18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

Topological Issues in Sensor Networks

Dekker, A.H. and A.T. Skvortsov

Defence Science and Technology Organisation, Australia
Email: dekker@acm.org

Abstract: Recent years have seen considerable interest in sensor networks. In this paper we study networks of chemical or similar sensors detecting a plume of contaminant, such as that in Figure (i), although our analysis also applies to sensors of other kinds.

In order to optimise power consumption, we assume that sensors are normally quiescent, but can "wake" each other if contaminant is detected. Consequently, a wave of activation spreads through the sensor network, much like the spread of infection through a population. In addition, exchanging information with other nodes allows the network to compensate for sensor detection errors.

We describe a Java-based simulation of such a network, which we use to study the performance of different sensor network topologies. In the model, the contaminant plume is simulated with a probability density function of chemical concentration, compatible with the Richardson-Obukhov theory of turbulent mixing. Sensor nodes are simulated with a simple agent-based model incorporating message-passing.

Of the six networks examined, the best-performing networks were a square grid and a network with short-range random links. These two networks most effectively compensated for sensor errors, while minimising the overall power consumption, as a result of not "waking up" unnecessary nodes. In general, random links in a sensor network appear to be effective, as long as the distance between linked nodes is small compared to the size of the contaminant plume.

The wave of activation which spreads through the network differs from traditional models of the spread of infection through a population, in that initial growth in the number of recently activated sensors is approximately linear, followed by an exponential decay. The linear phase corresponds to an expanding circle of activated nodes within the contaminant plume, while the decay phase occurs when the wave passes beyond the plume. Further work will be conducted to model this process in more detail.

Figure (i). An example contaminant plume. For many contaminants, such plumes are invisible (original photo by Mila Zinkova).

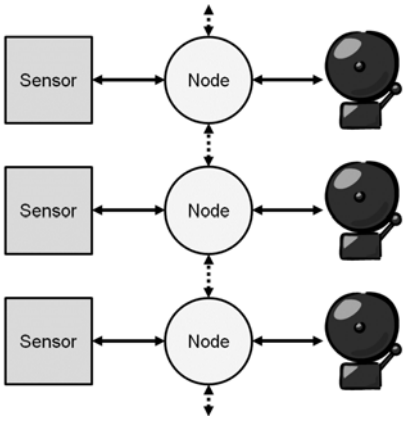

Figure (ii). A node in the sensor network raises an alarm based on the node's own sensor information together with information from other nodes.

Keywords: *sensor network, topology, sensor activation, epidemic, simulation*

1. INTRODUCTION

Chemical sensor networks have many potential applications. Airborne chemical contaminants such as volcanic emissions or industrial pollutants can cause human health problems, and may require government intervention. Effective intervention requires knowing the extent of the affected area, the source of the contaminant, and the speed and direction of movement of the contaminant plume.

Closely related issues occur with sensors for detecting airborne biological agents, for water-borne contaminants within the ocean, and even for the presence of animals or people over a geographical area.

In general, sensors are detecting a concentration field $C_t(x,y)$ at a given time t, where C_t is a smooth function with a value below some threshold C^* outside the contaminant area. Using a larger number of sensors provides a better estimate of the region for which $C_t(x,y) \geq C^*$, as well as permitting extrapolation of the field C_t forwards or backwards in time.

It is possible to consider all sensors to be directly connected to a central processing station which "fuses" the sensor data and performs the required computations. However, sensors are typically connected by radio links and, because of the inverse-square law for radio propagation, power consumption is minimised when messages are passed in multiple short-range "hops," rather than one large "jump" (Zhao and Guibas, 2004). For example, sensors can be arranged in a rough $n \times n$ grid, as in Figure 1, and communicate only with their closest neighbours in the grid. Data can be transferred to a specific point in the grid (such as the top-left node) from any other node in at most $2(n-1)$ "hops."

Sensor networks of this kind raise a number of issues regarding the most effective method for routing data through the network (Zhao and Guibas, 2004). However, our interest lies in *distributed processing* through the network. Rather than routing data to a central processing station, we are interested in the case where a node in the network uses its own (typically unreliable) sensor, together with information from adjacent nodes, to produce an integrated response. For example, a node could issue an audible alarm when its sensor data, together with data from nearby nodes, suggests a high probability that $C_t(x,y) \geq C^*$ in the immediately surrounding area. The ability to do this relies on the smoothness of the function C_t, together with the fact that connected nodes are physically close to each other. Figure 2 shows a sensor network of this kind.

In the present work, we concentrate attention particularly on establishing the region for which $C_t(x,y) \geq C^*$, and do not consider extrapolation of C_t forwards or backwards in time. Our main focus is on how the topology of connections between nodes affects the ability of the sensor network to determine the region for which $C_t(x,y) \geq C^*$.

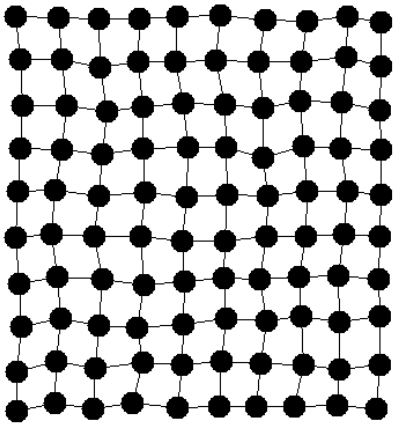

Figure 1. A 10×10 sensor grid.

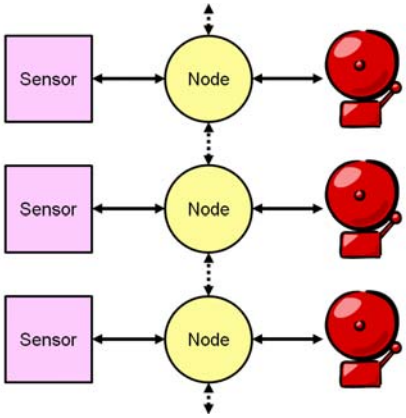

Figure 2. Detail of sensor network structure. Each node includes a sensor and an alarm.

Figure 3. Volcanic plume from the Halema'uma'u vent in Hawaii, showing local turbulence (photo by Mila Zinkova, under Creative Commons: commons.wikimedia.org/wiki/File:Sulfur_dioxide_emissions_from_the_Halemaumau_vent_04-08-1_1.jpg).

2. A SENSOR MODEL

We concentrate on contaminants where the field $C_t(x,y)$ is approximately Gaussian around a central point. However, for the simulations presented here, we approximate the Gaussian by a cosine curve for the region $-\pi\ldots\pi$, and a constant zero outside that range. This therefore puts a definite limit on the extent of the contaminant plume. Within the boundaries of the plume, local turbulence (such as that shown in Figure 3) may result in fluctuations around C_t. This results in a probability density function for the local concentration, given by equation (1):

$$\rho(C_t' \mid C_t) = (1-\omega)\delta(C_t') + \frac{\omega^2(\gamma-1)}{C_t(\gamma-2)}\left(1+\frac{\omega}{(\gamma-2)}\frac{C_t'}{C_t}\right)^{-\gamma} \quad (1)$$

In equation (1), $\gamma = 26/3$ would be compatible with the Richardson-Obukhov theory of turbulent mixing, and the intermittency factor ω ranges from close to 1 near the centre of the plume down to about 0.6 near the edge (Gunatilaka *et al*., 2008).

We can simulate the effect of local turbulence by choosing uniformly distributed numbers u in the range 0…1, and taking the local concentration to be zero for $u < 1-\omega$, and otherwise:

$$C_t\left(\frac{\gamma-2}{\omega}\right)\left[\left(\frac{1-u}{\omega}\right)^{\frac{1}{1-\gamma}}-1\right]$$

Areas within the contaminant plume may have very low local concentrations. However, these areas are still subject to threat, since the turbulent processes giving low local concentrations may also result in very rapid concentration changes within the general area of the plume. At the same time turbulent mixing may produce very sharp spikes of concentration near the edge of the plume.

We assume a sensor threshold C^*, so that sensors report the presence of contaminant whenever the local concentration is greater than or equal to C^*. Outside the plume, where the local concentration will always be below this threshold, we assume a false-positive failure rate of 10%. That is, activated sensors outside the plume falsely report the presence of contaminant 10% of the time.

We also assume that sensors require power to operate. They may include a fan for sampling air, for example. Sensors are therefore normally inactive. From time to time, sensor nodes "wake up" and sample their local environment. They are also "woken up" by their neighbours if a potential contaminant threat is detected.

The "waking up" process begins with a sensor which becomes active (for example, through a timer) within a contaminant plume, and registers a positive result. This node then broadcasts its sensor observations through the network to adjacent nodes, which fire up their own sensors and in turn broadcast their own observations (positive or negative). A wave of observations therefore sweeps through the network, resembling the spread of an infection through a population (Giesecke, 2002). We assume that this wave propagates rapidly compared to the evolution of the plume, so that the plume can be treated as effectively static.

The wave of activation stops when a node in the network receives a *negative* observation from a neighbour, and confirms the negative observation with its own sensor. In this case, the node deems the probability of contaminant to be so low that it does not wake any of its other neighbours.

After being woken, nodes integrate their own sensor observations with one or more observations from neighbouring nodes. A positive alarm is raised if two conditions are both met:

- there are no negative observations, or there are at least two positive observations; and
- at least 40% of the observations are positive.

3. SENSOR NETWORK PERFORMANCE CRITERIA

There are four criteria which we look at for assessing the performance of different sensor network topologies.

First, we consider the total number of nodes "woken up." This provides a measure of power consumption, since each of those nodes will incur the cost associated with operating its own sensor.

Second, we consider the total number of messages sent across the network. This provides a measure of power consumption due to communication.

Third, we consider the average time taken (in steps) before the wave of activation dies out. Each step involves a node firing up its sensor, waiting to obtain an observation, and notifying its neighbours in the network. As noted above, this process ends when the only notifications are negative ones, confirmed by the destination's own sensor.

Finally, and most importantly, we consider the accuracy with which the overall sensor network recognizes the boundaries of the detectable contaminant plume, i.e. the region for which $C_t(x,y) \geq C^*$. We do this by considering the sum of false-positive errors (where $C_t(x,y) < C^*$, but with probability 10% a positive reading has been recorded) and false-negative errors (where a node is within the plume, i.e. $C_t(x,y) \geq C^*$, but the local concentration is below the threshold due to turbulence). Communication within the network should reduce the chance of both kinds of error, either by neighbours failing to confirm a false positive, or by neighbours within the plume correcting a false negative. For a given network, we measure the errors as a percentage change compared to the baseline where all nodes are "woken up" but come to a decision based only on their own internal sensor. Thus –100% would mean that network communication eliminates all errors, while +100% would mean that network communication doubles the number of errors.

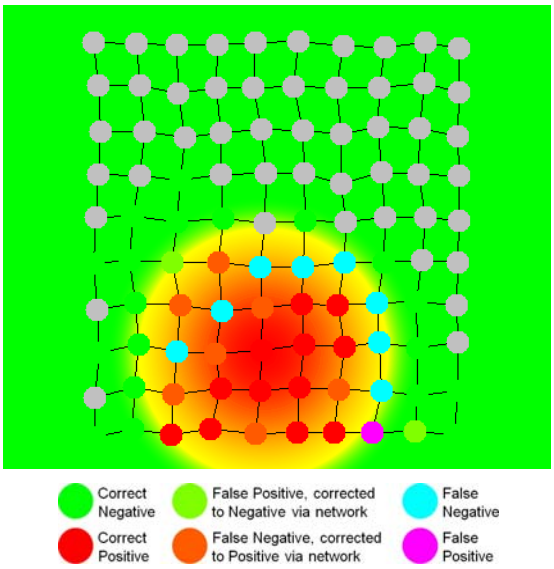

Figure 4. Experimental scenario. Background colour indicates the value of $C_t(x,y)$, with yellow being the threshold value C^*, orange and red above the threshold, and green below the threshold. Node colours are as per the legend, with grey indicating nodes that are never "woken up."

4. SIMULATION EXPERIMENTS

We have conducted a number of experiments involving a plume of contaminant as discussed above. We used a simple agent-based model, written in Java, and incorporating message-passing between sensor nodes. Figure 4 shows a typical run, with the contaminant plume in yellow, orange, and red. The background colour in this diagram indicates the value of $C_t(x,y)$, with yellow being the threshold value C^*, and green indicating values below this threshold. There are 100 sensor nodes, initially connected in a square grid network.

Experiments used the six network topologies shown in Figure 5. These included the square grid network N_G; four networks of the form N_d, with links connected at random between pairs of nodes at most a distance d apart; and a denser grid network N_X. Table 1 gives the average degree and average path length between nodes for these nets.

Table 1. Network characteristics (distances in N_d relate to a height of 400 for the area in Figure 4).

Net	Average degree	Average path length
N_G	3.6	6.67
N_{50}	3.6	6.03
N_{100}	3.6	4.53
N_{200}	3.6	3.75
N_{300}	3.6	3.73
N_X	6.84	4.68

For each network, we conducted 1000 simulation runs, and recorded the performance criteria from Section 3. Table 2 summarises the average results.

In terms of compensating for sensor errors, the best-performing networks were the square grid N_G and the network N_{50} with short-range random links. For these networks, over 40% of errors were corrected (on average) by sharing information with neighbours. Uncorrected errors were generally false positives just outside the plume area, or clusters of false negatives within the plume.

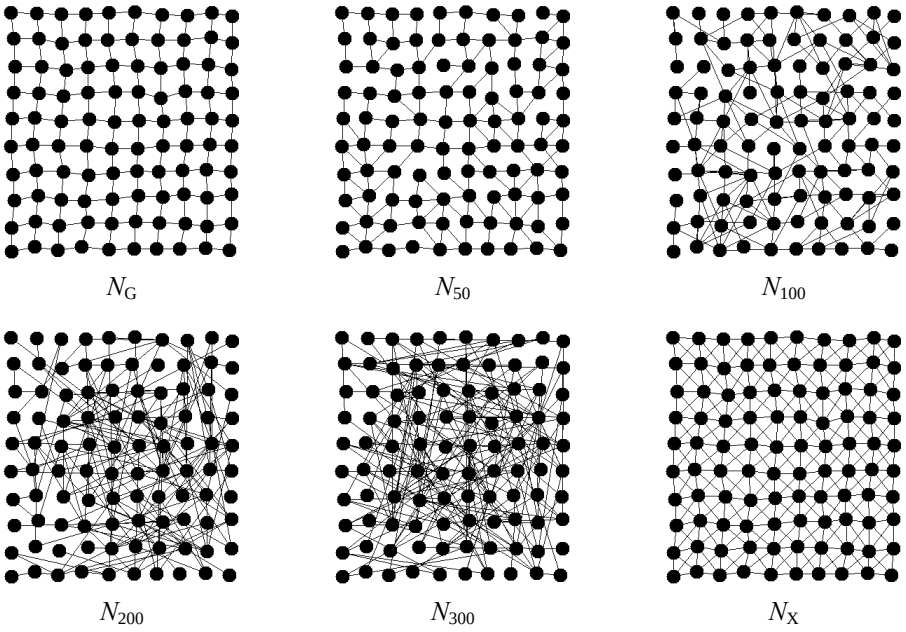

Figure 5. Network topologies used in experiment.

The two networks N_G and N_{50} also had the best power consumption, with (on average) about 54% of nodes awakened and (on average) 137 messages sent in total.

On the other hand, these two networks were also the slowest in obtaining a result. That is, the wave of activation took longest to die out on these networks. However, these results show that good performance can be achieved with random short-range links between nodes.

For the networks of the form N_d (with links connected at random between pairs of nodes at most a distance d apart), the number of errors increased with d. Indeed, for N_{200} and N_{300}, networking actually increased the number of errors above the baseline, because nodes were "nearby" in the network to physically distant nodes. Integrating sensor data from these physically distant nodes produced incorrect results.

Table 2. Simulation results, averaged over 1000 simulation runs for each network.

Net	Errors	Nodes awakened	Msgs sent	Total steps
N_G	–45.7%	55.2	139	9.11
N_{50}	–40.3%	53.1	135	8.55
N_{100}	–10.7%	66.2	168	7.48
N_{200}	58.1%	76.9	183	7.81
N_{300}	79.6%	78.2	177	8.35
N_X	–18.0%	76.1	336	7.82

The number of nodes awakened (and hence the total power consumption) also increased with d for the networks N_d. The time to obtain a result on these networks initially decreased with d (as a result of the decreasing average path length), and later increased (as a result of the increasing number of nodes awakened).

Performance of the denser grid network N_X was worse than expected. This shows that as long as a node has enough neighbours to provide reasonable error-correction, increasing the number of links does not necessarily improve performance. The practical application of this is that the radio transmission power of nodes can be limited to a value that ensures a sufficient average number of neighbours. More detailed simulation would be required to determine this threshold.

5. CHARACTERISTICS OF THE ACTIVATION WAVE

Figure 6 shows the average number of currently active and broadcasting (that is, just recently awakened) nodes over time for the six network topologies. Initial growth cannot be distinguished from linear growth for the first few steps (as the wave of activation moves through the plume), and is significantly faster for N_X, which has a higher average degree. The peak reached is lowest for N_G and N_{50}, and higher for the other four

networks. Activation decays exponentially for each case (as the wave of activation moves outside the plume).

Figure 6 suggests an analogy with the *SIR* equations (Giesecke, 2002) for modelling a disease epidemic, similar to that in De and Das (2008). Table 3 summarises this analogy.

The *SIR* equations are as follows:

$$\frac{dS}{dt} = -\alpha SI \tag{2}$$

$$\frac{dI}{dt} = \alpha SI - \beta I \tag{3}$$

$$\frac{dR}{dt} = \beta I \tag{4}$$

However, these equations result in an initially exponential growth of I, while Figure 6 shows approximately linear initial growth, as would be expected for an expanding circle of active nodes within the plume area. In fact, the data in Figure 6 fits the equation:

$$\frac{dI}{dt} = \zeta_t - \eta I \tag{5}$$

where the exponential decay factor η is approximately 0.4, and ζ_t is positive for the first few steps (as the "wave of activation" expands through the plume) and zero afterwards. There is a discontinuous change in ζ_t from its initial positive value down to zero as the wave reaches the plume boundary. The initial value of ζ_t is related to the average network degree, ranging from 3.2 (for the networks N_G and N_{50}) to 4.2 for N_{200} and 7.2 for N_X.

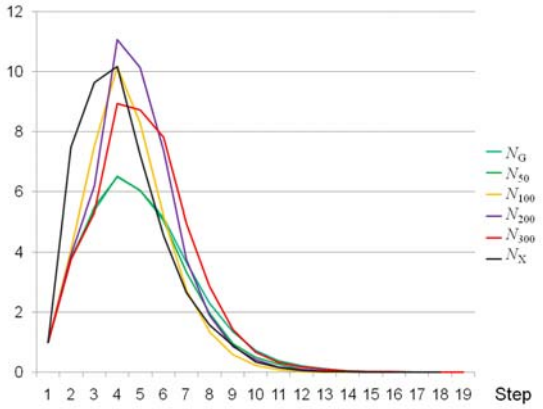

Figure 6. Average number of recently "waked" nodes over time for six network topologies.

Table 3. Sensor-network interpretation of the *SIR* epidemiological model.

	Epidemiological interpretation	**Sensor-network interpretation**
S	Susceptible people	Inactive nodes
I	Infected people (for a few days)	Recently "waked" nodes (for 1 timestep)
R	Recovered people	Activated nodes (and finished broadcasting sensor data to neighbours)

The curves in Figure 6 could also be modelled (a little less accurately) as differentials of sigmoid functions, and this might be the best fit for less perfectly circular plumes, where different parts of the activation wave reach the plume boundary over a wider time range.

An even more sophisticated model would include spatial dynamics of the sensor nodes, and therefore would include space derivatives (traditional diffusive terms) in the left-hand sides of equations (2) to (4). Analysis of such a model is outside the scope of this paper, and will be published elsewhere. However, it is worth noting here that, under the simplifying assumptions of homogeneity and isotropy, such a system can be reduced to the well-known Fisher-Kolmogorov equation, which also allows a quite general solution in the form of an activation wave with an exponential tail.

We plan to conduct further modelling along these lines, which will provide predictions of the configuration parameters required for the network to come to a correct decision for various sensor networks. This would include optimisation of average degree and average path length, and prediction of the time required for the network to come to a decision (i.e. for the wave of activation to die out).

6. DISCUSSION

We have presented an agent-based simulation of a sensor network in which, to optimise power consumption, sensors are normally quiescent, but can "wake" each other if contaminant is detected. A wave of activation therefore spreads through such a network, much like the spread of infection through a population. In addition, exchanging information with other nodes allows the network to compensate for sensor detection errors.

We have used this model to examine the performance of six different network topologies for connecting the sensor nodes.

The best-performing networks were the square grid N_G and the network N_{50} with short-range random links. These two networks most effectively compensated for sensor errors, while minimising the overall power consumption, as a result of not "waking up" unnecessary nodes. In general, random links in a sensor network appear to be effective, as long as the distance between linked nodes is small compared to the size of the contaminant plume. These are, of course, the kind of links that would result from making random connections within a limited radio range.

The wave of activation which spreads through the network differs from traditional models of the spread of infection through a population, in that initial growth in the number of recently activated sensors is approximately linear, followed by an exponential decay. We intend to conduct further modelling of the dynamics of this wave, in order to better predict the time required for the network to come to a decision.

REFERENCES

De, P. and Das, S.K. (2008), "Epidemic Models, Algorithms and Protocols in Wireless Sensor and Ad-hoc Networks," in Boukerche, A. (ed), *Algorithms and Protocols for Wireless Sensor Networks*, Wiley.

Dekker, A.H. (2008), "Network Effects in Epidemiology," Proceedings of Simulation Technology and Training (SimTecT) conference, Simulation Industry Association of Australia, pp 39–44, ISBN: 0-9775257-4-0.

Giesecke, J. (2002), *Modern Infectious Disease Epidemiology*, 2nd edition, Hodder Arnold.

Gunatilaka, A., Ristic, B., Skvortsov, A., and Morelande, M. (2008), "Parameter Estimation of a Continuous Chemical Plume Source," 11th International Conference on Information Fusion, DOI 10.1109/ICIF.2008.4632408.

Skvortsov, A., Connell, R., Dawson, P., and Gailis, R. (2007), "Epidemic Modelling: Validation of Agent-based Simulation by Using Simple Mathematical Models," in Oxley, L. and Kulasiri, D. (eds) MODSIM 2007 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, December. ISBN: 978-0-9758400-4-7. Available electronically at www.mssanz.org.au/MODSIM07/papers/13_s20/EpidemicModeling_s20_Skvortsov_.pdf

Vandermeer, J.H. and Goldberg, D.E. (2003), *Population Ecology: First Principles*, Princeton University Press.

Zhao, F. and Guibas, L. (2004), *Wireless Sensor Networks*, Elsevier.