Opportunity-Based Topology Control in Wireless Sensor Networks

Yunhuai Liu, *Member*, *IEEE*, Qian Zhang, *Senior Member*, *IEEE*, and Lionel M. Ni, *Fellow*, *IEEE*

Abstract—Topology control is an effective method to improve the energy efficiency of wireless sensor networks (WSNs). Traditional approaches are based on the assumption that a pair of nodes is either "connected" or "disconnected." These approaches are called connectivity-based topology control. In real environments, however, there are many intermittently connected wireless links called lossy links. Taking a succeeded lossy link as an advantage, we are able to construct more energy-efficient topologies. Toward this end, we propose a novel opportunity-based topology control. We show that opportunity-based topology control is a problem of NP-hard. To address this problem in a practical way, we design a fully distributed algorithm called CONREAP based on reliability theory. We prove that CONREAP has a guaranteed performance. The worst running time is $O(|E|)$, where E is the link set of the original topology, and the space requirement for individual nodes is $O(d)$, where d is the node degree. To evaluate the performance of CONREAP, we design and implement a prototype system consisting of 50 Berkeley Mica2 motes. We also conducted comprehensive simulations. Experimental results show that compared with the connectivity-based topology control algorithms, CONREAP can improve the energy efficiency of a network up to six times.

Index Terms—Topology control transitional region network reachability.

1 Introduction

Topology control is an effective method to improve the energy efficiency of Wireless Sensor Networks (WSNs) [17], [23]. In traditional methods, the network model is based on the assumption that a pair of nodes is either "connected" or "disconnected." When all nodes are connected to the network, the network is said to have the full connectivity. Traditional approaches aim to seek a topology of minimized energy cost and the full connectivity. We therefore call these approaches *connectivity-based topology control*. Connectivity-based topology control is a well-known NPC problem [12] and many heuristic algorithms have been proposed (e.g., [6], [11], [14], [22], [28]).

In real application environments, however, this connectivity-based model is not practical due to the *transitional region phenomenon*. Beyond the "connected" region, there is a *transitional region* that allows wireless links to be intermittently connected. Such links are called *lossy links* [3], [4], [7], [30], [32], [33], [34]. Via lossy links, a pair of nodes are reachable but not always connected. As reported in the literature (e.g., [3], [34]), lossy links usually account a major part of a wireless network. In a specific setup, lossy links are more than 90 percent.

An interesting observation is that lossy links can be used to construct more energy-efficient topologies if used properly. Compared with these reliable links, lossy links allow a transmitter to reach more nodes when they succeed. By taking the succeeded lossy links as an advantage, many transmissions can be saved and the cost is reduced.

The employment of lossy links in topology is not straightforward. When they are employed, the sink may not be able to reach all nodes for a given transmission. An important characteristic of WSNs is the large density and high data redundancy. For many applications, to require the sink reach all nodes for every transmission is not a necessity. As long as an expected percentage of the nodes can be reached, individual nodes are often less concern. This feature allows the lossy links to be employed during topology control by which some applications can trade certain part of the network for higher energy efficiency. We use network-wide broadcast to represent the routing protocol, which is the best that a routing protocol can do. No congestion or collision is considered. Under this transmission paradigm, we define the metric *network reachability* to quantify the percentage of the nodes that can be reached by the sink using a sink-initialized broadcast. Based on network reachability, we propose a novel *opportunity-based topology control*, targeting at more energy-efficient topologies.

An illustration is depicted in Fig. 1. The x-axis is the energy cost and the y-axis is the network reachability. Connectivity-based topology control only uses reliable links. It seeks the optimal topology of the minimized cost and the full connectivity (point A in the figure). As this is an NPC problem, different heuristic algorithms were proposed to approach this point. Different from connectivity-based algorithms, opportunity-based topology control employs lossy links during topology control. It aims at a topology of higher energy efficiency with certain degraded network reachability (constrained by an application-dependent threshold λ_{th}). In the example of Fig. 1, given the threshold λ_{th} (0.8 in this case), point B is the target topology of opportunity-based topology control. Point C is

• *The authors are with the Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong. E-mail: {yunhuai, qianzh, ni}@cs.ust.hk.*

Manuscript received 18 Sept. 2008; revised 22 Feb. 2009; accepted 24 Feb. 2009; published online 26 Mar. 2009.
Recommended for acceptance by X. Jia.
For information on obtaining reprints of this article, please send e-mail to: tpds@computer.org, and reference IEEECS Log Number TPDS-2008-09-0358.
Digital Object Identifier no. 10.1109/TPDS.2009.57.

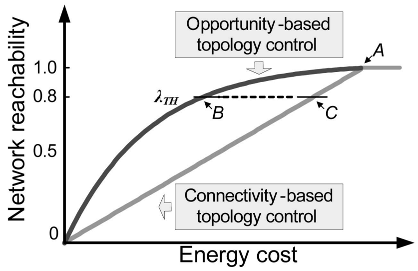


Fig. 1. Illustration of opportunity-based and connectivity-based topology control.

of connectivity-based one. The energy gap between B and C is the main motivation of this work. In some extreme applications of $\lambda_{th} = 1$, opportunity-based topology control is equivalent to the traditional connectivity-based one.

The key challenge in opportunity-based topology control is the computation of the network reachability. As we will see later (Section 2.2), it is an NP-hard problem. We do not yet have efficient algorithms to compute the network reachability accurately, even given the global network knowledge. The key issue then becomes how to guarantee a derived network topology can satisfy the optimization constraint (i.e., the network reachability is no less than the threshold λ_{th}). To address the problem, we explore the reliability theory during our algorithm design.

The main contributions of this paper are as follows. First, we identify and highlight the opportunity of using lossy links during topology control in WSNs. We formulate the novel opportunity-based topology control by defining the metric "network reachability." To the best of our knowledge, we are the first one to attempt to employ the lossy links for energy-efficient topology control purpose. Second, we prove that opportunity-based topology control is a problem of NP-hard. In order to address the problem in a practical manner, we design a fully distributed algorithm called *CONREAP* based on the reliability theory. We prove that the derived topology by CONREAP has a guaranteed network reachability, and the network energy cost is significantly reduced. The worst running time of CONREAP is $O(|E|)$, where E is the set of links in the original topology. The space requirement of individual nodes is $O(d)$, where d is the node degree. Third, in order to investigate the performance of CONREAP, we design and implement a prototype system consisting of 50 Berkeley Mica2 motes. We also conduct simulations to simulate a 200 node network. Our results show that CONREAP can improve energy efficiency of a network up to six times compared with the traditional connectivity-based algorithms.

The rest of the paper is organized as follows: In Section 2, we use an example to motivate our work. We then give an overview of opportunity-based topology control in Section 3. In Section 4, we present the proposed CONREAP algorithm and show some analytical results. Section 5 shows the implementation and simulation experiment results. We give an overview of the related work in Section 6. The last section concludes this paper and indicates possible future research directions.

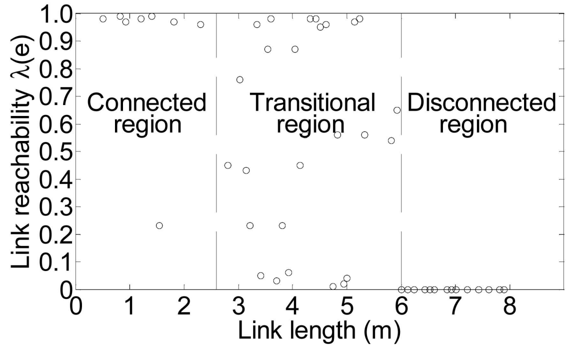


Fig. 2. Link reachability λ_e against the link length.

2 Motivation

In this section, we use an example to motivate opportunity-based topology control. We first show some empirical studies to unclose the transitional region phenomenon. We then describe the network model that takes the transitional region phenomenon into consideration. Finally, we use an example to motivate the opportunity-based topology control.

2.1 Transitional Region Phenomenon

In order to well understand the transitional region phenomenon in WSNs, we conduct a series of measurement experiments. Our testbed consists of one sink node and 50 Berkeley Mica2 nodes [8]. These nodes are uniformly deployed in an indoor environment. The sink node serves as the source and others are receivers. The transmitting power is set to be fixed -10 dbm so that the maximal reachable distance is about 5 meters. By counting the number of received packets, the receivers are able to calculate the packet reception rate of the link in between, denoted as *link reachability* λ. We do not apply any acknowledgement or retransmission which is typically in the routing design. Each run has 1,000 packets transmitted and each measurement is averaged from five independent runs.

The experiment results are shown in Fig. 2. Each plot represents the link reachability of one link. From these results, we can clearly identify the three distinctive regions. From 0 to 2.6 meters, it is the "connected" region that seven of eight links are quite reliable with link reachability more than 97 percent. Beyond 6 meters, it is the "disconnected" region that none of 15 receivers can receive any packets. The area from 2.6 to 6 meters is the transitional region which can be characterized as high variance in link reachability. In this region, 8 of 27 links are more than 95 percent reliable, six links are less than 5 percent reliable, and the left 13 links range from 20 to 90 percent. We can hardly observe a clear relation between the link reachability and the length of the corresponding link. We also conduct the experiments in outdoor environment and results are similar. These results are consistent with previous empirical studies [3], [4], [7], [30], [32], [33], [34]. It was shown [4] that in a specific setup, the transitional region was the area from 10 to 30 meters when the transmitting power is 0 dbm (1 mw). Since the transitional region is quit significant in size, many wireless links fall in the lossy link category.

2.2 Network Model

We model a WSN as a directed graph $G(V, E)$, where V is the set of nodes and E is the set of links.[1] For each link $e = \{u, v\} \in E$, we use $\lambda_e \in (0, 1]$ to indicate the probability that a packet can be successfully delivered from the node u to v, called *link reachability*. As links may fail, whether a node can be reached[2] by the sink via broadcast is a random event. Let V_r denote the set of reached nodes by the sink for a given broadcast. $|V_r|$ is thus a random variable indicating the number of such nodes. With these concepts available, we have the following definitions.

Definition (Network reachability $\lambda(G)$). *Given a network $G(V, E), \lambda(G)$ is defined as the ratio between the expected number of reached nodes by the sink using a sink-initialized broadcast, and the total number of nodes in V, i.e.,*

$$\lambda(G) = \frac{E(|V_r|)}{|V|}.$$

In order to obtain $E(|V_r|)$, we define node reachability as follows:

Definition (Node reachability $\lambda_G(v)$). *Given a node v in a network $G, \lambda_G(v)$ is the probability that the sink can reach v by a broadcast.*

By reliability theory [1], it is known that the computation of node reachability is a problem of NP-hard.

Theorem 2.1. *Network reachability $\lambda(G)$ is equal to the averaged node reachability of all nodes in the network, i.e.,*

$$\lambda(G) = \frac{\sum_{v \in V} \lambda_G(v)}{|V|}.$$

Proof. Define R_v as the indicator variable for one transmission from the sink to node v, where $R_v = 1$ if the transmission succeeds and $R_v = 0$ if otherwise. Then we have $|V_r| = \sum_{v \in V} R_v$. Thus,

$$\begin{aligned} E(|V_r|) &= E\left(\sum_{v \in V} R_v\right) \\ &= \sum_{v \in V} E(R_v) \\ &= \sum_{v \in V} \lambda_G(v). \end{aligned}$$

Note that the second equality is due to the linearity of random variable expectations. No independence is needed. □

Theorem 2.2. *The computation of the network reachability for a given network is NP-hard.*

Proof. Select any node u in G. By Theorem 2.1, we have

$$\begin{cases} \sum_{v \in V} \lambda_G(v) = |V| \cdot \lambda(G), \\ \sum_{v \in V \setminus \{u\}} \lambda_G(v) = (|V| - 1) \cdot \lambda(G \setminus \{u\}). \end{cases}$$

Combining these two, we obtain

$$\begin{aligned} \lambda_G(u) &= \sum_{v \in V} \lambda_G(v) - \sum_{v \in V \setminus \{u\}} \lambda_G(v) \\ &= (|V|) \cdot \lambda(G) - (|V| - 1) \cdot \lambda(G \setminus \{u\}). \end{aligned}$$

In other words, the computation of node reachability, a known NP-hard problem, can be reduced to the computation of network reachability in polynomial time. □

Definition (Transmitter set V_{Tx}). *Given a network G, V_{Tx} is the set of transmission nodes that have outgoing links in G, i.e.,*

$$V_{Tx} = \{u | \forall u \in V, \exists v \in V, \{u, v\} \in E\}.$$

Definition (Network energy cost $\varepsilon(G)$). *Given a network $G, \varepsilon(G)$ is the sum of the transmitting cost $\varepsilon_{Tx}(G)$ and receiving cost $\varepsilon_{Rx}(G)$, i.e.,*

$$\begin{aligned} \varepsilon(G) &= \varepsilon_{Tx}(G) + \varepsilon_{Rx}(G) \\ &= \sum_{v \in V_{Tx}} (d_G(v) + 1) \cdot \lambda_G(v). \end{aligned}$$

The network energy cost is composed of two parts: the transmitting cost $\varepsilon_{Tx}(G)$ and the receiving cost $\varepsilon_{Rx}(G)$. The energy consumption in WSNs is quite different from that in traditional wireless networks. On one hand, receiving a packet has a similar power level as transmitting it [21]. On the other hand, the consumptions of different transmission power levels are not significantly different [8], [9]. The electronic current of the minimal power is about half of the maximum power. To simplify the analysis, we assume that for any sensor node, each transmitting or receiving consumes one unit of energy. More information about the energy consumptions in WSNs can be found in [8], [9], [21], and the Tables 2 and 3.

The transmitting cost $\varepsilon_{Tx}(G)$ depends on $V_{Tx} \subseteq V$. It not only relates to the size $|V_{Tx}|$ but also the node reachability $\lambda_G(v), v \in V_{Tx}$. A transmitter of higher reachability is likely to consume more energy cost as it has more chances to forward packets. Therefore, we define: $\varepsilon_{Tx}(G) = \sum_{v \in V_{Tx}} \lambda_G(v)$.

The receiving cost $\varepsilon_{Rx}(G)$ depends on the receivers in G that have incoming links. Each transmitter v introduces $d_G(v) \cdot \lambda_G(v)$ receiving cost to the network, where $d_G(v)$ is the number of outgoing links of v. Therefore, $\varepsilon_{Rx}(G) = \sum_{v \in V_{Tx}} d_G(v) \cdot \lambda_G(v)$, where $d_G(v)$ is the node v's out degree in G.

Combining $\varepsilon_{Tx}(G)$ and $\varepsilon_{Rx}(G)$, we have the definition of the network energy cost. It is an expected value.

Definition (Reachability-energy ratio $\eta(G)$). *Given a network $G, \eta(G)$ is defined as the ratio of $\lambda(G)$ and $\varepsilon(G)$:*

$$\eta(G) = \frac{\lambda(G)}{\varepsilon(G)}.$$

1. In this paper, we use the term link and edge interchangeably, as well as node and vertex.

2. The term "reach" here means that there exists a data path from the sink to the node, along which all the links succeed during the broadcast.

TABLE 1
Notations Used in This Paper

Notation	Description
$G(V,E)$	A network topology, V is the set of nodes and E is the set of links
V_{Tx}	V_{Tx} is the set of transmitters in G: $V_{Tx} = \{u \vert \exists v \in V, \{u,v\} \in E\}$
$T_i(V,E_i)$	An intermediate tree topology; $E_i \cap E_j = \phi \Leftrightarrow i \neq j$;
$G_R(V,E_R)$	The derived topology; By CONREAP algorithm we have: $E_R = \bigcup E_i$
$\lambda_e, e \in E$	Link reachability, also written as $\lambda_{\{u,v\}}$ when $e = \{u,v\}$;
$\lambda_G(v)$	Node reachability for a node $v \in V$
$\tilde{\lambda}_G(v)$	An approximation of node reachability: $\tilde{\lambda}_G(v) = 1 - \prod(1 - \lambda_{T_i}(v))$
$\lambda(G)$	Network reachability of G: $\lambda(G) = \sum_{v \in V} \lambda_G(v)/\vert V\vert$
λ_{th}	A threshold to define the minimal required network reachablity
$\varepsilon(G)$	The energy cost of a network: $\varepsilon(G) = \sum_{v \in V_{Tx}} (d_G(v)+1) \cdot \lambda_G(v)$ where $d_G(v)$ is node v's out degree $\varepsilon(G) = \varepsilon_{Tx}(G) + \varepsilon_{Rx}(G)$
Nb_v	Node v and its neighboring set: $Nb_v = \{u \vert \forall u \in V, \{u,v\} \in E \vee \{v,u\} \in E\} \bigcup \{v\}$
$avg(\lambda_G(v))$	Local average of $\lambda_G(v)$ in Nb_v : $avg(\lambda_G(v)) = \frac{\sum_{u \in Nb_v} \lambda_G(u)}{\vert Nb_v\vert}$

The physical meaning of $\eta(G)$ is what network reachability is provided for each unit of energy. We use it to measure the network energy efficiency. This is an absolute value without normalization.

Table 1 lists some notations used in this paper.

2.3 Example

A simple example is illustrated in Fig. 3. G is the original topology with four nodes v_0, v_1, v_2, and v_3. The node v_0 serves as the sink. Solid lines denote reliable links and dash lines denote lossy links. The link reachability is labeled on top of the corresponding link.

Consider two topologies G_1 and G_2. Since the original topology G has no full connectivity, the best that connectivity-based algorithm can do is G_1 that connects v_1 and v_2 using reliable links. Different from G_1, G_2 connects v_2 and v_3 using lossy links.

In G_1, both v_1 and v_2 are reliably connected to the sink. The node reachability $\lambda_{G_1}(v_1) = 1, \lambda_{G_1}(v_2) = 1, \lambda_{G_1}(v_3) = 0$. Network reachability is thus $\lambda(G_1) = mean\{1,1,0\} = 0.67$. There are two transmitters v_0 and v_1 and two receivers v_1 and v_2. The energy cost is $\varepsilon(G_1) = \varepsilon_{Tx}(G_1) + \varepsilon_{Rx}(G_1) = 4$. The reachability-energy ratio is $\eta(G_1) = 0.67/4 = 0.167$.

In G_2, $\lambda(G_2) = mean\{1, 0.9, 0.9\} = 0.93$. There is only one sender v_0, and thus, $\varepsilon_{Tx}(G_2) = 1$. There are three receivers v_1, v_2, and v_3. Notice that a receiver cannot determine whether a transmission will succeed or not in advance. For any transmission, the receiver will pay the receiving cost no matter the transmission succeed or not. Therefore, $\varepsilon_{Rx}(G_2) = 3$, and $\varepsilon(G_2) = \varepsilon_{Tx}(G_2) + \varepsilon_{Rx}(G_2) = 3 + 1 = 4$. The reachability-energy ratio is $\eta(G_2) = 0.93/4 = 0.233$.

Certainly, there could be other topologies which obviously lack of efficiency and are out of discussion. From this example, we make following observations.

First, lossy links do have the opportunity to produce more energy-efficient topologies. Compared with G_1 of no lossy link, G_2 exhibits higher network reachability with the same energy cost. Therefore, it is the most energy-efficient one in this example. The energy efficiency $\eta(G_2)$ has about 40 percent improvement compared with $\eta(G_1)$. When

TABLE 2
Electronic Currents of Different Sensor Platforms [21]

Operations	Telos	Mica2	MicaZ
Minimum Voltage	1.8V	2.7V	2.7V
Standby (RTC on)	5.1 μA	19.0 μA	27.0 μA
Idle (DCO on)	54.5 μA	3.2 mA	3.2 mA
MCU Active	1.8 mA	8.0 mA	8.0 mA
MCU + RX	21.8 mA	15.1 mA	23.3 mA
TX (0dBm)	19.5 mA	25.4 mA	21.0 mA
Flash Read	4.1 mA	9.4 mA	9.4 mA
Flash Write	15.1 mA	21.6 mA	21.6 mA
MCU Wakeup	6 μs	180 μs	180 μs
Radio Wakeup	580 μs	1800 μs	860 μs

TABLE 3
Electronic Currents for Different Transmitting Power Levels

Modes	Power (mw)	Telos (mA)	Mica2 (mA)
-25 dbm	0.0032	8.5	N/A
-20 dbm	0.0100	N/A	5.3
-15 dbm	0.0316	9.9	7.4
-10 dbm	0.1000	11.0	7.9
-5 dbm	0.3162	14.0	8.9
0 dbm	1.0000	17.4	10.4

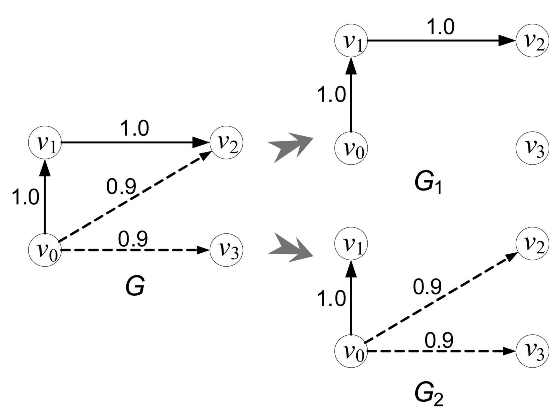


Fig. 3. An example of opportunity-based topology control; G is the original topology; G_1 is by connectivity-based and G_2 is by opportunity-based topology control.

packets are routed over such topologies, a routing algorithm of better energy utilization is expected.

Second, not all lossy links are energy-efficient. In this example, adding one more link $\{sink, v_4\}$ to G_2 will lead to a degraded reachability-energy ratio. The main reason is that although the transmitting cost can be saved by overhearing, receiving packets have the comparable energy cost as the transmitting. How to find the appropriate set of lossy links with the highest energy efficiency is one of the main issues in opportunity-based topology control.

Third, the computation of network reachability becomes critical. A practical algorithm shall be able to compute network reachability in a fully distributed manner. Centralized algorithms are costly.

3 Problem Overview

In this section, we give an overview of opportunity-based topology control. We first present the assumptions and the problem classification. We then make a formal statement of the problem, showing that this problem is NP-hard. Finally, we introduce the main challenges and research issues.

3.1 Assumptions and Classifications

In this work, we assume that individual link reachability is available, which can be obtained by sending periodic "Hello" messages, or by measuring the Link Quality Index (LQI). For the latter solution, many recent empirical studies have shown promising results (e.g., [20], [24]). We further assume that the link reachability is fixed. This assumption is reasonable as many empirical studies have shown that link reachability is pretty stable in a stationary environment. In real environments, links may be asymmetric as $\lambda_{\{u,v\}} \neq \lambda_{\{v,u\}}$. In this work, we consider only sink-to-sensors communications, which are quite useful for many applications such as query issuing, command dissemination, new network program upload [16], etc. Optimizations for other transmission patterns are left for future work. We do not consider the node failure as it can be easily transferred to the equivalent link failure. No node sleeping is considered in this work. No congestion or packet collision is considered either.

We assume that wireless links can be well controlled so that the unintended receivers are able to save receiving cost. In the example of Fig. 3, v_3 in G_1 will not pay receiving cost when the sink is transmitting, though v_3 may be in the sink's transmission range. To do so, a control field is needed in RTS/CTS packet header for CSMA-based MAC, or a scheduling-based MAC such as T-MAC is needed as we did in this paper.

It is worth noting that different applications may have different requirements on opportunity-based topology control, leading to different problems and optimal solutions. In this work, we only consider one of them, leaving more to future research work.

Here, we list three problems and focus on the first one. The others are left to future research.

Reachability-preserving problem (*RPP*). It is mainly for those data-critical applications such as disaster detections. In this problem, the objective is to minimize the energy cost while guaranteeing that the network reachability is not less than a given threshold $\lambda_{th} \in (0, 1]$.

Energy-preserving problem. This problem is mainly for those long-term monitoring and data gathering applications. In this problem, the objective is to maximize the network reachability while guaranteeing that the network energy cost is no greater than a given threshold.

Efficiency maximization problem. In this problem, the objective is to maximize the reachability-energy ratio. There is no constraint on energy or network reachability.

As mentioned, the focus of this paper is the first RPP. Now we make a formal statement of the problem.

3.2 Reachability-Preserving Problem (RPP)

The problem can be stated as follows:

Given a directed graph $G(V, E)$ and a threshold $\lambda_{th} \in (0, 1]$, assuming that $\lambda(G) \geq \lambda_{th}$, the problem is to find a subnetwork $G_R(V, E_R) \subseteq G$ such that:

$$minimize\colon \varepsilon(G_R), \tag{1}$$

$$subject\ to\colon \lambda(G_R) \geq \lambda_{th}. \tag{2}$$

Theorem 3.1. *RPP is NP-hard.*

Proof. We prove it by showing that: 1) Minimal Connected Dominating Set problem (MCDS) [31], a known NPC problem can be reduced to RPP in polynomial time and 2) $RPP \not\subseteq NP$.

1. An MCDS problem can be stated as follows. Given a connected graph $G(V, E)$, it is to find a subset $V' \subseteq V$ such that: a) V' is a dominating set and b) V' is minimal. Referring this MCDS problem instance, we construct RPP as follows: the graph G is the same; we assign reachability of every link $\lambda_e = 1$ and let the threshold $\lambda_{th} = 1$ as well. We argue that we can find the solution of the MCDS problem in polynomial time when the optimal solution of RPP G_R is available.

   a. G_R is a tree. Assume that G_R is not a tree. Then, there must be a cycle with two edges sharing the same endpoint. By removing any of the two edges, we can get another graph $G_R^{'}$. Since $G_R^{'}$ is still connected and every link is reliable, the constraint is satisfied. From G_R to $G_R^{'}$, we do not add any new transmitter. And the receivers are reduced by one. Therefore, we have

   $$\begin{aligned} \varepsilon(G_R^{'}) &= \varepsilon_{Tx}(G_R^{'}) + \varepsilon_{Rx}(G_R^{'}) \\ &\leq \varepsilon_{Tx}(G_R) + \varepsilon_{Rx}(G_R^{'}) \\ &= \varepsilon_{Tx}(G_R) + \varepsilon_{Rx}(G_R) - 1 \\ &< \varepsilon(G_R). \end{aligned}$$

   It contradicts that G_R has the minimal cost.

   b. The transmitter set V_{Tx} in G_R is the solution of the MCDS problem instance. Obviously, V_{Tx} is a dominating set. Assume that there exist other topologies with a smaller V_{Tx} than G_R. Let $G_R^{'}$ denote the one having the minimal V_{Tx}. Then, we have: $|V_{Tx}(G_R^{'})| < |V_{Tx}(G_R)|$. Obviously, $G_R^{'}$ is also a tree. Thus, we have

$$\begin{aligned}\varepsilon_{Rx}(G_R') &= \sum_{v \in V_{Tx}(G_R')} d_{G_R'}(v) \\ &= |V| - 1.\end{aligned}$$

Therefore,

$$\begin{aligned}\varepsilon(G_R') &= |V_{Tx}(G_R')| + |V| - 1 \\ &< |V_{Tx}(G_R)| + |V| - 1 \\ &= \varepsilon(G_R).\end{aligned}$$

It contradicts that G_R has the minimal cost.

2. The verification algorithm of an RPP is to verify whether a given network G can satisfy the constrain $\lambda(G) \geq \lambda_{th}$ or not. By Theorem 2.2, we know that the computation of network reachability is NP-hard. It cannot be done in polynomial time if $P \neq NP$. Therefore, the verification algorithm of RPP is not in P class and $RPP \not\subseteq NP$.

From 1) and 2), we state that RPP is NP-hard. □

3.3 Key Issue and Main Challenges

The main objective of RPP is to seek a subnetwork $G_R \subseteq G$ of the minimized cost in subject to the constraint of $\lambda(G_R) \geq \lambda_{th}$. The key issue is to guarantee that a derived topology can satisfy the constraint. To do so, it is critical to approximate the network reachability as precise as possible. It mainly involves two challenges.

The first one is to compute the node reachability. This is a known NP-hard problem for a general network [26]. The most difficult part when computing the node reachability is that the network is a so-called *coherent* system [1]. In a coherent system, any two components including inks and nodes are correlated to each other. Under such dependent scenarios, the inclusion-exclusion principle cannot be directly applied. How to compute the node reachability is the first challenge.

The second one is to compute the network reachability based on the node reachability. By Theorem 2.1, we know that the network reachability is the global average of the node reachability. Therefore, it is a challenging job to obtain the network reachability in a fully distributed manner.

These two challenges are the key problems we plan to address in the CONREAP design.

4 CONREAP Algorithm Design

In essential, CONREAP is a greedy algorithm. CONREAP starts from an empty topology. Each time, it adds the link that contributes the highest network reachability to the topology. A variation of Dijstra's shortest path algorithm is used to find such a link. CONREAP terminates when the constraint $\lambda(G_R) \geq \lambda_{th}$ is guaranteed to be satisfied. In other words, CONREAP only includes the links that are necessary to guarantee the constraint (though some links may not be really needed as CONREAP employs approximation algorithms).

In the next, we first introduce how to compute node reachability in a general network. We then present the two approximation algorithms employed by CONREAP, which are for the two challenges mentioned in Section 3.3. In the third part, we give a detailed description of CONREAP. In the last, we show some analytical results.

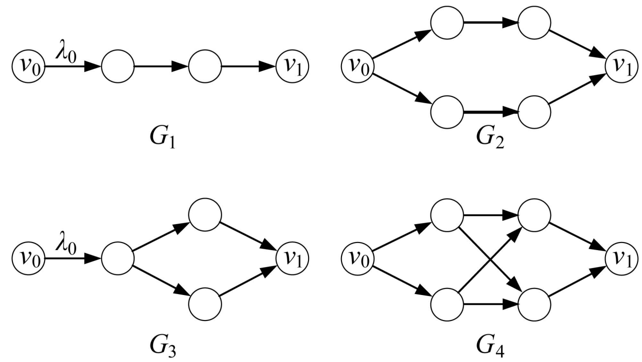


Fig. 4. Examples of node reachability computation; suppose v_0 is the sink and v_1 is the destination node.

4.1 Node Reachability Computation and Approximation in CONREAP

As mentioned before, the computation of node reachability for a general network is a known NP-hard problem.

Nevertheless, there are some specific topologies that can be computed in polynomial time [1]. By reliability theory, a general network must be converted into such forms to be practically computable. One is *series* topologies, in which there is a unique path from the source to the destination. As illustrated in Fig. 4 G_1, suppose v_0 is the sink and all the links have the same link reachability λ_0. Then, the node reachability of v_1 is $\lambda_{G_1}(v_1) = \lambda_0^3$. Note that trees are series topology as there is a unique path for any pair of nodes in a tree. Each node has only one unique path to every other individual node. Another form is *parallel* topology. An example of parallel topology is illustrated in Fig. 4 G_2. There are two link-disjoint paths from v_0 to v_1. Each path has the reachability λ_0^3. By the inclusion-exclusion principle in the probability theory, we have: $\lambda_{G_2}(v_1) = 1 - (1 - \lambda_0^3)^2$. Notice that two trees are parallel when they are link-disjoint. Currently, only simple combinations of series and parallel topologies, called *series-parallel networks*, can be computed in polynomial time. In the example of Fig. 4, G_3 is series-parallel and G_4 is not. More details of series-parallel networks can be found in [25].

There are two fundamental techniques to convert a general network into a series-parallel network. One is called pivotal decomposition or factoring. Since the running time is exponential in the worst case, this technique is not suitable for us. The other one is by approximation algorithms which try to find the maximum set of link-disjoint data paths from the source to the destination. We focus on this second technique in this paper. All the existing algorithms are, however, centralized based on the given network information. They are not applicable in fully distributed WSNs.

4.2 Approximations in CONREAP

The first approximation algorithm is to deal with the challenge of the computation of the node reachability. We compute an approximation of the node reachability instead

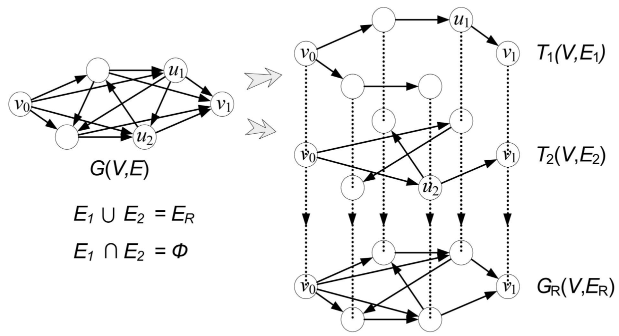


Fig. 5. An example of CONREAP; the left-up is the original topology; T_1 and T_2 are disjoint trees (no common link); they merge to derive the final topology G_R.

```
CONREAP(v)
 1. λ̃_{G_R}(v) = λ_{T_i}(v) = 0, i = 1, ...; E_R = ∅
 2. send ''Hello'' to probe neighbor set Nb_v
    measure the link reliability λ(u,v) for ∀u ∈ Nb_v
 3. do
 4.   receive tree information λ_{T_i}(u), ∀u ∈ Nb_v
 5.   u_i = arg max{λ_{T_i}(u) · λ_{{u_i,v}}} ∧ M_{{u_i,v}} = false.
 6.   λ_{T_i}(v) = λ_{T_i}(u_i) · λ_{{u_i,v}}, E_i = E_i ∪ {u_i, v}
 7.   T_k = arg max{λ_{T_i}(v), i = 1, ..}
 8.   E_R = E_R ∪ {u_i, v}; M_{{u_i,v}} = true
 9.   λ̃_{G_R}(v) = 1 − (1 − λ̃_{G_R}(v))(1 − λ_{T_k}(v))
10.   broadcast λ_{T_i}(v), i = 1...k to Nb_v
11.   exchange λ̃_{G_R}(v) with neighboring set Nb_v
12.   avg(λ̃_{G_R}(v)) = (Σ_{u∈Nb_v} λ̃_G(u)) / |Nb_v|
13. while (avg(λ̃_{G_R}(v)) < λ_{TH})
```

Fig. 6. CONREAP algorithm running on a node v in G.

of the accurate one. We uses link-disjoint trees as the building block during CONREAP design. The principle in behind is that trees are series topologies. The node reachability in a tree can be easily obtained by computing the probability along the path. When multiple trees are constructed, these paths are independent and parallel if trees are link-disjoint to each other. As such, the inclusion-exclusion principle can be applied.

To deal with the second challenge, we design another approximation algorithm to compute the global average in a fully distributed manner. The idea is that every node computes a local average from its neighbors. Observe that the global average is not less than the minimal of local averages (proved in Section 4.4). The global average must be above the threshold if all local averages are above the threshold. As such, the constraint $\lambda(G_R) \geq \lambda_{th}$ can be verified locally.

In the next, we show how we implement these two approximation algorithms in CONREAP.

4.3 Detailed Description

Fig. 6 shows CONREAP's pseudocode, and Fig. 5 gives an example of running CONREAP on a node v_1. Two trees T_1 and T_2 are constructed in this example. The union of T_1 and T_2 derives the final topology ($G_R = T_1 \cup T_2$). When deriving the final topology, CONREAP employs the greedy strategy to decide the order of the tree inclusion. A tree with higher network reachability is selected in prior to the tree of lower network reachability. In the next, we give a detailed description of CONREAP from an arbitrary node v's perspective.

Initially, v broadcasts a "Hello" message and initials its neighbor set Nb_v. At lines 3-13, v enters the main loop of CONREAP. At line 4, v receives the tree propagation information $\lambda_{T_i}(u), \forall u \in NB_v$ from Nb_v (Such information is originated from the sink, i.e., $\lambda_{T_i}(sink) = 1, i = 1...d_G(sink)$). Line 5 is the key step by which CONREAP implements the Dijstra's shortest path algorithm. The link reachability is used as the distance function. More specifically, from the known $\lambda_{T_i}(u)$, v selects a node u_i that provides v the highest tree reachability $\lambda_{T_i}(v)$ as its parent node in the tree, i.e., $u_i = \arg\max\{\lambda_{T_i}(u) \cdot \lambda_{\{u_i,v\}}\}$. One additional condition is that $M_{\{u_i,v\}} = false$, which means that the link $\{u_i, v\}$ has not yet been used. This additional condition ensures that different trees are link-disjoint. In the example of Fig. 5, v_1 selects u_1 as its parent in T_1 and u_2 in T_2. The node v then updates $\lambda_{T_i}(v) = \lambda_{T_i}(u_i) \cdot \lambda_{\{u_i,v\}}$ at line 6. The line 5 and line 6 are operated for every tree T_i. When $\lambda_{T_i}(v)$ is available, CONREAP is ready to select and include trees in G_R. Line 7 is the step of the operation. At line 7, node v selects the tree with the highest $\lambda_{T_i}(v)$ to join, denoted as $T_k = \arg\max\{\lambda_{T_i}(v), i = 1, ..\}$. Upon determining T_k, at line 8, v unions the sofar G_R with the link $\{u_i, v\}$ by storing this record in the local storage. Line 8 also marks the link $\{u_i, v\}$ as "used" ($M_{\{u_i,v\}} = true$). Line 9 is the essential part of the first approximation algorithm which implements the node reachability update function. By inclusion-exclusion principle, the node reachability is $\tilde{\lambda}_{G_R}(v) = 1 - \prod_{i=1}^{K}(1 - \lambda_{T_i}(v))$ when there are K trees. It can also be written in an iterative form (line 9) which is more suitable when K is not known in advance. At lines 10 and 11, v broadcasts all the tree information $\lambda_{T_i}(v)$ to Nb_v. v also exchanges the updated $\tilde{\lambda}_{G_R}(v)$ with Nb_v. Upon receiving the Nb_v node reachability, v updates $avg(\tilde{\lambda}_{G_R}(v))$ by line 12. This is the second approximation in Section 4.2. CONREAP terminates when $avg(\tilde{\lambda}(v)) \geq \lambda_{th}$.

It is possible that a node cannot satisfy the constraint when it joins all the available trees. In that extreme case, the original topology is kept.

4.4 Analytical Results

To show the correctness of CONREAP, in this section, we prove that $\lambda(G_R) \geq \lambda_{th}$ when CONREAP terminates. Based on CONREAP, we can state $\forall v \in V, avg(\tilde{\lambda}_{G_R}(v)) \geq \lambda_{th}$. Therefore, it is sufficient to prove that $\lambda(G_R) \geq \min\{avg(\tilde{\lambda}_{G_R}(v))\}$.

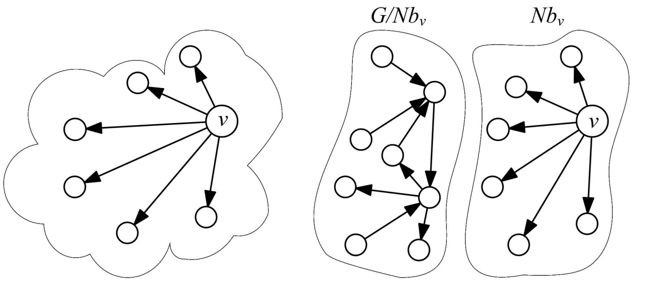
To simplify the presentation, we define *hub graph* as follows. Given an arbitrary graph $G(V, E)$, if G contains a node v that has links to every other node, v would have $avg(\lambda_G(v)) = \lambda(G)$. We call such a graph a hub graph. An example of a hub graph is illustrated in Fig. 7a.

Lemma 4.1. *$\lambda(G) \geq \min(avg(\lambda_G(v))), \forall v \in G$ when G is a hub graph.*

Proof. As G is a hub graph, $\exists u \in V, \lambda(G) = \lambda_G(v)$. If u is the minimal, then $\lambda(G) = \min(\lambda_G(v))$. Otherwise, $\lambda(G) = \lambda_G(u) > \min(\lambda_G(v))$. In both cases, the statement holds. □

Lemma 4.2. *$\lambda(G) \geq \min(avg(\lambda_G(v)))$ for any graph G.*

Proof. We use mathematic induction on the network size $|V|$.



(a) (b)

Fig. 7. Hub graph and a complement of hub graph. (a) A hub network. (b) A general network.

1. When $|V| = 1, G$ is a hub graph. Accordingly to Lemma 4.1, the statement stands.
2. Assume for any graph of size $|V| \leq k$, the statement stands. We state that for any graph G of $|V| = k+1$, the statement stands. Select any node in G denoted as v_0. Partition G to two parts: one is v_0's neighboring set Nb_{v_0} (including v_0) and the other is Nb_{v_0}'s complement set $G \setminus Nb_{v_0}$. An example of Nb_{v_0} and $G \setminus Nb_{v_0}$ is illustrated in Fig. 7b. There can be three cases:

   a. $Nb_{v_0} = \{v_0\}$ and, thus, $|V(G \setminus Nb_{v_0})| = k$. Based on the assumption, $\exists v_1 \in G \setminus Nb_{v_0}$, $\lambda(G \setminus Nb_{v_0}) > (avg(\lambda_{G \setminus Nb_{v_0}}(v_1)))$. Thus,

$$\begin{aligned}\lambda(G) &= \frac{\lambda(G \setminus Nb_{v_0}) \cdot |G \setminus Nb_{v_0}| + \lambda_G(v_0)}{|G \setminus Nb_{v_0}| + 1} \\ &\geq \min(avg(\lambda_{G \setminus Nb_{v_0}}(v_1), \lambda_G(v_0))) \\ &\geq \min(avg(\lambda_G(v))).\end{aligned}$$

   b. $Nb_{v_0} \supset \{v_0\}$, and $Nb_{v_0} \neq G$. In this case, we have $|V(G \setminus Nb_{v_0})| < k$. Thus,

$$\exists v_1 \in G \setminus Nb_{v_0}, \lambda(G \setminus Nb_{v_0}) > avg(\lambda_G(v_1)). \quad (3)$$

   As Nb_{v_0} is a hub graph, we have

$$\lambda(Nb_{v_0}) = avg(\lambda_G(v_0)). \quad (4)$$

   Combining (3) and (4), we have

$$\begin{aligned}\lambda(G) &= \frac{\lambda(G \setminus Nb_{v_0}) \cdot |V(G \setminus Nb_{v_0})|}{|V(G)|} \\ &\quad + \frac{\lambda(Nb_{v_0}) \cdot |V(Nb_{v_0})|}{|V(G)|} \\ &\geq \min(avg(\lambda_G(v_1)), avg(\lambda_G(v_0))) \\ &\geq \min(\lambda_G(v)).\end{aligned}$$

   The statement stands.

   c. $Nb_{v_0} = G$. In this case, G is a hub graph, and thus, the statement stands.

For all three cases, the statement stands. □

Lemma 4.3. $\forall v \in V, \lambda_{G_R}(v) \geq \tilde{\lambda}_{G_R}(v)$.

Proof. Recalling the definition, $\lambda_{G_R}(v)$ is the probability that the sink can reach a node v in G_R. It equals to the probability that at least one path from the sink to v works. According to our algorithm, $\tilde{\lambda}_{G_R}(v)$ is equal to the probability that at least one of the independent paths works. Therefore, $\tilde{\lambda}_{G_R}(v)$ is a subevent of $\lambda_{G_R}(v)$. According to the basic principle of probability theory, $\lambda_{G_R}(v) \geq \tilde{\lambda}(v)$. □

Theorem 4.4. $\lambda(G_R) \geq \min\{avg(\tilde{\lambda}_{G_R}(v)), v \in V\}$.

Proof. From Theorem 2.1 and Lemmas 4.2 and 4.3, we have

$$\begin{aligned}\lambda(G_R) &= \frac{\sum \lambda_G(v)}{|V|} \geq \frac{\sum \tilde{\lambda}_G(v)}{|V|} \\ &= \tilde{\lambda}(G_R) \geq \min\{avg(\tilde{\lambda}_{G_R}(v))\}.\end{aligned}$$

The two equalities are by Theorem 2.1. The first inequality is by Lemma 4.3 and the second inequality is by Lemma 4.2. □

For the worst running time of CONREAP, since different trees are disjoint to each other, a link can be added in at most one tree. In the worst case, all the links are added. Therefore, the worst running time is $O(|E|)$, where $|E|$ is the link set of the original network. The space requirement for individual nodes depends on the number of trees constructed. For each node, the maximum number of trees it can join is its degree d in the original network. Therefore, the space requirement is $O(d)$.

5 Performance Evaluation

In this section, we evaluate the performance of CONREAP. We design and implement a prototype system which consists of one sink node and 50 Berkeley Mica2 nodes [8]. These nodes are uniformly deployed in our laboratory at random (indoor). The transmission power is set to be -10 dbm so that the maximal distance is about 5 hops. Link reachability λ_e is measured using 1,000 "Hello" messages for each link. To measure the node reachability, the sink nodes broadcasts 1,000 packets and other nodes count the number of successfully received packets. Each measurement is averaged from five independent runs. In order to save the cost of unintended receivers, we adopt a scheduling-based MAC protocol T-MAC.

Our experiments apply following evaluation metrics:

- the network reachability $\lambda(G)$;
- the energy cost $\varepsilon(G)$; and
- the reachability-energy ratio $\eta(G)$.

We compared CONREAP with a recent Cone-Based Topology Control (*CBTC*) [13] algorithm which represents the connectivity-based topology control algorithm. The derived topology by CBTC is denoted as G_{CBTC}. And the topology of CONREAP is G_R.

One thing worth more words is that in practice, fewer wireless links are 100 percent reliable. If we implement CBTC using 100 percent reliable links only, G_{CBTC} can hardly satisfy the network reachability constraint because of the insufficient number of valid links. Since G_R is on the ground of satisfying the constraint, it is hard to compare the performance of G_R and G_{CBTC}. To deal with this problem, we employ a minor revised CBTC. The revised CBTC is allowed to use some high-quality lossy links based on a simple, threshold-based filtering algorithm. For instance,

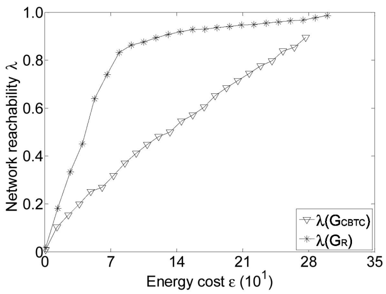


Fig. 8. $\lambda(G_{CBTC})$ and $\lambda(G_R)$ against network energy cost in the prototype system.

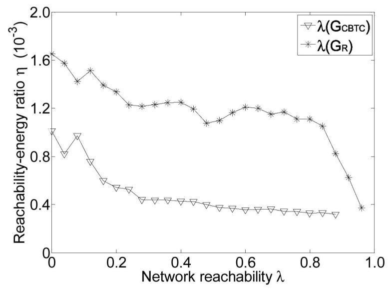


Fig. 9. $\lambda(G_{CBTC})$ and $\lambda(G_R)$ against the threshold λ_{th} in the prototype system.

CBTC with a threshold of 50 percent will consider all links of packet reception rate over 50 percent as reliable links and other links do not exist. With different threshold settings, CBTC can produce network topologies of different reachability. By adjusting the threshold setting, CBTC is able to satisfy the topology control constraint with different energy cost. Notice that in both CBTC and CONREAP, lossy links will always consume the receiving cost from receivers during a transmission no matter what threshold has been set. In order to obtain the link quality, in prototype experiments, we let receivers continuously measure the link quality and piggyback the results to senders (this is a standard technique in the current TinyOS revision). In our simulations, we assume that the link quality is available and fixed.

5.1 Prototype System Results

Fig. 8 depicts the network reachability of CONREAP and CBTC against the energy cost. As only reliable links are taken into account, CBTC has a steady performance regardless of the target network reachability. $\lambda(G_{CBTC})$ is almost linearly proportional to its cost. The upper limit of network reachability is about 90 percent since in a real setting, wireless links are hardly 100 percent reliable. In our experiments, we also implement CBTC using lossy links as reliable ones. In such settings, the network reachability can never exceed 0.2 (not shown) even when the whole network is connected. This is nature as lossy links have no guarantee of delivery. When they are used as reliable ones, the network will suffer from extremely low delivery rate. Since such low network reachability can hardly satisfy the constraint, we do not consider this approach in latter studies.

CONREAP constructs more than four trees to achieve $\lambda(G_R) \geq 0.98$. Each tree has the cost about 70 units of energy. The first tree provides $\lambda(G_R) = 0.82$, compared with $\lambda(G_{CBTC}) = 0.32$ under the same cost. To further increase the network reachability, the cost becomes high. The second tree improves $\lambda(G_R)$ from 0.82 to 0.91. And the third tree has only 0.03 improvement. These results are mainly because the first tree has used up most high reachability links. To further increase the reachability, these low reachability links have to be employed. It can also been seen from Fig. 9 that plots reachability-energy ratio against the network reachability. For $\lambda(G_R)$ from 0 to 0.8, $\eta(G_R)$ is steadily 1.2×10^{-3}, three times of $\eta(G_{CBTC}) = 0.4 \times 10^{-3}$. From 0.82 to 0.91, $\eta(G_R)$ decreases to 0.6×10^{-3} and when more than three trees are constructed, CONREAP is similar to CBTC.

Fig. 10 depicts network reachability $\lambda(G_{CBTC})$ and $\lambda(G_R)$ against λ_{th}. From this figure, we can investigate the performance of the two approximation algorithms in CONREAP. We see that since CBTC use only reliable links, the network reachability is nearly the same as the threshold. For CONREAP algorithm, first we see that it can guarantee a satisfied constraint when it terminates. However, there is a big gap between $\lambda(G_R)$ and λ_{th}. When more trees are constructed, the gap becomes larger. It implies that many links in G_R are not necessary. With better approximation algorithms, there is a further improvement space.

5.2 Simulation Results

To investigate the scalability, we also conduct simulations to evaluate CONREAP in a large scale of 200 nodes and present some representative results. We implemented a packet-level simulator which is able to scale up to thousands of nodes. We simplify the underlying MAC layer to be a probability-based transmission service provider. Wireless links have a fixed link reachability λ according to the transitional region phenomenon [34]. The nodes within $T_c = 10$ m have a reliable link. The area between T_c and $T_r = 30$ m is the transitional region that a link has a random link reachability between 0 and 1. Beyond T_c, it is the disconnected region. Notice that the simulated application field is in two dimension planes. As the connected range T_c is only $1/3$ of the reachable range T_r, only about $1/9$ of the simulated wireless links are reliable ones and the other $8/9$ are lossy links.

In our simulations, network nodes are distributed uniformly to a fixed-size field of $300 \text{ m} \times 300 \text{ m}$. The sink is set at the center. Each data are averaged from 10 independent runs. We measure the network reachability $\lambda(G_{CBTC})$ and $\lambda(G_R)$ by Monte Carlo method.

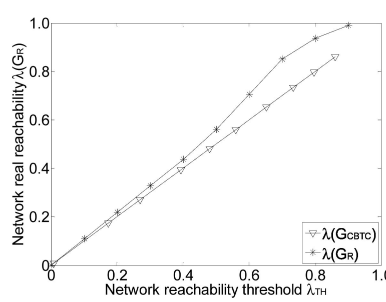


Fig. 10. $\lambda(G_{CBTC})$ and $\lambda(G_R)$ against threshold λ_{th} in the prototype system.

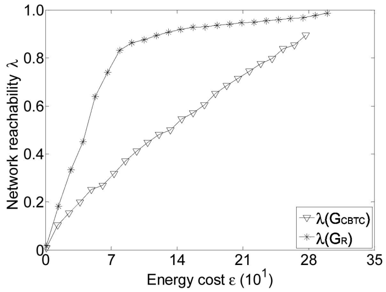


Fig. 11. $\lambda(G_{CBTC})$ and $\lambda(G_R)$ against energy cost, $|V| = 200$.

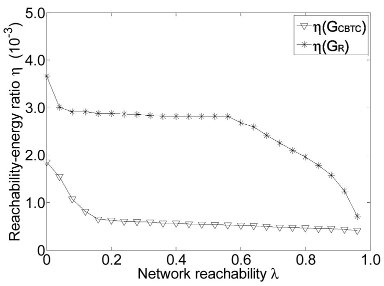


Fig. 12. $\eta(G_{CBTC})$ and $\eta(G_R)$ against network reachability, $|V| = 200$.

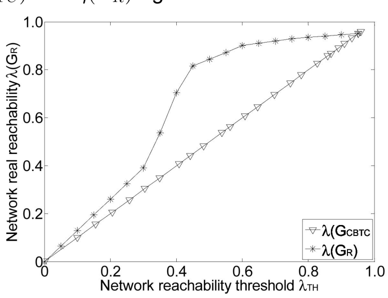


Fig. 13. $\lambda(G_CBTC)$ and $\lambda(G_R)$ against threshold $\lambda_{th}, |V| = 200$.

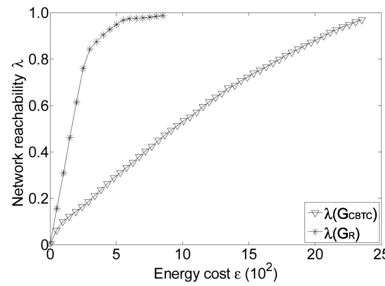


Fig. 14. $\lambda(G_{CBTC})$ and $\lambda(G_R)$ against energy cost, $(T_c, T_r) = (20, 60)$.

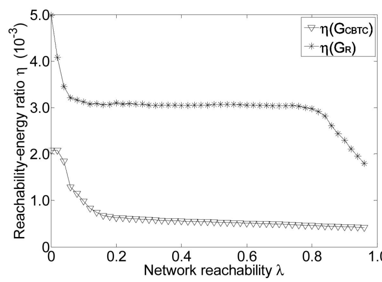


Fig. 15. $\eta(G_{CBTC})$ and $\eta(G_R)$ against reachability, $(T_c, T_r) = (20, 60)$.

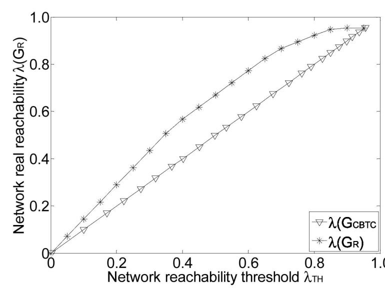


Fig. 16. $\lambda(G_{CBTC})$ and $\lambda(G_R)$ against threshold $\lambda_{th}, (T_c, T_r) = (20, 60)$.

5.2.1 Impact of Network Scale

Fig. 11 depicts $\lambda(G_{CBTC})$ and $\lambda(G_R)$ against energy cost with 200 nodes. Again, we see a nearly linear function between $\lambda(G_{CBTC})$ and its energy cost, while the performance is more stable. CONREAP exhibits a heavy-tailed-like function between $\lambda(G_R)$ and the energy cost. In the low-cost environments ($\varepsilon \leq 500$), the performance gain of CONREAP is up to four times. When high cost is allowed (e.g., $\varepsilon \geq 1{,}000$), the performance gain is much less, from 90 to 20 percent.

Fig. 12 shows reachability-energy ratio against reachability. Compared with the results in Fig. 9, we first find a higher efficiency gain. $\eta(G_R)$ is about four times of $\eta(G_{CBTC})$ in low-requirement environment ($\lambda(G_R) \leq 0.5$). However, this gain is sensitive to the required network reachability threshold. As $\lambda(G_R)$ increases from 0.5 to 0.9, $\eta(G_R)$ linearly decreases from 3×10^{-3} to 1×10^{-3}. In the latter case, CONREAP has only 90 percent gain compared with CBTC.

The results of Fig. 12 can, in some degree, be rooted to Fig. 13. In low-reachability cases from 0 to 0.4, the gap between the real and approximated network reachability is small. It implies that every link in G_R has the critical contribution to the network reachability. When the threshold is increased to more than 0.4, more trees are constructed to derive G_R. As the gap is large, many unnecessary links are in G_R. Consequently, the reachability-energy ratio has a sharp decrease.

From this set of experiments, we can conclude that CONREAP algorithm is more appropriate for low-cost, low-reachability requirement environment. When the requirement is high, CONREAP has a large gap between the really achieved and the approximation reachability.

5.2.2 Impact of Transitional Region

In this set of experiments, we vary the transitional region from $(T_c, T_r) = (10, 30)$ to $(T_c, T_r) = (20, 60)$. Fig. 14 shows $\lambda(G_R)$ against $\varepsilon(G_R)$. In this setting, we see more performance gains of CONREAP in high-reachability scenarios. When $\lambda(G_R) = 0.9$, the cost of $(20, 60)$ is about half of $(10, 30)$. It is mainly because increasing the transitional region enables more high reachability links. For CBTC, increasing the transitional region has a little gain of only 25 percent. This is consistence with the previous study of CBTC. As a result, we have Fig. 15 that depicts $\eta(G_R)$ against $\lambda(G_R)$. We also see that in the larger transitional region setting, the approximation algorithm performs better (Fig. 16). The gap between $\lambda(G_R)$ and λ_{th} is smaller than that in Fig. 13. Therefore, in this setting, CONREAP has the reachability-energy ratio gain up to six times from 0.26×10^{-3} to 1.6×10^{-3}. From this set of experiments, we can conclude that CONREAP is more appropriate for the environments of larger transitional regions.

6 Related Work

In this section, we give an overview of related works. As energy is the most concern in WSNs, many valuable efforts have been made in literature.

One category is the traditional connectivity-based topology control [23]. While the network connectivity is preserved, various energy-saving schemes have been adopted. One such scheme is by reducing the transmission power of individual nodes. As this problem is NPC, a lot of heuristics have been proposed such as MST-based [19], [22] and Cone-Based algorithms (CBTC) [13], [27]. Notice that [13] does not need location information such as [18]. In [14], the authors studied k-connectivity issues. Interference-aware topology control (e.g., [25]) focuses on interferences between consecutive hops which may result in a degradation of network throughput. All these works are based on the connectivity metric. Therefore, they are limited when leveraging the lossy links. Quality-based topology control (e.g., Xtc [29]) attempts to preserve links of higher quality. It is, however, based on connectivity metric but uses lossy links as reliable links. It will suffer from extremely low network reachability. Very interesting, a recent work [15] also noticed the impact of lossy links to the topology control. The authors studied the network reachability (but a different term) using the probabilistic model instead of the traditional deterministic model. The results are, however, still preliminary as only a specific topology is studied. The result cannot be applied to a general network. Further more, no energy issue is studied.

Another energy-saving scheme is to reduce the transmitter set by seeking the MCDS of the network. It is also a known NPC problem [10]. Many heuristic algorithms were proposed such as [31]. We prove that reachability-preserving problem is NP-hard based on MCDS problem. However, MCDS is also a connectivity-based topology control algorithm, and thus, no lossy link advantage is taken.

Designers of routing algorithms also noticed the transitional region phenomenon in WSNs. In particular, ExOR [2] and MORE [5] are proposed to explore packet overhearing along lossy links in routing. These opportunistic routing schemes are, however, based on the network topology that provided by topology control algorithm. Without appropriate support from the beneath topology control, routing algorithms can hardly find those energy-efficient routing schemes. For the example in Fig. 3, no opportunistic routing can be carried out when the beneath topology is G_2. In that sense, opportunity-based topology control aims at providing energy-efficient topology service for upper routing-layer protocols.

7 Conclusion and Future Work

In this paper, we identify and highlight the opportunity of lossy links during energy-efficient topology control. To seize this opportunity, we propose a novel opportunity-based topology control. We focus on the reachability-preserving problem and show that this problem is NP-hard. In order to address this problem in a practical way, we propose CONREAP algorithm by exploring reliability theory. We prove that CONREAP has the guaranteed network reachability and the energy cost can be significantly reduced. The worst running time is $O(|E|)$ and the space requirement is $O(d)$. Experimental results show that CONREAP is more appropriate for low-requirement, large transitional region environments. Compared with connectivity-based topology control, CONREAP can improve the energy efficiency up to six times.

Future work can be carried out along following directions. First, we study the reachability-preserving problem and leave other two problems open. Second, joint consideration with adjustable transmission power is also interesting. Third, we consider communications from sink to sensors. Optimizing topologies for other patterns are also important. Finally, some practical considerations are in concern. For example, we assume fixed-link reachability which is a function of time in practice. All these issues are calling for further investigations.

Acknowledgments

This research was supported in part by Hong Kong RGC Grants HKUST617908, China NSFC Grants 60933011 and 60933012, the National Basic Research Program of China (973 Program) under Grant No. 2006CB303000, the National Hi-Tech R&D Program of China (863 Program), and the National Science and Technology Major Project of the Ministry of Science and Technology of China under Grant No. 2009ZX03006-001.

References

[1] A. Agrawal and R. Barlow, "A Survey of Network Reliability and Domination Theory," *Operations Research,* vol. 32, pp. 298-323, 1984.

[2] S. Biswas and R. Morris, "Opportunistic Routing in Multi-Hop Wireless Networks," *Proc. ACM SIGCOMM,* 2005.

[3] A. Cerpa, J. Wong, L. Kuang, M. Potkonjak, and D. Estrin, "Statistical Model of Lossy Links in Wireless Sensor Networks," *Proc. Int'l Symp. Information Processing in Sensor Networks (IPSN),* 2005.

[4] A. Cerpa, J.L. Wong, M. Potkonjak, and D. Estrin, "Temporal Properties of Low Power Wireless Links: Modeling and Implications on Multi-Hop Routing," *Proc. ACM MobiHoc,* 2005.

[5] S. Chachulski, M. Jennings, S. Katti, and D. Katabi, "Trading Structure for Randomness in Wireless Opportunistic Routing," *Proc. ACM SIGCOMM,* 2007.

[6] J. Cheriyan, S. Vempala, and A. Vetta, "Approximation Algorithms for Minimum-Cost k-Vertex Connected Subgraphs," *Proc. Ann. ACM Symp. Theory of Computing (STOC),* 2002.

[7] K.-W. Chin, J. Judge, A. Williams, and R. Kermode, "Implementation Experience with Manet Routing Protocols," *ACM SIGCOMM Computer Comm. Rev.,* vol. 11, pp. 49-59, 2002.

[8] Chipcon. Cc1000 Data Sheet, http://focus.ti.com/lit/ds/symlink/cc1000.pdf, 2007.

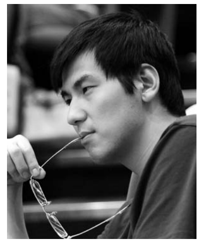
[9] Chipcon. Cc2420 Data Sheet, http://focus.ti.com/lit/ds/symlink/cc2420.pdf, 2007.

[10] V. Chvatal, "A Greedy Heuristic for the Set-Covering Problem," *Math. of Operation Research,* vol. 4, pp. 410-421, 1979.

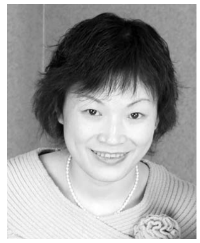
[11] A. Clementi, P. Crescenzi, P. Penna, G. Rossi, and P. Vocca, "A Worst-Case Analysis of an mst-Based Heuristic on Construct Energy-Efficient Broadcast Trees in Wireless Networks," *Proc. Ann. Symp. Theoretical Aspects of Computer Science (STACS),* pp. 121-131, 2001.

[12] A.E. Clementi, P. Penna, and R. Silvestri, "On the Power Assignment Problem in Radio Networks," *Mobile Networks and Applications,* vol. 9, pp. 125-140, 2004.

[13] R. D'Souza, D. Randall, C. Moore, and D. Galvin, "Global Connectivity from Local Geometric Constraints for Sensor Networks with Various Wireless Footprints," *Proc. Int'l Symp. Information Processing in Sensor Networks (IPSN)*, 2006.
[14] M. Hajiaghayi, N. Immorlica, and V. Mirrokni, "Power Optimization in Fault-Tolerant Topology Control Algorithms for Wireless Multi-Hop Networks," *Proc. ACM MobiCom*, 2003.
[15] M. Hefeeda and H. Ahmadi, "Network Connectivity under Probabilistic Communication Models in Wireless Sensor Networks," *Proc. IEEE Int'l Conf. Mobile Ad-Hoc and Sensor Systems (MASS)*, 2007.
[16] J.W. Hui and D.E. Culler, "The Dynamic Behavior of a Data Dissemination Protocol for Network Programming at Scale," *Proc. ACM Int'l Conf. Embedded Networked Sensor Systems (SenSys)*, 2004.
[17] M. Li and Y. Liu, "Iso-Map: Energy-Efficient Contour Mapping in Wireless Sensor Networks," *Proc. Int'l Conf. Distributed Computing Systems (ICDCS)*, 2007.
[18] M. Li and Y. Liu, "Rendered Path: Range-Free Localization in Anisotropic Sensor Networks with Holes," *Proc. ACM MobiCom*, 2007.
[19] N. Li, J.C. Hou, and L. Sha, "Design and Analysis of a mst-Based Distributed Topology Control Algorithm for Wireless Ad-Hoc Networks," *IEEE Trans. Wireless Comm.*, vol. 4, no. 3, pp. 1195-1207, May 2005.
[20] S. Lin, J. Zhang, G. Zhou, L. Gu, T. He, and J. Stankovic, "ATPC: Adaptive Transmission Power Control for Wireless Sensor Networks," *Proc. ACM Int'l Conf. Embedded Networked Sensor Systems (SenSys)*, 2006.
[21] J. Polastre, R. Szewczyk, and D. Culler, "Telos: Enabling Ultra-Low Power Wireless Research," *Proc. Int'l Symp. Information Processing in Sensor Networks (IPSN)*, 2005.
[22] R. Ramanathan, "Challenges: A Radically New Architecture for Next Generation Mobile Ad-Hoc Networks," *Proc. ACM MobiCom*, 2005.
[23] P. Santi, "Topology Control in Wireless Ad Hoc and Sensor Networks," *ACM Computing Surveys*, vol. 37, pp. 164-194, 2005.
[24] D. Son, B. Krishnamachari, and J. Heidemann, "Experimental Study of Concurrent Transmission in Wireless Sensor Networks," *Proc. ACM Int'l Conf. Embedded Networked Sensor Systems (SenSys)*, 2006.
[25] J. Tang, G. Xue, and W. Zhang, "Interference-Aware Topology Control and qos Routing in Multi-Channel Wireless Mesh Networks," *Proc. ACM MobiHoc*, 2005.
[26] L.G. Valiant, "The Complexity of Enumeration and Reliability Problems," *SIAM J. Computing*, vol. 8, pp. 410-421, 1979.
[27] R. Wattenhofer, L. Li, P. Bahl, and Y.-M. Wang, "Analysis of a Cone-Based Distributed Topology Control Algorithms for Wireless Multi-Hop Networks," *Proc. Symp. Principles of Distributed Computing (PODC)*, 2001.
[28] R. Wattenhofer, L. LI, V. Bahl, and Y. Wang, "Distributed Topology Control for Power Efficient Operation in Multihop Wireless Ad-Hoc Networks," *Proc. IEEE INFOCOM*, 2001.
[29] R. Wattenhofer and A. Zollinger, "XTC: A Practical Topology Control Algorithm for Ad-Hoc Networks," *Proc. Int'l Parallel and Distributed Processing Symp. (IPDPS)*, 2004.
[30] A. Woo, T. Tong, and D. Estrin, "Taming the Underlying Issues for Reliable Multihop Routing in Sensor Networks," *Proc. ACM Int'l Conf. Embedded Networked Sensor Systems (SenSys)*, 2003.
[31] J. Wu, M. Cardei, F. Dai, and S. Yang, "Extended Dominating Set and Its Applications in Ad-Hoc Networks Using Cooperative Communication," *IEEE Trans. Parallel and Distributed Systems*, vol. 17, no. 8, pp. 851-864, Aug. 2006.
[32] J. Zhao and R. Govindan, "Understanding Packet Delivery Performance in Dense Wireless Sensor Networks," *Proc. ACM Int'l Conf. Embedded Networked Sensor Systems (SenSys)*, 2003.
[33] G. Zhou, T. He, S. Krishnamurthy, and J. Stankovic, "Impact of Radio Irregularity on Wireless Sensor Networks," *Proc. ACM MobiSys*, 2004.
[34] M. Zuniga and B. Krishnamachari, "Analyzing the Transitional Region in Low Power Wireless Links," *Proc. Int'l Conf. Sensor and Ad Hoc Comm. and Networks (SECON)*, 2004.



Yunhuai Liu received the PhD degree from the Department of Computer Science and Engineering (CSE), Hong Kong University of Science and Technology (HKUST), in 2008. He is currently a research assistant professor at CSE, HKUST. His research interests include wireless sensor networks, pervasive computing, and wireless mesh networks. His work "Opportunity-based topology control in wireless sensor networks" was awarded the Outstanding Paper Award (1/638) in IEEE ICDCS 2008. He is a member of the IEEE.



Qian Zhang (M'00-SM'04) received the BS, MS, and PhD degrees from Wuhan University, China, in 1994, 1996, and 1999, respectively, all in computer science. She joined the Hong Kong University of Science and Technology in September 2005 as an associate professor. Before that, she was at Microsoft Research Asia, Beijing, China, from July 1999, where she was the research manager of the Wireless and Networking Group. She has published more than 200 refereed papers in international leading journals and key conferences in the areas of wireless/Internet multimedia networking, wireless communications and networking, and overlay networking. She is the inventor of about 30 pending patents. Her current research interests are in the areas of wireless communications and networking, IP networking, multimedia, P2P overlay, and wireless security. She has also participated many activities in the IETF Robust Header Compression (ROHC) WG Group for TCP/IP header compression. She is an associate editor for the *IEEE Transactions on Wireless Communications*, *IEEE Transactions on Multimedia*, *IEEE Transactions on Vehicular Technologies*, and *Computer Networks and Computer Communications*. She has also served as a guest editor for the *IEEE Wireless Communications*, *IEEE Journal on Selected Areas in Communications*, *IEEE Communication Magazines*, *ACM/Springer Journal of Mobile Networks and Applications* (MONET), and *Computer Networks*. She has also involved in the organization committee for many IEEE and ACM conferences. She received TR 100 (MIT Technology Review) world's top Young Innovator Award in 2004, the Best Asia Pacific (AP) Young Researcher Award elected by the IEEE Communication Society in 2004, and the Best Paper Award by the Multimedia Technical Committee (MMTC) of the IEEE Communications Society, and the Best Paper Award in QShine 2006, IEEE Globecom 2007, and IEEE ICDCS 2008. She received the Oversea Young Investigator Award from the National Natural Science Foundation of China (NSFC) in 2006. She is the chair of the Multimedia Communication Technical Committee of the IEEE Communications Society. She is also a member of the Visual Signal Processing and Communication Technical Committee and the Multimedia System and Application Technical Committee of the IEEE Circuits and Systems Society. She is a senior member of the IEEE.

Lionel M. Ni received the PhD degree in electrical and computer engineering from Purdue University, West Lafayette, Indiana, in 1980. He is a chair professor in the Department of Computer Science and Engineering at the Hong Kong University of Science and Technology (HKUST). He is also a distinguished professor at Shanghai Jiaotong University, an AF professor at Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, the director of the HKUST China Ministry of Education/Microsoft Research Asia IT Key Lab, and the director of the HKUST Digital Life Research Center. His research interests include parallel architectures, distributed systems, high-speed networks, and pervasive computing. He has chaired many professional conferences and received a number of awards for authoring outstanding papers. He is a fellow of the IEEE.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.