SIR Epidemic Model derived from Spatial Correlation for Worm Propagation in Event-driven Wireless Sensor Network

Sunil K Chaudhary, Rajeev K Shakya, Atul Kumar

Abstract: An event or a physical activity triggers many nodes for sensing the environment in wireless sensor networks for event monitoring applications. These triggered nodes then transmit the sensed data to the control center. To study the virus propagation behavior in event-driven WSNs, sensor coverage which is a quality of service parameter, can be considered in epidemic model design to get more insight. Existing epidemic models have global characteristics when it is considered a realistic behavior of WSNs. In this paper, a comparative analysis is carried out for spatial correlation feature found in sensor nodes based on sensing range. It is an extended version of our recent work. We present analysis using Susceptible-Infectious-Recovered (SIR) epidemic model with and without spatial correlation feature. Firstly, we show that the correlated nodes are occurred due to sensing ability for event monitoring applications. These nodes are responsible for transmitting correlated information when an event occurs in the field. A detailed comparative analysis with recent existing SIR epidemic models is presented with results and discussions. A comparison based on basic reproduction number is also discussed. Our Results show the impact of spatial correlation in the behavior of virus spread with time. Comparative study shows the effective use of our model is in designing prevention mechanisms for infection control. It can also be used to study the virus spreading behavior for event-driven scenarios in WSNs.

Keywords — Spatial Correlation; Comparative study, Epidemic model design analysis; Worm propagation

I. INTRODUCTION

In wireless sensor network (WSN), many nodes are distributed in unattended environment to observe physical conditions for example, temperature, sound, vibration, pressure etc. Sensor nodes work cooperatively to route the data to control centre or distant node called sink node or base station. Due to highly dynamic environment conditions, the sensing coverage (sensing range) and the network connectivity (communication range) are limited in WSNs. So, the densely deployment of nodes are necessary to provide the fine grain monitoring of the environment. It achieves sufficiently high level of data resolution when the information is observed by the sensor nodes. But, it leads to redundant data transmission collected at the sink. WSN system is mostly event-driven system where many to one communication paradigm is formed when an event is monitored by the sensor nodes. When an event occurs in the sensor field, many nearby nodes get activated for sensing and transmission. A common sensing area is formed by all the activated nodes covering the event sensing range. As a resultant, the spatial correlation exists based on common sensing area between the nodes [17, 18, 12]. A sensor node has processing unit, sensing unit, communication unit, and battery power unit. It can have sensors with sensing abilities within predefined range only. After sensing, the nodes create the packets or the reports to transmit over wireless channel. These reports or packets are routed through multi-hop communications towards the destination (i.e., Sink node). Since most of time there is no traffic due to the absence of physical activity, the nodes operate in sleep and wakeup mode. It saves the energy of nodes. In event-driven system, the nodes gather the information only when an event or physical activity is detected by them, otherwise they switch to sleep mode. In monitoring applications of WSN, spatial and the temporal correlation exist. Only spatial correlation tends to increase with increased value of inter-node distance (Euclidean distance) [12, 13, 23]. The correlation property in a densely deployed WSN can be used to study the worm or virus propagation behaviour in sensor networks. Alternatively WSNs have low defensive capabilities, so it can be easily targeted by the attacks. The security algorithms cannot be implemented due to limited memory and processing capabilities. For example, the malicious codes like a Cabir worm [6] is malicious codes that can affect sensor node’s operation in WSN. Also, blue-tooth-enabled devices are easily infected when comes in its proximity. Similar kind of Cabir warm is Mabir warm [7]. With the aim of study of these worms in WSN, for this, epidemiological models are recently used to predict the impact of worm spreading behaviour [9, 10].

According to epidemiological, different compartments such as susceptible, infectious, recovered classes etc are made from complete population of nodes in the network [2]. A group of healthy nodes or virus-free nodes are called Susceptible nodes. When a susceptible node gets infection from a virus, it is infectious node. These types of nodes are in infectious group. Infection increases using data communication into its neighbouring nodes. An infective node infects to its neighbours when it communicates to it.
Likewise, these infected neighbours will pass infections to others. Hence, virus propagates in whole network by one-by-one. We consider densely deployed node distribution, there is higher population. The virus spread behaviour tends to be exponential growth when infection time from one node to others is short. In literature, there many epidemiological models such as susceptible-infectious-susceptible (SIS) models, susceptible-infectious-recovered (SIR) models, SEIR models, SEIQR models, SEIRS-V models (with vaccinated class) etc [7, 2, 5, 3]. The most of these models are derived considering global behaviour of nodes in a network. These are not having consideration of correlation characteristics of WSNs that is important feature of sensor networks. There are also some feature of sensor network such as energy consumption of node, event activation status, transmission range of node etc. Recently, Tang [8] describes the modelling based on sleep-wakeup feature of nodes in WSNs. During sleep mode of the nodes, infected nodes can be cured using anti-virus scheme as proposed by Tang. An improved version of this model (i.e., SI model) has been derived that uses some protection mechanisms to control the further infection and also ensure the prevention around some region of the network. [13] 

In this paper, studies on dynamics based on the constraint of geometrical and spatial correlation are carried out. A detailed comparative analysis is presented. It is found in our recent work [23] that spatial correlation distance has impact on worm propagation. Based on this, a comparative analysis is presented with recent SIR epidemic models published in literature of virus or worm propagation in WSNs. Using the reasonable parameters of WSNs, we drive the dynamic of virus propagation.

The rest of the paper is organized as follows. Section II presents the basic theory on epidemic models and feature of wireless sensor network. In section III, spatial correlation aspect and its impact is described. To understand its impact, comparative analysis is presented with existing models. Results and discussions are highlighted in section IV. Finally, conclusion is given in section V. 

The sensor network features with existing epidemic models are described in section 2. In section 3, the proposed correlation based SIR Model is discussed with stability analysis. In this section, the calculation of the basic reproduction number and the threshold value based on spatial correlation are also discussed. Simulation results and analysis are presented with comparison in section 4. Finally, section 5 presents the conclusions.

II. EPIDEMIC THEORY AND WIRELESS SENSOR NETWORK

Epidemic theory includes the compartmental based design in terms of differential equations used to study the dynamics with time when the worm or virus is propagated [9, 2]. It is derived from investigated behaviour of infectious diseases (virus, bacteria etc.)

The dynamics with time are calculated by solving the differential equations that shows the system stability in terms of mathematical way. In research work [9, 15, 2, 14, 16], various models are proposed to determine the stability analysis such as SI model, SIRS model, SEIR etc. These include various network features for investigation. For example, SIRS model in [15] discussed the investigation of stability about communication radius. It describes reproduction number and threshold parameters in his study. Similarly, in [21, 22], stability analysis by taking the position for dysfunctional and IoT worm attack. These all models rely on global

Fig. 1. Spatial correlation between nodes $n_i$ and $n_j$ at their locations $s_i$ and $s_j$. 

Fig. 2. Correlated Clusters for $\xi = 0.25$ for $N = 60$, $rs = 20$. 

Behaviour of wireless sensor network as general deployed sensor nodes communications. It does not consider the local interactions of nodes, their remaining energy, their consumption cost etc [15, 19, 20].

In this paper, spatial correlation characteristics of sensor nodes is studied in the design the improved SIR model and its comparative analysis. In our analysis, susceptible-infectious recovered model is compared with modified version of SIR by Tang and Li [13]. In their model, sleep and wakeup mode of nodes are used in design. It shows epidemic model very much suited with our design given in [23].

A. Spatial Correlation in WSN

In our earlier work in [17, 18, 12], we have shown that the estimation of spatial correlation among nodes can be done using correlation coefficient parameter. Assume sink node knows the locations of nodes based on sensing range $r_s$, then correlation coefficient denoted by $\rho_{(i,j)}$ with distance $d_{(i,j)}$ given...
between nodes \( n_i \) and \( n_j \). The expression for correlation coefficient is given by Eq. (1).

\[
\rho_{ij} = \cos^{-1}\left(\frac{d_{ij}^2}{2r_s^2}\right) - \frac{d_{ij}}{4r_s} \sqrt{\left(4r_s^2 - d_{ij}^2\right)}
\]

For \( 0 \leq d_{ij} < 2r_s \),

\[
\rho_{ij} = 0
\]

To get more insight in spatial correlation between nodes, the results are produced using correlation coefficient parameter in Fig. 1, where sub-region is a set of points formed by overlapping correlation with strong effect. The line connected between the dotted nodes indicates the amount of correlation (i.e., how strong correlation impact with neighboring nodes). If there is strong correlation, then connected line is shorter, otherwise it is longer indicating weakly correlation. It also shows the fraction of overlapped sensing area of \( r_s \)-radius nodes centred at position of itself. So, correlation \( \rho_{ij} \) gives overlapped fraction of sensing area. For given density of nodes and amount of nodes number, a new parameter known as correlation threshold \( \xi \) (0 < \( \xi \) < 1) is defined.

**Table 1: Test results of produced correlated clusters \( r_{cc} \) for different values of \( r_s \) and \( \beta \) [12]**

<table>
<thead>
<tr>
<th>( \xi )</th>
<th>( r_s ) (m)</th>
<th>( r_{cc} ) (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2</td>
<td>5.780</td>
<td>4.110</td>
</tr>
<tr>
<td>0.4</td>
<td>8.030</td>
<td>5.500</td>
</tr>
<tr>
<td>0.5</td>
<td>10.040</td>
<td>6.230</td>
</tr>
<tr>
<td>0.6</td>
<td>12.070</td>
<td>7.050</td>
</tr>
<tr>
<td>0.8</td>
<td>14.000</td>
<td>9.600</td>
</tr>
</tbody>
</table>

The different values of \( r_s \) are shown with different values of \( \xi \) is given in Table 1. Using our design in previous paper, it is observed that if the overlap area between nodes is larger, spatial correlation will be strong [12]. If \( \rho_{ij} \geq \xi \), then there is strong correlation between node \( n_i \) and \( n_j \). If \( \rho_{ij} < \xi \), then weakly correlation between node \( n_i \) and \( n_j \). For defined threshold value \( \xi \), the correlated cluster size can be determined (denoted by \( r_{cc} \)) for \( d_{ij} = r_{cc} \) . By simplification, we get

\[
\xi = \cos^{-1}\left(\frac{r_{cc}}{2r_s}\right) - \frac{r_{cc}}{4r_s} \sqrt{\left(4r_s^2 - r_{cc}^2\right)}
\]

The Eq. (2) gives the relationship between \( \xi \) and \( r_{cc} \). This important equation also shows node separation property with sensing range and correlation.

To get more insight about spatial correlation in WSNs, 150 nodes are distributed randomly over the \( 180 \times 180 \) area with sensing range \( r_s = 20 \). The simulation results are revealed in Fig. 2 and Fig. 3 for different values of correlation threshold (i.e., \( \xi \)). It can be observed from these results that the different sized clusters can be formed with random node distribution. This type of behavior of node distribution can show different behavior of virus propagation. Based on the observations from results using spatial correlation, Table 1 gives simulated results for different values of parameters. By the motivations from these observations, we incorporate the characteristics of spatial correlation in existing epidemic models to study the virus spread dynamics in WSNs. By considering \( r_s \) and \( r_{cc} \), modified epidemic models are discussed in next section.

### III. COMPARATIVE STUDY WITH MODIFIED SIR MODEL

#### A. SIR model: without Spatial Correlation

According to Feng et al. [15], SIR model for WSNs has considered the node distribution and communication range. The following differential equations are derived by the Feng et al. [15] for distributed node in \( L \times L \) rectangular area.

\[
\frac{dS(t)}{dt} = \mu N - \beta \frac{\pi r_s^2}{L^2} S(t) I(t) - (\mu + \omega)S(t) + \epsilon R(t)
\]

\[
\frac{dI(t)}{dt} = \beta \frac{\pi r_s^2}{L^2} S(t) I(t) - (\mu + \gamma)I(t)
\]

\[
\frac{dR(t)}{dt} = \omega S(t) + \gamma I(t) - (\mu + \epsilon)R(t)
\]

In Eq. (3), infection possibility of a susceptible node (S) is \( b \). It is then become part of Infectious node I, so there is transition from S to I. By some anti-virus mechanism, I transitioned nodes go to Recovered state (R) with possibility \( g \). Nodes in R becomes healthy. It is then transitioned into S state with possibility from R state. Nodes in susceptible node (S) is transitioned into R with possibility w. For system equations (3), the basic reproduction number is given to achieve endemic equilibrium when \( R_0 > 1 \). It is given by Feng et al. [15] as

\[
R_0 = \frac{N\beta \pi r_s^2 (\mu + \epsilon)}{(\mu + \gamma)(\mu + \epsilon + \omega)}
\]

It can be observed that the basic reproduction number depends on given node distributed area, transmission range and rates of transitions from different compartments (i.e., S, I, R). Next, we consider the spatial correlation to investigate the impact on virus dynamics. A modified version is presented in next subsection.
SIR Epidemic Model derived from Spatial Correlation for Worm Propagation in Event-driven Wireless Sensor Network

B. Modified SIR model: with Spatial Correlation

Consider a node gets infected and hence it triggers a false event. In $L \times L$ sized monitoring area, the $r_e$-radius sized event area is formed by the false event as shown in Fig. 4. Assuming the number of susceptible sensor nodes inside the event area will get infection at time $t$. Here, each sensor node has $r_s$ sensing range and $r_t$ transmission range. From Eq. (2), the fraction of effective area will be $\frac{r_e^2}{r_s^2}$ for correlated nodes inside the event area when a virus attack is initiated as an event in sensor field. Using spatial correlation, we derive the modified SIR model.

$$S_e(t) = \left( \frac{\pi r_e^2}{\pi r_s^2} \right) S(t), \quad \text{(5)}$$

Assuming $\phi$ as given by following equation.

$$\phi = \left( \frac{\pi r_e^2}{\pi r_s^2} \right). \quad \text{(6)}$$

From Eq. (3), the modified system equations is given as

$$\frac{dS(t)}{dt} = -\mu S(t) - \phi S(t)I(t) - (\mu + \omega)S(t) + \epsilon R(t)$$

$$\frac{dI(t)}{dt} = \phi S(t)I(t) - (\mu + \gamma)I(t) \quad \text{(7)}$$

$$\frac{dR(t)}{dt} = \omega S(t) + \gamma I(t) - (\mu + \epsilon)R(t)$$

It can be observed from Eq. (7) that the virus propagation dynamics depends on event range and correlated regions of size $r_e^2$ of node based on spatial correlation between nodes. For system equations (7), the basic reproduction number based on spatial correlation is denoted by $R_{0cc}$ to achieve endemic equilibrium when $R_{0cc} > 1$. It is given as

$$R_{0cc} = \frac{N\beta(\mu + \epsilon)(\pi r_e^2 r_s^2 - L^2 r_{cc}^2)}{L^2 r_e^2(\mu + \gamma)(\mu + \epsilon + \omega)} \quad \text{(8)}$$

In next section, we investigate the impact of spatial correlation based parameters on virus spread dynamics in WSNs. The $r_e$ sensing range is directly related to $r_t$ transmission range. Typically, it is found to be as $r_e \geq 2r_t$.

IV. SIMULATION RESULTS AND ANALYSIS

In this section, impact of spatial correlation for WSNs and the dynamic characteristics of virus propagation is studied with simulation in MATLAB. We simulate different models with ODE function. Results are produced by changing various parameters for given initial values as $R(0) = 0$; $S(0) = 100$; $I(0) = 1$; $\mu = 0.0006$; $\epsilon = 0.06$; $\omega = 0.001$; $\gamma = 0.13$; $\beta = 0.01$. Feng et. al [15] considers the two different parameters such as transmission range $r_e$ and distributed node density $\rho$ for performance evaluation. We compare the existing SIR with our modified SIR in terms of transmission range and distributed node density while considering the different spatial correlation parameters using our system equation. For $I$, $S$, time evolution in the absence and presence of spatial correlation is plotted as given below. Fig. 5 shows the time evolution of the fraction of infected nodes in the absence and presence of spatial correlation with different transmission range $r_e = 5$, $r_e = 6$, $r_e = 8$, $r_e = 9$ considering the variation in spatial correlation parameters (i.e., $r_e$, $r_{cc}$). It is observed from all the four figures that the rate of infective nodes depends on change in transmission range. Because the spread of virus can only be possible by data communication between nodes in the field. When one node gets infected then it infects others using data communication based on transmission range. It can also be seen clearly from all four figures that there is considerable amount of impact on worm spreading dynamics in event-driven WSN. By changing the value of event sensing area (i.e., $r_e$) and correlated clusters (i.e., $r_{cc}$), time evaluation dynamics changes significantly.

Fig. 6 is the plot for time evolution of the fraction of susceptible nodes in the absence and presence of spatial correlation with SIR model. As discussed for the infectious node’s patterns with time in previous figure, the behavior of susceptible nodes with time is also according to different values of spatial correlation parameters.
Fig. 5. Time evolution of the fraction of Infectious nodes for 200 sensors with $t = 35; L = 10$ (a) $r_i = 5$ (b) $r_i = 6$ (c) $r_i = 8$ (d) $r_i = 9$ with different spatial correlation parameters, $r_e, r_{cc}$.

The rate of the susceptible nodes goes to minimum for SIR with spatial correlation whereas it is much larger for SIR without spatial correlation.

Fig. 6. Time evolution of the fraction of Susceptible nodes for 200 sensors with $t = 35; L = 10$; (a) $r_i = 5$ (b) $r_i = 6$ (c) $r_i = 8$ (d) $r_i = 9$ with different spatial correlation parameters, $r_e, r_{cc}$. 

SIR Epidemic Model derived from Spatial Correlation for Worm Propagation in Event-driven Wireless Sensor Network

Fig. 7. Time evolution of the fraction of Infectious nodes for 200 sensors with \( t = 35; L = 10 \) (a) \( \rho = 1 \) with different spatial correlation parameters, \( r_e, r_{cc} \); (b) \( \rho = 5 \) with different spatial correlation parameters, \( r_e, r_{cc} \).

Fig. 8. Time evolution of the fraction of Susceptible nodes for 200 sensors with \( t = 35; L = 10 \) (a) \( \rho = 1 \) with different spatial correlation parameters, \( r_e, r_{cc} \); (b) \( \rho = 5 \) with different spatial correlation parameters, \( r_e, r_{cc} \).

Fig. 7 is the graph of infective nodes when the total number of distributed nodes increases (i.e., variation in node density \( \rho \)). When the number of nodes increases with increase in the node density, more neighbors interact with an infective node. It also increases the correlated clusters with large number of members in a correlated cluster based on spatial correlation. It can be seen from Fig. 7(b) that increased node density indicates the frequent worm infection of nodes with respect to time. A lower values of \( r_e \) and \( r_{cc} \) can help in lesser variation of infection among nodes.

For similar settings of \( I \), Fig. 8 presents the graph of susceptible node variations with time. To get more insight in the behaviour of virus spreading dynamics, a comparison on calculated basic reproduction number (R0) with and without spatial correlation is shown in Table 2 for above graphs. The endemic equilibrium condition is obtained by \( R_0 \) to determine whether stability is achieved or not. Table indicates the lower reproduction number when spatial correlation is applied on existing SIR model. Thus the performance of SIR model can be improved using spatial correlation property of sensor nodes. These spatial correlation parameters are obtained from realistic scenario of sensor network.

V. CONCLUSIONS

In this work, a modified SIR model is proposed based on spatial correlation characteristics of WSNs. The existing SIR model given by Feng et. al [15] is used to integrate the correlation feature. SIR model by Feng et. al [15] does not consider the realistic parameters of sensor nodes (say sensing range). Firstly, the paper describes the characteristics of spatial correlation in two dimensional graph networks. It is then shown that how correlation exists with different sensing range. Our comparative study includes the performance analysis with existing SIR in terms of many parameters such as sensing range, event sensing area, correlated clustering.
Table 2: Comparison with basic Reproduction Number 

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Basic Reproduction Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_0 = 5$, $r_e = 5$, $R_{cc} = 0.6$</td>
<td>1183</td>
</tr>
<tr>
<td>$r = 5$, $r_e = 6$, $r_{cc} = 1$</td>
<td>1183</td>
</tr>
<tr>
<td>$r = 9$, $r_e = 3.83$, $R_{cc} = 0.9$</td>
<td>3834</td>
</tr>
<tr>
<td>$r = 9$, $r_e = 3.45$, $R_{cc} = 0.8$</td>
<td>3834</td>
</tr>
<tr>
<td>$r = 3$, $r_e = 3.77$, $R_{cc} = 0.75$</td>
<td>5750</td>
</tr>
<tr>
<td>$r = 3$, $r_e = 4.1$, $R_{cc} = 0.89$</td>
<td>5750</td>
</tr>
<tr>
<td>$r = 5$, $r_e = 1.55$, $R_{cc} = 0.35$</td>
<td>9584</td>
</tr>
<tr>
<td>$r = 5$, $r_e = 3.79$, $R_{cc} = 1.0$</td>
<td>9584</td>
</tr>
</tbody>
</table>

REFERENCES


