Information Quality Aware Co-design of Sampling and Transport in Wireless Sensor Networks

Vinay Sachidananda, Abdelmajid Khelil, David Noack and Neeraj Suri
TU Darmstadt, Hochschulstr. 10, 64289 Darmstadt, Germany
{vinay,khelil,noack,suri}@informatik.tu-darmstadt.de

***Abstract*—A key task in Wireless Sensor Networks (WSNs) is to deliver specific information about a spatial phenomenon of interest. To this end, a few Sensor Nodes (SNs) sample the phenomenon and transmit the acquired samples, typically multihop, to the application through a gateway called sink. Many applications require the spatial sampling to be accurate and the delivery to be reliable. However, providing a higher accuracy/reliability comes at the cost of higher energy overhead as additional messages are required: increasing the number of samples to increase the accuracy of sampling and increasing the number of retransmissions to increase the transport reliability. Existing design approaches overlook optimized spatial sampling accuracy and transport reliability in combination for minimizing energy consumption. This work aims in providing the optimized solution for sampling accuracy and transport reliability in composition for a maximized efficiency. Our approach features a message efficiency that optimally meets application requirements with the online adaptation and appropriate tradeoff between accuracy and reliability. The sampling and transport co-design proceeds by finding optimal number of SNs for the accuracy of the spatial sampling with the effect of reducing the number of retransmissions and still satisfying the application requirements. We validate the approach viability through analytical modeling and extensive simulations for a wide range of requirements.**

***Keywords*-Wireless Sensor Networks, Sampling, Transport, Optimization, Information Quality**

I. Introduction

In Wireless Sensor Networks (WSNs) delivering the gathered information with the application required quality is the main concern. To satisfy the required quality, it is crucial to carefully design the core functional blocks, such as (a) the sampling scheme in order to accurately represent the physical phenomena, and (b) the transport scheme in order to reliably deliver the information to the sink. In our work, we focus on the key operations of spatial sampling and transport along their quality attributes, i.e., accuracy and reliability respectively.

The user/application view considering the spatial phenomena of interest requires a certain *sensing task* (e.g., perimeter of the phenomenon area [17] on the spatial distribution of the phenomena [9]). Moreover, the perceived sensing accuracy should satisfy the application requirements (e.g., accurate form and location of the event perimeter). In addition, future WSN deployments should allow for varied concurrent applications. Usually, these applications need varied information and have evolvable requirements.

Research supported in part by DFG GRK 1362 (TUD GKmM).

Moreover, achieving the best possible sampling accuracy and transport reliability is related to a large resource overhead, particularly, because Sensor Nodes (SNs) rely on batteries. A higher quality level is often related with higher deployment costs and higher resource overhead. A higher accuracy of spatial sampling of a spatial physical phenomenon of interest is usually achieved through a higher number of active sampling SNs in the area of the physical phenomenon resulting in a higher energy/bandwidth overhead. On the other hand, the transport reliability usually is achieved through a higher number of retransmissions. Hence, besides attaining the required quality levels, it is indispensable to maximize energy/bandwidth efficiency. Considering the design view, the sampling accuracy can be tuned by injecting some redundancy (e.g., activating more SNs on the perimeter for higher accuracy) and using sampling protocols that allow for over-sampling such as [9]. Generally, transport reliability is tunable through the number of transmissions.

The state-of-the-art on Quality of Information (QoI) [25] and Quality of Service (QoS) [1][12] in WSN lacks the online composite adaptation of sampling accuracy and transport reliability to the network conditions and application requirements. In fact, available approaches usually target single functional blocks [8][9], assuming that other functional blocks are perfect. The performed sampling accuracy satisfies the application requirements only if the information transport is perfect, which is not true in WSNs. On the other hand, the transport reliability assumes the sampling block to be perfect while addressing the application requirements. The optimized co-design of sampling and transport that maximizes the energy efficiency while satisfying the requirements is lacking in available approaches. In particular, there are no efforts in WSNs addressing the composite tunability of sampling accuracy and transport reliability.

Usually, the sensing application (users, services, feedback controller, etc.) has a specific requirement on the sensing accuracy. The sensing accuracy experienced at the sink fundamentally depends on the transport reliability. The key challenge consequently is to tune transport reliability and sampling accuracy in composition so that the requirement is met. The naive approach of massive over-sampling and allowing an arbitrary number of retransmissions might indeed result in high sensing accuracy. However, such a solution would be highly ineffective as it is not required to provide higher quality than the application requirements. On the other hand, this naive solution results in unacceptable energy overhead. Our work emphasizes that sampling accuracy can not be considered without transport reliability for an optimized efficiency.

Common to all these observations is that the application requirement have to be exactly satisfied by considering the holistic view of functional blocks. In addition, the right trade-off between sampling accuracy and transport reliability should be considered in all real-world applications to ensure satisfied applications. The challenge in finding methods for combining these attributes and localized algorithms for implementing the co-design efficiently.

Achieving both sampling accuracy and transport reliability while maximizing efficiency requires a sophisticated tradeoff technique, which is the main contribution of this paper. In our solution, we aim to find an optimal tradeoff between sampling accuracy and transport reliability. The same user experience could be achieved by different combination of both attributes. For example, providing higher sampling accuracy would allow for lower transport reliability. As it is complex to provide the optimized solution, we progress stepwise to master the complexity. Our solution considers energy in terms of retransmissions and sampling accuracy in terms of samples needed at the sink. Using probabilistic analytical expressions for relating sampling accuracy, transport reliability and efficiency, the desired outcome is a composition of the number of retransmissions per hop and the number of nodes to sample the phenomenon. The key challenge relies on minimizing the overall number of retransmissions given the number of hops, samples required, the user-required sampling accuracy and the link quality.

In summary our contributions are as follows:

- We provide a mathematical model for composite investigation of accuracy, reliability and efficiency.
- We formulate and solve a constrained optimization problem to determine the optimal combination of sampling accuracy and transport reliability that maximizes efficiency. Our solution relies on the proposed analytical model and considers varied levels of fidelity w.r.t. exactly meeting the application requirements on achieving a certain sensing accuracy.
- Through extensive simulations, we confirm the tunability and optimized performance of our sampling and transport co-design approach.

The structure of the paper is as follows. In Section II we present the related work. Section III describes the preliminaries with system model, terminology and the problem statement. In Section IV we detail our approach on sampling accuracy and reliable information transport co-design, i.e., interlinking sampling accuracy and transport reliability for developing the optimal solution. We provide the performance evaluation results in Section V.

II. Related Work

Providing the optimized co-design of sampling and transport is not straightforward due to the dynamic requirements and operational conditions. Traditional network design investigates the sampling and communication co-design from the simplistic view that the application data rate usually exceeds the capacity of the network and therefore the rate should be adapted accordingly. The additive increase/multiplicative decrease of the Transmission Control Protocol (TCP) is a renowned example of these efforts. Further efforts focus on varied application requirements and provide a QoS based design of network transport that allocates varied data rates to varied users. Also in WSN, QoS provisioning [1][12][13][5][10] are focused on network capacity and consider simplistic model of sampling. In networked process control community, a co-design of sampling and transport has been addressed. This co-design has been driven by the limited capacity of the network. In WSN, in addition to the network capacity constraint, the co-design should take into consideration the energy constraint, which is of higher priority.

The state-of-the-art in WSN focus either on the sampling accuracy (e.g., [6][9][23][7]) or transport reliability (e.g., [11][2][4][8]). However, there is no prior work addressing a co-design of sampling and transport in composition along online adaptation to satisfy user evolvable requirements while maximizing energy efficiency.

In [6], the authors address the node selection for optimizing accuracy in WSN. However, the information transport is assumed to be reliable. In [9], the authors propose an adaptive sampling approach to achieve user required accuracy and to avoid over-/under-sampling. While this poses an efficient and adaptive approach to model sampling accuracy, reliable transport is not considered in this work and reliability is implicitly assumed to be perfect. In [23], the authors address the spatial correlation based on MAC protocol called Correlation based Collaborative Medium Access Control (CC-MAC). However, though the authors address the optimized solution for accuracy, the transport reliability and timeliness are neglected.

In [7], the sampling for convergecast applications is addressed. However, adapting the sampling rate is independent of the application requirements. In [11], the authors focus on bursty convergecast where the key challenges are reliable and real-time error control and the resulting contention control. However, [11] does not offer mechanisms to adapt to changing application requirements and neglect the aspect of sampling accuracy. In [2], probabilistic techniques are applied for service differentiation. However, the solution aims at providing strict conditions for messages. In [4], the authors propose multi-path forwarding to ensure end-to-end delays. Also [4] is not adaptable to fluctuating network conditions to make routing decisions. However, optimizing accuracy and reliability for maximizing efficiency are missing in [2][4]. In [19], the authors propose metrics to measure the quality of a path. However, they do not address tunability. GIT [8], aims at satisfying the end-to-end reliability by dividing the reliability per hop. The proposed transport protocol is tunable regarding the achievable reliability. Providing a solid basis for reliability, [8] yet has to be extended to consider sampling accuracy.

Considering accuracy and reliability, mutual dependencies are not as straightforward. Hence, for a co-design it is not sufficient to just superpose a tunable sampling scheme with another tunable transport scheme. The challenge is still to provide efficient composite tunability of both data operations. In [3], the authors propose accuracy-aware context data collection and queries for heterogeneous mobile ubiquitous computing environments. However, the approach overlooks the transport reliability. The authors in [18] present a transport protocol with tunable timeliness and reliability. However, the work is optimized for a specific domain, i.e., real-time control and ignore the sampling quality. In [14], we consider the tuning of transport reliability and timeliness in composition, but without addressing the sampling accuracy. In [20], the authors present a co-design of data aggregation and data transport in WSN, ignoring the sampling operation. Summarizing, to the best of our knowledge there is no prior work on sampling and transport co-design for providing application required quality with optimized tradeoffs spanning accuracy, reliability and energy efficiency in WSNs. In this work, we build first steps to fill this research gap.

III. Preliminaries

In this section, we discuss the system model and we provide the terminology as the preliminary requirement for next sections. We also formulate the problem as a constrained optimization problem.

A. System Model

Our system model consists of a homogeneous WSN with static SNs and one sink. We focus on WSNs with network sizes ranging from dozens to hundreds. Typically, each SN is equipped with short range radio, and shows limited processing, storage and energy capabilities. We allow the sink to be adequate in power, memory and processing capabilities. SNs communicate with each other and the sink via bi-directional (multihop) wireless links.

We consider a physical phenomenon of interest that spans a specific small sub area of the WSN field. In general, the application is interested in one specific information about this spatial phenomenon, e.g., the perimeter of its area. We consider the communication disruptions constitute the most frequent failures. *Collisions, contention and congestion constitute the major causes of message loss and hinder information transport in WSNs and altogether result in the* R_{link}. We assume that network conditions are dynamic and application requirements are evolvable. We assume that the most strict application requirements do not exceed the maximal capacity of the WSN [22]. We assume a default Carrier Sense Multiple Access (CSMA)-based MAC and an underlying link state routing protocol, which provides a path for all SN towards the sink. Each SN knows its direct neighbors, e.g., through beaconing and the number of hops from the source to the sink. The basic exchange of messages from the source to the sink and vice versa is reliable. We assume that our mechanism starts after any SN detected the phenomenon. In [24], inspired from the current basic model we are extending the model with more complexity and less assumptions by formulating the problem in a more intuitive way.

TABLE I: Important notations and their meanings

R_{link}	The achieved success probability of one message transmission on one link
R_{hop}	The achieved success probability of message transmissions on one Hop after specific number of retransmissions
R_{path}	Reliability of one path
R_{inf}	The achieved success probability of the information (S_{min} samples) to reach the sink
Fi_{acc}	The sensing accuracy fidelity, i.e., is the expectation that the perceived sensing accuracy is equal to the desired sensing accuracy
S_{min}	The application desired number of samples from the phenomenon area
S_{tx}	The number of samples transmitted from the phenomenon area
S_{rx}	The number of samples received at the sink
h	Number of hops from sampling nodes to the sink
$\#ret_h$	Total number of retransmissions on one hop
$\#ret_{total}$	Total number of retransmissions induced by the transport of S_{tx} samples

A minimum number of spatial samples S_{min} is required to reconstruct the information on the sink. To this end, S_{tx} SNs sample this spatial phenomenon and transmit the samples towards the sink. We assume that the S_{tx} sampling SNs have the same number of hops h to the sink. The hops are considered as the average hop count from all the active sources to the sink. On the other hand, as we are interested in the small sub area of the phenomenon, variations of 1 or 2 hops do not affect the model and the end result. This is the case if the phenomenon area is small compared to the WSN field which is often the case for event-driven applications. The application requirements can be distributed to the SNs via a standard dissemination mechanism. We consider that the number of sampling SNs S_{tx} can be controlled, e.g., through an existing duty cycling algorithm that interacts with the sampling scheme, e.g., [9] to decide on which nodes to keep active.

B. Terminology

In the following we define important terms. An overview on the used notations is in Table I.

1) **Transport Reliability** (R_{path}): We define the end-to-end transport reliability as the success rate of one sample from one specific sampling node to reach the sink. Moreover, considering R_{link} on the lowest level, varying number of retransmissions affects R_{path} directly.
2) **Sensing Accuracy** ($Acc_{Sensing}$): The sensing accuracy is the accuracy of sampling as perceived by the application/user/sink. Accordingly, $Acc_{Sensing}$ is the ratio of the number of samples received at the sink S_{rx} to the minimum required number of samples S_{min}. $Acc_{Sensing}$ = $\frac{S_{rx}}{S_{min}}$. The sensing accuracy depends on the optimized combination of transport reliability R_{path} and activating the right number of SNs S_{tx} for sampling accuracy.

C. Problem Formulation and Objectives

Providing a specific requirement of S_{min} samples, the application actually expects exactly S_{rx} = S_{min} samples to be delivered. However, this guarantee is hard to be satisfied in WSNs. Therefore, we assume the application requires to meet the requirements with certain fidelity $Fi_{acc} \in [0, 1]$. Furthermore, generating only S_{min} samples and delivering all of them to the sink would require a large number of retransmissions.

Preliminary investigations have shown that by slightly increasing the number of generated samples S_{tx} we can significantly reduce the total number of transmissions needed to deliver S_{min} samples to the application. However, sending too many additional samples will finally result in unnecessary high number of retransmissions. Hence, we aim to find the optimal number of additional samples and the optimal path reliability that result in a minimal number of total retransmissions. Such an optimization allows to co-design sampling and transport for a maximized message efficiency, which transforms into maximized energy efficiency, as usually radio is the most energy consuming module on a SN.

Summarizing, we formulate the problem as follows:

$$Minimize\{\#ret_{total} : P(S_{rx} \geq S_{min}) \geq Fi_{acc}\}$$

More precisely, $\#ret_{total}$ can be expressed depending on the network characteristics and the application requirements as we will elaborate in the next section. The expected result is to determine the optimal (S_{tx}, R_{path}) tuples for given network conditions (link reliability R_{link}, hop distance h) and application requirements (S_{min}).

IV. Sampling and Transport Co-design

Fig. 1 illustrates the two operations spatial sampling and transport. For readability, we emphasize one sample (S_1) and one path towards the sink. The main reasoning behind the targeted sampling and transport co-design is to online tune both operations using optimized S_{tx} and R_{path} values. To this end, we first solve the formulated optimization problem. This requires to analytically express the total number of retransmissions $\#ret_{total}$ as a function of the sampling accuracy S_{min} and transport reliability R_p and to select those pairs that globally minimize the $\#ret_{total}$.

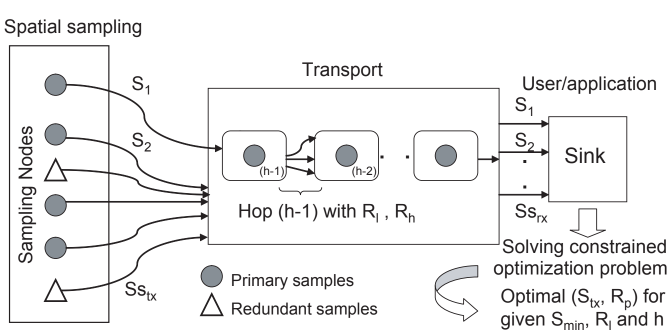


Fig. 1: A holistic view of sampling transport, and application interactions.

A. Computing the Total Number of Retransmissions as a Function of Sampling Accuracy and Transport Reliability

The total number of retransmissions occurring for a certain number of samples transmitted for the phenomenon area is the sum of all retransmissions on all traversed hops. Then, the expected maximum $\#ret_{total}$ can be computed as:

$$\#ret_{total} = S_{tx} * h * \#ret_h \tag{1}$$

The number of retransmissions per hop is determined by the achieved hop reliability R_{hop} and the underlying link quality R_{link} as shown in Fig. 1. The $\#ret_h$ is then computed as:

$$\begin{aligned} R_{hop} &= 1 - (1 - R_{link})^{\#ret_h} \\ \#ret_h &= \frac{log(1 - R_{hop})}{log(1 - R_{link})} \end{aligned} \tag{2}$$

Deriving Eq. (2) was the basic first step towards calculating the number of retransmissions per path. The achieved path reliability R_{path} depends on the achieved reliability at its hops as follows:

$$\begin{aligned} R_{path} &= R_{hop}^h \\ R_{hop} &= R_{path}^{\frac{1}{h}} \end{aligned} \tag{3}$$

In its turn, the achieved reliability R_{inf} depends on the achieved reliabilities of the paths the information (S_{tx} samples) traverses:

$$\begin{aligned} R_{inf} &= 1 - (1 - R_{path})^{S_{tx}} \\ R_{path} &= 1 - (1 - R_{inf})^{\frac{1}{S_{tx}}} \end{aligned} \tag{4}$$

Now we can express $\#ret_h$ as a function of h, S_{tx}, R_{inf} and R_{link} by substituting (4) in (3) and the resulting equation in (2):

$$\begin{aligned} 1 - (1 - R_{link})^{\#ret_h} &= R_{path}^{\frac{1}{h}} \\ 1 - (1 - R_{link})^{\#ret_h} &= (1 - (1 - R_{inf})^{\frac{1}{S_{tx}}})^{\frac{1}{h}} \\ (1 - R_{link})^{\#ret_h} &= 1 - (1 - (1 - R_{inf})^{\frac{1}{S_{tx}}})^{\frac{1}{h}} \\ \#ret_h * log(1 - R_{link}) &= log(1 - (1 - (1 - R_{inf})^{\frac{1}{S_{tx}}})^{\frac{1}{h}}) \\ \#ret_h &= \frac{log(1 - (1 - (1 - R_{inf})^{\frac{1}{S_{tx}}})^{\frac{1}{h}})}{log(1 - R_{link})} \end{aligned} \tag{5}$$

In Eq. (5) only R_{inf} is still not determined. As we have pointed out before, the combined accuracy and reliability application requirement consists of the number of samples S_{min} that have to be delivered to the sink. Our approach is to allow for a controlled degree of over-sampling and transport reliability that minimizes the total number of retransmissions while delivering the required S_{min} samples. In this work, we assume that any S_{min} samples from the generated S_{tx} samples fulfill the application requirement. The relation between reliability and the number of samples received at the sink accordingly can be defined as:

$$R_{inf} = \frac{S_{rx}}{S_{tx}}$$

which can be modeled as the expectation value of a Bernoulli process with S_{tx} trials and a success probability of R_{inf}. Hence, Eq. (5) becomes:

$$\#ret_h = \frac{log(1-(1-(1-\frac{S_{min}}{S_{tx}})^{\frac{1}{S_{tx}}})^{\frac{1}{h}})}{log(1-R_{link})} \quad (6)$$

Substituting (6) in (1), we obtain the total number of retransmissions as a function of sampling accuracy and transport reliability. This represents a fundamental basis for solving a crucial optimization problem.

B. Determining the Optimal Sampling Accuracy and Transport Reliability

Using the example of a Bernoulli process, the equation for R_{inf} can also be written as:

$$1-\left(\binom{S_{tx}}{0} R^0_{path}(1-R_p)^{S_{tx}-0}\right) = R_{inf}$$

and can be described as $P(at\ least\ 1\ out\ of\ S_{tx}\ samples\ is\ received)$ or according to the original notion as $1-P(all\ S_{tx}\ samples\ are\ lost)$. It is obvious that no information about the expected number of samples or the probability of receiving them can be given.

In order to be more flexible and to meet the application requirements we need to express the probability that at least S_{min} samples are received, which should be greater than or equal to the fidelity requirement Fi_{acc}. We describe it as the complementary probability of the event, where up to $S_{min}-1$ samples are lost:

$$1-\left(\sum_{i=0}^{S_{min}-1}\binom{S_{tx}}{i} R^i_{path}(1-R_p)^{S_{tx}-i}\right) \geq Fi_{acc} \quad (7)$$

Using this equation, the user can specify the reliability requirements more intuitively and precisely by providing Fi_{acc}. In order to use the above equation to derive our reliability requirements from the accuracy requirements, we need to solve it for R_{path}.

In the following we solve this by using the incomplete Beta function [15]. Obviously, the equation above describes the cumulative distribution function:

$$1-F(S_{min}-1) \geq Fi_{acc}$$

With the following relation of the distribution function to the Beta distribution:

$$\sum_{i=0}^{k}\binom{n}{i} \cdot S^i_{tx} \cdot (1-S_{tx})^{n-i} = I_{1-S_{tx}}(n-k, k+1)$$

where $I_x(a,b)$ is the regularized incomplete Beta function

$$I_{S_{tx}}(a,b) = 1 - I_{1-S_{tx}}(b,a)$$

we get the following derivation for R_{path}:

$$\begin{aligned}
1-\left(\sum_{i=0}^{S_{min}-1}\binom{S_{tx}}{i} R^i_{path}(1-R_p)^{S_{tx}-i}\right) &\geq Fi_{acc}\\
1-F(S_{min}-1) &\geq Fi_{acc}\\
1-I_{1-R_{path}}(1+S_{tx}-\lfloor S_{min}\rfloor, \lfloor S_{min}\rfloor) &\geq Fi_{acc}\\
I_{R_{path}}(\lfloor S_{min}\rfloor, 1+S_{tx}-\lfloor S_{min}\rfloor) &\geq Fi_{acc}\\
I^{-1}_{Fi_{acc}}(S_{min}, 1+S_{tx}-S_{min}) &= R_{path}
\end{aligned}$$

Therefore, the new expression for the number of retransmissions per hop, depending on the accuracy requirements is:

$$\begin{aligned}
1-(1-R_{link})^{\#ret_h} &= R^{\frac{1}{h}}_{path}\\
1-(1-R_{link})^{\#ret_h} &= I^{-1}_{Fi_{acc}}(S_{min}, 1+S_{tx}-S_{min})^{\frac{1}{h}}
\end{aligned}$$

$$\#ret_h = \frac{log(1-(I^{-1}_{Fi_{acc}}(S_{min}, 1+S_{tx}-S_{min}))^{\frac{1}{h}})}{log(1-R_{link})} \quad (8)$$

The optimal number of samples S_{tx} for a certain parameter setting can be found at the local minimum of $f(S_{tx}) = \#ret_{total}$, hence, the optimal number of active SNs is:

$$\begin{aligned}
\#ret_{total} &= \min\{f(S_{tx}) : S_{tx} \in \mathbb{N}\}\\
&= \min\{h \cdot S_{tx} \cdot\\
&\left\lceil \frac{log(1-(I^{-1}_{Fi_{acc}}(S_{min}, 1+S_{tx}-S_{min}))^{\frac{1}{h}})}{log(1-R_{link})} \right\rceil \\
&: S_{tx} \in \mathbb{N}\}
\end{aligned} \quad (9)$$

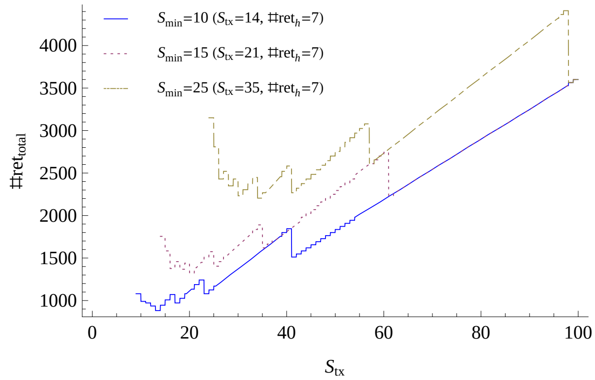
Note that $\#ret_{total}$ will always be an integer value due to the ceiling function applied to the number of retransmissions per hop (Eq. (8)), since non-integer values obviously can not be applied in practice.

C. Analytical Evaluation of the Sampling Accuracy and Transport Reliability

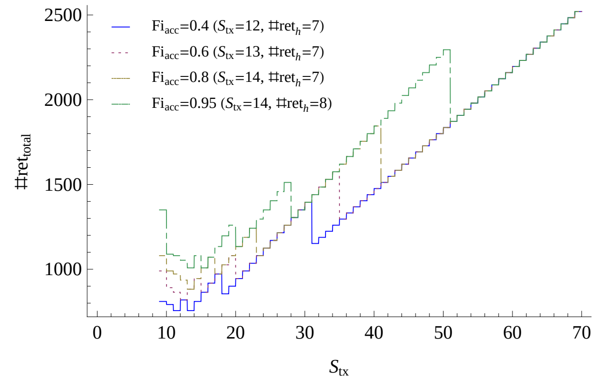
Based on the design goal, the objective function is to satisfy the application requirement given by the minimum number of samples S_{min} and a fidelity value Fi_{acc}, as indicated in the problem formulation and Eq. (9). Optimization and visualization of analytical results was conducted using Wolfram Mathematica [21]. Eq. (9) is plotted for selected settings in Fig. 2. Each graph consists of several linear segments resulting from the corresponding $\#ret_h$ value, which is highest for $S_{tx} = S_{min}$ and lowest as soon as so many samples have been added that retransmissions per hop are reduced to one (see Table II for examples of $\#ret_h$). Jumps from one segment to the next occur as soon as the reliability has been increased by redundant samples that much that Fi_{acc} is still satisfied when decreasing $\#ret_h$ by one. Note that there is always a small range where providing reliability using additional samples is more effective than using more retransmissions.

The main impact of the minimal number of samples required S_{min} is that at least S_{min} SNs need to sample the phenomenon. Furthermore, by increasing S_{min} the steps become larger until the $\#ret_h$ can be reduced. Higher requirements on Fi_{acc} obviously need high sampling/transmission redundancy. Hence, the threshold to reduce the $\#ret_h$ is higher for stronger requirements.

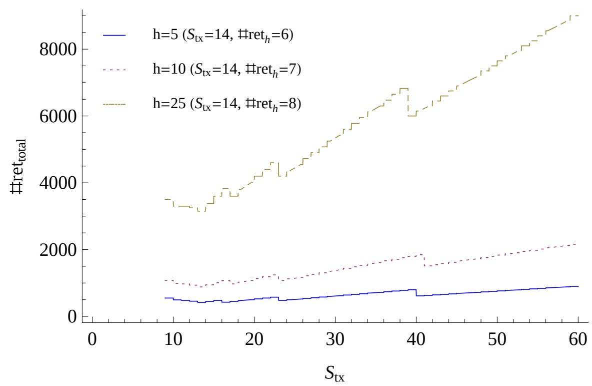
Furthermore, a higher requirement on Fi_{acc} generally decreases the potential gain in efficiency by activating more nodes. Besides the linear impact on $\#ret_h$, determination of the slope of the graph and of the initial number of retransmissions, the number of hops per path h also impacts how fast the next lower $\#ret_h$ can be achieved. Finally, the link quality has a significant impact especially on the $\#ret_h$,



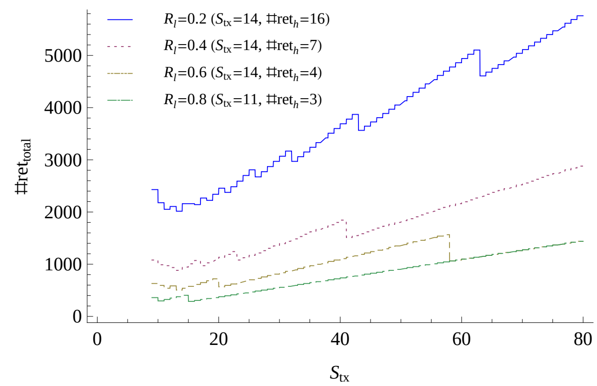
(a) Constant values are: $Fi_{acc} = 0.8$, $h = 10$, $R_{link} = 0.4$



(b) Constant values are: $S_{min} = 10$, $h = 10$, $R_{link} = 0.4$



(c) Constant values are: $Fi_{acc} = 0.8$, $S_{min} = 10$, $R_{link} = 0.4$



(d) Constant values are: $Fi_{acc} = 0.8$, $S_{min} = 10$, $h = 10$

Fig. 2: Graphs for different values of (a) S_{min}, (b) Fi_{acc}, (c) h and (d) R_{link}

since the low reliability has to be compensated by either more retransmissions or more samples. For a limited selection of WSN settings, we show the optimization results in Table II.

Fi_{acc} \ S_{min}		10		25	
	h \ R_L	0.4	0.6	0.4	0.6
0.8	10	$(S_{tx} \vert \#ret_h \vert)$	$(S_{tx} \vert \#ret_h \vert)$	$(S_{tx} \vert \#ret_h \vert)$	$(S_{tx} \vert \#ret_h \vert)$
		14 | 7	14 | 4	35 | 7	34 | 4
	15	14 | 8	14 | 5	31 | 9	31 | 5
0.95	10	14 | 8	16 | 4	33 | 8	37 | 4
	15	14 | 9	14 | 5	33 | 9	33 | 5

TABLE II: Optimal tuples of the number of samples (S_{tx}) and the number of retransmissions per hop ($\#ret_h$) for a selection of parameter settings.

D. Integrated Sampling and Transport

So far, we determined the optimal accuracy and reliability settings using global view. In the following, we present on preliminary efforts towards a localized integrated sampling and transport in generalized WSNs and its practicality.

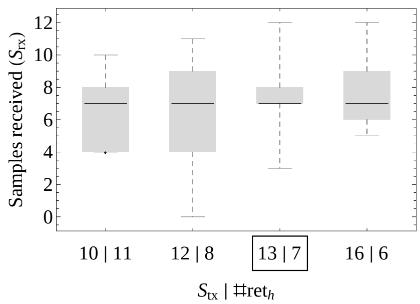
After the phenomenon detection and notification from the source to the sink, the sink immediately knows about the important properties such as link reliability and hop count. Fidelity and accuracy requirements are provided by the user or the application and always accessible to the sink. Having this information the sink can then solve Eq. (9) for the phenomenon area. The attained optimal values (S_{tx} and $\#ret_h$) are reliably transmitted to the sources in the phenomenon area. The overhead induced by the reliable communication is negligible since only a single message has to be transported reliable. After the source has received the values for the sampling and transport co-design, it can use the existing duty-cycling algorithm to a) activate the right number of SNs and b) to notify them about the number of retransmissions for the information transport. As soon as a SN is activated the user required sample is transported towards the sink with optimal number of retransmissions. As for the information transport each SN forwards the optimal number of retransmissions to upstream nodes by appending the number to the actual sample.

V. Performance Evaluation

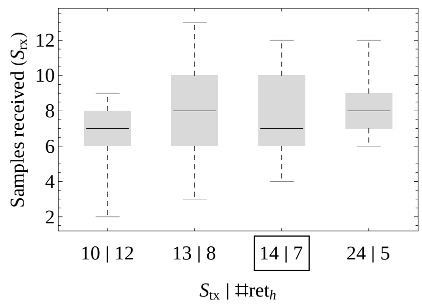
In order to evaluate our work, we first describe the simulation environment, simulation settings and the performance metrics. Next, we present our simulation results.

A. Simulation Settings and Performance Metrics

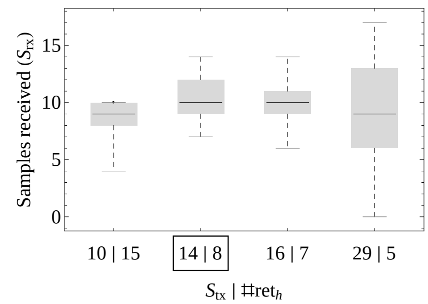
We simulate 100 SNs deployed in an area of 10×10 unit^2. The sink is located at one corner. The spatial samples are generated from one corner and transported towards the sink. We run the simulations in TOSSIM [16]. Simulation results are compared with the analytical benchmarks in order to verify if an analytical optimal solution indeed corresponds to an optimal solution in practice.



(a) $Fi_{acc} = 0.6$, $R_{link} = 0.4$

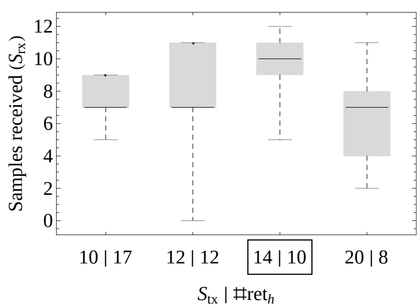


(b) $Fi_{acc} = 0.8$, $R_{link} = 0.4$

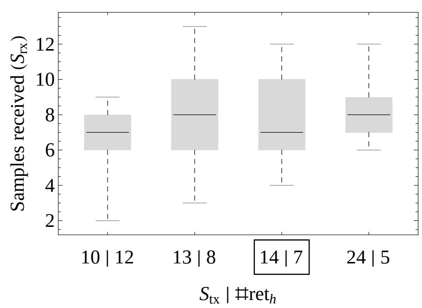


(c) $Fi_{acc} = 0.95$, $R_{link} = 0.4$

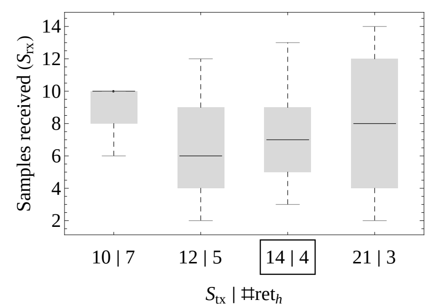
Fig. 3: Impact of varying fidelity requirement (Fi_{acc}) on sensing accuracy



(a) $R_{link} = 0.3$, $Fi_{acc} = 0.8$

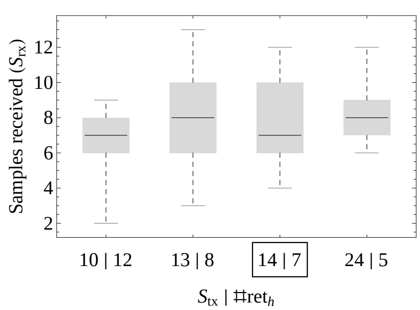


(b) $R_{link} = 0.4$, $Fi_{acc} = 0.8$

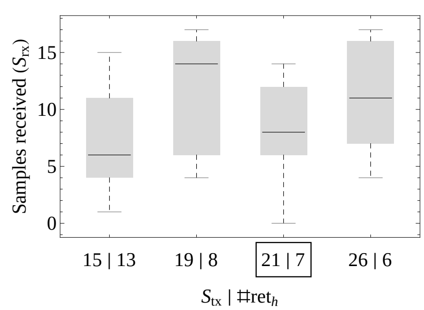


(c) $R_{link} = 0.6$, $Fi_{acc} = 0.8$

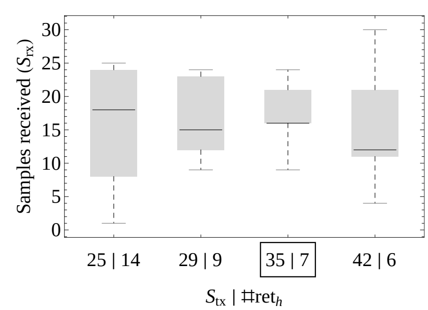
Fig. 4: Impact of varying link reliability (R_{link}) on sensing accuracy



(a) $S_{min} = 10$



(b) $S_{min} = 15$



(c) $S_{min} = 25$

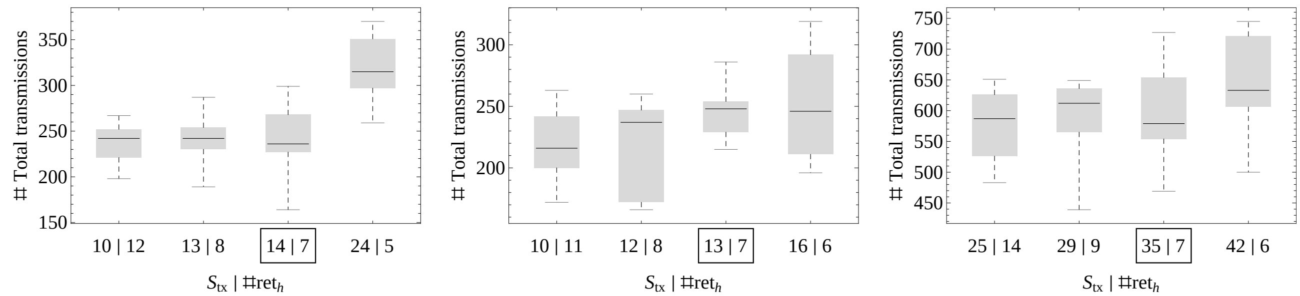
Fig. 5: Impact of varying samples application requirements (S_{min}) on sensing accuracy

The performance of our protocol is measured in terms of accuracy (samples received at the sink) and efficiency (average number of transmissions). Any of these combinations, computable by using optimal solution for the $\#ret_h$ are valid, i.e., they guarantee the delivery of S_{min} samples with a probability greater than or equal to Fi_{acc}. To show the online adaptation of our work, we first vary the application constraint Fi_{acc} keeping the number samples S_{min} and link reliability R_{link} constant. Later, we vary the link reliability R_{link}, keeping Fi_{acc} and S_{min} as constant. Finally, we vary S_{min} while keeping Fi_{acc} and R_{link} as constant. In all figures the optimal tuple is highlighted with a square box and have considered the median.

B. Simulation Results

In order to verify the analytical optimal solutions complying to optimal solutions in simulations, we selected combinations around the optimal combination, i.e., combinations with less samples and more $\#ret_h$ as well as combinations with more samples and less $\#ret_h$. Hence, regarding the total number of transmissions, if simulations for all combinations yield worse results than for the optimal ones, it is a strong indicator that this combination indeed is optimal. As for the samples received, we expect very similar results throughout all variations as an indicator for the tunability of our model. In some cases, when an optimal solution existed for more than one combination, the first one, i.e. using less active SNs, was chosen. Furthermore, in practice the number of transmissions is very likely to be less than the computed values, as it is dependent on whether delivery was successful or not.

Representative results of the simulations for varying the constraint Fi_{acc} are shown in Fig. 3. As expected, the choice of Fi_{acc} has influence on the delivery of S_{min} samples, as all combinations denote valid options to satisfy the requirements.



(a) $Fi_{acc} = 0.8$, $R_{link} = 0.4$, $S_{min} = 10$, $h = 10$

(b) $Fi_{acc} = 0.6$, $R_{link} = 0.3$, $S_{min} = 10$, $h = 10$

(c) $Fi_{acc} = 0.8$, $R_{link} = 0.4$, $S_{min} = 25$, $h = 10$

Fig. 6: Efficiency for varying S_{min}, Fi_{acc} and R_{link}

Furthermore, the delivery rate has been slightly higher for the simulations with higher requirements for Fi_{acc}, completely conforming to our expectations.

The impact of varying link reliability R_{link}, is shown in Fig. 4. In all the cases of varying R_{link} we achieve the desired number of S_{min} samples to reach the sink.

Representative results of the simulations for varying the S_{min} samples are shown in Fig. 5. The results are very close to our analytical solution. The impact of varying S_{min} samples does not affect the application requirements and completely satisfies the probability of satisfying the Fi_{acc} and S_{min}.

One more important aspect for the assessment of the validity of our model regarding optimal energy efficiency is the total number of transmissions (Fig. 6). Accordingly, we have varied Fi_{acc}, R_{link} and S_{min}. In all the cases, we achieve a desired optimal result that satisfies our requirement and conforms to our analytical model and solution. Finally, it remains to observe that the variations in hop length and link reliability introduced in simulations in contrast to the analytical model result in a noticeable noise throughout simulations, blurring the differences between the different variations.

VI. Conclusion

Through this paper we have achieved important steps towards the co-design of sampling and transport as per the application requirements. We have developed an analytical model for the case that no differences between sensor readings have to be regarded. This simplifies the problem of finding a specific subset of nodes to the problem of merely finding the optimal number of nodes that have to send a sample. Our analytical model gives the optimal number of SNs, so that the specific application requirements are satisfied. The optimized solution provided depending on the application requirements, reduces the total number of retransmissions by adding redundancy and sending more samples than required. This is the first instance of real time adaptation when an integrated sampling and transport solution is implemented. The present work is just focusing on the accuracy and reliability attributes and is further being extended for additionally considering the timeliness attribute.

References

[1] D. Chen et al., QoS Support in Wireless Sensor Networks: A Survey. In Proc. of ICWN 2004.

[2] E. Felemban et al., MMSPEED: multipath Multi-SPEED protocol for QoS guarantee of reliability and timeliness in wireless sensor networks. In Proc. IEEE Mobile Computing, 2006, pp. 738 - 754.

[3] Q. Han et al., Timeliness-accuracy balanced collection of dynamic context data. IEEE Transactions on Parallel and Distributed Systems, 2007, pp. 158 - 171.

[4] X. Huang et al., Multiconstrained qos multipath routing in wireless sensor networks. In Proc. of Wirel. Netw, 2008, pp. 465 - 478

[5] K. Karenos et al., A rate control framework for supporting multiple classes of traffic in sensor networks. In Proc. of RTSS, 2005.

[6] H. Li et al., Information-accuracy-aware jointly sensing nodes selection in wireless sensor networks. In Proc. of MSN, 2006, pp. 736 - 747.

[7] Y. Sankarasubramaniam et al., ESRT: Event-to-sink reliable transport in wireless sensor networks. In Proc. of MobiHoc, 2003, pp. 1003 - 1016.

[8] F. K. Shaikh et al., Generic information transport for wireless sensor networks. In Proc. of SUTC, 2010, pp. 27 - 34.

[9] P. Szczytowski et al., Asample: Adaptive spatial sampling in wireless sensor networks. In Proc. of SUTC, 2010, pp. 35 - 42.

[10] W. L. Tan et al., Performance evaluation of differentiated services mechanisms over wireless sensor networks. In Proc. of VTC, 2006.

[11] H. Zhang et al., Reliable bursty convergecast in wireless sensor networks. In Proc. of MobiHoc, 2005, pp. 266 - 276.

[12] J. Martnez et al., QoS in wireless sensor networks: Survey and approach. In Proc. of EATIS, 2007, pp. 1 - 8.

[13] M. Ash et al., Quality of Service in mobile ad hoc networks: a survey. In Proc. of IJAHUC, 2010, pp. 75 - 98.

[14] V. Sachidananda et al., Trading Transport Timeliness and Reliability for Efficiency in Wireless Sensor Networks. In Proc. of ICNSC, 2013.

[15] J. Dutka, The incomplete beta function a historical profile. Archive for History of Exact Sciences, 1981, pp.11 - 29.

[16] P. Levis et al., TOSSIM: Accurate and scalable simulation of entire tinyos applications. In Proc. of SenSys, 2003, pp. 126 - 137.

[17] A. Ghosh et al., Coverage and connectivity issues in wireless sensor networks: A survey. In Proc. of PMC, 2008, pp. 303 - 334.

[18] P. Park et al., Breath: An Adaptive Protocol for Industrial Control Applications Using Wireless Sensor Networks. In Proc. of IEEE Transactions on Mobile Computing, 2011, pp. 821 - 838.

[19] J. Wang et al., QoF: Towards comprehensive path quality measurement in wireless sensor networks. In Proc. of Infocom, 2011, pp. 775 - 783.

[20] H.X. Tan et al., Information Quality Aware Routing in Event-Driven Sensor Networks. In Proc. of Infocom, 2010, pp. 1 - 9.

[21] http://www.wolfram.com/

[22] P. Gupta et al., The capacity of wireless networks. In Proc. of IEEE Transactions on Information Theory, 2000, pp. 388 - 404.

[23] M. C. Vuran et al., Spatial Correlation-based Collaborative Medium Access Control in Wireless Sensor Networks. In IEEE/ACM Transactions on Networking, 2006.

[24] V. Sachidananda et al., On Co-modeling the Sampling and Transport in Wireless Sensor Networks. In Proc. of FGSN, 2012.

[25] V. Sachidananda et al., Quality of Information in Wireless Sensor Networks: A Survey. In Proc. of ICIQ, 2010.