Temporal-Spatial Surveillance Techniques from Non-Homogenous Random Geometric Graphs

James Dunyak[1], Ph.D., Mojdeh Mohtashemi[1, 2], Ph.D., Kenneth Mandl[3], M.D.
[1]The MITRE Corporation; [2]MIT AI and CS Lab; [3] Children's Hospital Boston

OBJECTIVE

This paper uses geometric random graph concepts to develop early detection algorithms for the real-time detection and localization of outbreaks.

BACKGROUND

Graph theory concepts are well established in epidemiology, with particular success as a description of agent-based modeling. An agent-based viewpoint leads to conclusions about the spatial distribution of links: infection is more likely among individuals in close proximity. In this analysis, we seek evidence of these temporal-spatial links though the properties of random geometric graphs.

Our investigation begins with the interpoint distance distribution (IDD) approaches in [1] and [2], which provide a promising approach to detect outbreaks that are localized in both space and time. Using a Mahalanobis-based metric, this distribution is compared to an expected distribution derived from historical records.

Unfortunately, when applied to a complex data set such as from Children's Hospital Boston (CHB), the IDD provides inadequate power. Emergency Department (ED) chief complaints from 1/1/2000-12/31/2004 were used to identify patients with infectious respiratory illness based on a triage process.

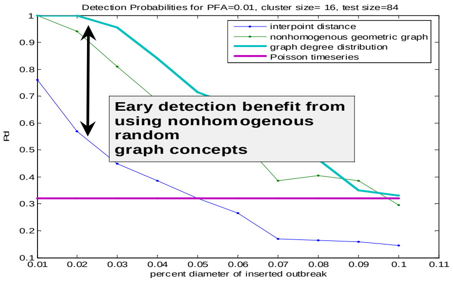
As in most realistic catchments, the historic density of patients varies greatly over the catchment area.

METHODS

To address this issue of varying density in early detection algorithms, we consider a random geometric graph viewpoint. In a geometric random graph, points are scattered across a region. All points within a specified radius are connected, and the number of graph edges at this radius is the cumulative IDD at this radius. Similarly, the degree of a node can be related to a k nearest neighbors statistic. To develop a non-homogenous graph model, we control the edge formation process based on local patient density in the historical record, in a similar fashion to k nearest neighbors.

RESULTS

The random geometric graph viewpoint lends itself to many detection approaches. The edges, node degrees, clustering coefficients, and other graph features provide useful tools for detecting localized outbreaks.



Eary detection benefit from using nonhomogenous random graph concepts

Fig. 1 – Early detection (<7 days) probabilities for a 16 patient outbreak, false positive probability=0.01

Detection and localization algorithms are derived for statistics based on edge count, node degree distribution, maximum node degree, clustering coefficient, and so on. Both the outbreak center and spatial extent are estimated and spatially homogenous and non-homogenous algorithms are compared. The increased power of non-homogenous graph-based techniques is clearly demonstrated, especially for early detection. Figure 1 demonstrates some of theses results.

The random graph approaches increase localized outbreak detection probabilities, during the first five-seven days of the outbreak, by more than 50% over interpoint distance metrics and 70% over time domain only approaches.

CONCLUSIONS

Concepts from random geometric graphs provide a useful tool in spatial-temporal syndromic surveillance. Using random geometric graphs, we unify previous viewpoints as well as develop new early detection and localization algorithms. In particular, we establish the increased detection power from algorithms that explicitly account for the varying patient density in a hospital's catchment.

REFERENCES

[1] Bonetti M, Olson K, Mandl K, Pagano M. Parametric models for interpoint distances and their use in biosurveillance. Proceedings of the American Statistical Association, Biometrics Section. 2003.

[2] Bonetti M, Pagano M. The interpoint distance distribution as a descriptor of point patterns, with an application to spatial disease clustering. Stat Med. Mar 15 2005;24(5):753-773.