Context Aware Routing in Sensor Networks

Melanie Hartmann[1], Holger Ziekow[2], and Max Mühlhäuser[1]

[1] Telecooperation Department
Darmstadt University of Technology
Hochschulstraße 10, D-64289 Darmstadt, Germany
[2] Institute of Information Systems
Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin, Germany

Abstract. Sensor networks are a prominent representative of ubiquitous computing technologies. The limited resources of sensor nodes and the low reliability of wireless communication pose special challenges for message routing in sensor networks. This paper discusses how context awareness can enhance routing in sensor networks. Further, we analyze the network properties which influence the benefit of context awareness in the routing process. The presented simulation results allow assessing which networks are best suited for such an approach.

1 Introduction

Over the past few years, a range of routing protocols tailored to sensor networks have been proposed [2]. In a number of protocols, routing paths are chosen dependent on the importance of a message [6,7,3]. Yet, little work has been done on the issue of how to determine the importance of a given message.

In this paper, we investigate the utilization of context information for assessing message importance. In section 2 of this paper, we review existing routing schemes for sensor networks and discuss how message importance can influence routing decisions. In section 3, we present our approach for using context awareness to assess message importance in the routing process. In section 4, we describe experiments that investigate the influence of network parameters on the benefit that can be achieved by using context aware routing and outline future work in section 5.

2 Importance Based Routing in Sensor Networks

In many applications, the spatial network size exceeds the communication range of a sensor node. Thus, messages need to be forwarded in multiple hops from node to node through the network. Numerous strategies for multi-hop message routing in sensor networks have been suggested in the last few years. Akkaya and Younis [1] categorize prominent approaches in data-centric, hierarchical, location-based as well as network flow and QoS-aware protocols. Several protocols that consider the network flow and some kind of QoS-parameters have been proposed. These approaches aim at finding routing graphs to optimize parameters such as overall network lifetime, message delay, or communication reliability while balancing those parameters against network resources.

The context aware approach that we present in this paper exploits that resource usage can be reduced by sending less important messages along paths with low communication costs, but probably also low reliability, or by even dropping them. We formalize the decision which path to choose by a payoff function $p_A(i) = r_A \cdot i - c_A$ with i being the importance of the message, r_A being the reliability of path A and c_A the cost of sending a message via A. With this formalization we lean on routing metrics as presented in [6]. The communication costs c are directly proportional to the path length counted in hops from sensor node to sensor node. The reliability r depends on the spatial distance between nodes on the path. At each hop, the link reliability is dependent on the distance between the communicating nodes [8]. As r depends on the spatial distances and c on the number of hops along a path, the payoff is determined by the geometric structure of the network. If p_A and p_B intersect at values for i with $i \in [0, i_{max}]$ (i_{max} being the maximal importance value), a choice between paths A and B is made, since either A or B is favourable depending on the message importance. However, if p_A and p_B intersect outside the interval $[0, i_{max}]$, one path dominates the other for all importance values. This raises the question in which kind of networks, messages would actually be routed along different paths if message importance is considered. Another question is which network properties influence the average cost between the considered paths. We address these questions in experiments presented in section 4.

3 Using Context to determine Message Importance

To determine the importance of a message to be routed, the sensor network needs knowledge about the usage of the data and about its semantics. The importance of a message can depend on a variety of different factors like the priority of the query, the current context or the importance of the cluster that reported the message. In this paper we focus only on context-dependent importance. In this section, we describe how a knowledge base needed for that purpose can be structured and how the message importance can be determined using context. A knowledge base is stored on every sensor node, so that the importance computation can be performed locally. Each node determines the route for the message or even drops it dependent on the estimated importance of the message.

3.1 Knowledge Base

The structure of the knowledge base leans on the theory of conceptual spaces coming from cognitive science [5]. Conceptual spaces deal with the problem of modeling representations and present an alternative to the two approaches currently dominating this domain: symbolic representations (e.g., used in the Semantic Web) and associationism (basing on artificial neuronal networks). Conceptual spaces are spanned by a number of quality dimensions, so that every entity belonging to the modeled concept is represented by a point in this space. For example, colors are represented by their hue, saturation and brightness.

In sensor networks, we most likely deal with linear quality dimensions like temperature or humidity. Deshpande et al. use correlations between attributes (which map to quality dimensions in our model) for estimating the selectivity of predicates in queries [4]. Our approach enables to use such correlations for estimating message importance. The sensor nodes store refinements of quality dimensions for the general scale depending on context data. For example, the global quality region for temperature comprises all possible temperature values and can contain subspaces depending on further context values like time. The available refinements depend on the accuracy needed and on the available memory of a sensor node. We assume, that context characteristics $\mathscr{C}_i$ for each subregion can be sufficiently characterized by stating its maximum and minimum value and by approximating its density distribution with the normal distribution specified by its mean and its standard deviation.

3.2 Determining Message Importance

For estimating the importance of a message we need a formalization for the context values and the message to be routed. We define the current context C as a set of context-types c_i that are associated with a domain D_i and a value x_i ($C = \{(c_i, x_i) | x_i \in D_i\}$, e.g., $C = \{(c_{temp}, 20°C)\}$). A message m consists of a context-type and the corresponding value ($m = (c_m, x_m)$ with $x_m \in D_m$). We illustrate the computation of the message importance in the following with an example: *Given a sensor network that can measure temperature and time data. We want to determine the coldest temperature within the next 10 hours (12:00 to 22:00) using a global median to reduce the influence of outliers.*

To specify a query, we have to state

- **Result Type** c_r**:** specifies the context-type of the query's result.
- **Selection criteria** S**:** Specifies when a sensor value is of interest and refers to selection predicates in query languages. The selection criteria S are defined as a set of restrictions on different context-types: $S = \{(c_i, a, b) | a,\, b \in D_i \cup \{\pm\infty\}\}$ with a, b being the upper and lower bound.
- **Importance function** imp**:** maps the content of the message $m = (c_m, x_m)$ and the current context C to the importance of m: $imp : x_m, C \mapsto [0, i_{max}]$. The user can define a function himself or use a predefined function, e.g., MIN, MAX, AVG, RAD_CHG (monitoring of radical changes) or DEV_AVG (monitoring of deviation from average).

In our example $c_r = c_{temp}$, $S = \{(c_{time}$, *12:00, 22:00*$)\}$ *and* $imp = imp_{MIN}$.

The knowledge base KB is defined as a set of functions s_i. Each s_i maps a set of tuples (c_j, x, y) to context characteristics $\mathscr{C}_k$ (with $j \neq k$). Tuples passed as parameters to a function s_i define a subspace in the conceptual space. Each tuple

(c_j, x, y) denotes the interval between x and y along the quality dimension c_j. The corresponding return value $\mathscr{C}_k$ is the context characteristics for this subspace. As stated before, we describe the characteristics of a context c_i with its mean, its standard deviation, its minimum and maximum value: $\mathscr{C}_k = (a_\mu, a_\sigma, a_{min}, a_{max})$. However, also any other context description could be used.

Having specified the query and the knowledge base, we have to determine how to compute the importance of a message m with respect to the current context. The current context is either sensed by the node itself or it is acquired by sniffing on forwarded data. If the message or the current context does not match the selection criteria S or is not of type c_r of any query currently processed in the network, the message is dropped. Otherwise, the importance of the message is determined by the importance function *imp*. If the importance is under a specified threshold, the message is also discarded. We can distinguish three ways to calculate the message importance:

- **Blind**: does not consider the content of the message, every message is assigned to the same importance (e.g., AVG)
- **Global view**: considers the expected overall run of the results for the query (including past and future values) (e.g., MIN, MAX). For that purpose, the characteristics of the expected result only have to be computed once for every query. Thus, the overall context defined by the selection criteria S is considered, but not the current context C of each particular value.
- **Local view**: rates the importance of the message only depending on the current values (e.g., DEV_AVG, RAD_CHG). The expected results have to be calculated for every rating, considering the current context C.

Within the function *imp*, the expected characteristics $\mathscr{C}_r$ are determined using the functions stored in the knowledge base KB considering S and C. This may require to use several characteristic measures because there is no single s_i that describes the expected values best. The (simplified) importance function for imp_{MIN} is defined as follows: $imp_{MIN}(x_m, C) = i_{max} - \frac{x_m - \mathscr{C}_r(a_{min})}{\mathscr{C}_r(a_{max}) - \mathscr{C}_r(a_{min})}$ for $\mathscr{C}_r(a_{min}) \leq x_m \leq \mathscr{C}_r(a_{max})$. The importance values calculated with respect to every considered function s_i have to be combined using a weighting function.

In our example, we need a global view for calculating the overall minimum. The selection criterion is limited to the context-type time, so we just consider s_1. *The result for* $s_1(\{(c_{time}, 12{:}00, 22{:}00)\})$ *is* $(16°C, 3, 12°C, 20°C)$. *Hence, if* $i_{max} = 1$, *the message* $m = (c_{temp}, 18°C)$ *would be assigned the importance 0.25.*

4 Experiments

In this section we describe experiments to investigate the influence of network characteristics on the utility of context awareness in routing. In particular, we analyse the relation between the spatial network size, node density and the potential benefits of evaluating the message importance. This should help to decide for which networks the context aware routing approach is beneficial. We measure the benefit as the average cost difference between reasonable routing paths. That is, each particular node may choose different paths with different communication costs for communicating messages of different importance to the sink. As indicator for potential benefits we take the maximal cost difference between paths from each node to the sink and calculate the average value for the whole network. Formally that is $\frac{\sum_{n \in N} C(n,i_{max}) - C(n,0)}{|N|}$ where N is the set of nodes in the sensor network and $C(n,i)$ is the communication cost from node n to the sink along the path with the best payoff for messages of importance i. We use this measure because the benefit of respecting message importance is limited to the maximal cost difference between considerable paths.

For the experiments, we created random networks in a simulation environment. Nodes of these networks were randomly spread across square virtual spaces. Connections and link reliabilities between the nodes were established based on their spatial distance. Therefore, we used a rough approximation of the measurements presented in [8].

We tested networks with 5 to 200 nodes created on virtual planes with edge sizes from 75 to 400 units of length. Overall 26000 randomly created networks were evaluated in the experiments. Figure 1 shows the benefit measures depending on the number of sensor nodes and the spatial size of the network. The tests show the average benefit changes dramatically within the analyzed scope. In general, the benefit increases with the node number and the spatial size of the network. Yet, it is crucial to consider both parameters in combination.

The right side of figure 1 shows cuts through the parameters space with a fixed node number (bottom) and fixed spatial size (top). Cuts at different

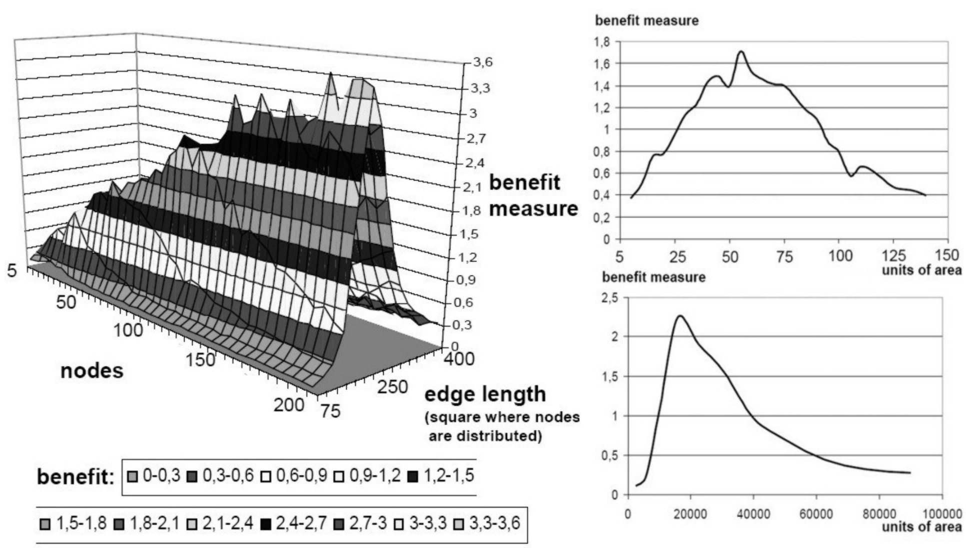

Fig. 1. Benefit measure depending on network parameters

positions in the parameters space result in curves of similar shapes. Given a fixed node number the benefit curve ascends rapidly with the spatial size of the network until a maximum is reached and descends less intense afterwards. For planning applications this means: the benefit of applications where sensors are distributed over large areas is easier to estimate. This is due to the fact that parameter changes have less effect for these networks than for dense networks. Changing the node number at a fixed network size results in a convex curve for the benefit measure as well. However, the curve does not show any crucial parts with dramatic changes in the benefit measure. Thus, confidence in benefit estimations is mainly dependent on the spatial network size and is generally higher for large scale networks.

5 Future Work

In our tests we reveal the dependencies of the spatial network size and the number of sensor nodes on the benefit that can potentially be obtained by context aware routing strategies. This enables us to identify networks which are suitable for the context aware routing model we propose. In future tests, we plan to apply our approach on sample applications with real world data. The tests will target the difference between the actually gained benefit and the predicted maximal benefit and will allow benchmarking our context aware routing approach against existing routing strategies.

References

1. Kemal Akkaya and Mohamed F. Younis. A survey on routing protocols for wireless sensor networks. *Ad Hoc Networks*, 3(3):325–349, 2005.
2. J. N. Al-Karaki and A. E. Kamal. Routing techniques in wireless sensor networks: a survey. *Wireless Communications, IEEE*, 11(6):6–28, 2004.
3. Budhaditya Deb, Sudeept Bhatnagar, and Badri Nath. Information assurance in sensor networks. In *2nd ACM international conference on Wireless sensor networks and applications*, pages 160–168, New York, NY, USA, 2003. ACM Press.
4. Amol Deshpande, Carlos Guestrin, Wei Hong, and Samuel Madden. Exploiting correlated attributes in acquisitional query processing. In *ICDE*, pages 143–154, 2005.
5. Peter Gärdenfors. *Conceptual Spaces*. MIT Press, Cambridge, Massachusetts, 2000.
6. Rajgopal Kannan, Sudipta Sarangi, and S. Sitharama Iyengar. Sensor-centric energy-constrained reliable query routing for wireless sensor networks. *J. Parallel Distrib. Comput.*, 64(7):839–852, 2004.
7. Yogesh Sankarasubramaniam, Özgür B. Akan, and Ian F. Akyildiz. Esrt: event-to-sink reliable transport in wireless sensor networks. In *MobiHoc*, pages 177–188, 2003.
8. Alec Woo, Terence Tong, and David E. Culler. Taming the underlying challenges of reliable multihop routing in sensor networks. In *SenSys*, pages 14–27, 2003.