Spatiotemporal Visual Analysis of Sensor Networks in the Wild

Qi Liao*, Lei Shi†, Yuan He‡, Zhong Su†, and Aaron Striegel*

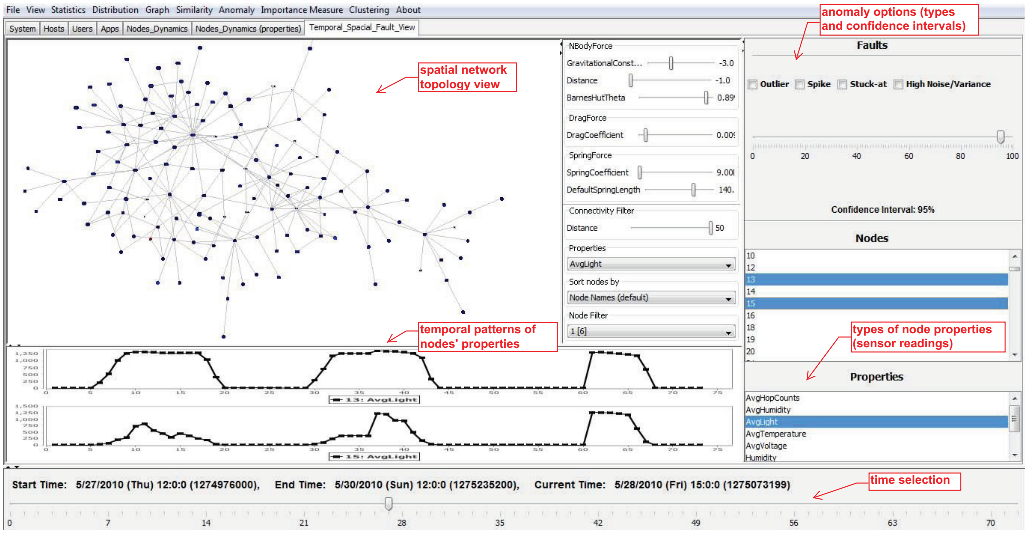

Figure 1: Visualization tool for analyzing spatiotemporal anomalies for network data. As an example, the average sensor readings of light over hourly-generated snapshot graphs are color-mapped into network topologies while multiple selection of nodes is enabled to observe temporal patterns.

Abstract

Diagnosing and managing a large-scale sensor network in the wild is a crucial but challenging task, due to the resource and bandwidth constraints on sensor nodes, the highly dynamic spatiotemporal network behaviors, and the lack of accurate models to understand such behaviors in a hostile environment. In this paper, we present a visualization tool that allows quick and effective navigation of the sensor data collected in real time. By tightly combining a spatiotemporal visualization framework and a novel construction of differential contour maps with existing sensor network failure models, our tool helps analyze and diagnose the sensor network in a way easy for ordinary users while customizable for domain experts.

1 Introduction

As the recent surge of sensor technology, networks have expanded from the traditional computer-based ones, e.g. the internet, to the vast object-based networks in the real world, and even to the science-fiction-like human body networks. In this paper, we are focusing on the diagnosis of sensor networks. A sensor network consists of a number of sensor nodes, each of which has the capability of sensing, computing, and wireless communication. We mainly focus on sensor networks in the wild, which are usually powered by batteries and resource-constrained. For example, in a sensor network deployed in the forest for carbon emission estimation and forest surveillance, we need to know in real time the portion of abnormal sensors to guarantee an accurate overall measurement. Generally, this kind of network is composed of a large number of resource-limited sensor nodes in the frequently-changing and unpredictable environments, vulnerable to both system failures and environmental abnormalities. As a direct result, sensor data are highly dynamic in both spatial and temporal dimensions, making singular model driven analytics hard to locate correlational patterns. These natures make the diagnostic tasks extremely difficult.

*Computer Science and Engineering, University of Notre Dame, USA. E-mail: {qliao,striegel}@nd.edu.

†IBM Research - China. E-mail:{shllsh,suzhong}@cn.ibm.com.

‡Computer Science and Engineering, Hong Kong University of Science and Technology, HK. E-mail:heyuan@cse.ust.hk

Traditionally, the diagnosis methods have been based on the network itself and its graph properties, such as degree distributions, subgraph isomorphism and graph edit distance [2]. While there are techniques dealing with time series data visualizations, e.g. the GrowthRingMaps [1], very few of them are designed specially for the diagnosis of performance issues in sensor networks and for the easy exploration of root causes for which there may be no prior established knowledge. In this work, we propose a visualization framework that incorporates 1) both the node properties (e.g. the sensor readings) and the physical/logical network topology; 2)correlation between both the spatial and temporal changes of the sensor network.

Our contributions are twofold. we develop an visual analytic tool (Figure 1) that allows quick and effective exploration of massive sensor datasets. The tool extends the data and anomaly analysis algorithms visually to allow network operators and analysts to interact with the data images and gain insights from a domain expert's perspective. Second, we propose a novel differential visualization method of topology-aware contour maps. Actual routing networks are encoded into contour maps using topological measurement such as the average hop counts to the sink. The proposed 2nd-order differential visualization of contour maps leverages our observations that the values of contour lines change at different rates. These changing rates, or gradients in short, can be used to construct a new contour map.

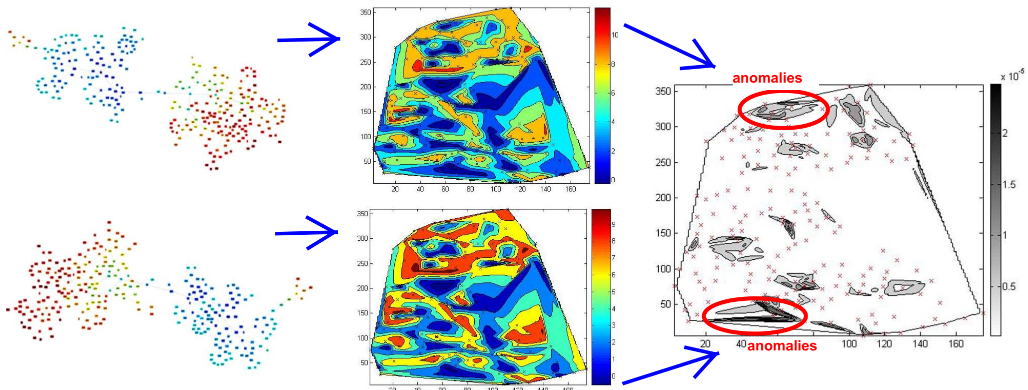

Figure 2: Differential visualization of topology-aware contour maps.

2 DATASET

The dataset we experiment with our visual analytic tool is collected in GreenOrbs project [3], which has more than 300 sensor nodes deployed in the wild and is among the largest real-life sensor network systems worldwide. Two types of data are measured in real time and aggregated at the network sink for investigation purpose: i) Network data, including the routing data, link state data, neighborhood list and the various counters at the sensor node indicating the network performance; ii) Sensor data, including the measured temperature, humidity, light, voltage, etc. From the raw sensor network data, we calculate the average hop count of each node going to the sink and construct their hop-count contour maps.

3 VISUALIZATION

In this section, we focus on a few examples illustrating the spatiotemporal visualization using interactive data exploration and anomalies detection modules as well as a novel differential visualization using topology-aware contour maps.

3.1 Interactive spatiotemporal visualization

Figure 1 shows a screenshot of the visual exploration tool that is under active development. After reading in data sent by the sensor networks in real time, it automatically generates snapshot graphs and visualizes them in the main window. Each node represents the actual sensor node and directed edges among nodes indicate the actual routing of packets from the source to the sink. The width of the edges reflects the magnitude of the connections. An investigator can drag the slider bar on the bottom to view snapshot network graphs at any specific time granularity with a customizable time history window.

Spatial anomalies can be visualized in a network topology view through standard color/shape/size coding on the nodes/edges based on the selection of node properties, e.g., average values of sensor readings such as temperature, humidity, light, voltage, etc. Nodes coded with a spectrum of colors are easy to visualize the spatial distribution of node property values across the entire network. This can be especially useful to detect potential correlations of network anomalies. The correlation among spatial anomaly can provide useful context information that can potentially reduce the false positives. For example, rapid changes of temperatures among the nodes in close vicinity would suggest events perhaps fire in a part of the forest, rather than merely due to sensor anomaly. Another example would be if many of the blue nodes (low voltages) have a large number of in-degrees, this could be evidence that these core nodes serve as the hubs that route many packets from sources to the sink node. Therefore, the rapid decrease in voltage of these nodes may not necessarily be anomalies.

Temporal anomalies are also very important because they are the key to capture the changing events (normal or abnormal) happening in the network. Depending on the selection of nodes and their properties, statistical charts representing the time evolution of sensor properties give investigators a quick view of the trend. As a result, any potential temporal anomalies can be easily picked up. For example, in Figure 1, one node (top) shows clear diurnal patterns of light readings over the three-day period while another selected node (bottom) also matches the overall trend of the first node. Misalignment could suggest temporal anomalies in this example. The detected abnormality patterns will be mapped into the network topology graphs to understand the spatiotemporal correlation of the anomalies (e.g., Type-I anomalies are observed exclusively in Cluster A).

3.2 Spatiotemporal differential visualization using contour maps

Understanding the *difference* among snapshot graphs is the crucial part of visualization for anomaly analysis. While comparing the network graphs at the level of nodes and edges gives the maximum of details, a novel topology-aware contour-map visualization is developed to reflect the balance of granularity and complexity. The differential visualization of topology-aware contour maps (Figure 2) shows the concept of generating such differential views of contour maps. Firstly, the spatial network topology is encoded into a contour map by making use of the average number of hops to the sink nodes derived from the actual routing paths. Secondly, we develop a novel differential view to compare dynamically temporal contour maps (e.g., compare the contour maps at time t and t'). The differences of any pair of gradient vectors (the direction and length) on the contour maps are then used as the values to construct the new differential contour map. The net result is that any potential anomalies are ready to be spotted, as highlighted in Figure 2. Based on the differential visualization, we can not only discover temporal patterns, but also gain insights into the correlations of the *content* (e.g., sensor readings such as temperature, humidity, light, voltage, etc) and the *context* (e.g., traffic loads, routing structures, etc).

REFERENCES

[1] P. Bak, F. Mansmann, H. janetzko, and D. A. Keim. Spatiotemporal analysis of sensor logs using growth ring maps. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):913–920, November/December 2009.

[2] H. Bunke, P. J. Dickinson, M. Kraetzl, and W. D. Wallis. *A Graph-Theoretic Approach to Enterprise Network Dynamics (Progress in Computer Science and Applied Logic (PCS))*, volume 24. Birkhäuser Boston, 2007.

[3] Y. Liu, Y. He, M. Li, J. Wang, *et al.* Does wireless sensor network scale? a measurement study on greenorbs. *To appear in IEEE INFOCOM*, April 10-15 2011.