Available online at www.sciencedirect.com
SCIENCE DIRECT®
J. Parallel Distrib. Comput. 64 (2004) 839–852
Journal of Parallel and Distributed Computing
http://www.elsevier.com/locate/jpdc

Sensor-centric energy-constrained reliable query routing for wireless sensor networks ☆

Rajgopal Kannan,[a,*] Sudipta Sarangi,[b] and S. Sitharama Iyengar[a]

[a] Department of Computer Science, Louisiana State University, Baton Rouge LA 70803, USA
[b] Department of Economics, Louisiana State University, Baton Rouge LA 70803, USA

Received 1 July 2003; revised 9 December 2003

Abstract

Standard wireless sensor network models emphasize energy efficiency and distributed decision-making by considering untethered and unattended sensors. To this we add two constraints—the possibility of sensor failure and the fact that each sensor must tradeoff its own resource consumption with overall network objectives. In this paper, we develop an analytical model of energy-constrained, reliable, data-centric information routing in sensor networks under all the above constraints. Unlike existing techniques, we use game theory to model *intelligent* sensors thereby making our approach *sensor-centric*. Sensors behave as rational players in an N-player routing game, where they tradeoff individual communication and other costs with network wide benefits. The outcome of the sensor behavior is a sequence of communication link establishments, resulting in routing paths from reporting to querying sensors. We show that the optimal routing architecture is the Nash equilibrium of the N-player routing game and that computing the optimal paths (which maximizes payoffs of the individual sensors) is NP-Hard with and without data-aggregation. We develop a game-theoretic metric called *path weakness* to measure the qualitative performance of different routing mechanisms. This sensor-centric concept which is based on the contribution of individual sensors to the overall routing objective is used to define the *quality of routing* (QoR) paths. Analytical results on computing paths of bounded weakness are derived and game-theoretic heuristics for finding approximately optimal paths are presented. Simulation results are used to compare the QoR of different routing paths derived using various energy-constrained routing algorithms.

Keywords: Reliable query routing; Sensor-centric; Game-Theory

1. Introduction

Recent engineering advances in micro-miniaturization along with robust low-power hardware for processing and wireless communication have led to the development of small multi-modal sensing devices. These sensors are capable of being deployed in large numbers in a variety of extreme environments, such as seismic zones, ecological contamination sites or battlefields. Equipped with compact energy-efficient operating systems, these devices (self-) organize to form distributed sensor networks that are capable of sensing and in situ processing of spatial as well as temporally dense data over the deployment zone [1]. Information is extracted from the network through the dissemination of 'interest' queries originating from control nodes (called sinks) and resulting in responses from those sensors (called sources) whose sensed information satisfy the query attributes [9]. Sensors within the network collaborate to route queries and responses to/from sink and source nodes [3,16]. Here, the sensor network can be perceived as a reverse multicast tree with information aggregated or fused at intersecting nodes and routed to the sink node at the root. The technique of data aggregation is used to solve the problems of data implosion and overlap [15].

Sensors in wireless sensor networks operate under a set of unique and fundamental constraints which

☆This work was supported in part by NSF-ITR under IIS-0312632 and IIS-0329738 and DARPA/AFRL Grant F30602-02-1-0198.
*Corresponding author. Fax: +1-225-578-1465.
E-mail addresses: rkannan@csc.lsu.edu (R. Kannan), sarangi@lsu.edu (S. Sarangi), iyengar@csc.lsu.edu (S.S. Iyengar).

0743-7315/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jpdc.2004.03.010

make collaborative information routing challenging. These are:

1. Sensors have limited and unreplenishable power resources making energy management a critical issue in wireless sensor networks. In particular, routing protocols must be energy-aware and designed to prolong the lifetime of individual sensors (and indirectly network lifetime in terms of network connectivity/information utility). When a sensor receives a packet to be forwarded, the selection of the next-hop node must be based in part on the communication energy costs. For example, routing on the basis of minimizing aggregate energy costs on the path is one possible metric [10,17,23].
2. Sensors are unattended. Nodes must make decisions independently without recourse to a central authority because of the energy needed for global communication and latency of centralized processing. In particular, sensors must have the capacity to independently decide whether to participate in a routing path and if so, select the next-hop destination based on some (local) energy considerations.

Clearly, the untethered and unattended nature of sensors constrain their actions as individual devices, since they must independently and efficiently utilize their limited energy resources. However, designing sensor network solutions that only optimize energy consumption will not always lead to efficient architectures, since the above constraints do not account for collaborative tradeoffs between groups of sensors. Note that collaborative interaction among sensors provides some network-wide benefits (as opposed to 'energy' benefits to individual sensors), where network-wide is a semantic term referring to overall goals of the entire network or a sufficiently large group of sensors. Consider for example, collaborative data mining/information fusion among sensors to respond meaningfully to queries [2,5]. Too many sensors simultaneously participating in the collaborative decision making required for aggregation of mined data will lead to excessive routing paths in the network, thereby increasing energy consumption and competition for communication resources. On the other hand, too little collaborative data aggregation will make distributed mining inaccurate and ineffective.

This motivates the primary theme of our paper: the information utility of the sensor network (in terms of data collecting and processing ability) decreases as nodes die out. Thus sensors are implicitly constrained by a third factor: to increase information utilization of the network, sensors must cooperate to maximize network-wide objectives while maximizing their individual lifetimes. We label this paradigm for broad sensor network operation as *sensor-centric*.

The choices for untethered and unattended sensors under this paradigm are a natural fit for a game-theoretic framework. Thus in this paper, we provide an analytical model of sensor node actions in which sensors are modeled as rational/intelligent agents cooperating to find optimal network architectures that maximize their payoffs in a network game, where sensor payoffs are defined as benefits to the network of this sensor's action minus individual costs (as opposed to aggregate path costs). It is to be noted that the sensor-centric paradigm is general enough to model sensor payoffs under a variety of network game scenarios. However, in this paper we restrict ourself to routing interactions between sensors, where the network-wide benefits of the routes formed in the game correspond to quality of service metrics such as path reliability, as explained below.

There are many applications where sensors are deployed in hazardous and hostile environments in which they can fail to operate or be destroyed with certain probabilities. Wireless sensor networks are also extremely vulnerable to data loss under denial of service (DoS) attacks [25]. In these cases the task of routing a query response from observing sensors to querying nodes should not be compromised by the inhospitability of the environment. Consider sensor networks for monitoring environmentally toxic situations, or seismic sensor networks in earthquake or rubble zones or even sensors in military battlegrounds under enemy threat. For such networks to carry out their tasks meaningfully, sensors must route strategic and time-critical information via the most reliable paths available. Hence, we can introduce an additional constraint on sensor operation.

3. Sensor s_i can fail with probability $q_i = 1 - p_i$.

In this paper, we model the problem of finding maximally reliable, energy-constrained routes/trees in the network and use the sensor-centric paradigm to develop a game-theoretic metric called *path weakness*. This path performance metric is used to evaluate the quality of reliable routing trees obtained. As described later, we use the probability measure p_i primarily for ease of analytical treatment. Our results can be easily modified to use overall path reliability measures (which can be either measurement based/estimated or probabilistic) rather than using node failure probabilities.[1]

[1] Note that there are many possible alternative interpretations of p_i for different network optimization objectives. For example, path security can be modeled by assuming p_i to be the probability that a sensor node is compromized by an adversary. This compromization may mean that data passing through the sensor can be undetectably corrupted or even deleted. Thus p_i represents the degree of preference for routing information through node s_i, with $\prod p_i$ on a path the probability that the data is not received or received in a compromized fashion at the receiving end. Similarly, aggregate route quality of service measure such as delay can be used to compute payoffs and

We summarize the contributions of this paper below:

- A game-theoretic model of routing in sensor networks is developed. Rational, intelligent sensors select routing paths by evaluating the tradeoffs between reliability and the costs of communication.
- A sensor-centric paradigm for evaluating the quality of routing trees (QoR: also called tree/path weakness) for data-aggregated routing in sensor networks, is proposed. This QoR concept captures the participation suboptimality of a node on the given tree, i.e., how much would a node gain by deviating from the current tree to an optimal one. Routing heuristics based on team versions of the routing game called *Team-RQR* are presented.
- Analytical results on the complexity of computing paths with bounded weakness are derived along. Sufficient conditions on costs and probabilities for well known routing algorithms (such as most reliable path and least cost neighbor) to be congruent to the optimal sensor-centric route are also presented.
- Simulation results comparing the QoR of paths obtained using some well known routing algorithms and identifying ranges of costs and probabilities in which they perform favorably are shown.

The paper is organized as follows: Section 2 motivates the idea of path weakness by considering the problem of energy-constrained reliable query reporting[2] (RQR) in sensor networks along with some previous work in sensor network routing. Section 3 describes the details of our game-theoretic model set-up. Section 4 contains analytical as well as complexity results on path congruence and optimal path computability. Section 5 explains the quality of routing (QoR) paradigm and some theoretical QoR complexity results. Simulation results comparing the QoR of different algorithms are also presented in Section 4. Finally, Section 5 concludes the paper.

2. Reliable query reporting

Energy-constrained routing (of queries and query responses) is essential for increasing sensor network lifetime. The problem of sensor energy conservation can be addressed at multiple layers in the protocol stack. For example, [11,14,21,24] describe techniques for minimizing energy losses at the medium access control (MAC) layer. The primary goal of most energy-conserving MAC protocols is to allocate the shared wireless channels among sensor nodes as fairly as possible and ensure that no two interfering nodes transmit at the same time. In this paper, we focus on the routing layer problem of finding good (i.e. energy-efficient) routes in the network, independent of the underlying MAC layer protocol. A good choice of next-hop nodes is critical for energy-efficiency, since nodes consume transmission power proportional to physical distance to the selected next-hop neighbor [1]. Moreover, nodes that are not part of any routing path can save energy by sleeping during those periods.

While the energy-efficiency of routes is an important parameter, maximizing network information utility and lifetime implies that the *reliability* of a data transfer path from reporting to querying sensor is also a critical metric. This is especially true given the susceptibility of sensor nodes to DoS attacks and intrusion by adversaries who can destroy or steal node data [26]. The possibility of sensor node failure due to operation in hazardous environments cannot be discounted, especially for environmental monitoring and battlefield sensor network applications. From an abstract point of view, path reliability can be modeled by assigning probabilities to the compromizability of data passing through a node/node failure probabilities. Path reliability can also be measurement based, using periodic observation of DoS patterns with statistical inference tools, to determine reliability at each node.

In datacentric information routing [9,15], interest queries are disseminated through the sensor network for retrieving named data, i.e., data satisfying specific attributes. Further, data can be aggregated or combined at intersecting nodes along the route to reduce data implosion. Thus the sensor network can be perceived as a reverse multicast tree rooted at the originating node. There are many popular datacentric routing algorithms for minimizing energy consumption such as MECN [19] and diffusion routing [9], which use local gradients to identify paths for sending information. LEACH [18] proposes a clustering-based protocol that accomplishes load balancing by rotating local cluster heads. The underlying assumption of this protocol is that the cluster heads directly talk to the gateway node and the transmission power is adjustable at each node. Another protocol similar in spirit is GEAR [23], which uses an energy-aware metric and also the geographical position of each node to determine a route. Most of these algorithms attempt to minimize overall energy consumption costs. This may result in uneven energy consumption patterns across sensor nodes. Consequently, some nodes could deplete their energy resources sooner than necessary, thereby reducing the information utility of the sensor network.

In [22,20], the authors describe elegant routing algorithms for sensor networks that take energy constraints and quality of service considerations into

(*footnote continued*)
derive optimal routes/trees in the network game for certain classes of sensor networks, as we have shown in [12].

[2] We use the term query reporting and query routing interchangeably in the paper.

account. The authors in [20] show that the lowest energy path may not always be optimal for long-term network connectivity. Their scheme probabilistically uses sub-optimal paths to provide substantial gain. However, these models contrast from ours in not being sensor-centric. They do not analytically model optimal route reliability in conjunction with minimizing communication costs.

2.1. Sensor-centric reliable query reporting

In this paper, we formalize the concept of relating network-wide path performance metrics to communication energy costs in sensor networks by developing a new model of information routing. Unlike existing techniques, we use game theory to model rational/intelligent sensors thereby making our approach sensor-centric. Sensors route over the most reliable paths while minimizing their own power/energy consumption, rather than some aggregate path energy criterion. In effect, each sensor independently assumes itself as critical to the network's survival and therefore attempts to reduce its energy costs, while still satisfying network-wide objectives.

The sensor-centric paradigm of reliable energy-constrained routing has two intuitive benefits: *First*, it is in the interests of long-term network operability that nodes survive even at the expense of somewhat longer (but not excessively so!) paths. The network will be better served when a critical sensor can survive longer by transmitting via a cheaper link rather than a much costlier one for a small gain in reliability or delay. *Second*, it takes the cost distributions of individual sensors into account while choosing good paths. The advantages of modeling rational, self-interested sensors can be seen easily from the following example. Given a path involving three sensors with absolute communication costs in the low, medium and high ranges, respectively, choosing a reliable path subject to minimizing overall costs might lead to the first two nodes having to select their highest cost links as the third node is dominant in the overall cost. This would run counter to the long-term operability goal of the network.

We can now formally define the problem of reliable query reporting (RQR) in a sensor network in game-theoretic terms: Given that data transmission in the network is costly and nodes are not completely reliable, how can we induce the formation of a maximally reliable data aggregation tree from data reporting sensors (sources) to the query originating node (sink), where every sensor is 'smart', i.e., it can tradeoff individual costs with network wide benefits. This optimally reliable data aggregation tree (henceforth the optimal RQR tree) will naturally be distinct from standard multicast trees, such as the Steiner tree [6] or shortest path trees, which minimize overall network costs, and therefore cannot represent the outcome of self-interested sensors. The solution to this problem lies in designing a routing game with payoff functions, such that its Nash equilibrium [6] corresponds to the optimal RQR tree. In what follows, we define the components of this game using a model of additive data aggregation at intersecting nodes, based on information value quantification. We show that computing the optimal paths/tree (which maximizes payoffs of the individual sensors) is NP-Hard with and without data-aggregation.

This leads us to consider two important questions. First, are there easily computable routing algorithms which produce approximately optimal routing paths/trees? Secondly, in a sensor-centric network what is an approximately optimal routing path? There is as yet no formal framework for quantifying and comparing the merits of different routing algorithms in terms of the QoR paths obtained. We use the term QoR path from the game-theoretic or individual sensor's perspective rather than the well-known quality of service (QoS) based path (least cost or least delay path, for example) which is an end-to-end concept. Given the increasing prevalence of networks with 'smart' components, it is necessary to evaluate the performance gain of individual components within the overall objective. Traditional measures such as quality of service do not suffice in capturing this concept. Therefore, we require new techniques for computing the QoR of routing paths, i.e. ranking them. At a more specific level, given that the optimal path is a vector of payoffs of individual nodes, how do we characterize approximately optimal paths?

In this paper, we derive a game-theoretic path performance metric labeled *path weakness*. We use this to evaluate the suboptimality of *any* routing path in the network from the point of view of individual sensor payoffs. We propose a heuristic called Team-RQR which has low path weakness and address the following issues: How well do standard routing algorithms perform when compared to the optimal analytical solution and are there distributions of costs, probabilities and values under which some routes are 'less weaker' than others.

3. Game-theoretic RQR model

We model reliable data-centric routing with data-aggregation in sensor networks. In data-centric routing, interest queries are disseminated through the network to assign sensing tasks to sensor nodes. Attribute based naming is used to resolve these queries by using the attributes of the phenomenon to trigger responses from appropriate sensor nodes. Further, data aggregation at intersecting nodes can be used to reduce implosion and overlap problems in the network. With data-aggregation, the sensor network can be perceived as a reverse multicast tree with information fused at intersecting nodes and routed to the sink node at the root.

Let $S = \{s_1, \ldots, s_n\}$ denote the set of sensors, modeled as players in the routing game defined below, with generic members i and j. For ordered pairs $(i,j) \in S \times S$, the shorthand notation ij is used. Assume that a query has been sent from the sink node $s_q = s_n$ to the nodes in S. The query may match the attributes of data stored at each s_i to varying degrees. This data has to be reported back to s_q and aggregated along the way, if feasible. Information is routed to s_q through an optimally chosen set $S' \subseteq S$ of intermediate nodes who form neighbor communication links. Sensors in S may be initially deployed in an arbitrary topology after which they self-organize themselves into groups of neighbors [22]. Communication between neighboring sensors is implemented via an underlying MAC protocol. There is an associated energy cost to transmitting as well as receiving packets. The energy cost of transmission is proportional to the distance between sensors [1]. We abstract the ij transmission link cost metric by $c_{ij} > 0$ ($c_{ij} = \infty$ if s_i and s_j do not belong to the same neighborhood group). Note that alternate link cost metrics such as delay at the next node or link cost inversely proportional to remaining battery life can also be used. Also, for ease of presentation of our model, we assume that packet reception costs are zero. As will be seen later, incorporating non-zero reception costs is straightforward.

Our model should select data transfer paths based on the *importance* of the data being reported. For example, popular data items representing successful query matches must be treated differently and routed over more reliable paths even at higher costs, as the penalty for non-delivery is more severe. We abstract this idea of information retrieval by attaching a value $v_i \in \Re$ to the data retrieved from each sensor s_i, $1 \leqslant i < n$, ($v_i = 0$ for nodes whose sensor data does not satisfy the specified attributes of the query). Thus link formation in the network occurs by a process of simultaneous reasoning[3] at each node, leading to a path from each s_i with non-zero value v_i to s_q. We assume that node s_i can fail with a probability $(1 - p_i) \in [0, 1)$. We make no assumptions about correlations in these probabilities while formulating our abstract model, since the model primarily requires the values of path reliability, which we assume can be obtained.[4] For ease of calculation in our simulations (Section 5), we do assume independent failure probabilities. Also, for simplicity, we assume that the sink node s_q never fails.

Thus the graph $G = (S, E, P, C)$ represents an instance of a data-centric sensor network in which data of value v_i is to be optimally routed from node s_i to node s_q, with S the set of sensors interconnected by edge set E, $P(s_i) = p_i$ the node success probabilities and $C(s_i, s_j) = c_{ij}$, the cost of links in E. We denote a path from any node s_a to s_b in G by the node sequence $(s_a, s_2, \ldots, s_b)$.

There are several possible ways to model payoffs to sensor nodes, based on deterministic, probabilistic, or distributed learning algorithms for measuring path reliability, delay or data security. These have different implications on the type of resultant query reporting architectures [12]. Here we describe a simple reliability payoff model for clarity. We now describe the different components of the strategic RQR game.

Strategies: Each node's strategy is a vector $l_i = (l_{i1}, \ldots, l_{ii-1}, l_{ii+1}, \ldots, l_{in})$ and $l_{ij} \in \{0, 1\}$ for each $j \in S \backslash \{i\}$. The value $l_{ij} = 1$ means that nodes i and j have a link initiated by i whereas $l_{ij} = 0$ means that sensor i does not send information to j. The set of all pure strategies of player i is denoted by $\mathcal{L}_i$. We focus only on pure strategies in this paper. In general, node i has the option of forming or not forming a link with each of $d_i - 1$ nodes, where d_i is the degree of i in G, the number of strategies available to node i is $|\mathcal{L}_i| = 2^{d_i - 1}$. The strategy space of all nodes is given by $\mathcal{L} = \mathcal{L}_1 \times \cdots \times \mathcal{L}_n$. Notice that there is a one-to-one correspondence between the set of all directed networks with n vertices or nodes and the set of strategies $\mathcal{L}$. In order to keep the analysis tractable, in this model we assume that each sensor can only establish one link to a neighboring node. Note that while diffusion routing based algorithms start off with nodes sending query responses to the sink over multiple paths [9], eventually a single route is established once interest gradients are determined. Our objective in this paper is to compare and evaluate these final routing paths from the game-theoretic optimality point of view and hence our restriction is valid. Further, the overall strategy space $\mathcal{L}$ obviously includes routing loops. These can be avoided by ensuring that strategies resulting in a node linking to one of its ancestors yield a payoff of zero and are thus inefficient. Under these assumptions each meaningful strategy profile $l = (l_1, \ldots, l_n)$ becomes a reverse tree $\mathcal{T}$, rooted at the sink s_q. We now proceed to model the payoffs in this game.

A standard noncooperative game assumes that players are *selfish* and are only interested in maximizing their own benefits. This poses a modeling challenge as we wish to design a decentralized information network that can behave in a collaborative manner to achieve a joint goal while taking individual operation costs into account. Since the communal goal in this instance is reliable data transmission, the

[3] It can be shown for this particular game that sequential reasoning by nodes in order of selection will also produce exactly the same equilibrium paths.

[4] While we assume static failure probabilities in developing our model, a dynamic extension would view the network in terms of failure probability snapshots in successive operational periods.

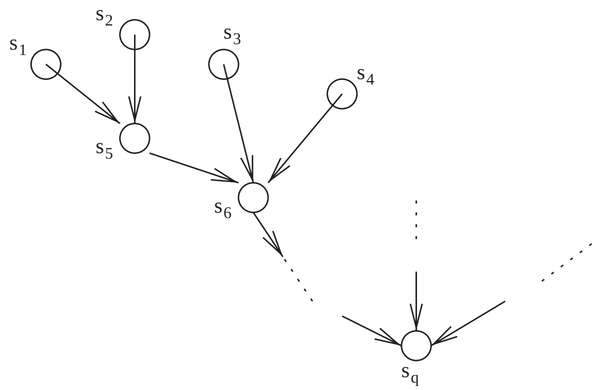

Fig. 1. Payoffs with data aggregation.

benefits to a player must be a function of path reliability but costs of communication need to be individual link costs.

Payoffs: Consider a strategy profile $l = (l_i, l_{-i})$ resulting in a tree $\mathcal{T}$ rooted at s_q, where l_{-i} denotes the strategy chosen by all the other players except player i. Since every sensor that receives data has an incentive in its reaching s_q, the benefit to any sensor s_i on $\mathcal{T}$ must be a function of the path reliability from s_i onwards. Since the network is unreliable, the benefit to player s_i should also be a function of the expected value of information at s_i. Hence we can write the payoff at s_i as

$$\Pi_i(l) = \begin{cases} g_i(v_1, \ldots, v_{n-1})R_i - c_{ij} & \text{if } s_i \in \mathcal{T}, \\ 0 & \text{otherwise,} \end{cases}$$

where R_i denotes the path reliability from s_i onwards to s_q and g_i the expectation function, is explained below.

Consider the data-aggregation tree shown in Fig. 1. Let $\mathcal{V}_i = g_i(v_1, \ldots, v_{n-1})$ denote the expected value of the data at node i and $F(i)$ the set of its parents. Then $\mathcal{V}_i = v_i + \sum_{j \in F(i)} p_j \mathcal{V}_j$, i.e., s_i gets information from its parents only if they survive with the given probabilities. The expected benefit to sensor s_i is given by $\mathcal{V}_i R_i$, i.e., i's benefits depend on the survival probability of players from i onwards. Hence the payoff to s_i is $\Pi_i = R_i \mathcal{V}_i - c_{ij}$. For example, the payoff to sensor s_5 in the figure is $\Pi_5 = R_5(v_5 + p_1 v_1 + p_2 v_2) - c_{56}$.

Definition 1. A strategy l_i is said to be a best response of player i to l_{-i} if

$$0 \leqslant \Pi_i(l_i, l_{-i}) \geqslant \Pi_i(l_i', l_{-i}) \quad \text{for all } l_i' \in \mathcal{L}_i.$$

Let $BR_i(l_{-i})$ denote the set of player i's best response to l_{-i}. A strategy profile $l = (l_1, \ldots, l_n)$ is said to be an *optimal RQR tree* $\mathcal{T}$ if $l_i \in BR_i(l_{-i})$ for each i, i.e., sensors are playing a Nash equilibrium. In other words, the payoff to a node on the optimal tree is the highest possible, given optimal behavior by all other nodes. A node may get higher payoffs by selecting a different neighbor on another tree, however, it can only do so at the cost of suboptimal behavior by (i.e reduced payoffs to) some other node(s).

Note that under the definitions above, although each sensor can form only one link, multiple equilibrium trees can exist.[5] However, it can be shown formally that *restricting* the strategy of each sensor to one link (that does not form any routing loops in the network) will eliminate trivial scenarios where any (short) path/tree forms a Nash equilibrium [12] Such equilibria are meaningless from the routing context. Thus the optimal strategy requires each node to select that node as next-neighbor, the optimal tree through which it gets the highest payoff. Given the additive nature of data aggregation, note that many of the results that hold for multiple sources are also true when considering a single source, routing to the sink. Hence, we present our results mainly in terms of single source-sink paths and when necessary the result is stated in terms of trees.

4. Results

This section contains results on two aspects of the RQR problem. We first analyze the complexity of computing the optimally reliable (or equilibrium) data aggregation tree in a given sensor network. Note that the overhead in terms of collecting network state for this protocol is quite high, since each node needs to know global path reliability and link cost values. Moreover, finding the optimal RQR tree is computationally intensive. While it can be shown that polynomial time solutions requiring limited state information for computing optimal RQR trees for large classes of graphs exist [12], in this paper, we derive some analytical results that establish congruence between the optimal RQR path and other well-known path metrics such as the most reliable path and other energy conserving paths, along with some good heuristics for almost optimal paths.

4.1. Complexity results

We begin with the following general result.

Theorem 1. *Given an arbitrary sensor network G with sensor success probabilities P, communication costs C, and data of value $v_i \geqslant 0$ to be routed from each sensor s_i to the sink s_q, computing the optimaly reliable data aggregation tree $\mathcal{T}$ (the RQR tree) is* NP-*Hard.*

[5] In the case of routing paths, payoff ties at a node can be broken by selecting the edge that lead to higher reliability. However, this is not always possible in the case of trees.

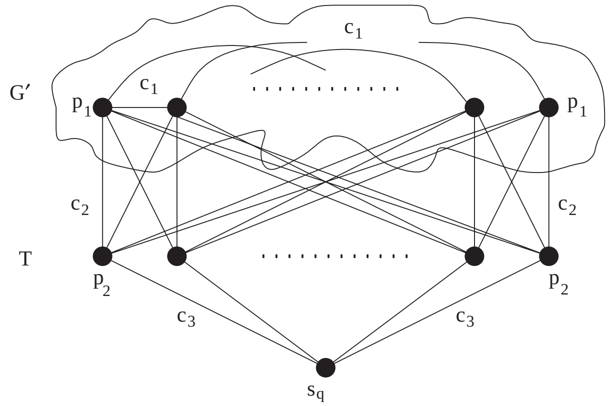

Fig. 2. Reduction from Hamiltonian path.

Proof. Given any solution $\mathcal{T}'$ to the RQR problem, verifying the optimality of the successor for each node in $\mathcal{T}'$ requires exhaustively checking payoffs via all possible trees to s_q. Thus RQR does not belong to NP. That the RQR problem is NP-Hard follows by reduction, using the following lemma which considers the special case of finding an optimal path, given a single source. (Note that this is equivalent to finding routing trees without data-aggregation.) □

Lemma 1. *Let $\mathcal{P}$ be the optimal RQR path for routing data of value v_r from a single reporting sensor s_r to the sink node s_q in a sensor network G where $v_i = 0\ \forall i \neq r$. Computing $\mathcal{P}$ is NP-Hard.*

Proof. As before, verifying the optimality of the successor for each node in $\mathcal{P}$ requires exhaustively checking payoffs via all possible paths to s_q. Thus the RQR path problem does not belong to NP.

We show that the problem is NP-Hard by considering a reduction from Hamiltonian path [7]. Let $G' = (V', E')$ be any graph in which a Hamiltonian path is to be found, where $|V'| = n$. We convert G' into another graph $G = (S, E, P, C)$ on which an instance of RQR path with value $v_r = 1$, must be computed[6] as shown in Fig. 2.

Introduce $n+1$ new vertices to form $S = V' \bigcup T \bigcup s_q$, where $|T| = n$ and s_q is the other new vertex. The new edge set E consists of the original edge set E' along with n^2 new edges from $E_2 = T \times V'$ and n new edges from $E_3 = T \times s_q$. Edges in E', E_2 and E_3 are assigned costs c_1, c_2 and c_3, respectively. All vertices $u \in V'$ and $w \in T$ are assigned success probabilities p_1 and p_2 respectively. The relationships between the probabilities and costs are as follows:

$$p_1 p_2 > \left(\frac{3}{4}\right)^{\frac{1}{n-1}}, \tag{1}$$

$$c_1 = \frac{(p_1 p_2)^n}{3}, \tag{2}$$

$$c_2 = \frac{2(p_1 p_2)^n}{3}, \tag{3}$$

$$c_3 = (p_1 p_2)^n. \tag{4}$$

Let s_r and s_t be any two nodes in V'. We claim that there exists an optimal RQR path of reliability $p_1^n p_2$ from s_r to s_q in G if and only if there exists Hamiltonian path from s_r to s_t in G'.

For the first part of the claim, assume there is a Hamiltonian path $\mathcal{H} = (s_r, \ldots, s_t)$ in G'. Consider the path $\mathcal{H}$ followed by the edges (s_t, x) and (x, s_q) in G', where x is any node in T. This path has reliability $R(\mathcal{H}) = p_1^n p_2$. The payoff of node s_t is $R(\mathcal{H}) - c_2$ obtained by linking to node x, which is optimal since there does not exist any other unvisited node in V'. Similarly the payoff of node x is also optimal since it can only link to s_q. Now consider the kth node in $\mathcal{H}$, $1 \leqslant k \leqslant n-1$. The two choices for this node are either to link to some node $x \in T$ or the node in G' that lies on the Hamiltonian path $\mathcal{H}$. If the first option is chosen, the most reliable alternate path (and hence the maximum possible alternate payoff) is given by $p_1^k p_2 - c_2$ which is less than $R(\mathcal{H}) - c_1$ by conditions (1)–(3). Thus, the second choice is optimal for this node.

For the second part of the claim, we need to show that if no Hamiltonian path exists in G', there cannot be an optimal RQR path of reliability $p_1^n p_2$. Note that linking to any available node in V' with cost c_2 is always preferable for any node $s_i \in T$. The worst case payoff to s_i via a link of cost c_2 is $p_1^n p_2^n - c_2$, which outweighs the best possible payoff via a link of cost c_3 which is $p_1 p_2 - c_3$. So the optimal path *must* visit all nodes in V'. To maximize payoffs, the optimal path must have the shortest possible length. This will require minimizing visits to T. The optimal path will thus consist of sequences of long paths in V' (the longest possible since any node in V' will always prefer to link to another node in V', if feasible), interspersed with visits to T. Since G' does not contain a Hamiltonian path there will be at least two visits to nodes in T and hence the reliability of such a path will be at most $p_1^n p_2^2$ which is less than $p_1^n p_2$ as claimed. □

It can be seen easily that the above reduction is still valid when all nodes in V' and T have the same success probability p. Consequently, the RQR path and tree problems remain NP-Hard for the special case when nodes have equal success probabilities. The case when

[6] We set $v_r = 1$ for notational simplicity since results for any v_r can be obtained by scaling edge costs appropriately.

all edges have the same cost is much simpler, however, as will be shown below.

4.2. Analytical results

Given the complexity of computing the optimal RQR tree, we try to analytically derive conditions that establish congruence between the optimal and other well known, easily computable trees, such as the most reliable tree (i.e., the union of the most reliable paths) from sources to the sink and energy-conserving trees. Identifying these conditions on network parameters will save the overhead of computing optimal (or approximately optimal) RQR trees in these cases. For simplicity, we present these congruence results in terms of paths from a single source to the sink; the results can be easily extended to trees.

Let G be an arbitrary sensor network with a single source node having data of value v_r ($v_i = 0$ for all other nodes). Then the following results hold. Note that the results describe only sufficient conditions for congruence with the optimal path.

Observation 1. *Given $p_i \in (0, 1]$ and $c_{ij} = c$ for all ij, then the most reliable path (tree) always coincides with the optimal RQR path (tree). For uniform p_i, the equilibrium RQR path is also the path with least overall cost.*

Before proceeding further, we now introduce some notation. For any node s_i, let $c_i = \{c_{ij}\}$, $c_i^{\max} = \max\{c_{ij}\}$ and $c_i^{\min} = \min\{c_{ij}\}$. Also $c^{\max} = \max_i\{c^i_{\max}\}$ and $c^{\min} = \min_i\{c^i_{\min}\}$. We use $\mathcal{P}_i^l$ to denote a path of length l from s_i to s_q and benefits along this path by $\mathcal{P}_i^l$.

Proposition 1. *Given G and $P(s_i) = p \in (0, 1]$, for all i, the most reliable path from s_r to s_q will also be the optimal path if*

$$c_i^{\max} - c_i^{\min} < v_r p^m (1 - p)$$

for all s_i on the most reliable path $\mathcal{P}_r^m$.

Proof. Consider an arbitrary node s_i at a distance i from s_r. Since we have uniform p, reliability is now inversely proportional to path length. Let l be the length of the shortest path from s_i to s_q, on which s_{i+1} is the next neighbor of s_i. For s_i, $\mathcal{P}_i^l$ is optimal if

$$v_r p^{i+l} - c_{ii+1} > v_r p^{i+l+\lambda} - c_{ij}, \quad \lambda = 1, 2, \ldots$$
$$\Rightarrow \frac{c_{ij} - c_{ii+1}}{v_r} < p^{i+l}(1 - p^{\lambda}),$$

where s_j is a neighbor of s_i through which there is a simple path of length $l + \lambda$. Since $m = i + l$ on $\mathcal{P}_r^m$, the reliability term above is minimized for $\lambda = 1$, whereas the cost term is maximized at $c^i_{\max} - c^i_{\min}$. □

Note that the above result identifies sufficient constraints on costs for the most reliable path to also be optimal. The result shows that while the most reliable path can be costlier than other paths, to be optimal it cannot be 'too' much more expensive. From the above result, it also follows that when $c^{\max} - c^{\min} < p^m(1 - p)$ this path coincides with the optimal, thereby providing a global bound on costs for congruence. The equivalent result for the most reliable tree can be obtained by substituting $\mathcal{V}_i$, the expected aggregated data value at s_i, for v_r in the above proposition.

We now look at the situation when the probabilities of node survival are non-uniform. Let s_i and s_{i+1} be subsequent nodes on the most reliable path. Denote by R_i, the reliability of the most reliable path from s_i to s_q with R_i' being the reliability along any alternative path from s_i. Let $\Delta c_i = c_{ii+1} - c_{ij}$ where s_j is any neighbor not on the optimal path and ΔR_i is defined similarly.

Proposition 2. *Given G and $P(s_i) = p_i \in (0, 1]$, the most reliable path from s_r to s_q will be optimal if*

$$\frac{\Delta c_{i+1}}{\Delta c_i} < \frac{\Delta R_{i+1}}{\Delta R_i}$$

for all s_i and s_{i+1} on the most reliable path.

Proof. Let $\bar{R}_i$ represent the reliability on the portion of the most reliable path $\overline{\mathcal{P}}$ from s_r to s_i. Since $\overline{\mathcal{P}}$ is optimal, s_i cannot benefit by deviating if

$$v_r \bar{R}_i R_i - c_{ii+1} > v_r \bar{R}_i R_i' - c_{ij}$$
$$\Rightarrow v_r \bar{R}_i > \frac{\Delta c_i}{\Delta R_i}.$$

It follows that $v_r \bar{R}_{i+1} > \frac{\Delta c_{i+1}}{\Delta R_{i+1}}$. Since $\bar{R}_{i+1} = p_{i+1}\bar{R}_i$, we have $v_r p_{i+1} \bar{R}_i > \frac{\Delta c_{i+1}}{\Delta R_{i+1}}$. This can be rewritten as $1 \geqslant p_{i+1} > \frac{\Delta c_{i+1}}{\Delta c_i} \frac{\Delta R_i}{\Delta R_{i+1}}$, which gives us $\frac{\Delta c_{i+1}}{\Delta c_i} < \frac{\Delta R_{i+1}}{\Delta R_i}$ as desired. □

The easiest way to interpret this result is by rearranging the terms so that we can write it as $\frac{\Delta c_{i+1}}{\Delta R_{i+1}} < \frac{\Delta c_i}{\Delta R_i}$. Then each fraction can be interpreted as the marginal cost of reliability of deviating from the optimal path. Since each subsequent node on the optimal path has lower expected value of information, this results suggests that the marginal cost of deviation in terms of reliability must be higher for each node's ancestor where the expected value of information is also higher.

We define the cheapest neighbor path (CNP) from s_r to s_q as the simple path obtained by each node choosing its successor via its cheapest link (that connects to s_q). In a sense, this path reflects the route obtained when each node has only limited network state information (about neighbor costs and probabilities), and in the absence of gradient information or route quality feedback, should merely minimize its local communication costs. The following proposition identifies when CNP will coincide with the optimal path.

Proposition 3. *Given G and* $P(s_i) = p \in (0, 1)$, *for all i, the optimal RQR path is at least as reliable as the cheapest neighbor path. Furthermore, the CNP will be optimally reliable if*

$$\min\{c_k \backslash c_k^{\min}\} - c_k^{\min} > v_r p^l (1 - p^{t-l}),$$

where l is the length of the shortest path from s_r *to* s_q *and t is the length of the CNP.*

Proof. Consider an arbitrary node s_k which is k hops away from s_q on the CNP. Clearly, for the CNP to be optimal s_k should not get higher payoff by deviating to an alternative path. Also, we do not need to consider alternative paths that have lengths greater than k to s_q since that would decrease benefits and the CNP already has the lowest cost edges. Let m be the path length along the CNP from s_r to s_k. For alternative paths of length $i = 1, \ldots, k-1$, from s_k to s_q to be infeasible, we need

$$c_i > c_o + v_r p^{m+i}(1 - p^{k-i}),$$

where c_o is the edge cost along the CNP, and c_i the edge cost along alternative paths. By definition, for any node on the CNP $m + i \geqslant l$. Also at s_k we have $c_o = c^k_{\min}$, with c_i being at most $\min\{c^k \backslash c^k_{\min}\}$. Thus, when $\min\{c^k \backslash c^k_{\min}\} - c^k_{\min} > v_r p^l (1 - p^{t-l})$, the CNP will coincide with the optimal path. □

The above proposition illustrates that the CNP does not have to be the most reliable in order to be optimal, it only needs to be sufficiently close. For networks in which some paths (edges) are overwhelmingly cheap compared to others, routing along CNPs may be reasonable. However, in networks where communication costs to neighbors are similar, routing based on local cost gradients is likely to be less reliable.

5. Quality of routing

We divide this section into two subsections. In the first subsection, we present our route evaluation metric and some theoretical results. The second half provides heuristics with low path weakness followed by simulation results about the quality of routes obtained using different routing algorithms. Throughout this section, we assume that there is a single source and destination pair. Thus results are presented in terms of paths instead of trees.

5.1. Evaluation metric

In an ideal sensor-centric network, optimal RQR paths are computed by individually rational sensors who maximize their own payoffs. On the other hand, traditional routing algorithms optimize using a single (end-to-end) distinguishing attribute such as total cost or overall latency.[7] From a sensor-centric perspective these approaches are inadequate and sub-optimal since they use a single network wide criterion. How then do we compare different suboptimal paths? For example, one path may yield high payoffs for sensor i with low payoffs for sensor j, while the exact opposite situation may prevail on another path. Clearly in a framework where rational, independent sensors maximize their own payoff subject to the overall network objective, we need a new metric for evaluating the quality of different paths from an individual sensor's point of view. We introduce a metric called path weakness which captures the suboptimality of a node on the given path, i.e., how much a node would have gained by deviating from the current path to an optimal one. We believe this provides a new sensor-centric paradigm for evaluating the quality of routing in sensor networks.

We formally define our QoR metric as follows: Let $\mathcal{P}$ be any given path from the source sensor s_r to the sink node s_q. Assume that the source contains information of value v_r and all other nodes have value $v_i = 0$. Consider any node s_i on $\mathcal{P}$ with ancestors $\{s_r, \ldots, s_{i-1}\}$. Let $\hat{\mathcal{P}}_{iq}$ be the optimal RQR path for routing information of value $\mathcal{V}_i = v_r \prod_{t=r}^{i} p_t$ (i.e., the expected value) to s_q from s_i in the subgraph $G \backslash \{s_r, \ldots, s_{i-1}\}$, assuming such a path exists. Thus $\hat{\mathcal{P}}_{iq}$ represents the best that node s_i can do, given the links already established by nodes $s_r, \ldots, s_{i-1}$ and assuming optimal behavior from nodes s_i onward, downstream. Define $\Delta_i(\mathcal{P}) = \Pi_i(\hat{\mathcal{P}}_{iq}) - \Pi_i(\mathcal{P})$ as the payoff deviation for s_i under the given strategy profile (path) $\mathcal{P}$. A negative deviation represents the fact that s_i is benefiting more from this path (perhaps at the expense of some other sensor). Conversely, a positive deviation indicates s_i could have done better. We set $\Delta_i(\mathcal{P}) = v_r$ whenever $\Pi_i(\mathcal{P})$ is negative. This positive deviation from the optimal payoff is intended to represent the fact that s_i is participating in a path which is giving it negative payoffs, i.e., the communication cost on the edge out of s_i in $\mathcal{P}$ outweighs the benefits to s_i of participating in this route. Also note that it is possible that no optimal path from s_i exists, even if its payoff on $\mathcal{P}$ is positive. For example, all of s_i's neighbors might have very high communication costs and cannot participate in any optimal path, making s_i in a sense isolated. In such cases, we set $\Delta_i(\mathcal{P}) = -\Pi_i(\mathcal{P})$.

$\bar{\Delta}(\mathcal{P}) = \max_i \Delta_i(\mathcal{P})$ represents the payoff deviation at the node which is 'worst-off' in $\mathcal{P}$. What can be said

[7] See [22] however, for an elegant model in which the authors develop data-centric routing algorithms for sensor networks that take both energy constraints and quality of service considerations into account. However, the model contrasts from ours in not being sensor-centric.

about this parameter for optimal and sub-optimal paths?

Observation 2. *$0 < \bar{\Delta}(\mathcal{P}') \leqslant v_r$ for all non-optimal paths $\mathcal{P}'$.*

However, observe that $\Delta_i(\mathcal{P}')$—the weakness of individual nodes on sub-optimal paths can take both positive and negative values. On the other hand, $\bar{\Delta}(\mathcal{P}) = 0$ if and only if $\mathcal{P}$ is the Nash equilibrium path of the game. Thus from a global point of view, $\bar{\Delta}(\mathcal{P})$ identifies the maximum degree to which a node on the path can gain by deviating. This allows us to rank the 'vulnerability' of different paths, which embodies the idea that a path is only as good as its weakest node. We label this QoR measure *path weakness*.

Note that the weakness metric can be similarly defined for data-aggregation trees. Given a sensor on any tree $\mathcal{T}$, its weakness can be calculated as its payoff deviation from the optimal tree that would have been obtained, given the expected value at that sensor along with the distribution of values in the remaining nodes in the graph. As mentioned before, we focus on single-source single-destination paths in the rest of this paper.

We now present bounds for finding paths with low path weakness. We state in the following theorem that there exist networks not containing paths of bounded weakness.

Theorem 2. *For an arbitrary sensor network, there exists no polynomial time algorithm to compute approximately optimal RQR paths of weakness less than $(\frac{v_r}{3} - \varepsilon)$ unless $P = \mathrm{NP}$.*

The proof relies on constructing a specific sensor network whose best polynomial time computable sub-optimal paths satisfy the above weakness characteristics. Details of the proof are in [13].

5.2. Path weakness heuristics

Theorem 2 indicates the feasibility of finding approximately optimal RQR paths of bounded weakness. While this problem still remains open for arbitrary sensor networks, it can be shown that polynomial time solutions requiring limited state overhead exist for computing optimal RQR paths/trees for geographically routed sensor networks [12]. Here, we present some easy to compute heuristics based on a team version of the RQR game (called TRQR), for finding approximate RQR paths. Simulation results presented in the next subsection verify that the team-RQR heuristic has low path weakness and compares favorably with other standard routing algorithms.

The TRQR path can be interpreted game-theoretically as a 'team' version of the RQR game in which all nodes on the path share the payoff of the worst-off node on it. Rather than selecting a neighbor to maximize their individual payoffs as in the original game, nodes in the team-RQR model compromize by maximizing their least possible payoff. As before, each sensor's strategy is to select at most one next-neighbor (if the payoffs exceed its participation cost). Choices resulting in routing loops have zero payoffs. Formally, the payoffs to nodes in the network are defined as follows:

$$\Pi_i(l) = \begin{cases} v_r R(\mathcal{P}) - \max\limits_{(s_i,s_j)\in\mathcal{P}} c_{ij} & \text{if } s_i \in \mathcal{P}, \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

where $R(\mathcal{P})$ is the reliability of path $\mathcal{P}$ from s_r (with value v_r) to s_q formed under strategy choice l. The Nash equilibrium of the TRQR game is the path from source to destination containing the node with the highest least cost-reliability tradeoff over all paths. In case of multiple equilibria, the path with highest reliability is selected.

Formally, let $\overline{\mathcal{P}}_c$ represent the most reliable path from s_r to s_q that does not traverse any link exceeding cost c. Then the optimal TRQR path $\overline{\mathcal{P}}$ is given by

$$\overline{\mathcal{P}} = \arg\max_{c_i \in C}\{v_r R(\overline{\mathcal{P}}_{c_i}) - c_i\} \tag{6}$$

for each distinct edge cost c_i in C. $\overline{\mathcal{P}}$ can be computed by repeatedly determining the most reliable path in the graph that is obtained by successively removing edges of decreasing distinct cost. In the worst case m most reliable path calculations are made, where m is the number of distinct edge costs in the network.

5.3. Experimental results

In this section, we simulate the performance of different routing algorithms to answer the following question: What are the quality of paths compared to that of the optimal RQR path? This allows us to identify the different ranges of node reliabilities and edge costs in which a particular algorithm performs better than the others.

The setup for our experiments is as follows: In every iteration a random graph with 20 nodes and edge density of 30% is generated. The source and destination pair are randomly chosen and the value of data at the source node is normalized to one. For each run, we choose a node survival probability, which is identical for all nodes. Communication costs over each edge are drawn randomly from a given parameter range in every iteration. For each set of node success probabilities and edge costs, we have presented results for 15 different source and destination pairs (we have verified that this is a representative sample). In each simulation run, for a particular source and destination pair, routing paths are generated by several algorithms and the corresponding path weakness (QoR) is calculated. The data have been used to construct graphs which are presented at the end of the paper. We have used the following algorithms:

1. *Most reliable path* (MRP): This produces the most reliable path from source to the sink. Since, in our setup, each node has the same success probability the MRP is always the shortest path as evaluated by Djikstraa's standard shortest path algorithm.

2. *Overall cheapest path* (MCP): This algorithm is also Dijkstraa's shortest path algorithm, with the weight of each edge being the communication cost.

3. *Cheapest next node path* (CNP): This provides a path where each node chooses its cheapest available edge leading to the sink node.

4. *Team RQR path* (TRQR): This is calculated as defined previously.

5. *Genetic algorithm path* (GA): Here, we use a genetic algorithm for solving the optimal RQR problem based on the GA for the bicriteria shortest path problem provided in [8]. A path has been encoded according to the priority-based method. In this procedure, a set of n random numbers (n being the total number of sensor nodes) is generated so that the ith random number is the priority of the ith node. A path is sequentially constructed led by the highest priority feasible nodes, i.e., nodes which do not lead to a dead end or a cycle. The genetic operators used here are position-based crossover and swap mutation. A next generation is chosen by tournament method. We stop if the difference between the fitness values of the best paths of two adjacent generations is equal to zero.

The first three algorithms are standard routing algorithms. The fourth algorithm is our heuristic derived from a game theoretic point of view. Genetic algorithm is a standard technique applied to problems which are NP-complete or NP-hard. We have used it here to check if there is any range of node success probabilities and costs where it does well.

6. Interpretation of results

Our simulation results are illustrated in Figs. 4–7.

In the first five graphs, nodes are assigned very high success probabilities. Edge costs are low and chosen from a distribution such that every path is feasible (all node payoffs are positive). In case I and II, we keep the node success probability fixed at 0.99 and vary the maximum edge cost from 0–0.05 to 0–0.01, respectively (Fig. 3).

In case I, the path weakness ranges from 0 to 0.6. MCP and TRQR have average weaknesses 0.08 and 0.05, respectively. Since the cost range and hence the cost differences among various edges are not significantly large, all three cost-based algorithms (TRQR, MCP and CNP) that try to reduce the overall cost in different ways behave reasonably well. However, the range of path weakness of MRP (0–0.4) suggests that the cost range is so high that a path which relies solely on maximizing reliability (MRP) cannot perform well.

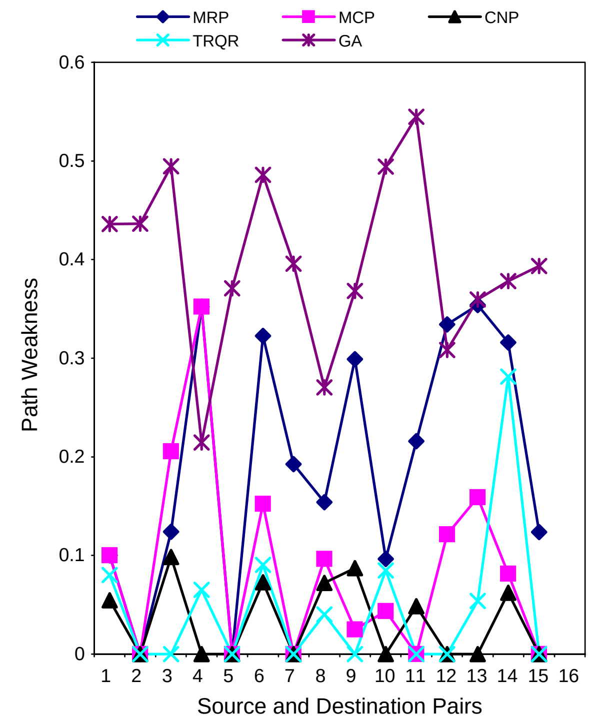

Fig. 3. Case I: $p = 0.99$, $c \leqslant 0.05$.

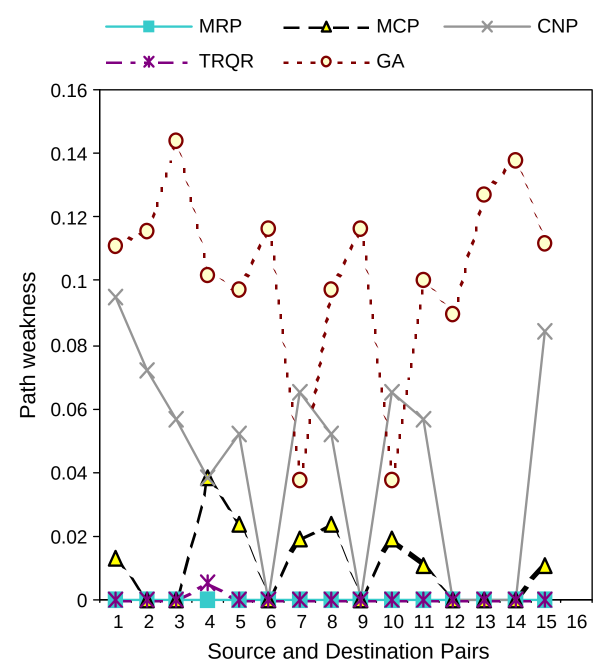

Fig. 4. Case II: $p = 0.99$, $c \leqslant 0.01$.

In case II, the maximum edge cost is reduced to 0.01. Consequently, the overall range of path weakness reduces to 0–0.14. Significant improvement takes place in the behaviour of MRP and TRQR as they coincide

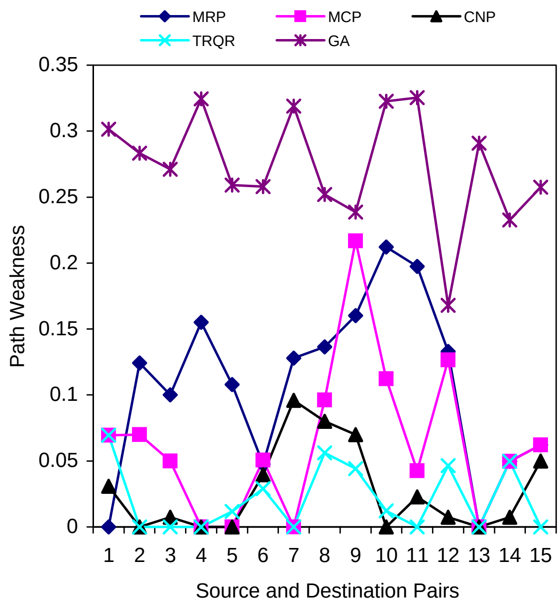

Fig. 5. Case III: $p = 0.992$, $c \leqslant 0.12$.

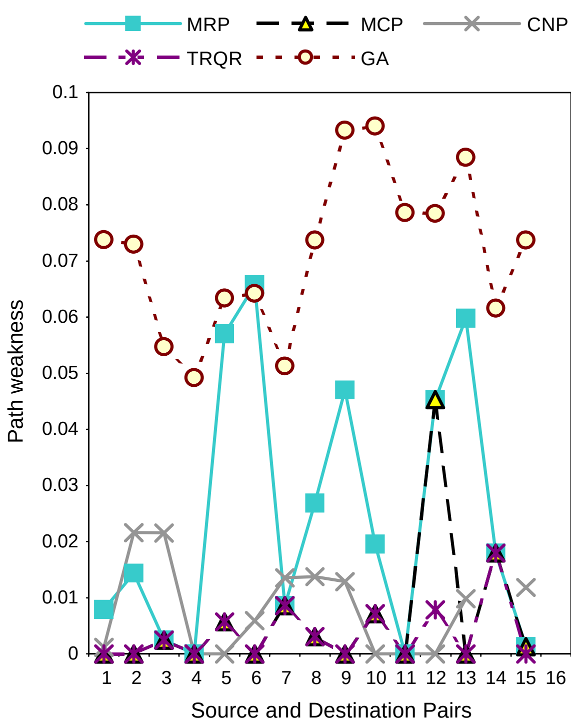

Fig. 6. Case IV: $p = 0.998$, $c \leqslant 0.058$.

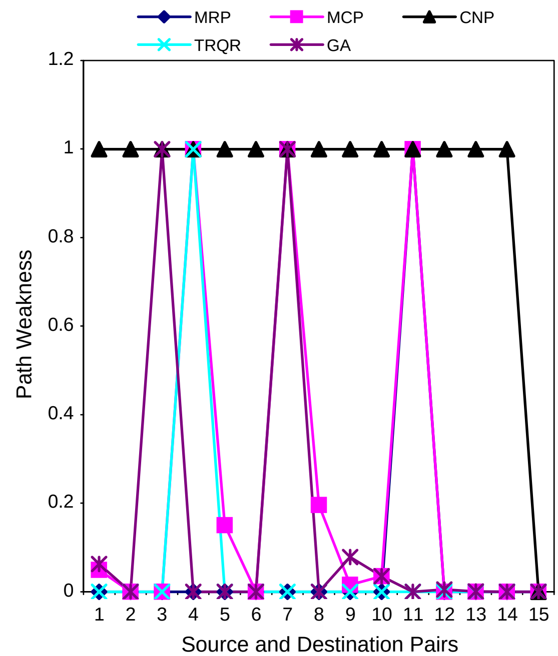

Fig. 7. Case V: $p = 0.5$, $c \leqslant 0.065$.

with the optimal path for more than 90% of the source and destination pairs. The fact that MRP always coincides with the optimal path indicates that the very high node success probability and very small cost range together have reduced the length of the optimal path. The diminished variation within different edge costs allows MCP to perform well. Since the behaviour pattern and the range of path weakness of CNP do not vary significantly from case I to case II, we can conclude that performance of CNP is invariant over a large cost range when reliability is kept very high.

For Cases III–V, we make the maximum edge cost a decreasing function of the node success probability. Then, we slowly increase node success probability to observe the impact. In case III, where the node success probability is 0.992 and the cost range is 0–0.12, the range of path weakness is quite high (0–0.35). When we raise the value of the success probability, the optimal paths can have longer lengths without sacrificing too much reliability. Therefore CNP, which tends to have a longer length, has lower path weakness now (average weakness being 0.035 approximately). The TRQR heuristic, which tradesoff both the overall path reliability and the overall cost performs as well as CNP producing an average path weakness of 0.32. The above mentioned feature of the optimal path can also explain MRP's unstable pattern and the high range of path weakness in spite of very high node success probabilities. In case IV, the success probability is increased to 0.998 and the cost range is reduced to 0–0.058. This accounts not only for the relatively small range of path weakness (0–0.1) but also for the good performance of MCP, CNP and TRQR. The congruence of TRQR and MCP is well explained by the significantly large difference between the success probability and the maximum edge cost. In case V, we explore the consequences of restricting the likely optimal path length using one low node success probability (0.5) and maximum edge cost 0.065. MRP, the shortest path, always coincides with the optimal path even though the success probability is quite low. So do TRQR and MCP. However, since the CNP usually has longer path lengths, its QoR is quite weak, in most cases.

When we compare the first 5 graphs, we observe that the increment in the node success probabilities together with the decrement in the maximum edge costs gradually leads to improvements in the behavior of all five algorithms. In general, MRP will be a good heuristic for obtaining good QoR paths only when path reliabilities are low. The behaviors of TRQR and MCP are quite stable (with a little variation in the weakness ranges) in all the ranges of our experiment and on average, provide better QoR. CNP provides good QoR when the success probability increases and the maximum edge cost decreases accordingly.

7. Conclusion

In this paper, we formulate a sensor-centric model of intelligent sensors using game theory. The problem of routing data in such a network is studied under the assumption that sensors are rational and act to maximize their own payoffs in the routing game. Further, nodes in our model are susceptible to failure and each node has to incur costs in routing data. To evaluate the contribution of individual nodes in the routing tree, we develop a metric called path weakness. This individual-sensor oriented evaluation criteria provides a new paradigm for examining paths, which we label QoR. While the optimal routing problem has high state overhead and is computationally hard, our experimental results show that standard path routing mechanisms like MRP and MCP usually find reasonably good paths. Our game-theoretically oriented algorithm—Team RQR compares favorably to the other standard routing algorithms.

For future work, we plan to develop bounded, approximately optimal RQR paths/trees for general sensor networks (the problem is still open), along with extensions using distributed and cooperative game models. Polynomial time solutions for the optimal RQR and delay constrained paths/trees are presented in [12] for special classes of sensor graphs. We also plan to investigate the efficiency and practicality of implementing optimal RQR protocols in hierarchical (clustered) sensor networks. Is it beneficial to compute optimal paths within each cluster for routing to gateway nodes (that handle inter-cluster routing).

Acknowledgments

The authors are thankful to Ms. Lydia Ray for assisting with the simulation programs and figures.

References

[1] I.F. Akyldiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, Wireless sensor networks: a survey, Comput. Networks 38 (4) (March 2002) 393–422.

[2] R.R. Brooks, C. Griffin, D. Friedlander, Self-organized distributed sensor network entity tracking, Internat. J. High-Performance Comput. Appl. 16 (3) (2002).

[3] R.R. Brooks, S.S. Iyengar, Fundamentals of Multisensor Fusion, Prentice-Hall, Englewood Cliffs, NJ, 1997, p. 458.

[4] M. Chu, H. Haussecker, F. Zhao, Scalable information-driven sensor querying and routing for ad hoc heterogeneous sensor networks, Xerox-PARC Technical Report P2001-10113, July 2001.

[5] D. Fudenberg, J. Tirole, Game Theory, MIT Press, Cambridge, MA, 1991.

[6] M.R. Garey, D.S. Johnson, Computers and Intractability, Freeman Publishing, New York, 1979.

[7] M. Gen, R. Cheng, Genetic Algorithms and Engineering Optimization, Wiley/Interscience Publication, New York, December 1999.

[8] C. Intanagonwiwat, R. Govindan, D. Estrin, Directed diffusion: a scalable and robust communication paradigm for sensor networks, in: Proceedings of the Sixth Annual International Conference on Mobile Computing and Networks (MobiCom 2000), August 2000, Boston, MA.

[9] D.B. Johnson, D.A. Maltz, Dynamic source routing in ad hoc wireless sensor networks, Mobile Computing, Kluwer, Dordrecht, 1996.

[10] C.E. Jones, K.M. Sivalingam, P. Agrawal, J.C. Chen, A survey of energy efficient network protocols for wireless networks, Wireless Networks 7 (4) (July 2001) 343–358.

[11] R. Kannan, S.S. Iyengar, Game-theoretic models for reliable path-length and energy-constrained routing with data aggregation in wireless sensor networks, IEEE J. Selected Areas Comm., in press.

[12] R. Kannan, S. Sarangi, S.S. Iyengar, L. Ray, Sensor-centric quality of routing in sensor networks, in: Infocom 2003 San Francisco, CA, April 2003.

[13] R. Kalidindi, R. Kannan, L. Ray, S.S. Iyengar, Energy efficient MAC protocols for wireless sensor networks, in: International Conference on Wireless Networking (ICWN03), Las Vegas, June 2003.

[14] B. Krishnamachari, D. Estrin, S. Wicker, Modeling data-centric routing in wireless sensor networks, in: Proceedings of the INFOCOM 2002, New York, June 2002.

[15] G. Pottie, Hierarchical information processing in distributed sensor networks, in: Proceedings of the International Symposium on Information Theory, August 1998, pp. 163–168.

[16] C. Perkins, E. Royer, Ad hoc on demand distance vector routing, Proceedings of 2nd IEEE Workshop Mobile Computing Systems and Applications, February 1999.

[17] W. Rabiner Heinzelman, A. Chandrakasan, H. Balakrishnan, Energy-efficient communication protocol for wireless microsensor networks, Proceedings of the 33rd International Conference on System Sciences, 2000.

[18] V. Rodoplu, T.H. Meng, Minimum energy mobile wireless networks, IEEE J. Selected Areas Comm. (JSAC) 17 (8) (August 1999) 1333–1344.

[19] Rahul C. Shah, Jan M. Rabaey, Energy aware routing for low energy ad hoc sensor networks, Proceedings of IEEE Wireless Communications and Networking Conference (WCNC), Orlando, FL, March 2002.

[20] S. Singh, C.S. Raghavendra, PAMAS: power aware multi-access protocol with signaling for ad hoc networks, ACM Comput. Comm. Rev. 28 (3) (July 1998) 5–26.

[21] K. Sohrabi, J. Gao, V. Ailawadhi, G. Pottie, Protocols for self-organization of a wireless sensor network, IEEE Personal Comm. October 2000, pp. 16–27.
[22] Yan Yu, R. Govindan, D. Estrin, Geographical and energy aware routing: a recursive data dissemination protocol for wireless sensor networks, UCLA Computer Science Department Technical Report UCLA/CSD-TR-01-0023, May 2001.
[23] Wei Ye, J. Heidemann, D. Estrin, An energy-efficient MAC protocol for wireless sensor networks, Proceedings of 12th IEEE International Conference on Computer Networks, INFOCOM 2002, New York, USA, June 2002.
[24] A. Wood, J. Stankovic, Denial of service in sensor networks, IEEE Comput. (October 2002) 54–62.
[25] L. Zhou, Z. Haas, Securing ad-hoc networks, IEEE Network 13 (6) (1999) 24–30.