A Heterogeneous Spatial Correlation Model for WSN Applications

deo kumar (✉ deojiet@gmail.com)

Birla Institute of Technology <https://orcid.org/0000-0002-5743-1181>

Sanjeet kumar

Birla Institute of Technology

Research Article

Keywords: Heterogeneous Spatial Correlation, Clustering Algorithm, wireless sensor networks

Posted Date: February 11th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1313847/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

A Heterogeneous Spatial Correlation Model for WSN Applications

Deo Kumar and Sanjeet Kumar
Birla Institute of Technology, Mesra, Ranchi, India
deojiet@gmail.com, sanjeet@bitmesra.ac.in

Abstract— In densely deployed Wireless Sensor Network, a large number of sensor nodes in an event area observes the event information and sends information to the sink. Due to physical properties related to sensed event, the sensor observations are highly spatially correlated with respect to location of sensor nodes. In this paper we have proposed a heterogeneous correlation model that exploits the spatial correlation between each sensor nodes. Heterogeneous WSNs consists of normal nodes and advance nodes with different sensing range. With the help of sensor coverage model and location of sensor nodes, heterogeneous correlation functions are derived to analyze the correlation characteristics of observed information's by sensor nodes. The case studies using heterogeneous correlation functions are performed to study the spatial relationship among sensor nodes. Finally, a correlated clustering algorithm is proposed for efficient medium access and clustering protocol for heterogeneous WSNs.

Keywords—*Heterogeneous Spatial Correlation; Clustering Algorithm; wireless sensor networks*

1 Introduction

The recent advancement in wireless sensor networks consist of low cost, low power enabled sensor nodes, have ability to sense, transmit, receive, and computing the observed data by collaborative monitoring of desired application specific event or environment. A typical sensor node is consists of a battery, sensor, processor, A/D convertor, transceiver and antenna. If sensor nodes are GPS enabled, then they can referred to as Location aware nodes [1].

Usually, the WSNs are densely random deployed and the sense data are spatially correlated because of common sensing area between the sensor nodes. In event-driven WSNs, all nodes in an event area work as a source node and try to transmit their data to sink and the other nodes work as a router node and cooperatively help to send these data to sink [2].

The spatial correlation increases as the distance between the sensor nodes decreases. The highly correlated nodes have more redundant data (i.e., common data between nodes) than the less correlated nodes. One of the common approaches to reduce the redundant data is round wise iterative selection of some nodes within the event area. the iterative selected nodes can be different set of nodes as followed by round are allowed to be active for sensing and other nodes are act as router nodes[2,3].

For example in figure 1, only the three black nodes required as active nodes to detect the event in the event area. Since these three nodes are spatially correlated with their neighbor nodes and remain uncorrelated with each other, so, they are sufficient to detect the event. The remaining all nodes are act as router nodes and they help to send the observed data to the sink. So, a considerable amount of energy can be saved by exploiting the spatial correlation. Hence, taking advantage of spatial correlation an improved energy-efficient communication protocol can be developed.

RELATED WORKS

In WSN, correlation is one of the interested topic /domain among the researchers. Various works on correlation is available in literature.

Mathematical expression for correlated data has been analyzed by many researchers using (a) data correlation model and (b)the common spatial area by two nodes. Data correlation model introduced by Berger et al. [3] .it considered JGRV(Joint Gaussian Random Variable) signal characteristic to derive analytical expression for four different

correlation model(spherical, power exponential, rational quadratic, matérn). Vuran et al [4, 5] introduced and explain the fundamental of Spatio-temporal correlation for wireless sensor networks.

Figure 1: Example of System Model

Data correlation among nodes by considering power exponential correlation model, the system behavior and performance are analyzed. Analytical mathematical equations are derive to find the data correlation among nodes.

The correlation using common area is considered by Shakya et al. [6, 7]. Derive the homogeneous spatial correlation mathematical model by considering the correlation of sense area among the two nodes. These papers are considering only the homogeneous networks where all nodes have equal initial energy and equal sensing and communication range.

Leach protocol is a well-known hierarchical clustering protocol, which uses the probabilistic approach for dynamic clustering. The concept of leach protocol is firstly proposed by Heinzelman et al. [10]. The variety of leach protocols homogeneous as well as heterogeneous network has been proposed by different authors. Younis et al.[12] proposed HEED protocol which uses residual energy for hierarchical clustering to enhance the lifetime of homogeneous WSN. H-HEED [13] which is modified version of HEED protocol which uses multi-level heterogeneity of nodes to enhance the network lifetime. Smaragdakis[8] proposed SEP protocol which is a heterogeneous-aware protocol to enhance the lifetime of network. This protocol uses weighted election probability according to the initial energy of nodes for hierarchical clustering. Qing [9] proposed DEEC protocol which is a heterogeneous-aware protocol to enhance the lifetime of network. This protocol uses weighted election probability according to the residual energy of nodes for hierarchical clustering. In these paper heterogeneous means nodes which have higher capability than a normal node. A heterogeneous node is a node which has higher initial energy than normal nodes but have same sensing and communication range.

In this paper we consider heterogeneous nodes with more constraints. The heterogeneous node have high initial energy as well as more sensing and communication range than normal node.

2 Architecture and correlation model

In WSNs, the sensor observations are highly spatial correlated within the event area. In event area, suppose there are N sensor nodes that can sense the event, some of them are advance node and some of them are normal node.

A. Sensor Deployment Model

Consider a heterogeneous sensor network application where large numbers of sensor nodes are deployed randomly over the monitoring area. An example of random deployed heterogeneous WSN is shown in fig 1. The event area is indicated by Circle; the sensing area of a node is represented by dashed circle; Sink node is represented by Black node; the random event is represented by star. Two types of nodes represented by white node can be clearly visualized. The small white nodes are the normal node and larger white nodes are the advance nodes. The capacity of a node in WSN can be characterized by their sensing range(for sensing the event) and transmission range (for transmitting data to sink or next node).generally, the sensing range of sensor node may be larger or smaller then transmission range depending on type of sensor or application. some reasonable assumption are considered are as follows. (i) The normal nodes and advance nodes follow the Boolean disk coverage model. They have fixed sensing radius and coverage area is represented by a disk centered at the spatial location of node. An event can be detected only if the event occurs within the coverage area of nodes.(ii)the transmission range is much larger than the sensing range. (iii) all nodes are static after deployment w.r. t. their location, so the spatial co-ordinates of each node is known and distance between each node can be calculated.

B. The correlation Model

In this section we derive heterogeneous correlation function to describe the degree of correlation among sensor nodes in the event area. Consider a Boolean disk coverage model with sensing range r (for normal node) and sensing range R (for advance node). Let us consider the two nodes, one is normal node (n_i) and another is advance node (n_j) have spatial location s_i and s_j respectively. The spatial correlation between them is defined by $k_v||s_i-s_j||$ is a function of distance $||s_i-s_j||$ between these two nodes. The spatial correlation represent the correlation property of event information acquired/gathered/collected by these two sensor nodes.

To find the spatial correlation between the sensor nodes, each node calculates the Euclidean distance with other nodes using their spatial co-ordinates/location in the event area. Then a mathematical model can be constructed to compute the spatial correlation among the sensor nodes using their sensing range/radius and distance between them. The correlation among sensor nodes can be estimated by analyzing their observed information's. The covariance among these observed information's can be described as a random matrix between these nodes.

Consider a random event area S where N number of nodes are randomly deployed with spatial locations $\{s_1, s_2, s_3, \dots, s_N\}$. Since a heterogeneous network is chosen, the two types of nodes normal node and advance node are randomly deployed over the event area S . note that the total number of nodes is equal to sum of number of normal nodes and number of advance nodes and the advance nodes are small fraction of total nodes. The observed information's gathered by N sensors are send to sink node and with help of these information sink estimate the source. Let the k normal nodes and $N-k$ advance node are deployed in the event area. The physical phenomena can be modeled as joint Gaussian random variables (JGRV) at each sensor nodes whose mean is given by:

$$E\{S_i\} = m_i, i = 1, 2, \dots, k \quad E\{S_j\} = m_j, i = k + 1, k + 2, \dots, N$$

Variance is given by:

$$\text{var}\{S_i\} = \sigma_{s_n}^2, i = 1, 2, \dots, k$$

$$\text{var}\{S_j\} = \sigma_{s_a}^2, i = k + 1, k + 2, \dots, N$$

$$\text{Correlation is given by: } \text{corr}\{S_i, S_j\} = \frac{\text{cov}[(S_i - m_i)(S_j - m_j)]}{\sqrt{\text{var}\{S_i\} \text{var}\{S_j\}}} = K_v(d_{i,j})$$

Where, $d_{i,j} = \|S_i - S_j\|$ denotes the distance between nodes n_i and n_j having spatial locations S_i and S_j , respectively, and $K_v(.)$ is the correlation coefficient which decreases monotonically with respect to distance has values maximum at $d=0$ and minimum at $d=\text{infinity}$.

3 Proposed correlation model

1) Symbols and notation: To design a heterogeneous mathematical model, consider an advance nodes n_i and a normal node n_j has spatial location/coordinates s_i and s_j respectively. The symbols and notations used in fig. 3 are listed in Table I.

Case 1: if both nodes are different, one is advance node with radius R and other is normal node with radius r . Let there are two nodes S_i, S_j having different transmission range R and r respectively. We assume that the nodes have different radius of transmission range and also they are circular in nature.

Figure 2: Heterogeneous Correlated area

The heterogeneous correlation model: This model is based on the intersection point of two circular sensing regions. The radius of advance node's sensing region is greater than the radius of normal node's sensing region.

Table 1: Symbols and Notations

symbol	Description
S_i	Sensing area of Advance node n_i with radius R
S_j	Sensing area of normal node n_j with radius r
A	Area of sensing region of n_i
$d_{(i,j)}$	Euclidean distance between two nodes n_i and n_j
P_1, P_2	Intersection point of two circular sensing region of S_i and S_j
a	A common chord length connected line segment between two intersection points P_1 and P_2
(x, y)	Co ordinates of intersection point P_1 or P_2
A_i^j	Area of region between the arc P_1P_2 and common chord 'a' (Shadded region)
A_j^i	Area of region between the arc P_1P_2 and common chord 'a'
$K_v(.)$	Spatial correlation coefficient between two nodes n_i and n_j

Using equation of circle:

$$\text{For } S_i, (x-0)^2 + (y-0)^2 = R^2 \Rightarrow x^2 + y^2 = R^2$$

$$\text{For } S_j, (x-d)^2 + (y-0)^2 = r^2 \Rightarrow (x-d)^2 + y^2 = r^2$$

From equation (1) and (2), we get

$$x = \frac{R^2 - r^2 + d^2}{2d}$$

Note: In the fig, all points (x, y) that are 'R' and 'r' distance away from the center of Si and Sj, respectively. Here, the intersection point of two circles Si and Sj are P1 and P2, which have same x- coordinates, so, the point 'O' are 'x' distance from center of Si and (d-x) from center of Sj and OP1 = OP2 = y.

Therefore, using Pythagoras theorem,

$$y^2 = R^2 - x^2 = \frac{4d^2 R^2 - (d^2 - r^2 + R^2)^2}{4d^2} \Rightarrow y = \frac{1}{2d} \sqrt{4d^2 R^2 - (d^2 - r^2 + R^2)^2}$$

$$\text{From figure, we can see that the chord, } a = 2y = \frac{1}{d} \sqrt{4d^2 R^2 - (d^2 - r^2 + R^2)^2}$$

For region Si P1 P3,

$$\text{Using area of sector, } A_{\alpha_1} = \frac{1}{2} R^2 \cdot \alpha_1$$

$$A_{P_1 P_3} = A_{\alpha_1} - \frac{1}{2} y \cdot R = \frac{1}{2} R [R \cdot \alpha_1 - y]$$

Similarly, for region Sj P1 P4,

$$A_{\alpha_2} = \frac{1}{2} r^2 \cdot \alpha_2$$

$$A_{P_1 P_4} = A_{\alpha_2} - \frac{1}{2} y \cdot r = \frac{1}{2} r [r \cdot \alpha_2 - y]$$

Now, the area A_i^j will be,

$$A_i^j = 2A_{P_1 P_3} + 2[\Delta P_1 P_3 O] = R^2 \alpha_1 - yx$$

Similarly,

$$A_j^i = 2A_{P_1 P_4} + 2[\Delta P_1 P_4 O] = r^2 \alpha_2 - dy + xy$$

The spatial correlation can be defined as,

$$K_v(d_{i,j}) = \frac{A_i^j + A_j^i}{A}$$

Where, $K_v(d_{i,j})$ is the decreasing function with distance d, following the limiting value of 1 at $d(i,j) = 0$ and of 0 to $d(i,j) \geq (R+r)$.

Case A: calculating from advance node, condition for existence of intersection point only if $r \leq d_{(i,j)} \leq r+R$. then, Using equation (8) and (9), we can write the above equation as,

$$K_\theta(d_{i,j}) = \frac{R^2\alpha_1 - yx + r^2\alpha_2 - dy + yx}{\pi R^2}$$

$$= \frac{R^2 \cos^{-1}\left(\frac{d^2 + R^2 - r^2}{2dR}\right) + r^2 \cos^{-1}\left(\frac{d^2 - R^2 + r^2}{2dr}\right) - d\sqrt{\frac{4d^2 R^2 - (d^2 - r^2 + R^2)^2}{4d^2}}}{\pi R^2}$$

Where,

$$\alpha_1 = \cos^{-1}\left(\frac{S_i O}{S_i P_1}\right) = \cos^{-1}\left(\frac{x}{R}\right) = \cos^{-1}\left(\frac{R^2 + d^2 - r^2}{2dR}\right)$$

$$\alpha_2 = \cos^{-1}\left(\frac{S_j O}{S_j P_3}\right) = \cos^{-1}\left(\frac{d-x}{r}\right) = \cos^{-1}\left(\frac{d^2 - R^2 + r^2}{2dr}\right)$$

$$y = \frac{1}{2d} \sqrt{4d^2 R^2 - (d^2 - r^2 + R^2)^2}$$

We get,

$$K_v(d_{i,j}) = \frac{R^2 \cos^{-1}\left(\frac{d^2 + R^2 - r^2}{2dR}\right) + r^2 \cos^{-1}\left(\frac{d^2 - R^2 + r^2}{2dr}\right) - \frac{1}{2} \sqrt{4d^2 R^2 - (d^2 - r^2 + R^2)^2}}{\pi R^2} \quad (11)$$

The Heterogeneous Spatial Correlation can be performed using the ratio constraint between the Advance node and Normal node. Let ratio constraint, $n = \frac{R}{r}$

, then, Substituting $R=nr$ in (10) can be rewritten as

$$K_v(d_{i,j}) = \frac{r^2 \left[n^2 \cos^{-1} \left(\frac{d^2 + r^2(n^2 - 1)}{2ndr} \right) + r^2 \cos^{-1} \left(\frac{d^2 + r^2(1 - n^2)}{2dr} \right) \right] - \frac{1}{2} \sqrt{[(r(1-n))^2 - d^2][d^2 - (r(1+n))^2]}}{\pi n^2 r^2} \quad (12)$$

Also, if $0 \leq d < r$, then the normal node's sensing area is coincide into the advance node's sensing area, therefore, $A_i^j + A_j^i = \text{area of normal node} = \pi r^2$. then, the spatial correlation becomes, $K_v(d_{i,j}) = \frac{1}{n^2}$.

Case B: similarly, calculating from normal node, condition for existence of intersection point only if $r \leq d_{(i,j)} \leq r + R$. then, Using equation (8) and (9), we can write the above equation as,

$$K_v(d_{i,j}) = \frac{r^2 \left[n^2 \cos^{-1} \left(\frac{d^2 + r^2(n^2 - 1)}{2ndr} \right) + r^2 \cos^{-1} \left(\frac{d^2 + r^2(1 - n^2)}{2dr} \right) \right] - \frac{1}{2} \sqrt{[(r(1-n))^2 - d^2][d^2 - (r(1+n))^2]}}{\pi r^2} \quad (13)$$

Also, if $0 \leq d < r$, then the normal node's sensing area is coincide into the advance node's sensing area, therefore, $A_i^j + A_j^i = \text{area of normal node} = \pi r^2$. then, the spatial correlation becomes, $K_v(d_{i,j}) = 1$.

Case II: when both nodes are Advance node (i.e., $r=R$), then ratio constraint $n=1$ and for $0 \leq d_{(i,j)} \leq 2R$, equation (12) can be rewritten as

$$k_v(d_{i,j}) = \frac{2 \cos^{-1} \left(\frac{d}{2R} \right)}{\pi} - \frac{d}{\pi R^2} \sqrt{R^2 - \frac{d^2}{4}} \quad \text{for } 0 \leq d_{(i,j)} \leq 2R$$

$$= 0 \quad \text{for } d > 2R$$

Case III: when both nodes are normal node (i.e., $R=r$), then ratio constraint $n=1$ and for $0 \leq d_{(i,j)} \leq 2r$, equation (13) can be rewritten as

$$k_v(d_{i,j}) = \frac{2 \cos^{-1} \left(\frac{d}{2r} \right)}{\pi} - \frac{d}{\pi r^2} \sqrt{r^2 - \frac{d^2}{4}} \quad \text{for } 0 \leq d_{(i,j)} \leq 2r$$

$$= 0 \quad \text{for } d > 2r$$

Note that equation (14) is similar as sakya[7].

If correlation coefficient is equal to 1 means the two nodes are highly correlated and if correlation coefficient is equal to 0 means these nodes have no correlated data.

4 System model

E_0 represent the initial energy of normal nodes. $E_0(1+\alpha)$ represent the energy of Advance nodes. $r=20m$ is the sensing range of normal nodes. $R=40 m$ is the sensing range of Advance nodes. m is the percentage of nodes act as Advance node. $P_{opt}=0.1$.

Total energy of network = $nE_0(1-m) + nmE_0(1+\alpha) = n E_0 (1+\alpha m)$

- The weighted probability

$$P_N = \frac{P_{opt} E_i(r) SCavg(i)}{(1 + m \alpha) \bar{E}(r)} \quad \text{for normal nodes}$$

$$P_A = \frac{P_{opt}(1 + \alpha) E_i(r) SCavg}{(1 + m \alpha) \bar{E}(r)} \quad \text{for Advance nodes}$$

Where, $SCavg(i)$ is the average Spatial correlation of i th node

$$\text{Threshold for CH selection } T(S_i) = \frac{P_i}{1 - P_i (r \bmod \frac{1}{P_i})}$$

Proposed clustering algorithm

```

begin
for r=1 to rmax do
    CH==0
    calculate p(N), p (A)
    calculate T(N) and T(A)
    CH= CH+1
    for i=1 to n do
        update Eres for each nodes
    if CH(i).E> TH(i)
        then retain as CH and cluster members are associated with them for next round.
    else
        participate in CH formation in next round
    end if
    end for
end for
end

```

5 Results and Discussion

Figure 3: Network life time ($m=0.2$ and $\alpha=3$)

- The simulation result shows that the proposed scheme outperforms SEP and DEEC protocols with an improvement of 150% in lifetime (TND).
- No of Packets received at Base station of the proposed scheme is also improved.

Figure 4: Data packet received ($m=0.2$ and $\alpha=3$)

The above section describes one approach to extract the correlation characteristics of the observed information's of nodes using the sensing area of nodes. The proposed model is a generalized correlation model which can be applicable for homogeneous as well as heterogeneous Wireless Sensor Network. This model follows the traditional Omni-directional model for simplicity.

To fulfill the requirement of application of WSNs, the sensor nodes consist of temperature sensor, humidity sensor, image sensor, bio sensor, level sensor, pressure sensor etc. All sensor nodes follows the disk coverage model means they can sense the event within a certain radius around 360° from the center of the nodes. The proposed correlation model consider the real network conditions of WSNs by deploying two different sensor nodes which have different sensing range as well as initial energy.

In addition, the proposed correlation model can useful for designing an energy efficient communication protocol. For example, the correlation function can also be calculated for three or more types of sensor nodes deployed in event area which have different sensing range as well as initial energy with each other. This correlation function can compared with the correlation function used by vurun[4] and shakya[7] to design an energy efficient heterogeneous MAC protocol.

Declarations:

Funding: No funding received for this research work

Conflicts of interest/ computing interests: There is no conflict of interests regarding the publication of this paper.

Availability of data and material: All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Code availability: The authors have no financial or proprietary interests in any material discussed in this article.

Author's contributions: 1st author are involved in the study conception, design, data collection, analysis and interpretation of results. 2nd author is project guide and involve in manuscript preparation. All authors reviewed the results and approved the final version of the manuscript.

Ethics approval: This paper does not contain any studies with human participants or animals performed by any of the authors.

Consent to participate: The authors have no any objection.

References

1. I. F. Akyildiz, W. Su, Y.Sankarasubramaniam, and E.Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, 2002.
2. J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer Networks*, vol. 52, no. 12, pp. 2292–2330, 2008.
3. J. O. Berger, V. de Oliveira, and B. Sanso, Objective bayesian analysis of spatially correlated data," *J. American Statistical Assoc.* 96, 1361-1374, 2001.
4. M.C. Vuran ,O.B. Akan, I.F. Akyildiz," Spatio-temporal correlation: theory and applications for wireless sensor networks", *Comput. Netw.J.(Elsevier)*,2004,45,pp.245-256.
5. M.C. Vuran ,O.B. Akan, I.F. Akyildiz," Spatio correlation-based collaborative medium access control in wireless sensor networks", *IEEE/ACM Trans. Netw.* 2006, 14, (2),pp. 316-329.
6. R.K. Shakya, Y.N.Singh, N. k. Verma," A correlation model for Mac protocols in event –driven wireless sensor networks",*Proc. 2012 IEEE Region 10 Conf.,(TENCON 2012)*, november 2012,pp. 1-6.
7. R.K. Shakya, Y.N.Singh, N. k. Verma," Generic correlation model for wireless sensor network applications", *IET wirel.sens. Syst.*,2013, vol3,issue 4,pp.266-276.
8. G. Smaragdakis, I. Matta, A. Bestavros, "SEP: A stable election protocol for clustered heterogeneous wireless sensor networks," 2004. second International workshop on sensor and actor network protocol and application(SANPA).
9. L. Qing, Q. Zhu, and M. Wang, "Design of a distributed energy efficient clustering algorithm for heterogeneous wireless sensor networks," *Computer Comm.*, vol 29, no 12, pp. 2230-2237, 2006.
10. Heinzelman, W. B., Chandrakasan, A. P., Balakrishnan, H.: Energy-efficient communication protocol for wireless microsensor networks. *Proc. Of 33rd Annual Hawaii Int. Conf. on System sciences*, Maui, HI, USA, vol. 2, pp. 1-10 (2000).
11. Heinzelman, W. B., Chandrakasan, A. P., Balakrishnan, H.: An Application-Specific Protocol Architecture for Wireless Microsensor Networks. *IEEE Trans. on Wireless Communications*, 1(4), (2002).
12. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient distributed clustering approach for ad hoc sensor networks," in *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 366-379. Oct-Dec.
13. Chand, S., Singh, S. & Kumar, B. Heterogeneous HEED Protocol for Wireless Sensor Networks. *Wireless Pers. Commun.* 77, 2117–2139 (2014).