Quality of Information in Wireless Sensor Networks: A Survey[1]

(**Completed paper**)

Vinay Sachidananda, Abdelmajid Khelil, Neeraj Suri.
Technische Universität Darmstadt
Dependable, Embedded Systems and Software Group
Hochschulstr. 10, 64289 Darmstadt, Germany.
{vinay,khelil,suri}@cs.tu-darmstadt.de
Category: Methods, concepts, and tools for IQ (IQ in Sensor Networks and Information Fusion)

Abstract: In Wireless Sensor Networks (WSNs) the operating conditions and/or user requirements are often desired to be evolvable, whether driven by changes of the monitored parameters or WSN properties of configuration, structure, communication capacities, node density, and energy among many others. While considering evolvability, delivering the required information with the specified quality (accuracy, timeliness, reliability etc) defined by the user constitutes a key objective of WSNs. Most existing research efforts handle fluctuations of operation conditions in order to deliver information with the highest possible specified quality. In this paper, we take these aspects into consideration and survey existing work on Quality of Information (QoI). As a contribution, we categorize WSN information into a set of abstract classes for generality across varied application types. Our survey shows that currently QoI is usually addressed in isolation by focusing on discrete data processing operations/building blocks such as raw data collection, in-network processing (compression, aggregation), information transport and sink operations for decision making. This survey comprehensively explains the different views of QoI on attributes, metrics and WSN functional operations mapped with existing approaches. The survey also forms the basis for specifying needed QoI research issues.

Keywords: Sensor Data, Information, Quality of Information, Wireless Sensor Networks.

INTRODUCTION

Wireless Sensor Networks (WSNs) are a distributed collection of sensor nodes having potential in domains such as monitoring, automation, health-care etc. The core functionality of WSNs is (a) to generate the minimal necessary raw data, (b) to process this data in-network (close to its source) in order to extract the relevant information specified by the user, and (c) to deliver the information to the user. It is noteworthy to distinguish the terms data and information. Data refers to basic monitored facts/chunks (e.g., sensor readings) and information is the collated and interpreted data systematized by purposeful acumen and processing required for an application (e.g., event occurrence). There exist a few classifications on information [37] [38] which are unfortunately specific to certain WSN functionalities. Therefore, we present a comprehensive classification of WSN information in this paper.

[1] Research supported in part by DFG GRK 1362 (GKmM), EU INSPIRE and EU CoMiFin.

In WSNs, the various applications and users drive the specific information needs. The user requirements regarding information are evolvable having specified information with a certain quality. Also the achievable information quality is evolvable according to the operating conditions such as network and environmental conditions. Accordingly, the WSN functional operations should be designed while considering the fluctuating operating conditions and the user's evolvable requirements on information quality. We refer to quality as the degree or grade of excellence, and to Quality of Information (QoI) as the quality experienced/perceived by the user concerning the received information, which (may) fully accomplish the user evolvable requirements while saving valuable resources such as energy and bandwidth.

Similar to Quality of Service (QoS) in traditional networks, QoI is significant in WSNs and considered as the center of attraction for users, designers, decision makers, application planners etc. There exists no survey detailing the attributes/metrics/techniques related to QoI. We take the opportunity to review the snapshot of the state-the-art of this emerging research field, and to discuss the pros and cons of the different existing QoI approaches.

Paper Objectives: Overall, the paper targets the ongoing research activities that attempt to address QoI in a manner which provides the foundation for the design, deployment and operation of WSNs. To build a common understanding and overcome the ambiguity in different existing definitions, we provide a generic definition of QoI. We also accommodate the discussion and comparison of QoI classification and identify the properties and characteristics of QoI for WSN. To this end, we classify the WSN operations/building blocks into different classes and then map the existing QoI approaches to them. Accordingly, we briefly summarize the existing approaches mentioning the building blocks they are concentrating and what the effects of neglecting other blocks are. In addition, we determine the way in which functional properties depend on and can be affected by various other features like deployment. Hereby, we provide an account, analysis of the design features, solutions, pros and cons that have been adopted by current QoI frameworks and methods [47] [12] [48] [37] [6].

Currently, QoI is addressed isolated by focusing on well-separate data processing operations/blocks comprising raw data collection, in-network processing (compression, aggregation etc.), information transport and sink operation for decision making. These blocks are present from the source (raw data creation) to the sink (information delivery to the user). We argue that QoI can satisfy the user evolvable requirements when all or combinational blocks are considered. Considering the different blocks as whole, the challenge lies in delivering the information just not by having the best techniques in the different blocks to deliver high QoI, but sometimes require tuning the techniques to deliver only required quality. Accordingly, this survey paper briefly outlines a future research map for QoI in WSN beyond the survey/review contribution. We mainly (a) propose and argue for a holistic view for QoI, and (b) we propose to quantify the QoI as the user evolvable requirements may be not satisfied while processing the data/information from the source to the sink.

Paper Organization: In next section, we present a generic architecture and system model, followed by classification of information, providing information attributes, defining QoI, other Qo* approaches related to QoI and assessment of QoI. The paper later focuses on explaining building blocks with different views of QoI followed by the survey and comparison of the existing approaches in relation to the information class, building blocks, attributes and metrics. Finally, we provide a preliminary research roadmap and conclude the paper.

ARCHITECTURE AND SYSTEM MODEL

Usually, in WSNs the raw data collected undergoes in-network processing and then transported to the sink for decision making. To define decision making let's first define decision, meaning a selection between alternatives. *Decision making is a choice between one or more paths of action.* One important factor in data processing operations is sensor data/decision fusion [11]. The value fusion is the operation, where raw data from different local nodes are fused, i.e., filtered, aggregated, etc. The decision fusion is the operation where the local decisions from many sensor nodes are fused [46] [7]. The sensor nodes completing the fusion are known as fusion centers. Usually, fusion centers fuse local decisions from *n* sensor nodes into one user relevant decision for detecting a certain event. We assume a feedback channel to transport information back to sensor nodes as show in Figure 1. Feedback channel is also used for user requirements dissemination. We refer to the different operations on raw data and then on the constructed information by the building blocks of the WSN.

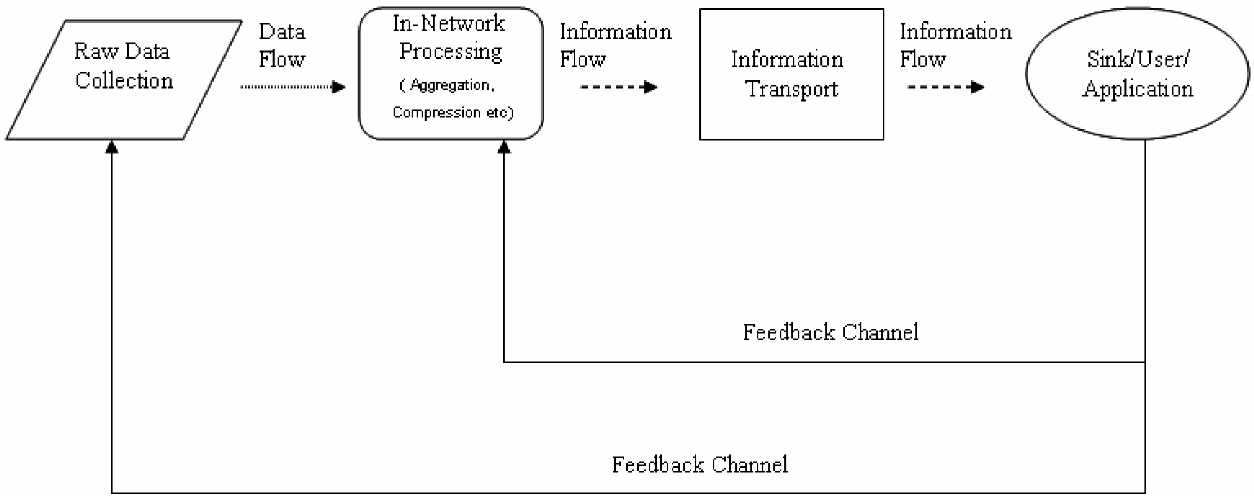

Figure 1: WSN Functional Blocks and Communication Channels

QoI: Definition, Assessment and Provisioning

In this section, we first classify the information into different classes and provide information attributes. Next, we propose a novel definition of QoI that is based on user evolvable requirements, information attributes and resources. Then, we discuss the aspects of metrics to quantitatively evaluate the QoI delivered by a WSN. Finally, we review the QoI provisioning approaches.

Information Classification

The existing classifications of information in literature [37] [38] are more specific towards information transport and service level. For instance, the classification in [37] is concentrating on reliable information transport itself, and not on the other building blocks. However, we require a classification of information which is general in order to suit to the holistic QoI approach, i.e., considering all the building blocks.

Information is usually present in different forms depending on the user evolvable requirements. For example there might be different user's pertaining to obtain information from the same event of interest. However, the information can be different depending on the requirements such as perimeter, wind, temperature, humidity etc. Without loss of generality, we identify three classification criteria's for common types of information that a WSN can deliver: Usage time of data (Criteria 1 or C1), type of the information (Criteria 1 or C2), and construction location of the information (Criteria 3 or C3). In Table I, we summarize all the resulting classes.

Prediction Information [C1-P]: Refers to the information that can be forecasted before its occurrence. Examples are predicted important events such as, user events and forecasted network partitioning [54].

Real-Time Information [C1-R]: Is the information which is created just after (during) the occurrence of event. It is required that this information reach the sink/user with a best effort latency.

Historic Information [C1-H]: Refers to the information that is of interest after their occurrence/creation. It is usually stored at the sink. Examples of applications that require such information are forensics, statistics etc.
Query Result Information [C2-Q]: This information is the result of a user query to the deployed network. Example is the average temperature in a certain region of the network.
Event Information [C2-E]: Is a set of attributes of an event of interest that occurred in the sensor field. Examples are: the type, time and location of the detected event.

Information Classifications	**Types of Information**
Criteria 1: Usage time of Information [C1]	Prediction Information [P] Real-Time Information [R] Historic Information [H]
Criteria 2: Type of the Information [C2]	Query Result Information [Q] Event Information [E]
Criteria 3: Construction location of the Information [C3]	Information Constructed In-Network [N] Information Constructed at the Sink [S]

Table I: Information Classifications

From the above generalized classification, almost all the building blocks, characteristics, applications and services fit into one of the categories. However, the first classification is the future perspective class. Whereas, the second classification is well established and know. On the other hand, we are more interested in the construction of information, because of its relation to the in-network processing, information creation, information transport and sink operations. Therefore, we further classify it. Without loss of generality, in WSNs information is created from raw data either:
Information Constructed In-Network [C3-N]: the information is created within the network.
Information Constructed at the Sink [C3-S]: the information is created at the sink.

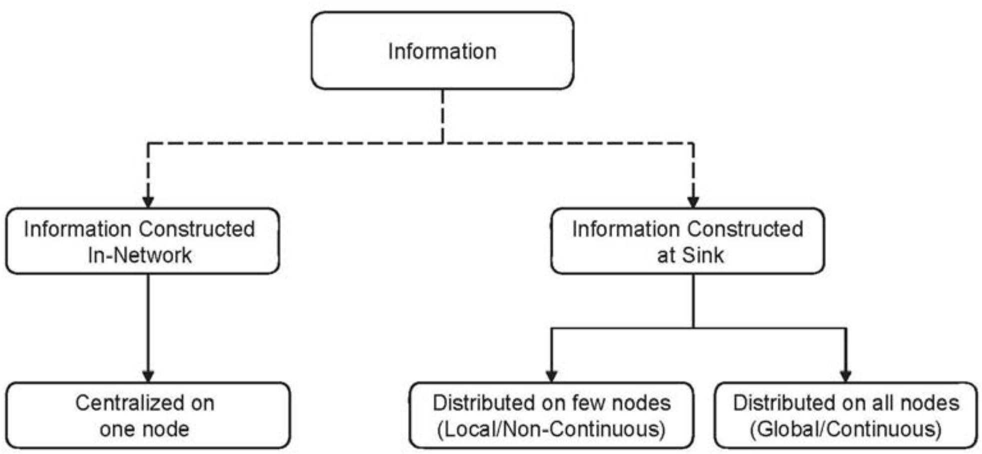

Figure 2: The Adopted Classification of Information

We know further sub classify the classification according to Criteria 3 as follows:
Information constructed in-network on a centralized sensor node. The centralized node has received/processed the necessary raw data or local decisions from different sensor nodes.
Information constructed at the sink with a few distributed sensor nodes which are local and non-continuous. The raw data is sent to the sink from a few sensor nodes from the region of event of interest.
Information constructed at the sink with a distribution on all sensor nodes which are global and continuous. The raw data is sent to the sink from all the sensor nodes and the information is created at the sink. Example, to construct the information about the perimeter of the region of deployed WSN.

Information Attributes

In order to assess the QoI, we first need to understand the information attributes. This section surveys existing and proposes some new attributes of information. There exist many attributes for information, but we choose only those which are relevant and useful for QoI in WSN. There also exists information model in defining information attributes [51], which benefits to define the existing attributes. To plan an application and use it in an operational perspective, one needs to give more importance on various attributes concerning QoI. We define some of the existing attributes below based on QoI and building blocks shown in Figure 1.

Accuracy: is the degree of correctness which provides the level of detail in the deployed network. It is the value which is the close imitation of the real world value.

Precision: is the degree of reproducibility of measured values which may or may not be close to real world value.

Completeness: is the characteristic of information which provides all required facts for user during the construction of information at the sink such as the perimeter of the event region.

Timeliness: is an indicator for the time needed when the first data sample generated in the network till the information reaches the sink for decision making.

Throughput: is the maximum information rate at which information is provided to the user after raw data collection.

Reliability: is the characteristic of information, in which information is free from change or no variation of information from all the blocks of the source to the sink.

Usability: is the ease use of information that is available after raw data collection undergone in-network processing and can be applied to the application based on user's evolvable requirements.

Certainty (Uncertainty): is the characteristic of information from the source to the sink with desired level of confidence helping the user for decision making. *Uncertainty* is the condition of information which leads to risk minimization (information free from doubt).

According to our knowledge from [15] [39] [1] [40] [50] there are still some missing attributes in WSNs for QoI, these attributes play a vital role and are useful in WSNs. The following attributes are similarly interwoven to the existing ones in the literature and also used in other fields like database management, machine learning and management studies. The following defined attributes are applicable to WSNs and also required, because of their sensible aspect in information processing.

Tunability: is the characteristic of information, where the information can be modified and undergo in-network processing based on user's evolvable requirements. Information is tunable, if the user requirements are changing to collect raw data or information at sink needs to be tailored. User can take the advantage of feedback channel in order to tune the information.

Affordability: is the characteristic of information to know the cost for measuring, collecting and transporting the data/information. It is the expensiveness of information. Affordability can be of raw data, and how cost effective it is to measure raw data. Affordability can be characterized to all the blocks based on user's requirements.

Reusability: is the characteristic of information, where the information is reusable during its lifetime or as long as it is relevant (in the time domain) for future use in the context of WSN. In [50], the timeliness attribute gives the similar meaning to reusability. However, the term timeliness in [50] is mainly with information manufacturing systems. In WSN due to resource constraints and user requirements evolvability, timeliness and reusability gives separate meaning.

*Generalized QoI Definition and Relation to other Qo**

In current trends QoI is being addressed in terms of network design, deployment and applications but all in early stages [23]. There exist some definitions of QoI, but still QoI definition is an inconclusive endeavor. QoI defined by S. Zahedi et al. [47] is more application oriented view as representing “the level of confidence that a sensor-data-dependent application may place on information derived from the WSNs that (may) support the information needs of the application". QoI in regard to monitoring environment [12] is "the difference between the data that the output of the WSNs produces concerning some environment that is being monitored, and the actual events in that environment which one wishes to observe or track". The multidimensional concept [48] of QoI on application view is "the collective effect of the available knowledge regarding sensor derived information that determines the degree of accuracy and confidence by which those aspects of the real world (that are of interest to the user of the information) can be represented by this information." The design process of an application is unaware of QoI delivery function. Accordingly, QoI in a tactical WSN is defined [41] in terms of "the extent to which the WSN fulfill the requirements of command and in its contribution to the picture of ground truth in operation". To deal with a target mission, there are always requirements of quality in context that can meet the specific information requirement. Some concepts of information processing also define "QoI as a characterization of the goodness of the data captured by and following through the WSNs and the information derived from processing those data along the way" [3]. It could also be argued that QoI provides a foundation for the design, deployment and operation of WSNs. The user defines QoI in his/her own way depending on the mission targeted.

In this paper, we define QoI in regard to information attributes, user evolvable requirements, and resource efficiency and application dependency. There is always scope and need for adaptable network characteristics with more confidence that it works with similar consistency as throughout the network. We therefore propose the following definition:

"Quality of Information (QoI) is the quality experienced/perceived by the user concerning the received information, while considering the evolvable user requirements, the overhead of information acquisition and operational quality perturbations".

This definition encompasses the various quality characteristics/attributes, accounts for the cost of acquiring the information, and considers the perturbation factors of the achievable and desirable quality.

QoI Relation to other Qo*: There exists other Qo* which are intending to be used in different isolated blocks to satisfy user's evolvable requirements and targeted application. These are the qualities which intend the flow from the time of detection of event till decision making [36] [28]. From Figure 1, the quality we present here refers to different techniques used at different blocks. We generalize that other qualities like detection, routing, and decision making to be part of QoI. Quality of Detection (QoD) [15] is of main interest in the aspect of event detection and is tightly related to the quality of the detected event. The main aspect of detection in WSNs is any event, and it's also related with user application. The application targeted by the user also needs some network services during information transport. This is mainly known as the Quality of Service (QoS) [2] provisioning in WSN. The QoS management refers to the systematic approaches to measuring and managing the delay, jitter, and/or throughput of information/data transport in WSN, usually, investigated in terms of bandwidth allocation, prioritization etc. Quality of Routing (QoR) is a fundamental task in WSNs [8] [33] [16] which is tightly related to QoS as it focus only on routing reliability and timeliness. QoS and QoR are handled in the third building block of information transport (Figure 1). For decision maker QoI is one of the vital factors as Quality of Decision making (QoDm) [1] [31] is mainly based on QoI.

Information Quality Assessment

Usually, the quality of delivered/achieved information should be assessed according to the required/expected quality. For a quantitative assessment, metrics play a major role. In the following, we briefly discuss the user requirements on QoI as well as the QoI metrics.

User Requirements*:* As mentioned before, QoI is complied with a set of attributes. These attributes are measured by relevant metrics to give the level of detail of QoI. Hence, we consider that user requirements are information based on some set of attributes. The user requirements can be further regarded as measured information based on a specific set of attributes. User is not necessarily a human and can be application planner, end user, decision maker, consumer, intelligent system etc. The use of feedback channel is important here for user requirements dissemination.

QoI Metrics: Measuring the information quality is quantifying the information attributes. Metrics are valuable at both design and deployment time as the user requirements are evolvable and the user would benefit from knowing the level of QoI of received information entities for safer decision making. Measuring an information attribute is either completed in-network or/and at the sink. A metric is a standard of measurement stated in quantitative term which captures the performance in relative to standard on the occurrence of event. The quality of a system, such as its energy-efficiency, information attributes such as accuracy, timeliness etc. and the evaluation criterion of these qualities are judged by the term *metric*. Measure can be classified as "happening" and "valuing". For example of fire detection in the forest, there is fire is the true state of event happening, there is fire with 95% accuracy is the valuing of the event. If the metric is well defined it has to lead to actionable performance to satisfy the deployed system and also needs a capable system model to measure it. This doesn't mean to have high rate of data collection or reliable protocol or having non-effective metrics satisfy the user evolvable requirements. Hence, we can define that a metric is acceptable with certain performance measure only if it has some opening limit, meaning it is a limit which is likely near/above to threshold value or real world value. There are quite a few metrics defined in the literature such as probability of error [49], Peak Signal Noise Ratio (PSNR) [12], recall and false-positive rate [13], path weakness [33], transient information level [35] etc., but not for all the attributes mentioned above such as tunability and affordability. However, these metrics are not the sole to measure other attributes such as accuracy, precision and reliability.

QoI: Provisioning Across the WSN Functional Blocks

Currently, there exists justifiable work about varying aspects of WSNs activities such as data collection/sampling [4] aggregation, resource allocation of the data to further operation etc. Currently, information (quality) is addressed in different roof, i.e., regarding fusion [20], data impact, decision making [31], degradation, miss association in fused information and data level acquisition [34] [26] [9]. When QoI is viewed in these aspects there are different dimensions and we classify these aspects in four different views of raw data collection, in-network processing, and information transport and sink operations. This classification shows that information should be considered as one important aspect throughout all operations.

We consider WSNs in a holistic view and consider isolated blocks spanning raw data collection, in-network processing (compression, aggregation etc.), information transport and sink operations as whole. We consider these building blocks because we expect that QoI should satisfy user requirements from the time of raw data collection till operations at sink are conducted.

Raw Data Collection and Dissemination: is the process of collecting raw data samples at sensor nodes (sampling in time and space domain) based on user evolvable requirements [27] [52], this operation is very important for monitoring, localization [17] [22] among others. Temporal sampling quality depends on the sensor quality and the sampling frequency. Generally, an adaptive temporal sampling is required for varying user requirements. Sensor node failures and duty cycling affect the accuracy of spatial sampling and consequently the QoI level. After local temporal sampling of the physical phenomena on node level, sensor nodes usually share their samples in order to detect events through in-network processing. This local dissemination/sharing are usually limited to neighboring nodes.

In-Network Processing (Information Extraction): in WSNs data collection is done by collecting raw data samples, and this raw data is processed in order to extract useful information (Figure 2). According to the established communication-computation trade-off principle in WSN, the extraction of useful information from raw data is usually achieved through in-network processing [44]. Aggregation is one of the most popular in-network processing techniques [24] [45] [25]. Another technique is to use dynamic Bayesian network [10] for sensor selections, which improves the information quality. Further common in-network processing techniques are filtering, compression, suppression etc.

Information Transport: is the end-to-end transport/routing of the information from the source (where it is generated/extracted) to the sink [21] [53] Most of the current existing routing protocols and their techniques assume that the information coming from the source is trustworthy and reliable [30]. However, still it's not sure how much % of confidence can be placed on this information.
Sink Operations: information is managed at the sink for further operations such as decision making. Information Risk Management (IRM) [32] is one of the approaches to minimize the risks such as information misunderstanding and insufficiencies of metric which may affect learning quality. The most important fact is how and why the user needs to place confidence on the information through all the process of data collection, data processing, information transport and information management at the sink.

SURVEY OF EXISTING APPROACHES ADDRESSING QOI IN WSN

In the literature, QoI is being addressed by frameworks [47] [19], methods [34], models [29] and decision making techniques [11]. This section presents an overview of the existing approaches in QoI. In this section, we first classify the existing approaches. Next, we briefly describe them. Then, we qualitatively compare them. Table II presents an overview of these approaches based on application criteria, i.e., user, application, information class and the system model used. In Table III, we define different classes of building blocks. Table IV compares the same approaches w.r.t. the covered building blocks, information attributes and QoI metrics.

Classification According to Application Domain

In this section, we classify the existing approaches that address QoI in WSN. Our classification criteria's are the considered information class, user, system and application models.
We follow a step-by-step process to classify the existing approaches. First, we identify the source of information is from sensor nodes and also from the user/application. As we have already identified the functional building blocks as raw data collection, in-network processing, information transport and the sink operations, we take this step for granted. On the other hand, we have also identified the information classes, we also take this step to be done.

Approaches	**User**	**Application**	**Information**
QoI Analysis for Detection Oriented Systems [47]	Application Planner	Detection system	C1[R], C2[E], C3[S]
QoI in DTN [35]	End User	Resource allocation	C1[R], C2[E, Q], C3[S]
Context Aware QoI Computation [18]	End User	Context-aware computing	C1[R], C2[Q], C3[S]
Characterization of Information Quality [11]	Decision Maker	Decision making	C1[R], C2[Q], C3[S]
QoI Rate Control [49]	End user	Event detection	C1[R], C2[E, Q], C3[S]
Information Fusion Analytical Framework [34]	End User	Decision making	C1[H], C2[E], C3[N]
Data Model Framework [19]	End User	Data modeling	C1[P, H], C2[Q,E], C3[S]
Resource Management [42]	End User	Monitoring	C1[P, H], C2[Q], C3[S]
Data Driven Sensor Reporting [14]	Application Planner	Event detection	C1[R], C2[E, Q], C3[S]
Information Quality Management [43]	Health Systems	Health care	C1[R], C2[E, Q], C3[S]
QoI Inspired Analysis for Sensor Network Deployment [48]	Application Planner	Deployment planning	C1[R, H], C2[E], C3[S]
QoI an Empirical Approach[12]	End User	Context-aware computing	C1[R], C2[Q, E], C3[S]
Information Risk Minimization [32]	End User (Patients)	Health care system	C1[P], C2[E], C3[S]
Sensor Sampling and QoI [4]	Decision	Detection system	C1[R], C2[E], C3[S]

	Maker		
Detection Performance [15]	End User	Detection system	C1[R], C2[E], C3[S]
Data Cleaning [13]	End User	Data quality analysis	C1[H], C2[E], C3[S]
Quality of Routing [33]	End User	Sensor centric modeling	C1[P, R, H], C2[Q], C3[S]
Information Awareness [1]	Decision Maker	Control system design	C1[R, H], C2[E, Q], C3[S]
Dynamic Target Tracking [40]	Military	Tracking	C1[R, H], C2[E, Q], C3[S]
Letter Soup for QoI [5]	End User	Operational context	C1[P,R, H], C2[E, Q], C3[I, S]

Table II: Mapping of QoI Approaches based to the Information Classes and System Models

Discussion and Comparison of Existing QoI Approaches

Now, we compare the existing approaches based on the building blocks, the information attributes and QoI metrics. In Table III, we define different classes of building blocks. Usually, the approaches focus on a few selected attributes. We have gathered most of the state-the-art related to QoI approaches in the following subsections. Usually existing approaches just focus on a single block and only a few consider more than one block.

Building Blocks Classifications	Types of Building Blocks
Class 1: Single Building Blocks [C1S]	Data Collection [D] In-network processing [I] Information transport [T] Sink operations[S]
Class 2: Combinational Building Blocks [C2C]	Data collection and In-network processing [D, I] Data collection and Sink operations [D, S] In-network processing and Information transport [I, T] In-network processing and Sink operations [I, S]
Class 3: All Building Blocks [C3A]	Data collection, In-network processing, Information transport and Sink operations [D, I, T, S]

Table III: Building Blocks Classifications

Qualitative Comparison of Existing Approaches based on Functional Building Blocks

The approaches classified in this subsection are based on building blocks.

Approaches Addressing a Single Building Block: Most of the existing QoI approaches just concentrate on a selected building block like data collection or information transport or sink operations etc. However, they still lack to identify the effects of neglecting other blocks.

[C1S] [D]: Though there are many approaches concentrating just on data collection block, we just brief two [42] [14] here as to highlight the issue on neglecting other blocks. The aims of these approaches are resource allocation and sensor selection. The HYBRID [14]approach is more prominent as the variance of application requests or data change across motes increases, the model is based on push and pull method, the model dynamically switches between push and pull techniques based on system condition. However, the HYBRID model is setup for aspect of data collection, it neglects other blocks leading to negotiation with QoI attributes. The challenge of [42] lies in considering phenomena state distribution while making application admission decision. The framework acts as an admission control scheme to decide if the WSN is able to provide the required service. Though the vital fact in this approach is data collection, the user can not be sure of the acquired data till it reaches the sink.

[C1S] [T]: In the block of information transport the attributes reliability and timeliness are highly regarded. Though most of the routing protocols always assume the data/information coming from sensor nodes are accurate enough, but can't place a certain level of confidence on this data/information. Hence, in this regard, the attributes such as accuracy, precision are ignored. This issue is identified in the approach [35] relevant to QoI which is focusing on information transport. Disregarding the fusion process and not focusing on the sensor fusion aspects assuming that those processes have been completed, the framework as in [35] handles the QoI assigned message in the network. However, in [35] the mere aspects of in-network processing and sink operations and attributes related to this blocks are violated.

[C1S] [S]: Information plays different roles and has different values for decision makers at different levels. For characterizing the information quality spectrum the techniques like fuzzy values [11] are used. Here the approach is just concentrating on sink operations. Authors discuss the issues of uncertain data, imprecision. The main aim is in determining acceptance regions, similarity functions to determine the similarities between components and the confidence measures to rate attributes. Therefore, exploiting the tolerance for imprecision and uncertainty when precise information carries a cost or unavailable in the decision making process.

Approaches Addressing Combinational Building Blocks: Some of the approaches combine at least two building blocks like data collection with in-network processing or in-network processing with sink operations or data collection with sink operations. However, still they lack to identify the after effects of neglecting other blocks. We have identified such approaches here with combination of building blocks.

[C2C] [D, I]: Some approaches adopting the combination of data collection and in-network processing are presented here. The approaches [47] [48] [49] [43] tend to overlook the effects of information transport and sink operations. As from our information attributes Section, information transport and sink operations hold some specific attributes specific to own block. Hence, in this case when the user is not considering these blocks or not respecting the attributes pertaining to these blocks there is violation of QoI or the evolvable user requirements are not satisfied.

Now, we briefly discuss these approaches and their challenges. The current state of the art on layered framework for decomposing the deployment evaluation is done in three steps of input processing, core analysis and result post-processing [47]. The given framework facilitates the decoupling of the three steps, the mix and match analysis and modeling approaches. It serves as a computational aid for a sensor system designer to evaluate the performance of users design based on deployment and QoI constraints provided by the application planner.

The results in [49] demonstrate the benefit of using prior information of event location on the probability of error. In this case the data collection phase should be accurate and also it should be relevant to the evolvable user requirements for decision making at the sink. However, here the approach neglects the aspect of sink operations. The approach is very similar to the content centric networking, which endows the networking stack with knowledge of the intent of the communication transaction. On the same basis in [49], new greedy rate control algorithm selects rates based on each node’s contribution to the QoI, but the drawback is that the proposed greedy rate control algorithm is unable to handle errors in wireless links.

One important factor in WSNs is the process of sensor deployment and sensor selection. The later process of sensor selection using Bayesian model is not appropriate for sensor selection as there is no notion of time in the Bayesian network and many sensor reading has to be taken to provide desired QoI. However, in [43] this is achieved by using a dynamic Bayesian network model that provides the information quality to WSNs. The dynamic Bayesian network models optimize with one application and use very little resources in order to not to address the aspects of losses of data in the network. The setback of these kinds of models is assuming that every model is actually a complex, able of doing online data processing, which is not always true. However, what happens to data after sensor selection or to achieve information quality when it reaches the sink is not discussed.

In WSNs data models can help to combine readings from different sensor nodes to assess the QoI or to minimize energy consumption and thus maximize the lifetime of the WSNs while still respecting QoI attributes. In [19] the model is based on data collection and in-network processing. Here the framework allows several data models to run in parallel. The framework runs in offline mode, but for on-line the authors still propose future work and still lack to explain the QoI attribute factors within the framework. In this approach authors neglect the aspect of information transport and sink operations and lack to provide required QoI. The cons in this approach are, the work in [19] takes into consideration data

collection and in-network processing, and neglects information transport and sink operations. However, this leads to non-confident information reaching the sink.
[C2C] [I, T]: In [37] the authors showed preliminary results on the benefit of considering the type and the reliability attribute of the information in order to enhance the information transport. The work proposed to manage redundant, atomic and composite information for efficient information transport with in-network processing. However, the fact of raw data collection and sink operations are neglected.
[C2C] [I, S]: In-network processing is the next block after data collection (Figure 1). In Table IV, the information fusion [34] approaches are based on blocks of in-network processing and sink operations. Identifying good candidates for information fusion is presented in [34] combining with sink operations. This analytic framework is to study information fusion competition between the negative effect of disassociations and the positive effect of synthesis, to demonstrate and analyze their interplay quantity. The generic model used here is to demonstrate the varying degrees in fusion, namely increased QoI versus decreased QoI. In [34] the approach concentrates on in-network processing and sink operation is not sure about what are the data collected and how the information is transported. Though one can assume a good underlying routing protocol, but the facts of violating attributes with accurately collected data and saving resources makes the approach still primitive.

Qualitative Comparison of Existing Approaches based on QoI Attributes and Metrics
This sub-section is the classification of QoI approaches based on attributes and metrics, which are used to characterize and quantify QoI. We always argue that to have achievable QoI pertaining to user evolvable requirements, user needs to respect the characteristics of QoI. Moreover, we present some of the approaches concentrating only on some attributes and measuring them.
The principle based framework [47] is a strategy of principles and steps to achieve ideology of deployment planning, decision making, and quality enhancement. The current state of art on layered framework for decomposing the deployment evaluation is done in three steps of input processing, core analysis and result post-processing. The framework uses probability of error to measure detection probability and false alarm rate. The main aspect of detection in WSNs is any event, in [15] detection performance is measured with average sampling rate with characteristics such as accuracy and timeliness.
The QoI aware route control in [49] uses probability of error as metric to measure accuracy. It explicitly optimizes application relevant QoI metrics during network resource allocation decision. The QoI approach presented in [12] focuses on accuracy and measure it with Peak Signal Noise Ratio (PSNR). However, though the QoI here is measured, other attributes such as timeliness for timely arrival of information for decision making have been not discussed.
Exploiting the tolerance for characterization of information quality using fuzzy logic [11] some attributes such as accuracy, completeness, relevance, timeliness and usability are explored. However, though the work considers some attributes relevant for QoI, never quantifies it. In QoI with characterization of information, Information Risk Management (IRM) [32] is also proposed in the literature to minimize the risks such as information misunderstanding and insufficiencies of metric which may affect learning quality. Dimension extension (DIME) is a framework to accommodate local and prior knowledge into learning coarse by measuring accuracy through dot product as metric.
To achieve better results, data processing is used in current trends of QoI. Usually, in resource constrained framework a real good data processing is a key precondition for analysis decision and data integration. One of the frameworks [13] addressing this is based on rule base, scheduling and log-management. The attributes such as consistency, accuracy, extensibility and interactivity are used for data cleaning and measured by metrics such as recall and false-positive rate. The overall design fully shows the features of extensibility and interactivity, meaning the framework allows users to add rules, and at the same time allows user to form strategies in the needs of different data cleaning.
The concept of operational context to ease the dynamic binding of sensor resources to applications represents QoI needs of an application and the capabilities of the sensor resources by the 5WH (why,

where, when, what, who, how) principle [5]. With the interpretation of the 5WH primitives provided, spatial and temporal relevance is used as a metric to measure data completeness.

The evolution of the context may be used to adjust dynamically the weights of the sensor nodes that ease selecting the right set of sensor nodes given the dynamic context change as the one in [18]. Some attributes such as certainty, accuracy/confidence and timeliness are used for context aware QoI computation. Still here the information is not measured. Relative to this the selection of sensors can be made by using metrics such as information gain and using other attributes missing in [18].

By targeting all the building blocks and attributes related to each block, we now brief a strategy [33] that develops a game-theoretic metric called path weakness to measure the qualitative performance of different routing mechanisms. The approach uses qualitative performance as a QoI characteristic and uses sensor-centric concept.

Considering information transport, prioritizing traffic has been studied for a long time. Disregarding the fusion process and not focusing on the sensor fusion aspects, if those processes have been completed, the framework as in [35] handles the QoI assigned message in the network. Based on this the key metric transient information level is defined, which is the product of information and projected physical distance of that information from destination node. This approach is very relevant to QoI information transport block as attribute related to information transport such as timeliness of information are used. The QoI level is also measured, but the approach neglects the effects of other building blocks and some attributes.

Approaches	Building Blocks	Attributes	Metrics
QoI Analysis [47]	[C2C], [D,I]	Detection probability and false positive rate	Probability of error
QoI in DTN [35]	[C1S], [T]	Timeliness, integrity and consistency	Transient information level
QoI Computation [18]	[C2C], [D,S]	Certainty, accuracy, timeliness, integrity	___
Characterization of Information Quality [11]	[C1S], [S]	Accuracy, completeness, relevance, timeliness, usability	___
QoI Rate Control for Sensor Networks [49]	[C2C], [D,I]	Accuracy	Probability of error
Information Fusion [34]	[C2C], [I,S]	Uncertainty	___
Data Model Framework [19]	[C2C], [D,I]	Accuracy, reliability	___
Resource Management [42]	[C1S], [D]	Completeness, uncertainty, accuracy	___
Data Driven Sensor Reporting [14]	[C1S], [D]	___	___
QoI Management [43]	[C2C], [D,I]	Certainty, completeness, accuracy	Entropy
QoI Inspired Analysis for WSN Deployment [48]	[C2C], [D,I]	Accuracy	Probability of detection and false positive rate
QoI an Empirical Approach [12]	[C1S], [D, I, S]	Accuracy	Peak Signal to Noise Ratio
Information Risk Minimization [32]	[C1S], [I]	Accuracy	Dot product
Sensor Sampling and QoI [4]	[C2C], [D,I]	Accuracy, timeliness and confidence	___
Detection Performance [15]	[C2C], [D,I]	Accuracy, timeliness and robustness	Average sampling rate
Data Cleaning [13]	[C2C], [I, S]	Consistency, accuracy, extensibility, interactivity	Recall and False-Positive Error
Quality of Routing [33]	[C1S], [T]	Qualitative performance	Path weakness
Information Awareness [1]	[C2C], [I, S]	Precision, quality and usability	___
Dynamic Target Tracking [40]	[C1S], [S]	Accuracy, timeliness	Entropy, information gain, residual likelihood
Letter Soup for QoI [5]	[C1S], [D, I, S]	Accuracy, completeness, timeliness	Spatial and temporal relevancy

Table IV: Classification of QoI Approaches Based on Building blocks, attributes and metrics

RESEARCH ROADMAP

This survey has pointed some drawbacks of the state-of-the-art of QoI research, which represent emerging research directions that we briefly discuss in this section.

As discussed before, most of existing approaches address one or more different blocks. Unfortunately, the approaches neglect also the importance of the one or the other block as none of them considers all blocks holistically. They indeed assume the other blocks to be perfect in maintaining the highest possible QoI contribution. In addition, some of the approaches just concentrate on a few information attributes such as accuracy and ignore other attributes. Consequently, our first research direction is to identify the cross fertilization when considering all blocks together. We have defined our holistic view in this paper, one can try to take one or more blocks and later try to cross fertilize them together.

The existing research based on these blocks always try to enhance the information leading to negotiate with quality attributes like accuracy and use more resources leading to energy depletion or activities affecting the deployed network [12]. It is not efficient to use the "best" solutions [47] [48] (processing techniques, protocols etc.) from the source to the sink to deliver information with the required QoI. It is not always necessary to increase the quality, but sometimes to decrease it to save valuable resources such as energy and bandwidth, and increase timeliness of information delivery without under-performing the required quality indicators/metrics such as accuracy. To design and deploy a WSN, one should consider the holistic view to achieve the required QoI level while maximizing efficiency. Accordingly, we propose to review and improve the tuning capabilities of single blocks. It is important to use the concepts presented in this survey paper and understand how one can enhance/tune each block to achieve QoI satisfying user evolvable requirements. The attribute tunability can be considered to extend the work in [37] [38] by figuring out how to reliably transport information in the aspect of timeliness as the main concern. Tuning the timeliness of the information which reaches the sink with complex system model consisting of mobile nodes can be a stride forward contribution.

Though information attributes are relatively well discussed, QoI metrics definition and their efficient computation are still in their infancy. Accordingly, the future research directions may progress on the aspects of defining metrics and the techniques to efficiently compute them on the fly in all information extraction stages. However, as one need to narrow research into fewer attributes, one will take some must considered attributes during the flow of information from the source to the sink. One can define and defend how it is relevant and necessary to use these attributes and violation of this lead to QoI which does not satisfy the user requirements. Akaike's information criterion, is a measure of the goodness of fit of an estimated statistical model grounded on the concept of entropy, in effect offering a relative measure of the information lost when a given model is used to describe reality. As an example, Akaike's information criterion can be used to measure QoI when certain information is lost from the source to the sink. Metrics and their run-time quantification represent a powerful tool to assess the dependability of WSN, which allows for efficient and tunable QoI provisioning.

CONCLUSIONS

Users are mainly interested in information from the WSN. This highlights the importance of understanding the quality of data collection, in-network processing, information transport and decision making, or shortly QoI. This paper has focused on the factors of QoI. We introduced new QoI attributes namely tunability, affordability and reusability. In the literature there exist user defined QoI definitions, which are unfortunately specific for certain application domains. Therefore, we proposed a new generic QoI definition that considers the aspect of application, attributes and saving resources such as energy. Consequently, we provided a comprehensive survey/review of most of the existing work related to QoI. Our study was performed on the QoI attributes and its application perspective. QoI across all WSN functional operations from the raw data generation to the information extraction and delivery to the user was the base of our study. We have classified information and building blocks into different classes.

Based on these classes we have mapped the existing QoI frameworks, methods and models, approaches and highlighted the pros and cons. Moreover, our classification technique is more general to be adapted to most of the applications. On the other hand, we also mapped the existing work with QoI related attributes and metrics.

Nevertheless, this survey presents an overview on information based on building blocks, attributes, and metrics and information classes. However, we have identified the QoI related problems, i.e., to focus on all or combinational building blocks and satisfy user requirements. We have also mentioned that violating of least required attributes may affect QoI. We built our holistic view on these factors and proposed that information should satisfy user evolvable requirements saving resources.

REFERENCES

[1] Arnborg S., Artman H., Wallenius K. *Information awareness in command and control: precision, quality, utility*. In Proc. of FUSION, 2000.

[2] Basaran C., Kang K. D. *Quality of service in wireless sensor networks.* In Book Series Computer Communications and Networks, 2009.

[3] Bisdikian C., Branch J., Leung K.K., Young R.I. *A Letter Soup for the Quality of Information in Sensor Networks*. In Proc. of 1st Int. Workshop on IQ2S (part of IEEE PerCom), 2009.

[4] Bisdikian C. *On sensor sampling and quality of information: A starting point*. In PerCom, 2007.

[5] Bisdikian C., Branch J., Leung K. K, Young R. I. *A letter soup for the quality of information in sensor networks.* In Proc. of PerCom, 2009.

[6] Bisdikian C., Kaplan L. M, Srivastava M. B, Thornley D. J, Verma D, Young R. I. *Building principles for a quality of information specification for sensor-originated information.* In Proc. of ITA, 2009.

[7] Chi H. L, Leung K. K., Branch J. W., Bisdikian C. *A QoI-aware middleware for task-oriented applications in wireless sensor networks.* In Proc. of ITA, 2009.

[8] Costas B., Kannan R., Vasilakos A. V. *Quality of routing congestion games in wireless sensor networks.* In Proc. of WICON, 2008.

[9] Chong C. Y. *Sensor networks: evolution, opportunities, and challenges.* In IEEE proceedings, 2003.

[10] Duncan G., Thornley D., Bisdikian C. *Probabilistic approaches to estimating the quality of information in military sensor networks.* In Proc. of the computer Journal, 2009.

[11] Ehikioya, S. *A Characterization of Information Quality Using Fuzzy Logic*. In Proc. NAFIPS 1999.

[12] Gelenbe E., Hey L. *Quality of information: An empirical approach*. In Proc. of MASS, 2008.

[13] Hao Y., Diao X. C., Li K. Q. *Research on information quality driven data cleaning framework*. In Proc. of FITME, 2008

[14] Hakkarinen D., Han Q. *Data quality driven sensor reporting.* In Proc. of MASS 2008.

[15] He T., Zafer M. *Adaptive Sampling for Transient Signal Detection in the Presence of Missing Samples.* In Proc. of. 1st Int.QoISN, (Part of IEEE MASS), 2008.

[16] Heinzelman W. B., Murphy A. L., Carvalho H. S., Perillo M. A. *Middleware to support sensor network applications.* In Proc. of IEEE Network, 2004.

[17] Hossain M. A., Atrey P. K., Saddik A. E. *Modeling quality of information in multi-sensor surveillance systems*. In Proc. of Data Engineering Workshop, 2007.

[18] Hossain M. A., Atrey P. K., Saddik A. E. *Context-aware QoI computation in multi-sensor systems*. In Proc. of MASS, 2008.

[19] Hunkeler U., Scotton P. *A quality-of-information-aware framework for data models in wireless sensor networks*. In Proc. of MASS, 2008.

[20] Huijia L., Lu M., Nikola M., Gao J. *Composable information gradients in wireless sensor networks*. In Proc. of IPSN, 2008.

[21] Jamal N., Kamal A. E. *Routing techniques in wireless sensor networks: A survey.* In journal IEEE Wireless Communications, 2004.

[22] Kaplan L., Damarla T., Pham T. *QoI for passive acoustic gunfire localization*. In Proc. of MASS, 2008.

[23] Khelil A., Reinl C., Ayari B., Shaikh F.K., Szczytowski P., Ali A., Suri N. *Sensor Cooperation for a Sustainable Quality of Information*", Book chapter in "Pervasive Computing and Networking", (To appear 2010).

[24] Krishnamachari B., Estrin D., Wicker S. *The impact of data aggregation in wireless sensor networks*. In Proc. of Inter. Conference on Distributed Computing Systems, 2002.
[25] Krishnamachari B., Estrin D., Wicker S. *The impact of data aggregation in wireless sensor networks*. In Proc. of the 22nd International Conference on Distributed Computing systems, 2002.
[26] Markus W., Braun T. *Efficient signal processing and anomaly detection in wireless sensor networks*. In Proc. of EvoCOMNET, 2009.
[27] Meliou, A., Chu D., Guestrin C., Hellerstein J. Hong, W. *Data gathering tours in sensor networks*. In Proc. of IPSN, 2006.
[28] Mo L,, Yang B. *A Survey on Topology issues in Wireless Sensor Networks*. In Proc. of ICWN, Nevada, 2006.
[29] Moody D. L. *Evaluating the quality of information models: empirical testing of a conceptual model quality framework.* In Proc. of Software Engineering, 2003.
[30] Omprakash G,, Fonseca R,, Jamieson K. *Collection tree protocol.* In Proc. of SENSYS, 2009.
[31] Prasanth R., Cabera J., Mehra A. R., Purtell R., Smith R. *Quality of information measures for autonomous decision-making.* In Proc. of American Control Conference, 2004.
[32] Qun C., Wang X., Wang X-L. *Local information embedding by dimension extension and information risk minimization.* In Proc. of Machine Learning and Cybernetics, 2005.
[33] Rajgopal K., Sarangi S., Iyengar S. S., Ray L. *Sensor-centric quality of routing in sensor networks.* In Proc. of IEEE, INFOCOM, 2003.
[34] Ronald K. *The dynamics of information fusion: synthesis versus misassociation*. In Proc. of FUSION, 2006.
[35] Robert A. N. *QoI in DTN-based directional networks.* In Proc. of SARNOFF, 2009.
[36] Sandip B., Arora A. *Stabilizing reconfiguration in wireless sensor networks*. In Proc. of SUTC, 2006.
[37] Shaikh F.K., Khelil A., Suri N. *Generic information transport for wireless sensor networks.* In Proc. of SUTC, 2010, pp: 27-34.
[38] Shaikh F.K., Khelil A., Suri N. *AReIT: adaptive reliable information transport protocol for wireless sensor networks*. In Proc. of ICWN, 2009.
[39] Srinivasan J. *The role of trustworthiness in information service usage: The case of parry information kiosks.* In Proc. of ICTD, 2007.
[40] Tan C. H et. al. *Generating reliable quality of information metrics for target tracking*. In Proc. of SPIE, 2009.
[41] Thornley D.J.,Young R., Richardson J. *From Mission Specification to Quality of Information Measures Closing the Loop in Military Sensor Networks.* In Proc. of ITA 2008.
[42] Tolstikov A., Tham C K., Biswas J. *Quality of information assurance using phenomena-aware resource management in sensor networks* . In Proc. of 2nd Workshop on COQODS-II (ICON2006).
[43] Tolstikov A., Xiao W., Biswas J., Zhang S., Tham C K. *Information quality management in sensor networks based on the dynamic bayesian network model*. In Proc. of ISSNIP, 2007.
[44] Xingbo Y., Mehrotra S., Venkata S N. *SURCH:distributed aggregation over wireless sensor networks*. In Proc. of IDEAS, 2006.
[45] Yonggang R. G., Zhao J., D. Estrin. *Computing aggregates for monitoring wireless sensor networks*. In Proc. of The First IEEE International Workshop on SNPA , 2003.
[46] Yu B., Sycara K. *Learning the quality of sensor data in distributed decision fusion*. In Proc. of the Ninth International Conference on Information Fusion, 2006.
[47] Zahedi S., Srivastava M. B., Bisdikian C. *A computational framework for quality of information analysis for detection-oriented sensor networks*. In Proc. of MILCOM, 2008.
[48] Zahedi S., Bisdikian C. *A framework for QoI-inspired analysis for sensor network deployment planning*. In Proc. of 3rd international conference on Wireless internet, 2007.
[49] Zainul C., Zahedi S., Cho Y. H. *Toward quality of information aware rate control for sensor networks*. In 4th Inter. Workshop on Feedback Control Implementation and Design in Computing Systems and Networks, 2009.
[50] Ballou B., Wang R., Pazer H., Tayi G K. *Modeling Information Manufacturing Systems to Determine Information Product Quality*. In Proc. of Management Science, 1998.
[51] Wang R., Strong D. *Beyond Accuracy: What Data Quality Means to Data Consumers*. In Journal of Management Information Systems, 1996.
[52] Szczytowski P., Khelil A., Suri N. *ASample: Adaptive Spatial Sampling in Wireless Sensor Networks*. In Proc. of SUTC, 2010.
[53] Shaikh F.K., Khelil A., Suri N. *On modeling the reliability of data transport in wireless sensor networks*. In Proc. of PDP, 2007.
[54] Ali A., Khelil A., Suri N. *MPM: Map based Predictive Monitoring for Wireless Sensor Networks*. In Proc. of Autonomics 2009.