Data-Adapted Temporal Alerting Algorithms for Routine Health Monitoring

Howard Burkom[1], John Copeland[2], Sean Murphy[1], Rick Hu[3], Jerry Tokars[2]
[1]The Johns Hopkins Applied Physics Lab, [2]Centers for Disease Control and Prevention
[3]Science Applications International Corporation

Objective

This paper discusses selection of temporal alerting algorithms for syndromic surveillance to achieve reliable detection performance based on statistical properties and the epidemiological context of the input data. We used quantities calculated from brief data history to derive criteria for algorithm selection.

Background

The data streams selected for biosurveillance depend upon syndromic, spatial, and temporal aggregation decisions. These decisions are driven by medical knowledge and by the available data. The need for relevant monitoring capability dictates that these decisions drive, rather than be driven by, the selection of alerting algorithms.

Given the selection of these data streams, alerting algorithms involve four steps: preconditioning, computation of expected values, computation of test statistic, and application of threshold criteria for alerting. Most methods do not apply all four steps explicitly. Control-chart-based methods such as the EARS algorithms [1] use raw counts without preconditioning, regression approaches derive residual values and assume a zero expectation, etc. Syndromic data streams evolve with changes in informatics systems, in syndromic definitions, and in the population behavior generating the data. Published studies generally focus on a specific background dataset and evaluate one or several algorithms in that context; there is no universally accepted algorithm in any data environment, and no algorithm will perform well in all contexts. A classification methodology is required to determine an appropriate algorithm for detecting a relevant class of signals within a current data stream, allowing for dynamic case definition.

Methods

Syndromic data from the BioSense program of the Centers for Disease Control and Prevention were chosen for this study. These data were aggregated at the treatment facility level and included 2.5 years of daily counts. The syndrome groupings and facility types and sizes yielded time series of varying scale, serial correlation, seasonality, and day-of-week behavior.

For each of these time series, we tabulated a set of descriptors including the mean, median, coefficient of variation, and selected autocorrelation coefficients. We chose a set of detection algorithms to represent several modeling and control-chart strategies. The modeling methods included Poisson and negative binomial regression using sliding, fixed-length baselines.

Following a Monte Carlo procedure used in previous studies [2], we applied the algorithms repeatedly to selected time series with injected stochastic signals. Sensitivity values were recorded for practical background alert rates to compare algorithm detection performance. We then used the descriptors to group the time series according to detection performance in search of simple criteria for algorithm selection.

Results

The combination of the tabulated descriptors with the detection performance results allowed us to distinguish regimes of time series behavior such that specific algorithms are indicated for best detection performance in each regime. For data streams beyond a median value of 20 counts/day and with significant day-of-week and seasonal behavior, the preconditioning step improved the ability to detect the simulated outbreak effects, and careful covariate selection sharpened it further. For other data regimes, control-chart-based methods, especially when modified to manage autocorrelation, were more effective. For the regime of sparse data with daily median counts equal to zero, temporal scan statistics gave the best detection results. Performance differences were not uniform over all time series in each regime, but the clustering of performance results by descriptors was distinct.

Conclusions

From these results, simple data classification schemes may be automated to choose an effective algorithm for a given data stream. For most data regimes, this classification requires only 4-8 weeks of data and is readily improved with heuristic knowledge. When indicated, models for preconditioning may be improved as more data history becomes available. Such schemes may be used for both initialization and regular maintenance. The derived criteria are subject to further refinement, but for robust monitoring, a balance should be maintained between improved sensitivity and overfitting.

References

[1] Hutwagner L, Thompson W, Seeman GM, Treadwell T. (2003).The Bioterrorism Preparedness and Response Early Aberration Reporting System (EARS). J Urban Health. 2003 Jun 1;80 Supplement 1:I89-I96.

[2] Burkom HS, Murphy S, Coberly J, Hurt-Mullen K (2005). "Public health monitoring tools for multiple data streams." Morbidity Mortality Weekly Report 2005 Aug 26;54 Suppl:55-62.

Further Info.: Howard Burkom, Howard.Burkom@jhuapl.edu