Available online at www.sciencedirect.com
ScienceDirect
Information Sciences 177 (2007) 2188–2205
INFORMATION SCIENCES
AN INTERNATIONAL JOURNAL
www.elsevier.com/locate/ins

Energy and quality aware query processing in wireless sensor database systems

Qingchun Ren *, Qilian Liang

Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX 76019-0016, USA

Received 7 February 2006; received in revised form 23 October 2006; accepted 20 December 2006

Abstract

Query processing has been studied extensively in traditional database systems. However, few existing methods can be directly applied to wireless sensor database systems (WSDSs) due to their characteristics, such as decentralized nature, limited computational power, imperfect information recorded, and energy scarcity in individual sensor nodes. This paper proposes a quality-guaranteed and energy-efficient (QGEE) algorithm. QGEE utilizes in-network query processing method to task WSDSs through declarative queries, and confidence interval strategy to determine the accuracy of query answers. In QGEE, the correlation between a query and a node is calculated by vector space model (VSM), and a query correlation indicator (QCI) is designed to quantify the priority of becoming active for individual nodes. Given a query, the QGEE algorithm will adaptively form an optimal query plan in terms of energy efficiency and quality awareness. This approach can reduce disturbance from measurements with extreme error and minimize energy consumption, while providing satisfying service for various applications. Simulation results demonstrate that QGEE can reduce resource usage by about 50% and frame loss rate by about 20%. Moreover, the confidence of query answers is always higher than, or equal to, the users' pre-specified precision.
© 2007 Elsevier Inc. All rights reserved.

Keywords: Wireless sensor database system; Query processing; Energy-efficiency; Space vector model; Imagine chain; *k*-Partial cover set problem; Energy reservation; Query optimization; Multipath routing

1. Introduction

Recent developments in integrated circuit technology have allowed the construction of low-cost sensor nodes that are generally equipped with sensing capability, wireless communication, as well as limited power supply and memory. Embedded those devices into environment, an emerging new type of network is created: wireless sensor network (WSN) [34]. Moreover, most of high-level tasks of WSNs, such as monitoring specific events, collecting and processing information, are accomplished by cooperation of multiple sensor nodes.

* Corresponding author. Tel.: +1 8172723488.
E-mail addresses: ren@wcn.uta.edu (Q. Ren), liang@uta.edu (Q. Liang).

0020-0255/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ins.2006.12.017

WSNs are being intended for a broad range of environmental sensing applications from weather data collection to vehicle tracking to habitat monitoring [44,8].

In WSNs, the widespread deployment of sensor nodes is transforming the physical world into a computing platform. Sensor nodes not only respond to physical signals for producing data, but also are equipped with computing and communicating capabilities. They are thus able to store, process locally, and transfer data produced. From data storage and process point of view, a WSN can be regarded as a kind of database – distributed wireless sensor database system (WSDS). Compared with traditional database systems, WSDSs store data within system and allow queries to be injected anywhere through query processing operators.

In a WSDS, the query execution usually starts from front-end nodes that issue queries into the system. Sensor nodes are viewed as data sources that provide relevant information for query processing operators. Generally, queries in a WSDS can be classified into three categories depending on the type of data (past, present, or future) requested:

- Historical queries
  This type of query is mainly used for the analysis of historical data stored at front-end nodes. For example, "what was the temperature two hours ago in the northwest quadrant?"
- One-time queries
  This type of query gives a snapshot view of a system. For example, "what is the temperature in the northwest quadrant?"
- Persistent queries
  This type of query is mainly used to monitor a system over a time interval with respect to some parameters. For example, "report the temperature in the northwest quadrant for next two hours".

Warehousing and on-demand approaches are two conventional ways of handling queries [25]. In the warehousing approach, base stations collect and store data, periodically, depending on a set of predefined parameters. There are two specific limitations for this approach: users can only query base stations, and query processing is very expensive as it utilizes valuable resources like channel bandwidth and energy. In the on-demand way, data is collected based on users' requests. However the drawbacks are that the delay for queries is unacceptable for time critical data, and the flooding of the entire system might be wasteful for one-time queries.

Lessons learned from developing network protocols for WSNs in the last couple of years show that using traditional layered networking approach has several drawbacks in system performance and efficiency. Quite often, significant improvements are possible for network protocols, but they require a great amount of information to be passed along the layers of system. Although this approach, in principle, allows independence among various protocols, it incurs significant overhead in parameter transfer. Moreover, improvements performed in a specific layer can cause impairments, or even be counterproductive for other layers. Therefore, optimization can be more effective when taking into account overall system and using all available knowledge. In other words, cross-layer design approach is a viable approach for WSNs' energy and quality problems.

The goal of monitoring through sensor nodes is to infer information about objects from measurements made from remote locations. Moreover, inference processes are always less than perfect. Consequently, the problem of uncertainty, which stands for the quality of query answer, is central to monitoring applications. Hence, to build useful information systems, it is necessary to learn how to represent and reason with imperfect information [43]. As a result, Motro [28] is interested in how imperfect information may be represented in a database system. In [42] and [31], they proposed methods of dealing with the uncertainty of moving objective databases. In [41], they presented a unified fuzzy-probabilistic framework for modelling processes of medical diagnosis. In their work, the belief computation is related to diagnostic inference. The final conclusion of inference is the diagnosis with the greatest belief value. It is also shown how their membership functions and basic probability assignments are estimated on the basis of experimental data.

In general, imperfect information is typically handled by attaching a number to it, which represents a subjective measure of the certainty according to the observer. The way, in which the number is manipulated, depends upon the theory that underlies the number. There are possibilistic databases [35] and probabilistic databases [7,32,33]. Probabilistic approach has begun to be used by WSDSs to process query with limited information [2]. In [38], they discussed how to handle aggregate operations in probabilistic databases. They

Download English Version:

https://daneshyari.com/en/article/396166

Download Persian Version:

https://daneshyari.com/article/396166

Daneshyari.com