An Overview and Classification of Approaches to Information Extraction in Wireless Sensor Networks

Tariq A. A. Alsbouí, Mohammad Hammoudeh, Zuhair Bandar, Andy Nisbet

School of Computing, Mathematics & Digital Technology
Manchester Metropolitan University
Manchester, UK
Email: tariq.al-sboui@stu.mmu.ac.uk, {m.hammoudeh,z.bandar,a.nisbet}@mmu.ac.uk

***Abstract*—Recent advances in wireless communication have made it possible to develop low-cost, and low power Wireless Sensor Networks (WSN). The WSN can be used for several application areas (e.g., habitat monitoring, forest fire detection, and health care). WSN Information Extraction (IE) techniques can be classified into four categories depending on the factors that drive data acquisition: event-driven, time-driven, query-based, and hybrid. This paper presents a survey of the state-of-the-art IE techniques in WSNs. The benefits and shortcomings of different IE approaches are presented as motivation for future work into automatic hybridisation and adaptation of IE mechanisms.**

***Keywords*-Wireless Sensor Networks; Information Extraction**

I. Introduction

The main purpose of a WSN is to provide users with access to the information of interest from data collected by spatially distributed sensors. In real-world applications, sensors are often deployed in high numbers to ensure a full exposure of the monitored physical environment. Consequently, such networks are expected to generate enormous amount of data [1]. The desire to locate and obtain information makes the success of WSNs applications, largely, determined by the accuracy and quality of the extracted information. The principal concerns when extracting information include the timeliness, accuracy, cost, and reliability of the extracted information and the methods used for extraction. The process of IE enables unstructured data to be retrieved and filtered from sensor nodes using sophisticated techniques to discover specific patterns [2]. Practical constraints on sensor node implementation such as power consumption (battery limits), computational capability, and maximum memory storage, make IE a challenging distributed processing task.

In terms of data delivery required by an application, IE in WSNs can be classified into four broad categories: event-driven, time-driven, query-based, and hybrid. In event-driven, data is only generated when an event of interest occurs , while, in the time-driven, data is periodically sent to a sink every constant interval of time. With query-based, the data is collected according to end user's demand. Finally, the hybrid approach is a combination of one or more of the above.

The rest of this paper is organised as follows: Section 2 identifies what types of information needs to be reported to end users. Section 3 looks at event-driven IE approaches and presents sample developments. Section 4, describes time-driven IE and recent successful deployments. Section 5, describes query-based IE and present some of the recent approaches. Section 6, describes and identify recent advances in hybrid IE methods. In section 7, a summary on future research direction for IE is discussed. Section 8 concludes this paper.

II. What needs to be reported?

IE is one of the most vital efforts to utilise the ever burgeoning amount of data returned by WSNs for achieving detailed, often costly task of finding, analysing and identifying needed information. The process of IE involves the classification of data based on the type of information they hold, and is concerned with identifying the portion of information related to a specific fact. In the context of WSNs, the notion of fact can be defined as a property or characteristic of the monitored phenomenon at a certain point in time or during a time interval. Fact can also refer to an event or action. An event is a pattern or exceptional change that occasionally appears in the observed environment [3]. Events have some distinct features that can be used as thresholds, e.g. temperature > 50, to make a distinction between usual and unusual environmental parameters.

An event may arise in many other forms. It can be a continuous, gradually occurs over time (e.g. temperature does not change instantly), and has obvious limit with normal environment parameters. In [4], complex events are defined as sequences of sensor measurements over a period of time indicating an unusual activity in the monitored environment. In WSNs, the network owners may be unaware in advance what type of events may occur. This is because one of the ultimate goals of such networks is to discover new events and interesting information about the monitored phenomenon. For this reason, threshold based event detection methods are not always efficient to identify and extract event-based facts. From this deficiency arise the need for periodic, query-based, and hybrid IE approaches.

Events can be further classified into two categories: system events and environmental events. System events are concerned with architectural or topological changes, e.g. a mobile node entered a cluster area. Environmental events are concerned

 ISBN: 978-1-61208-144-1

Table I
OVERVIEW OF THE SELECTED APPROACHES TO IE

Event-driven	[3], [7]–[15]
Time-driven	[9], [16]–[20]
Query-based	[21]–[29]
Hybrid	[30]–[34]

with the occurrences of unusual changes across the monitored environment, e.g. spotting a moving target [4].

Nodes organisation plays an important role in IE because it defines, among other factors, the cost (amount of energy required to collect raw data), accuracy (level of coverage), reliability (e.g. timeliness) of extracted information. The organisation of nodes can be either centralised or hierarchical. In the centralised approach, data collected by all nodes are sent towards a sink node using single or multi-hop communication [5]. However, this approach does not provide scalability, which is a main design factor for WSN. Also, it causes communication bottlenecks and transmission delays due to congestions especially in areas around the sink [6]. To overcome the problems in the centralised approaches, hierarchical techniques has been proposed as an effective solution for achieving longer network lifetime and better scalability.

Since the number of existing IE approaches is significantly large, it will not be feasible to provide a detailed description of each approach. Instead, we have selected recent approaches that particularly represent directions of future research without focusing on the details of these approaches. However, characteristics of various approaches that are common for the approach they apply will be presented. Table I lists the reviewed approaches and some older approaches for the more interested readers. In Sections III to VI the approaches are presented based on the categorisation so as related subcategories are discussed in the common context. To make the analysis of different approaches more logical and to set up a common base for their comparison and connection we consider some qualitative criteria.

III. EVENT-DRIVEN IE

A. Description and Operation

In event-driven approaches to IE, the initiative is with the sensor node and the end user is in the position of an observer, waiting for incoming information. Any node may generate a report when a significant event (e.g, a change of state) or an unusual event (e.g, fire) occurs. Event-driven is valuable for detecting events as soon as they occur over a specified region. In the simplest form, sensor nodes are pr-econfigured with threshold values that when exceeded indicate an event.

Event-driven approaches incur low power consumption and requires low maintenance. Among the benefits of this class of approaches are: they reduce the amount of communication overhead by applying local filtering on collected data to determine whether to send new data or not; they implement local mechanisms to prevent multiple nodes reporting the same event; they exploit redundancy to reduce the number of false alarms; they allow timely responses to detected events; they are easy to implement and configure; they allow distributed processing at the node level or within a group of node to collaboratively detect an event; and they are suitable for time critical applications, e.g. forest fire monitoring or intrusion detection.

However, there are a number of limitations to the event-driven IE. First, it is difficult to capture events of spatio-temporal characteristics. Second, detecting complex event may require non trivial distributed algorithms, which require the involvement of multiple sensor nodes [9]. Third, due to the fact that events occur randomly, some nodes generate higher rates of data than other nodes. This will lead to unbalanced workloads among sensor nodes. Fourth, it is not suitable for continuous monitoring applications, where sensed measurements change gradually and continuously. Finally, due to sensors measurement inaccuracies, event-driven approaches may potentially generate false alarms.

B. Event-driven Approaches to IE

In earlier studies, events were detected with a user-defined threshold values [7], [8]. In such approaches sensor nodes are pr-econfigured with a static threshold value. When the sensor node reading deviates from the pre-defined thresholds, this indicates an event, which triggers the node to convey it is data back to the sink. To overcome some of the inherent problems in the threshold-based event-detection, [13] presented a data fusion tool to increase resilience of event detection techniques. They introduced two levels for event detection: at the first level, each sensor node will individually decide on detecting event using classifier (naive bayes). At the second level, fusion technique is placed in a higher level (e.g, cluster head) and used to distinguish between outliers. Outliers are measurements that differ from the normal pattern of sensed data occurring at individual nodes and events that more nodes agree upon [15]. The data fusion approach reduces the number of transmission, thus extends the network life time. It also reduces the number of false alarms, since cluster heads are able to distinguish between anomalies and event. Because of its distributed nature, this approach is scalable. However, processing data at the cluster head introduces delays in reporting an event. Moreover, the efficiency of the approach depends on the efficiency of cluster formation methods. For instance, many clustering algorithms result in unbalanced clusters.

Another threshold-based approach [12] introduced double decision mechanisms. A sensor may decide about the presence of an event of interest either directly or asking for additional data from nearby nodes. This approach minimises the energy consumption since the final process detection is activated only when it is needed, and there is no need for fusion centre to process the data as a fixed number of nodes will take the responsibility to make decisions about the occurrences of an event. However, it is always difficult to determine node's neighbours. Although these approaches can reduce communication overhead and report events promptly, however, it is difficult to define the optimal threshold values. Also, the

 ISBN: 978-1-61208-144-1

complications of implementing an efficient sleep-wake cycles may dissipate the gained energy savings.

More advanced approaches, such as SAF [35] and Ken [18], exploit the fact that physical environments frequently exhibit predictable stable and strong attribute correlations to improve compression of the data communicated to the sink node. The basic idea is to use replicated dynamic models to reflect the state of the environment being monitored. This is done by maintaining a pair of dynamic probabilistic models over the WSN attributes with one copy distributed in the network and the other at the sink. The sink computes the expected values of the WSN attributes according to the defined prediction model and uses it extract information. When the sensor nodes detect anomalous data that was not predicted by the model within the required certainty level, they route the data back to the sink. This approach is subject to failure as basic suppression. It does not have any mechanism to distinguish between node failure and the case that the data is always within the error bound. Ken is not robust to message loss; it relies on the Markovian nature of the prediction models to presume that any failures will eventually be corrected with model updates, and the approximation certainty will not be affected by the missed updates. They propose periodic updates to ensure models can not be incorrect indefinitely. This approach is not suitable for raw value reconstruction; for any time-step where the model has suffered from failures and is incorrect, the corresponding raw value samples will be wrong. Finally, as the approach presented in [17], SAF and Ken can only handle static network models.

A decentralised, lightweight, and accurate event detection technique is proposed in [36]. The technique uses decision trees for distributed event detection and a reputation-based voting method for aggregating the detection results of each node. Each sensor node perform event detection using it is own decision tree-based classifier. The classification results, i.e. detected events, from several nodes are aggregated by a higher node, e.g, a cluster head. Each node sends it is detected events, called detection value, to all other nodes in it is neighbourhood. The detection value will be stored in a table. Finally, tables are sent to the voter (e.g. cluster head), which in turns decides to make a final decision among different opinion. The decision tree approach provides accurate event detection and characterised by low computational and time complexities. However, the processing of data at the cluster head will introduce further delays in reporting an event.

IV. Time-driven IE

A. Description and Operation

In time-driven approaches to IE, a sensor node periodically generates a report from the physical environment to give the end-user its current status. The reporting period may be preconfigured or set by the end-user depending on the nature of the monitored environment and applications requirements.

Time-driven approaches have the ability to enable arbitrary data analysis, they provide continuous monitoring of the WSN to reflect environmental changes, they scale to handle millions of nodes (through aggregation), they extend network life time by sending nodes to sleep between transmissions, they can reduce congestion and improve system reliability by scheduling nodes to transmit at different times, they explicitly incorporate resource capacity, and highlights unused resources. However, there are a number of limitations to the time-driven approaches. First, they are limited to specific set of applications where consistent changes occur across the network, e.g. agricultural applications. Second, a large portion of the returned data might be redundant and not useful for the end-user thereby resulting in wastage of resources. Third, nodes have to maintain global clock and deal with synchronisation issues. Finally, it is extremely difficult to define optimal time intervals.

B. Time-driven Approaches to IE

In time-driven IE, most of the published work in the literature is based on probabilistic models that attempt predict the next value that the sensor is expected to acquire. For example, Ken's [18] model exploits the spatio-temporal data correlations while guaranteeing correctness. It involves placing a dynamic probabilistic model on the sensor node and on the sink, and these models are always kept in synchronisation for periodic updates. Similar approaches to Ken have been suggested in [9], [37]. In contrast to Ken, these approach uses dynamically changing subset of the nodes as samplers where the sensor readings of the sampler nodes are directly collected, while, the values of non sampler nodes are predicated through probabilistic models that are locally and periodically constructed. All approaches in [9], [18], [37] save energy by reducing the number of transmitted messages. However, the additional cost to maintain models synchronised is not negligible.

Another approach called Cascading Data Collection (CDC) is presented in [19]. In CDC only a subset of sensor nodes are selected randomly to periodically transfer data back to the sink node. The mechanism is distributed and only utilises local information of sensor node. The CDC reduces communication cost by allowing only a subset of sensor nodes to periodically transmit readings back to the sink. However, the CDC uses packet aggregation at an intermediate node, which introduces undesirable communication delays. The work presented in [38] takes CDC one step further by enabling each node to use its local and neighbourhood state information to adapt its routing and MAC layer behaviour.

V. Query-based IE

A. Description and Operation

Query-based approaches to IE, typically involve request-response interactions between the end-user or application components and sensor nodes. End users issue queries in an appropriate language, and then each query is disseminated to the network to retrieve the desired data from the sensors based on the description in the query.

Query based approaches provide a high level interface that hides the network topology as well as radio communication from end users. Queries can be sent on demand or at fixed intervals. They provide a solution if the data needs to be retrieved from the entire network.

However, there are a number of limitations to the query-based approaches. First, most of existing query languages do not provide suitable constructs to easily articulate spatio-temporal sense data characteristics. Second, it is difficult to formulate queries using current languages that represent higher level behaviour, or specify a subset of nodes that have significant effect on the query answer. This may result in generating large amount of data of which big portion is not useful for the end user. Third, to the best of our knowledge, there is no published work that fully exploits all the potentials of different heterogeneous resources in WSN applications in a context-aware manner. Forth, approaches that take a database view of the network are inclined more towards the extraction of the reactive behaviour of the WSN and suggestions were made that the active database should be viewed as two end-points of the range of rule-based languages in databases [39]. Finally, though declarative languages are came into view in WSNs settings, the trigger that are the fundamental means for specifying the reactive behaviour in a database have not yet been maturely developed.

B. Query-based Approaches to IE

Query-based systems, applies techniques used in traditional database systems to implement IE. A query is sent to the network and data is collected according to the description in the query. COUGAR [21] was the first project that attempted to introduce the concept of WSN as a distributed database. It allows the end user to issue a declarative query (SQL) for retrieving information. The authors introduced a query layer between the application layer and the network layer. The query layer comprises a query proxy, which is placed on each sensor node to interact with both the application layer and the networking layer. The goal of the query proxy is to perform in-network processing. In-network processing increases efficiency in terms of power consumption, and reduces the amount of data that needs to be sent to the gateway node. The user does not need to have knowledge about the network, or how the data is retrieved or processed. However, COUGAR is incapable of capturing complex events, e.g. of spatio-temporal nature, or a produce queries that targets only a subset of the network [22].

A similar approach to COUGAR is proposed in [23]. TinyDB is a query processing system, which extracts information from the data collected by the WSN using the TinyOS operating system. TinyDB maintains a virtual database table called SENSORS. It disseminates the queries throughout the network by maintaining a routing tree (spanning tree) rooted at the end point (usually the user's physical location). Every sensor node has its own query processor that processes and aggregates the sensor data and maintains the routing information. TinyDB is extensible and complete framework with effective declarative queries. In-network processing reduces the amount of data that is required to be sent to the sink, thus, energy consumption is reduced. However, data does not include the georeferencing of sensor nodes for spatial quires, and tight correlation among routing and queries.

In [24], a new data collection algorithm that aims on reduce energy consumption by focusing on selective aggregate queries. The proposed algorithm, named, PDT (Pocket Driven Trajectories) deals with queries that aggregate data only from a subset of all network nodes. PDT is based on the logical assumption that spatial correlation in sensor values coupled with query selectivity gives rise to a subset of participating nodes formed by one or more geographically clustered sets (pockets). The algorithm starts by discovering the set of pockets for a given query. Then, the aggregation tree to the spatially optimal path connecting these pockets is aligned. The PDT algorithm reduces the amount of communication and is scalable for large WSNs. However, PDT introduces a delay in reporting data to a sink, because data is processed at an intermediate node.

In [25], a mobile sink moving through the sensing field issues a query to a specific area. The sensor node that is closest to the centre of the area of interest elects itself as a cluster head. The cluster head performs data collection and aggregation, then the aggregated data is sent back to the mobile sink. The proposed mobile sink approach saves energy by choosing an optimal time and location to disseminate query. The area-based querying and the mobile sink makes the approach scalable for large-scale WSNs. However, the proposed approach is limited to a set of applications, specifically, intelligent transportation system and environmental monitoring. Also, it introduces undesirable delay since the data is aggregated at the mobile sink.

The authors in [26], proposed a query processing algorithm, that allows the user to specify a value and time accuracy constraints based on an optimised query plan. Using these optimisation constraints, the algorithm can find an optimal sensing and transmission of attribute readings to sink node. Rather than sending sensors readings directly to the sink, the proposed algorithm report only updates. This results in considerable reduction in communication costs. However, the algorithm does not support dynamic adjustment of accuracy constraints.

More recently in [27], the authors designed and implemented a distributed in-network query processing, called Corona. Corona is composed of three components: the query engine that is executed on the sensors; a host system on the clients PC that is connected to the sink; and GUI that is connected to the host system via TCP/IP. The Corona query processing provides multi-tasking capabilities by running multiple queries concurrently, which in turns reduces processing delays and communications cost by applying data aggregation. However, the language can not easily capture spatio-temporal events.

VI. Hybrid-based IE

A. Description and Operation

A hybrid approach is an approach that combines the functionality of two or more algorithms from different IE categories. Hybrid approaches aims to minimise the effect of the disadvantages of individual IE categories described above.

B. Hybrid Approaches to IE

Many hybrid approaches to IE have been recently proposed in the literature. In [34], the authors proposed a hybrid protocol that adaptively switches between time-driven and event-driven data collection. A sensor node is triggered to detect an event of interest, and from the point when an event detected to the point when the event becomes no longer valid, the protocol switches to behave as a time-driven protocol. During this period sensor nodes continuously report data to the sink. This protocol reduces unnecessary data transmission and minimise event notification time. However, it is not guaranteed to work well for all applications due to limitations of the PAD algorithm, such as if sensor nodes detecting an event are located at the border between clusters, those nodes in other clusters can be included only when clusters at the same level have used time-driven data dissemination

More recently, in [33], the authors proposed a hybrid framework similar to [9], [18], which deploy both of event-driven and query-based approaches to IE. The idea is to process continuous group-by aggregate queries, and to allow each sensor node to check whether sensor readings satisfy local predicates based on a predefined thresholds. Then, nodes send only data that satisfy local predicates to their cluster heads, which in turns process the data to answer the query as accurate as possible. The proposed hybrid framework is able to target a subset of the network by using the group-by clause. It reduces communication cost by using one dimensional haar wavelets. However, it introduces a delay in reporting events since the data is processed at the cluster head.

In [32], the authors proposed energy-efficient hybrid data collection architecture similar to [25]. The aim is to enhance the network performance and reduce the total energy consumption by introducing mobile node entities. A mobile node is moving through the network deployment region to collect data from the static nodes over a single hop radio links. The mobile node visits the sink periodically to drop off the collected data. The proposed solution reduces energy consumption and communication overhead by moving the sink node near to the nodes to collect data. However, the mobile node introduces latency in transferring the data as it has to travel back to a sink.

VII. Discussion and possible future directions

Before concluding this paper, this section provides a discussion about research issues, and future directions in the area of IE in WSNs. This short survey revealed that most of the existing approaches to IE suffer from inherent problems that limit their applications including: they are application specific; characterised by poor spatio-temporal IE capabilities; consume high power; many approaches trade the amount and quality of returned information by energy consumption; they lack appropriate high-level interfaces that allow the user to set thresholds and issue queries; and the tight coupling between IE algorithms, applications, and hardware stacks leads to lack of code reuse. The lack of development frameworks means each new application has to be tackled from the ground up. These issues limit the usefulness of the developed IE approaches, making it hard to use them on anything other than the application it was designed for.

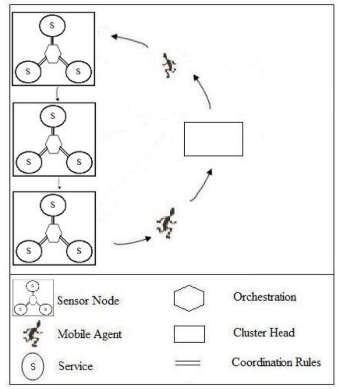

Figure 1. An integrated IE framework for WSNs.

The problems and limitations presented above are the opportunities we intend to follow in our future work. Possible solutions that we are currently investigating for the integration of the three IE approaches will be achieved through the use of coordination rules [40] and mobile agents [41].

Coordination rules are a set of modelling primitives, design principles and patterns that deal with enabling and controlling the collaboration among a group of software distributed agents performing a common task. If each algorithm in each IE category is viewed as a service, then the composition of these services will result in a complete IE framework. Service composition provides new services by combining existing services. The coordination rules specifies the order in which services are invoked and the conditions under which a certain service may or may not be invoked.

The mobile agent paradigm will be adopted to facilitate cooperation among services on different nodes. Mobile agent is a piece of software that performs data processing autonomously while migrating from node to node [41]. The agent can collect local data and perform any necessary data aggregation. Mobile agents can make decision autonomously without user input. They provide flexibility in terms of decision making, and reliability in terms of node failure [42].

Figure 1 shows an illustration of the described hybrid framework. It shows how services on one node are connected and how a service can access other services on remote node.

VIII. Conclusion

The main objective of this paper is to provide an understanding of the current issues in this area for better future academic research and industrial practice of WSNs IE. We have presented a review of the state of the art for IE approaches in WSNs. We discussed various approaches to IE. We also discussed the challenges as well as future research directions in developing a complete integrated WSNs IE framework.

 ISBN: 978-1-61208-144-1

REFERENCES

[1] M. Hammoudeh, R. Newman, and S. Mount, “An approach to data extraction and visualisation for wireless sensor networks,” *International Conference on Networking*, vol. 0, pp. 156–161, 2009.

[2] S. Sarawagi, “Information extraction,” *Found. Trends databases*, vol. 1, pp. 261–377, March 2008.

[3] C. Zhang, C. Wang, D. Li, X. Zhou, and C. Gao, “Unspecific event detection in wireless sensor networks,” *Communication Software and Networks, International Conference on*, vol. 0, pp. 243–246, 2009.

[4] R. Bhargavi, V. Vaidehi, P. Bhuvaneswari, P. Balamurali, and G. Chandra, “Complex event processing for object tracking in wireless sensor networks,” *Web Intelligence and Intelligent Agent Technology, IEEE/WIC/ACM International Conference on*, vol. 3, pp. 211–214, 2010.

[5] A. Iranli, M. Maleki, and M. Pedram, “Energy efficient strategies for deployment of a two-level wireless sensor network,” in *Proceedings of the 2005 international symposium on Low power electronics and design*, 2005, pp. 233–238.

[6] K. Akkaya and M. Younis, “A survey on routing protocols for wireless sensor networks,” *Ad Hoc Networks*, vol. 3, no. 3, pp. 325 – 349, 2005.

[7] M. L. Segal, F. P. Antonio, S. Elam, J. Erlenbach, and K. R. De Paolo, “Method and apparatus for automatic event detection in a wireless communication system,” no. 6124810, September 2000.

[8] G. Werner-Allen, K. Lorincz, M. Welsh, O. Marcillo, J. Johnson, M. Ruiz, and J. Lees, “Deploying a wireless sensor network on an active volcano,” *IEEE Internet Computing*, vol. 10, pp. 18–25, 2006.

[9] B. Gedik, L. Liu, and P. S. Yu, “Asap: An adaptive sampling approach to data collection in sensor networks,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 18, pp. 1766–1783, 2007.

[10] K. Pripužić, H. Belani, and M. Vuković, “Early forest fire detection with sensor networks: Sliding window skylines approach,” *Proceedings of the 12th international conference on Knowledge-Based Intelligent Information and Engineering Systems, Part I*, pp. 725–732, 2008.

[11] H. bin Wang, Z. Yuan, and C. dong Wang, “Intrusion detection for wireless sensor networks based on multi-agent and refined clustering,” in *Communications and Mobile Computing, 2009. CMC ’09. WRI International Conference on*, vol. 3, 2009, pp. 450 –454.

[12] S. Sardellitti, S. Barbarossa, and L. Pezzolo, “Distributed double threshold spatial detection algorithms in wireless sensor networks,” in *Signal Processing Advances in Wireless Communications, 2009. SPAWC ’09. IEEE 10th Workshop on*, 2009, pp. 51 –55.

[13] N. H. P. Bahrepour, M. Meratnia, “Sensor fusion-based event detection in wireless sensor networks,” *Mobile and Ubiquitous Systems: Networking Services, MobiQuitous, 2009. MobiQuitous ’09. 6th Annual International*, pp. 1–8, 2009.

[14] M. Bahrepour, N. Meratnia, and P. Havinga, “Use of ai techniques for residential fire detection in wireless sensor networks,” in *AIAI 2009 Workshop Proceedings*, vol. 475, July 2009, pp. 311–321.

[15] Y. Zhang, N. Meratnia, and P. Havinga, “Outlier detection techniques for wireless sensor networks: A survey,” *Communications Surveys Tutorials, IEEE*, vol. 12, no. 2, pp. 159 –170, 2010.

[16] J. Polastre, R. Szewczyk, A. Mainwaring, D. Culler, and J. Anderson, *Analysis of wireless sensor networks for habitat monitoring*, 2004, pp. 399–423.

[17] A. Deshpande, C. Guestrin, S. R. Madden, J. M. Hellerstein, and W. Hong, “Model-driven data acquisition in sensor networks,” in *Proceedings of the Thirtieth international conference on Very large data bases - Volume 30*, 2004, pp. 588–599.

[18] D. Chu, A. Deshpande, J. M. Hellerstein, and W. Hong, “Approximate data collection in sensor networks using probabilistic models,” in *Proceedings of the 22nd International Conference on Data Engineering*, 2006, pp. 48–.

[19] L. L. Li. Hong, Yu. Hongyi and A. Liu, “On the cascading data collection mechanism in wireless sensor networks,” *Wireless Communications, Networking and Mobile Computing, 2007*.

[20] L. Wang and A. Deshpande, “Predictive modeling-based data collection in wireless sensor networks,” in *Proceedings of the 5th European conference on Wireless sensor networks*, 2008, pp. 34–51.

[21] Y. Yao and J. Gehrke, “The cougar approach to in-network query processing in sensor networks,” *SIGMOD Rec.*, vol. 31, pp. 9–18, 2002.

[22] M. Bestehorn, K. Böhm, E. Buchmann, and S. Kessler, “Energy-efficient processing of spatio-temporal queries in wireless sensor networks,” in *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ser. GIS ’10, pp. 340–349.

[23] S. R. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, “Tinydb: an acquisitional query processing system for sensor networks,” *ACM Trans. Database Syst.*, vol. 30, pp. 122–173, 2005.

[24] M. Umer, L. Kulik, and E. Tanin, “Optimizing query processing using selectivity-awareness in wireless sensor networks,” *Computers, Environment and Urban Systems*, vol. 33, no. 2, pp. 79 – 89, 2009.

[25] L. Cheng, Y. Chen, C. Chen, and J. Ma, “Query-based data collection in wireless sensor networks with mobile sinks,” in *Proceedings of the 2009 International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly*, 2009, pp. 1157–1162.

[26] J.-Z. Sun, “An energy-efficient query processing algorithm for wireless sensor networks,” in *Ubiquitous Intelligence and Computing*, 2008, vol. 5061, pp. 373–385.

[27] R. Khoury, T. Dawborn, B. Gafurov, G. Pink, E. Tse, Q. Tse, M. Gaber, U. Rohm, and B. Scholz, “Corona: Energy-efficient multi-query processing in wireless sensor networks,” in *Database Systems for Advanced Applications*, 2010, vol. 5982, pp. 416–419.

[28] P. Bonnet, J. Gehrke, and P. Seshadri, “Towards sensor database systems,” in *Proceedings of the Second International Conference on Mobile Data Management*, 2001, pp. 3–14.

[29] C. Chien-Chung Shen. Srisathapornphat, C. Jaikaeo, “Sensor information networking architecture and applications,” *Personal Communications, IEEE*, vol. 8, pp. 52–59, 2001.

[30] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energy-efficient communication protocol for wireless microsensor networks,” in *Proceedings of the 33rd Hawaii International Conference on System Sciences-Volume 8 - Volume 8*, 2000, pp. 8020–.

[31] T. E. Daniel, R. M. Newman, E. I. Gaura, and S. N. Mount, “Complex query processing in wireless sensor networks,” in *Proceedings of the 2nd ACM workshop on Performance monitoring and measurement of heterogeneous wireless and wired networks*, 2007, pp. 53–60.

[32] A. Rasheed and R. Mahapatra, “An energy-efficient hybrid data collection scheme in wireless sensor networks,” in *Intelligent Sensors, Sensor Networks and Information, 2007. ISSNIP 2007. 3rd International Conference on*, 2007, pp. 703 –708.

[33] C.-H. Lee, C.-W. Chung, and S.-J. Chun, “Effective processing of continuous group-by aggregate queries in sensor networks,” *J. Syst. Softw.*, vol. 83, pp. 2627–2641, December 2010.

[34] B.-D. Lee, “Adaptive data dissemination protocol for wireless sensor networks,” in *Security-Enriched Urban Computing and Smart Grid*, 2010, vol. 78, pp. 188–195.

[35] D. Tulone and S. Madden, “An energy-efficient querying framework in sensor networks for detecting node similarities,” in *Proceedings of the 9th ACM international symposium on Modeling analysis and simulation of wireless and mobile systems*, 2006, pp. 191–300.

[36] M. Bahrepour, N. Meratnia, M. Poel, Z. Taghikhaki, and P. J. Havinga, “Distributed event detection in wireless sensor networks for disaster management,” in *Proceedings of the International Conference on Intelligent Networking and Collaborative Systems, INCoS 2010*, November 2010, pp. 507–512.

[37] G.-S. Ahn, “Information-driven tracking and access control in wireless ad hoc and sensor networks,” Ph.D. dissertation, Columbia University, 2009.

[38] R. Jurdak, P. Baldi, and C. Videira Lopes, “Adaptive low power listening for wireless sensor networks,” *IEEE Transactions on Mobile Computing*, vol. 6, pp. 988–1004, 2007.

[39] G. Trajcevski and P. Scheuermann, “(reactive /+ proactive behavior) situation awareness in sensor networks,” in *Workshop on Research Directions in Situational-aware Self-managed Proactive Computing in Wireless Adhoc Networks*, 2009.

[40] J. Abreu and J. Fiadeiro, “A coordination model for service-oriented interactions,” in *Coordination Models and Languages*, 2008, vol. 5052, pp. 1–16.

[41] M. Chen, T. Kwon, Y. Yuan, and V. Leung, “Mobile agent based wireless sensor networks,” *Journal of Computers*, vol. 1, no. 1, 2006.

[42] D. Massaguer, C.-L. Fok, N. Venkatasubramanian, G.-C. Roman, and C. Lu, “Exploring sensor networks using mobile agents,” in *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, ser. AAMAS ’06, 2006, pp. 323–325.