Computer Networks 139 (2018) 81–91

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Accurate per-link loss tomography in dynamic sensor networks

Chenhong Cao, Yi Gao*, Wei Dong, Jiajun Bu

College of Computer Science, Zhejiang University, China

ARTICLE INFO

Article history:
Received 15 September 2017
Revised 11 March 2018
Accepted 11 April 2018
Available online 17 April 2018

Keywords:
Wireless sensor network
Loss tomography
Arithmetic coding

ABSTRACT

Wireless Sensor Networks (WSNs) have been successfully applied in many application areas. Understanding the wireless link performance is very helpful for both protocol designers and network managers to improve the network performance and prolong the network lifetime. Loss tomography is a popular approach to infer the per-link loss ratios from end-to-end delivery ratios. Previous studies, however, are usually targeted for networks with static or slowly changing routing paths. In this work, we propose Dophy, a **D**ynamic l**o**ss tomogra**phy** approach specifically designed for dynamic WSNs where each node dynamically selects the forwarding nodes towards the sink. The key idea of Dophy is based on an observation that most existing protocols use retransmissions to achieve high data delivery ratio. Dophy employs arithmetic encoding to encode the number of retransmissions along the paths compactly. Dophy incorporates two mechanisms to optimize its performance. First, Dophy intelligently reduces the size of the symbol set by aggregating the number of retransmissions, reducing the encoding overhead significantly. Second, Dophy periodically updates the probability model to minimize the overall transmission overhead. We implement Dophy on the TinyOS platform and evaluate its performance extensively using large-scale simulations. Results show that Dophy achieves both high encoding efficiency and high estimation accuracy. Comparative studies show that Dophy significantly outperforms traditional loss tomography approaches in terms of accuracy.

1. Introduction

As an emerging technology that bridges cyber system and the physical world, wireless sensor networks (WSNs) are envisioned to support numerous applications that are unthinkable before [1–3]. In these applications, data packets are usually delivered towards a central sink through a multi-hop wireless network. Reliable packet delivery is one of the most important goals for a deployed sensor network. However, the link performance still suffers from network dynamics due to environmental factors, posing difficulties for achieving good packet delivery performance.

Lossy links can cause routing loops which degrade the network performance quickly. In particular, routing loops can frequently be formed with lossy links, and once the loop happens, the energy is wasted on forwarding packets involved in the loop [4]. Understanding the wireless link performance is very helpful for both protocol designers and network managers. For protocol designers, routing performance improvement can be achieved by blacklisting bad links [5]. For network managers, the availability of link-level metrics can greatly facilitate anomaly detection and diagnosis. For example, the network manager can easily detect the problematic node with lossy links due to low battery or physical obstacles [6]. Moreover, the link state information can also be used in the offline analysis for security or simply enhancing our collective understanding of the network as a complex system.

A wireless link is usually characterized by its loss ratio. Loss tomography is a popular approach to inferring the per-link loss ratios from only end-to-end delivery ratios. Due to its small overhead, loss tomography is widely studied in the literature [7–11].

Previous studies, however, are usually targeted for networks with static or slowly changing routing paths. By establishing a linear system with *known* path metrics and *unknown* link metrics, we can infer the link metrics including per-link loss ratios. Such approaches are not suitable for dynamic WSNs for two reasons. First, links in WSNs are changing frequently. To obtain fine-grained and time-varying link metrics, traditional approaches require frequent probing, which incurs too much overhead for sensor networks. Second, WSNs usually employ dynamic routing protocols to adapt to link dynamics. CTP (Collection Tree Protocol) [12] is an instance of dynamic routing protocols which delivers data to the sink node via multihop wireless. Due to dynamic routing, the number of *unknown* link metrics may be significantly larger than the number of *known* end-to-end metrics. Therefore, traditional approaches cannot effectively infer link metrics by solving the linear system.

* Corresponding author.
E-mail addresses: chcao@zju.edu.cn (C. Cao), gaoyi@zju.edu.cn (Y. Gao).

https://doi.org/10.1016/j.comnet.2018.04.007
1389-1286/

As an example, we consider a WSN consisting of N sensor nodes, delivering data to a sink node via multihop wireless. Each node selects its parent for data forwarding. With static routing, there will be $N-1$ routing links as well as $N-1$ node-to-sink metrics (e.g., delivery ratios from each node except for the sink). Hence, we can easily perform loss tomography. With dynamic routing, however, there will be $N' \gg N$ routing links. Still, we can only obtain $N-1$ node-to-sink metrics. Hence, the link metrics can no longer be inferred with dynamic routing.

In order to address the above two challenges, we propose Dophy, a **D**ynamic l**o**ss tomogra**phy** approach in contrast with existing approaches designed with static or slowly changing routing paths. The key idea of Dophy is based on an observation that most existing protocols use retransmissions to achieve high data delivery ratio towards the sink. Hence, we could encode into data packets the number of retransmissions along the forwarding paths which can be efficiently inferred by recent path tracking techniques [13,14]. With such per-packet information, we can easily decode the *fine-grained* link metrics.

A practical challenge is how to achieve optimal encoding, i.e., encoding all information with the smallest overhead, or, encoding information as much as possible with a given overhead. A basic idea for compression is to use a short code value to represent frequently occurring symbols which refer to the number of retransmissions in our case. Dophy employs arithmetic encoding since it achieves close-to-optimal compression rates. Dophy maintains a separate probability model for each node (at both the node itself and the sink node) to track the symbol frequencies. Dophy incorporates two mechanisms to optimize its performance in dynamic WSNs. First, Dophy intelligently reduces the size of the symbol set by aggregating the number of retransmissions. For example, retransmitting a packet for 20 times and 21 times both indicate a very lossy link, so we can aggregate these numbers into one category. By this mechanism, the encoding overhead can be significantly reduced. Second, Dophy periodically updates the probability model to minimize the overall transmission overhead including both the encoding overhead and the updating overhead. Note that there is an inherent tradeoff between the encoding overhead and the updating overhead: frequent updating incurs much overhead but yields an accurate probability model which allows small encoding overhead.

We implement Dophy on the TinyOS platform and evaluate its performance extensively using large-scale simulations. Results show that Dophy achieves high encoding efficiency and estimation accuracy. Comparative studies show that Dophy significantly outperforms traditional loss tomography approaches in terms of accuracy.

The contributions of this paper are summarized as follows:

- We propose Dophy, a fine-grained loss tomography approach for dynamic WSNs, exploiting arithmetic encoding to compactly encode per-hop retransmissions for instantly and accurately inferring per-link loss rates.
- We propose a symbol distribution updating mechanism to maintain the retransmission probability model at each node. By this mechanism, the overall transmission overhead can be optimized.
- We implement Dophy on the TinyOS platform and extensively evaluate its performance by large-scale simulations. Results show that our approach significantly outperforms existing approaches.

The rest of this paper is structured as follows. Section 2 describes the related work. Section 3 introduces the background of arithmetic coding. Section 4 presents the design of Dophy. Section 5 describes the implementation details. Section 6 shows the evaluation results, and finally, Section 7 concludes this paper.

2. Related work

2.1. Loss measurement in wired IP networks

Packet loss measurement is widely studied in wired IP networks to provide the foundation for many network protocols and systems. Bolot [15] and Paxson [16] evaluate end-to-end probe measurements and report features of packet loss over a selection of paths on the Internet. ZING [17] measures end-to-end packet loss between two participating end hosts in one-direction. Besides the end-to-end loss measurement approaches, there are also approaches aiming at inferring loss rate on individual links. It has been demonstrated to be effective to use Network Tomography based on both multicast and unicast probes for inferring loss rates of internal links on end-to-end paths [18]. However, these loss inference approaches for IP network are based on the assumption that the routing path is stable and the link metrics are constant during the measurement period. Hence the above works cannot be easily applied to WSNs due to the high environmental and routing dynamics.

2.2. Loss measurement in wireless sensor networks

Per-link loss rate inference. There are many approaches [7,10,11] for loss rate inference in WSNs. In [7], network tomography was firstly applied in WSNs for link loss inference. They formulate the problem as a Maximum-Likelihood Estimation (MLE) problem and use the Expectation-Maximization algorithm to solve it. [19] investigates active loss tomography on mesh topologies and uses a linear algebraic approach to estimate link loss rates. LIPM [11] is a representative approach to inferring link loss performance based on network tomography. However, these approaches are based on WSNs where the aggregation trees remain static for a long time. In practice, many WSNs employ dynamic routing protocols, e.g., CTP [12], for multihop data forwarding. In these networks, the performance of existing approaches degrades significantly. In contrast, Dophy is designed specifically for such dynamic networks, and its performance is significantly better than existing approaches as shown in this paper.

Lossy links identification. In WSNs, there are several approaches for identifying the lossy links [6,20]. Compared with loss inference approaches, these approaches usually incur less overhead since they only require coarse-grained information. LLIS [6] is proposed to isolate links with high loss rate based on a set cover heuristic. However, this approach lacks real-time feature. Link Scanner [20] is a probe-based and rule-free approach for detecting faulty links. It depends on a relatively stable network topology which may, however, be changed due to a dynamic environment and routing policy. [21] addresses loss inference in wireless networks under network coding traffic and proposes inference algorithms based on Bayesian principles to discover the set of highly lossy links in sensor networks. Different from these approaches, Dophy aims at inferring fine-grained per-link loss rates.

3. Background on arithmetic coding

In this section, we give some background information about arithmetic coding. Arithmetic coding [22] is a lossless data compression method which achieves close-to-optimal compression ratios. It maps a sequence of data symbols to a real number in the interval [0, 1). This real number is called “code value”. We denote the total number of possible symbols in the data alphabet as n. We use an example to illustrate how to perform encoding and decoding with arithmetic coding. In this example, the symbol set contains three symbols $\{a, b, c\}$ with occurring probabilities