

Energy-Efficient Methods for Highly Correlated Spatio-Temporal Environments in Wireless Sensor Network Communications

Mohammad Abdul Azim¹, Zeyar Aung¹, Sofiane Moad², Nizar Bouabdallah²,
Mario E. Rivero-Angeles³, Israel Leyva-Mayorga⁴

¹Masdar Institute of Science and Technology, Abu Dhabi, UAE

²DYONISOS, INRIA, Rennes, France

³National Polytechnique Institute, CIC-IPN, Mexico City, Mexico

⁴National Polytechnique Institute, SEPI ESCOM-IPN, Mexico City, Mexico

Email: azim@ieee.org, zaung@masdar.ac.ae, smoad@ieee.org, nbouabdallah@ieee.org, mriveroa@ipn.mx,
ileyvam0600@alumno.ipn.mx

Received 2 April 2014; revised 2 May 2014; accepted 9 May 2014

Copyright © 2014 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Continuous-monitoring (CM) of natural phenomenon is one of the major streams of applications in wireless sensor networks (WSNs), where aggregation and clustering techniques are beneficial as correlation dominates in both spatial and temporal aspects of sensed phenomenon. Conversely, in Event Driven Reporting (EDR), the efficient transmission of sensitive data related to some predefined alarm cases is of major importance. As such, reporting latency is a more important performance parameter. However, in some applications, the transmission of both CM and EDR data is encouraged or even required. For either CM or EDR applications, system performance can be greatly improved when both the number of packets to be transmitted as well as the packet size is reduced. This is especially true for highly dense sensor networks where many nodes detect the same values for the sensed phenomenon. Building on this, this paper focuses on studying and proposing compression techniques to improve the system performance in terms of energy consumption and reporting latency in both CM and EDR applications. Furthermore, we extend our analysis to hybrid networks where CM and EDR are required simultaneously. Specifically, this paper presents a simple aggregation technique named smart aggregation (SAG) for the CM applications and an event driven scheme named compression cluster scheme in spatial correlated region (CC_SCR). The proposed SAG exploits both spatial and temporal correlations where CC_SCR exploits the spatial correlation of such networks by data compression. Rationalizing the developments is attained by simulations that compare energy efficiency of the proposed SAG with k -hop

How to cite this paper: Azim, M.A., et al. (2014) Energy-Efficient Methods for Highly Correlated Spatio-Temporal Environments in Wireless Sensor Network Communications. *Wireless Sensor Network*, 6, 67-92.

<http://dx.doi.org/10.4236/wsn.2014.65009>

aggregation and CM based event driven reporting (CMEDR) schemes. Results of CC_SCR show that the technique may reduce the energy consumption drastically. In some specific cases the reduction becomes more than 10 times compared to a classical clustering scheme. Two different strategies for the transmission of event reports through the CM infrastructure are incorporated: PER and NPER protocols. Both strategies take advantage of the cluster-based architecture which assigns a TDMA schedule for the CM data transmission while using NP/CSMA for the transmission of the event information. Consequently, no extra energy is consumed for separate event clusters. As such, the number of packets to be transmitted is greatly reduced.

Keywords

Wireless Sensor Networks, Aggregation, Continuous-Monitoring, Event Driven Reporting, Data Compression, Clustering

1. Introduction

Tiny computers equipped with sensors and radios known as sensor nodes are the basic building blocks of wireless sensor networks (WSN) [1]. Well-diversified applications and numerous challenges due to the resource constraints make the WSN research vibrant.

1.1. Diversified Monitoring Paradigm

A range of impending applications is envisaged in the area of WSNs (e.g., surveillance, tracking, environmental monitoring etc.), where WSN can be broadly categorized as continuous-monitoring (CM), event driven monitoring and hybrid.

CM applications: In CM approach, data is generated continuously with a predefined reporting frequency. To aggregate data in such networks cluster-based architectures can be used such as low energy adaptive clustering hierarchy (LEACH) [2] protocol. Following cluster formation, nodes send the sensed data to the cluster heads (C-Hs), which in turn aggregate and forward the collected data to the sink.

Event driven monitoring: In event driven monitoring applications sensor generates data packets only when an event is triggered. Threshold sensitive energy efficient sensor network protocol (TEEN) [3] is an event driven data generation approach where energy is saved by limiting generation of the packets. TEEN introduces two thresholds of sensed values namely a hard threshold (T_h) and a soft threshold (T_s), in its mechanism of limiting the number of packets. Data is only sent to the sink if the current value is greater than T_h or if the difference between the current and the previous value is greater or equal to T_s .

Hybrid protocol: As mentioned earlier, CM and event driven applications are different types of protocols where the former collects the periodic data proactively while the later reports time critical data reactively. CM applications send data to the sink even if there is no significant importance of the sensed values thereby resulting in wastage of resources. Conversely, if the thresholds are not reached, event driven application may not deliver data to the sink for a long period of time. Moreover continuous sensing required with event driven applications might cause energy wastage. To resolve this issue a hybrid protocol named adaptive periodic threshold-sensitive energy efficient sensor network (APTEEN) [4] [5] has been proposed. This protocol can responds to the queries on the historic data, forwards the periodic data and also responds to any sudden change on the sensed values.

1.2. Resource Constraints

Being tiny and inexpensive devices, resource limited in computation, communication and energy pose challenges in design and developing of the applications for WSNs. Among the resource limitations the most important challenge is the energy efficiency as the battery lifetime is considered as the network lifetime in the arena of WSNs [1]. This poses a requirement of efficient hardware and software solutions for all the perspectives of design in the sensor nodes and networks. Power efficient hardware and software are required to tackle the resource constraints in WSN. Consequently Telos [6] like hardware platform, energy efficient path selection applications, sleep-enabled medium access control mechanism in the context of WSN have been developed [7].

The key objective of a deployed WSN is to reliably report a detected event based on the cooperative effort in densely deployed sensor networks. Limited information provided by individual report is compensated by the collective sensing notion, which is brought about by the dense deployment of sensor nodes. The collaborative nature of the WSN has significant improvements over traditional sensor nodes including greater accuracy and a larger coverage area; it induces high correlation among sensor measurements. Exploiting this correlation is a powerful tool, which can drastically enhance overall network performance, and especially network lifetime. In-network processing includes data aggregation and data compression. We focus on fusion approach of in-network data aggregation, which is widely used to adapt to the diversity that exists in WSNs. With this in-network aggregation, sensor nodes carry out part of the aggregation rather than just propagate raw readings. This reduces power consumptions because sensor power usage is dominated by transmission costs. Thus, being able to transmit less data results in reduce energy consumption at node level.

1.3. Our Contributions

Exploiting the collaborative nature in the correlated information paradigm of the WSN and achieving energy efficiency is the key criteria of the proposed techniques for both the CM and event driven reporting (EDR) applications perspective.

- We investigate CM applications that aggregate data with a bounded error in energy-constrained sensor networks. A simple aggregation technique named Smart Aggregation (SAG) [8] is presented for CM applications in WSNs. SAG maintains a tolerable deviation *i.e.*, limited error in the aggregated data while aggregating both spatially and temporarily.
- Contrary to CM an EDR approach of sensor application exploiting spatial correlation is derived by a novel compression cluster scheme in spatial correlated region (CC_SCR) [9] [10]. In this scheme, the cluster head selects a single value from all the available values transmitted by the cluster members and it is chosen as reference called reference value; this parameter is transmitted to all the nodes inside the cluster in such an order that, the other active nodes transmit only the difference between the reference and sensed value. The reference value selection mechanism aims at obtaining the reference node in such a way that the reduction of the packet size becomes the highest compared to all the other active nodes in the vicinity.
- We also focus on hybrid WSNs where both CM and EDR capabilities are needed. In particular, we consider a CM WSN where data is reported periodically to the sink node. Also, once an event occurs, the sensor nodes in the event area reports data. In doing so, the end user acquires more information from regions with turbulence. Two strategies for the event reporting in clustered based architecture are proposed for Time Division Multiple Access (TDMA) based Medium Access Control (MAC) mechanism.
- Both strategies are based on assigning a specific period for the event related data through already formatted clusters for the CM applications. As such, there is no extra energy drains for the formation of separate clusters for the event reporting. Hence, the number of required packets to transmit the EDR data is reduced. In this study, we show that by assigning this event driven period, the performance of the network, in terms of energy consumed, is improved compared to other protocols that are specifically designed either for CM only or EDR only applications. In the first proposed strategy, NPER, the nodes inside the event area report their data at the end of the CM transmissions. The second strategy, PER, considers the transmission of event related data at the beginning of each TDMA slot assigned to each cluster member for CM transmissions.
- It is important to mention that in highly correlated environments, it is reasonable to consider that the nodes inside a cluster share the same information for both CM and event reporting data. In particular, for the hybrid mechanism, it is still possible to perform aggregation or compression techniques such as the ones proposed in this work.

The remainder of this paper is organized as follows. Section 2 presents the related works regarding the in-network aggregation, the compression and the clustering schemes. Section 3 describes the proposed aggregation protocol, while Section 4 presents the proposed compression clustering protocol, and Section 6 covers the proposed event reporting and CM protocols. Section 6 reports the simulation results, followed by some concluding remarks in Section 7.

2. Related Work

We review some of the related works regarding the in-network aggregation, the compression and the clustering

schemes presented in this work.

2.1. In-Network Aggregation Schemes

Based on the aggregation concept, extensive research efforts have been conducted in the literature on designing correlation-avoidance mechanisms in WSNs, such as in-network processing [11] [12]. Tiny DB [13] embraces directed diffusion DD [14] approach where the network is considered as a distributed database. The aggregation service to the Tiny DB named tiny aggregation (TAG) [15] maintains an epoch time specified by the user and determines the arrival rate at which the user receives the sensed data. Data is collected through the aggregation tree. Unfortunately construction of a single tree may not be sufficient in case of unreliable links. To address the unreliable wireless channel, directed acyclic graph based in-network aggregation (DAG) for sensor network monitoring is proposed [16]. DAG enables each node having multiple parents to collect the data through a fault tolerant directed acyclic graph.

Temporal coherency-aware in-network aggregation (TiNA) [17] suppress the amount of data by exploiting temporal correlation to conserve transmit energy. Unlike TAG, TiNA transmits the data based on the present and past sensed values and avoids monotonic periodical transmissions. A sensed value is only transmitted if $t_{ct} < (V_{\text{new}} - V_{\text{old}})/V_{\text{old}}$; where t_{ct} is user specified desired tolerance to the particular query. And V_{old} and V_{new} are the old and new sensed values respectively. Attempting to further reduction of transmit power clustered aggregation (CAG) [18] proposes a technique that exploits temporal as well as special correlation of the physical sensed values. The cluster in CAG is formed based on a predefined threshold as well as the sensed value of sensors. As such, the spatial correlation is exploited. A reformation of a cluster only triggers when the sensed value exceeds a threshold over the time. Rather than sending values from each node like TAG, CAG sends a single value from each cluster to the sink thereby achieves a high-energy efficiency.

Isoline aggregation (IA) [19] supports CM of environmental conditions by detecting isolines. Note that isolines are the lines that make up a contour map of the sensed values. In IA, only the nodes that detect the isolines generate data and achieves energy efficiency.

Layered clustering structure based aggregation technique provides delay and accuracy sensitive reporting in sensor data networks [20]. Depending on the application requirements and location information the network is divided into several layers. Each layer is further divided into clusters. Accordingly, each layer belongs to different delay and accuracy. Based on the application requirement a query is sent to the right layer. The shortcoming of the application is twofold. Such centralized protocol is often inappropriate for a large distributed network. Furthermore localization dependent aggregation often considered not very practical in randomly deployed networks as localization itself is expensive.

Data centric adaptive in-network aggregation provides an algorithm that dynamically calculates the accepted delay based on past values therefore limiting the error due to delay of reporting [21]. The mean and standard deviation of previous set of values are the deciding factor of the accepted delay. While aggregating, in case that immediate past value is within the accepted delay limit, the value is incorporated. Conversely, if the value is older than the limit then it is ignored. Though the proposal handles the freshness of the values it still cannot handle the error incurred due to the aggregation function itself. Tiny aggregation algorithm with pre-processing in local cluster (TAGPP) [22] provides an aggregation technique to deal with the measured error and noise disturbance on sensed data due to the interferences in harsh environment. Cluster heads in the local cluster eliminate outlier and derive clean and reliable data. Additionally data fusion is performed by weighted fuzzy clustering algorithm. To this end this paper proposes a simple aggregation technique well suited for resource constrained WSN.

2.2. Compression Scheme

Compression algorithms in WSN are studied intensely in the literature. Unfortunately, the study is confined to the so-called classical networks and is not suitable to deploy in the context of WSN [23] [24]. The underlying reason of unsuitability comes from the fact that the nodes are of limited memory and processing power. In case of memory, the size of bzip2 and LZO are 219 KB and 220 KB respectively a bit too high for WSN. With 4 - 8 MHz of processing power in WSNs while embedding classical compression, tiny nodes are rather incapable of handling the operations required in such schemes. Therefore, nodes require fresh design of low-complexity and small sized data compression schemes to be developed and be deployed with. A number of the compression

schemes are reviewed in this literature. For further detailed of such compression techniques please refer to the reference [25].

In coding-by-ordering data-compression scheme [26] an aggregation node combines data and among the data it is combining some of the combined data is transmitted implicitly rather than explicitly. Replacing the data transmission of certain nodes by a certain order *i.e.*, the order of the aggregated packets by the aggregation node is the underlying idea of such compression scheme.

The pipelined-in-network compression algorithm [27] combines data from different nodes and make it smaller compared to the original algorithm. The aggregator node collects and stores the data for a certain time. The stored data then is combined by the node into a single packet. Therefore the total transmission number is decreased and energy expense is reduced. Note that, this is a simple compression scheme where the shared prefix system is utilized for both node ID and timestamp fields as a result more data-compression can be achieved.

The distributed compression scheme uses side information to encode source information [28]. For example, if there are two data sources: X and Y , which are correlated and chosen from a discrete alphabet. Then, X can be compressed at the theoretical rate of its condition entropy. The receiving node maintains the cosets and can decode X knowing Y 's code-vector, and with partial information from the source X .

The main advantage of the proposed CC_SCR algorithm, compared to those previously mentioned is that CC_SCR takes into account the physical characteristics of the sensed data in order to compress data and also considers clustering communication in order to reduce energy consumption. It is important to notice that in WSNs with low node density, the difference in power can be almost negligible under certain conditions. As such, this proposal is effective in most dense WSNs.

2.3. Clustering Schemes

In addition to the compression techniques, there have been increased interests in the area of energy efficient clustering algorithms. As a result extensive clustering algorithms have been proposed for WSNs. Hereafter, we briefly review the most relevant energy efficient clustering algorithms. For more details, the reader can review [29]-[32].

Hybrid Energy-Efficient Distributed clustering (HEED) [33] protocol operates in two main phases: the set-up phase where clusters are formed and the steady phase where the sensor nodes transmit their data using the Time Division Medium Access (TDMA) frames. HEED set-up phase operates in three sub-phases. The first sub-phase is the initializing. Nodes exchange *hello* messages to discover their neighborhoods. The second sub-phase consists of a competition process. The third sub-phase is the finalizing and it allows nodes to join their corresponding C-H based on the connectivity degree. Power-Efficient Gathering in Sensor Information Systems (PEGASIS) [34] steady phase consists of a formation of chains instead of clusters. In the chain formation, the BS and sensor nodes are connected via a chain using a greedy algorithm. One of the nodes, in the chain, is selected (by turns) to represent the head. In data gathering phase, each node delivers the sensing data to the nearest neighbor node until the data reach the head node, which aggregates and delivers the sensing data to the BS.

In Energy Efficient Clustering Scheme (EECS) protocol [35], C-H candidates compete for the ability to elevate to a C-H with a certain probability. This competition involves candidates broadcasting their residual energy to neighboring candidates. If a given node does not find a node with more residual energy, it becomes a C-H.

The main difference between the aforementioned clustering algorithms and the CC_SCR algorithm is that nodes uses the physical characteristic of the sensed data to elects C-Hs. The benefit of our proposed scheme is that nodes use the compression scheme to reduce the energy consumption.

In this work, we combine the benefits of using both the clustering and the compression techniques to reduce the energy consumption in the network. Indeed our proposed scheme takes into account the symptomatic of physical surveyed event. The scheme exploits the energy unconstrained Base Station (BS) take part into C-H selection mechanism. The theme of our developed technique is to take advantage of the fact that all the sensor nodes detecting of a certain event indicating the involving nodes resides at the close proximity of each other. That entails a high correlation between sensed data and provides scope reduction of size of the data packet that communicating through the network. As a result, BS selects C-H efficiently therefore minimize total transmitting bits on the air.

2.4. Hybrid Protocols for CM and EDR

Finally, the concept of an *event cluster* is presented in event reporting on continuous monitoring wireless sensor

networks [36]. The basic idea is to form a separate cluster where the cluster members are the nodes inside event area. By doing this, the compression algorithm used [26] takes advantage of the spatial correlation of the event data in order to increase the compression ratio. However, this protocol is only suited for applications where the event rate (number of events per unit of time) is low, since the separate cluster formation consumes an important amount of energy. Also, it is only suited for single events, *i.e.*, only one event can occur in the supervised area since multiple events generate a high collision probability in the separate cluster formation procedure. As opposed to [36], the proposed strategies in our work focus on general event rates and number of events in the network. Since the occurrence of events in the system do not entails separate clusters, no extra energy is consumed when an event is detected. This is done by taking advantage of the already formed clusters for the CM data transmission. As such, these proposed strategies can be used for any system's conditions. Also, the proposed strategies presented in this work can be benefited by the introduction of compression algorithms, such as the one presented in [26] which is considered for future work. **Figure 1** depicts the proposed protocols in the highly correlated spatio-temporal paradigm in WSNs.

3. Proposed Aggregation Protocols

The applications of interest are monitoring applications deployed in a specific geographic area for a single or multiple natural phenomenon(s). In-network aggregation techniques intend to reduce the amount of redundant data through various techniques. The underlying idea is to reduce the transmission energy for resource constrained WSNs with fewer amounts of data while preserving the quality of detection. Building on this, three different methods for data compression and efficient data transmission are proposed: SAG and CC_SCR.

3.1. Network Model

In the area of interest (where the WSN is deployed) the sensors measure the phenomenon repeatedly. The measurement is rather discrete than continuous even for CM applications. We consider a CM application with a predefined sampling frequency of f_1 in normal conditions. However with a trigger of an event f_1 sampling rate might not be adequate enough to support the altered environmental condition. Therefore in an intensified situation the application intends to monitor the environment with a greater interest by using a higher sensing frequency

Proposed Protocols

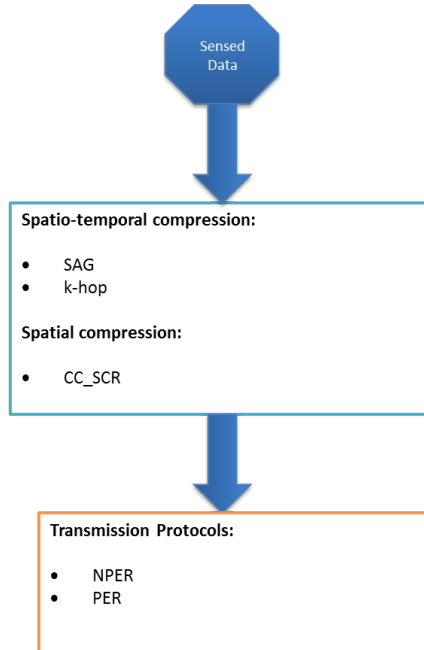


Figure 1. The proposed protocols in the spatio-temporal WSN networked systems.

f_2 . Therefore the sensing takes place with two different frequencies f_1 and f_2 , where $f_1 < f_2$. Whenever an event is detected, the goal is to perform a monitoring by the leaf nodes in an intensively fashion while nodes halts the aggregation process.

An example of the CM application is depicted in **Figure 2**. A network with tree-like topology is the underlying structure of the example, where fifteen sensor nodes are responsible for collecting temperature of a particular geographic space with a sampling rate of $f_1 = 2$ values/minute. Alternatively, each sensor node is sampling its own environment periodically each 30 second. Node ID and the current sensed temperature values are presented in such figure. When there is a trigger of an event at a particular point *i.e.*, in a sensor node, an aggressive monitoring approach is made. For example if an event is defined as a temperature more than 70 degree. When temperature at node 10 senses a temperature of 71 degree, an event is triggered. With the occurrence of an event, node 10 starts to observe the environment with a sampling rate $f_2 = 4$. Therefore, from this point and until the temperature goes below the triggered value *i.e.*, temperature > 70 degree, node 10 keeps monitoring the environment every 15 seconds which is twice of the normal sampling frequency. In addition the node in its every message sets a flag that indicates the message is in priority. Therefore the intermediate nodes (node 6 and node 2 in **Figure 2**) can forward the messages with a priority basis. As soon as the intermediate node receives a message with such flag set, the node forwards the message straightaway. An aggregation is avoided as it incurs additional delay.

Fundamentally sensing and reporting data depends on the sampling frequency. Nodes that do not have any childnode send data to its parent node without incurring in any delay. Conversely, the nodes with a child wait for a predefined threshold amount of time before sending its value to its parent. A child link breakdown or reconstruction and redefinition of a tree may cause no reporting to a specific parent from a particular child. Routing protocol is responsible to maintain the tree fresh. Any necessary change in the tree construction is revealed within a reasonable timeframe. Therefore the parent never waits for a child for a long time. A predefined waiting period is sufficient and waiting for a long period of time is rather unnecessary because a particular child may not necessarily remains a child forever due to the dynamic nature of the network. This waiting period is a configurable parameter and basically is a function of refreshing frequency of hello messages of a routing protocol.

3.2. Smart Aggregation (SAG)

This section describes the proposed smart aggregation scheme. Under the scheme aggregation only takes place if the sensed values are within a predefined deviation. Let D_{\max} and D_{\min} are the maximum and minimum sensed values respectively. The D_{\max} and D_{\min} are defined among all the values to be aggregated at a particular point. The allowable predefined deviation by the scheme is defined as Δ . According to SAG, aggregation only takes place if and only if $\Delta \geq (D_{\max} - D_{\min})$. An explicit flag in the message define the characteristic of the message *i.e.*, whether the message is aggregate able or not. In case that the flag is set to false the data at the intermediate node

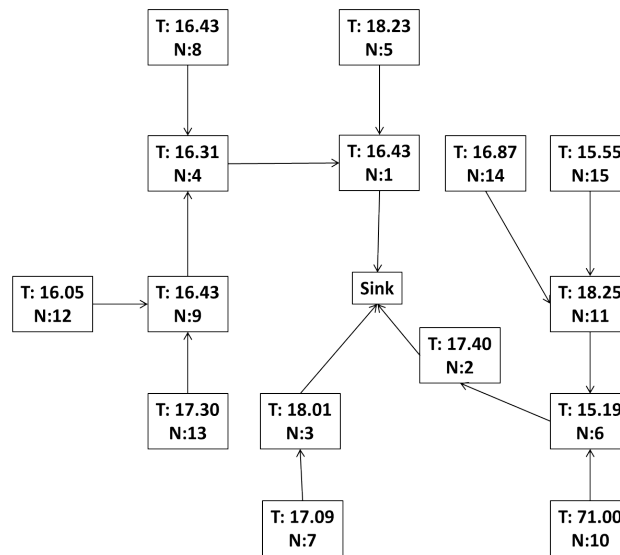


Figure 2. Tree structure for a sensor networks.

is forwarded instantly without any alteration and without any delay. On the other hand when the flag is set to true the same procedure depicted above is applied at the next node and so on until the collection procedure finishes at the sink.

The underlying idea of the technique is to control the accuracy of the monitoring application. The smaller the Δ is the more accurate the monitoring technique is with a cost of additional energy consumption. The rationale behind the proposed scheme is smaller Δ freezes aggregation on a large number of messages contributes to a larger data transition cost.

Node 9 in **Figure 2** has its own sensed value 16.43 and it receives values 16.05 and 17.30 from its two children. If Δ is set to 2, node 9 will aggregate the values as the deviation here is 1.25 (less than 2). On the other hand, node 11 will not perform any aggregation as the deviation at node 11 is 2.7 (greater than 2).

3.3. k -Hop Aggregation

Unlike SAG, k -hop aggregation relies on hop count for aggregation decision rather than the deviation of the sensed values. Therefore aggregation takes place up to the k -hop of transmission and from $k + 1$ hop and on aggregation is suspended. Let such k -hop aggregation technique considering the network depicted in **Figure 2**. Data from node 13 flows through nodes 9-4-1 and finally at the sink. According to this technique, if k is set to 2 then the intermediate nodes 9 and 4 will aggregate the data. On the other hand node 1 will send the data to the sink without any aggregation. Likewise SAG node wait for the child node sensed data for a predefined threshold. As soon as a node collects data from all of its children, it aggregates and forwards the data to the upper layer. Nodes that do not have any child simply send their data as soon as it samples the medium.

The further we go the more deviation in the sensed data is expected. Limiting aggregation with the hop counts may limit the error due to aggregation. But an exact control on error like SAG is unlikely by this approach. Dual frequency approach to closely monitoring the event is preserved in the k -hop aggregation. A special case of k -hop aggregation with single frequency where aggregation takes place only at the first hop is presented in [37].

3.4. Continuous-Monitoring Based on Event Driven Reporting (CM-EDR)

In CM-EDR [38] aggregation only considers locally generated data where sensed values from child node have no impact on aggregation. Here, sensor nodes compare current data with its last transmitted data sent towards the sink. A data packet is only generated when there is a change between the two data.

Let a node sensing the temperature commencing at t_0 with a sampling interval δ . Consider a scenario where the temperature sensed by a given sensor node evolves as follows: at t_0 , $T_{emp}(t_0) = 68^\circ\text{F}$; at $t_0 + \delta$, $T_{emp}(t_0 + \delta) = 68^\circ\text{F}$ and finally at $t_0 + 2\delta$, $T_{emp}(t_0 + 2\delta) = 70^\circ\text{F}$. Considering an event driven monitoring application with a threshold equal to 100°F , no report is transmitted to the sink node. With a classical CM application, the sensor node transmits its data periodically, *i.e.*, three reports are transmitted to the sink node. Enabling the CM-EDR option, the sensor node transmits only two reports, at t_0 and at $t_0 + 2\delta$. As SAG and k -hop aggregation, two-sampling frequency scenario is used with the CM-EDR scheme. **Figure 3** and **Figure 4** portray the protocols and algorithms in pseudo codes and flow diagrams.

3.5. Complementary Techniques

Making the entire WSN functional we need additional techniques in various layers along with the proposed SAG as it is a bare aggregation technique. Efficient data forwarding from source to destination requires a suitable routing protocol. Besides routing we need a MAC suitable for WSN that enables energy-efficient access to the shared channel. Note that energy efficiency and simplicity remains our essential criteria of selecting the routing and MAC protocols, which is aligned to our design goal of the proposed aggregation technique. As such, the following routing and MAC protocols are used: LEPSM [39] and BMAC [40] respectively. Note that these protocols are not part of the proposed compression mechanisms previously described.

4. Proposed Compression Clustering Protocol

In order to further reduce the energy consumption, we incorporate an in-network compression within sensor nodes. WSN are typically densely deployed. As a result nodes in the field typically reside in close proximity [41]. Unlike CM applications, EDR applications' active nodes reside in relatively small region and sensed values


```

Algorithms

if(timer fired)
  sense the environment
  call appropriate aggregation function

SAG

  if(no child for the current node)
    pass sensed value to parent node
  else
    if (data from all childs reaches the node)
      find maximum value  $D_{max}$ 
      find minimum value  $D_{min}$ 
      if ( $\Delta \geq (D_{max} - D_{min})$ )
        where  $\Delta = Threshold$ 
        aggregate and send
      else
        send without aggregation
    else
      wait for the value from the child(s)

k-hop
  if(no child for the current node)
    pass the sensed value to the parent node
  else
    if(hop count is less than k)
      aggregate and send
    else
      send without aggregation
    else
      wait for the values from the child node

CM-EDR

if(current value = previous value)
  pass the value to the parent node

```

Figure 3. Algorithmic description of the aggregation protocols.

from these nodes are anticipated to be having small variations. Based on the underlying idea we propose a clustering technique exploiting the spatial-correlation of event-related data. The goal is to reduce the data packet size. To be more specific we use a cluster-based architecture where the cluster-members (C-Ms) do not send the sensed values directly. C-Ms rather send the differences between a reference data and their sensed values of that instance. Note that this reference value is defined by the sensed value of the corresponding cluster head (C-H).

4.1. Data Sensing Model

This proposition utilizes the diffusion property while modeling spatially correlated data [42] rather than real data as processed in the SAG protocol. Generally the value at any point with a distance d from the event center can

be defined as $D = \frac{C}{(d+1)^\alpha}$. Here C is a constant and represents the value at the event center. α denotes at the

diffusion parameter. Note that α is based on the particular phenomenon and environment under surveillance. For example, for heat $\alpha = 1$ and for light $\alpha = 2$. We take reference [43] notion while modeling our data in evaluation. Therefore, a node sensing data at a distance d of the center of an event while sensing temperature would see a value given by $D = \frac{1}{d}$.

4.2. Energy Consumption Model

Modeling of the energy consumption is carried out utilizing Henizelman's energy consumption model [2].

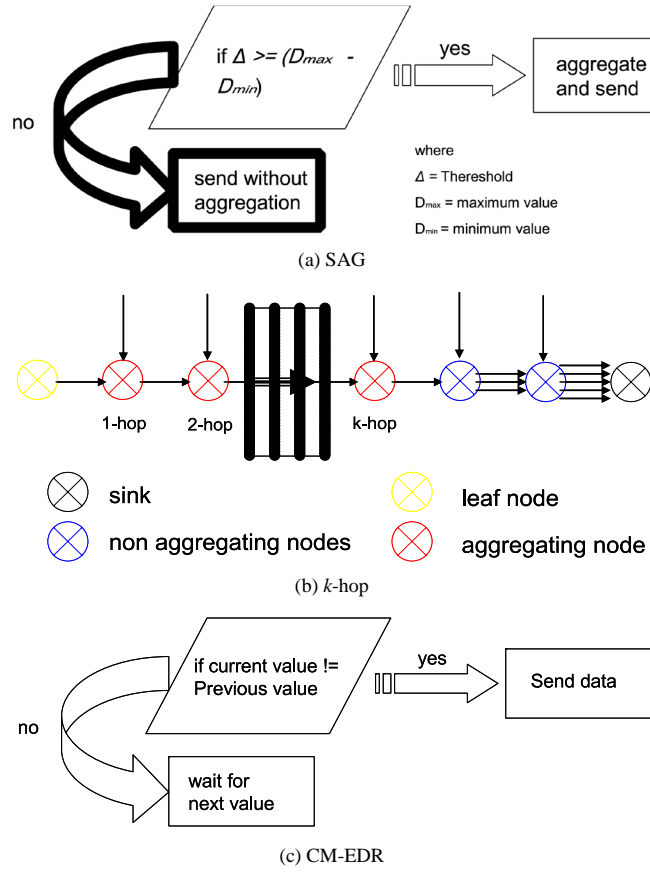


Figure 4. Illustration of the aggregation techniques.

Where the transmit energy consumption is defined as

$$E_{tx}(sz, d) = \begin{cases} sz \times E_{elec} + sz \times d^2, & \text{if } (d \leq d_0) \\ sz \times E_{elec} + sz \times E_{mp} \times d^4, & \text{otherwise} \end{cases} \quad (1)$$

Here sz denotes the size of the data packet. E_{elec} denotes the energy consumption of Tx/Rx circuitry and E_{mp} denotes the amplifier power consumption. $d_0 = \sqrt{\frac{E_{fs}}{E_{mp}}}$ defines the distance threshold of Tx to Rx. If the distance is beyond the threshold a multipath fading channel model is utilized. At Rx the energy consumption is defined as $E_{rx}(sz) = sz \times E_{elec}$. The aggregation energy consumption of the data is denoted as E_{DA} .

4.3. Initialization

The initialization process for the classical compression approach and CC SCR are described as follows: The proposed clustering scheme CC SCR consists of two phases like the classical approaches. The phases are 1) set up phase and 2) steady state phase though the details of the phases are different. In classical approaches every time an event occurs the set up phase goes on. Contrarily, in CC SCR, all the active nodes transfer the first sensed value to the BS therefore they are not confined to transmitting the control packets only. The underlying reasoning behind transmitting the sensed value for the first time is to utilize the sensed value in selecting C-H. We acknowledge the extra consumption of energy at the set up phase while comparing with the classical approach but argue that the procedure provides superior selection of C-H and therefore saves more energy at the steady state phase which is often reasonably much longer a period.

As a result, the CC SCR protocol is best suited for scenarios where the event conditions are fairly stable, *i.e.*,

where once an event occurs, the sensed parameter has no high variations. The rationale behind is the C-H is created based on the first sensed data at the setup phase. Therefore if the event conditions varies drastically during the event duration the originally selected CH may no longer provide desirable energy saving.

For example an application for forest fire surveillance, in case of an event of a fire the temperature remains stationary during the event of fire at that particular region. Target tracking applications have also similar characteristics as the targets are the source of the measured data (light/temperature) of the nodes. During the presence of the target at the same place the observed data likely to remains the same.

4.4. Set Up Phase and Steady State Phase

During the set up phase, the BS receives first data packets from all the active nodes. It then calculate the differences of the received values of node n_i and node n_j (where, $(i \neq j, \text{ and } i \leq N, j \leq N)$). The differences are summed over and denoted as S_i . BS therefore determines the C-H. The C-H is the node that provides the minimum value of S_i . After selecting the C-H the BS then broadcast the control message and notifies all the cluster members about the C-H selection. Nodes not selected as BS are known as cluster members (C-Ms). The cluster members do not need to send any more control messages as the BS knows the information about the active nodes.

4.5. Steady State Phase

After the set up phase, the compressed values of the data packets are sent by all the C-Ms. Here, they do not send the sensed value $value_CM_i$ they rather send the difference between the sensed data value and the data value of the corresponding C-H. Let the compressed value denoted as Δ_i . The i -th C-Ms data value denoted as $value_CM_i$ and the corresponding C-H data value denoted as $value_CH$. Therefore, $\Delta_i = |value_CM_i - value_CH|$. Note that, at the beginning of each occurrence, the C-H sends the complete set of the sample data values of all the C-Ms and based on the information the compression becomes achievable. As a result only Δ_i from the cluster members is sufficient to be delivered to the C-Hs. As the BS is not an energy constrained device and the calculation of S_i is centralized the low power C-Ms are not involved in expensive calculations is a strength of the proposed algorithm.

We illustrate the proposed protocol operation and our classical counterpart in **Figure 5(a)** and **Figure 5(b)** respectively. In these figures we illustrate the case of five active nodes in the event region e and the coverage of

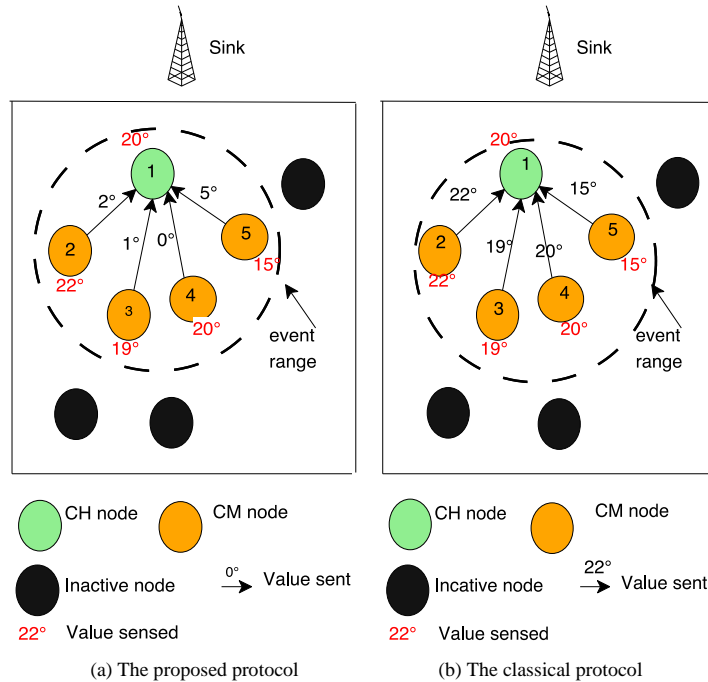


Figure 5. Example of the system operation.

region range is R_{event} . In the case of the proposed protocol, nodes transmit their values to the BS. Upon receiving the values from the nodes the BS then calculates S_i for each n_i . Therefore BS may select either node n_1 or node n_4 as a C-H as both nodes minimize the total difference value measured. Other nodes in turn become C-Ms. And during the steady state phase the all the C-Ms transit only the Δ_i value to the C-H rather than the complete temperature values. As a result a compressed value with less coded bits compared to the complete data in the classical clustering scheme. With is reduction on transmitting bits energy saving is achieved in a great extent demonstrated in simulation results.

It is important to notice that we assumed that nodes perform the compression instantaneously. Although this may not be accurate in the sense that in practice, it may take some time to perform the compression algorithm, it is difficult to compare to other similar protocols since LEACH for instance do not consider this time neither. Additionally, it is not straight-forward to calculate this time theoretically since, in many cases, commercial nodes do not specify the time to perform each instruction. In the future, we plan to implement these protocols in a real environment in order to calculate these factors from a real network. In any case, we believe that the compression time in the proposed protocols should not take considerably more time than typical aggregation mechanisms since we are basically comparing two values in order to compress data.

5. Proposed Event Reporting and Continuous Monitoring Protocols

In this section, the hybrid protocols used for both CM and EDR applications are described. The main idea of both proposals is to take advantage of the clustered architecture established for the CM transmissions in order to also transmit data events. By doing so, there is no need to form different clusters for the event reporting as it is done in [36]. As it is shown in [36], the extra energy consumed in order to form *event clusters* (clusters specifically formed in order to transmit event data reporting) can lead to an excessive energy drain especially when events occur with a high rate even if data compression is enabled.

By using the already formed clusters for both the CM for the transmission of event driven data transmission, the network's lifetime is extended as opposed to using a separate protocol for each application, for instance, LEACH for CM data and TEEN for EDR. Indeed, the event reporting of the nodes inside the event area by means of the NP-CSMA random access protocol, avoids the extra energy consumption drained for separate clusters or multi hop transmissions in case of a non-clustered architecture. Note that these proposals are well suited for highly correlated environments where many nodes inside the clusters detect the events. As such, it is expected that the reporting data has a high amount of redundancy that can be reduced by means of aggregation or compression techniques such as the ones previously discussed. For reasons of simplicity, the LEACH protocol is considered as the protocol for the CM data transmission. For these results, the NP-CSMA technique is used at the cluster formation phase.

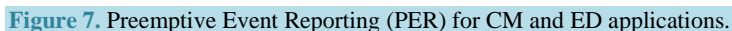
5.1. Network Model

The following assumptions and system parameters are considered: The total number of sensor nodes in the system is $N = 100$. Sensor nodes are uniformly distributed in an area between (0, 0) and (100, 100) meters (*i.e.*, square 100×100 area). The sink node is situated outside the supervised area at the coordinate (200, 0). Hence, transmissions to the sink node, which is situated at least 100 meters from the closest node in the supervised area, are highly energy consuming. Each C-H uses a distinct code to transmit the gathered data to the sink node using a CDMA technique. As such, no collisions among C-Hs are possible. All sensor nodes have the same amount of initial energy. For the CM application, all sensor nodes in the supervised area transmit f_1 data packets per second using the LEACH protocol. The event can be sensed by all the sensors inside the sensing range. Events are supposed to occur uniformly distributed in the supervised area. All nodes can transmit with enough power to reach directly the sink node. Additionally, nodes can use power control to vary the amount of transmit power. The size of the data packet (2 kbits) comprises the data payload, the identification field, Id , and a *type* field to specify the type of packet: control packet and CM data packet. The size of the control packet is considered to be of 1 kbits. The energy consumed follows the model explained in section 4.2. Each C-H dissipates energy in receiving and transmitting the signals received from the C-Ms. Unlike LEACH [2], the C-Hs do not perform ideal data aggregation. The steady state phase is considered to be of 20 seconds. A Uniform back off protocol with window size of $w = 15$ time slots is considered. The rest of the parameters are listed in **Table 1**.

Figure 6 presents the basic operation of NPER. First, the clusters are formed using the LEACH protocol, depicted by the Cluster Formation (CF) interval. Then, the CM data is transmitted using the TDMA structure, depicted by the CM interval. Then, the C-Hs transmit this information to the sink node using a CDMA based scheme in order to avoid collisions among C-Hs. At this point, the nodes inside the event area transmit their information to the appropriate C-Hs using the random access (RA) protocol, and it is denoted as the ED period. At the end of this period, the C-Hs transmit this information to the sink node again using a specific code avoiding collisions among C-Hs. This transmission to the sink node is not shown in **Figure 6** since there are not always such transmissions. Indeed, there are such transmissions only when an event is detected. Therefore, most of the time there are no event-related transmissions to the sink node. This procedure (TDMA/CDMA/RA) is repeated until the end of the steady state phase. At this point, clusters are formatted again in order to rotate the C-Hs in the system. Note that, the clusters used for the CM data transmission are also used for the event reporting data. Also, it is important to notice that in the RA period, not all nodes inside the cluster transmit event-related information but only the nodes that detect an event transmit in the RA period. Hence, the collision probability is not expected to be high.

As in the previous strategy, the clusters formed by the LEACH protocol are also used for the event reporting. However, the difference consists basically in that the event reporting is performed at each TDMA slot in the steady phase. Specifically, at the beginning of each TDMA slot, the nodes inside the event area transmit their data to their respective C-Hs using the RA protocol. Then the nodes in their CM activities transmit their packet in a collision-free manner. By doing this, the nodes reporting the event data do not have to wait until the end of the CM period. However, if there are many nodes that sensed the event in a given cluster, the collision probability may be high; in which case, the nodes have to retransmit in further time slots. This is explained in [Figure 7](#). In this protocol, CM data is lost when an event is being reported. The rationale behind this is to reduce delay for

Parameter	Value
E_{fs}	10 pJ/bit/m ²
E_{elec}	50 nJ/bit
Idle power	13.5 mW
Sleep power	15 μ W
Initial energy per node	2 J
Transmission bit rate	40 kbps



the event related data. Also, the loss of information for the CM application may be acceptable in case of emergency, typical of event occurrence.

6. Simulation Results

We evaluate the performance of the proposed SAG along with its modified version k -hop and CM-EDR. Tiny OS based network simulator TOSSIM [44] is taken as our platform. Mica2 is chosen as node therefore the available memory is 8 kb RAM (program and data) and 128 kb flash (external) respectively. We take real data from [45] for both sensed and network connectivity and feed into our simulation. For all the concerned three aggregation techniques BMAC and LEPSM are taken as medium access control and routing protocol respectively. Default setting is used for both BMAC and LEPSM. In these setting both transmit mode (preamble length) and listening mode (check interval) is set to 100 ms also known as mode 4. Route update is set 10 seconds.

Energy consumption with the interested aggregation algorithms using different configurations and node distributions is evaluated. Figure 8 and Figure 9 show the average energy consumptions of CM-EDR, k -hop and SAG protocols over the time (1200 - 4000 seconds) with 40-node deployments with sampling frequency f_1 only and both f_1 and f_2 respectively. Similarly Figure 10 and Figure 11 show the average energy consumption with three aforementioned algorithms using different number of node (10 - 50) deployments after 36,000 seconds

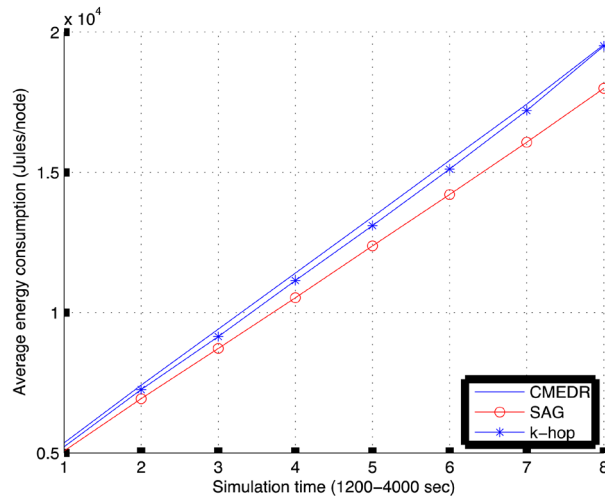


Figure 8. Energy consumption over simulation runtime with f_1 .

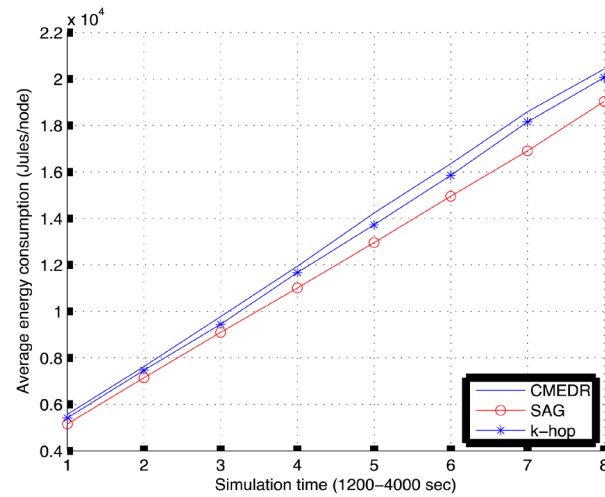


Figure 9. Energy consumption over simulation runtime with f_1 and f_2 .

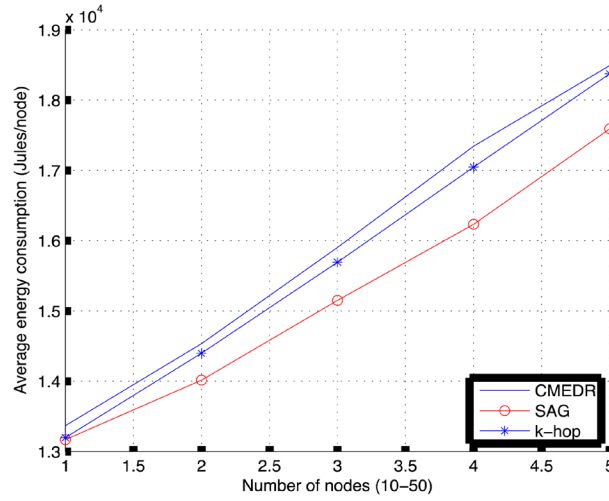


Figure 10. Energy consumption over number of node deployments with f_1 .

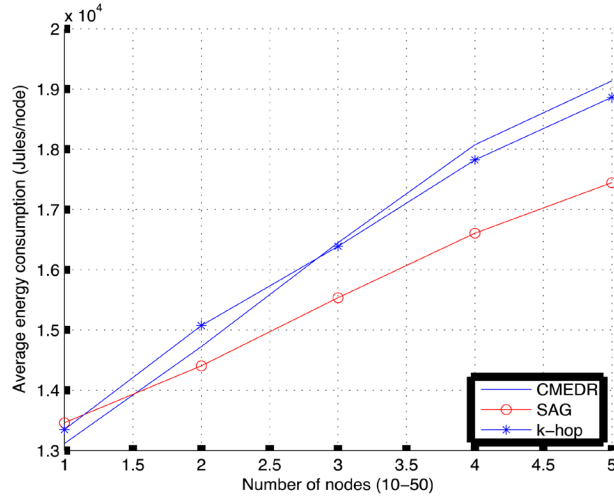


Figure 11. Energy consumption over number of node deployments with f_1 and f_2 .

with f_1 only and both f_1 and f_2 respectively. For k -hop and SAG aggregation approaches, k and \mathcal{A} are set to one hop and one degree respectively. Sampling frequency f_1 and f_2 are set to one and two sample/minute respectively. An event is defined as the temperature more than 70 degree. The superiority of the proposed SAG is obvious and demonstrated by the minimum average energy consumption in joules depicted in the figures.

Supremacy of the k -hop and CM-EDR over each other is undecided. This is due to the fact that k -hop exploits spatial correlation and CM-EDR exploits the temporal correlation of sensed data. Though in **Figure 8** and **Figure 9** k -hop performs better compared to CM-EDR this performance is specific to the data itself. With different sets of data the performance would be different. Evidently k -hop performance is independent of the temporal correlation on the other hand CM-EDR performance is independent of the spatial correlation of the data. In a scenario where there exists more spatial correlations compared to the temporal correlations among the sensed data k -hop aggregation protocol may outperform the performance of CM-EDR aggregation technique.

The argument is backed by the result depicted in **Figure 12** where the performance of k -hop and CM-EDR keep crossing each other. The best performance of the SAG aggregation achieves due to the fact that it exploits both spatial and temporal correlations.

The impact of the aggregation and sleep strategy on the proposed SAG is depicted by **Figure 13**. Employing

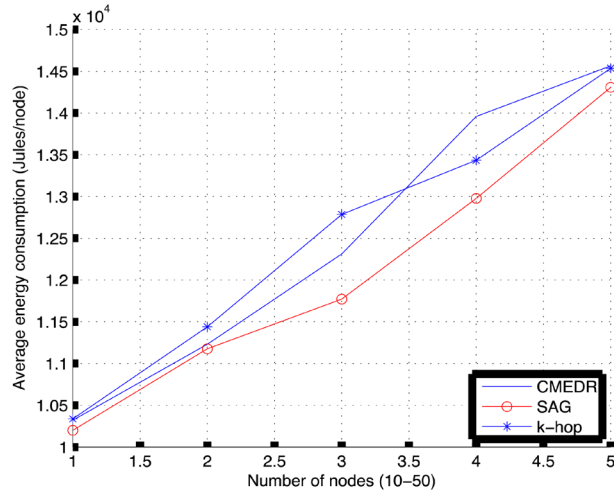


Figure 12. Impact of aggregation over the time. Supremacy of k -hop and CM-EDR depends on data.

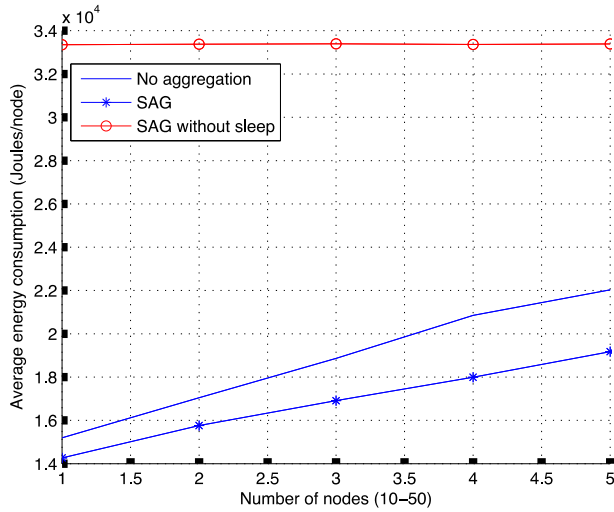


Figure 13. Impact of aggregation and sleep over SAG.

BMAC in conjunction with the protocols in terms of energy saving is huge thereby justifies the deployment of the specific medium access control mechanism in this scenario. Energy saving due to the proposed aggregation function is also significant as depicted in the figure.

Impact of tuning parameters k and Δ on the performance are worthy to investigate. **Figure 14** shows the impact of k on the energy consumption on the k -hop protocol. The bigger the k is the more energy it is saved utilizing the protocol. On the other hand **Figure 15** shows the impact of Δ over the SAG protocol the greater the deviation is allowed the lesser the energy is used. Intuitively in both the cases the energy efficiency is achieved on the expense of losing data accuracy. There is no bound on the error in k -hop aggregation technique. On the contrary SAG error will be limited as Δ is predefined where no aggregation takes place for highly deviated samples.

In order to quantify the benefits of the CC_SCR protocol, it is compared with a classical clustering protocol, in this case LEACH. Recall that a classical clustering protocol is composed of two different phases: 1) setup phase and 2) steady state phase. In the case that an event occurs at a random location in the sensor network, the sensors in the vicinity wake up. At the same time the clustering process initiates. At this stage the nodes in the event region competes among each other to become a C-H. In this competition the active nodes through the random medium access protocol, send the control packets towards the BS. We consider the carrier sense multiple

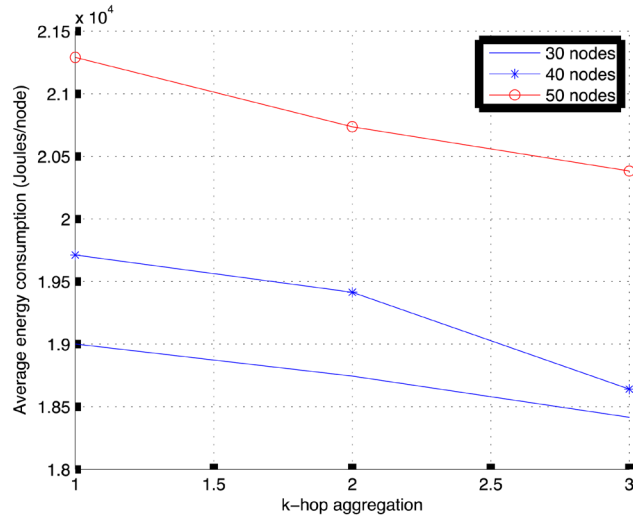


Figure 14. Impact of k on k -hop aggregation (for $k = 1, 2, 3$).

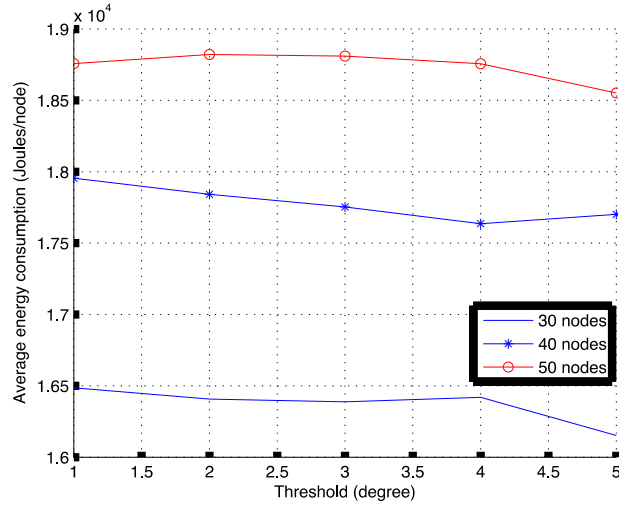


Figure 15. Impact of threshold on SAG aggregation (for $\Delta = 1 - 5$).

access (CSMA) as the access control mechanism in our evaluation.

Note that the control packet does not perform any data compression since it only carries the node ID at this stage and the first node that able to successfully transmit the packet becomes the C-H and report to all the active nodes using an explicit messaging over the network. Upon reception of this C-H selection notification non-C-H nodes become C-Ms. During the steady state phase C-Ms send the sensed values to the corresponding C-Hs utilizing time division multiple access (TDMA) protocol, one of the most popular scheduled access control mechanism. Upon receiving all the data from all the C-Ms, C-H aggregates the received data with its own data and transmits the resultant data to the BS. As SAG, CC_SCR uses TOSSIM simulator. The parameters used for this set of results are presented in Table 2.

Figure 16 presents the average energy consumption in the network per unit of time for different number of nodes in the range of 10 to 70. Twenty events are simulated and the average energy consumption is shown in the figure. The results clearly demonstrate the superiority of the proposed CC_SCR over the classical clustering approach in terms of energy consumption in unit time.

Increasing the number of nodes, in the system increases the energy consumption. This is due to the fact that increasing the number of nodes in the network results increases the node density in the network. Hence, in case of an event, the number of nodes detecting the event also increases. With a higher number of nodes detecting

events, increases the number of both control and data packets in the network impacting on both the CC_SCR and classical approaches. Though the increment in the classical approach is much higher compared to the proposed CC_SCR as the slope of classical approach is higher. The energy saving comes from the fact that the proposed algorithm does not transmit the sensed values; rather the difference Δ_i is transmitted during the steady state phase.

For the high-density network scenario the difference of the consumption increases between the classical and CC_SCR approach. This is due to the fact that increasing the node density reduces distances of the nodes. As such, it increases the correlation degree among the sensed data. And finally increasing correlation allows a higher compression rate.

With classical schemes where no compression takes place, a higher node density increases the number of packets to be sent to the destination. And such transmissions are quite costly in terms of energy consumption.

Figure 17 attempts to explore further into the single hop scenario where all the nodes are directly connected to the BS. The scenario is designed in order to explore the advantages of the cluster-based network architecture. In case there is no C-H, a reference node is chosen to transmit data. BS receives all the initial data packets from the active nodes and upon reception it carries out computation to identify the node that minimizes the data differences (explained in the previous section). Upon deciding on the reference node BS then notify the nodes by transmitting control packets using the node ID of the reference node.

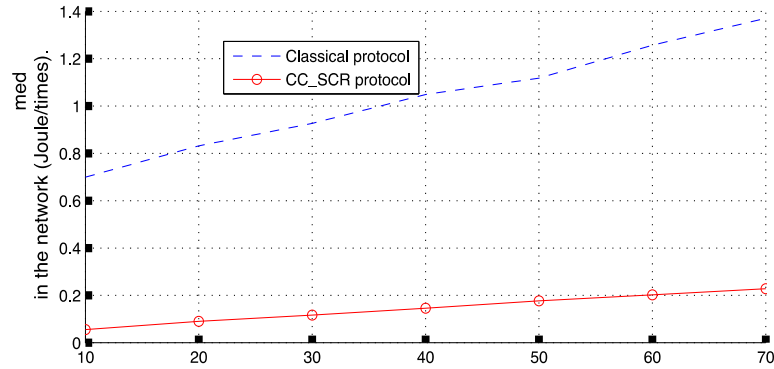


Figure 16. Average energy consumed by unit of time vs number of nodes (classical vs CC_SCR).

Table 2. Simulation parameters.

Parameter	Value
E_{fs}	10 pJ/bit/m ²
E_{np}	0.0013 nJ/bit/m ⁴
E_{elec}	50 nJ/bit
E_{DA}	5 nJ/bit/signal
Signaling packet length S	24 bit
Data value at the center of the event C	250°
Initial energy per node E_0	10 J
T_{event}	200 s
R_{event}	60 m
R_c	100 m
R_t	400 m
Area A	100 × 100 m ²

Figure 17 depicts the average energy consumption for 60 nodes over 20 rounds. Results clearly show that the proposed CC_SCR outperforms the classical approach by conserving a significant amount of energy. This is due to the fact the choice of the clustering techniques provides improvement on energy performance over the single hop counterpart.

We further provide the gain ratio of the proposed algorithm over classical and single hop scenarios in **Figure 18**. The gain ratio may reach up to 11 times and 119 times for the classical and single hop scenarios respectively. The aforementioned gain ratios therefore provide the benefits of the proposal in this aspect. The average energy consumption for different R_{event} regions is depicted in **Figure 19**. When R_{event} varies, the active node

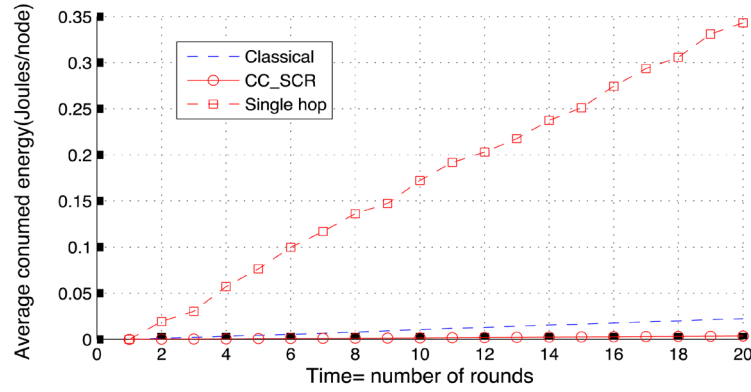


Figure 17. Average energy consumed over time (classical vs CC_SCR).

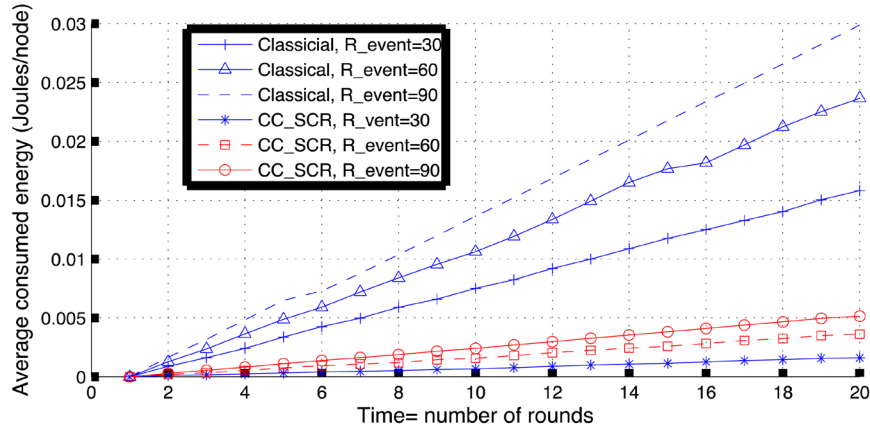


Figure 18. The ratio gain of the energy for each event occurrence (classical vs CC_SCR).

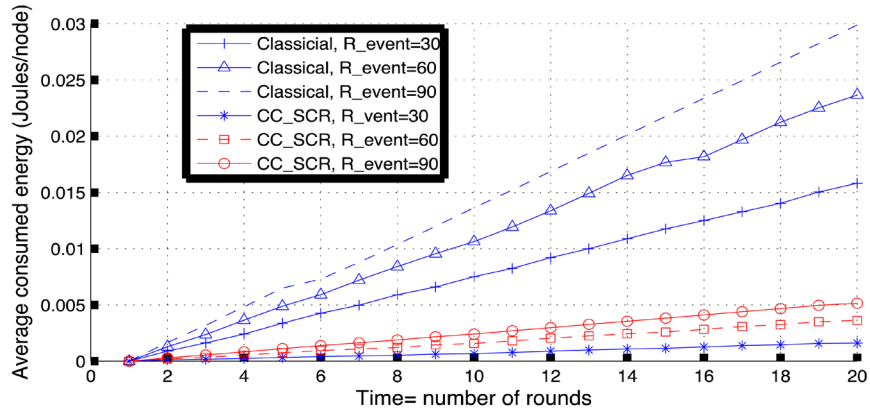


Figure 19. Average energy consumed with varying the R_{event} region (classical vs CC_SCR).

numbers per event varies accordingly.

The active number of nodes per event is depicted in **Figure 20**. In both classical and proposed scheme the number of active nodes is approximately the same. The proposed scheme therefore has no impact on the number of active nodes. And with the increase of the active number of nodes the energy consumption also increases. Observe that the energy consumption of the classical approach is far too high. For instance, the energy consumption of the classical approach when $R_{event} = 30$ is much higher than the energy consumption of the proposed scheme when $R_{event} = 90$. And using the proposed approach can easily extend the lifetime of the resource constrained sensor network significantly.

The average energy consumption for different values of the T_{event} period is depicted in **Figure 21**. Increasing T_{event} increases the steady state period. As a result the number of packets for the data reporting increases. These extra packets in turn increase the power consumption on the network.

It is worthy to note that for the high event duration time the proposed algorithm is more efficient compared to the low event duration time. With the increasing of the reporting duration the C-Ms eventually generate much more data for both the classical and the proposed approaches. But with enabling compression the proposed scheme send data with much smaller packet sizes and saves energy. On the other hand the classical approach transmits the full length packets and dissipates a large amount of energy.

The impact of the energy consumption due to the aggregation is depicted in **Figure 22**. Numerical results clearly demonstrate the performance improvement of the energy consumption due to the aggregation procedure. By enabling aggregation, the C-H only transmits a single packet to the BS no matter how many C-Ms are forwarding the data to the C-Hs. On the other hand the aggregation scheme is not enabled the C-Hs need to forward each packet from all the C-Ms to the BS.

We now evaluate the performance of the hybrid proposals PER and NPER for event reporting in CM WSNs.

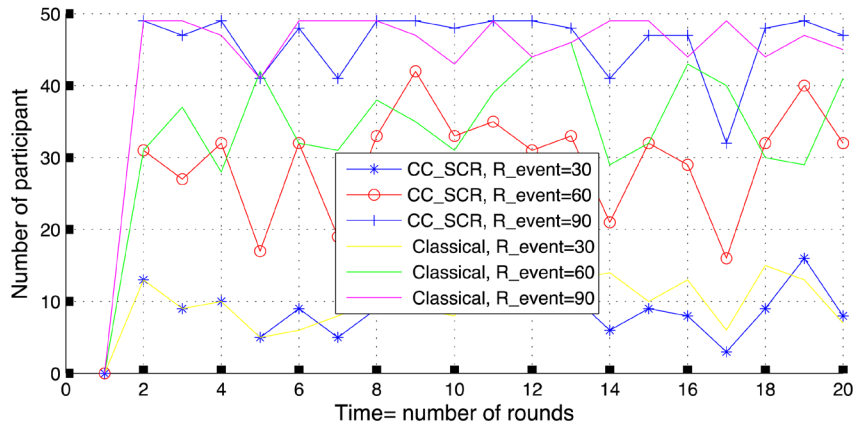


Figure 20. Number of active nodes for each event occurrence (classical vs CC_SCR).

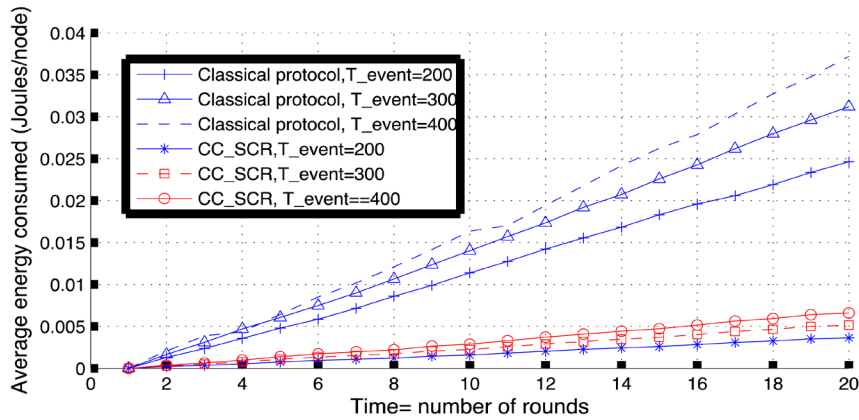


Figure 21. Average energy consumption vs T_{event} (classical vs CC_SCR).

We study the gain introduced by using the different proposals compared to LEACH and TEEN for similar event rates in order to have a fair comparison. Since LEACH is mainly used for CM applications and TEEN is mainly used for EDR, it is shown in this section that the proposed strategies outperform these protocols when both CM and event reporting are present in the same network. For all results presented in this section, the WSNs system was implemented as a discrete event simulation. Also, it is considered that events occur uniformly in the complete supervised area. Hence, all nodes in the network have the same probability of detecting a given event. It is assumed that there are 100 nodes in the network. The following results are obtained using our own discrete event simulator in C++. We programmed the LEACH, TEEN, PER and NPER protocols in order to compare their performance under the settings described in Section 5.2.

First, the system performance is studied for a high event rate generation in the system. **Figure 23** shows the network lifetime, which is an indication of the energy consumption in the network for an event rate of 22.5 events per second in the supervised area and different event area radius. As mentioned earlier, the proposed strategies are compared to LEACH for the transmission of CM and event reporting data. Specifically, LEACH allows the formation of 5 clusters (5% of the nodes becomes C-Hs in each cluster formation phase). Since there are 100 nodes in the network, then, there are in average 19 cluster members per cluster. Hence, considering the event rate of 22.5, it can be considered that in average, each node in the cluster reports an event. As such, there are 40 transmissions per cluster. Then, the C-H schedules 45 time slots in the TDMA structure in order to accommodate all possible transmissions. Conversely, since TEEN does not have a structure for the CM transmission, and in order to have a fair comparison in terms of the number of transmissions per second in the system, it is considered a transmission rate of 45 events per second (20 transmissions correspond to the CM information and 25 transmissions correspond to the event reporting). As for the proposed strategies, there are also 20 C-M transmissions and 22.5 event reporting transmissions in the respective transmissions: CM and EDR respectively. From these results, it can be seen that LEACH and TEEN have a very similar behavior. In this case, the ED periods of the proposed mechanisms experience a high collision probability, which leads to a higher energy drain. Note that LEACH has a very good behavior since all transmissions are well scheduled and there are no retrans

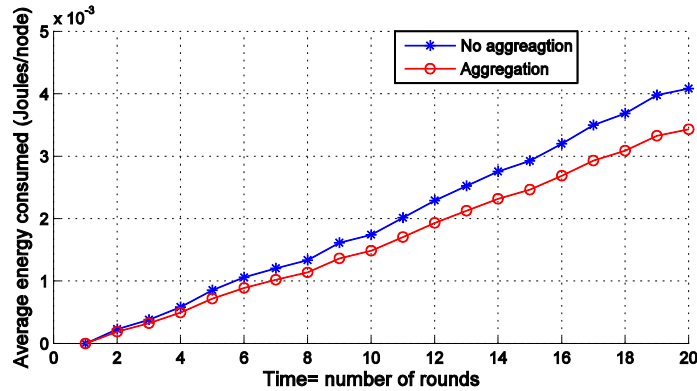


Figure 22. Energy consumed with aggregation and no aggregation (classical vs CC_SCR).

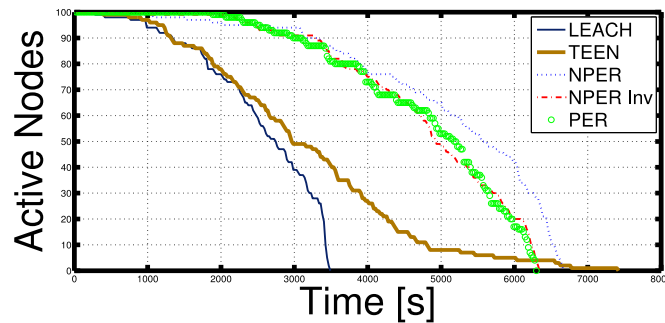


Figure 23. Network lifetime for high event generation rate.

missions due to collisions. TEEN on the other hand consumes a high amount of energy at the beginning of the system operation but as time goes by, there are less nodes colliding among each other, which extends the life of the protocol at the end of the simulation.

Now, the system is studied for different events area radius as shown in **Figure 24**. It is clear that when the event radius is small; the energy consumption is low, increasing the network's lifetime. As the radius increases, also the number of reporting nodes increases and the network's lifetime decreases. It is important to note that Non-Preemptive Event Reporting (NPER) always achieves higher lifetime for any event radius compared to LEACH, with is not affected by the event radius since clusters are formed independently of the coverage of the event. Preemptive Event Reporting (PER) also achieves acceptable results but is only higher than LEACH for $R = 5$ and $R = 10$. The reason for this is that in NEPR there is a contention for the use of the common channel only among nodes reporting the event while in proposal PER there can be collisions among event reporting data and CM data. On a more detail analysis, **Figure 25** presents the number of collisions during the simulations. It can be seen that there are much more collision in proposal PER than in proposal NPER.

However, PER entails a much lower transmission delay for the vent data since nodes do not have to wait until the end of the CM data. This can be seen in **Figure 26**. As such, a network operator can use PER if the event reporting requires a very low average transmission time at the cost of higher energy consumption. Also note that, the NPER and PER protocols have the same delay as the LEACH protocol for the CM since it is used as the basis of this work. On the other hand, the event reporting delay is also bounded by the steady state phase of LEACH. Finally, **Figure 27** presents the number of detected events and transmitted events for both proposals. Note that in PER, the number of detected events is almost the same as the number of reports transmitted since nodes that detect the event contend in different occasions during the CM transmissions increasing the successful transmission rate. On the other hand, in NPER, nodes that detect an event contend among each other at the end of the CM transmission, which leads to a more restricted time to report the events diminishing the number of event transmissions compared to the number of detected events. However, this difference is not significant.

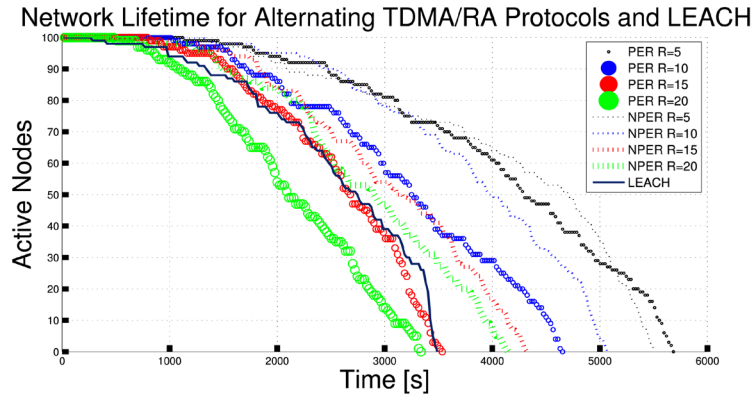


Figure 24. Network lifetime for different event radius.

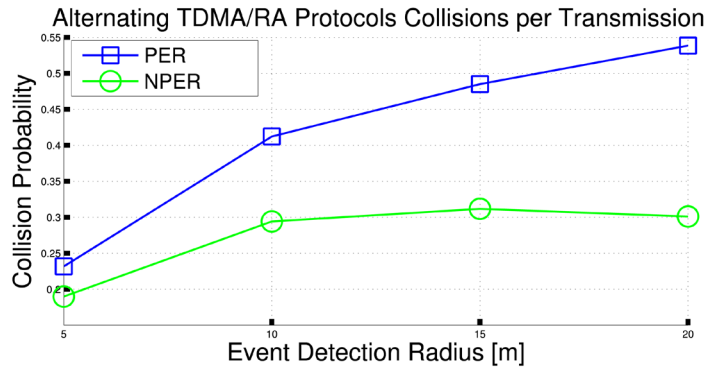


Figure 25. Collisions in the system for different event radius.

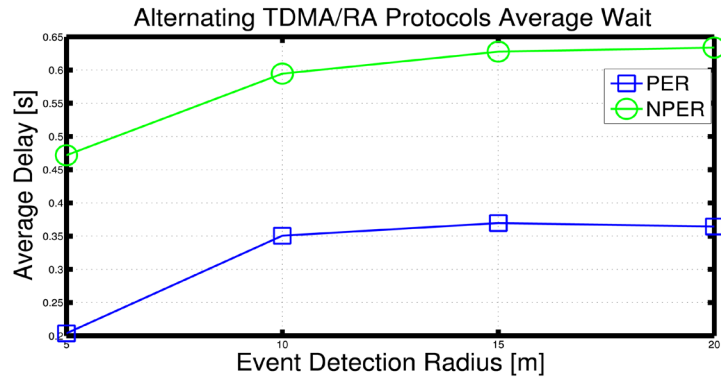


Figure 26. Average waiting time for different event radius.

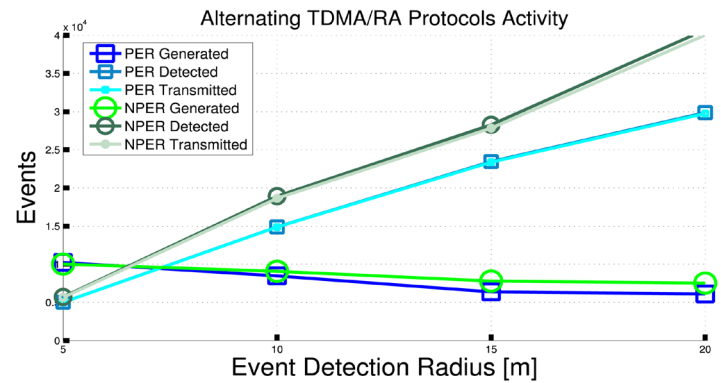


Figure 27. Number of detected and transmitted events for different event radius.

7. Conclusion

Prolonged network lifetime is one of the most focused and challenging research issues in WSNs. It requires that any design and development efforts must follow the energy conservation criteria in sensor networks. Power conservation techniques must be implemented in such way that the reliability and dependability of the system preserve. Every aspect of the protocol and system must provide a desired quality. To this end, this paper presents three distinct compression algorithms: SAG, k -hop and CC_SCR. The SAG protocol is a simple aggregation technique that maintains a bounded error while utilizing spatio-temporal correlation of data. It gracefully provides a balance between the energy conservation and error present in the aggregated data with trivial processing. In the context of WSN this paper further presents a compression clustering technique, named CC_SCR in the paradigm of event-driven applications. While the SAG protocol focuses on the spatio-temporal correlation, the CC_SCR rather focuses on the spatial correlation of data present in event driven applications. As such, each of these compression protocols is aimed at different conditions and system's parameters. The presented aggregation technique is compared with two other aggregation techniques named k -hop aggregation and CM-EDR algorithms. The potential improvement gained by using our proposal is demonstrated by the simulation results presented. As an additional feature of this work, a hybrid WSN with both CM and event reporting capabilities is considered, namely the PER and NPER schemes are proposed to this end. While CM is performed using the well-known LEACH protocol, and event reporting is achieved through the assignment of special time periods where event reporting is based on NP-CSMA with uniform backoff. The benefits of such assignment have been studied in the context of highly dense WSNs. As such, the PER and NPER schemes are focused on reducing the number of control packets required to transmit both types of data: CM and EDR. By observing the numerical results, it can be seen that either reducing the number of packets (through the PER and NPER protocols) or reducing the packet size in the network (through SAG or CC_SCR protocols) is an efficient strategy to reduce energy consumption in WSNs.

Acknowledgements

This research was partially funded by the Government of Abu Dhabi, United Arab Emirates through its sponsorship of Masdar Institute of Science and Technology's research project on "Wireless Sensor Data Analytics".

References

- [1] Al-Karaki, J.N. and Kamal, A.E. (2005) Routing Techniques in Wireless Sensor Networks: A Survey. *IEEE Wireless Communications*, **11**, 6-28. <http://dx.doi.org/10.1109/MWC.2004.1368893>
- [2] Heinzelman, W., Chandrakasan, A. and Balakrishnan, H. (2002) An Application-Specific Protocol Architecture for Wireless Microsensor Networks. *IEEE Transactions on Wireless Communications*, **1**, 660-670. <http://dx.doi.org/10.1109/TWC.2002.804190>
- [3] Manjeshwar, A. and Agarwal, D.P. (2001) TEEN: A Routing Protocol for Enhanced Efficiency in Wireless Sensor Networks. *IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, San Francisco, 23-27 April 2000, 2009-2015.
- [4] Manjeshwar, A. and Agarwal, D.P. (2002) APTEEN: A Hybrid Protocol for Efficient Routing and Comprehensive Information Retrieval in Wireless Sensor Networks. *IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, Miami, 15-19 April 2001, 195-202. <http://dx.doi.org/10.1109/IPDPS.2002.1016600>
- [5] Manjeshwar, A., Zeng, Q. and Agrawal, D.P. (2002) An Analytical Model for Information Retrieval in Wireless Sensor Networks Using Enhanced APTEEN Protocol. *IEEE Transactions on Parallel and Distributed Systems*, **13**, 1290-1302. <http://dx.doi.org/10.1109/TPDS.2002.1158266>
- [6] Polastre, J., Szewczyk, R. and Culler, D. (2005) Telos: Enabling Ultra-Low Power Wireless Research. *ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, Los Angeles, 25-27 April 2005, 364-369.
- [7] Demirkol, I., Ersoy, C. and Alagoz, F. (2006) MAC Protocols for Wireless Sensor Networks: A Survey. *IEEE Communications Magazine*, **44**, 115-121. <http://dx.doi.org/10.1109/MCOM.2006.1632658>
- [8] Azim, M.A., Moad, S. and Bouabdallah, N. (2010) SAG: Smart Aggregation Technique for Continuous-Monitoring in Wireless Sensor Networks. *IEEE International Conference on Communications (ICC)*, Cape Town, 23-27 May 2010, 1-6.
- [9] Moad, S., Rivero, M., Bouabdallah, N. and Langar, R. (2011) *CC_SCR*: A Compression Cluster-Based Scheme in a Spatial Correlated Region for Wireless Sensor Networks. *IEEE International Conference on Communications (ICC)*, Kyoto, 5-9 June 2011, 1-6.
- [10] Moad, S., Rivero-Angeles, M.E., Bouabdallah, N., Sericola, B. and Aoul, Y.H. (2012) Performance Analysis of a Compression Scheme for Highly Dense Cluster-Based Wireless Sensor Network. In: Matin, M.A., Ed., *Wireless Sensor Network—Technology and Applications*, InTech Press, Rijeka, Croatia, 207-230. <http://dx.doi.org/10.5772/48370>
- [11] Fasolo, E., Rossi, M., Widmer, J. and Zorzi, M. (2007) In-Network Aggregation Techniques for Wireless Sensor Networks: A Survey. *IEEE Wireless Communications*, **14**, 70-87. <http://dx.doi.org/10.1109/MWC.2007.358967>
- [12] Rajagopalan, R. and Varshney, P.K. (2006) Data-Aggregation Techniques in Sensor Networks: A Survey. *IEEE Communications Surveys and Tutorials*, **8**, 48-63. <http://dx.doi.org/10.1109/COMST.2006.283821>
- [13] Intanagonwiwat, C., Govindan, R., Estrin, D., Heidemann, J., Silva, F., et al. (2003) Directed Diffusion for Wireless Sensor Networking. *IEEE/ACM Transactions on Networking*, **11**, 2-16. <http://dx.doi.org/10.1109/TNET.2002.808417>
- [14] Madden, S.R., Franklin, M.J., Hellerstein, J.M. and Hong, W. (2005) TinyDB: An Acquisitional Query Processing System for Sensor Networks. *ACM Transactions on Database Systems*, **30**, 122-173. <http://dx.doi.org/10.1145/1061318.1061322>
- [15] Madden, S., Franklin, M.J., Hellerstein, J.M. and Hong, W. (2002) TAG: A Tiny AGgregation Service for Ad-Hoc Sensor Networks. *Proceedings of the 5th Symposium on Operating Systems Design and Implementation*, Boston, 9-11 December 2002, 131-146.
- [16] Motegi, S., Yoshihara, K. and Horiuchi, H. (2006) DAG Based in-Network Aggregation for Sensor Network Monitoring. *IEEE/IPSJ International Symposium on Applications and the Internet (SAINT)*, Phoenix, 23-27 January 2006, 292-299. <http://dx.doi.org/10.1109/SAINT.2006.20>
- [17] Sharaf, M.A., Beaver, J., Labrinidis, A. and Chrysanthis, P.K. (2003) TiNA: A Scheme for Temporal Coherency-Aware In-Network Aggregation. *ACM Data Engineering for Wireless and Mobile Access*, San Diego, 19 September 2003, 69-76.
- [18] Yoon, S. and Shahabi, C. (2007) The Clustered AGgregation (CAG) Technique Leveraging Spatial and Temporal Correlations in Wireless Sensor Networks. *ACM Transactions on Sensor Networks*, **3**, Article No. 3.

- [19] Solis, I. and Obraczka, K. (2005) Efficient Continuous Mapping in Sensor Networks Using Isolines. *IEEE International Conference on Mobile and Ubiquitous Systems (MobiQuitous): Networking and Services*, San Diego, 17-21 July 2005, 325-332.
- [20] Mirian, F. and Sabaei, M. (2009) A Delay and Accuracy Sensitive Data Aggregation Structure in Wireless Sensor Networks. *IEEE International Conference on Information Management and Engineering (ICIME)*, Kuala Lumpur, 3-5 April 2009, 231-235.
- [21] Weerasinghe, H., Elhajj, I.H., Krsteva, A. and Najm, M.A. (2007) Data Centric Adaptive In-Network Aggregation for Wireless Sensor Networks. *IEEE International Conference on Advance Intelligent Mechatronics (AIM)*, Zurich, 4-7 September 2007, 1-6.
- [22] Yang, X.F., Wu, X.B. and Huang, J.A. (2009) TAGPP: A Tiny Aggregation Algorithm with Preprocessing in Local Cluster. *IEEE International Conference on Networks Security, Wireless Communications and Trusted Computing (NSWCTC)*, Wuhan, 25-26 April 2009, 390-393.
- [23] Barr, K.C. and Asanović, K. (2006) Energy-Aware Lossless Data Compression. *ACM Transactions on Computer Systems*, **24**, 250-291. <http://dx.doi.org/10.1145/1151690.1151692>
- [24] Sadler, C.M. and Martonosi, M. (2006) Data Compression Algorithms for Energy-Constrained Devices in Delay Tolerant Networks. *ACM International Conference on Embedded Networked Sensor Systems (SenSys)*, Boulder, 1-3 November 2006, 265-278.
- [25] Marcelloni, F. and Vecchio, M. (2010) Enabling Energy-Efficient and Lossy-Aware Data Compression in Wireless Sensor Networks by Multi-Objective Evolutionary Optimization. *Information Sciences*, **180**, 1924-1941. <http://dx.doi.org/10.1016/j.ins.2010.01.027>
- [26] Petrovic, D., Shah, R.C., Ramchandran, K. and Rabaey, J. (2003) Data Funneling: Routing with Aggregation and Compression for Wireless Sensor Networks. *IEEE International Workshop on Sensor Network Protocols and Applications (SNPA)*, Anchorage, 11 May 2003, 156-162.
- [27] Arici, T., Gedik, B., Altunbasak, Y. and Liu, L. (2003) PINCO: A Pipelined In-Network Compression Scheme for Sata Collection in Wireless Sensor Networks. *IEEE International Conference on Computer Communications and Networks (ICCCN)*, Dallas, 20-22 October 2003, 539-544.
- [28] Pradhan, S.S., Kusuma, J. and Ramchandran, K. (2002) Distributed Compression in a Dense Micro-Sensor Network. *IEEE Signal Processing Magazine*, **19**, 51-60. <http://dx.doi.org/10.1109/79.985684>
- [29] Wei, D., Kaplan, S. and Chan, H.A. (2008) Energy Efficient Clustering Algorithms for Wireless Sensor Networks. *IEEE International Conference on Communications (ICC) Workshops*, Beijing, 19-23 May 2008, 236-240.
- [30] Younis, O., Krunz, M. and Ramasubramanian, S. (2006) Node Clustering in Wireless Sensor Networks: Recent Developments and Deployment Challenges. *IEEE Network*, **20**, 20-25. <http://dx.doi.org/10.1109/MNET.2006.1637928>
- [31] Abbasi, A.A. and Younis, M. (2007) A Survey on Clustering Algorithms for Wireless Sensor Networks. *Computer Communications*, **30**, 2826-2841. <http://dx.doi.org/10.1016/j.comcom.2007.05.024>
- [32] Amaxilatis, D., Chatzigiannakis, I., Koninis, C. and Pyrgelis, A. (2011) Component Based Clustering in Wireless Sensor Networks. ACM Computing Research Repository.
- [33] Younis, O. and Fahmy, S. (2004) HEED: A Hybrid, Energy-Efficient, Distributed Clustering Approach for Ad Hoc Sensor Networks. *IEEE Transactions on Mobile Computing*, **3**, 366-379. <http://dx.doi.org/10.1109/TMC.2004.41>
- [34] Lindsey, S. and Raghavendra, C.S. (2002) PEGASIS: Power-Efficient Gathering in Sensor Information Systems. *IEEE Aerospace Conference*, Big Sky, 9-16 March 2002, 1125-1130.
- [35] Ye, M., Li, C., Chen, G. and Wu, J. (2007) EECS: An Energy Efficient Clustering Scheme in Wireless Sensor Networks. *IEEE International Performance, Computing, and Communications Conference (IPCCC)*, New Orleans, 7-9 April 2005, 535-540.
- [36] Rivero-Angeles, M.E. and Boubdallah, N. (2009) Event Reporting on Continuous Monitoring Wireless Sensor Networks. *IEEE Global Telecommunications Conference (GLOBECOM)*, Honolulu, 30 November-4 December 2009, 1-6.
- [37] Jurdak, R., Baldi, P. and Lopes, C.V. (2007) Adaptive Low Power Listening for Wireless Sensor Networks. *IEEE Transactions on Mobile Computing*, **6**, 988-1004. <http://dx.doi.org/10.1109/TMC.2007.1037>
- [38] Bouabdallah, N., Rivero-Angeles, M.E. and Sericola, B. (2009) Continuous Monitoring Using Event-Driven Reporting for Cluster-Based Wireless Sensor Networks. *IEEE Transactions on Vehicular Technology*, **58**, 3460-3479. <http://dx.doi.org/10.1109/TVT.2009.2015330>
- [39] Hohlt, B.A. (2005) The Design and Evaluation of Network Power Scheduling for Sensor Networks. Ph.D. Thesis, University of California, Berkeley.
- [40] Polastre, J., Hill, J. and Culler, D. (2004) Versatile Low Power Media Access for Wireless Sensor Networks. *ACM*

- Conference on Embedded Networked Sensor Systems (SenSys)*, Baltimore, 3-5 November 2004, 95-107.
- [41] Akyildiz, I.F., Su, W., Sankarasubramaniam, Y. and Cayirci, E. (2002) A Survey on Sensor Networks. *IEEE Communications Magazine*, **40**, 102-114. <http://dx.doi.org/10.1109/MCOM.2002.1024422>
 - [42] Jindal, A. and Psounis, K. (2004) Modeling Spatially-Correlated Sensor Network Data. *IEEE International Conference on Sensor and Ad Hoc Communications and Networks (SECON)*, Santa Clara, 4-7 October 2004, 162-171.
 - [43] Faruque, J. and Helmy, A. (2004) Rugged: Routing on Fingerprint Gradient in Sensor Network. *IEEE International Conference on Pervasive Service (ICPS)*, Beirut, 19-23 July 2004, 179-188. <http://dx.doi.org/10.1109/PERSER.2004.27>
 - [44] Levis, P., Lee, N., Welsh, M. and Culler, D. (2003) TOSSIM: Accurate and Scalable Simulation of Entire TinyOS Applications. *ACM Conference on Embedded Networked Sensor Systems (SenSys)*, Los Angeles, 5-7 November 2003, 126-137.
 - [45] Intel Lab Data. <http://db.csail.mit.edu/labdata/labdata.html>