# Understanding Routing Dynamics in a Large-scale Wireless Sensor Network

Tong Zhu[1,2], Wei Dong[3], Yuan He[2], Qiang Ma[2], Lufeng Mo[4], Yunhao Liu[2]

[1]Department of Computer Science and Engineering, HKUST
[2]School of Software and TNLIST, Tsinghua University
[3]Zhejiang Key Laboratory of Service Robot, College of Computer Science, Zhejiang University
[4]School of Information Engineering, Zhejiang Agriculture & Forest University

*Abstract*—**Routing dynamics are intrinsic characteristics of operational wireless sensor networks (WSNs). We present the measurement and analysis results for routing dynamics in a large-scale WSN. We seek to answer several fundamental questions: How dynamically are current routing protocols performing? What causes routing dynamics? What is the impact of routing dynamics? Answers to the above questions are critical to understanding the interactions among multiple network elements, evaluating protocol design strategies, and improving system performances. However, measurements in large-scale WSNs are challenging due to the lack of dedicated log information (be analogous to configuration files, syslog messages used in Internet). We propose an approach to identify the routing dynamics based on limited information and correlate them with system events to find out the root causes. The key findings of our study include: 1) parent change events mainly affect local nodes, i.e. they do not cause routing instability on far-away nodes; 2) environment dynamics and routing loops have large impact on routing; 3) small portion of parent changes might not be necessary, while a large portion of parent changes are effective in improving network performance.**

## I. INTRODUCTION

Wireless sensor networks (WSNs) are deployed for various applications, e.g., event detection [1], [2], target tracking [3] and environment monitoring [4], [5]. Many WSNs interact closely with the environment. Therefore, how a WSN system operates is intuitively affected by the environmental dynamics, including weather condition changes, urban traffic flows and signal interference from other applications like WiFi. These unknown dynamics impose serious challenges on predicting the quality of wireless links for the routing protocols and subsequently lead to the fluctuation of system performances.

To adapt to environment dynamics, dynamic routing protocols are widely adopted in wireless ad-hoc networks [6], [7], [8]. In those protocols, the routing path is regularly updated to maintain consistent network performance. As dynamic routing protocols over large-scale wireless sensor networks become popular and they are significantly affected by dynamic environments, the expectation of good routing performance has dramatically increased: large fluctuation of system performances caused by environmental dynamics is not acceptable. To meet this challenge, protocol designers have made

tremendous efforts to improve routing performance. Most of those efforts are based on measurement-driven insights from small real systems [9], [10]. There have been several studies with controlled lab settings to evaluate how wireless links vary over time and space [11], [12], [13]. Very few work on understanding how the dynamic environment affects the routing performance in the wild and at scale.

Currently, lots of large-scale outdoor systems have attracted people's attention and provide real services. In this aspect, it becomes important to systematically understand routing dynamics in a large WSN. This knowledge can help designers to better focus on optimizing the performance metrics(e.g., packet delivery ratio and delay). Specifically, answers to the following fundamental questions are common interests:

- How dynamically are current routing protocols performing?
- What causes routing dynamics?
- What is the impact of routing dynamics?

However, understanding routing dynamics in large-scale WSNs is challenging. First, the deployment of a large scale WSN is non-trivial. We cannot be aware of locations with highly environment dynamics. In advance, given limited information, WSN operators cannot make precise judgements on the correctness and effectiveness of routing decisions in the network. Second, it proves difficult to estimate and verify the impact scope of a network event in large-scale WSNs. Third, better measurements require better equipments and workload. Nonetheless, local log systems and external monitoring are hardly achieved for resource-constrained WSNs. To address these issues, we propose an approach to characterize routing dynamics by correlating system events and seek to identify the underlying causes of routing dynamics.

Our goals are two-fold. Firstly, we seek to characterize network routing dynamic patterns in large-scale WSNs. Secondly, we want to understand the root causes and the impact of those dynamics to the design of future routing protocols in large-scale WSNs. Since our system is based on the Collection Tree Protocol (CTP) [6], the routing dynamics are essentially caused by the route selection, which refers to the dynamics of

IEEE computer society

parent change events in CTP. Therefore we will focus on the study and analysis of parent change behaviors in this work.

To better understand routing dynamics in large-scale WSNs, we study this problem from three perspectives:

- **Characterizing the spatial-temporal characteristics of routing dynamics.** Route selection is a key functionality of dynamic routing protocols. Characterizing routing dynamics can help to understand the interactions of multiple network elements. To this end, we extract path change events and analyze when, where, and under what kind of circumstances routing dynamics occur.
- **Revealing the root causes of routing dynamics.** Root causes are explored to understand the key factors that affect routing dynamics. However, extracted path change events are insufficient to disclose the underlying causes. Therefore, we also extract other system events such as packet loss events and routing loop events. Examining their spatial-temporal correlations, one can better understand their causal relationships.
- **Estimating the impact of routing dynamics.** In managing a WSN, operators need to focus on system performance metrics like packet reception ratio (PRR). Understanding the impact of routing dynamics on system performance can evaluate the effectiveness of routing protocols. In this study, we analyze the effectiveness of CTP routing selection mechanism by comparing the PRR before and after making routing decisions.

We obtain several key observations from our study:

- **Parent change events mainly affect the local area.** A routing decision might affect local nodes while the probability of affecting far-away nodes is fairly low.
- **Variations in humidity affect the link qualities to a large extent.** Therefore, environment has a large impact on routing dynamics. Variations in humidity may cause huge packet losses and even no possibility to find an available route.
- **Routing loops and parent change events show strong correlation.** Routing loops have a close relationship with routing dynamics. The majority of loops can be recovered within a short duration while others need a much longer time to be recovered.
- **Most route selections are effective.** While a large portion of parent changes is effective in improving network performance, a small portion of parent changes might not be necessary. 30% of the parent changes might not be effective in improving network performance.

The rest of this paper is organized as follows. Section II presents the related works. System architecture and data sources are described in Section III. In Section IV we define and extract the parent change events. We characterize temporal and spatial distributions of parent change events in Section V. We correlate system events to identify the root causes of routing dynamics in Section VI. In Section VII, we estimate the effectiveness of current routing selection mechanism in CTP. Finally, our work is concluded in Section VIII.

## II. RELATED WORKS

Recently, the research community has made lots of efforts on deploying real-world sensor networks and conducting performance measurements of these systems. These studies show valuable observations and guidelines to the design and implementation of network protocols. However, most of those results are based on controlled testbeds or small networks, and thus fail to consider many practical and critical factors for large-scale sensor networks. In this section, we review studies on sensor network deployments and related measurement studies.

### A. Large-scale WSN deployments

Deploying sensor networks is always a multi-faceted job since during the pre-deployment test we cannot simulate all conditions which would happen in a real environment. Tolle *et al.* [14] conduct an experiment to monitor the microclimate of a redwood tree. Unfortunately more than a half of sensors cannot form a routing tree after the deployment, and nearly 15% of the remaining nodes die in one week by exhausting their battery power. Though this study gives valuable guidance to early sensor network design, the network scale is small and cannot reflect intrinsic characteristics in large-scale sensor networks. Viginet [15] uses a 200-node network to support long-term military covering of $100 * 100$ square meters. ExScal [16] is an attempt to deploy a sensor network at an extreme scale. The system consists of about 1000 sensor nodes and 200 backbone nodes, covering $1300 * 300$ square meters. SenseScope [17] is a real-world deployment on a rock glacier, consisting of about 100 sensor nodes. They provide many practical guidelines for large-scale WSN deployments, but few of them focus on routing dynamics.

### B. Sensor network measurements

Considering unpredictable performances in real environments, researchers design indoor controlled testbeds to conduct experiments for exploring more nuances. Testbeds help us to understand in-depth network behaviors. MoteLab developed by Harvard university provides a web interface for users to easily program sensor applications and conduct experiments remotely. Maheshwasi *et al.* [18] conclude that the physical interference model is most accurate comparing to other interference models in two testbeds. Although those testbeds bring benefits to explore specific aspects in sensor networks, testbeds can hardly simulate real environment.

In sensor network systems, measurement studies are limited by the system scale and constrained resources on sensor nodes. Werner-Allen *et al.* [9] use a wireless sensor network to monitor volcanic eruptions within a 19-day deployment and collect over 54 hours of continuous data which include at least 9 large explosions. They analyze the packet loss performance

Fig. 1. The overview of prototype system deployment



Fig. 2. Component graph of GreenOrbs implementation

and propose some hypotheses of the causes, e.g., equipment dropouts, weather conditions and temperature fluctuations. Due to lack of extra information, those hypotheses cannot be well validated by the data analysis.

Packet delivery performance is one of the most important indicator for wireless communications and thus attracts many research efforts. Zhao *et al.* [10] report packet delivery performance measurement on medium-sized sensor networks in three different environments. They obtain some interesting findings, e.g., gray areas of radio communications and a high percentage of asymmetric links. Srinivasan *et al.* [19] conduct an experiment to identify root causes of packet losses. They show that packet losses are highly correlated over short-time periods, but are independent over longer periods. They also show that though RSSI value change over time,it still can provide useful information for the link estimation if the receiver's noise floor is known.

We have conducted a measurement study of a large-scale network in the forest [20], which mainly focuses on the evaluation of overall system performance, especially the packet delivery ratio.

### C. Internet measurements

Traditional Internet measurements always rely on experienced operators and dedicated instrumentations broadly deployed across the network. Maropooulou *et al.* [21] use passive optical taps and high-speed packet capture hardware to study failure in the Spring backbone. To avoid significant capital and operational expense, Turner *et al.* [22] extract sufficient knowledge from router configuration files, syslog archives, and operational mailing list announcements. However, measurement on large-scale sensor networks is even more challenging due to complex wireless behaviors.

### III. NETWORK ARCHITECTURE AND DATA SETS

#### A. GreenOrbs system

Our research project "GreenOrbs" was started from April 2009. We want to build a long-term large-scale WSN system which aims to provide services to many forestry applications such as canopy estimation, fire risk prediction, etc. The system currently uses TelosB motes with msp430f1611 processor and CC2420 radio. We develop programs based on TinyOS 2.1.1 [23]. Each node periodically collects environment data including temperature, humidity, carbon dioxide concentration. Each node also records system status, and delivers three
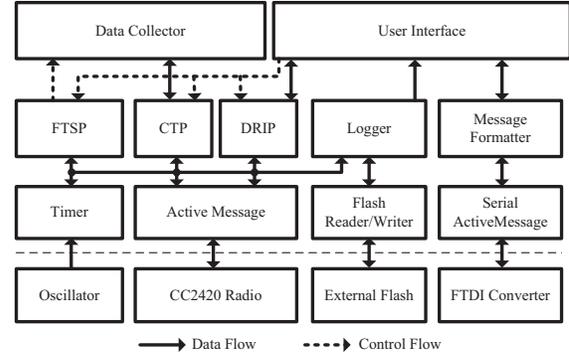
TABLE I
PACKET CONTENT

| C1 | Sensing Data | temperature,humidity light and voltage |
|----|--------------|------------------------|
|    | Path Data | parent node ID routing path length node ID along the path |
| C2 | Routing Table | neighbor ID neighbor's RSSI value link ETX value path ETX value |
| C3 | Statistics Counter | parent change counter loop counter drop_no_ack counter |

packets to the sink node every 10 minutes. Figure 1 is the deployment overview of our prototype system. The component graph of the software design is shown in Figure 2. Many protocols are used, like [6], [24].

In CTP protocol each node maintains a routing table which records the next hop information associated with the destination. Packets are delivered to the next hop as recorded in the routing table and the process repeats until the packets reach the sink. The route selection is based on the ETX metric which is a measure of the transmission cost. ETX estimation is updated by both the control plane traffic and the data plane traffic. CTP controls the beacon rate by increasing Trickle time when a node reboots or detects a loop. When the network is steady, it can also reduce unnecessary control plane traffic.

#### B. Data sets

We use a data set containing 343 nodes from December 19, 2010 for 10 days. The data includes three types of packets: "Sensing Packets" (C1 packet), "Network Status Packets" (C2 packet) and "Statistic Packets" (C3 packet). Table I shows the format of these three types of packets. In particular, the parent change counter records the accumulated number of parent changes, the loop counter records the accumulated number of detected loops and the drop_no_ack counter records the accumulated number of packet drops due to be exceeded the retransmission threshold (i.e., 30 in CTP).
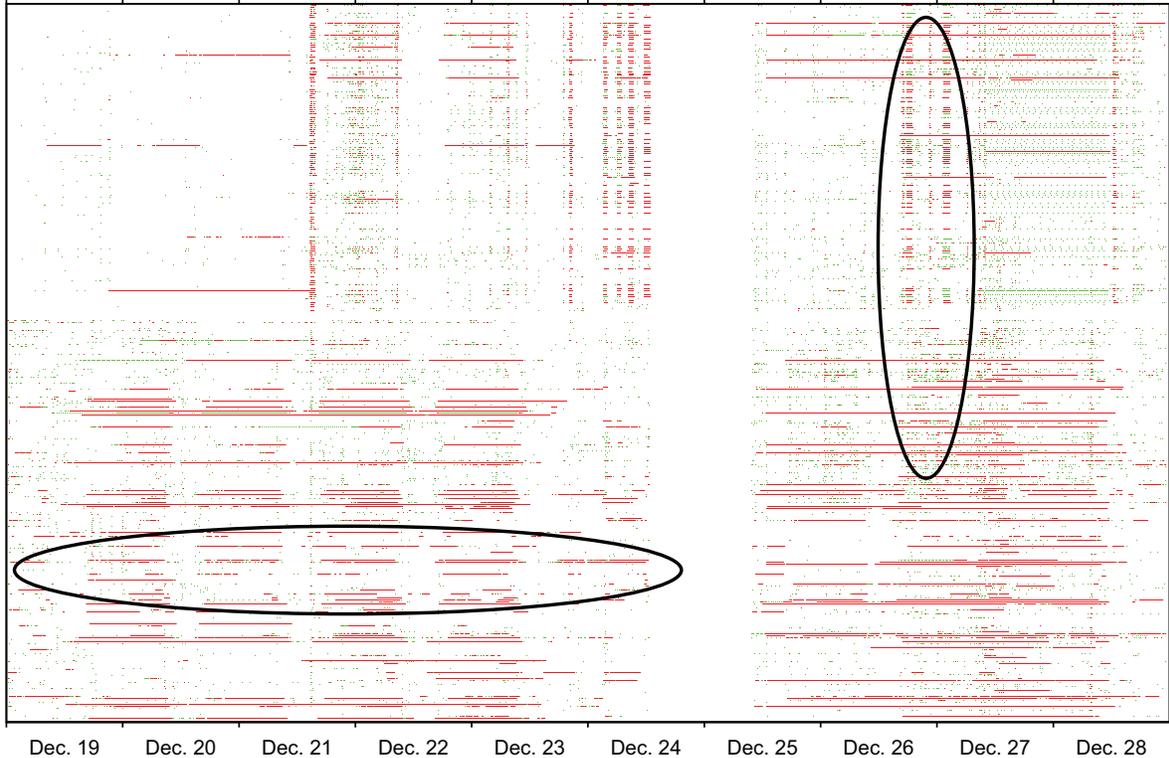
Fig. 3. Parent change at a glance. The y-axis denotes the node ID in ascending order. Each parent change event is represented by a line on the plot, located according to the start of the event. A red line indicates a parent change event with high frequency. A green line indicates a parent change event with low frequency.

## IV. TERMINOLOGIES

### A. Defining and identifying routing dynamics

In CTP protocol, each node dynamically chooses one parent node as its forwarding node, so routing dynamics are essentially caused by the parent selection. To study the routing dynamics, we first define parent change events.

We define a parent change event as a quadruple: $\langle$ $id$, $t$, $p_a$,$p_b$ $\rangle$, which indicates that node $id$ changes its parent from node $p_a$ to node $p_b$ at time $t$. Detecting all parent change events in the system, however, is challenging since limited external flash cannot support massive event logs. Therefore, we use a parent change counter to track parent change events and the accumulated counter value is periodically reported to the sink node. We define a detectable parent change event as $\langle$ $id$, $t_a$, $p_a$, $t_b$, $p_b$, $counter_{pc}$, $freq$ $\rangle$. It means that node $id$ has $counter_{pc}$ times of parent changes between time interval $t_a$ and $t_b$, as well as the parent of node $id$ is $p_a$ at $t_a$ and $p_b$ at $t_b$ respectively. The $freq$ value denotes the parent change event frequency which is the average number of parent changes every 10 minutes. For example, if the duration of the parent change event are 50 minutes and the $counter_{pc}$ records 30 parent changes in this interval, then the frequency of the parent change event is 6. Similarly, we define a tuple $\langle$ $id$, $t_a$, $t_b$, $counter_{loop}$ $\rangle$ as a loop event. It means

node $id$ encounters $counter_{loop}$ loops between $t_a$ and $t_b$. The definition of no_ack_drop events is a quadruple $\langle$ $id$, $t_a$, $t_b$, $counter_{loss}$ $\rangle$ as a no_ack_drop event, which means the node $id$ drops $counter_{loss}$ packets between time interval $t_a$ and $t_b$ due to no ack received from parent nodes after exceeding the retransmission threshold.

### B. Event extraction

17,644 parent change events are extracted from the data set. We find that the sink was down during 14:40 December 24, 2010 to 8:20 December 25, 2010. So when excluding the sink down period, the number of remaining parent change events is 17,281 with 201,734 parent changes.

## V. CHARACTERIZING ROUTING DYNAMICS

We ask several fundamental questions regarding the protocol strategy for routing dynamics. Especially, we consider:

- How often do parent changes occur? Where and when do they frequently?
- How long do parent changes last? Are routing links near to sink stable than others?

### A. Event history at a glance

Figure 3 shows the spatial and temporal distributions of parent change events. Each parent change event is represented

by a line on the plot, located according to the start of the event. A red line indicates a parent change event with high frequency. A green line indicates a parent change event with low frequency. We have two observations from Figure 3:

- **Horizontal banding**
  Some nodes experience periodic horizontal bandings indicates the parent change events exhibit the feature of periodicity. At these nodes, HPC events and LPC events are interleaved regularly. We will further look into such a phenomenon in section VI-B.
- **Vertical banding**
  We observe vertical bandings covering a subset of nodes. In this case, a subset of nodes experience parent change events simultaneously. This is mostly caused by routing loops. A further discussion will be in section VI-C.

### B. Aggregate statistics

To have a first look at the routing dynamics in our system, we investigate three aggregated statistics: the number of parent changes experienced by a node, the frequency of parent change events, and the lifetime for each routing link.

*1) Number of parent changes per node:* We present a global statistics of the number of parent changes for each sensor node. The overall statistics is shown in Figure 4, where the $y$-axis represents the number of parent changes and the $x$-axis represents the node IDs sorted according to the $y$-axis value. As shown in Figure 4, $80\%$ of nodes whose number of parent changes are fewer than 630. Routing links of most nodes are quite stable, while remaining ones exhibit frequent parent changes. Why do those nodes exhibit significantly different parent change behaviors? **Is hop count a key factor that causes the parent changes?** Figure 5 shows the correlation between the number of parent changes for each node and hop counts to the sink node. According to the theoretical collection tree based model, nodes near the sink should carry more traffic loads which may cause potential interferences. However, the number of parent changes reaches the peak at hop distance 4 and decreases towards to the sink. So we cannot simply conclude that the large number of parent changes is caused by the interference or heavy traffic.

*2) Frequency of parent changes:* Figure 6 shows the CDF of the frequency of parent change events. The $x$-axis indicates the frequency. Here the frequency is the total number of parent change events divided by the number of link quality updates (including control beacons and data packet updates). From Figure 3 we see that the parent change events widely exist in the network. However, the value of $freq$ of over $60\%$ parent change events is only one in Figure 6. Based on these observations, we formally define all these parent change events which $freq$ value are larger than one as "high frequency parent change events" (HPC events) and remaining ones as "low frequency parent change events" (LPC events). HPC events are paid more attentions since they will potentially cause network dynamics.

*3) The lifetime of routing link:* Though a part of nodes frequently switch their parents, the size of their forwarding sets are limited. Are some nodes more likely chosen as parent nodes? To figure it out, we check the lifetime of routing links. We define the lifetime of routing link from $A$ to $B$ is the time interval during which node $A$'s parent is node $B$. However it is difficult to measure the accurate lifetime of the routing links since we lack of detailed parent switching timing. In order to approximately estimate the lifetime of the routing links, we provide an alternative approach by discretization in time: each ten minutes which is our sampling period notes a timing point. If we observe that no parent change event happens within the interval between $t_i$ and $t_j$, and the parents at $t_{i-1}$ and $t_{j+1}$ are different from the parent in the above interval, we consider $t_j$-$t_i$+1 as the lifetime of a valid routing link. In Figure 7, we plot CDF of the lifetime of the routing link. Around $33\%$ of routing links have a lifetime of below 10 minutes, $70\%$ of routing links have a lifetime of below 100 minutes, and around $10\%$ of routing links' lifetime are over 1000 minutes. It represents that many routing links are temporarily established and then die out quickly. Note that short duration dominates the lifetime for most of routing links, which means that the protocol seeks for better routing performance by frequently switching the parents. **Does the routing performance increase after the routing selection?** The evaluation will be provided in the section VII.

## VI. CAUSES OF ROUTING DYNAMICS

According to the above analysis, the interference caused by the heavy traffic may not be the only reason for HPC events. Other reasons, such as link failures, node failures, and the link recovery, may affect the routing dynamics. In this section we focus on identifying root causes of those HPC events.

### A. Distinguishing correlated and independent parent change events

Firstly parent change events are grouped. Here we name to the parent change events which cause other parent change events as *trigger parent change events* (TPC events). Accordingly, the parent change events caused by others are called *influenced parent change events* (IPC events). Grouping these events can help narrow down the suspect set, as it allows us to focus on root causes of TPC events. However, inappropriately grouping two events are undesirable. Hence, the problem in this subsection is, how to partition all events into disjoint groups, and each element in one group is caused by the same event.

We will discuss our strategy of grouping parent changes based on spatial and temporal correlations in detail.

Based on the principle of the routing selection in CTP, a parent change event will occur if both of these two requirements are satisfied:

1) At least one neighbor other than the current parent is found with an acceptable link quality and that neighbor is not congested.
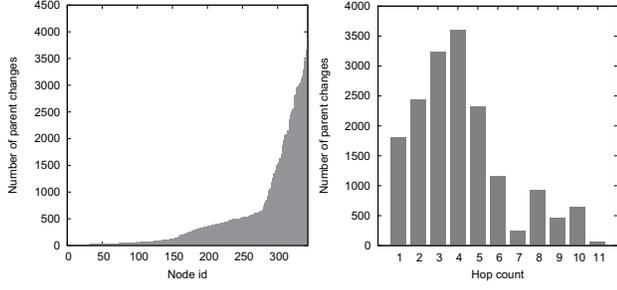
Fig. 4. Number of parent changes per node

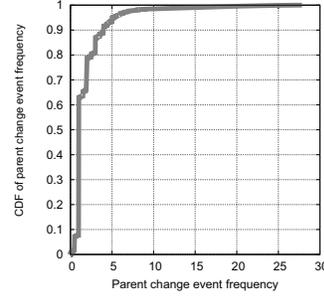Fig. 5. Number of parent changes on different hops

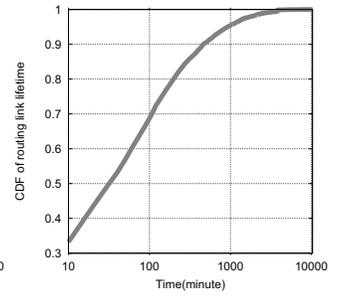Fig. 6. CDF of the frequency of parent change events

Fig. 7. Routing link Lifetime

2) The current parent is congested and the second best route is at least as good OR the current parent is not congested and the link quality of the neighbor is better than a threshold. The default ETX threshold is 1.5.

A sidenote of this principle is that if the current parent is congested, then in order to avoid forming loops, we will select a node which is not a descendant of the current parent.

Parent change events are correlated across two dimensions, *time* and *location*. The *time* dimension observes the set of parent change events close together in time. The *location* dimension takes into account the topology relationship between nodes in parent change events.

**Correlation across Time:** When a node updates its ETX value, other nodes will not be aware of its change immediately. Updating messages need to be passed by beacons. The beacon rate is always in a low level in LPL model. Unless the node detects an inconsistency of one neighbor's ETX value, the beacon rate will be adaptively changed. Hence, we say two parent change events $\langle\ x,\ xt_a,\ xp_a,\ xt_b,\ xp_b,\ xcounter_{pc},\ x_{freq}\ \rangle$ and $\langle\ y,\ yt_a,\ yt_a,\ yt_b,\ yt_b,\ ycounter_{pc},\ y_{freq}\ \rangle$ are correlated across the time dimension if:

$$|min\{xt_a, yt_a\} - max\{xt_b, yt_b\}| \leq |hop_x - hop_y| * period_{beacon}$$
(1)

where $hop_x$ and $hop_y$ are hop counts of node $x$ and $y$ respectively. $period_{beacon}$ refers to the beacon period. Each node broadcasts the beacon containing its ETX value every 8 minutes at most.

**Correlation across Location:** Two kinds of impact areas of a parent change event, *descendant nodes* and *neighboring nodes* are considered. When correlating parent change events across the location dimension, we need to handle two categories separately. Correlating neighboring nodes is easier, for we just need to check the parent change events on neighboring nodes within a pre-defined time lag, e.g. the maximum beacon period. To group upstream nodes, path information extracted from C1 packet is utilized. When a node and its closet downstream node switch their parents, if their timings satisfy Equation 1, we say that they are correlated across the location

dimension. It is high likely that the parent change of upstream node is caused by the parent change of the downstream node. We group two parent change events and define the parent change event of downstream node is TPC events.

Note that inappropriately handling grouping issues will lead to a large number of false positive or negative groups. For example, due to the loss of packets, we cannot know the exact timing when those parent change events happen. Counting all of those parent change events within a long time duration will significantly increase the number of false positive groups. We set a threshold which is the maximum delay for beacon announcement from the sink to the farthest nodes in the network. If and only if a parent change event lasts no more than this threshold, it can be grouped. For the other parent change events with duration larger than the threshold, we call them "fuzzy parent change events" (FPC events) and analyze their root causes in the following sections.

After employing the grouping strategy, we now have 6,132 groups consist of 16,014 correlated parent change events excluding 1,267 FPC events. Among the 16,014 parent change events, 13,861 of them are TPC events and 3,233 of them are IPC events. Figure 8 and Figure 9 show the temporal and spatial distributions of grouped parent change events respectively. The blue line represents the number of TPC events and the red line indicates the number of IPC events. The figures show that in both of the temporal and spatial distributions, the number of IPC is positively proportional to the number of TPC events and the ratio between them is below 15%. It implies that the parent change event might only have a local impact and will not cause topology changes in a large area. To confirm our conjecture, we plot the CDF of the group size in Figure 10. It can be observed that 94% of groups have only one parent change event and only 0.2% of groups have more than 10 parent change events. Most of the groups are relatively small in size, so they usually have small effects and will not affect the routing strategy globally. The reason might be the parent change event will only occur when the ETX difference between the node's best routing and second best routing is more than a threshold. This rule prevents the network topology from oscillating. Besides, a moving average
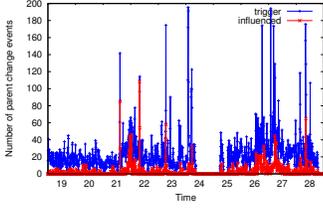
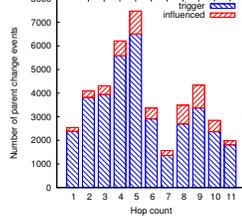Fig. 8. Temporal distribution of grouped parent change events

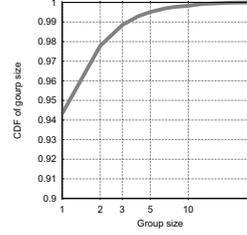Fig. 9. Spatial distribution of grouped parent change events

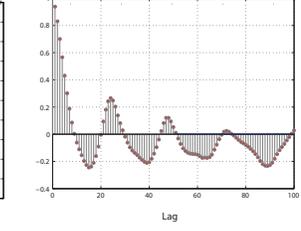Fig. 10. The size of grouped parent change events

Fig. 11. Auto-correlation function on the long duration parent change events

method is employed on the link estimation and a large weight is assigned to history value, they lead to that the estimation value of link qualities will not change too much within a short time.

### B. Causal relationship between parent change events and routing link degradation

Considering FPC events' long durations, we cannot precisely correlate these events with other parent change events. Then, we would like to know whether long duration packet losses provide some hints in identifying the root causes. As shown in Figure 3, for nodes close to the sink, they have many periodic high HPC events (periodic red lines in the figure) with long durations. Do these events share the same cause? FPC events in each hour of the first 6 days are aggregated and correlated with several other network events, e.g., no_ack_drop events, loop events, reboot events and so on. Finally, no_ack_drop events have strong correlations with FPC events. Figure 12 shows the temporal and spatial distributions of $counter_{pc}$ of those aggregated FPC events correlated with no_ack_drop events. From Figure 12a, aggregated number of FPC events exhibits a rough periodic patten and the number of correlated parent change events reach its bottom at noon. As shown in Figure 12b, most of correlated FPC events locate near the sink.

Based on the observation, we are interested in the intrinsic characteristics of the periodic patten. Auto-correlation method is employed to evaluate temporal correlations for correlated FPC events. Auto-correlation function is a standard tool in time series to analyze repeated patterns at different period lengths. As shown in Figure 11, a high degree of temporal correlation of the number of FPC events. Specifically, the crest occur at a period of 24 hours and the trough occur at a period of 18 hours. It implies that the period of 24 reaches the maximum probability to be the repeated pattern. According to our experiences, no_ack_drop events always refers to the link degradations and is highly likely caused by either environmental issues or wireless interferences. Since the system is deployed in a forest and there rarely exist interference signals, we conclude a conjecture that environmental factors may have significant impact on the periodic phenomenon.

After examining the link qualities from neighbor tables of typical nodes, we find that most of them become extremely poor during 8pm to 10am. They cannot find good neighbors to forward packets. Further inspections show that the humidity during these time intervals is quite high. As indicated in [25], water accumulated on foliage or the antenna is more likely a contributing factor.

### C. Causal relationship between parent change events and loops

In the above subsection, we have correlated a part of FPC events with no_ack_drop events. However, some FPC events are not caused by link degradations. In order to identify root causes of remaining FPC events, we review Figure 3 and find that those remaining events constitute some vertical bandings (vertical red bandings in the figure). The vertical bandings imply that those parent change events co-occur on multiple nodes. In the design of the CTP, when detecting an inconsistency of ETX value, a node reselects the parent to break the loop. The process usually involves several nodes simultaneously. So another conjecture is concluded that **do those remaining *fuzzy parent change events* have strong correlation with loop events?**

Figure 14 shows the temporal distribution of $counter_{pc}$ of remaining FPC events and $counter_{loop}$ of loop events. The blue line is the number of parent changes and the red line indicates the number of loops. As shown in Figure 14, many nodes with large loop counters also experience more parent changes. It suggests that two kind of events have high correlations. Figure 13 shows the temporal and spatial distribution of $counter_{PC}$ of those remaining aggregated FPC events correlated with loop events. From Figure 13b, most correlated events locate at far-away nodes. To quantitatively analyze their relationship, we calculate the pearson correlation between the remaining FPC events and loop events for each node. Figure 15 shows that all nodes have correlation coefficients larger than 0.51. It implies that loop events have strong correlation with remaining FPC events. Many packets are lost due to loops when we check no_ack_drop counters. However, if loops can be broken down in a short duration, many packets will not be lost. Intuitively, a question is asked: **how effective is parent**
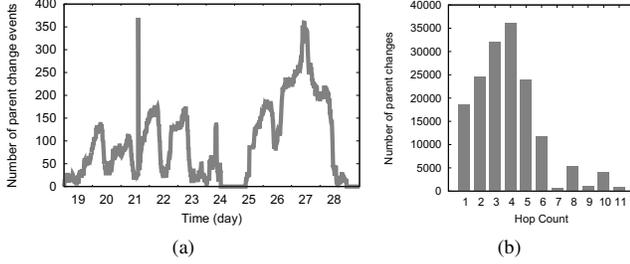
Fig. 12. (a) Temporal distribution of parent change events correlated with no_ack_drop events. (b) Spatial distribution of parent change events correlated with no_ack_drop events.
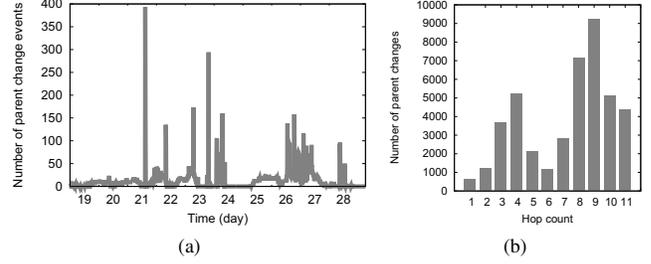


Fig. 13. (a) Temporal distribution of parent change events correlated with loop events. (b) Spatial distribution of parent change events correlated with loop events.
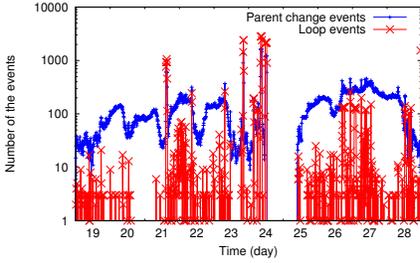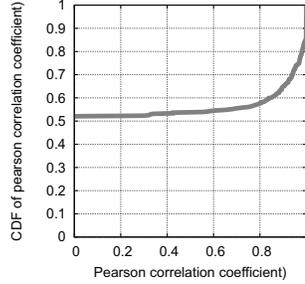


Fig. 14. Loop events vs. parent change events

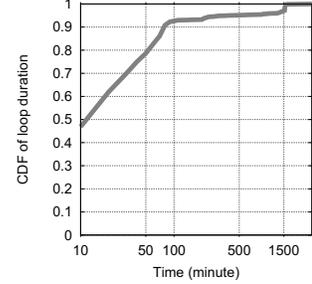Fig. 15. CDF of pearson correlation coefficient
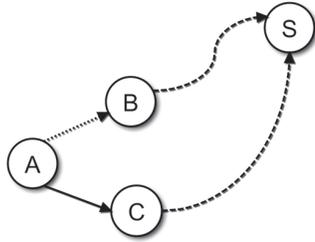
Fig. 16. CDF of loop duration



Fig. 17. Illustration of parent change effectiveness detecting algorithm

**change mechanism for breaking loops?** Figure 16 shows the duration of the loops correlated with FPC events. Though $40\%$ of correlated loops can be dispersed within 10 minutes, nearly $20\%$ of loops which at least last 50 minutes to be broken down. Long duration loops are harmful to the protocol performance. A lot of packets will be dropped if buffering queues overflow. Therefore, effective methods for breaking down loops or loop-free methods are necessary to avoid packet losses.

## VII. IMPACT OF ROUTING DYNAMICS

After analyzing some possible causes of routing dynamics, we view whether parent changes provide the optimal solution for routing under CTP. We compare the packet reception ratios between the nodes choosing old parent and those choosing new parent. Assume node $A$ 's previous parent is $B$ at time $t_1$ and its current parent is $C$ at time $t_2$. As shown in Figure 17, we

compare the packet reception ratio of $A$, using $C$ as the parent for a given time interval like 50 minutes after time $t_2$. This term is described as $PDR_C$ for short. The packet delivery ratio of $A$, still using $B$ as the parent for the same amount of time interval after $t_2$. This term is $PDR_B$ for short. If $PDR_B$ is better than $PDR_C$ then it implies that this parent change may not be necessary.

In practice, we are unable to obtain the ground truth of $PDR_B$. Here an alternative approach is provided: If the RSSI values of $link_{AB}$ are larger than a threshold at both of time $t_1$ and $t_2$ (normally -85 dBm according to our experience threshold. Since when the link RSSI is higher than -85 dBm, we always observe good communication performances), we consider that the link quality does not change significantly. Under these situations we take the $PDR$ value of node $B$ after time $t_2$ as an approximation of $PDR_B$. If the approximated value $\hat{PDR_B}$ is better than $PDR_C$, we consider that parent change as ineffective. We define the ratio of the number of ineffective parent change events and total parent change events for each node as "Routing selection effectiveness coefficient". By counting it, we obtain Figure 18. Around $30\%$ of nodes suffers from ineffective routing decisions. Their ineffective routing selections take up more than $60\%$. It is becaused in CTP, if we only consider the ETX value and exclude other factors like retransmission upper bound and forwarding quality of the node, stable links might not be chosen. Therefore ineffective routing oscillation will be incurred.
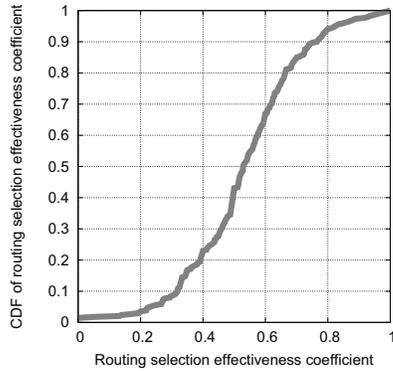
Fig. 18. CDF of routing selection effectiveness coefficient

## VIII. CONCLUSION

In this paper, we conduct extensive measurements of routing dynamics in a large-scale WSN. The measurement results and the analysis provide answers to several fundamental questions regarding to routing dynamics: To what extent does the routing protocol exhibit dynamics? What causes routing dynamics? What is the impact of routing dynamics? The answers we provided in this paper give new understanding of the interactions among multiple network elements, evaluating protocol design strategies, and improving system performance.

The key findings of our study include: 1) Parent change events mostly impact only local nodes, *i.e.* they do not result in routing instability on distant nodes; 2) Environment and routing loops have large impact on routing dynamics; 3) While a large portion of parent changes are effective in improving network performance, a few changes might not be necessary and may even hurt the network routing performance. All those findings will provide guidelines to understand the routing dynamics and design more robust and adaptive routing protocols in dynamic environments.

## IX. ACKNOWLEDGEMENT

## REFERENCES

[1] R. Tan, G. Xing, J. Chen, W.-Z. Song, and R. Huang, "Fusion-based volcanic earthquake detection and timing in wireless sensor networks," *ACM TOSN*, vol. 9, no. 2, p. 17, 2013.

[2] Q. Cao, T. F. Abdelzaher, T. He, and J. A. Stankovic, "Towards optimal sleep scheduling in sensor networks for rare-event detection," in *Proceedings of ACM/IEEE IPSN*, pp. 20–27, 2005.

[3] B. Kusy, Á. Lédeczi, and X. D. Koutsoukos, "Tracking mobile nodes using rf doppler shifts," in *Proceedings of ACM SenSys*, pp. 29–42, 2007.

[4] L. Mo, Y. He, Y. Liu, J. Zhao, S. Tang, X.-Y. Li, and G. Dai, "Canopy closure estimates with greenorbs: sustainable sensing in the forest," in *Proceedings of ACM SenSys*, pp. 99–112, 2009.

[5] Q. Ma, K. Liu, X. Miao, and Y. Liu, "Sherlock is around: Detecting network failures with local evidence fusion," in *INFOCOM*, pp. 792–800, 2012.

[6] O. Gnawali, R. Fonseca, K. Jamieson, D. Moss, and P. Levis, "Collection tree protocol," in *Proceedings of ACM SenSys*, pp. 1–14, 2009.

[7] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: a scalable and robust communication paradigm for sensor networks," in *Proceedings of ACM MOBICOM*, pp. 56–67, 2000.

[8] A. Woo, T. Tong, and D. E. Culler, "Taming the underlying challenges of reliable multihop routing in sensor networks," in *Proceedings of ACM SenSys*, pp. 14–27, 2003.

[9] G. Werner-Allen, K. Lorincz, J. Johnson, J. Lees, and M. Welsh, "Fidelity and yield in a volcano monitoring sensor network," in *Proceedings of USENIX OSDI*, pp. 381–396, 2006.

[10] J. Zhao and R. Govindan, "Understanding Packet Delivery Performance In Dense Wireless Sensor Networks," in *Proceedings of ACM SenSys*, 2003.

[11] G. Zhou, T. He, S. Krishnamurthy, and J. A. Stankovic, "Impact of radio irregularity on wireless sensor networks," in *Proceedings of ACM MobiSys*, 2004.

[12] K. Srinivasan, M. A. Kazandjieva, S. Agarwal, and P. Levis, "The beta-factor: measuring wireless link burstiness," in *Proceedings of ACM SenSys*, pp. 29–42, 2008.

[13] K. Srinivasan, M. Jain, J. I. Choi, T. Azim, E. S. Kim, P. Levis, and B. Krishnamachari, "The kappa factor: inferring protocol performance using inter-link reception correlation," in *Proceedings of ACM MOBICOM*, pp. 317–328, 2010.

[14] G. Tolle, J. Polastre, R. Szewczyk, D. Culler, N. Turner, K. Tu, S. Burgess, T. Dawson, P. Buonadonna, D. Gay, and W. Hong, "A Macroscope in the Redwoods," in *Proceedings of ACM SenSys*, 2005.

[15] T. He, P. Vicaire, T. Yan, Q. Cao, G. Zhou, L. Gu, L. Luo, R. Stoleru, J. A. Stankovic, and T. F. Abdelzaher, "Achieving Long-Term Surveillance in VigilNet," *ACM/IEEE TON*, vol. 5, no. 1, pp. 1–39, 2009.

[16] A. Arora, R. Ramnath, E. Ertin, and et al., "ExScal: Elements of an Extreme Scale Wireless Sensor Network," in *Proc. of IEEE RTCSA*, 2005.

[17] G. Barrenetxea, F. Ingelrest, G. Schaefer, M. Vetterli, O. Couach, and M. Parlange, "SensorScope: Out-of-the-Box Environmental Monitoring," in *Proc. of ACM/IEEE IPSN*, 2008.

[18] R. Maheshwari, S. Jain, and S. R. Das, "A measurement study of interference modeling and scheduling in low-power wireless networks," in *Proceedings of ACM SenSys*, pp. 141–154, 2008.

[19] K. Srinivasan, P. Dutta, A. Tavakoli, and P. Levis, "Understanding the Causes of Packet Delivery Success and Failure in Dense Wireless Sensor Networks," tech. rep., Stanford University, 2006.

[20] Y. Liu, Y. He, M. Li, J. Wang, K. Liu, L. Mo, W. Dong, Z. Yang, M. Xi, J. Zhao, and X.-Y. Li, "Does wireless sensor network scale? a measurement study on greenorbs," in *Proceedings of IEEE INFOCOM*, pp. 873–881, 2011.

[21] A. Markopoulou, G. Iannaccone, S. Bhattacharyya, C.-N. Chuah, and C. Diot, "Characterization of failures in an ip backbone network," in *Proceedings of IEEE INFOCOM*, 2004.

[22] D. Turner, K. Levchenko, A. C. Snoeren, and S. Savage, "California fault lines: understanding the causes and impact of network failures," in *Proceedings of ACM SIGCOMM*, pp. 315–326, 2010.

[23] TinyOS, "http://www.tinyos.net."

[24] G. Tolle and D. E. Culler, "Design of an application-cooperative management system for wireless sensor networks," in *Proceedings of EWSN*, pp. 121–132, 2005.

[25] A. Markham, N. Trigoni, and S. A. Ellwood, "Effect of rainfall on link quality in an outdoor forest deployment," in *Proceedings of WINSYS*, pp. 148–153, 2010.