Immune System Based Distributed Node and Rate Selection in Wireless Sensor Networks

Barış Atakan   Özgür B. Akan

Next Generation Wireless Communications Laboratory
Department of Electrical and Electronics Engineering
Middle East Technical University, 06531, Ankara, Turkey
Email:{atakan,akan}@eee.metu.edu.tr

***Abstract*— Wireless sensor networks (WSNs) are event-based systems that rely on the collective effort of dense deployed sensor nodes. Due to the dense deployment, since sensor observations are spatially correlated with respect to spatial location of sensor nodes, it may not be necessary for every sensor node to transmit its data. Therefore, due to the resource constraints of sensor nodes it is needed to select the minimum number of sensor nodes to transmit the data to the sink. Furthermore, to achieve the application-specific distortion bound at the sink it is also imperative to select the appropriate reporting frequency of sensor nodes to achieve the minimum energy consumption. In order to address these needs, we propose the new Distributed Node and Rate Selection (DNRS) method which is based on the principles of natural immune system. Based on the B-cell stimulation in immune system, DNRS selects the most appropriate sensor nodes that send samples of the observed event, are referred to as *designated nodes*. The aim of the designated node selection is to meet the event estimation distortion constraint at the sink node with the minimum number of sensor nodes. DNRS enables each sensor node to distributively decide whether it is a designated node or not. In addition, to exploit the temporal correlation in the event data DNRS regulates the reporting frequency rate of each sensor node while meeting the application-specific delay bound at the sink. Based on the immune network principles, DNRS distributively selects the appropriate reporting frequencies of sensor nodes according to the congestion in the forward path and the event estimation distortion periodically calculated at the sink by Adaptive LMS Filter. Performance evaluation shows that DNRS provides the minimum number of designated nodes to reliably detect the event properties and it regulates the reporting frequency of designated nodes to exploit the temporal correlation in the event data whereby it provides the significant energy saving.**

***Index Terms*— Wireless Sensor Networks, Immune system, Spatio-Temporal Correlation, Energy Efficiency.**

I. INTRODUCTION

Wireless Sensor Networks (WSN) are generally comprised of densely deployed sensor nodes collaboratively observing and communicating extracted information about physical phenomenon [1]. Due to the dense deployment of sensor nodes, sensor observations are the spatially correlated according to the spatial location of sensor nodes. This results in transmission of highly redundant sensor data which is not necessary to estimate the event properties at the sink. Therefore, due to the energy constraint of sensor nodes it is necessary to select the minimum number of designated node which can provide the sink to accurately estimate the event properties. Furthermore, the nature of the energy-radiating physical phenomenon constitutes the temporal correlation between each consecutive observation of a sensor node [2]. Therefore, it is necessary to regulate the reporting frequency of sensor nodes to achieve minimum energy consumption while achieving the event detection reliability.

There has been some research efforts to study about correlated data gathering in WSN according to the different methods [3], [4], [5], [6], [7]. These works provide the great deal of capability to exploit the correlation at the sensor data. In [10], the model for point and field sources are introduced and their spatio-temporal characteristics are derived along with the distortion functions. In [9], exploiting spatial correlation at the MAC layer is achieved by collaboratively regulating medium access so that redundant transmissions from correlation neighbors are suppressed. In [9], the Iterative Node Selection (INS) algorithm is given to select the representative nodes which represent the set of nodes generating spatially correlated data. Since INS selects the representative nodes with centralized manner at the sink and it computes the distortion according to the Wide Sense Stationary (WSS) assumption on the sink which is neither scalable nor realistic. These works consider and capture the spatio-temporal model of physical phenomenon observed by sensor nodes for the realization of advanced efficient communication protocols. However, in these works since the event data is assumed to be Gaussian and wide-sense stationary (WSS) and the most process encountered in practice are non-stationary, these works are not realistic. Moreover, in the literature, there exist some distributed solutions [11], [12] to exploit the spatio-temporal correlation in WSN. However, these studies neither select the sensor nodes which communicate with the sink node nor regulate the data rate to the sink node according to the estimation distortion at the sink node.

On the other hand, the natural Immune System has given the great inspiration for several researches from robotic to network security [13], [14], [15]. Because of its ability to self and non-self discrimination, it has been used for the data clustering and the computer security as an inspiration. In [13], an effective artificial immune system is presented, which is used as a simple classification tool. Using the clonal selection mechanism in Immune System, it attempts to group similar data items according to relationship between them. In [16], the problem of protecting computer systems is addressed as the

1-4244-0463-0/06/$20.00

problem of learning to distinguish self from other. It proposes a method for change detection which is based on the generation of T-cells in the immune system.

In this paper, we propose the Distributed Node and Rate Selection **(DNRS)** method which is based on the principle of natural immune system. Based on the B-cell stimulation model in Immune System, **DNRS** selects the most appropriate sensor nodes that send samples of observed event, are referred to as designated nodes. The aim of the designated node selection is to meet the event estimation distortion constraint at the sink node with the minimum number of sensor node. **DNRS** enables sensor nodes to distributively decide whether it is a designated node or not according to its correlation with its neighbors and the event source. In addition, **DNRS** distributively regulates the reporting frequency rate of each sensor node to provide the event detection reliability with minimum energy consumption. Based on the immune network models, **DNRS** selects the appropriate reporting frequencies of sensor nodes according to the congestion in the forward paths and the event estimation distortion periodically calculated at the sink node by Adaptive LMS Filter.

The remainder of this paper is organized as follows. In Section II, after we briefly introduce the Biological Immune System, we introduce the relationship between Wireless Sensor Networks (WSN) and Immune System. In Section III, we present the **DNRS** method. We firstly introduce the Distributed Node Selection scheme and then, introduce our Distributed Rate Selection mechanism. In Section IV, we evaluate the performance of **DNRS**. Finally, we conclude this paper with Section V.

II. IMMUNE SYSTEM AND WIRELESS SENSOR NETWORKS

In this section, we first briefly introduce the immune system and its basic operation principles. Then, we discuss the similarities and the relation between the Immune System and Wireless Sensor Networks.

A. Biological Immune System

The human immune system is a complex natural defense mechanism. It has the ability to learn about foreign substances (pathogens) that enter the body and to respond to them by producing antibodies that attack the antigens associated with the pathogen [13]. The adaptive immune system is made up of lymphocytes which are white blood cells, B and T cells. Each of B-cells has distinct molecular structure and produces 'Y' shaped antibodies from its surfaces. The antibody recognizes antigen that is foreign material and eliminates it. This antigen-antibody relation is innate immune response [14]. Most antigens have various antigen determinants that are called epitope. In order to grab and latch onto antigen, antibody possesses a structure called paratope. Furthermore, each antibody has a unique epitope called idiotope. Note that different antibody types may have many epitopes in common. An epitope that is unique to a given antibody type is called an idiotope, hence the name idiotypic network for any scheme of regulation that works through the recognition of idiotopes [17]. An antibody type is thought to be stimulated when its paratopes recognize other types, and suppressed if its epitopes are recognized by other antibody paratopes.

The surface of a B-cell contains antibodies for that B-cell. When an antibody for a B-cell binds to an antigen, the B-cell becomes stimulated. The level of B-cell stimulation depends not only on the success of the match to the antigen, but also on how well it matches other B-cells in the immune networks. As the response of the antigen, the stimulated B-cells secrete the antibody to eliminate the antigen. B-cells are also affected by T-cells during immune response. The helper T-cells activate B-cells when the antigen invades it. The suppressor T-cells prevent the activation of B-cell when the antigen was eliminated.

Immune system has the ability to process information, to learn and memorize, to discriminate between self and non-self, and to keep up harmony of the whole system. Because of these abilities, it gives the great inspiration to many kind of researches in computer sciences, robotics and signal processing. Therefore, Artificial Immune System (AIS) models have been introduced. In [17], an AIS model is proposed, which consists of a set of B-cells and the links between those B-cells. Each B-cell object can represent a data item which is being used for learning. According to this model, each B-cell is capable of responding to an antigen specified by a data item when it is stimulated by that antigen specific data item. In the natural immune system, the level of the B-cell stimulation relates that how well its antibody binds to the antigen and its affinity (or enmity) to its neighbors in the network. In [13], B-cell stimulation is modeled in terms of these three influences. The primary stimulus is defined as the affinity between the B-cell and the pathogen i.e. how well they match. The second stimulus for a B-cell is the affinity to its neighbors. Third is the suppression (enmity) factor from the loosely connected neighbor of this B-cell.

The stimulated B-cells start to produce more lymphocytes (i.e. to clone) and to secrete free antibodies. To model the antibody secretion of the stimulated B-cell, In [18], idiotypic network hypothesis is proposed, which is based on mutually stimulus and suppression between antibodies. In [17], an analytical immune network model is proposed in terms of idiotypic network hypothesis introduced in [18]. In [14], to improve the adaptation ability of the system a modified immune network model is proposed by adding helper and suppressor T-cell model. In Immune System, the helper T-cell activates B-cell when the antigen invades it, and the suppressor T-cell prevents the activation of B-cell when the antigen was eliminated. The antigen concentration is kept in the desired level by the antibody concentrations depending on the stimulus value of antibody and T-cells. More specifically, when the stimulus value of antigen $(g_i(t))$ is high and the stimulus value of antibody $(S_i(t))$ is small, the concentration of T-cell $(c_i(t))$ is small. Therefore, in this case $c_i(t)$ take a role of helper T-cell that stimulate B-cell and increases the antibody concentration. On the contrary, the stimulus value of antigen is small and the stimulus value of antibody is big, the $c_i(t)$ is big. So, it take part in suppressor T-cell and decreases the antibody concentration. In this model, the stimulus value of antibody $(S_i(t))$ is expressed by

$$\frac{dS_i(t+1)}{dt} = (\alpha \sum_{j=1}^{N} m_{ij} s_j(t) - \alpha \sum_{k=1}^{N} m_{ki} s_k(t) + \\ + \beta g_i - c_i(t) - k_i) s_i(t) \quad (1)$$

where m_{ij} is mutual stimulus coefficient of antibody i and j, g_i is affinity between antibody i and antigen, k_i is the natural extinction of the antibody i, α, β, η are constants. $s_i(t)$ is the concentration of antibody i at time t and given as follows

$$s_i(t+1) = \frac{1}{1 + exp(0.5 - S_i(t+1))} \quad (2)$$

Furthermore, depending on the $S_i(t)$ and $g_i(t)$, $c_i(t)$ which denotes the concentration of T-cells and controls the concentration of antibody i is expressed by

$$c_i(t) = \eta(1 - g_i(t)) S_i(t) \quad (3)$$

In Section III, we adopt the immune network model given above to derive an effective communication model between sensor nodes and the sink. In the next section, we introduce some analogies between Immune System and Wireless Sensor Networks (WSNs).

B. Immune System Based Sensor Networks

Although Immune System seems different from the Wireless Sensor Networks (WSN), they have the great deal of analogies when considered the tasks they accomplish. For example, when the Immune System encounters a pathogen, some B-cells are stimulated and secrete antibodies in different densities to eliminate the antigens produced by the pathogen. Similarly, when an event occurs in WSN environment, some sensor nodes referred to as designated node (DN) sense the event information and send this information to sink with a reporting frequency to achieve a certain reliability/distortion constraint at the sink.

TABLE I
RELATIONSHIP BETWEEN IMMUNE SYSTEM AND WIRELESS SENSOR NETWORKS (WSN).

Immune System	**Wireless Sensor Networks (WSN)**
B-cells	Sensor nodes
Antibody	Sensor data
Antibody density	Reporting frequency rate f
T-cells	Rate control parameter
Pathogen	Event source
Antigen	Estimation distortion
Natural extinction	Packet loss

In WSN, since multiple sensors record information about a single event in the sensor field, sensor observations are spatially correlated with respect to spatial location of sensor nodes. This results in transmission of highly redundant sensor data which is not necessary to estimate the properties of the event at the sink node. Therefore, it may not be necessary for every sensor node to transmit its data to the sink; instead, a smaller number of DN measurements might be adequate to communicate the event features to the sink within a certain reliability/fidelity level [8]. **DNRS** use the B-cell stimulation model given in [13] to determine the minimum number of DNs while meeting the application specific distortion bound at the sink.

For a fixed number of DNs, the minimum distortion can be achieved by choosing these nodes such that (i) they are located as close to the event source as possible and (ii) are located as farther apart from each other as possible [8]. Similarly, as given in [13], the stimulation of a B-cell depends on the distance between the pathogen and the B-cell and distance between the B-cell and its neighbors which stimulates or suppress it. That is, the B-cells which is nearest to the pathogen and suppressed as least as possible by their neighbors become the stimulated B-cells. The stimulated B-cells start to secrete the antibodies in different densities which are controlled by T-cells to keep the antigen densities in desired level. Similarly, after the selection of DNs they send the event information to sink node with a reporting frequency to achieve the distortion constraint at the sink node. In Table I, we summarize this relationship between Immune System and Wireless Sensor Networks (WSN). According to this relationship, **DNRS** distributively selects the DNs by using the B-cell stimulation model and distributively regulates the reporting frequency of the DNs based on immune network principles.

III. IMMUNE SYSTEM BASED DISTRIBUTED NODE AND RATE SELECTION

In this section, we first introduce the DN selection scheme and then, the reporting frequency regulation mechanism in **DNRS**.

A. Distributed Designated Node Selection

According to the application, the physical phenomenon can be modeled by a point source or by a field source. For example, while the case of object tracking can be modeled by a point source, the case of monitoring magnetic field and seismic activities can be modeled by field source. Here, in order to exploit the spatio-temporal correlation with Immune System based DNRS mechanism, we follow the spatio-temporal correlation framework given in [10]. The sink node is interested in reconstructing the source signal at a specific location (x_0, y_0). Although the reconstruction at the sink node depends on application, the source signal can be modeled as the spatio-temporal process $s(t, x, y)$ where t and (x, y) denote time and spatial coordinates respectively. For reconstruction of the event features, the sink node estimates the event source, S, according to the observations of the sensor nodes, n_i $(i = 1, ..., N)$, in the event area. Each sensor node sends its encoded information $Y_i[n]$ to sink node. Encoded information $Y_i[n]$ is the delayed and noisy version of the event data. The sink decodes the received data $Y_i[n], i = 1...N, n = n_1...n_\tau$ to reconstruct an estimation $\widehat{S}$ of the source S in each time period $\tau = t_{n_\tau} - t_{n_1}$. We give the decoded data at the sink node as $Z_i[n], i = 1...N, n = n_1...n_\tau$. The sink is interested in estimating the expected value of the event during a decision interval τ. Assuming N sensor nodes send information at a rate of f $samples/sec$, this estimation can simply be found by

$$\widehat{S}(\tau, f, N) = \frac{1}{\tau f N} \sum_{i=1}^{N} \sum_{k=1}^{\tau f} Z_i[k] \quad (4)$$

As discussed in Section II, the stimulation of a B-cell depends on three influences. First influence is the affinity between the B-cell and the pathogen. Second influence is the affinity between the B-cell and its neighbors which make it stimulated. Third influence is the affinity between the B-cell and its neighbors which suppress it [13]. According to the B-cell stimulation model given in [13], for a B-cell if the summation of these three influences specified by some equations is greater than a certain threshold value, this B-cell becomes stimulated. In WSN, the minimum distortion can be achieved by choosing the DNs such that (i) they are located as close to the event source as possible and (ii) are located as farther apart from each other as possible [8]. In fact, as B-cell stimulation the selection of DNs depends on the three influences given as follows:

1) First influence is the affinity between the sensor node (B-cell) and event source (pathogen).
2) Second influence is the affinity between the sensor node (B-cell) and its uncorrelated neighbor nodes (stimulating B-cells).
3) Third influence is the affinity between the sensor node (B-cell) and its correlated neighbor nodes (suppressing B-cells).

To determine the correlated and uncorrelated neighbors of a sensor node, we define the application-specific correlation radius r. For a sensor node, while the neighbor nodes in its correlation radius r are the correlated neighbors for this sensor node, the neighbor nodes which are not in correlation radius r are uncorrelated neighbors for this sensor node. Since the spatial correlation between the sensor nodes depends on the variance and the mean of the observed event signal, we assume that correlation radius r is an application specific value, which can be determined based on the statistical probability of the physical phenomena observed in a given sensor network application [8]. We also assume that each sensor node is aware of its position and event source location by means of existing localization techniques [21]. Each sensor knows the relative position of its neighbors, as all nodes periodically broadcasts its position to neighbors.

For the determination of the affinity between the sensor node and its correlated and uncorrelated neighbor nodes, the correlations between the nodes and event source are needed. To express the correlations between event source and the nodes, we use the correlation coefficients, $\rho_{s,i}$ and $\rho_{i,j}$. $\rho_{s,i}$ states the correlation between a node n_i sending information and the event source S. $\rho_{i,j}$ states the correlation between the nodes n_i and n_j. We use the power exponential form [19] to model the correlation coefficients $\rho_{s,i}$ and $\rho_{i,j}$ and give these coefficients as follows.

$$\rho_{s,i} = K_{\vartheta}(d_{s,i}) = e^{(-d_{s,i}/\theta_1)^{\theta_2}}; \theta_1 > 0, \theta_2 \in (0, 2] \quad (5)$$

$$\rho_{i,j} = K_{\vartheta}(d_{i,j}) = e^{(-d_{i,j}/\theta_1)^{\theta_2}}; \theta_1 > 0, \theta_2 \in (0, 2] \quad (6)$$

where $d_{s,i}$ and $d_{i,j}$ denote the distances between event source S and node n_i and between nodes n_i and n_j respectively. For $\theta_2 = 1$, the model becomes exponential, while for $\theta_2 = 2$ squared exponential. The covariance function is assumed to be nonnegative and decrease monotonically with the distance, with limiting values of 1 at $d = 0$ and of 0 at $d = \infty$.

Now, as the B-cell stimulation model given in [13], we mathematically model the DN selection. B-cell stimulation model given in [13] depends on three influences (i) the matching between the pathogen a B-cell (ii) the affinities between the B-cell and its neighbor B-cells (iii) the enmity between the B-cell and its neighbor B-cells. According to this model, when the summation of these three influences exceeds certain threshold for a B-cell, this B-cell becomes stimulated. Using this B-cell stimulation scheme, we model the DN selection depending on the similar three influences as follows:

1) We model the first influence which is the affinity between the sensor node n_i (B-cell) and event source s (pathogen) as $\rho_{s,i}$. Here, $\rho_{s,i}$ indicates the correlation between event source s and sensor node n_i. As the distance between sensor node n_i and event source s decreases, $\rho_{s,i}$ increases. Hence, it is more possible to become a DN for a sensor nodes nearest to the event source.
2) We model the second influence which is the affinity between the sensor node n_i (B-cell) and its uncorrelated neighbor nodes (stimulating B-cells) $n_j, \forall j$ as follows. n_j is selected as the neighbor node which is not in the correlation radius of sensor node n_i.

$$\sum_{j} (1 - \rho_{i,j}) \quad (7)$$

Here, $\rho_{i,j}$ indicates the correlation between sensor node n_i and sensor node n_j. As the distance between n_i and n_j increases, $\rho_{i,j}$ decreases. Therefore, $\sum_j (1 - \rho_{i,j})$ is big for a sensor node having the more uncorrelated neighbors (out of r). Hence, it is more possible to become a DN for such sensor nodes.
3) We model the third influence which is the affinity between the sensor node and its correlated neighbor nodes (in r) $n_k, \forall k$ such that n_k is in the correlation radius of sensor node n_i and give as

$$\sum_{k} (-\rho_{i,k}) \quad (8)$$

where $\rho_{i,k}$ indicates the correlation between sensor node n_i and sensor node n_k. As the distance between n_i and n_k decreases, $\rho_{i,k}$ increases. Therefore, $\sum_k (-\rho_{i,k})$ is small for a sensor node having the more correlated neighbors. Hence, it is the least possible to become a DN for such sensor nodes.

Here, we define the DN selection weight of node n_i (T_i) as the combination of these three influences as follows:

$$T_i = \rho_{s,i} + \sum_j (1 - \rho_{i,j}) + \sum_k (-\rho_{i,k}) \tag{9}$$

When T_i exceeds the certain threshold t_{dns} called the DN selection threshold, sensor node n_i becomes a DN. It is clear that while t_{dns} increases, the number of DN decreases because the number of nodes which exceed t_{dns} decreases. t_{dns} is determined at the sink node by means of distortion periodically computed by the Adaptive LMS Filter. Adaptive LMS Filter is typically applied in environments where signals with unknown or non-stationary statistics are involved and appear for this reason particularly suited to be used in highly dynamic systems such as sensor networks [20]. Using the Adaptive LMS Filter, the sink node predicts the estimated version of the event signal sample $\widehat{S}(\tau_k, f, N)$ given in (4) from the previous samples $\widehat{S}(\tau_{k-r}, f, N)$ $r = 1...p$, where p is the order of the filter. Hence, the estimation error in the Adaptive LMS Filter is expressed by

$$e[\tau_k] = \widehat{S}(\tau_k, f, N) - \bar{\widehat{S}}(\tau_k, f, N) \tag{10}$$

where $\bar{\widehat{S}}(\tau_k, f, N)$ denotes the prediction of $\widehat{S}(\tau_k, f, N)$ with Adaptive LMS Filter. Depending on this error, $e[\tau_k]$, we give the estimation distortion D as follows.

$$D = \frac{1}{M} \sum_{k=1}^{M} e[\tau_k]^2 \tag{11}$$

where M is the number of iteration.

For the event data regions having the temporally correlated characteristic, since $\widehat{S}(\tau_k, f, N)$ can be inferred from the previous samples $\widehat{S}(\tau_{k-r}, f, N)$, $(r = 1...p)$ at the sink node, the estimation distortion of the Adaptive LMS Filter decreases to the certain interval at the sink node. With this decrease, starting from the minimum value of t_{dns} which makes all nodes in the event area DN and hence provides the minimum event estimation distortion, the sink node starts to increase t_{dns} to determine the minimum number of DN.

To increase t_{dns}, the sink node adds a number (u) to t_{dns} in every decision interval (τ_k) and broadcasts the updated designated node selection threshold $(t_{dns}+u)$ to sensor nodes. As the t_{dns} increases in each decision interval τ_k, the number of DN decreases. Unless the distortion level of Adaptive LMS Filter at the sink is increased, the sink continues to increase t_{dns} because the increase in the distortion means that henceforth, the selected DNs can not represent the event data gathering with all sensor nodes in the event area. Therefore, when the distortion is increased, the sink node stops increasing the t_{dns} and it sets t_{dns} to the last updated value of t_{dns} before the increase of the estimation distortion. After the determination of maximum allowable t_{dns}, the sink node broadcasts this t_{dns} to all sensor nodes. Therefore, each sensor node can distributively decides whether it is a DN or not. If $T_i > t_{dns}$, sensor node i is a DN. If $T_i < t_{dns}$, sensor node i is not a DN.

Now, we give outline the entire DN node selection scheme of **DNRS** with the algorithmic manner as follows.

- *Step 1.* When an event occurs, to decide whether it is a DN or not, each sensor node computes its DN selection weight $(T_i, \forall i)$ according to its correlated and uncorrelated neighbor nodes and the distance to the event location.
- *Step 2.* Sink node sets the DN selection threshold t_{dns} as low as possible value such that at the beginning all nodes in the event area become DN $(T_i, > t_{dns}, \forall i)$. This provides the minimum event estimation distortion at the sink node at the beginning.
- *Step 3.* When the temporal correlation in the event data is observed, the estimation distortion D given in (11) decreases to a certain interval.
- *Step 4.* When the sink node observes the decrease on the distortion D, it starts to increase t_{dns} at each decision interval τ_k without disturbing the distortion D which decreases to a certain interval. It broadcasts the updated threshold value to all sensor nodes in each decision interval τ_k. Thus, at each decision interval τ_k, the number of DN decreases because the number of nodes which exceed the updated threshold t_{dns} decreases.
- *Step 5.* Until the distortion which decreases to a certain interval is increased, sink node continues to increase the DN selection threshold t_{dns} and to broadcast the updated t_{dns}.
- *Step 6.* When the distortion is increased at a decision interval, the sink node stops increasing the DN selection threshold t_{dns} at this decision interval. Thus, sink node sets the DN selection threshold t_{dns} to the last updated value of t_{dns} before the increase of the estimation distortion and it broadcasts this t_{dns} to all sensor nodes.
- According to the final t_{dns}, the all sensor nodes distributively decide whether it is a DN not or not as follows.
- *Step 7.* If $T_i > t_{dns}$, sensor node i is a DN.
- *Step 8.* If $T_i < t_{dns}$, sensor node i is not a DN.

Thus, when an event occurs in WSN environment, **DNRS** distributively selects the DNs by using above algorithm. The selected DNs represent all nodes in the event area for this event. However, when the location of the event changes, **DNRS** again employs the DN selection algorithm given above to determine the new DNs. After the determination of the DNs, it is important to regulate their reporting frequency f for exploiting the temporal correlation in the event data. In the next section, we give our distributed frequency rate selection scheme in **DNRS**.

B. Distributed Frequency Rate Selection of Designated Nodes

The reporting frequency of a sensor node f is defined as the number of samples taken and hence packets sent out per unit time by that node for a sensed phenomenon. Hence, the reporting frequency f controls the amount of traffic injected to the sensor field while regulating the number of temporally-correlated samples taken from the phenomenon. This, in turn, affects the observed event distortion, i.e., event detection reliability [8].

According to the event signal characteristics, there exist the upper and lower bound for the estimation distortion given (11).

Since the temporally correlated data can be easily inferred from the previous samples of the data, for the temporally correlated region in the event data the estimation distortion approaches the lower bound. Inversely, while the temporal correlation in the event data decreases, the estimation distortion at the sink node approaches the upper bound. To exploit the temporal correlation in the event data, the appropriate reporting frequency regulation can be given as follows:

- For the case in which the event data is temporally correlated and therefore, the estimation distortion is lower, the reporting frequency should be decreased.
- For the case in which the event data is not temporally correlated and therefore, the estimation distortion is upper, the reporting frequency should be increased.
- Furthermore, since the increase of the reporting frequency results in increasing the contention in the wireless channel, possible congestion in the forward path should also be addressed with the appropriate congestion control mechanism depending on the frequency rate regulation.

As discussed in Section II-A, in human immune system when a pathogen enters a body, some B-cells are stimulated and they start to secrete the antibody with the appropriate density to eliminate the antigens produced by the pathogen. This natural mechanism is modeled with the immune network models given in (1), (2), (3). Similarly, in WSN when an event occurs, some sensor node should be selected as the DNs and they should send the sensed information with the appropriate reporting frequency to sink node to reliably detect the event without the congestion in the forward path. In order to address the reporting frequency regulation and the congestion in the forward path, we exploit the relationship between the immune system and Wireless Sensor Networks (WSN) and we adopt the immune network equations given in (1), (2), (3). According to the relation between Immune System and WSNs, which is summarized in Table I, we give the adaptation as follows:

- We consider the each sensor node as a B-cell which secretes only one kind of antibody.
- We consider the sensor data as the antibodies which are secreted by B-cells.
- In (1), (2) and (3), the stimulus value of antibodies (S_i) and the concentration of T-cells (c_i) are the control parameters which belong to antibody i to control the concentration of antibody i. Therefore, we consider S_i and c_i as the rate control parameter of DN i and denote with $F_i(\tau_k)$ and $c_i(\tau_k)$ respectively. We give $F_i(\tau_k)$ as follows:

$$\frac{dF_i(\tau_{k+1})}{dt} = (\alpha \sum_{j=1}^{K} \gamma f_j(\tau_k) + \beta D - c_i(\tau_k) - L_i(\tau_k)) f_i(\tau_k) \tag{12}$$

where K is the number of DN i neighbors, j at the $k-th$ decision interval, which is a neighbor of DN i, $\alpha, \beta, \eta, \gamma$ are the constants. In Immune system, the antibody concentrations are regulated to kept the antigen concentration in the desired level. Therefore, we consider the estimation distortion D given in (11), which needs to be kept in the desired level, as the antigen concentration g given in (1). We consider the packet loss of DN i at the $k-th$ decision interval $(L_i(\tau_k))$ as the natural extinction of the antibody i (k_i) given in (1). We assume that the packet loss may result from any link error or possible congestion. Depending on the estimation distortion D and $F_i(\tau_k)$, we give $c_i(\tau_k)$ as follows:

$$c_i(\tau_k) = \eta(1 - \beta D)F_i(\tau_k) \tag{13}$$

- We consider the reporting frequency of DN i at the $k-th$ decision interval, $(f_i(\tau_k))$ as the concentration of antibody i (s_i), which is secreted by B-cell i. $f_j(\tau_k)$ denotes the reporting frequency rate of DN j in 12. We give $(f_i(\tau_k))$ as follows:

$$f_i(\tau_{k+1}) = \frac{1}{1 + exp(0.5 - F_i(\tau_{k+1}))} \tag{14}$$

In each decision interval (τ_k), each sensor node distributively evaluates these mechanism given in (12), (14) and (13) by means of the coordination between DNs and sink as follows:

- In each decision interval (τ_k), each DN broadcasts its reporting frequency $(f_i(\tau k))$ to its neighbor DNs.
- In each decision interval τ_k, each DN computes the number of its packet loss $(L_i(\tau_k))$ by the feedback from the sink node.
- D is computed at the sink node by Adaptive LMS Filter and it is broadcasted to all sensor nodes in each decision interval.

Although the Immune System based rate control mechanism given in (12), (14) and (13) appears like very complicated, it provides the useful results to distributively regulate the reporting frequency f of each sensor node. These results can be given as

- While the distortion D at the sink node increases, the control parameters F_i and c_i, $\forall i$ decreases. Therefore, in this case the control parameters c_i, $\forall i$ take a role of increasing the reporting frequency of all DNs to achieve the desired distortion constraint.
- When the distortion D at the sink node decreases, the control parameters F_i and c_i, $\forall i$ increases. Therefore, in this case the control parameters c_i, $\forall i$ take a role of decreasing the reporting frequency of all DNs.
- When the packet loss occurs at DN i packets due to the congestion or the link error, L_i for DN i becomes big and in this case L_i take a role of decreasing the reporting frequency of DN i. This provides the distributed congestion control to each node without disturbing the distortion constraint.

IV. Performance Evaluation

In this section, we give the numerical simulation results of **DNRS** for the DN selection and the regulation of DN reporting frequency. In order to evaluate the performance of **DNRS**, we develop a simulation environment using MATLAB. In this environment, 100 sensor nodes in 50×50 sensor field were randomly positioned. An event was generated with a point

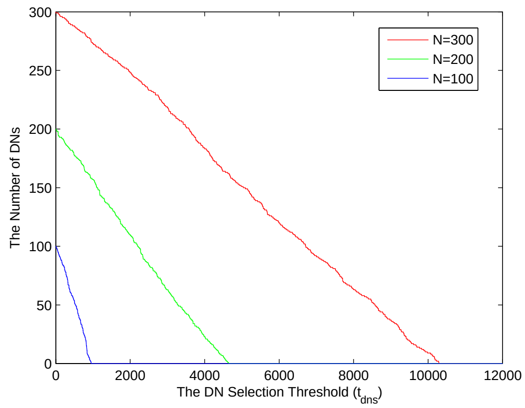

Fig. 1. The designated node selection Threshold t_{dns} vs. the number of designated nodes.

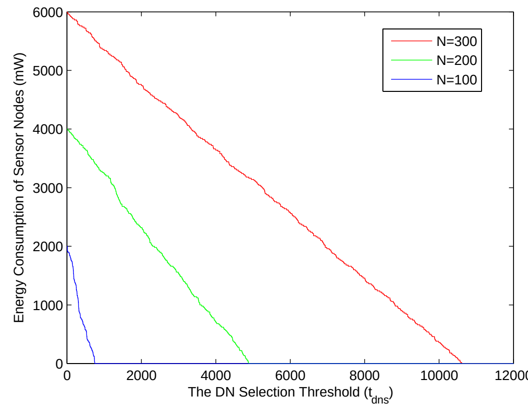

Fig. 2. The designated node selection Threshold t_{dns} vs. the average total energy consumption.

source randomly positioned in this environment. To investigate the energy consumption of the **DNRS**, we assume that on average each DN consumes $20mW$ cumulatively for sensing the event and communicating the observed information to the sink node during the overall event monitoring period. The parameters used in our simulation are given in Table II.

TABLE II
SIMULATION PARAMETERS

Area of sensor field	50×50 m^2
Number of sensor nodes (N)	100-300
Radio range of a sensor node	20 m
Correlation radius r	5 m
Correlation coefficients θ_1	0.3
Correlation coefficients θ_2	0.3

DNRS determines the DN selection threshold t_{dns} at the sink node and the sink node broadcasts t_{dns} to all sensor nodes whereby each sensor node distributively decides whether it is a DN or not. To show the effect of the t_{dns} on the number of DN, we change the t_{dns} from a certain minimum value which make the all nodes the DN to a maximum value which make the number of DNs minimum. As observed in Fig. 1, while the t_{dns} increases, the number of DN decreases. This provides to the sink node to regulate the number of DN by changing the DN selection threshold t_{dns}.

As observed in Fig. 1, the number of DN is regulated by changing the DN selection threshold t_{dns}. The total energy consumption of sensor nodes depends on the DN selection threshold t_{dns}. To show this dependence, we change the t_{dns} in same interval used for last simulation. As observed in Fig. 2, while the t_{dns} increases, the average total energy consumption decreases. According to this result, we state that the sink can regulate the total energy consumption of sensor nodes by dealing with keeping it maximum since maximum value of t_{dns} makes the number of DN minimum.

The DN selection algorithm of **DNRS** find the t_{dns} to distributively determine the DNs by capturing the increase of the estimation distortion computed in each decision interval by the Adaptive LMS Filter at the sink. Therefore, to evaluate

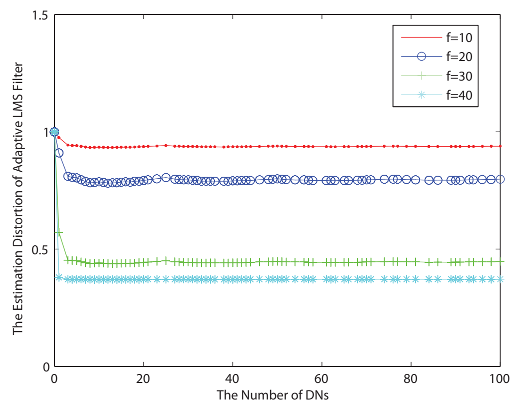

Fig. 3. According to the different reporting frequencies, the number of DN vs. the estimation distortion.

the DN selection algorithm of **DNRS** we run the simulation to observe the estimation distortion D and the number of DN according to the different reporting frequencies from $f = 10$ to $f = 40$. As observed in Fig. 3, the estimation distortion of Adaptive LMS Filter (D) stays relatively constant when the number of designated nodes is increased from 10 to 100 for all reporting frequencies from $f = 10$ to $f = 40$. For this case, we can say that the estimation distortion D increases for the less than 10 DN according to all reporting frequencies and while the reporting frequency increases, the estimation distortion decreases. Therefore, at the point where the estimation distortion is increased, the sink stops increasing the DN selection threshold t_{dns} and sets the maximum allowable t_{dns} providing minimum number of designated nodes without diverging the estimation distortion. Thus, **DNRS** determines the minimum number of DN. This provides the significant energy saving to Wireless Sensor Networks (WSN). For our case in this simulation, since 100 sensor node can be represented with the 10 DN, **DNRS** provides the ninety percent of energy saving.

In addition to the designated node selection, **DNRS** regulates the reporting frequency of the designated nodes to exploit the temporal correlation in the event data. As discussed in Section III-B, temporally correlated event data can be inferred

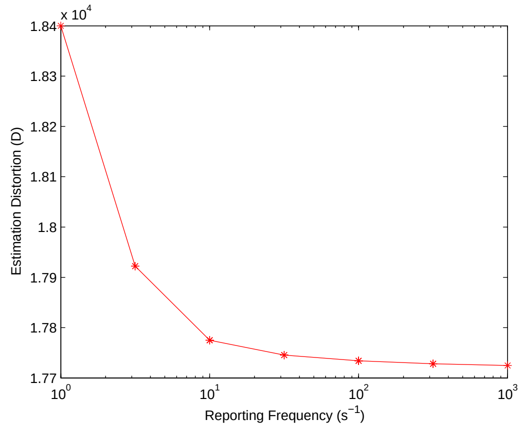

Fig. 4. Estimation distortion of Adaptive LMS Filter for varying normalized reporting frequency.

from its previous sample, the less information transmitted to the sink node by the designated nodes is enough to reliably detect the event properties. Therefore, **DNRS** decreases the reporting frequency of designated nodes when the temporally correlated event data is observed at the sink node. To evaluate the effect of the reporting frequency of designated nodes on the estimation distortion D we increase the normalized reporting frequency f and observe the estimation distortion D for the temporally correlated event data. As observed in Fig. 4, while the reporting frequency increases, the estimation distortion decreases and after the reporting frequency 10^1 estimation distortion D stays relatively constant. Therefore, a significant energy saving can be achieved by selecting small enough f which does not result in the increase of the estimation distortion D.

V. Conclusion

In this paper, we propose our method Distributed Node and Rate Selection **(DNRS)** which based on the principles of natural Immune System. Based on the B-cell stimulation in immune system, **DNRS** selects the most appropriate sensor nodes that send samples of observed event, are referred to as designated node (DN). Using the method, each sensor node can distributively decide whether it is a designated node or not according to its correlation with its neighbors and event source. In addition to the determination of the DNs, to exploit the temporal correlation in the event data based on the immune models, **DNRS** selects the appropriate reporting frequencies of sensor nodes according to the congestion in the forward paths and the event estimation distortion periodically calculated at the sink node by Adaptive LMS Filter. With the selection of the minimum number of designated nodes and the regulation of the reporting frequency of designated nodes, **DNRS** provides the significant energy saving to Wireless Sensor Networks (WSN). **DNRS** is to enable the development of efficient transport protocol which exploits the spatio-temporal correlation in WSN.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, Wireless Sensor Networks: A Survey, Computer Networks (Elsevier) Journal, vol. 38, no. 4, pp. 393-422, March 2002.
[2] J. Kusuma, L. Doherty, K. Ramchandran, Distributed compression for sensor networks, in: Proceedings of the IEEE Image Processing 2001, vol. 1, October 2001, pp. 82-85.
[3] H. Gupta, V. Navda, S. R. Das, V. Chowdhary, Efficient Gathering of Correlated Data in Sensor Networks, in: Proc. ACM MOBIHOC, Urbana-Champaign, Illinois, May 2005, pp. 402-413.
[4] P. V. Rickenbach, R. Wattenhofer, Gathering Correlated Data in Sensor Networks, in: Proc. ACM DIALMPOMC'04, October 2004, Philadelphia, Pennsylvania, USA.
[5] R. Cristescu, B. Beferull-Lozano, M. Vetterli, On network correlated data gathering, in Proc. IEEE INFOCOM 2004, vol. 4, pp. 2571-2582, March 20.
[6] R. Cristescu and M. Vetterli, On the optimal density for real-time data gathering of spatio-temporal processes in sensor networks, in Proc. IPSN 2005, April 2005.
[7] D. Ganesan, R. Cristescu, B. Beferull-Lozano, Power-efficient sensor placement and transmission structure for data gathering under distortion constraints, in Proc. IPSN 2004, pp. 142 - 150, April 2004.
[8] M. C. Vuran, O. B. Akan, and I. F. Akyildiz, Spatio-Temporal Correlation: Theory and Applications for Wireless Sensor Networks, Computer Networks Journal (Elsevier), vol. 45, no. 3, pp. 245-261, June 2004.
[9] M. C. Vuran, I. F. Akyildiz, Spatial Correlation-based Collaborative Medium Access Control in Wireless Sensor Networks, IEEE/ACM Transactions on Networking, June 2006.
[10] M. C. Vuran, O. B. Akan, Spatio-temporal Characteristics of Point and Field Sources in Wireless Sensor Networks, IEEE ICC 2006, İstanbul, June 12-15.
[11] S. S. Pradhan, K. Ramchandran, Distributed source coding using syndromes (DISCUS): design and construction, IEEE Transactions on Information Theory, Volume:49, pp.626-643, March 2003.
[12] C. Guestrin, P. Bodi, R. Thibau, M. Paski, S. Madde, Distributed regression: an efficient framework for modeling sensor network data, IPSN'04, April 26-27, 2004, Berkeley, California USA.
[13] J. Timmis, M. Neal, J. Hunt, An artificial immune system for data analysis, Elsevier, BIOSYSTEMS 2000, 143-150.
[14] J. H. Jun, D. W. Lee, K. B. Sim, Realization of Cooperative Strategies and Swarm Behavior in Distributed Autonomous Robotic Systems using Artificial Immune System, IEEE SMC'99, 614-619.
[15] J. Timmis, M. Neal, J. Hunt, Data analysis using artificial immune systems, cluster analysis and Kohonen networks: some comparisons IEEE SMC'99 Systems Man and Cybernetics.
[16] S. Forrest, A. S. Perelson, L. Allen, R. Cherukuri, Self-Nonself Discrimination in a Computer, 1994 IEEE Symposium on Security and Privacy.
[17] J. D. Farmer, N. H. Packard, A. S. Perelson, The immune system, Adaptation, and Machine Learning, Physica 22D 1986, 187-204, North-Holland, Amsterdam.
[18] N.K. Jerne, Idiotypic Network and Other Preconceived Ideas, Immunological Rev., vol. 79, 5-24, 1984.
[19] J.O. Berger, V. de Oliviera, B. Sanso, Objective bayesian analysis of spatially correlated data, J. Am. Statist. Assoc. 96 2001 1361-1374.
[20] S. Santini, K. Römer, An Adaptive Strategy for Quality-Based Data Reduction in Wireless Sensor Networks, Institute for Pervasive Computing ETH Zurich.
[21] J. Hightower, G. Borriello, Location systems for ubiquitous computing, IEEE Computer, Aug. 2001.