Multi-hop Network Tomography: Path Reconstruction and Per-hop Arrival Time Estimation from Partial Information

Matthias Keller, Jan Beutel, and Lothar Thiele
Computer Engineering and Networks Laboratory
ETH Zurich, Switzerland
{kellmatt,beutel,thiele}@tik.ee.ethz.ch

ABSTRACT

In the context of low-power wireless sensor networks, this paper presents multi-hop network tomography (MNT), a novel, non-intrusive algorithm for reconstructing the path, the per-hop arrival order, and the per-hop arrival time of individual packets at runtime. While explicitly transmitting this information over the radio would negatively impact the performance of the system under investigation, information is instead reconstructed after packets have been received at the sink.

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design—*Wireless communication*

Keywords

Wireless Sensor Networks, Data Analysis, Tomography

1. OVERVIEW

Wireless sensor networks (WSNs) are networks of small, low-power sensor nodes that autonomously build and maintain a network structure. Running at very low duty-cycles, battery-powered sensor nodes can continuously operate for several years without being serviced. Those characteristics make WSNs well-suited for data collection applications, *e.g.*, environmental monitoring [5].

However, as significantly increasing the size or amount of packets transmitted threatens the system performance, *e.g.*, increases the packet loss rate [8], available performance metrics are usually limited to the most important health indicators, *i.e.*, end-to-end data yield, end-to-end packet delays, and battery voltage readings [7]. While those metrics are well-suited for coarsely estimating system health, understanding the root causes of an observed performance requires yet unavailable measurements taken inside the network.

For addressing this problem, this paper presents the MNT algorithm for reconstructing the path, the per-hop order, and the per-hop arrival time of individual packets. Information is reconstructed after packets arrived at the sink. Thus, the MNT algorithm allows to analyze the system performance both at runtime and at any later point in time. Without limiting the scope of this new tool, the MNT algorithm is firstly intended for the analysis of deployed systems that were continuously extended by adding nodes and new sensing modalities, *e.g.*, sensors that sample with an in the order of magnitude higher rate. Here, the interest lies in making more informed decisions when modifying an existing system configuration.

SIGMETRICS'12, June 11–15, 2012, London, England, UK.
ACM 978-1-4503-1097-0/12/06.

2. SYSTEM MODEL

We assume a multi-hop wireless sensor network that consists of a number of static sensor nodes that report to one or multiple sink nodes. All nodes produce and relay data in the sense of a data collection application. Nodes do not have access to global timing information and rely on a local clock.

Lossy channel. Communication between nodes may fail, *i.e.*, a packet transmitted is not received by the intended receiving side.

Multi-hop data collection. Each sink is the root of a multi-hop tree topology, a tree structure that consists of concatenated 1:n one-hop parent-child relations. The topology is autonomously maintained by the routing protocol used, *e.g.*, CTP [3].

FIFO send queue. All sensor nodes have the dual functionality of (i) generating data locally and (ii) forwarding packets received from other nodes. Nodes maintain a finite FIFO send queue. A packet is immediately added to this queue after generation in the local application or after arrival on the radio. If connected, a sensor nodes transmits the contents of its send queue to its parent.

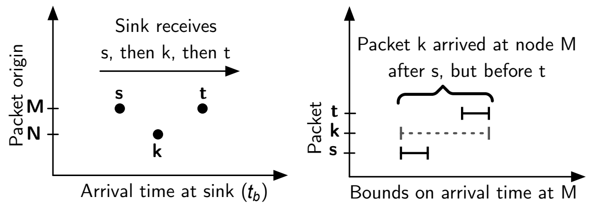
Available per-packet information. For each packet k, the first-hop receiver address $p(k)$ is the address of the node to which the source node $o(k)$ forwarded the packet. The time of arrival at the sink $t_b(k)$ is measured on a precise, perfect clock, thus the order of $t_b(k)$ corresponds to the order of arrival at the sink. In contrast, the packet generation time $t_g(k)$ of a packet k is measured on clocks with imperfections, *e.g.*, low time resolution and drift, and can thus only be estimated. The maximum error of the generation time estimate $\tilde{t}_g(k)$ is bounded by $\tilde{t}_g(k) - \Delta^l(k) \leq t_g(k) \leq \tilde{t}_g(k) + \Delta^u(k)$. Worst-case clock error bounds $\Delta^{u,l}(k)$ are specific to a certain clock implementation. The unique packet index $id_M(k)$ is defined so that its order yields the order of packet generation at a node M: $id_M(u) > id_M(v)$ iff $t_g(u) > t_g(v)$. Please note that commonly found packet sequence numbers might not fulfill the required properties [6].

Observability of parent changes. Nodes can decide to switch their parent node during operation. Path changes can only be safely detected if the parent of a node is not changing more than once between the successful transmission of two consecutively, locally generated packets. Concretely, there must have been no parent change if consecutively generated packets u and v report equal first-hop receivers $p(u) \equiv p(v)$.

3. INFORMATION RECONSTRUCTION

Given a packet k, we want to reconstruct the following:

Packet path $\mathcal{N}_k$: Starting at the packet source $o(k)$, the ordered set $\mathcal{N}_k$ contains all nodes that packet k visited until arriving at the sink node S. The order of $\mathcal{N}_k$ reflects the order of visited nodes.



(a) Selection of s and t (b) Arrival time estimation

Figure 1: Information of a forwarded packet k is reconstructed using information from locally generated packets s and t.

Queue index $qid_N(k)$: For all nodes N that k visited, *i.e.*, $\forall N \in \mathcal{N}_k$, we want to build a queue index $qid_N(k)$ so that the order of $qid_N(k)$ reflects the order of packet arrival at N: The queue index is larger, *i.e.*, $qid_N(m) > qid_N(n)$ iff packet m arrived at N after another packet n, *i.e.*, $t_a(N,m) > t_a(N,n)$. In contrast to the already known packet index $id_N(k)$, the queue index $qid_N(k)$ provides not only the sequence of locally generated packets, but also that of forwarded packets.

Bounds on queue arrival time $t_a^{u,l}(N,k)$: For all nodes N that k visited, *i.e.*, $\forall N \in \mathcal{N}_k$, we want to bound the unknown queue arrival time $t_a(N,k)$ so that $t_a^l(N,k) \leq t_a(N,k) \leq t_a^u(N,k)$.

The idea is to extract missing information from locally generated packets while following the path of k through the network: While the first-hop receiver $p(k)$ of packet k is already known, *e.g.*, $p(k) := M$, the next hop after node M is extracted from two packets s and t that were generated at M. Likewise, packets s and t are used for assigning a queue index of packet k at M and for bounding the arrival time of k at M. Concretely, the queue index $qid_M(k)$ of packet k at node M must be chosen so that $qid_M(s) < qid_M(k) < qid_M(t)$. The arrival time $t_a(M,k)$ of k at M is bounded using the generation times of packets s and t, *i.e.*, $\tilde{t}_g(s) - \Delta^l(s) \leq t_a(M,k) \leq \tilde{t}_g(t) + \Delta^u(t)$. Ideally, this procedure is repeated until packet k has been traced up to the sink.

The question arises how can we determine corresponding packets s and t that were generated at M and that can be used to extract missing information: Packets from different sources are correlated based on their order of arrival at the sink. In the exemplary situation in Figure 1, packets s and t were consecutively generated at a node M, whereas packet k was generated at another node N. The observed order of arrival at the sink is shown in Figure 1(a), *i.e.*, packet k arrived at the sink after packet s, but before packet t. This matches with the order in which packets s, k and t arrived at the queue of the intermediate node M, see Figure 1(b).

It becomes apparent, that this approach can only yield correct information if the order of arrival at the queues of intermediate nodes matches with the observable order of arrival at the sink. The question arises how we can decide only with information available at the sink if corresponding sequences match, *i.e.*, extracting information is safe, or if packets might have been reordered while traveling through the network, and thus cannot be used. Defining $\mathcal{P}$ as the set of all packets that were received at the sink, the answer of this problem is to determine a set $\mathcal{R} \subseteq \mathcal{P}$ of "reliable" packets. The definition of a "reliable" packet is as follows: *A packet k is reliable,* i.e.*, $k \in \mathcal{R}$, if it fulfills two properties: From our observations at the sink, we can guarantee that (i) packet k can only have travelled along exactly one path $\mathcal{N}_k$, and that (ii) the order relation between packet k and any other packet $m \in \mathcal{R}$ is consistent along all packet queues in the network including the sink.*

While a detailed explanation including proofs would clearly exceed the format of this paper, we can only hint on the determination of a set $\mathcal{R} \subseteq \mathcal{P}$ of "reliable" packets: First, we need to determine if there are ambiguities in the observed path of a packet, *e.g.*, a packet arrived at a node M while the parent of this node M changed. Second, we need to exclude packet reordering at each hop that k visited. Parent changes are the single cause for packets arriving at the sink out of order, *i.e.*, packets that were sent out before a parent change experience a higher delay than packets that were transmitted after the parent change. Evidence for packet reordering due to a path change at a node M is found when locally generated packets of M arrive at the sink out of order.

4. VALIDATION AND EVALUATION

For validation and evaluation purposes, the MNT algorithm is applied to traces from two state-of-the-art data collection protocols, CTP [3] and Dozer [2]. Corresponding tests are executed on up to 91 Tmote Sky nodes (TI MSP430, CC2420 radio) that are part of the TWIST testbed [4], and 25 TinyNode nodes (TI MSP430, Semtech SX1211 radio) that are part of the FlockLab testbed [1]. Ground truth information is transmitted via a second communication channel, *i.e.*, written to the serial port of the sensor nodes. When compared to ground truth, reconstructed information is correct in all cases. The reconstruction performance of the MNT algorithm is presented in Figure 2. A packet has been fully traced if its complete path was reconstructed.

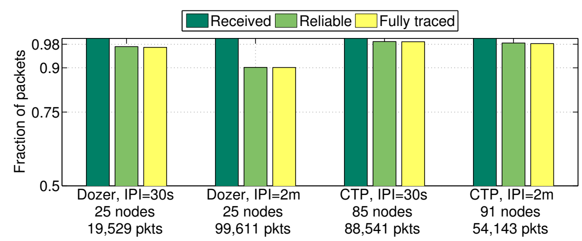


Figure 2: Performance evaluation using two well-known communication protocols with varying inter-packet intervals (IPI).

Acknowledgements. The work presented was supported by NCCR-MICS, a center supported by the Swiss National Science Foundation under grant number 5005-67322, and from the Swiss Nano-Tera.ch initiative.

5. REFERENCES

[1] J. Beutel, R. Lim, A. Meier, L. Thiele, C. Walser, M. Woehrle, and M. Yuecel. The flocklab testbed architecture. In *Proc. 7th ACM Conf. Embedded Networked Sensor Systems (SenSys '09)*, pages 415–416, 2009.

[2] N. Burri, P. von Rickenbach, and R. Wattenhofer. Dozer: ultra-low power data gathering in sensor networks. In *Proc. 6th Int'l Conf. Information Processing Sensor Networks (IPSN '07)*, pages 450–459, 2007.

[3] O. Gnawali, R. Fonseca, K. Jamieson, D. Moss, and P. Levis. Collection tree protocol. In *Proc. 7th ACM Conf. Embedded Networked Sensor Systems (SenSys '09)*, pages 1–14, 2009.

[4] V. Handziski, A. Köpke, A. Willig, and A. Wolisz. Twist: A scalable and reconfigurable testbed for wireless indoor experiments with sensor networks. In *Proc. 2nd Int'l Workshop on Multi-hop Ad Hoc Networks (REALMAN '06)*, pages 63–70, 2006.

[5] A. Hasler, I. Talzi, J. Beutel, C. Tschudin, and S. Gruber. Wireless sensor networks in permafrost research - concept, requirements, implementation and challenges. In *Proc. 9th Int'l Conf. on Permafrost (NICOP 2008)*, volume 1, pages 669–674, 2008.

[6] M. Keller, L. Thiele, and J. Beutel. Reconstruction of the correct temporal order of sensor network data. In *Proc. 10th Int'l Conf. on Information Processing in Sensor Networks (IPSN '11)*, pages 282–293, 2011.

[7] M. Keller, M. Woehrle, R. Lim, J. Beutel, and L. Thiele. Comparative performance analysis of the permadozer protocol in diverse deployments. In *Proc. of the 6th Int'l Workshop on Practical Issues in Building Sensor Network Applications (SenseApp '11)*, pages 969–977, 2011.

[8] Y. Sankarasubramaniam, I. Akyildiz, and S. McLaughlin. Energy efficiency based packet size optimization in wireless sensor networks. In *Proc. 1st Int'l Workshop on Sensor Network Protocols and Applications (SNPA '03)*, pages 1–8, 2003.