

**National Syndromic Surveillance Conference  
New York City  
September 2002**

**-Next Generation of Potential Outbreak Detection and Prioritization System-**

**Comment and Discussion: G. P. Patil**

Success of syndromic surveillance initiative rests on potential outbreak detection capability. This capability largely appears to rely today on circle-based and ellipse-based SaTScan, a very novel tool indeed. But clusters are clusters. They can be of any shape, and cannot be captured only by circles and ellipses. This is likely to give more of false alarms and more of false sense of security.

What we further need is capability to detect arbitrarily shaped clusters. The proposed upper level set scan statistic system innovation is expected to fill the need and provide a timely next generation of hotspot detection and delineation system for geoinformatic surveillance for something unusual of interest. The partial order set prioritization system provides the ranking innovation based on multiple indicators without having to integrate the indicators.

# **Geographic and Network Surveillance for Arbitrarily Shaped Hotspots -Next Generation of Potential Outbreak Detection and Prioritization System-**

**G. P. Patil, W. L. Myers, C. Taillie, and D. Wardrop**  
**Pennsylvania State University, University Park, PA 16802**  
**Email: gpp@stat.psu.edu**

We present a version of the spatial scan statistic that is intended to address the following shortcomings of circle-based scans:

- Circles are able to capture only compactly shaped clusters. In many applications, clusters can have a very irregular shape, for example, in a disease outbreak along a winding river valley or other natural landscape feature. Similarly, cylindrical zones can yield poor hotspot delineation in a space-time scan.
- The circle-based zonation relies on Euclidean distance and is inappropriate for data defined along a network, such as a stream or subway system.
- The spatial scan statistic yields a maximum likelihood point estimate for the hotspot but provides no assessment of the uncertainty or variability in the delineation of that hotspot. One would like to have alternative plausible delineations of the hotspot expressed as a hotspot confidence set.

Our version of the scan statistic employs the notion of an upper level set (ULS) and is accordingly called the ULS scan statistic. Consider a syndromic response rate defined over the cells of a tessellation (or over the nodes of a network). An upper level set is the collection of cells (or nodes) at which the response rate exceeds some specified level. Upper level sets do not have to be connected, but each can be written as the disjoint union of connected components. In our approach, candidate zones for hotspots consist of all connected components of all upper level sets. These zones comprise a rooted tree under set inclusion and the likelihood function is defined over this ULS tree. The likelihood function determines not only the maximum likelihood point estimate for the hotspot but also a confidence set of hotspots on the tree by the usual duality between confidence sets and hypothesis testing. A hotspot-membership rating is assigned to individual cells (or nodes) according to the proportion of hotspot zones in the confidence set that include the cell (or node) in question.

Attractive features of the ULS scan statistic include:

- Identifies arbitrarily shaped clusters
- Data-adaptive zonation of candidate hotspots
- Applicable to data on a network
- Provides both a point estimate as well as a confidence set for the hotspot
- Uses hotspot-membership rating to map hotspot boundary uncertainty
- Computationally efficient
- Applicable to both discrete and continuous syndromic responses
- Identifies arbitrarily shaped clusters in the spatial-temporal domain
- Provides a typology of space-time hotspots with discriminatory surveillance potential

We also present a prioritization innovation. It lies in the ability for prioritization and ranking of hotspots based on multiple indicator and stakeholder criteria without having to integrate indicators into an index, using Hasse diagrams and partial order sets.

## Website Links

1. Prospectus 8: Synoptic Surveillance  
<http://www.stat.psu.edu/~gpp/PDFfiles/Prospectus-8.pdf>
2. Prospectus 11: Network-Based Surveillance  
<http://www.stat.psu.edu/~gpp/PDFfiles/Prospectus-11.pdf>
3. Prospectus 10: Classification and Prioritization  
<http://www.stat.psu.edu/~gpp/PDFfiles/Prospectus-10.pdf>
4. Prospectus 9: Crop Surveillance  
<http://www.stat.psu.edu/~gpp/PDFfiles/Prospectus-9.pdf>
5. Prospectus Abstract Syndromic Surveillance  
<http://www.stat.psu.edu/~gpp/PDFfiles/prospectus-12.pdf>
6. Proof-of-Concept Paper-1  
<http://www.stat.psu.edu/~gpp/PDFfiles/TR2002-0501.pdf>
7. Proof-of-Concept Paper-2  
<http://www.stat.psu.edu/~gpp/PDFfiles/TR2002-0801.pdf>
8. Proof-of-Concept Paper-3  
<http://www.stat.psu.edu/~gpp/PDFfiles/TR2001-1204.pdf>
9. Background Biographics 1  
<http://www.stat.psu.edu/~gpp/PDFfiles/Patil-3-page%20bio.pdf>
10. Background Biographics 2  
<http://www.stat.psu.edu/~gpp/PDFfiles/GP%20NSF%20Bio.pdf>

You are invited to participate in a manner you like in this promising cross-disciplinary cross-fertilization of ecosystem health geoinformatics and public health geoinformatics to accomplish next generation of potential outbreak detection and prioritization system needed for synoptic and network syndromic surveillance for arbitrarily shaped clusters.

G. P. Patil

Distinguished Professor of Mathematical and Environmental Statistics

Director, Center for Statistical Ecology and Environmental Statistics

Editor-in-Chief, Environmental and Ecological Statistics

Department of Statistics

421 Thomas Building

The Pennsylvania State University

University Park, PA 16802

Webpage: <http://www.stat.psu.edu/~gpp>

Telephone: 814/865-9442

Fax: 814/865-1278

Email: [gpp@stat.psu.edu](mailto:gpp@stat.psu.edu)