STCS-GAF: Spatio-Temporal Compressive Sensing in Wireless Sensor Networks- A GAF-Based Approach

Mohammad Reza Ghaderi¹, Vahid Tabataba Vakili²,*, and Mansour Sheikhan³

¹,²,³Department of Electrical Engineering, Islamic Azad University, South Tehran Branch, Tehran, Iran.
²Department of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran.
*Corresponding Author’s Information: vakily@iust.ac.ir

ARTICLE INFO

ARTICLE HISTORY:
Received 20 May 2019
Revised 12 July 2019
Accepted 13 July 2019

KEYWORDS:
Compressive sensing
GAF protocol
Spatio-temporal
Wireless sensor network

ABSTRACT

Routing and data aggregation are two important techniques for reducing communication cost of wireless sensor networks (WSNs). To minimize communication cost, routing methods can be merged with data aggregation techniques. Compressive sensing (CS) is one of the effective techniques for aggregating network data, which can reduce the cost of communication by reducing the amount of routed data to the sink. Spatio-temporal CS (STCS), with the use of spatial and temporal correlation of sensor readings, can increase the compression rate in WSNs, thereby reducing the cost of communication. In this paper, a new method of STCS technique based on the geographic adaptive fidelity (GAF) protocol is proposed which can effectively reduce the communication cost and energy consumption in WSNs. In the proposed method, temporal data is obtained from random selection of temporal readings of cluster head (CH) sensors located in virtual cells in the clustered sensors’ area and spatial data will be formed from the data readings of CHs located on the routes. Accordingly, a new structure of sensing matrix will be created. The results show that the proposed method as compared to the method proposed in [29], which is the most similar method in the literature, reduces energy consumption in the range of 22% to 43% in various scenarios which were implemented based on the number of required measurements at the sink (M) and the number of measurements in the routes (m_r).

1. INTRODUCTION

The limitation of energy resources in the sensors is the main challenge in implementing wireless sensor networks (WSNs). Data aggregation and routing are two basic methods to reduce communication cost and energy consumption. Routing methods can be merged with data aggregation techniques to minimize energy consumption in the network. Recently, various applications of compressive sensing (CS) have been proposed for data aggregation in WSNs, suggesting that CS can reduce communication cost and results in network lifetime increase [1]. CS introduces a structure for development of methods for aggregation of correlated data in multi-hop WSNs [2, 3].

In most of CS techniques, only the spatial readings of the sensors and their correlation are used to compress the sensor readings. In other words, spatial correlation of data read by sensors located in different places of the network is used to compress signals that have been read in a given time slot. While data read by any sensor in a sampling period, which includes multiple time slots, can also be correlated (temporal correlation). In addition, another concept of sensing based on spatio-temporal correlation was introduced. The physical phenomena sensed by a WSN often
shows compressibility in both space and time domains. The sensory data in a WSN from natural phenomena generally exhibit correlations in both spatial and temporal (spatio-temporal) domains [4].

The concept of spatio-temporal sensing in a sensors' area is shown in Figure 1. Several studies have shown that the sensors observations are both spatially and temporally correlated. Typically, the observed sensor readings of natural phenomena have both temporal and spatial dependency. Spatial and temporal correlations along with the collaborative nature of WSNs bring significant potential advantages for the development of efficient communication protocols well-suited for the WSN paradigm.

![Figure 1: Concept of the spatio-temporal sensing.](image)

In general, the improvement of data gathering approaches that are based on CS in WSNs is originated from three aspects: (a) the improvement of the reconstruction algorithms of compressive sensing signals; (b) the production of the transformation base matrices; and (c) the production of the sensing matrices. In this paper, we propose a method for improving the sensing matrix. Moreover, routing techniques are also used to reduce energy consumption in WSNs. The communication cost in a WSN is the main source of energy consumption [5] and is a serious challenge in designing routing protocols [6, 7]. Routing protocols can be categorized in two general types: flat routing and hierarchical routing (or clustered routing). In the flat type, all nodes have the same function and send data to the sink in multi-hop paths. In the hierarchical routing protocols, sensor nodes are distributed in clusters and a sensor node is selected as cluster head (CH) node in each cluster. The task of sending data to the sink is performed by the CH nodes. In large-scale WSNs, it is recommended to use hierarchical routing methods [8]. In the clustered routing protocol, sensors know their geographic locations with the help of GPS-like equipment which is embedded in them. Consequently, these protocols also refer to geographic routing protocols. In the geographic routing protocols, sensors are addressed by their geographic locations. Moreover, sensors use their geographic location to determine the distance to other neighbors which can use them as a hop sensor in the route [9-11].

Geographic adaptive fidelity (GAF) [12] is one of the most well-known geographic routing protocols. In the GAF protocol, the sensors area is divided into square virtual cells. In this protocol, each sensor in each cell can communicate with sensors in its four neighboring square cells (horizontally and vertically). In GAF, the size of square cells is defined in such a way that the farthest sensors in the adjacent cells can communicate with each other depending on the sensor range (R). In this protocol, only a sensor in each cell is selected as the active sensor and the radio of the remaining sensors goes off in order to save energy in each sampling period. Selected sensors in the cells are basically CH sensors that are responsible for sending data and routing to the sink. Due to the limitations of the GAF protocol, which allows sending data only in two vertical and horizontal directions in square cells, another version of the GAF which is called diagonal GAF (DGAF) was proposed [13]. In DGAF, sensors in two adjacent diagonal cells can also communicate with each other. In this paper, we propose a new spatio-temporal CS (STCS) method for aggregating data based on the GAF protocol.

In this proposal, similar to DGAF, we first divide the sensors area into virtual squared cells and then lay the sensors area based on the cells locations to improve the routing. In each sampling period, only one sensor of each cell will be selected as the CH sensor. Data read by CHs are compressed at various time intervals based on the proposed STCS technique and then this data is sent to the sink based on a routing algorithm.

The continuation of this article is as follows: In the second section, theoretical foundations of CS theory will be discussed and in the third section, we will review the related works. The proposed method will be presented in the fourth section. In the fifth section, the proposed method will be evaluated and in the last section we will conclude the proposed method.

## 2. Compressive Sensing

### A. Spatial CS

Based on the CS theory, instead of sending \( N \) sensors readings to the sink, the sink will only require \( M \) compressed measurements which \( M \) is much smaller than \( N \) [14-16]. To explain the CS theory, consider a sensor network with \( N \) sensor nodes. If \( x \) is the vector of the spatially reading signals of \( N \) sensors in a sampling period and \( \Psi \) is a suitable transformation base, then the vector \( \alpha \in \mathbb{R}^N \), which is a sparse signal, can be obtained as \( \alpha = \Psi x \). But the \( \alpha \) vector length still remains equal to the number of network sensors \( (N) \). If we consider the matrix
Φ ∈ ℝ^{N×N} as the measurements matrix in the CS technique, then i-th column in this matrix belongs to i-th sensor (Φ_i) in the network. In this case, each sensor multiplies its spatial reading (α_i) by the vector of its own column (Φ_i). Thus, the set y_i = α_iΦ_i will be generated for all sensors in which α_i is the data read by the i-th sensor and Φ_i is the i-th column of the matrix Φ. If all sensors send their y_i values to the sink, then the vector y = αΦ will arrive at the sink. The length of y is equal to M, which is less than the sparse reading vector of the sensors (x). So, less energy will be used to send data to the sink. The sink can easily recover the reading data vector (x) from the measurements vector (y) [17]. Using the method of minimizing L-norm [14, 18-20], or other recovery methods such as the orthogonal matching pursuit (OMP) algorithm [21], the original vector (x) can be obtained from the vector of measurements (y). However, there are two basic conditions for the Ψ and Φ matrices which play important roles in the stable recovery of the compressed signal. These two conditions are as follows: (a) the restricted isometric property (RIP) for the matrix A [16, 22, 23] and (b) the mutual coherence between the Ψ and Φ matrices [16, 22, 24]. The matrix A is defined as: A = ΨΦ ∈ ℝ^{M×N} [20, 25]. To implement the CS technique, M must follow M ≥ C.K. logN/M(Ψ,Ψ) [24] where C is a constant value greater than 1, k is the sparsity and μ(Ψ,Ψ) is the coefficient of coherence between Φ and Ψ matrices.

B. STCS

To describe the STCS technique, assume that the X ∈ ℝ^{T×N} signal represents a two-dimensional signal which includes T data read at t time slots (1 ≤ t ≤ T) by N sensors in each sampling period. If i represents any of sensors (i ∈ N), then (t, i)-th input of the signal X represents the t-th data read by the i-th sensor. It can be shown that the X signal in the form of X = [x_1, x_2, ..., x_T], in which the columns x_t ∈ ℝ^N represent the data read by i-th sensor in T sampling period. We can also show X as X = [x_1, x_2, ..., x_T]^T where, x_t ∈ ℝ^N represents the data read in a time slot t (1 ≤ t ≤ T) by all the sensors. It is assumed that X has spatial-temporal correlation. We assume that there is suitable transformation base in spatial and temporal domains which X has a sparse representation on them. These transformations can be represented in spatial and temporal domains by Ψ_s ∈ ℝ^{N×N} and Ψ_t ∈ ℝ^{T×T}, respectively. Therefore, each vector of data (the data read by all sensors at time slot t), namely x_t; t = 1, 2, ..., T has a sparse representation as x_t = Ψ_sα_s, in which α_s, ∈ ℝ^N is the transformation coefficients in the spatial domain.

With the accumulation of coefficients as α_s = [α_s,1, α_s,2, ..., α_s,T], the transformation of X in the spatial domain can be represented by X^T = Ψ_sα_s. It means that:

\[ x_1, x_2, ..., x_T = \Psi_s [\alpha_{s,1}, \alpha_{s,2}, ..., \alpha_{s,T}] \]  

Similarly, x_t will have a sparsity representation as x_t = Ψ_tα_{t,i}, where α_{t,i} ∈ ℝ^T is the transformation coefficient in the temporal domain. By the definition of α_t = [α_{t,1}, α_{t,2}, ..., α_{t,N}], the transformation of X in the temporal domain can also be represented by X = Ψ_tα_t or:

\[ x_1, x_2, ..., x_N = \Psi_t [\alpha_{t,1}, \alpha_{t,2}, ..., \alpha_{t,T}] \]  

The Kronecker sparse bases can combine different patterns of correlation in two dimensions to form a matrix [26-28]:

\[ \Psi = (\Psi_t \otimes \Psi_s) \in ℝ^{TN×TN} \]  

Due to the irregular distribution of sensors in the sensors area, Ψ_s and Ψ_t can be constructed using Laplacian graph vectors and discrete cosine transform (DCT), respectively [29]. In the conventional methods of STCS, the compressed measurement of the spatial readings of sensors in time slot t is shown by y_t ∈ ℝ^{M_t} where M_t is the size of measurements vector (M_t < N). The measurements vector can be represented by y_t = φ_t x^T where φ_t ∈ ℝ^{M_t×N} is a matrix with M_t rows and N columns. Considering all time slots in a sampling period (T), we can define the sensor matrix Φ as a diagonal-block matrix as follows [30]:

\[ \Phi = \begin{bmatrix} \phi_1 \\ \vdots \\ \phi_T \end{bmatrix} \in ℝ^{M×T} \]  

As a result, the general vector of measurements is defined as y = Φ x where y = [y_1, y_2, ..., y_T]^T ∈ ℝ^M represents the measurement vector and M = ∑_{t=1}^{T} M_t represents the total number of measurements involved in the STCS process during a sampling period T. Given the compression capability of the spatial domain, we can obtain the multi-dimensional signal X by solving the recovery problem using conventional decoding methods in CS technique. One can get any x^T; t = 1, 2, ..., T from the measurements separately by the following Eq. [26]:

\[ \hat{\alpha}_{s,t} := \arg\min_{\alpha_s} ||\alpha_s||_1 \]  

As a result, we can write \[ \hat{x}^T = \Psi_s \hat{\alpha}_{s,t} \]  

3. RELATED WORKS

Various CS-based data gathering methods have been proposed for data aggregation in WSNs. In some of these works, routing methods are implemented based on clustered routing protocols and in others, routings are presented based on tree-based routing approaches. Some studies have shown that clustered routing methods have better performance than tree-
based routing algorithms in terms of energy consumption and traffic load balancing [30]. In [1], a CS-based data aggregation method with a random walk-based routing algorithm (which is a tree-based routing) was suggested in order to reduce communication cost. In [1], the CS-based data was routed based on random routes called random walks. In addition, a method was presented using a CS technique and graph theory in which random measurements were aggregated by random walks. The random sensing matrix in this method was defined by random walks. Similar to [1], a method was proposed for data aggregation based on CS with a binary sensing matrix in [29], where random walk and random sampling were used jointly in a clustered network. That sensing matrix as a structured matrix was an unbalanced expander graph adjacent matrix (UEG-AM) if the following parameters were selected correctly: (a) random walk length; (b) random sampling probability; (c) number of measurements; and (d) number of clusters. In [32], a method called neighbor-aided compressive sensing (NACS) was proposed in which data aggregation was performed using a STCS-based method. In this method, the sensor nodes only send the rows of data of their temporal readings to the randomly selected neighbor sensors in each sampling period. Then, the sensor measurement which was produced by the neighboring sensor was sent directly to the sink. Sensing matrix in this method was also produced on this basis. In [29], a data aggregation algorithm was proposed based on random sampling and random walk-based routing in the spatio-temporal domain. In [29], the UEG-AM [31] has been also used as the sensing matrix. In many CS-based proposed methods in WSNs, tree-based routing algorithms were employed. It has been shown that clustered routing methods show better performance than tree-based routing algorithms for energy consumption and traffic load balancing [30].

In [33], the authors exploited the mobility pattern for spatio-temporal mobile data gathering method by employing an improved random walk algorithm for a “mobile sink” to collect data from a sensors area. The proposed scheme in [33] exploited Kronecker compressive sensing (KCS) for spatio-temporal correlation of sensory data by allowing the mobile sink to gather temporal compressive measurements from a small subset of randomly selected nodes along a random routing path. Authors in [34] proposed an adaptive sampling method based on spatio-temporal correlation of sensor readings for clustered WSNs. In [34], a clustering method according to spatial correlations of sensor nodes was proposed. In addition, sensors area was divided into clusters so that each CH kept a prediction model for sensors reading data which was derived from historical data in the temporal domain. In this method, redundant data transmission was reduced by adjusting temporal sampling frequency. Some sensor sets were selected within each cluster following intra-cluster correlation, and only one collection was needed to be activated at each sampling round time. Sensors reading data of non-sampler can be substituted by those of sampler due to strong spatial correlation between them. In [35], the authors proposed a hierarchical adaptive spatio-temporal data compression (HASDC) method. The method proposed in [35] explores the temporal correlation of sensory data by employing the discrete cosine transform (DCT) and adaptive threshold compression algorithm (ATCA). The CH explores the spatial correlation among the compressed temporal readings by utilizing discrete wavelet transform (DWT) and ATCA. The proposed method in [35] combines three techniques including: (a) data sorting; (b) ATCA; and (c) spatio-temporal data gathering method. At the same time, according to the correlation of sensory data and the adaptive threshold value, the compression ratio can be adaptive-controlled. In [36], for improving the accuracy of reconstructed data, a method based on weighted spatio-temporal compressive sensing was proposed. In the proposed method in [37], the sensors area was clustered and in order to reduce the temporal redundancy. In addition, dual prediction was used in the intra-cluster transmission and hybrid compressive sensing (HCS) technique was employed for reducing the inter-cluster spatial transmission redundancy.

4. PROPOSED ALGORITHM

A. Network Model

We have some assumptions in our model:
I) Impact of packet loss during data transmission is not considered, since it can be handled at the lower layer.
II) Only energy consumption due to communication processes is considered. The energy consumption of computation processes depends on the type of sensor node; so, it is neglected. In general, it is noted that in a typical WSN, energy consumption during data acquisition (sensing) and signal processing depend on the type of sensor nodes; so, this work only considers energy consumption during communication between sensors. Moreover, communication between sensors has the highest energy consumption in WSNs. Transmitting and receiving signals consume about two-thirds of the total energy in a typical sensor. So, we only consider energy consumption during the transmitting and receiving data packets in this paper.

In our proposed model, the sensors are distributed uniformly and randomly in a square region. Similar to the DGAF, we divide the sensors area into equal virtual squared cells, so those sensors in adjacent cells
(in addition to vertical and horizontal directions) can also communicate in a diagonal direction. In this model, the sink is located at the center of the sensors area. In this paper, we use the geographic sink placement (GSP) strategy for sink placement in the sensors area. In GSP, sink is placed at the center of gravity of an area. The GSP strategy is intended for uniformly distributed networks when there is no information about sensor nodes’ locations. In this paper, the sensor nodes are distributed uniformly in sensors area based on the network model, so our strategy is same as GSP. The sink placement at the center of gravity gives pretty good results for uniform node distributions. For applications whose nodes are uniformly distributed, GSP is a good option in order to minimize the delay. Furthermore, GSP strategy is obviously very computationally efficient. In many researches in the field of WSNs, the sink is located at the center of sensors area.

In order to reduce energy consumption in the network, similar to the GAF protocol, one of the sensors in each cell was randomly selected as an active sensor (CH sensor) at each sampling period. Each sampling period \( T \) is divided into equal time slots \( t \). The CH sensors read and record the environmental data in time slots \( 1 \leq t \leq T \). Since both spatial and temporal readings of the sensors in the time slots and sampling period may be correlated, it is possible to aggregate the data recorded in the CH sensors based on the STCS technique. The proposed routing algorithm in this paper is implemented based on the geographic location of the sensors. In this model, each cell has an address in the form of \( C(I,J) \), where \(-m \leq I,J \leq m\) and \( m = \sqrt{G} \). \( A \) and \( r_G \) are introduced in Tables 2 and 3, respectively.

To reduce the energy consumption in WSNs, various routing protocols have been introduced by the researches. Furthermore, based on the network structure class, routing approaches are divided into flat and hierarchical protocols. In the flat routing (similar to the routing method which is proposed in [29]-random walk routing), all sensor nodes cooperate with each other through multi-hop routing. In this type of routing, nodes have the same role. On the other hand, the hierarchical routing is classified into two categories: cluster-based and grid-based clustering techniques.

Figure 2 shows the proposed network model. In order to improve the routing, we lay the cellular sensors area. In Figure 2, cells in different layers are shown in different colors. The lowest layer is the closest layer to the sink which is considered as the first layer. In proposed model, there is a routing constraint by which routing is implemented only from the higher layer to the lower one. Therefore, if CH sensors are located in the diagonal cells of the sensors area, there is only one path to send data from the higher layer to the lower one. If the CH sensors are located in non-diagonal cells, depending on which cells they are located, there are 2 or 3 paths from higher layer to the lower one. The decision to select a CH from 2 or 3 candidate cells is based on the shortest distance of the candidate cells. In other words, the CH of candidate cells is chosen as a relay sensor which has a lower distance with the source sensor. This process is shown in the routes of Figure 2. The routes with solid lines are implemented routes and dashed lines are candidate routes that are not selected due to more distance to the source sensor. In summary, the effects of clustering in our results are as follows: (a) simplification of the node management; (b) reduction of energy consumption; (c) achieving scalability; and (d) improving load balancing, robustness and data aggregation.

### B. Data Gathering

In a sampling period \( T \), a sensor will be randomly selected as a CH sensor in each cell. The CH sensors will read the environmental data temporally in \( t \) time slots. In a sampling period, a number of \( R \) routes are implemented in the cellular network. Temporal data is gathered based on temporal readings of CH sensors in \( t \) time slots, and spatial data is gathered based on CH readings on the routes. This process can be described as follows: According to the CS theory, assume that the sink requires \( M \) measurements to accurately recover the temporally read signals by the CHs. In addition, assume that \( R \) routes are generated in each sampling period. Therefore, \( m_r \) measurements will be performed in each route such that \( M=rm_r \).

To explain how to compress and send compressed data to the sink, consider \( r \)-th route. Assume that the length of this route (the number of hops) is equal to \( h \) (Figure 3).

If the temporal data is represented by \( x_i \), then the temporal data read by the first sensor in the route (source sensor \( n_1 \)) can be given by \( x_{n_1} \in \mathbb{R}^T \). Now the compressed measurement of the temporal readings of the sensor \( n_1 \) in a time slot can be shown as \( y_r = \phi_r^T x_{n_1} \) where \( y_r \in \mathbb{R}^{m_r} \) and \( m_r \) is the length of the measurements on the \( r \)-th route \( (m_r<T) \). This relation, \( \phi_r^T \in \mathbb{R}^{m_r \times T} \) is the measurement matrix with \( m_r \) rows and \( T \) columns. The length of the measurements on the \( r \)-th route \( (m_r) \) is formed by random selection of the temporal readings of the sensor \( n_1 \). The matrix elements \( \phi_{n_1}^r \) are represented by \( \phi_{i_r}^r \) with the values of 0 or 1. If \( n_1 \), temporal readings are shown by \( x_{n_1,t} \), then we can select a number of temporal readings \( x_{n_1,t} \) randomly with probability of \( q \). In general, for \( i \)-th sensor in \( r \)-th route, the elements of the matrix \( \phi_i^r \) can be defined as [29]:

---

In general, for \( i \)-th sensor in \( r \)-th route, the elements of the matrix \( \phi_i^r \) can be defined as [29]:

---


157
The sensor will relay its compressed data to the next hop sensor in the route (i.e., sensor). As before, the vector of compressed measurements in the sensor will be generated as where is the vector of compressed data. The sensor will send the total sum to the next sensor on the route. So, all the sensors in the route send their data to the sink (Fig. 3). Finally, after hop counts, the data will arrive at the sink. Consequently, the vector of measurements in the -th route will be

\[ y_r = \sum_{i=1}^{n_h} \phi_i^r x_i \]

(6)

In this case, the number of transmitted data packets in the route is of the order of \( O(m_r h) \) [18]. When all the routes (R routes) in the network run such a CS data gathering approach, a measurement vector as \( y = \phi x \) is created. The sensing matrix \( \phi \in \mathbb{R}^{R \times N_{CH}} \) has R rows and \( N_{CH} \) columns.

Here we describe the structure of the matrix \( \phi \) by an example based on the proposed routing algorithm. We consider a sensors area consisting of 16 square cells in two layers and a total of 16 CH sensors in each sampling period (Figure 4). For simplicity, assume that CHs are located in the center of the cells. According to the proposed routing algorithm, a
number of $R = 12$ routes can be created in a sampling period.

These routes are illustrated by continuous lines of different colors. It should be noted that other routes can be created, as well. These routes are shown in Figure 4 by the red dashed lines. Accordingly, the matrix $\phi$ in this example will have 12 rows and 16 columns. In all of the routes $h = 2$, so in each row of the matrix $\phi$ only two non-zero elements exist and the remaining elements in each row are zero. In this example, the $\phi$ matrix elements are the matrices as $\phi_i^r \in \mathbb{R}^{16 \times 12}$. The zero elements in the matrix $\phi$ are matrices of $0 \in \mathbb{R}^{16 \times 12}$. The zero elements shown in red represent the dashed red lines. In this example, only 2 non-zero elements exist in each of the routes. Therefore, it can be concluded that the matrix $\phi$ is a sparse matrix. In the $r$-th route, a vector of length $m_r$ is created. Therefore, each of the CH sensors that located in the $r$-th route will equally send $m_r$ data packets to the next CH sensor on the route. The number of "inter-cell" data packets in each route (e.g., in the $r$-th route) is $TR_r = m_r h_r$ where $h_r$ is the number of sensors located on the route $r$. As a result, the number of transmitted packets to all routes in the sensors area ($R$ routes) is $TR = \sum_{r=1}^{R} m_r h_r$. The number of received data packets per route is equal to the number of transmitted data packets in that route minus one. The source sensor in the route will not receive any data packet. In other words, all sensors on the route, except the source sensor, will send and receive data packets. So the number of received packets per route is $RE_r = m_r (h_r - 1)$ and the total number of received packets in all routes in each sampling period is obtained by $RE_R = \sum_{r=1}^{R} m_r (h_r - 1)$.

C. Energy Consumption Model

Energy consumption in a WSN node can be categorized into two groups: (a) energy consumption due to computation and (b) energy consumption due to communication. It is noted that in a typical WSN, energy consumption during data acquisition (sensing) and signal processing depend on the type of employed sensor nodes and is based on the application; so, this work only considers energy consumption during communication between sensors. In WSNs, communication between sensors has the highest energy consumption. Transmitting and receiving signals consume about two-thirds of the total energy in a typical sensor. However, energy consumption during sensing and processing depends on the type of sensors. So, we only consider energy consumption during the transmitting and receiving data packets in this paper.

Based on the energy consumption model proposed in [33], the energy consumed by a sensor for sending and receiving a $b$-bit data packet are obtained by

$$E_{Tx} = (E_{elec} + \varepsilon_{amp} d^l) b$$
and

$$E_{Rx} = E_{elec} b,$$
respectively where $d$ is the distance between transmitter and receiver sensors, $E_{elec}$ is the amount of energy consumed per bit in the transmitter and receiver circuits, and $\varepsilon_{amp}$ is the transmitter amplifier parameter. It is assumed that for a free space model $l = 2$, and for a multi-path fading model $l = 4$ [33]. This model is shown in Fig. 5.

$$E_{Tx} = (E_{elec} + \varepsilon_{amp} d^l) b$$
and

$$E_{Rx} = E_{elec} b,$$
respectively where $d$ is the distance between transmitter and receiver sensors, $E_{elec}$ is the amount of energy consumed per bit in the transmitter and receiver circuits, and $\varepsilon_{amp}$ is the transmitter amplifier parameter. It is assumed that for a free space model $l = 2$, and for a multi-path fading model $l = 4$ [33]. This model is shown in Fig. 5.

$$E_{Tx} = (E_{elec} + \varepsilon_{amp} d^l) b$$
and

$$E_{Rx} = E_{elec} b,$$
respectively where $d$ is the distance between transmitter and receiver sensors, $E_{elec}$ is the amount of energy consumed per bit in the transmitter and receiver circuits, and $\varepsilon_{amp}$ is the transmitter amplifier parameter. It is assumed that for a free space model $l = 2$, and for a multi-path fading model $l = 4$ [33]. This model is shown in Fig. 5.

$$E_{Tx} = (E_{elec} + \varepsilon_{amp} d^l) b$$
and

$$E_{Rx} = E_{elec} b,$$
respectively where $d$ is the distance between transmitter and receiver sensors, $E_{elec}$ is the amount of energy consumed per bit in the transmitter and receiver circuits, and $\varepsilon_{amp}$ is the transmitter amplifier parameter. It is assumed that for a free space model $l = 2$, and for a multi-path fading model $l = 4$ [33]. This model is shown in Fig. 5.
consumption for receiving data packets for all routes in each sampling period is obtained by

$$E_{\text{RER}} = \sum_{r=1}^{R} E_{\text{RER}}$$

5. Performance Evaluation

For evaluation of the proposed method, we consider a square sensors area with a side of 2000 m in which 1000 sensors are distributed. Moreover, energy parameters were obtained for the network model simulation. The energy parameters are assumed as listed in Table 1. The other parameters which are adjusted during the algorithm execution are given in Table 2. The equations that are used to plot some of the following figures in this article are listed in Table 3 (Figures 7, 9, 11, and 13). It is noted that the complete set of equations that are needed to design and implement the proposed method are listed in Table 4.

A. Sensors Area Division and Layering

In the proposed model, the sensors area is divided into equal virtual square cells (grid-based clustering).

### Table 1
Pre-Defined Network and Energy Parameters

<table>
<thead>
<tr>
<th>Description of parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy to send a bit</td>
<td>$E_{\text{elec}}$</td>
<td>50 nJ/bit</td>
</tr>
<tr>
<td>Free space transmitter power amplifier parameter</td>
<td>$e_{fs}$</td>
<td>100 pJ/bit/m$^2$</td>
</tr>
<tr>
<td>Length of data packets</td>
<td>$b$</td>
<td>64 bits</td>
</tr>
<tr>
<td>Average distance between the nodes</td>
<td>$d$</td>
<td>63 m</td>
</tr>
</tbody>
</table>

### Table 2
Parameters Adjusted During the Algorithm Execution

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>Area of sensors’ region</td>
</tr>
<tr>
<td>$Cell$</td>
<td>Number of cells</td>
</tr>
<tr>
<td>$M$</td>
<td>Number of required measurements in the sink</td>
</tr>
<tr>
<td>$m_r$</td>
<td>Number of measurements in the $r$-th route</td>
</tr>
<tr>
<td>$R$</td>
<td>Sensor transmission range</td>
</tr>
<tr>
<td>$CH$</td>
<td>Number of cluster heads</td>
</tr>
</tbody>
</table>

The cell sensors can send data along the diagonal direction in addition to vertical and horizontal directions. As a result, the maximum size of the cell is obtained by $r_{G(\text{max})} = \frac{R}{2\sqrt{2}}$, where $R$ is the sensor range (Figure 6). Figure 7 shows the relationship between the cell size and sensor range. This figure shows that it is necessary to increase the range of sensor to create larger cells in the sensors area. However, considering that in large-scale WSNs, the routes are multi-hop paths and the energy consumed for sending data is proportional to the $d^l$ ($l=2$ or 4) where $d$ is the distance between the hop sensors in the route. So, increasing the size of the cells will increase the energy consumption of sending data considerably.

### Table 3
Equations Used to Plot Figs. 7, 9, 11, and 13

<table>
<thead>
<tr>
<th>Figure No.</th>
<th>Equation No.</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>7</td>
<td>$r_{G(\text{max})} = \frac{R}{2\sqrt{2}}$</td>
</tr>
<tr>
<td>9</td>
<td>13</td>
<td>$h = \frac{\sqrt{N_{\text{CH}}}}{2}$</td>
</tr>
<tr>
<td>11</td>
<td>16</td>
<td>$TR_r = m_r h_r = m_r \frac{N_{\text{CH}}}{2} R \sqrt{\frac{h_r}{M \sqrt{N_{\text{CH}}}}} = \frac{M \sqrt{N_{\text{CH}}}}{8(\sqrt{N_{\text{CH}}} - 1)}$</td>
</tr>
<tr>
<td>13</td>
<td>17</td>
<td>$TR_e = TR_r \times R = \frac{M (N_{\text{CH}} + \sqrt{N_{\text{CH}}})}{2(\sqrt{N_{\text{CH}}} - 1)} = 2m_r (N_{\text{CH}} + \sqrt{N_{\text{CH}}})$</td>
</tr>
</tbody>
</table>

Figure 6: The farthest two sensors in the square cells.

Moreover, the number of cells in the sensors area depends on the range of sensor. The number of cells in the sensors area is obtained by $Cell = \frac{8A}{R^2}$

Figure 8 shows the relationship between the sensor range and the number of cells for different dimensions of sensors area. In a fixed-dimension sensors area, decreasing the number of cells requires an increase in the range of sensors, as the number of cells is decreased, the distance between the sensors in the adjacent cells is increased. In this case, the energy consumed will be increased for sending data.

In the proposed model, the sensors area is layered to improve the routing. Based on the proposed routing algorithm, routes are formed from higher layer to
lower one. The number of layers and routes depend on the number of cells in the sensors area. The relationship between the number of routes and the number of cells in the network is \( R = 4\sqrt{\text{Cell} - 1} \). The length of the routes, or the number of hop sensors in the route, depends on the number of cells in the sensors area. In the proposed model, the relationship between the length of the routes with the number of cells in the sensors area is \( h = \sqrt{\text{Cell}/2} \). Figure 9 shows the relationship between the number and length of routes with the number of the cells in the sensors area. It can be seen that as the number of cells is increased, the length and number of the routes is also increased. But as seen in Figure 9, an increase in the number of cells by one hundred times, only increases the length of the paths by ten and increases slightly more than ten times in the number of routes.

For example, in point A of Figure 9, the number of cells is 100. If the number of cells is increased by 100 times (i.e., 10000 that is shown by point B), then the length of the route changes from 5 to 50 and the number of routes changes from 40 to 400 (points C and D). This means that an increase in the number of cells by one hundred times, only increases the length and number of routes by ten. So, the proposed method is scalable and may be used in a wide range of WSNs with different cell sizes.

B. Sensor Density in the Cell

In the proposed model, the sensors are distributed randomly and uniformly in the cells. Therefore, the sensor density (the number of sensors in the cell) will depend on the number of cells. The relationship between the number of sensors in a cell with the sensor range can be represented by \( N_{\text{cell}} = \frac{N_R^2}{B_A} \). Figure 10 shows the number of sensors in a cell for different dimensions of sensors area and a certain number of CHs \((N_{\text{CH}} = 1000)\).

C. Communication Cost

The communication cost in a WSN is considerably dependent on the number of transmitted and received data packets, as well as the length of the route. In the proposed model, to reduce communication cost, both data aggregation and routing techniques are used to reduce the number of data packets and the length of the routes, respectively. In the proposed model, reducing the size of the STCS measurements in the routes \((m_r)\) and sending them in multi-hop routes (based on geographic routing algorithm) result in communication cost reduction in the network. In this model, the number of transmissions in each route is obtained by \( TR_r = \frac{m_r \sqrt{N_{\text{CH}}}}{2} \). The communication cost of sending data on all routes in each sampling period is obtained by \( TR_R = 2m_r (N_{\text{CH}} + \sqrt{N_{\text{CH}}}) \).

Figures 11 and 12 show the relationship between the number of transmitted data packets with the number of active sensors in each sampling period (the number of CHs) for a route and for all routes, respectively. The length of the vector of measurements in the route \((m_r)\) is determined by the number of required measurements in the sink \((M)\). According to the relation \( M = m_A \sqrt{N_{\text{CH}} - 1} \). Figures 11 and 12 show that an increase of 100 times in the number of CHs (or the number of cells) only increases the number of transmitted packets by 10. Therefore, the proposed model will be very useful for large-scale WSNs. By increasing the number of cells in the sensors area, the distance between the sensors in multi-hop routes will be decreased and thus reduce the amount of energy consumed for sending and receiving data packets which is proportional to the second (or fourth) power of distance between hop sensors in the routes.

In the proposed model, the total number of transmitted packets in a sampling period is obtained by \( TR_R = \frac{M (N_{\text{CH}} + \sqrt{N_{\text{CH}}})}{2(N_{\text{CH}} - 1)} \). Figure 13 shows the relationship between the total number of transmitted packets and the number of measurements required in the sink for a number of cells. The results obtained from this figure confirm the results of Figures 11 and 12. These results generally indicate the superior performance of this model for large-scale WSNs.

In this section, a simulation for obtaining the energy efficiency is performed. It is noted that energy efficiency is defined as the ratio of transmitted data packets to average energy consumption in the route. We consider a square sensors area in which a number of sensors are distributed randomly and independently. The energy parameters are shown in Table 1. We perform our analysis for various route lengths: \( h_r = 50 \) to 500. For simplicity, assume that the average distance between sensor nodes in the routes is 63 m. Analytic evaluation is done based on Eqs. (16) and (20) of Table 4. Energy parameters are listed in Table 1. Figure 14 shows the energy efficiency which is defined as average energy consumption per transmitted data packet in the route.

D. Comparative Evaluation

In order to evaluate the proposed method in comparison with a similar method, we compare the proposed method with the method presented in [29], which is a STCS-based method. Given that, the routing performed in [29] was based on the random walk approach, while the routing is based on a geographic routing protocol in our proposed method. For this purpose, we consider a square sensors area in which sensors are distributed randomly and uniformly similar to [29]. We divide the sensors area into virtual square cells so that only one sensor is located in each.
cell. The comparison is performed in three scenarios based on the number of required measurements in the sink \( (\overline{M}) \) and the number of measurements in the routes \( (\overline{m}_r) \). The network and energy parameters and also scenario specifications are given in Table 5. Based on the proposed model, when the number of cells in the sensors area is considered to be 1000 cells, the number of generated routes will be equal to 50 routes. On the other hand, there is a relationship between the number of required measurements in the sink and the number of the routes \( (R = \frac{\overline{M}}{\overline{m}_r}) \). The results of this comparison for three scenarios are shown in Figure 15. It can be seen that the amount of energy consumed in each sampling period in the proposed model shows 22.73% reduction in the first scenario, 35.57% reduction in the second scenario, and 43.53% reduction in the third scenario as compared to the proposed method in [29].

6. CONCLUSION

In this paper, a new method of STCS technique based on the GAF protocol is proposed. In the proposed model, the sensors area is divided into square virtual cells and then the sensor area is layered to facilitate routing. In this method, temporal data is obtained from random selection of temporal readings of each \( CH \) in the sensors area and spatial data will be formed from the data readings of \( CHs \) located on the routes. Based on this model, a new structure for the sensor matrix is created which reduces the energy consumption of the network. The results of evaluation of the proposed model indicate that this model will be suitable for large-scale WSNs.

We compared the proposed method with the method presented in [29] which is an STCS-based method. Although, routing in [29] was performed based on the random walk approach, while routing in our proposed method is done based on a geographic routing protocol. Comparison is performed in three scenarios based on the various values of required measurements in the sink \( (\overline{M}) \) and the number of measurements in the routes \( (\overline{m}_r) \). The results show that the proposed method reduces energy consumption in the range of 22% to 43% in various scenarios as compared to the method proposed in [29].
Table 4

<table>
<thead>
<tr>
<th>Eq. No.</th>
<th>Description</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Cell size</td>
<td>( r_G(\text{max}) = \frac{R}{2\sqrt{2}} )</td>
</tr>
<tr>
<td>8</td>
<td>Cell area</td>
<td>( A_{\text{cell}} = r_G^2 = \frac{R^2}{B} )</td>
</tr>
<tr>
<td>9</td>
<td>Total number of cells</td>
<td>( \text{Cell} = \frac{A}{A_{\text{cell}}} = \frac{A}{r_G^2} = \frac{8A}{R^2} )</td>
</tr>
<tr>
<td>10</td>
<td>Number of sensors in the cell</td>
<td>( N_{\text{cell}} = \frac{N}{\text{Cell}} = \frac{NR^2}{8A} )</td>
</tr>
<tr>
<td>11</td>
<td>Total number of CHs in sensors’ area</td>
<td>( N_{\text{CH}} = \text{Cell} )</td>
</tr>
<tr>
<td>12</td>
<td>Number of sensors’ area layers</td>
<td>( l = \frac{\sqrt{\text{Cell}}}{2} )</td>
</tr>
<tr>
<td>13</td>
<td>Length of route</td>
<td>( h = \sqrt{\frac{N_{\text{CH}}}{2}} = \frac{\sqrt{\text{Cell}}}{2} )</td>
</tr>
<tr>
<td>14</td>
<td>Total number of routes</td>
<td>( R = 4(2L - 1) = 4(\sqrt{\text{Cell}} - 1) = 4(\sqrt{N_{\text{CH}} - 1}) = 4(2h - 1) = \frac{M}{m_r} )</td>
</tr>
<tr>
<td>15</td>
<td>Number of required measurements in the sink</td>
<td>( M = m_r A(\sqrt{N_{\text{CH}} - 1}) = m_r A(\sqrt{\text{Cell}} - 1) )</td>
</tr>
<tr>
<td>16</td>
<td>Number of transmitted packets per route</td>
<td>( TR_r = m_r h_r = m_r \frac{\sqrt{N_{\text{CH}} \times M}}{R} h = \frac{M}{8} \frac{\sqrt{N_{\text{CH}}}}{(\sqrt{N_{\text{CH}} - 1})} = \frac{m_r \sqrt{N_{\text{CH}}}}{2} )</td>
</tr>
<tr>
<td>17</td>
<td>Number of transmitted packets in all routes</td>
<td>( TR_R = TR_r \times R = \frac{M}{8} \frac{(N_{\text{CH}} + \sqrt{N_{\text{CH}}})}{(\sqrt{N_{\text{CH}} - 1})} = 2m_r (N_{\text{CH}} + \sqrt{N_{\text{CH}}}) )</td>
</tr>
<tr>
<td>18</td>
<td>Number of received packets in the route</td>
<td>( RE_r = m_r h_r - 1 = \frac{M}{8} \frac{\sqrt{N_{\text{CH}}}}{(\sqrt{N_{\text{CH}} - 1})} - 1 = \frac{m_r \sqrt{N_{\text{CH}}}}{2} - 1 )</td>
</tr>
<tr>
<td>19</td>
<td>Number of received packets in all routes</td>
<td>( RE_R = RE_r \times R = 2m_r \frac{\sqrt{N_{\text{CH}} - 4(\sqrt{N_{\text{CH}} - 1})}}{2} = 4(\sqrt{N_{\text{CH}} - 1})(2m_r \sqrt{N_{\text{CH}} - 1}) )</td>
</tr>
<tr>
<td>20</td>
<td>Energy consumption for data transmission per route in the sensors’ area</td>
<td>( E_{TR_r} = \sum_{i=1}^{h-1} (E_{\text{elec}} + \epsilon_{\text{amp}} d_{(n_i, n_i+1)}) bTR_r )</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( = \sum_{i=1}^{h-1} (E_{\text{elec}} + \epsilon_{\text{amp}} d_{(n_i, n_i+1)}) bm_r h_r )</td>
</tr>
<tr>
<td>21</td>
<td>Energy consumption for data transmission to all routes in the sensors’ area</td>
<td>( E_{TR_R} = \sum_{r=1}^{R} E_{TR_r} )</td>
</tr>
<tr>
<td>22</td>
<td>Energy consumption for data reception per route in the sensors’ area</td>
<td>( E_{RE_r} = (E_{\text{elec}} b)(m_r(h_r - 1)) )</td>
</tr>
<tr>
<td>23</td>
<td>Energy consumption for data reception from all routes in the sensors’ area</td>
<td>( E_{RE_R} = \sum_{r=1}^{R} E_{RE_r} )</td>
</tr>
</tbody>
</table>

Table 5

<table>
<thead>
<tr>
<th>Scenario</th>
<th>M</th>
<th>m_r</th>
<th>A</th>
<th>N</th>
<th>( E_{\text{elec}} )</th>
<th>( \epsilon_{fs} )</th>
<th>b</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>100</td>
<td>2</td>
<td>2000×2000 m²</td>
<td>1000</td>
<td>50 nJ/bit</td>
<td>100 pJ/bit/m²</td>
<td>64 bits</td>
<td>63 m</td>
</tr>
<tr>
<td>Second</td>
<td>200</td>
<td>4</td>
<td>2000×2000 m²</td>
<td>1000</td>
<td>50 nJ/bit</td>
<td>100 pJ/bit/m²</td>
<td>64 bits</td>
<td>63 m</td>
</tr>
<tr>
<td>Third</td>
<td>500</td>
<td>6</td>
<td>2000×2000 m²</td>
<td>1000</td>
<td>50 nJ/bit</td>
<td>100 pJ/bit/m²</td>
<td>64 bits</td>
<td>63 m</td>
</tr>
</tbody>
</table>
TABLE 6
ENERGY CONSUMPTION REDUCTION IN THE PROPOSED METHOD AS COMPARED TO THE METHOD REPORTED IN [29]

<table>
<thead>
<tr>
<th>Scenario</th>
<th>$M$</th>
<th>$m_r$</th>
<th>Energy reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td>100</td>
<td>2</td>
<td>22.73%</td>
</tr>
<tr>
<td>Second</td>
<td>200</td>
<td>4</td>
<td>35.57%</td>
</tr>
<tr>
<td>Third</td>
<td>500</td>
<td>6</td>
<td>43.53%</td>
</tr>
</tbody>
</table>

Figure 11: Number of transmitted data packets per route versus the number of CHs.

Figure 12: Number of transmitted data packets in all routes in a sampling period versus the number of CHs.

Figure 13: Communication cost versus the number of required measurements in the sink for different number of cells.

Figure 14: Average energy consumption per transmitted data packets in the route.

Figure 15: Comparison of average energy consumption in each sampling period in the proposed method and the method presented in [29].

REFERENCES


**BIographies**

Mohammad Reza Ghaderi is currently a Ph.D. student in communication department of Islamic Azad University South Tehran Branch, Tehran, Iran. He received his B.Sc. degree in Electrical Engineering from Kajeh Nasir Toosi University of Technology, Tehran, Iran, in 1992 and his M.Sc. degree in Communication Engineering form Tehran University in 2014. His research interests include wireless sensor networks and intelligent systems.
Vahid Tabataba Vakili received the B.S. degree from Sharif University of Technology, Tehran, Iran, in 1970, the M.S. degree from the University of Manchester, Manchester, UK, in 1973, and the Ph.D. degree from the University of Bradford, Bradford, UK, in 1977, all in Electrical Engineering. He is currently a Full Professor in Communication Engineering Department of Iran University of Science and Technology. His research interests are in the areas of mobile cellular systems, interference cancellation for CDMA systems, ultra wideband communication system and space–time processing and coding.

Mansour Sheikhan received the B.S. degree in Electronic Engineering from Ferdowsi University, Mashad, Iran, in 1988 and M.Sc. and Ph.D. degrees in Communication Engineering from Islamic Azad University (IAU), Tehran, Iran, in 1991 and 1997, respectively. He is currently a Full Professor in Communication Engineering Department of IAU-South Tehran Branch. His research interests include intelligent systems, speech signal processing, security in communication networks, and neural networks.

How to cite this paper:
DOI: 10.22061/JECEI.2019.5302.209
URL: http://jecei.sru.ac.ir/article_1104.html