2005 International Symposium on Nonlinear
Theory and its Applications (NOLTA2005)
Bruges, Belgium, October 18-21, 2005

Designing the Structure of Sensor Networks with the Highly Optimized Tolerance Model

Takaya Niyano† and Takako Tsutsui‡

†Department of Micro System Technology, Ritsumeikan University
Noji-higashi 1-1-1, Kusatsu, Shiga 525-8577 Japan
‡National Institute of Public Health
Minami 2-3-6, Wako, Saitama 351-0197, Japan,
Email: tmiyano@se.ritsumei.ac.jp, tsutsui@niph.go.jp

Abstract—The highly optimized tolerance model is applied to generating the secure structure of a sensor network as a defensive resource to a hypothetical bioterrorism. The performance of the network is assessed in terms of the power law between the number of victims and its probability of occurrence under the prior information about where to set to the attack.

1. Introduction

Power laws between the magnitude of events and their cumulative probability are conspicuous traits of real-world complex, interconnected systems. They can be viewed as a signature of robustness of complex systems to uncertainties [1]. Recently, Carlson and Doyle discovered that the power laws can come out from broader classes of complex systems, introducing a new concept for generating robust systems, referred to as highly optimized tolerance (HOT) [2]–[5]. This approach requires neither specific interactions between the constituents nor specific geometrical symmetry to generate the structure of a system. Instead, appropriate resources whose utilities are designed in relation to the size of event are requested to prevent failures or attacks from damaging the system. The optimal structure is captured by searching the minimum of an empirical risk functional in much the same way as many other optimization problems. Its performance is measured in terms of the scaling property between the magnitude of event and its cumulative probability. Thus, a highly structured system is derived to guarantee high performance and robustness to designed-for uncertainties. The chosen state concurrently has hypersensitivity to design flaws and unanticipated uncertainties. These features are characteristic of the HOT state.

The present paper is concerned with a simplified model of bio-terrorism as a toy problem settled to develop a practical model for designing secure systems by embodying the HOT state. In this problem, a network of hypothetical bio-sensors is given as the resource to avert the persons staying in an hypothetical airport terminal from being infected with pathogenic virus or bacteria planted by a terrorist. The secure structure of the network is determined so as to minimize the size of infection with the prior information about the spread in the likely position of the pathogenic source. The performance of the searched structure is assessed in terms of the achievement of the HOT state. The present toy problem might also be interesting in terms of current concern of the society, when thinking of the SARS epidemic in Eastern Asia and the bio-terrorism with anthrax in the United States [6, 7].

2. A Simplified Model of Bio-terrorism

Given a square floor partitioned by an $N \times N$ grid representing a hypothetical airport terminal, let us denote the number of persons staying at the (i, j) node as $n(i, j)$ where i and j express the position of the node in the x and y directions, respectively. Every variable and function is defined on each node of the grid in this model. For simplicity, $n(i, j)$ are assumed to be independent of time. Prior information about the spread in the likely position where to plant the pathogenic source may be given by the probability density function of the form:

$$\begin{aligned} p_v(i,j) &= p_v(i)p_v(j) \\ &= \frac{1}{p_{total}} \exp\left[-\frac{(i-i_v)^2}{2\sigma_x^2}\right] \exp\left[-\frac{(j-j_v)^2}{2\sigma_y^2}\right] \end{aligned} \quad (1)$$

$$p_{total} = \sum_{i,j=0}^{N-1} p_v(i,j) \quad (2)$$

where (i_v, j_v) expresses the position where the source is the most likely to be planted, and σ_x^2, σ_y^2 are the variances in the x and y directions, respectively. The probability density function is assumed to be independent of time.

Let us express the density of pathogen in the source planted at (i, j) as m_0. The diffusion is assumed to be isotropic for simplicity. The density of pathogen at

(x, y) may be given by:

$$m(x,y) = m_0 \exp\left(-\frac{d}{d_0}\right) \quad (3)$$

$$d = \sqrt{(x-i)^2 + (y-j)^2} \quad (4)$$

where d_0 is the critical distance of diffusion and d is the distance from the source.

The probability of infection at (x, y) is dependent on the density $m = m(x, y)$ $(0 \leq m \leq m_0)$, being assumed to be of the form:

$$\begin{aligned} p_{inf}(x,y) &= p_{inf}(m) \\ &= \frac{p_\infty}{1+\exp[-\alpha(m-m_{th})]} \\ &\quad - \frac{p_\infty}{1+\exp(\alpha m_{th})} \quad (5) \\ p_\infty &= \left[\frac{\exp(\alpha m_{th})}{1+\exp(\alpha m_{th})}\right]^{-1} \quad (6) \\ m \to \infty &\Rightarrow p_{inf}(m) \to 1 \quad (7) \end{aligned}$$

where α and m_{th} are constants representing infectiousness. Figure 1 illustrates the functional dependence of p_{inf} on m for $\alpha = 6$, $m_0 = 1$, and $m_{th} = m_0/2$. With these parameters, $p_{inf}(m_0) = 0.95$ (95 % of persons would be infected with the pathogen of the density m_0.) and $p_{inf}(m_0/2) = 0.475$. This set of parameters is utilized in this work.

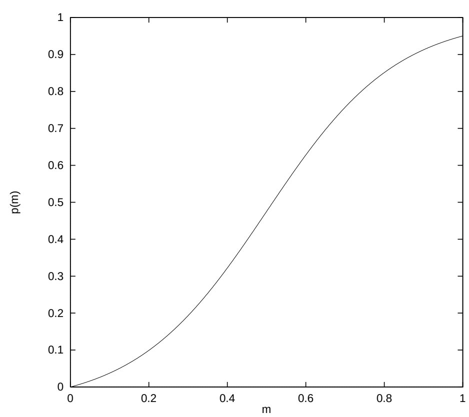

Figure 1: Probability of infection as a function of the density of pathogen. $\alpha = 6$, $m_0 = 1$, and $m_{th} = m_0/2$.

Under these assumptions, the number of persons at (x, y) infected with the pathogen coming from the source at (i, j) is given by

$$N_{inf} = p_v(i,j) p_{inf}[m(x,y)] n(x,y) \ . \quad (8)$$

Hypothetical bio-sensors are utilized as the resources to avert epidemic outbreak. They can be settled at any node in the floor. Immediately after a sensor makes response, a bulkhead will be shut down at the corresponding node to isolate the subarea surrounded by the bulkheads and to prevent the diffusion of pathogen. For simplicity, we may impose the restriction on the arrangement of the sensors that they are settled in line crossing the full span of the floor to generate a horizontal (in the x direction) or a vertical (in the y direction) alignment. Let us assume Q_x horizontal and Q_y vertical alignments be available. Two horizontal and vertical alignments are fixed on the periphery of the whole floor. The optimal positions of the remaining $Q_x - 2$ and $Q_y - 2$ alignments are to be determined in this toy problem. Let us denote the set of rectangular subareas as $R = \{A_1, \cdots, A_S\}$ where $S = (Q_x - 1) \times (Q_y - 1)$. The expected value of total infected persons over the whole floor is estimated by

$$E(R) = \sum_{s=1}^{S} \sum_{(i,j)\in A_s} p_v(i,j) \sum_{(x,y)\in A_s} p_{inf}[m(x,y)] n(x,y) \quad (9)$$

where $A_s \in R, s = 1, \cdots, S$. The set R that achieves the minimum of E, i.e., argmin $E(R)$, can correspond to the HOT state of the sensor network.

Let us express the probability density function for u infected persons as $P(u)$ and the cumulative probability function for $u > U$ as

$$P_{cum}(U) = \int_{u>U} P(u) du \ . \quad (10)$$

The HOT state provides the following relationship of power law:

$$P_{cum}(U) \propto U^{-\gamma} \quad (11)$$

where γ is a power-law exponent proper to the system [2]. In order for equation (11) to hold, it is sufficient to satisfy

$$P(u) \propto u^{-\gamma-1} \ . \quad (12)$$

The achievement of the HOT state can thus be confirmed with the scaling property inferred from the log-log plot of $P(u)$ versus u.

3. Numerical Experiments

We numerically solved the toy problem under the prior information about where to plant the pathogenic source. A floor of 64×64 grid $(N = 64)$ was given, being divided into four subfloors. To each subfloor, its population was assigned as follows:

$$\begin{aligned} 0 \leq i < X_{mid}, 0 \leq j < Y_{mid} &\rightarrow N_{11} \\ 0 \leq i < X_{mid}, Y_{mid} \leq j < N &\rightarrow N_{12} \\ X_{mid} \leq i < N, 0 \leq j < Y_{mid} &\rightarrow N_{21} \\ X_{mid} \leq i < N, Y_{mid} \leq j < N &\rightarrow N_{22} \end{aligned}$$

where X_{mid} and Y_{mid} indicate the boundaries between subfloors in the x and y directions, respectively, and

N_{ij} $(i, j = 1, 2)$ are the subtotal populations, given as $X_{mid} = 25, Y_{mid} = 30$, and $N_{11} = 550, N_{12} = 300, N_{21} = 100, N_{22} = 50$ to express a heterogeneous floor. The density of population was assumed to be uniform in each subfloor. The pathogen was supposed to diffuse in the long distance: $d_0 = N/4$ to simulate a dangerous situation. The most likely position of the pathogenic source (i_v, j_v) and its variances σ_x, σ_y were given as prior information. The number of horizontal and vertical alignments was set to $Q_x = 5, Q_y = 5$.

Figure 2 displays the optimal alignments determined by the HOT model for $(i_v, j_v) = (18.5, 21.5)$ and $\sigma_x = \sigma_y = N/16$. The most likely position of the source is in the subfloor having the largest density of population. The variances simulate low uncertainty in the prior information. The HOT model generates alignments intensively surrounding (i_v, j_v). To estimate the scaling property between $P(u)$ and u, $P(u)$ is approximated by summing up $p_v(i, j)$ over $(i, j) \in A_s$ for each subarea A_s, and u by the mean over the infected persons within A_s:

$$u = \sum_{x,y \in A_s} p_{inf}[m(x, y)]n(x, y) \qquad (13)$$

$$P(u) = \sum_{i,j \in A_s} p_v(i, j) \; . \qquad (14)$$

These equations may provide the worst-case estimation. The log-log plot of $P(u)$ versus u is shown in Fig.3. It seems to suggest the achievement of the HOT state in the network structure. Linear fitting of the plot provided an estimate of the power-law exponent as $\gamma = 1.579$.

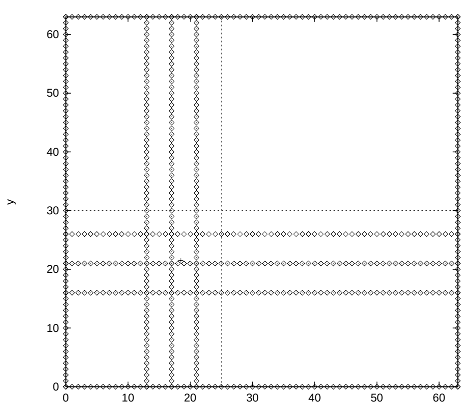

Figure 2: Optimal alignments of the bio-sensors determined by the HOT model (◇). The most likely position to plant the pathogenic source is indicated by +: $(i_v, j_v) = (18.5, 21.5)$ and $\sigma_x = \sigma_y = N/16$.

Figure 4 illustrates the corresponding results for $(i_v, j_v) = (38.5, 40.5)$ and $\sigma_x = \sigma_y = N/4$. The most likely position of the source is in the subfloor

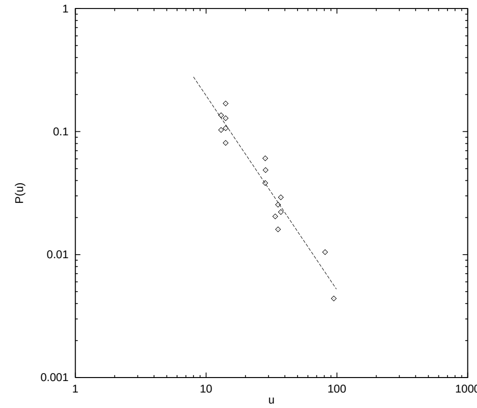

Figure 3: Number of infected persons versus its probability.

having the smallest density of population. The variances simulate high uncertainty in the prior information. The HOT model allows the chosen alignments to shift away from (i_v, j_v) to the subfloors of higher density of population. The log-log plot is shown in Fig.5. The HOT state seems to have been achieved in the chosen structure. The power-law exponent was estimated as $\gamma = 0.870$.

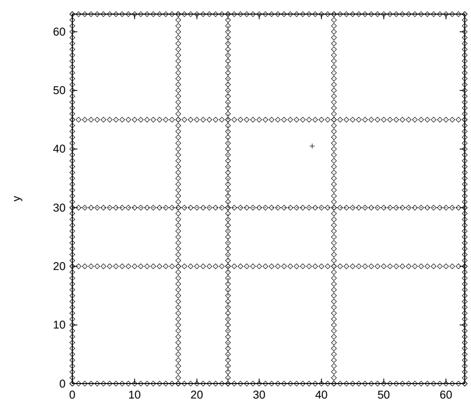

Figure 4: Optimal alignments of the bio-sensors determined by the HOT model (◇). The most likely position to plant the pathogenic source is indicated by +: $(i_v, j_v) = (38.5, 40.5)$ and $\sigma_x = \sigma_y = N/4$.

4. Discussion

When thinking of that bio-terrorism intends to cause personal damage, $(i_v, j_v) = (18.5, 21.5)$ with small variances may represent reasonable prior information for system design, since it can be interpreted as that the source of pathogen is likely to be planted in a crowded place. Then, the HOT model appropriately selects the alignments of bio-sensors to intensively

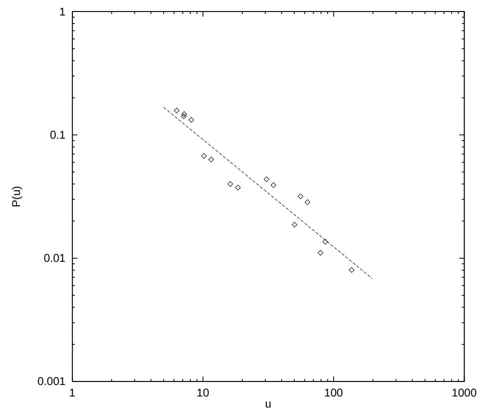

Figure 5: Number of infected persons versus its probability.

monitor crowded areas. The estimated power-law exponent suggests that the chosen structure is effective to avert large events.

On the contrary, the prior information $(i_v, j_v) = (38.5, 40.5)$ saying that the source of pathogen is likely to be planted in the least crowded place seems to be contradictory to the intention of bio-terrorism. Hence, the prior information may simulate questionable information. This can be mitigated by introducing the large variances ($\sigma_x = \sigma_y = N/4$) that address a suspicion to the most likely position of the source. This allows circumventing inappropriate choice in which the defensive resources are intensively settled in the least crowded area. The estimated power-law exponent suggests, however, that the chosen structure is less effective to avert large events. This may reflect the inappropriateness of the prior information about (i_v, j_v).

The present results, as a whole, suggest that power laws between the size of an event and its probability of occurrence as functions of the structure of a sensor network can be a useful measure to assess the performance of the network in terms of robustness to uncertain attacks, when the network is utilized as a defensive resource for the security of the society and the environment. In such situations, the HOT model is a powerful tool for generating the optimal structure of sensor networks.

5. Conclusion

In this work, we solved the toy problem of bioterrorism to generate the secure structure of a sensor network as the defensive resource. The numerical solutions have demonstrated how the HOT model works in the problem. When the prior information is reasonable despite containing some degrees of uncertainty, a highly structured robust network is chosen. These suggest a novel approach to designing sensor networks of various use in uncertain environments.

In the present case studies, we have imposed the restriction that bio-sensors be arranged in line to form horizontal and vertical alignments. This restriction can be removed to allow the position of a sensor to be determined independently of its neighboring sensors. In such a deregulated situation, computational burden becomes immense to search the optimal structure. A fast algorithm for searching the minimum of $E(R)$ will be necessary.

Acknowledgments

This study was supported by grants-in-aid (No.15500126) from the Japanese Ministry of Education, Science and Culture.

References

[1] R. Albert, H. Jeong, and A. -L. Barabasi, “Error and Attack Tolerance of Complex Networks,” *Nature*, vol.406, pp.378–382, 2000.

[2] J. M. Carlson and J. Doyle, “Highly Optimized Tolerance: A Mechanism for Power Laws in designed Systems,” *Phys. Rev.* E, vol.60, pp.1412–1427, 1999.

[3] J. M. Carlson and J. Doyle, “Highly Optimized Tolerance: Robustness and Design in Complex Systems,” *Phys. Rev. Lett.*, vol.84, pp.2529–2532, 2000.

[4] J. Doyle and J. M. Carlson, “Power Laws, Highly Optimized Tolerance, and Generalized Source Coding,” *Phys. Rev. Lett.*, vol.84, pp.5656–5659, 2000.

[5] T. Zhou and J. M. Carlson, “Dynamics and Changing Environments in Highly Optimized Tolerance,” *Phys. Rev.* E, vol.62, pp.3197–3204, 2000.

[6] C. Dye and N. Gay, “Modeling the SARS Epidemic,” published on-line 23 May 2003; 10.1126/science.1086925, 2003.

[7] L. G. Kun, “Homeland Security: The Possible, Probable, and Perils of Information Technology,” *IEEE Engineering in Medicine and Biology*, vol.21, pp.28–33, 2002.