International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 03 Issue: 06 | June-2016 www.irjet.net p-ISSN: 2395-0072

A Distributed framework for mobile data gathering with concurrent data uploading & Data Retrieval Scheduling in wireless sensor

J.Sadhana
Research Scholar, Department of Computer Science
Shrimati Indira Gandhi College, Trichy,
R.Indra,M.Sc.,M.phil.,M.C.A.,Assistant Professor,
Department of M.C.A Shrimati Indira Gandhi college,
Trichirappalli - 620002

ABSTRACT:- *Data uploading time constitutes a large portion of mobile data gathering time in wireless sensor networks. a new data gathering cost minimization framework for mobile data gathering in wireless sensor networks. we first propose a data gathering cost minimization (DaGCM) framework with concurrent data uploading, which is constrained by flow conservation, energy consumption, link capacity, compatibility .One of the main features of this framework is that it allows concurrent data uploading from sensors to the mobile collector to sharply shorten data gathering latency and reduce energy consumption.. To maximize the number of downloads given a deadline, we define a problem called largest number data retrieval .We prove the decision problem of LNDR is NP-hard, which aims at downloading a set of requested data items with the least response time and energy consumption.Furthermore, we present a distributed algorithm composed of cross-layer data control, routing, power control and compatibility determination subalgorithms with explicit message passing.. Finally, we provide numerical results to show the convergence of the proposed DaGCM algorithm and its advantages over the algorithm without concurrent data uploading and power control in terms of energy consumption.*

Keywords: ***Multi-channel, wireless data broadcast, multi-item request, data retrieval scheduling***

1. INTRODUCTION

MOBILE ad hoc networks (MANETs) are originally designed for military tactic environments. Communication anonymity is a critical issue in MANETs, which generally consists of the following aspects: 1) Source/ destination anonymity—it is difficult to identify the sources or the destinations of the network flows. 2) End-to-end relationship anonymity—it is difficult to identify the end to-end communication relations. To achieve anonymous MANET communications, many anonymous routing protocols ,traffic analysis models have been widely investigated for static wired networks. Recently, statistical traffic analysis attacks have attracted broad interests due to their passive nature, i.e., attackers only need to collect information and perform analysis quietly without changing the network behavior. 1) The broadcasting nature: In wired networks, a point-to-point message transmission usually has only one possible receiver. While in wireless networks, a message is broadcasted, which can have multiple possible receivers and so incurs additional uncertainty? 2) The ad hoc nature: MANETs lack network infrastructure, and each mobile node can serve as both a host and a router. Thus, it is difficult to determine the role of a mobile node to be a source, a destination, or just a relay. 3) The mobile nature: Most of existing traffic analysis models do not take into consideration the mobility of communication

peers, which make the communication relations among mobile nodes more complex. we investigate the data retrieval problem in both push-based and pull-based broadcasts. When users only retrieve one data item per request, the retrieving process is straightforward. In the last decade, how to allocate data items onto multiple channels to minimize the expected response time has become a hot research topic which captured a great deal of attentions. In a multi-channel system, retrieving multiple data items probably needs switching among the channels, which not only consumes additional energy, but also causes possible conflicts Therefore, different retrieving Order may result in different response time and different energy consumption. Assuming the arriving time and channel locations of requested data items are already known from the index information, our work will mainly focus on the data retrieval scheduling from the client's point of view, in which multiple data items have to be downloaded from different channels.

2. LITERATURE SURVEY

1) Algorithms for Mining the Evolution of Conserved Relational States in Dynamic Networks,

AUTHORS: R. Ahmed and G. Karypis

Dynamic networks have recently being recognized as a powerful abstraction to model and represent the temporal changes and dynamic aspects of the data underlying many complex systems. Significant insights regarding the stable relational patterns among the entities can be gained by analyzing temporal evolution of the complex entity relations. This can help identify the transitions from one conserved state to the next and may provide evidence to the existence of external factors that are responsible for changing the stable relational patterns in these networks. This paper presents a new data mining method that analyzes the time-persistent relations or states between the entities of the dynamic networks and captures all maximal non-redundant evolution paths of the stable relational states. Experimental results based on multiple datasets from real-world applications show that the method is efficient and scalable.

2) Novel Approaches to Crawling Important Pages Early

AUTHORS: M.H. Alam, J.W. Ha, and S.K. Lee

Web crawlers are essential to many Web applications, such as Web search engines, Web archives, and Web directories, which maintain Web pages in their local repositories. In this paper, we study the problem of crawl scheduling that biases crawl ordering toward important pages. We propose a set of crawling algorithms for effective and efficient crawl ordering by prioritizing important pages with the well-known PageRank as the importance metric. In order to score URLs, the proposed algorithms utilize various features, including partial link structure, inter-host links, page titles, and topic relevance. We conduct a large-scale experiment using publicly available data sets to examine the effect of each feature on crawl ordering and evaluate the performance of many algorithms. The experimental results verify the efficacy of our schemes. In particular, compared with the representative RankMass crawler, the FPR-title-host algorithm reduces computational overhead by a factor as great as three in running time while improving effectiveness by 5 % in cumulative PageRank

3) Identifying Influential and Susceptible Members of Social Networks

AUTHORS: S. Aral and D. Walker

Identifying social influence in networks is critical to understanding how behaviors spread. We present a method that uses in vivo randomized experimentation to identify influence and susceptibility in networks while avoiding the biases inherent in traditional estimates of social contagion. Estimation in a representative sample of 1.3 million Facebook users showed that younger users are more susceptible to

influence than older users, men are more influential than women, women influence men more than they influence other women, and married individuals are the least susceptible to influence in the decision to adopt the product offered. Analysis of influence and susceptibility together with network structure revealed that influential individuals are less susceptible to influence than noninfluential individuals and that they cluster in the network while susceptible individuals do not, which suggests that influential people with influential friends may be instrumental in the spread of this product in the network.

4) Big Privacy: Protecting Confidentiality in Big Data

AUTHORS: A. Machanavajjhala and J.P. Reiter

A tremendous amount of data about individuals – e.g., demographic information, internet activity, energy usage, communication patterns and social interactions – are being collected and analyzed by many national statistical agencies, survey organizations, medical centers, and Web and social networking companies. Wide dissemination of microdata (data at the granularity of individuals) facilitates advances in science and public policy, helps citizens to learn about their societies, and enables students to develop skills at data analysis. Often, however, data producers cannot release microdata as collected, because doing so could reveal data subjects' identities or values of sensitive attributes. Failing to protect confidentiality (when promised) is unethical and can cause harm to data subjects and the data provider. It even may be illegal, especially in government and research settings. For example, if one reveals confidential data covered by the U. S. Confidential Information Protection and Statistical Efficiency Act, one is subject to a maximum of $250,000 in fines and a five year prison term.

3. Existing system:

When users only retrieve one data item per request, the retrieving process is straightforward. a user requests multiple data items at a time (e.g., a user may submit a query for 10 stocks at a time). In wireless communication technologies such as OFDM (Orthogonal Frequency Division Multiplexing) makes efficiently broadcasting through multiple channels possible. To allocate data items onto multiple channels to minimize the expected response time has become a hot research topic which captured a great deal of attentions. In a multi-channel system, retrieving multiple data items probably needs switching among the channels, which not only consumes additional energy, but also causes possible conflicts. Therefore, different retrieving order may result in different response time and different energy consumption.

Drawbacks:

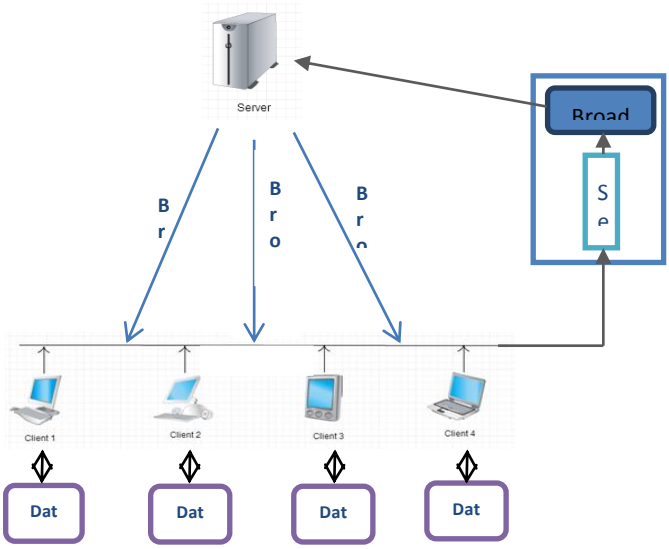
- Clients will retrieve all the necessary indices before downloading data, which is not related to the data retrieval scheduling
- The mobile clients can only access a single channel at any particular time. So a client cannot download data from two or more channels.
- The retrieval does not require downloading the data items in a specific order.

Proposed System:

The data retrieval scheduling problem for multi-item requests in multi-channel broadcast environments. To maximize the number of downloads given a deadline; we define a problem called largest number data retrieval (LNDR) and another problem, namely minimum cost data retrieval (MCDR), with the objective of minimizing the response time and energy consumption. Responsetime is defined as the time interval between the momentsthat a request is

submitted to the moment that all therequested data items are downloaded.

4. SYSTEM MODEL



5.IMPLEMENTATION:

Service Broadcasting:

In this module, to transmit the data we use number of channel for multiple broadcasting. The server maintained the collection of request and services. The down-link, clients will send requests to the server through an up-link, and the server will decide the broadcast schedule based on the requests received.

MCDR:

The MCDR is to reduce the response time and energy consumption. The index information is assumed to be obtained before data retrieving; hence the tuning time for index retrieval is not considered when calculating the energy consumption for data retrieving. we develop a greedy heuristic for MCDR. It combines the benefits of both channel scheduling to reduce the energy consumption in channel switching and data item scheduling to reduce the response time and the energy consumption.

LNDR:

To maximize the number of downloads before a deadline; we define a problem called largest number data retrieval (LNDR). LNDR problem takes the "deadline" into consideration and therefore also describes the time-critical scenario. We propose a polynomial time optimal algorithm for LNDR. When channels are unsynchronized, we prove LNDR is NP-hard. For pull-based broadcast, the data items are broadcasted timely.

Pull Based System:

Pull-based broadcast is different from push-based broadcast in that data items are scheduled at the server side in real time. After submitting a request, a client at any particular time t may only get the broadcast schedule for the next one or several time slots. Assuming a client can get the broadcast schedule for the next r time slots, we show that a greedy data receival strategy can provide a (1/2+2/r) factor approximation solution.

Data Scheduling:

For data scheduling at the server side. We implemented a dynamic programming approach. The data items are placed on channels timely according to the R _W value, where R denotes the number of requests for a data item and W is the longest waiting time for a data item. It is worthy to mention that the algorithms are designed for data scheduling for multi-channel environments.

6. CONCLUSION AND FUTURE WORK

In this paper, we have designed a cross-layer optimization framework for mobile data gathering in WSNs considering elastic link capacity and power control on sensors. By enabling concurrent data uploading, we first formulate the problem to minimize total data collection cost under the constraints of flow

conservation, energy budget (power control), elastic link capacity and compatibility among sensors. Then, by introducing auxiliary variables, we transform the non-convex problem into a convex one and further decompose it into several sub problems of data control and data split at the transport layer, routing at the network layer, and power control and compatibility decisions at the physical layer. We employ sub gradient iterative approach to solve the problem and present several distributed sub algorithms with explicit message passing. there are some interesting issues to be further explored in future. First, the performance gains compared to system complexity using MMSE receivers should be further studied. Second, the cost function used in this paper may not completely reflect the overall pricing structure in the network. Therefore, a more comprehensive model that accounts for aspects from sensor's transmission/reception energy, buffer, encode/decode, moving energy and human administration cost may be considered in future

8. REFERENCES:

[1] C. Hua and T.-S. P. Yum, "Optimal routing and data aggregation for maximizing lifetime of wireless sensor networks," IEEE/ACM Trans. Netw., vol. 16, no. 4, pp. 892–903, Aug. 2008.

[2] Y. Wu, Z. Mao, S. Fahmy, and N. Shroff, "Constructing maximum-lifetime data-gathering forests in sensor networks," IEEE/ ACM Trans. Netw., vol. 18, no. 5, pp. 1571 1584, Oct. 2010.

[3] Z. Zhang, M. Ma, and Y. Yang, "Energy-efficient multihop polling in clusters of two-layered heterogeneous sensor networks," IEEE Trans. Comput., vol. 57, no. 2, pp. 231–245, Feb. 2008.

[4] N. Ahmed, S. Kanhere, and S. Jha, "The holes problem in wireless sensor networks: A survey," ACM SIGMOBILE, vol. 9, no. 2, pp. 4–18, Apr. 2005.

[5] R. W. N. Pazzi and A. Boukerche, "Mobile data collector strategy for delay-sensitive applications over wireless sensor networks," Comput. Commun., vol. 31, no. 5, pp. 1028–1039, Mar. 2008.

[6] M. Gatzianas and L. Georgiadis, "A distributed algorithm for maximum lifetime routing in sensor networks with mobile sink," IEEE Trans. Wireless Commun., vol. 7, no. 3, pp. 984–994, Mar. 2008.

[7] M. Ma and Y. Yang, "SenCar: An energy-efficient data gathering mechanism for large-scale multihop sensor networks," IEEE Trans. Parallel Distrib. Syst., vol. 18, no. 10, pp. 1476–1488, Oct. 2007.

[8] M. Ma, Y. Yang, and M. Zhao, "Tour planning for mobile data gathering mechanisms in wireless sensor networks," IEEE Trans. Vehicular Technol., vol. 62, no. 4, pp. 1472–1483, May 2013.

[9] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," IEEE J.Sel. Areas Commun., vol. 22, no. 6, pp. 1089–1098, Aug. 2004