Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2012, Article ID 656231, 9 pages
doi:10.1155/2012/656231

Research Article

Modelling and Performance Analysis of a Network of Chemical Sensors with Dynamic Collaboration

Alex Skvortsov[1] and Branko Ristic[2]

[1] *HPP Division, Defence Science and Technology Organisation, 506 Lorimer Street, Fishermans Bend, VIC 3207, Australia*
[2] *ISR Division, Defence Science and Technology Organisation, 506 Lorimer Street, Fishermans Bend, VIC 3207, Australia*

Correspondence should be addressed to Alex Skvortsov, alex.skvortsov@dsto.defence.gov.au

Received 14 July 2011; Revised 4 November 2011; Accepted 4 November 2011

Academic Editor: Frank Ehlers

The problem of environmental monitoring using a wireless network of chemical sensors with a limited energy supply is considered. Since the conventional chemical sensors in active mode consume vast amounts of energy, an optimisation problem arises in the context of a balance between the energy consumption and the detection capabilities of such a network. A protocol based on "dynamic sensor collaboration" is employed: in the absence of any pollutant, the majority of sensors are in the sleep (passive) mode; a sensor is invoked (activated) by wake-up messages from its neighbors only when more information is required. The paper proposes a mathematical model of a network of chemical sensors using this protocol. The model provides valuable insights into the network behavior and near optimal capacity design (energy consumption against detection). An analytical model of the environment, using turbulent mixing to capture chaotic fluctuations, intermittency, and nonhomogeneity of the pollutant distribution, is employed in the study. A binary model of a chemical sensor is assumed (a device with threshold detection). The outcome of the study is a set of simple analytical tools for sensor network design, optimisation, and performance analysis.

1. Introduction

Development of wireless sensor network (WSN) for a particular operation scenario is a complex scientific and technical problem [1, 2]. Very often this complexity resides in establishing a balance between the peak performances of the WSN prescribed by the operational requirements (e.g., minimal detection threshold, size of surveillance region, detection time, rate of false negatives, etc.) and various resource constrains (e.g., limited energy supply, limited number of sensors, limited communication range, fixed detection threshold of individual sensors, limited budget for the cost of hardware, maintenance, etc.). The issue of resource constraints becomes even more relevant for a network of chemical sensors that are used for the continuous environmental monitoring (air and water pollution, hazardous releases, smoke, etc.). The reason is that a modern chemical sensor is usually equipped with a sampling unit (a fan for air and a pump for water), which turns on when the sensor is active. The sampling unit usually requires a significant amount of energy to operate as well as frequent replacement of some consumable items (i.e., cartridges, filters). This leads to the critical requirement in the design of a WSN to reduce the active (i.e., sampling) time of its individual sensors.

One attractive way to achieve an optimal balance between the peak performance of the WSN and its constraints in resources mentioned above is to exploit the idea of dynamic sensor collaboration (DSC) [3, 4]. The DSC implies that a sensor in the network should be invoked (or activated) only when the network will gain information by its activation [4]. For each individual sensor, this information gain can be evaluated against other performance criteria of the sensor system, such as the detection delay or detection threshold, to find an optimal solution in the given circumstances.

While the DSC-based approach is a convenient framework for the development of algorithms for optimal scheduling of constrained sensing resources, the DSC-based algorithms involve continuous estimation of the state of each

sensor in the network and usually require extensive computer simulations [3, 4]. These simulations may become unpractical as the number of sensors in the network increases (e.g., "smart dust" sensors). Even when feasible, the simulations can provide only the numerical values for optimal network parameters, which are specific for an analysed scenario, but without any analytical framework for their consistent interpretation and generalisation. For instance, the scaling properties of a network (the functional relationship between the network parameters) still remain undetermined, which prevents any comprehensive optimisation study.

This motivates the development of another, perhaps less rigorous, but certainly simpler approach to the problem of network analysis and design. The main idea is to phenomenologically employ the so-called bioinspired (epidemiology, population dynamics) or physics-inspired (percolation and graph theory) models of DSC in the sensor network in order to describe the dynamics of collaboration as a single entity [5–10]. Since the theoretical framework for the bio- or physics-inspired models is already well established, we are in the position to make significant progress in the analytical treatment of these models of DSC (including their optimisation). From a formal point of view, the derived equations are ones of the "mean-field" theory, meaning that instead of working with dynamic equations for each individual sensor we only have a small number of equations for the "averaged" sensor state (i.e., passive, active, faulty, etc.), *regardless of the number of the sensors in the system*. A reveling example of the efficiency of this approach is the celebrated SIR model in epidemiology [11]. For any size of population, the SIR model describes the spread of an infection by using only three equations, corresponding to three "infectious" classes of the population: susceptible, infectious, and recovered.

The analytic or "equation-based" approach often leads to valuable insights into the performance of the proposed sensor network system by providing simple analytical expressions to calculate the vital network parameters, such as detection threshold, robustness, responsiveness, and stability and their functional relationships.

In the current paper, we develop a simple model of a wireless network of chemical sensors, where dynamic sensor collaboration is driven by the level of concentration of a pollutant (referred to as the "external challenge") at each individual sensor. Our approach is based on the known analogy [10] between the information spread in a sensor network and the epidemics propagation across a population. In this analogy, the infection transmission process corresponds to message passing among the sensors. A chain reaction in transmission of an infection is called the epidemic. In the context of a sensor network, a chain reaction will trigger the network (as a whole) to move from the "no pollutant" state to the "pollutant present" state, which will indicate the presence of an external challenge.

The paper shows that the adopted epidemics or population-inspired approach can provide a reliable description of the dynamics of such a sensor network. The simple analytical formulas (scaling laws) derived from the model express the relationships between the parameters of the network (e.g., number of sensors, their density, sensing time, etc.), the network performance (probability of detection, response time of a network), and the parameters of the external challenge (environment, pollutant). As an example of application of the proposed framework, we performed a simple optimisation study. Numerical simulations are carried out and presented in the paper in support of analytic expressions.

Although the model presented in this paper is specific to a network of chemical sensors, the underlying analytical approach can be easily adapted to other applications and other types of networks by a simple change of the model of environment and sensor.

2. The Model of Environment

The external challenges are modeled by a random time series which mimics the turbulent fluctuation of concentration at each sensor of the network. In this approach, the fluctuations in concentration C are modeled by the probability density function (pdf) of C with the mean C_0 as a parameter (i.e., C_0 is a mean concentration of the tracer in the area) [12]:

$$f(C \mid C_0) = (1-\omega)\delta(C) + \frac{\omega^2}{C_0}\frac{(\gamma-1)}{(\gamma-2)}\left(1+\frac{\omega}{(\gamma-2)}\frac{C}{C_0}\right)^{-\gamma}. \tag{1}$$

Here, the value $\gamma = 26/3$ can be chosen to make it compliant with the theory of tracer dispersion in Kolmogorov turbulence (see [12]), but it may vary with the meteorological conditions. The parameter ω, which models the tracer intermittency in the turbulent flow, can be in the range $[0, 1]$, with $\omega = 1$ corresponding to the nonintermittent case. In general, it also depends on a sensor position within a chemical plume; thus, ω is in the range 0.95–0.98 near the plume centroid and may drop to 0.3–0.5 near the plume edge. For $\omega \neq 0$, the pdf f of (1) has a delta impulse in zero, meaning that the measured concentration in the presence of intermittency can be zero on some occasions. It can be easily shown that the pdf of (1) integrates to unity, so it is appropriately normalized.

The measured concentration time series can be generated by drawing random samples from the probability density function given in (1) at each time step. The random number generator is implemented using the *inverse transform* method based on the following steps [13]:

(1) draw a sample u from the standard uniform distribution: $u \sim U[0, 1]$;

(2) compute the value of C that satisfies $F(C) = u$, where $F(\cdot)$ is the cumulative distribution function (cdf) of the distribution of interest;

(3) the value of C computed in the previous step is a random sample drawn from the desired probability distribution.

The cdf $F(\cdot)$ needed for inverse transform sampling is obtained by integrating the pdf in (1) and is given by

$$F(C \mid C_0) = 1 - \omega\left[1+\left(\frac{2}{\gamma-2}\right)\frac{C}{C_0}\right]^{1-\gamma}. \tag{2}$$

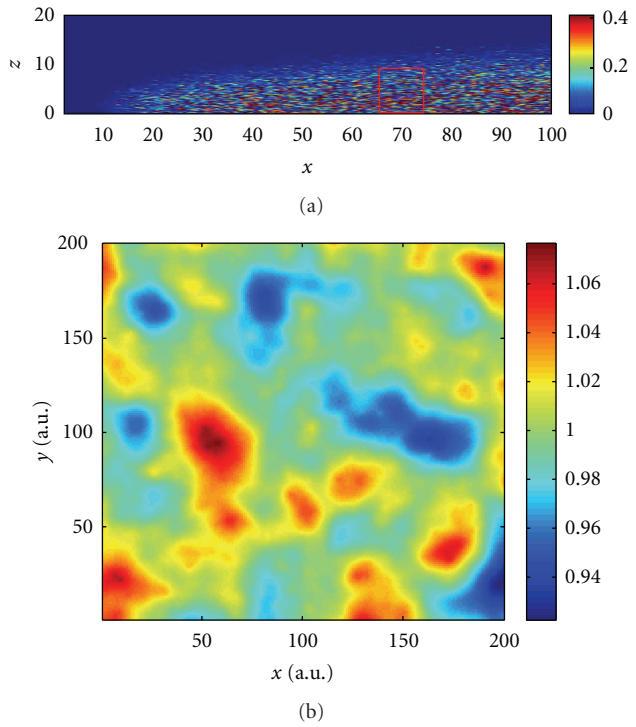

Figure 1: Example for concentration realisation for a plume-like flow (a) and the selected area of the flow with higher resolution (b).

The use of this cdf in the inverse transform sampling procedure generates the value of concentration

$$C = \begin{cases} C_0\left(\frac{\gamma-2}{\omega}\right)\left[\left(\frac{1-u}{\omega}\right)^{-1/(\gamma-1)} - 1\right], & u \geq 1-\omega, \\ 0, & u < 1-\omega, \end{cases} \tag{3}$$

where u is again the standard uniform distribution $u \sim U[0,1]$.

In order to produce spatial correlations that comply with the well-known scaling properties of turbulent dispersion, a special "swapping" algorithms was implement. This recursive algorithm mimics the chaotic fluctuations occurring in the real turbulent flows (for details, see [14]).

The proposed framework allows to implement a reasonably realistic model of the contaminated environment (i.e., to generate the concentration realisation at each sensor over time), see Figure 1. Due to a universal nature of turbulence, it can be used to simulate performance of WSN in detection of either airborne and waterborne releases. The parameters γ and ω are typically estimated from geophysical observation (meteorological and organological) and will be assumed known.

The geometrical complexity of the turbulent flow can be incorporated in the theoretical framework (2) by assuming a temporal and spatial variability of the mean concentration filed $C_0 \equiv C_0(\mathbf{r}, t)$. This way we can simulate various morphologies of the flow (jet, wake, boundary layer, compartment flow, etc.) as well as various scenarios of hazardous release (plume, puff), for details see [13, 15]. For the sake of simplicity in the current paper, we consider only case $C_0 =$ const. This assumption corresponds to the approximation when the size of WSN is less than the width of hazardous plume (see Figure 1), or to an important practical case of a "highly distributed" source of pollutant (traffic, extended industrial site, or urban area [16]).

3. The Model of a Chemical Sensor

We adopt a simple binary (or "threshold") model of a sensor, with the sensor reading V given by

$$V = \begin{cases} 1, & C \geq C_*, \\ 0, & C < C_*. \end{cases} \tag{4}$$

We emphasize that threshold C_* is an internal characteristic of the sensor, unrelated to C_0 in (1). This threshold is another important parameter of our model. A chemical sensor with bar readings, which includes many subsequent levels for concentration thresholds mapped into a discrete sensor output, is an evident generalisation of (4).

Using (3) and (4), it is straightforward to derive the probability of detection for an individual sensor embedded in the environment characterised by (2)

$$p = 1 - F(C_* \mid C_0). \tag{5}$$

This aggregated parameter links the characteristics of a specific sensor C_*, the parameter of the external challenge C_0, and the environment $(F(\cdot), \gamma, \omega)$.

4. Modeling and Analysis of Network Performance

Our focus is a wireless network of chemical sensors with dynamic collaboration. We assume that N identical sensors (i.e., with the same detection threshold C_* and sampling time τ_*) are uniformly distributed over the surveillance domain of area S with density $\rho = N/S$.

We will model the following network protocol for dynamic collaboration. Each sensor can be only in one of the two states: *active* or *passive*. The sensor can be activated only by a message it receives from another sensor. Once activated, the sensor remains in the active state during an interval of time τ_*; then it returns to the passive (sleep) state. While being in the active state, the sensor senses the environment, and if the chemical tracer is detected (binary detection), it broadcasts a (single) message. If a sensor receives an activation message while it is in the active state, it will ignore this message. The broadcast capability of the sensor is characterized by its communication range r_*, which is another important parameter of the model. The described protocol assumes that certain sensors of the network are permanently active. The number of permanently active sensors in the network is fixed, but the actual permanently active sensors vary over time in order to equally distribute the energy consumption of individual sensors.

The WSN following this protocol can be considered as a system of agents, interacting with each other (by means of message exchange) and with the stochastic environment (by means of sampling and probing). The interactions can change the state of agents (active and passive). From this perspective, this WSN is similar to the epidemic SIS (susceptible-infected-susceptible) model [11], in which an individual can be in only two states (susceptible or infected), and the change of state is a result of interaction (mixing) between the individuals (which corresponds to the exchange of messages in our case). Thus, a dynamic (population) model for our system [11] is as follows:

$$\frac{dN_+}{dt} = \alpha N_+ N_- - \frac{N_+}{\tau_*}, \tag{6}$$

$$\frac{dN_-}{dt} = -\alpha N_+ N_- + \frac{N_+}{\tau_*}, \tag{7}$$

where N_+, N_- denote the number of active and passive sensors, respectively. The nonlinear terms on the RHS of (6) and (7) are responsible for the interaction between individuals (i.e., sensors), with the parameter α being a measure of this interaction. The population size (i.e., the number of sensors) is conserved, that is, $N_+ + N_- = N = \text{const}$.

The next step is to express α in terms of the parameters of our system by invoking physics-based arguments used in population dynamics [11]. It is well known that parameter α in (6) describes the intensity (contact rate) of social interaction between individuals in the community, so we can propose (see [11, 17])

$$\alpha \propto \frac{mp}{N\tau_*}, \tag{8}$$

where m is the number of contacts made by an "infected" sensor during the infectious period τ_* (i.e., the number of sensors receiving a message from an alerting sensor). In our case, we have $m = \pi r_*^2 \rho$. Then using $N = S\rho$, we can write

$$\alpha = G\frac{\pi r_*^2}{\tau_* S} p, \tag{9}$$

where G is a constant calibration factor, being of order unity (it must be estimated during the network calibration); p was defined by (5). In order to simplify notation, from now on, we will assume that G is absorbed in the definition of r_*.

It is worth noting that by introducing nondimensional variables $n_+ = N_+/N$, $n_- = N_-/N$, and $\tau = t/\tau_*$, the system (6)-(7) can be rewritten in a compact nondimensional form

$$\frac{dn_+}{d\tau} = R_0 n_+ n_- - n_+, \quad n_- = 1 - n_+, \tag{10}$$

with only one nondimensional parameter

$$R_0 = \alpha \tau_* N. \tag{11}$$

The parameter R_0 is well known in epidemiology where it has the meaning of a *basic reproductive number* [11].

The system (6)-(7) combined with the condition $N_+ + N_- = N$ can be reduced to one equation for $y = N_+$,

$$\frac{dy}{dt} = \alpha y(N - y) - \frac{y}{\tau_*} = y(b - \alpha y), \tag{12}$$

where

$$b = \alpha N - \frac{1}{\tau_*} = \frac{(R_0 - 1)}{\tau_*}. \tag{13}$$

By simple change of variables $z = \alpha y/b$, this equation can be reduced to the standard logistic equation

$$\frac{dz}{dt} = bz(1 - z), \tag{14}$$

which has the well-known solution

$$z(t) = \frac{z_0}{(1 - z_0)\exp(-bt) + z_0}, \tag{15}$$

where $z_0 = z(0)$.

We can see that if $b < 0$, then $z \to 0$ as $t \to \infty$ for any z_0, so any individual sensor activation in the network will "die out," that is, the network will not be able to detect the external challenge. The same is valid for $b = 0$ when $z = z_0 = \text{const}$ (no response to external challenges). Only if the condition $b > 0$ is satisfied, then $z \to 1$ as $t \to \infty$ (independently of z_0). In this case, after a certain transition interval, the network will reach a new steady state with

$$\frac{N_+}{N} = 1 - \theta, \quad \frac{N_-}{N} = \theta, \quad \theta = \frac{1}{\alpha \tau_* N} \equiv \frac{1}{R_0}. \tag{16}$$

A fraction of active sensors N_+ at this new state is a measure of the network (positive) response to the event of chemical contamination. From (15), it is clear that the time scale for the network to reach the new state can be estimated from the condition $e^{-bt} \ll 1$, so

$$\tau \geq \frac{1}{b} = \frac{\tau_*}{R_0 - 1}. \tag{17}$$

This equation provides the relationship between the scale of activation time and parameter R_0. One can see that this scale decreases as R_0 increases.

From (14), (17), it follows that an "epidemic threshold" for the sensor network is simply $\alpha \tau_* N > 1$ or in terms of the "basic reproductive number" (11),

$$R_0 = \alpha \tau_* N = pN\frac{\pi r_*^2}{S} > 1. \tag{18}$$

Observe that sensor sampling time τ_* has disappeared from the expression for R_0. This means that it is possible to create an information epidemic (i.e., detect a chemical pollutant) for any value of τ_*, provided this time is long enough for a sensor to detect the chemical tracer. But according to (17), the responsiveness of the whole network to the external challenges (i.e., the time constant of detection) is, indeed, strongly dependent on the sensor sampling time $\tau = \tau_*/(R_0 - 1)$.

The expressions (16), (17), and (18) are the main analytical results of the paper. For a given level of external challenges (i.e., C_0) and meteorological conditions (i.e., γ, ω), these expressions provide a simple yet rigorous way to estimate how a change in the network and sensor parameters (i.e., N, C_*, τ_*) will affect the network performance (i.e., N_+, τ).

We can also see that for a given external challenge the network of chemical sensors will respond in the most effective way when its parameters are selected in the combination which meets the criterion for "information epidemic" (18).

The final analytical expressions enable us to maximize the network information gain and optimize other parameters. For example, from (16), we can readily infer the important scaling properties of the network performance:

$$\frac{N_-}{N} \sim \frac{1}{r_*^2}, \qquad \frac{N_-}{N} \sim \frac{1}{N}, \qquad \frac{N_-}{N} \sim \frac{1}{p}. \tag{19}$$

For instance, if we double the communication range of an individual sensor r_*, the fraction of inactive sensors in the network will drop four times. Likewise, if we need to reach a specified fraction of active sensors $(1 - N_-/N)$ to be able to reliably detect a given level of pollutant concentration, these formulas describe all possible ways of changing the parameters of the model in order to achieve this goal.

5. Information Gain of Collaboration

We have explained earlier that the concept of DSC is important for a network with limited energy/material resources. But the question remains will a network with DSC be inferior (in terms of detection performance) in comparison with a benchmark network where all sensors operate independently of each other and only report their (positive) detections of chemical pollution to the central processor for decision making? Clearly, such a benchmark network would be very expensive to run (all sensors would have to be active all the time), but could provide excellent detection performance.

In this section, we show that, under a certain condition, the network with DSC can provide superior detection performance compared to the benchmark network. Let us assume that we have δN sensors continuously operating ($0 \le \delta \le 1$). For a benchmark network, on average, we have $p\delta N$ sensors detecting pollutant. For the network with DSC, the same quantity can be estimated as $p(1-\theta)N$ (as we have seen the saturation level of N_+ does not depend on initial conditions). From here, we can then deduce that the network with DSC will provide more information (for detection of chemical pollution) than the benchmark network if the following condition is satisfied:

$$\theta = \frac{1}{\alpha\tau_* N} \le (1 - \delta), \tag{20}$$

which is eventually reduced to the condition of "epidemic threshold" (18) for the small value of δ.

The value of the parameter δ can be also estimated based on the following arguments. Let us assume that our aim is to detect a level concentration C_0 associated with a hazardous release within the time T (the constraint on time is driven by the requirement to mitigate the toxic effect of the release). Then, we can write a simple condition for the information "epidemic" in the WSN to occur during time T,

$$\frac{\delta pNT}{\tau_*} \ge 1, \tag{21}$$

where p is given by (5), that is, $p = 1 - F(C_* \mid C_0)$. Evidently, for information epidemic to be observable, the number of continuously active sensors should be less than the number of sensors activated due to the hazardous release. Thus, from (20), we can write the following "consistency" condition for the minimum value of δ:

$$\delta_{\min} \approx \frac{\tau_*}{pNT} \le \left(1 - \frac{1}{\alpha\tau_* N}\right), \tag{22}$$

or by rewriting it in terms of R_0, see (16),

$$\delta_{\min} \approx \frac{\tau_*}{pNT} \le \left(1 - \frac{1}{R_0}\right). \tag{23}$$

It can be seen that with other conditions being equal, the fraction of "stand-by" sensors $\delta_{\min}$ can be made however small (since $R_0 \ge 1$). It implies that only a small fraction of WSN will be active most of the time and is a clear demonstration of the energy consumption gain associated with the "epidemic" protocol.

Another important criteria for epidemic protocol can be derived by comparison of amplitude of "detectable events" for the *same number of sensors* in the network with DSC with the system of N-independent sensors. For the network with DSC, it is $(1 - \theta)N$ (since we use N_+ to retrieve information about the environment), and for the system of the *same independent sensors*, it is still pN (since N_+ is simply equal to N). Then instead of (20), we can write

$$\theta < (1 - p). \tag{24}$$

Under this condition, more detectable events will occur in the presence of chemical pollution by the described network with DSC (activation messages) then in a network of stand alone sensors (signals of positive detection). This leads to the interesting threshold condition on the number of sensors in the network

$$N > \frac{S}{\pi r_*^2} \frac{1}{p(1-p)}. \tag{25}$$

The last term in RHS $(p(1-p))^{-1}$ has an obvious minimum 4 corresponding to $p = 1/2$, so finally, we arrive at the simple universal condition

$$N > N_* = \frac{4}{\pi} \frac{S}{r_*^2}. \tag{26}$$

This condition reads that if the number of sensors in the system is greater than N_*, then networking with DSC *can* provide an information gain over the benchmark network. Under this condition, the network with DSC is not only desirable from the aspect of energy conservation, but also provides better detection performance through the information gain.

The condition $p = 1/2$ minimizing RHS of (25) can be considered as a criterion for an "optimal" sensor for a given network with DSC and for a given concentration of pollutant to be detected. Namely, from the equation $F(C_* \mid C_0) = 1/2$ and using (2), we can write

$$C_* = C_0 \left(\frac{\gamma - 2}{2}\right) \left[\left(\frac{1}{2\omega}\right)^{1/(1-\gamma)} - 1\right]. \tag{27}$$

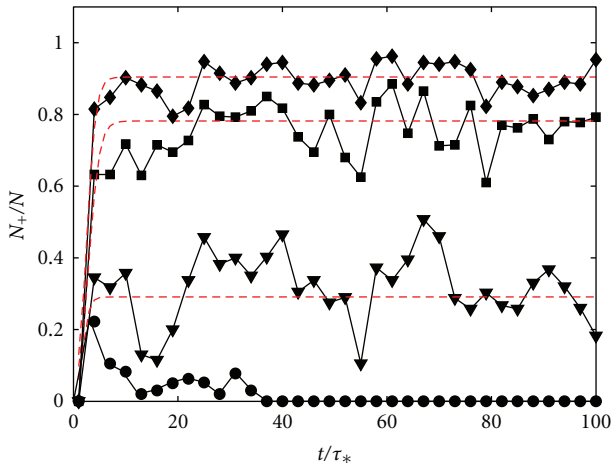

FIGURE 2: Results of numerical simulations: the fraction of active sensors in the network over time for the different communication range r_*: ♦—$r_* = 40$ m, ■—$r_* = 30$ m, ▼—$r_* = 27$ m, and •—$r_* = 20$ m; $C_*/C_0 = 1.03$, $N_+(t = 0) = 10$. The dashed red line corresponds to the analytical predictions (15). It is clearly seen that in the case $r_* = 20$ m, the information epidemic in WSN dies off.

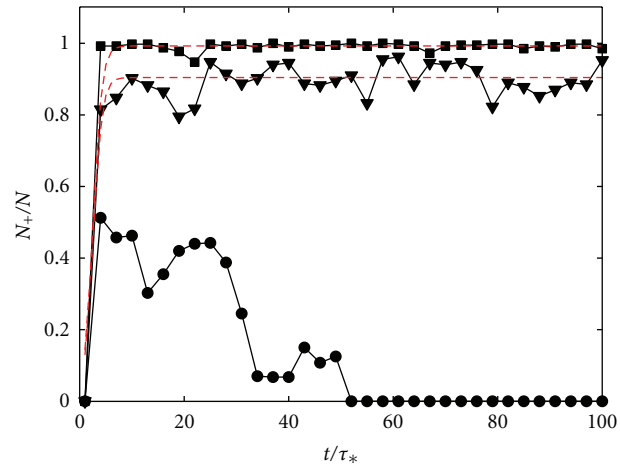

FIGURE 3: Results of numerical simulations: the fraction of active sensors in the network over time for the different threshold of individual sensor C_*: ■—$C_*/C_0 = 1.05$, ▼—$C_*/C_0 = 1.02$, and •—$C_*/C_0 = 1.00$; $r_* = 40$ m, $N_+(t = 0) = 10$. The dashed red line corresponds to the analytical predictions (15). It is clearly seen that in the case $C_*/C_0 = 1.00$, the information epidemic in WSN dies off.

Given environmental parameters (γ, ω) and given the level of concentration to be detected (C_0), formula (27) also specifies a simple condition on detection threshold for an individual sensor to maximize an information gain by being networked.

6. Numerical Simulations

In support of analytical derivations presented above, a network of chemical sensors operating according to the adopted protocol for dynamic collaboration was implemented in MATLAB. A comprehensive report with numerical simulations result will be published elsewhere; here, we present only some illustrative examples.

For consistency, a 1000 m × 1000 m surveillance region populated by $N = 400$ sensors with a uniformly random placement was assumed in all tests. In each run, chemical pollution with concentration $C_0 = 150$ is applied, and the simulation starts when a single randomly selected sensor (which has detected the presence of chemical contamination in its vicinity) starts broadcasting. Due to this random initiation and the fact that the probability of detection of individual sensors is less than unity ($p < 1$), each run of the computer program results in a slightly different outcome. Figures 2 and 3 show the average evolution of the ratio N_+/N in the network over time. The curves were obtained by using the following parameters: $\omega = 0.98$, $\gamma = 26/3$. Figure 2 demonstrates the changes in dynamics of the WSN for different values of communication range r_*, and Figure 3 depicts the similar plots for changes of the detection threshold of individual sensor C_*. For all plots in Figures 2 and 3, the initial number of active sensors is $N_+(t = 0) = 10$.

Overall, we found that the simulation output is much more sensitive to the changes of communication range than to the threshold of an individual sensor (see range of parameters depicted in Figures 2 and 3). In all cases, we observed the transition of N_+ from the initial steady state (where N_+ is very small indicating the absence of the pollutant) to the new steady state (high value of N_+), so information "epidemic" in the network of chemical sensors does occur. By direct substitution into (18), it was also validated that in all cases presented in Figures 2 and 3 the condition for an information "epidemic" was satisfied. In general, the saturation value of N_+ derived from these plots was in an agreement with theoretical prediction (16), but the estimated standard deviation of N_+ (not shown in Figure 2) could be very high (up to 30%) for some combination of parameters. The relative standard deviation (normalized by mean value N_+) usually gradually decreased over time and quite rapidly decays with the increase of communication range r_*. The occasional high variability of the output of the sensor network is undesirable and motivates further analysis. We also used the data from the plots in Figures 2 and 3 to calibrate our model. The calibration was performed by extracting the steady-sate (or saturation) values of N_+ from the plots and by adjusting the "free" constant G in the analytical expressions (16) to achieve the best match between the analytical predictions and simulations. The value $G \approx 0.7$ seems to provide an optimal agreement with the presented simulations.

In order to validate our simple model for parameter α, we performed the following study. For each simulation, we derived the value of α_s from (16) and then compared it with the value of α_t calculated from the theoretical expression (9) using the calibration value $G \approx 0.7$. The results of this study are presented in Figure 4. The red dashed line corresponds to the perfect agreement between the theory and simulations. Considering the high variability of N_+ and a rather simple model for α, the agreement between the theory and simulations is acceptable.

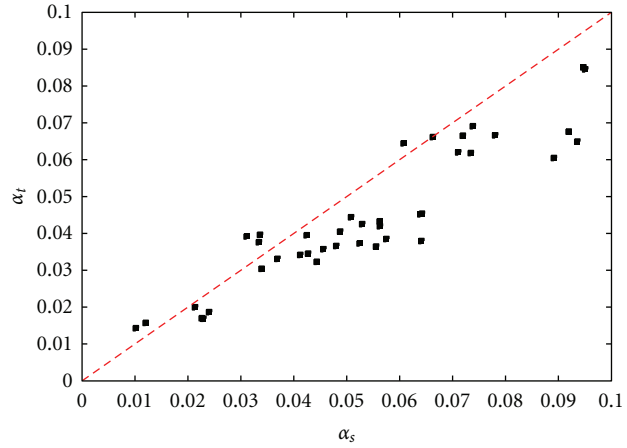

FIGURE 4: Simulation and theoretical predictions of parameter α: α_t is the theoretical value (9), α_s is the results of simulations. The red dashed line corresponds to the perfect agreement.

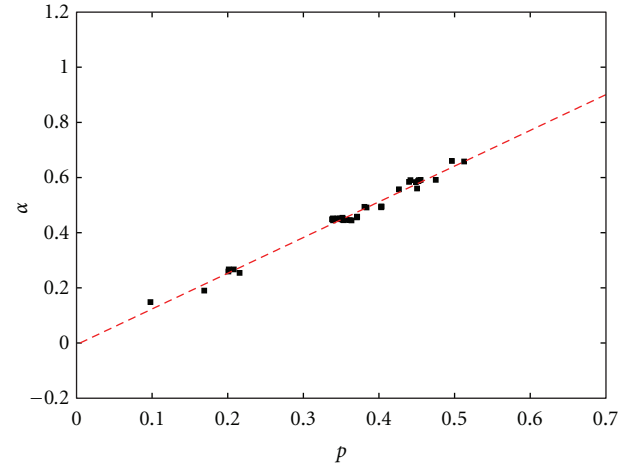

FIGURE 5: Parameter α as a function of p extracted from numerical simulations in log-log scale. The dashed line corresponds to the power-law fit $\alpha \propto p^q$, $q = 1.27$, and theoretical prediction corresponds to $q = 1$, see (8).

To validate further the alignment between the computer simulations and the proposed mathematical model, we numerically estimated some scaling properties of the network system (i.e., (9), (19)). Firstly, we derived the scaling properties from computer simulations and then compared them to the theoretical predictions. In general, we found that all trends of the derived scaling do agree with theoretical expressions in (19), but the quantitative agreement may significantly vary from case to case. As an illustration, in Figure 5, we present the plot of dependency of α against p in log-log scale. The extracted exponent corresponds to $\alpha \propto p^q$, where $q = 1.27$, while the theoretical value according to (8) is $q = 1$. This indicates that while our analytical model is very simple and fast to compute, for higher accuracy it may need further refinements as discussed below.

The results of numerical simulations presented above serve to verify that the "information epidemic" does occur in the wireless network of chemical senors. This also implies that the proposed theoretical framework may lead to a gain in the energy consumption that may result in the significant advantages in operational deployment of such systems. More detailed analysis of the optimal values of parameters satisfying threshold conditions (18), (23), and (26) and lead to the optimal performance of WSN will be reported in separate publications.

7. Refinements of the Model

The disagreement described above is due to the implicit assumption of "homogeneous mixing" which we made in equations (6)-(7). The homogeneous mixing manifests itself in the bilinear form of the interaction terms on the RHS of (6)-(7). This bilinearity means that the number of new "infected" sensors is proportional to the product of the number which is currently "infected" and the number which is currently "susceptible." Effectively it means that all passive sensors are equally likely to be activated. This assumption holds only if the majority of activated ("infected") sensors are far away from each other (i.e., at the distances $\gg r_*$). At some stage of the sensor "epidemic," this assumption can be violated, because the secondary "infected" sensors will be at the shorter distances from the "infectious" parents (see Figure 6). The broadcasted messages in overlapping areas become duplicated and the rate of new "infections" will be no longer proportional to the number of their parents. The fraction of "infected" sensors in the overlapping areas will depend on the new equilibrium state of the sensor system (i.e., N_+/N as $t \to \infty$) and may not be small for some scenarios. To overcome this restriction, we again invoke an approach successfully implemented in epidemiology (see [17]). Instead of (6)-(7), we now write

$$\frac{dN_+}{dt} = \alpha N_+^\nu N_- - \frac{N_+}{\tau_*}, \qquad \frac{dN_-}{dt} = -\alpha N_+^\nu N_- + \frac{N_+}{\tau_*}, \tag{28}$$

where a new parameter $0 \le \nu \le 1$ depends on the packing density of "infected" sensors (or on the ratio N_+/N). For a "sparse" network configuration, we have $\nu \approx 1$ (no overlapping areas), and for an extremely "dense" network, $\nu \approx 0$ (all sensors are located around the same point), see Figure 6. In general, ν can be used as a fitting parameter of the model [8] or estimated based on the mathematical theory of packing. For a specific network configuration, a value $\nu = 1/2$ was derived in [7] based on some simplified assumptions. By employing new parameter ν, we can significantly improve agreement between analytical model and simulation at the initial stage of information epidemic, since here we can assume that $N_- \approx N = \text{const}$, so $dN_+/dt \propto N_+^\nu$. An example of improved fitting is presented in Figure 7.

Similarly to the epidemiological models (see [11]), incorporation of the spatial inhomogeneity can be achieved by

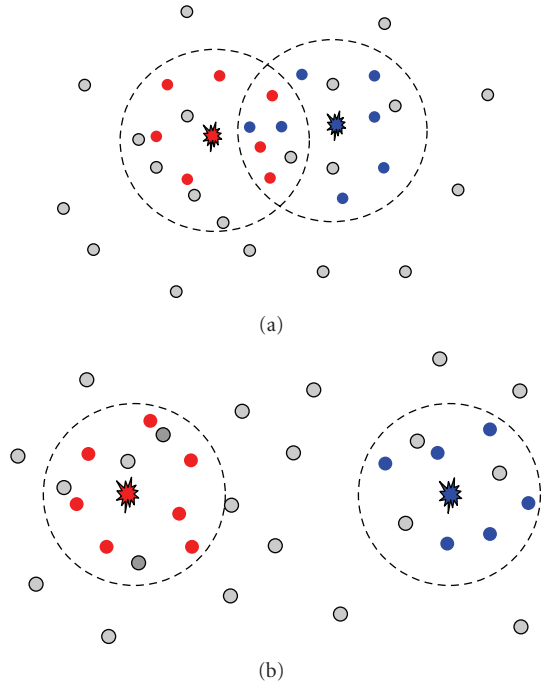

Figure 6: Examples of dense (a) and sparse (b) wireless sensor networks.

adding the appropriate diffusion terms on the LHS of (6) and (7):

$$\frac{\partial N_+}{\partial t} - D\Delta N_+ = \alpha N_+ N_- - \frac{N_+}{\tau_*},$$
$$\frac{\partial N_-}{\partial t} - D\Delta N_- = -\alpha N_+ N_- + \frac{N_+}{\tau_*}, \quad (29)$$

where D is diffusivity in the sensor system which can be estimated as $D \approx r_*^2/\tau_*$ and Δ is the Laplace operator. At the same time, the inhomogeneity of pollutant distribution can be easily incorporated in $\alpha(\mathbf{r})$ with nonuniform $C_0(\mathbf{r})$ (see (2), (5), (8)).

An important property of the system (29) is the existence of analytical solutions in the form of traveling waves, propagating with the velocity $v_0 \sim \sqrt{\alpha D}$ [11]. In our case, these waves correspond to the switching fronts between active and passive sensors. If pollutant is advected by the wind flow with a characteristic velocity v_*, then a simple synchronisation condition $v_0 \geq v_*$ or $\alpha \geq v_*^2 \tau_*/r_*^2$ provides an important criteria for network optimisation.

Another interesting extension of the proposed model is the introduction of the concept of a *faulty* sensor, a sensor which is no longer available for sensing and networking. This state of a sensor would correspond to the *removed* population segment in the epidemiological framework and can be attributed to any kind of faults (flat battery, software malfunction, hardware defects, etc.). As in the celebrated SIR epidemiological model [11], a new state results in the third equation for N_0 in the system (6)-(7) with a new temporal parameter—an average operational time (the lifespan) of a sensor. The total number of sensors will be still conserved:

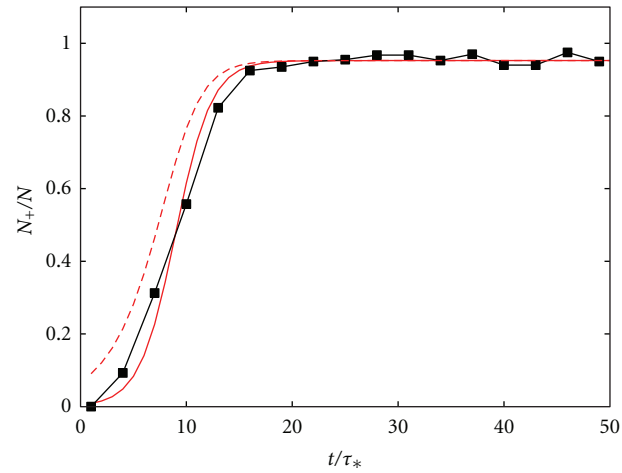

Figure 7: Effect of parameter ν in the model (28) on the simulation data fit: dashed line: $\nu = 1$, solid line: $\nu = 0.7$.

$N = N_+ + N_- + N_0 = \text{const}$. This model provides a more realistic representation of an operational sensor systems and allows us to estimate such important parameters as the operational lifetime of the network and the reliability of the network.

8. Conclusions

We developed a "bioinspired" model of a network of chemical sensors with dynamic collaboration for the purpose of energy conservation and information gain. The proposed model leverages on the existing theoretical discoveries from epidemiology resulting in a simple analytical model for the analysis of network dynamics. The analytical model enabled us to formulate analytically the conditions for the network performance. Thus, we found an optimal configuration which, within the underlying assumptions, yields a balance between the number of sensors, detected concentration, the sampling time, and the communication range. The findings are partly supported by numerical simulations. Further work is required to address the model refinements and generalisations.

Acknowledgments

The authors would like to thank Ralph Gailis, Ajith Gunatilaka, and Chris Woodruff for helpful technical discussions and Champake Mendis for his assistance in the software implementation of the proposed model.

References

[1] C. S. Raghavendra, K. M. Sivalingam, and T. Znati, *Wireless Sensor Networks*, Springer, New York, NY, USA, 2005.

[2] P. E. Bieringer, A. Wyszogrodzki, J. Weil, and G. Bieberbach, "An evaluation of propylene sampler grid designs for the FFT07 field program," Tech. Rep., National Center for Atmospheric Research, Boulder, Colo, USA, 2006.

[3] E. Ertin, J. W. Fisher, and L. C. Potter, "Maximum mutual information principle for dynamic sensor query problems," in *Information Processing in Sensor Networks*, vol. 2634 of *Lecture Notes in Computer Science*, pp. 91–104, 2003.

[4] F. Zhao, J. Shin, and J. Reich, "Information-driven dynamic sensor collaboration for tracking applications," *IEEE Signal Processing Magazine*, vol. 19, no. 2, pp. 61–72, 2002.

[5] J. Mathieu, G. Hwang, and J. Dunyak, "The state of the art and the state of the practice: transferring insights from complex biological systems to the exploitation of netted sensors in command and control enterprises," MITRE Technichal Papers, MITRE Corporation, Bedford, Mass, USA, 2006.

[6] A. Khelil, C. Becker, J. Tia, and K. Rothermel, "An epidemic model for information diffusion in MANETs," in *Proceedings of the 5th ACM international workshop on Modeling, Analysis and Simulation of wireless and mobile systems (MSWiM '02)*, pp. 54–60, Atlanta, Ga, USA, 2002.

[7] P. De, Y. Liu, and S. K. Das, "An epidemic theoretic framework for evaluating broadcast protocolsin wireless sensor networks," in *Proceedings of the Proceedings of IEEE Internatonal Conferenceon Mobile Adhoc and Sensor Systems (MASS '07)*, pp. 1–9, Pisa, Italy, 2007.

[8] B. Ristic, A. Skvortsov, and M. Morelande, "Predicting the progress and the peak of an epidemic," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '09)*, pp. 513–516, Taiwan, April 2009.

[9] A. Dekker and A. Skvortsov, "Topological issues in sensor networks," in *Proceedings of the Modelling and Simulation Society of Australia and New Zealand International Congress on Modelling and Simulation (MODSIM '09)*, Cairns, Australia, 2009.

[10] S. Eubank, V. S. A. Kumar, and M. Marathe, "Epidemiology and wireless communication: tight analogy or loose metaphor?" in *Bio-Inspired Computing and Communication*, vol. 5151 of *Lecture Notes in Computer Science*, pp. 91–104, Springer, New York, NY, USA, 2008.

[11] J. D. Murray, *Mathematical Biology*, vol. 1-2, Springer, New York, NY, USA, 2002.

[12] V. Bisignanesi and M. S. Borgas, "Models for integrated pest management with chemicals in atmospheric surface layers," *Ecological Modelling*, vol. 201, no. 1, pp. 2–10, 2007.

[13] A. Gunatilaka, B. Ristic, A. Skvortsov, and M. Morelande, "Parameter estimation of a continuous chemical plume source," in *Proceedings of the 11th International Conference on Information Fusion (FUSION '08)*, pp. 1–8, Cologne, Germany, July 2008.

[14] A. Gunatilaka, A. Skvortsov, and R. Gailis, "Progress in DSTO CBR simulation environment development," in *Proceedings of the Land Warfare Conference (LWC '08)*, pp. 62–68, Brisbane, Australia, 2008.

[15] A. Skvortsov and E. Yee, "Scaling laws of peripheral mixing of passive scalar in a wall-shear layer," *Physical Review E*, vol. 83, no. 3, part 2, Article ID 036303, 8 pages, 2011.

[16] M. Jamriska, T. C. DuBois, and A. Skvortsov, "Statistical characterisation of bio-aerosol background in an urban environment," *Atmospheric Environment*. In press, http://arxiv.org/abs/1110.4184.

[17] A. T. Skvortsov, R. B. Connell, P. D. Dawson, and R. M. Gailis, "Epidemic spread modeling: alignmentof agent-based simulation with a simple mathematical model," in *Proceedings of the International Conference on Bioinformatics and Computational Biology (BIOCOMP '07)*, vol. 2, pp. 487–490, CSREA Press, Las Vegas, Nev,USA, 2007.