ISDS
INTERNATIONAL SOCIETY
FOR DISEASE SURVEILLANCE

Tracking Dynamic Water-borne Outbreaks with Temporal Consistency Constraints

Skyler Speakman*, Yating Zhang and Daniel B. Neill

Carnegie Mellon University, Pittsburgh, PA, USA

Objective

We describe a novel graph-based event detection approach which can accurately identify and track dynamic outbreaks (where the affected region changes over time). Our approach enforces soft constraints on temporal consistency, allowing detected regions to grow, shrink, or move while penalizing implausible region dynamics. Using simulated contaminant plumes diffusing through a water distribution system, we demonstrate that our method improves both detection time and spatial-temporal accuracy when tracking dynamic water-borne outbreaks.

Introduction

Space-time scan statistics are often used to identify emerging spatial clusters of disease cases [1,2]. They operate by maximizing a score function (likelihood ratio statistic) over multiple spatio-temporal regions. The temporal component is typically incorporated by aggregating counts across a given time window, thus assuming that the affected region does not change over time. To relax this hard constraint on spatial-temporal "shape" and increase detection power and accuracy when tracking spreading outbreaks, we implement a new graph-based event detection approach which enables identification of dynamic clusters while enforcing temporal consistency constraints between temporally-adjacent spatial regions.

Methods

In the subset scanning framework, temporal consistency constraints may be interpreted as influencing the prior probability p_i^t of location i being included in the optimal spatial subset at time t. We model this prior probability for each location as

$\log(p_i^t/(1-p_i^t)) = B_0 + B_1 X_i^{t-1}$ where X_i^{t-1} is 1 if location i was included in the previous time step and 0 otherwise, and maximize the penalized log-likelihood ratio over dynamic spatio-temporal regions. Our efficient algorithm incorporates these constraints into the Graph-Scan method [3] by iteratively optimizing the spatial subset for each time slice conditioned on the previous and next slices. Each individual optimization step is made possible by expressing the score function as an additive function (conditioned on the relative risk), which enables the priors to be included while maintaining computational efficiency.

Results

Outbreak plumes were simulated in a water distribution system for 12 one-hour periods. We assumed noisy binary sensors (with 10% false positive and 90% true positive rates) observed hourly at each pipe junction. Our method ("Dynamic") was compared to the "Static" method, which aggregates counts across time for each spatial region and is therefore constrained to only return temporal cylinders, and the "Independent" method, which separately optimizes the spatial subset for each time slice without taking temporal consistency into account. The methods were evaluated on spatial-temporal overlap (Figure 1), defined as the number of sensors contained in both the detected and affected space-time regions divided by the number of

sensors in either the detected or affected space-time regions. A measure of 1 is a perfect match of spatial subsets across each time window and 0 would reflect disjoint space-time regions. Additionally, average time to detect an outbreak (at a fixed false positive rate of 1/month) was 4.24, 4.56, and 6.65 hours for the dynamic, static, and independent methods respectively.

Conclusions

Relaxing constraints on spatial-temporal region shape must be done carefully. Allowing independent selection of spatial regions loses important temporal information while hard constraints on the spatial-temporal region will fail to capture the dynamics of the outbreak. Our approach for detecting dynamic space-time clusters, while incorporating temporal consistency constraints, addresses these issues and results in higher spatial-temporal accuracy and detection power.
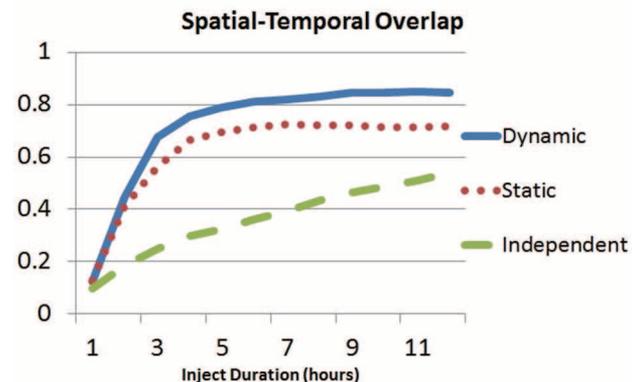
Spatial-temporal overlap for three competing detection methods.

Keywords

outbreak detection; space-time scan statistics; dynamic event tracking; penalized likelihood ratio

Acknowledgments

References

[1] Kulldorff M. A spatial scan statistic. Communications in Statistics: Theory and Methods, 1997, 26(6): 1481-1496.
[2] Neill DB. Fast subset scan for spatial pattern detection. Journal of the Royal Statistical Society (Series B), 2012, 74(2):337–360.
[3] Speakman S and Neill DB. Fast graph scan for scalable detection of arbitrary connected clusters. Advances in Disease Surveillance, 2010.

*Skyler Speakman
E-mail: skylerspeakman@gmail.com