

Collaborative Query Processing among Heterogeneous Sensor Networks

Yuan He, Mo Li, Yunhao Liu
Hong Kong Uni. of Sci. and Tech.
{heyuan, limo, liu}
@cse.ust.hk

Jizhong Zhao
Xi'an Jiaotong University
zjz@mail.xjtu.edu.cn

Wei Lan Huang, Jian Ma
Nokia Research Center, China
{Jennifer.Huang, Jian.J.Ma}
@nokia.com

ABSTRACT

Demands on better interacting with physical world require an effective and comprehensive collaboration mechanism among multiple heterogeneous sensor networks. Previous works mainly focus on improving each single and specific sensor network, thus fail to address this newly emerged issue. In this paper, we study the issue of collaborative query processing among multiple heterogeneous sensor networks and formulate it into an optimization problem with respect to energy efficiency, called EE-QPS. To the best of our knowledge, we are the first one considering the collaborative query processing among heterogeneous sensor networks. By utilizing the implications among sensor networks, we design a heuristic approach named IAP to resolve EE-QPS. The experimental results validate our scheme and show that IAP achieves optimized energy efficiency under various environments.

Categories and Subject Descriptors

C.2.1 [Computer Communication Networks]: Network Architecture and Design – *Distributed networks*; C.2.4 [Computer Communication Networks]: Distributed Systems – *Distributed applications*;

General Terms

Algorithms, Performance, Design.

Keywords

Sensor Network, Collaboration, Query Processing.

1. INTRODUCTION

The recent advances in wireless communication and micro-electronic technologies have boosted the popularization of sensor networks. Sensor networks nowadays range from personal to mission critical systems including scientific observation, digital life, home automation, environment surveillance, traffic monitoring, and so on [1, 2, 3, 4].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

HeterSANET'08, May 30, 2008, Hong Kong SAR, China.

Copyright 2008 ACM 978-1-60558-113-2/08/05...\$5.00.



Figure 1. Oil production map in Northeast China

Meanwhile, we have seen an expansion in the use of various sensing devices such as widely equipped cameras for traffic monitoring, gas and temperature sensors deployed to ensure safe coal mining production, tens of thousands of PDAs and smartphones capable of capturing the acoustics or images disseminated in publics. They are developed and promoted by governments, enterprises, and public organizations, offering continuous collection of real-time information, satisfying the requirement of people's daily lives.

In the foreseeable future, we expect to witness the proliferation of sensor networks with a variety of functions that require a comprehensive collaboration mechanism among them. Previous studies in sensor networks, however, mainly focus on the performance and efficiency inside a single sensor network [1, 2, 3, 4, 5]. In this work, we broaden the research into the scope of multiple heterogeneous sensor networks.

This study is indeed motivated by a practical application of Qinhuangdao Oilport [7], which is a hinge city in the oil production and transportation of Northeast China. In order to keep the whole production and transportation flow safe and efficient, timely planning is necessary. And it is related to many factors such as oil supplies in oil fields, the flux capacity of oil pipeline

transmission, landway traffic, and environments of harbors. Formerly, we can only make relatively static decisions based on coarse estimations on these factors. The output of decisions often suffers from the dynamics of these factors, causing un-redemptive loss of profit and even serious accidents. Therefore a number of wireless sensor networks (WSNs) are deployed for the application to obtain live environmental data, as shown in Figure 1.

To truly utilize WSNs in the above application, however, many challenges need to be addressed. Previous studies mainly concentrate on data collection and query processing in a single sensor network [6]. Using these approaches, we can only obtain isolated and incomplete results, inevitably leading to unilateral and even incorrect decisions. The WSNs in the above application have heterogeneous sensor nodes, functionalities, and are dispersedly deployed. It is necessary but challenging to enable interoperations and information integration among them. Also, the sensor networks continuously generate huge volumes of data with various attributes, simply gathering all the data and processing them in a centralized manner is obviously infeasible. Distributed sensing and collaborative query processing among them are indispensable. Moreover, the sensor networks are likely to receive substantive complex queries, while the sensors are usually energy-constrained and not easily rechargeable. Therefore energy-efficient query processing among multiple sensor networks is a crucial issue but has never been studied before.

To address the above challenges, we propose the scheme of collaborative query processing among heterogeneous sensor networks. The major contributions of this work are as follows.

- First, we introduce the system model of collaborative sensor networks and then formalize the concept of collaborative query processing among them.
- Second, implications among different sensor networks are utilized to minimize the total energy cost of query processing. We formulate the optimization problem of collaborative query processing with respect energy efficiency, called EE-QPS, which is proved to be NP-hard.
- Finally, we propose an efficient approach to schedule the pipeline of query processing to achieve optimized energy efficiency for the involved sensor networks.

The remainder of this paper is organized as follows. Section 2 introduces the background of our study. We elaborate the detailed design in Section 3, including the formalization of collaborative query processing, the formulation of the EE-QPS problem, and the implication-aware heuristic approach to resolve it. Section 4 presents the methodology and results of experiments. In Section 5, we conclude the paper and point out future research directions.

2. BACKGROUND

We have seen remarkable success in the research field of sensor networks. However, the state of arts mainly concentrates on the in-network issues, such as in-network sensing control, data processing, and protocol design [9, 10, 11, 12, 13, 14]. There is not many existing works concerning the intra-sensor-network issues such as exploring mechanisms to manage, share, analyze, and understand the data among different sensor networks. As the background of our study in this paper, we discuss some related models and systems in this section. We also compare some mature research fields with collaborative sensor networks,

showing the special characteristics and challenges we should emphasize and address in this emerging area.

A preliminary model of networking heterogeneous sensors is initially proposed by Kevin A. Delin et al [15]. In their proposal named Sensor Web, different types of sensing devices owned by a common authority, are geographically distributed and connected as a unified networking system. The feature of this unified model lies in the ability of the individual pieces to act and coordinate as a whole. Heterogeneous sensors are contained in the system and interact with each other through wireless communications. This immediately allows the system to be synchronous throughout, unlike many other networks. Data sensed by a sensor are delivered to and utilized by some other sensors such that all the sensors can act a whole for their common purposes. By definition, a Sensor Web is an autonomous, stand-alone, sensing entity, which does not address the issue of query processing, nor manipulate the collaboration mechanism of multiple sensor networks. It differs greatly from collaborative heterogeneous sensor networks that we study in this paper.

The term “SenseWeb” is sometimes used to refer to sensors connected to the Internet. S. Nath et al propose SensorMap [16], which represents a new class of applications that relies on real-time sensor data and its mash-up with the geocentric web to provide instantaneous environmental visibility and timely decision support. The platform also transparently provides mechanisms to archive and index data, to process queries, to aggregate and present results on the web interface based on Windows Live Local. SensorMap adopts a fully centralized architecture. Data collection, aggregation, visualization, sensor indexing, and query processing are all executed by a central server locally. SensorMap works reasonably well currently when only a few sensor networks are included and user queries are not too frequent. But it will probably be argued that such architecture lacks scalability.

IrisNet proposed in [17] is another software infrastructure for worldwide web of sensors that lets users query globally distributed collections of high-bit-rate sensors. IrisNet appears similarly with SensorMap in the sense that they both adopt tools like XML to describe sensor data and provide visualized results for queries on a Web portal. But IrisNet differs much from SensorMap because it adopts a decentralized architecture. Heterogeneous sensors are integrated in IrisNet. Despite differences between sensor types, developers need a generic data acquisition interface to access sensors. In IrisNet, the nodes that implement this interface are called sensing agents (SAs). On the other hand, services must store the service-specific data the SAs produce in a distributed database. In IrisNet, the nodes that implement this distributed database are called organizing agents (OAs). In general, SAs take charge of sensing controls and data aggregation from sensor networks, while OAs are in charge of storing data and resolving queries. They two, together with the underlying sensor networks, and a web portal construct the main body of IrisNet.

Another prototype system is SensorNet [18]. SensorNet proposes a data architecture and infrastructure that supports plug-and-play sensors of various types, archival storage of sensor data, standards-based publication of sensor data, and sensor control services. It allows for the integration of dissimilar sensor systems

into one system. It focuses on high speed, reliable access and delivery of sensor data inside the infrastructure.

The above models and systems (SensorMap, IrisNet, and SensorNet) mainly concentrate on the data collection, aggregation, and direct exhibition of sensor data from heterogeneous sensor networks as well as the query processing based on the collected data afterwards. Although different types of sensor networks are contained, they are still independent during the process to resolve queries, not supporting any intra-network collaborations among each other.

We may also find some similarities between distributed databases [19] and collaborative sensor networks. Both these two types of systems serve users with distributed data sources. The data in distributed databases are less dynamic and often replicated among the hosts, while the sensor data are frequently updated and reside in different sensor networks. More importantly, the intra-network bandwidth consumption is the first concern when we design the query processing in a distributed database. On the contrary, we must regard the energy costs of sensing and communication inside each sensor network as the first concern when we design the collaborative query processing among heterogeneous sensor networks. Hence it is inapplicable to migrate the existing approaches in the field of distributed databases into the context of collaborative sensor networks.

3. DESIGN

In this section, we first formalize the concept of collaborative query processing among heterogeneous sensor networks. We elaborate how to utilize the implications among sensor networks to save the total energy cost of query processing. The so-called EE-QPS problem is then formulated. Due to the NP-hardness of EE-QPS, we design a heuristic approach to resolve it.

3.1 Collaborative Query Processing

We describe a system supporting collaborative query processing among heterogeneous sensor networks as follows: first, it is an information system providing live sensing and query processing services; second, multiple heterogeneous sensor networks are integrated in; third, instead of directly uploading their data for aggregation and exhibition at any dedicated server, sensor networks in this system collaborate with each other in accomplishing the querying tasks.

A typical system diagram of collaborative sensor networks is shown in Figure 2. Generally it consists of the following components: sensor networks, data and query processor, and the portal for query input and data output. Sensor networks, usually via their sink nodes, connect with each other over the Internet. A central managing server (CM), which can be web-based or web-free, is in charge of accepting external queries and outputting the queried results. The core component in the middle manages sensing control, query scheduling, and data aggregation, which is named as data and query processor.

Our design aims at a mechanism for CM to schedule a complex query that involves multiple sensor networks. We look back to the motivated application. To retrieve information from those WSNs, a straightforward but inefficient method is that we independently query every sensor network and gather all corresponding information for local analysis. Then each query will potentially

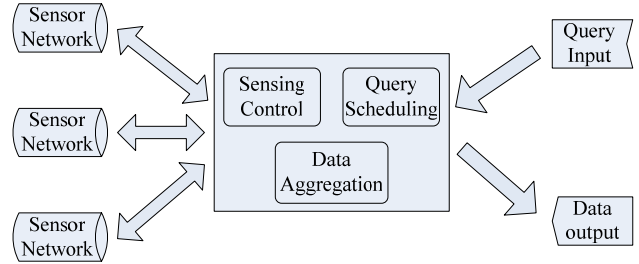


Figure 2. The diagram of a system that integrates collaborative sensor networks

involve all the sensor nodes in each sensor network. Such a blind querying scheme incurs excessive energy consumption, which obviously ruins the sensor networks for long-term uses.

Consequently, a crucial yet challenging issue is how to efficiently query the sensor networks and obtain the desired information with the minimum total energy cost in all the sensor networks.

Due to the natural interdependence in the physical world, data of different sensor networks are usually correlated with each other.

For example, sensor data of temperature and humidity, Ultra-violet Index and illumination, the road traffic and the busyness of parking lots, the flux of oil pipelines and the available capacity of oil tanks. We can partially infer the data of a sensor network based on the data of another one, as long as they are correlated. Existing data correlations have been validated by both practical deployments and theoretical models of sensor networks [2, 3, 8]. We call such correlations among sensor networks *implication*.

Implications can be utilized to save the total energy cost of query processing. Specifically, when we process a query involving multiple sensor networks, the data from previously queried sensor networks can be used to partially infer the data of the subsequently queried sensor networks. Therefore it costs the subsequent ones fewer operations (including sensing and communication) to obtain the necessary data. The total energy cost to process this query is thus reduced. For example in the motivating application, when setting a schedule for the oil production flow, after we obtain the information from the harbor surveillance sensor networks, we may sweep off those harbors under infeasible condition. In the subsequent stages, we only query the status of traffic and oil pipelines from only a portion of sensor networks related to the feasible harbors, saving unnecessary operations. Therefore, in the context of collaborative sensor networks, it is of great significance to schedule the sequence of query processing to achieve the optimized overall energy efficiency by fully utilizing the implications among sensor networks.

For a complex query, the involved sensor networks are correlated with each other. The cost reduction in a subsequently queried sensor network is an accumulative effect caused by all the previously queried ones. Thus it is a natural choice to process a complex with a pipeline of all the involved sensor networks. Suppose a query Q concerns a subset of sensor networks, say $\{W_1, W_2 \dots W_N\}$. As shown in Figure 3, we denote a sensor network as a node, ignoring the concrete structures. The pipeline to resolve a complex query can be depicted as a directed cycle in the graph, starting and ending at CM. At the beginning, query Q is received from the web portal and interpreted. Then CM selects

sensor network W_1 to forward query Q . After corresponding operations (data sensing and transmission etc.), W_1 passes query Q as well as the filtered sensor data to W_2 . Similarly, after W_2 finishes its work, it passes query Q and the accumulative filtered sensor data to W_3 . The process continues until all the N sensor networks have been accessed by query Q one by one. In the end, the complete result of query Q is returned from the last sensor network W_N to CM, and then output to the user.

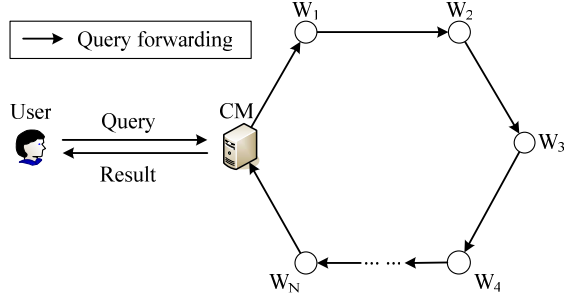


Figure 3. The pipeline of collaborative query processing

3.2 The Query Optimization Problem

We address the issue of query optimization by emphasizing the implication-aware collaboration among the sensor networks. Instead of digging into the concrete operations (such as data sensing, filtering, aggregation, and caching) and correlation patterns of sensor data, we focus on the impact of implication and the methodology to utilize it, so that we can minimize the total energy cost of query processing. We name the query optimization problem as EE-QPS (Energy-Efficient Query Processing among heterogeneous sensor networks).

The EE-QPS problem can be modeled with a directed weighted graph $G=(V, E)$. Suppose we have N sensor networks involved. V is the set of nodes. Let nodes $v_1, v_2 \dots v_N$ represent the sensor networks $W_1, W_2 \dots W_N$. E is the set of edges representing implications. Edge e_{ij} is the edge from v_i to v_j , and s_{ij} is the weight of e_{ij} ($i \neq j$). We define s_{ij} as the index of implication from W_i to W_j . It denotes the proportion of information in W_j that remains uncertain when the data of W_i are known. Note that, s_{ij} is not necessarily equal to s_{ji} . When W_j is completely independent from W_i , $s_{ij}=1$; when W_j can be completely inferred from W_i , $s_{ij}=0$. We use information entropies to quantify the indices of implications from sensor network X to sensor network Y .

$$s_{XY} = \frac{H(Y|X)}{H(Y)} = \frac{-\sum_i \sum_j P(x_i, y_j) \log P(y_j | x_i)}{-\sum_j P(y_j) \log P(y_j)}$$

We directly use X and Y to represent the data sets of X and Y while x_i, y_j are the corresponding sensor data respectively. $H(Y)$ is the original entropy, denoting the original uncertainty of data set Y . $H(Y|X)$ is the conditional entropy, denoting the uncertainty of data set Y when data set X is already known. The conditional probability $P(y_j|x_i)$ is calculated with those x_i and y_j falling into the same *Period of validity*. For instance, $P(y_j=m|x_i=n)$ is the probability of y_j to be m , given the current value of x_i to be n .

We define C_i as the original energy cost in sensor network W_i (measured by nJ) incurred by a query, when no information of W_i is inferred from other sensor networks. It can be quantified as

follows: $C_i = \text{Unit cost} \times \text{Scale}$ of W_i , where *Unit cost* is the cost for a single sensor to respond to a query and *Scale* is the number of sensors in W_i .

It is difficult to accurately estimate the cumulative effect of implications among sensor networks, especially in dynamic and unpredictable environments. As a simplified case, we assume all s_{ij} are independent from each other. Thus the aforementioned accumulative effect can be quantified by multiplying the indices of implications from all the upstream sensor networks along the pipeline to the current one.

Taking Figure 3 as an example, we process a query through the pipeline ($W_1 \rightarrow W_2 \rightarrow \dots \rightarrow W_N$). The total cost P incurred in all the N sensor networks is calculated by:

$$P = \sum_{i=1}^N P_i = \sum_{i=1}^N (C_i \times \prod_{1 \leq j \leq i} s_{ji})$$

For convenience, we set $s_{ii}=1$. Clearly, we have $N!$ options to schedule the pipeline of query processing. Meanwhile, due to the natural heterogeneity, the indices of implications probably vary a lot with different pairs of sensor networks. Therefore different pipelines present great difference in the total energy costs. Towards the same queried result, there exists an optimal pipeline scheduling, which incurs the minimum total energy cost. Formally, the EE-QPS problem is formulated as follows:

INSTANCE: A sequence of positive constants ($C_1, C_2 \dots C_n$), where C_i denotes the normalized original cost in sensor network W_i incurred by a query. Correspondingly, there is an implication matrix

$$S_{N,N} = \begin{bmatrix} s_{11} & s_{12} & \dots & s_{1N} \\ s_{21} & s_{22} & \dots & s_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ s_{N1} & s_{N2} & \dots & s_{NN} \end{bmatrix}$$

where $0 \leq s_{ij} \leq 1$, $s_{ii}=1$ for all integers i and j in $[1, N]$.

SOLUTION:

($a_1, a_2 \dots a_N$), which is a permutation of $(1, 2 \dots N)$.

MEASURE:

$$P = \sum_{i=1}^N (C_{a_i} \times \prod_{1 \leq j \leq i} s_{a_j a_i})$$

which is the total cost of all the involved sensor networks. The optimal solution of the problem minimizes the value of P , i.e., achieves the best energy efficiency. The optimization problem of EE-QPS is NP-hard, so we need to design a heuristic approach to resolve it. Due to the limit of pages, we do not present the detailed proof here.

3.3 The Heuristic Approach

First of all, sensor networks periodically exchange their latest sample data with each other so that the implications among them can be quantified. The updated indices of implications are reported to CM. Thus, CM obtains a global view of implications while needs not collect the detailed sensor data.

We design a greedy approach called IAP (Implication-Aware Processing) to find a close-to-optimal scheduling, the temporal complexity of which is only $O(N^2)$. Given an instance of EE-QPS

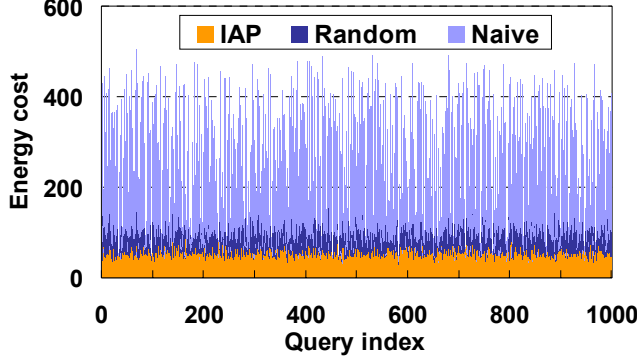


Figure 4. Comparison of approaches (1)

with N sensor networks ($W_1, W_2 \dots W_N$), we suppose ($V_1, V_2 \dots V_N$), a permutation of ($W_1, W_2 \dots W_N$), is the final scheduled pipeline. Then the process of scheduling can be divided into $N+1$ states $T_0, T_1 \dots T_N$, where T_i refers to the state when the first i sensor networks of the pipeline have been selected. Correspondingly, we define two sets R and R' . Given state T_i , R contains the first i sensor networks that have been determined in the pipeline, while R' contains the $(N-i)$ unselected ones. We define a heuristic function as follows to select the $(i+1)$ th sensor network of the pipeline.

$$U(v) = C_v \prod_{x \in R} s_{xv} + \sum_{y \in R' - \{v\}} (C_y \times s_{vy} \times \prod_{x \in R} s_{xy})$$

The parameters C_v , C_x and C_y are the original costs, which can be known from the basic information of the sensor networks. s_{xv} , s_{xy} and s_{vy} are the indices of implications quantified as above. $U(v)$ is the sum of two parts. The first part is the energy cost in v if it is selected as the $(i+1)$ th sensor network. The second part is the upper bound of total energy cost in the remaining $(N-i-1)$ unselected sensor networks, if v is selected as the $(i+1)$ th sensor network. $U(v)$ denotes the maximal energy cost incurred in the remaining $(N-i)$ sensor networks if v is selected as the $(i+1)$ th sensor network of the pipeline.

Therefore the $(i+1)$ th sensor network of the pipeline should be sensor network v which minimize $U(v)$, expressed as follows:

$$V_{i+1} = \arg \min_{v \in R'} U(v)$$

Subsequently, V_{i+1} is removed from R' and added into R . After N rounds of selection, the pipeline is finally decided. As soon as the pipeline is scheduled, the query is passed from CM to sensor network V_1 , then V_1 to V_2 , then V_2 to V_3 and so on. In the end, the query is finished on V_N . The final result is then returned to CM and output to the user.

4. PERFORMANCE EVALUATION

We conduct several groups of simulations to evaluate the performance of the proposed IAP approach based on the data we collected from the prototype system. The parameters used in the simulations are listed below:

N_W : Total number of sensor networks;

N_Q : Total number of queries injected;

$c[1..N_W]$: Original costs in each sensor network, measured by nJ ;

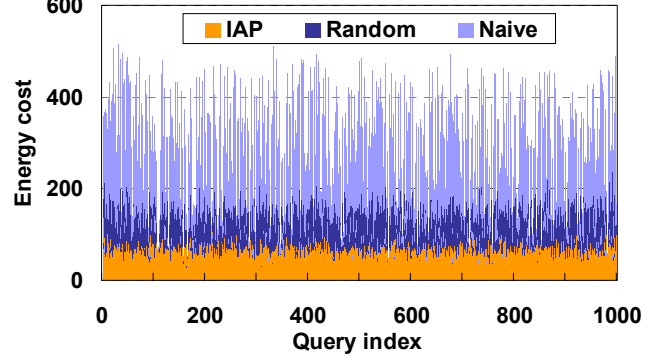


Figure 5. Comparison of approaches (2)

$s[1..N_W][1..N_W]$: Indices of implications;

$n[1..N_W]$: Number of sensor networks involved in each query.

The following basic settings apply to all the simulations: $N_W=30$. $N_Q=1000$. The sensor networks involved in each query are randomly chosen from the 30 simulant sensor networks. Other relevant parameters are varied in the simulations for comparisons.

4.1 Comparison among Approaches

Let $n[1..N_W]$ conform to a uniform distribution on $[3, 15]$. $c[1..N_W]$ conform to a uniform distribution on $[10, 50]$. The first group of simulations is divided into two rounds. In round 1, $s[1..N_W][1..N_W]$ conform to a uniform distribution on $[0, 1]$. In round 2, $s[1..N_W][1..N_W]$ conform to a uniform distribution on $[0.7, 1]$.

We compare IAP with two other approaches for query processing. One of them is the naive approach, which broadcasts a query to all the involved sensor networks simultaneously and all the sensor networks process the query independently. Obviously the total energy cost of that approach should be the sum of original costs in all the involved sensor networks. The other is the random approach, which processes a query in a pipeline and the pipeline is randomly scheduled. The random approach partially utilizes the implications among sensor networks but does not optimize the pipeline scheduling.

Figures 4 and 5 plot the energy costs of all 1000 queries using three different approaches. The statistical result says:

1) Compared with the naive approach, the percentage of cost saved by the random approach has mean 51.3% and standard deviation 19.6% in round 1, and has mean 38.9% and standard deviation 17.1% in round 2. This shows the benefit of transformation among sensor networks on improving the energy efficiency of query processing. It also supports the necessity of the implication-based sink-overlay construction, as introduced in Section 3.3.

2) Compared with the random approach, there is a further save of cost in IAP. The percentage of cost saved by IAP has mean 37.6% and standard deviation 17.2% in round 1, and has mean 41.6% and standard deviation 19.1% in round 2. IAP always outperforms the random approach in all instances. This validates the rewarding effect of IAP, which optimizes pipeline scheduling by fully utilizing the implications among sensor networks.

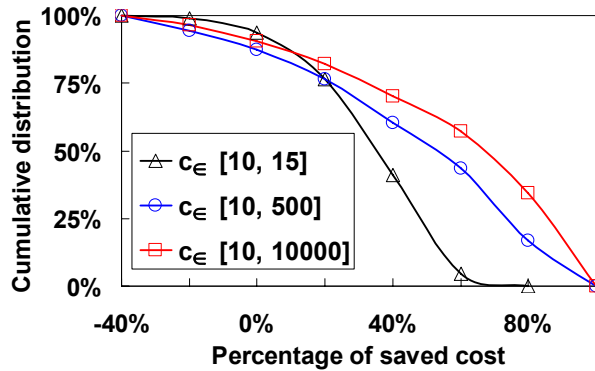


Figure 6. Performance gain vs. sensor network heterogeneity

It is worth noticing that in the two rounds of simulations, we conduct the comparisons using different settings of implications. The experimental results suggest that as long as the sensor networks are correlated with each other, it is always necessary and beneficial to utilize the implications among sensor networks.

4.2 Performance Gain vs. Heterogeneity

In this group of simulations, we evaluate the performance gain of IAP when the original energy costs of sensor networks become heterogeneous. Here $s[1..N_H]$ $[1..N_H]$ conform to a uniform distribution on $[0, 1]$. $n[1..N_H]$ conform to a uniform distribution on $[3, 15]$. In order to simulate sensor network heterogeneity, we let $c[1..N_H]$ respectively conform to a uniform distribution on $[10, 15]$, $[10, 500]$ and $[10, 10000]$. Since $c[1..N_H]$ is uniformly distributed, a larger interval of $c[1..N_H]$ leads to stronger heterogeneity.

As we can see from Figure 6, the performance gain of IAP, compared with the random approach, apparently increases as the sensor networks become more and more heterogeneous. The mean percentages of saved cost are respectively 33.1%, 43.7%, 55.5%. Consider the practical applications, which generally integrate many heterogeneous sensor networks, the cost of sensing in the sensor networks must be quite diverse. The simulation result reveals that IAP is especially suitable and efficient in such environments.

5. CONCLUSION

Demands on better interacting with physical world require an effective and comprehensive collaboration mechanism among heterogeneous sensor networks. In this paper we formalize the concept of collaborative query processing and study the issue of query optimization among heterogeneous sensor networks. Accordingly we design a heuristic approach IAP to minimize the energy costs by utilizing the implications among sensor networks. The simulation results validate our design, which show that IAP greatly enhances the energy efficiency of query processing. Further compared with the random approach, the performance gain of IAP keeps consistent with various types of queries and becomes even greater in heterogeneous environments.

Unified systems of web and sensor networks present a promising direction for integrating various sensor networks to achieve powerful and intelligent functionalities. In the next step, we plan to progress the research in both theoretical and systematical aspects based on our OceanSense [10] project. We also consider the user-oriented optimization as a potential direction of our future work.

6. Acknowledgements

This work is supported in part by NSF China Grant No. 60573140 and Nokia APAC research grant.

7. REFERENCES

- [1] T. Gao, T. Massey, L. Selavo, M. Welsh, and M. Sarrafzadeh, "Participatory User Centered Design Techniques for a Large Scale Ad-Hoc Health Information System," in Proceedings of *HealthNet*, 2007.
- [2] R. Szweczyk, A. M. Mainwaring, J. Polastre, J. Anderson, and D. E. Culler, "An analysis of a large scale habitat monitoring application," in Proceedings of *ACM SenSys*, 2004.
- [3] M. Li, Y. Liu, and L. Chen, "Non-Threshold based Event Detection for 3D Environment Monitoring in Sensor Networks," in Proceedings of *IEEE ICDCS*, 2007.
- [4] P. Zhang, C. M. Sadler, S. A. Lyon, and M. Martonosi, "Hardware design experiences in ZebraNet," in Proceedings of *ACM SenSys*, 2004.
- [5] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cyirci, "Wireless sensor networks: A survey," *Computer Networks*, vol. 38, pp. 393-422, 2002.
- [6] E. F. Nakamura, A. A. F. Loureiro, and A. C. Frery, "Information fusion for wireless sensor networks: Methods, models, and classifications," *ACM Computing Surveys*, vol. 39, 2007.
- [7] Qinhuangdao Oilport, <http://www.qhdport.com>
- [8] N. Xu, S. Rangwala, K. K. Chintalapudi, D. Ganesan, A. Broad, R. Govindan, and D. Estrin, "A wireless sensor network For structural monitoring," in Proceedings of *ACM SenSys*, 2004.
- [9] Z. Yang, M. Li, and Y. Liu, "Sea Depth Measurement with Restricted Floating Sensors," in Proceedings of *IEEE RTSS*, 2007.
- [10] OceanSense, <http://www.cse.ust.hk/~liu/Ocean/>
- [11] F. Ye, G. Zhong, S. Lu, and L. Zhang, "PEAS: A Robust Energy Conserving Protocol for Long-lived Sensor Networks," in Proceedings of *IEEE ICDCS*, 2003.
- [12] J. Hill and D. Culler, "Mica: A Wireless Platform For Deeply Embedded Networks," *IEEE Micro*, vol. 22, pp. 12 - 24, 2002.
- [13] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks," in Proceedings of *ACM MobiCom*, 2000.
- [14] L. Gu and J. A. Stankovic, "t-kernel: Providing Reliable OS Support to Wireless Sensor Networks," in Proceedings of *ACM SenSys*, Boulder, Colorado, USA, 2006.
- [15] K. A. Delin, "The Sensor Web: A Macro-Instrument for Coordinated Sensing," *Sensors*, vol. 2, pp. 270 - 285, 2002.
- [16] S. Nath, J. Liu, and F. Zhao, "SensorMap for Wide-Area Sensor Webs," *IEEE Computer*, vol. 40, pp. 90 - 93, 2007.
- [17] P. B. Gibbons, B. Karp, Y. Ke, S. Nath, and S. Seshan, "IrisNet: An Architecture for a World-Wide Sensor Web," *IEEE Pervasive Computing*, vol. 2, pp. 22 - 33, 2003.
- [18] SensorNet, <http://www.sensornet.gov/>
- [19] M. T. Ozsu and P. Valduriez, *Principles of Distributed Database Systems*, second ed: Prentice Hall, 1999.