Exposé zur Diplomarbeit
Processing Multiple Aggregation Queries in Dynamic Wireless Sensor Networks

betreut von: Timo Gläßer, Prof. Ulf Leser
bearbeitet von: Björn Schümann

September 2008 - Februar 2009

1 Motivation

In recent years Wireless Sensor Networks have become an important and useful tool for a wide area of applications. A Wireless Sensor Network (WSN) consists of a bulk of sensor nodes, small battery powered devices equipped with a set of sensors and a wireless communication interface. WSNs have been used in applications like monitoring the habitats of birds [SPMC04], the structural health of buildings [KPC+07] or the areas near eruptive volcanoes [WAJR+05]. First commercial applications outside of the scientific world start to use wireless meshed sensor networks to support us in daily life problems, like searching for the next parking spot [Dou07].

The networks are getting bigger. An increasing number of sensor nodes provides a finer resolution of data points and the prices for the nodes are falling. Sharing workload between nodes provides opportunities to extend the lifetime of the network and reduces the risk of data loss in case of failing nodes. Large multipurpose networks have to be flexible enough to deliver the required data for several users with different needs. Using the sensor network as a distributed query processing system, that uses the processing power of the nodes in the network, provides the flexibility needed to cope with the changing requirements without reprogramming the nodes.

Most of the the application scenarios need aggregated results over the sensor network. Users disseminate a declarative aggregation query into the network from sink nodes, e.g. to get the average or an extremum of all sensor readings. Those queries are usually restricted to a specific region of the network. Nodes inside the region of interest start to measure a specific environment parameter using their sensor devices and work together with the neighbour nodes to calculate the aggregated result before propagating this result to the appropriate sink node. Multiple queries might have overlapping query regions making it possible to share partial results. Since senor nodes are battery powered, saving energy becomes an important factor. Radio communication is usually the most important consumer. Hence minimizing the communication needed to get and forward the correct results of all queries to the users becomes a important consideration in designing an efficient query processing system.

Large-scale networks have to be flexible enough to handle changes in the network topology, caused for example by failing and moving nodes. Furthermore recent studies [ZHKS04] have shown that radio communication is highly irregular, which results in non-isotropic radio properties and

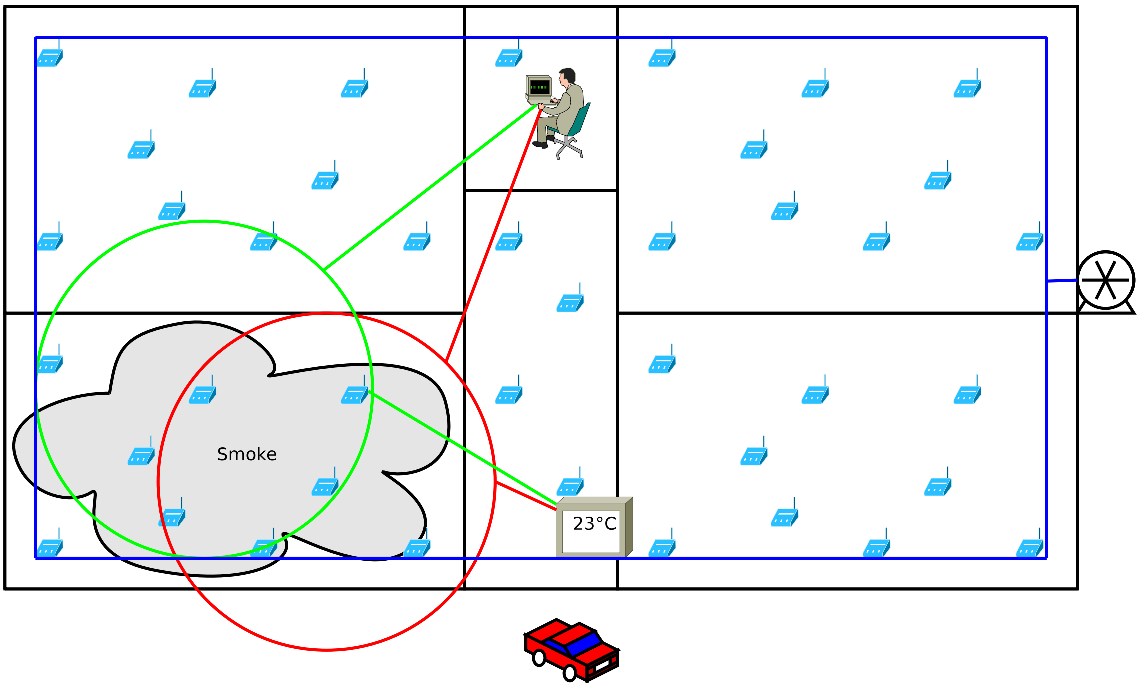

Figure 1: Temperature monitoring during a fire alarm in a modern office building.

changing link qualities. An efficient query processing system has to use protocols which are able to deal with these conditions.

Most existing solutions like for processing aggregation queries in wireless sensor networks, like [MFHH02] and [TGS06], assume a static network with a single gateway node used to disseminate all queries over over static links. This is not only unrealistic for real world applications with multiple users, but it also ignores the opportunities offered by a wireless communication medium.

1.1 Scenario

A possible application scenario might be a fire alarm in a modern building, as shown in figure 1. The building is equipped with a network of smoke detectors and temperature sensors. Under normal conditions these temperature sensors can be used for climate regulation. In case of an alarm by a smoke detector, further information about the environment can help to validate the fire alarm. Radio conditions between the nodes are irregular because of walls obstructing the line of view. Furthermore smoke and high temperatures are causing even more dynamics in the radio communication and failing nodes.

```
QUERY by Climate Control:
SELECT AVG(TEMP) FROM Nodes
  WHERE LOC IN (0,0,3000,2000)
  SAMPLE PERIOD 10min;
```

```
QUERY by Facility Manager:
ON EVENT smoke-detect(loc):
  SELECT AVG(temp), event.loc FROM sensors AS s
  WHERE dist(s.loc, even.loc) < 10m
  SAMPLE PERIOD 30s FROM 60min;

QUERY by Fire Fighters:
ON EVENT smoke-detect(loc):
  SELECT AVG(temp), event.loc FROM sensors AS s
  WHERE dist(s.loc, even.loc) < 10m
  SAMPLE PERIOD 30s FROM 60min;
```

Example queries: calculating the average temperature in different subareas of the network

The climate control needs to know the overall average to regulated the supply of cold air. In case of an alarm by a smoke sensor the temperature sensors in the vicinity of this sensor start monitoring the local temperature average. The results have to be submitted to the office of the facility manager and an information panel for the fire fighters located in the lobby. It is not uncommon that multiple smoke events are detected. MQO can exploit the spatial overlap between these queries.

The query processing system should be able to answer such queries using as little energy as possible in the sensor nodes. A single alarm should not drain the batteries in the sensors.

2 Background

TinyDB [MFHH03], [MFHH02] proposed the first comprehensive solution of a declarative query processing system running on wireless sensor nodes. TinyDB uses a fixed routing tree for query propagation and in-network aggregation. Multiple queries are processed independently which leads to duplicate sampling of sensors and transmitting of results.

Multi-query optimization has long been studied by the database community. Several heuristics have been proposed to solve this NP-hard problem efficiently [RSSB00]. Optimizing the processing of multiple queries in WSNs differs from traditional MQO. In WSNs the data and the queries are referenced with geographic information, so that the logical and the physical distribution of the data is highly correlated. Therefor optimizing the logical level is critical in this situation.

Several recent works propose solutions to address the problem of MQO in sensor networks [XBKlZ07], [TYDG05]. Some of them address especially aggregation queries [TGS06], [yYcPhL07], but all of them rely on a fixed routing tree with a single fixed gateway. Each node has a predefined parent node, which aggregates the data to an intermediate result. MQO mostly takes place in the unique gateway node. This gateway can be used as a proxy collecting the different user queries to find common expressions and to generate and disseminate artificial queries. The results of these queries can then be used locally to calculate the results of the user queries. Most of the solutions assume stable links between the nodes inside of the tree and use special maintenance messages to repair failing nodes or links.

Special routing protocols like CDF[FWMH03], IGF[HBC+07] and [HBBA04] have been designed to address the challenges of geographical routing under the dynamic conditions a real sensor network usually encounters. Instead of proactively maintaining a fixed routing structure, which requires additional control messages, they defer the binding of the routing states to the physical network topology until an individual packet forwarding operation actually happens at a sending node. They use the location, remaining energy and other factors for dynamically choosing the optimal next hop in the routing path under the current conditions.

Using a combination of dynamic routing with traditional aggregation procedures should outperform existing algorithms.

3 Objective

The objective of this thesis is to design, analyze and evaluate algorithms for processing multiple concurrent aggregation queries using a combination of robust routing protocols and techniques for multi-query optimization on a sensor network.

4 Procedure

The first task will be further research into existing ideas for MQO of aggregation queries in wireless sensor networks. That needs a survey the recent work in this topic and analysing how these ideas can be applied to a dynamic and multi-sink scenario.

A study of existing routing protocols, that have been designed for dynamic environments with irregular radio conditions, will be necessary to analyse and compare how they would perform under the given conditions. Designing a possible solution might be necessary if non of the existing protocols is appropriate for in network aggregation.

Afterwards, different approaches, including a baseline algorithm using the traditional static tree approach, will be implemented in a prototype to simulate a sensor network executing such queries. JiST/SWANS [BHvR05] is a scalable wireless network simulator built atop of a discrete event simulation engine that offers different radio models and failing nodes to simulate a dynamic environment.

A simulation of the different approaches in various scenarios will provide the data needed for an evaluation of how the new different approaches perform compared to the existing solutions under different query load and radio conditions.

Finally writing a report about the ideas, procedures and evaluation results will provide a documentation of the findings.

References

[BHvR05] Rimon Barr, Zygmunt J. Haas, and Robbert van Renesse. Scalable wireless ad hoc network simulation. In Jie Wu, editor, *Handbook on Theoretical and Algorithmic Aspects of Sensor, Ad hoc Wireless, and Peer-to-Peer Networks*, chapter 19, pages 297–311. CRC Press, 2005.

[Dou07] Conor Dougherty. The parking fix. *The Wall Street Journal*, 3, Februar 2007.

[FWMH03] Holger Füßler, Jörg Widmer, Martin Mauve, and Hannes Hartenstein. A novel forwarding paradigm for position-based routing (with implicit addressing. In *In IEEE Computer Comm. Workshop (CCW 2003*, pages 194–200, 2003.

[HBBA04] Marc Heissenbüttel, Torsten Braun, Thomas Bernoulli, and Markus Wälchli A. Blr: Beacon-less routing algorithm for mobile ad-hoc networks. *Elsevier's Computer Communications Journal (Special Issue*, 27:1076–1086, 2004.

[HBC+07] Tian He, Brian M. Blum, Qing Cao, John A. Stankovic, Sang H. Son, and Tarek F. Abdelzaher. Robust and timely communication over highly dynamic sensor networks. In *Real-Time Systems Journal, Special Issue on Real-Time Wireless Sensor Networks, 2007*, 2007.

[KPC+07] Sukun Kim, Shamim Pakzad, David Culler, James Demmel, Gregory Fenves, Steven Glaser, and Martin Turon. Health monitoring of civil infrastructures using wireless sensor networks. In *IPSN '07: Proceedings of the 6th international conference on Information processing in sensor networks*, pages 254–263, New York, NY, USA, 2007. ACM Press.

[MFHH02] Samuel Madden, Michael J. Franklin, Joseph Hellerstein, and Wei Hong. Tag: a tiny aggregation service for ad-hoc sensor networks. In *In OSDI*, 2002.

[MFHH03] Samuel Madden, Michael J. Franklin, Joseph M. Hellerstein, and Wei Hong. The design of an acquisitional query processor for sensor networks. In *SIGMOD '03: Proceedings of the 2003 ACM SIGMOD international conference on Management of data*, pages 491–502, New York, NY, USA, 2003. ACM.

[RSSB00] Prasan Roy, S. Seshadri, S. Sudarshan, and Siddhesh Bhobe. Efficient and extensible algorithms for multi query optimization. *SIGMOD Rec.*, 29(2):249–260, 2000.

[SPMC04] Robert Szewczyk, Joseph Polastre, Alan Mainwaring, and David Culler. Lessons from a sensor network expedition. In *Proceedings of the First European Workshop on Sensor Networks (EWSN)*, pages 307–322, jan 2004.

[TGS06] Niki Trigoni, Alexandre Guitton, and Antonios Skordylis. Routing and processing multiple aggregate queries in sensor networks. In *SenSys '06: Proceedings of the 4th international conference on Embedded networked sensor systems*, pages 391–392, New York, NY, USA, 2006. ACM.

[TYDG05] Niki Trigoni, Yong Yao, Alan Demers, and Johannes Gehrke. Multi-query optimization for sensor networks. In *In DCOSS*, pages 307–321, 2005.

[WAJR+05] G. Werner-Allen, J. Johnson, M. Ruiz, J. Lees, and M. Welsh. Monitoring volcanic eruptions with a wireless sensor network. In *European Workshop on Sensor Networks*, pages 108–120, 2005.

[XBKlZ07] Shili Xiang, Hock Beng, Lim Kian-lee, and Tan Yongluan Zhou. Two-tier multiple query optimization for sensor networks. In *In ICDCS*, 2007.

[yYcPhL07] Huei you Yang, Wen chih Peng, and Chia hao Lo. Optimizing multiple in-network aggregate queries in wireless sensor networks?, 2007.

[ZHKS04] Gang Zhou, Tian He, Sudha Krishnamurthy, and John A. Stankovic. Impact of radio irregularity on wireless sensor networks. In *in MobiSYS '04: Proceedings of the 2nd international conference on Mobile systems, applications, and services*, pages 125–138. ACM Press, 2004.