Refinement of a Population-Based Bayesian Network for Fusion of Health Surveillance Data

Howard Burkom*, Yevgeniy Elbert, Liane Ramac-Thomas, Christopher Cuellar and Vivian Hung

Johns Hopkins Applied Physics Laboratory, Laurel, MD, USA

Objective

The project involves analytic combination of multiple evidence sources to monitor health at hundreds of care facilities. A demonstration module featuring a population-based Bayes Network [1] was refined and expanded for application in the Department of Defense Electronic Surveillance System for Community-Based Epidemics (ESSENCE).

Introduction

The ESSENCE demonstration module was built to help DoD health monitors make routine decisions based on disparate evidence sources such as daily counts of ILI-related chief complaints, ratios of positive lab tests for influenza, patient age distribution, and counts of antiviral prescriptions [1]. The module was a population-based (rather than individual-based) Bayesian network (PBN) in that inputs were algorithmic results from these multiple aggregate data streams, and output was the degree of belief that the combined evidence required investigation. The module reduced total alerts substantially and retained sensitivity to the majority of documented outbreaks while clarifying underlying sources of evidence. The current effort was to advance the prototype to production by refining components of the fusion methodology to improve sensitivity while retaining the reduced alert rate.

Methods

The multi-level approach to sensitivity improvement included expanded syndromic queries, more data-sensitive algorithm selection, improved transformation of algorithm outputs to alert states, and hierarchical training of Bayesian networks. Components were tested individually, and the net result was iteratively refined with performance using documented outbreaks.

We examined time series of classes of prescribed drugs and laboratory tests during known events and discussed outbreak-associated elements with domain experts to liberalize data queries. Algorithms were matched to data streams with injection testing applied to 4.5 years of data from 502 outpatient clinics. A hierarchical approach was applied for improved training and verification of PBNs for events related to categories of Influenza-like Illness, Gastrointestinal, Fever, Neurological, and Rash, chosen both for public health importance and for availability of multiple supporting data types. Hierarchical, modular training was applied to common subnetworks, such as a severity indicator PBN depending on case disposition, acute case indicators, complex evaluation/management codes, and patient bounce-backs, depicted in Figure 1. Conversion of individual algorithm outputs to belief states (e.g. "at least two red alerts/past 7 days") was broadened using analysis of lags between data sources. With data from the known events, we calculated decision support thresholds for the parent-level PBN decision nodes with a stochastic optimization technique maximizing the ratio of alert rates during outbreak to non-outbreak periods.

Results

The expanded data queries, more stream-specific algorithm selection, generalized state transformation, and hierarchical PBN training detected 22 of an expanded collection of 24 documented outbreaks, with incremental improvement ongoing. The mean alert rate drop achieved by the Bayes Net was 87% (minimum of 85%) compared to the combined alerts of all component algorithms across syndromes and facilities.

Conclusions

Expansion and further technical validation upheld the PBN approach as a user-friendly means of analytic decision support given multiple, variably weighted evidence sources. The PBN affords not only sharply reduced alerting, but also transparent indication of evidence underlying each alert. The older algorithm approach remains available as backup. Beta testing of the resulting production system will drive further modification.

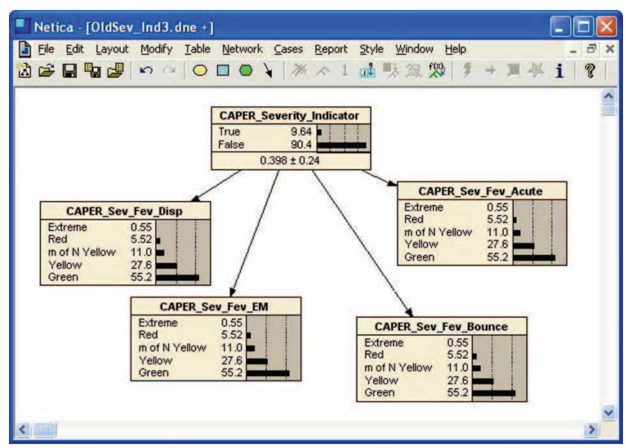

Figure 1: PBN Subnetwork for Event Severity, based on Outpatient Data Fields

Keywords

Fusion; Bayesian Network; Multivariate; Decision Support

Acknowledgments

Drs. Julie Pavlin and Rhonda Lizewski of the Armed Forces Health Surveillance Center for data- and event-related consultation and Joe Lombardo and Wayne Loschen of Johns Hopkins APL for consultation on production enhancement.

References

Burkom H, Elbert Y, Ramac-Thomas L et al., Analytic fusion of ESSENCE clinical evidence sources for routine decision support, Emerging Health Threats Journal Supplements, eISSN 1752-8550, ISSN 2001 1350 (print).

*Howard Burkom
E-mail: howard.burkom@jhuapl.edu