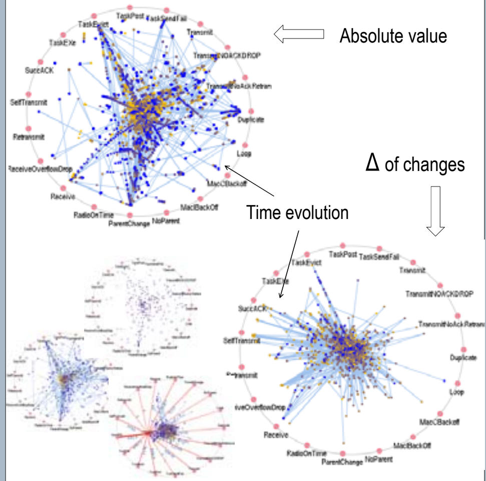
SAVE has been deployed into a real-world large-scale Wireless Sensor Network (WSN) system – GreenOrbs. Based on long-term iterative studies with domain experts, the visual analytic tool can effectively identify silent failures and provide critical guidance for the WSN operators and researchers to drill-down to the root causes of these anomalies. In most cases, SAVE indeed saves a significant amount of time and effort for the operators in completing their tasks.

Correlation-based Radial Projection

High-dimensional sensor nodes: outlier and dynamics

- Spatial anomalies: How does each sensor node differ from other nodes in terms of its dimensions?
- Temporal anomalies: How does each sensor node evolve from its previous status?

Insight: Dimension anchors should be arranged by their correlation scores.



Correlation Graphs

- Correlation scores between property values of sensor nodes using Pearson's product-moment coefficient:

$$\frac{|p_1|\cdot\sum_{i=1}^{|p_1|} p_{1_i}\cdot p_{2_i} - \sum_{i=1}^{|p_1|} p_{1_i}\cdot\sum_{i=1}^{|p_2|} p_{2_i}}{\sqrt{|p_1|\cdot\sum_{i=1}^{|p_1|} p_1^2 - \left(\sum_{i=1}^{|p_1|} p_1\right)^2}\cdot\sqrt{|p_2|\cdot\sum_{i=1}^{|p_2|} p_2^2 - \left(\sum_{i=1}^{|p_2|} p_2\right)^2}}$$

- Edge lengths = inverse of correlation scores
- +Δ = green, +Δ = red