Detecting Coherent Structures in Syndromic Time Series using Wavelets

James Dunyak[1], Miriam Nuño [2]

[1]MITRE Corporation, Bedford MA; [2]Harvard School of Public Health, Boston MA

OBJECTIVE

Temporally localized outbreaks occur in the presence of a complex background, greatly complicating both retrospective and real-time detection. Numerous techniques have been proposed for adjusting thresholds to account for this variable background. In this paper, we apply wavelet transforms to detect localized structures in health care time series, using a generalization of many of these viewpoints. A rigorous, nonparametric approach is applied in a general setting to identify coherent outbreaks.

BACKGROUND

By localizing signals in time and scale, the wavelet transform has become a popular tool for analyzing coherent structures in many settings. With time series $x(t)$ and analyzing function $\psi(t)$ we define wavelet coefficients at time *b* and scale *a* as

$$W_x(b,a) = a^{-1/2} \int_{-\infty}^{+\infty} x(t)\psi\left(\frac{t-b}{a}\right)dt.$$

Note that the wavelet coefficients are matched filters between the signals and shifted and scaled versions of the analyzing function, which is chosen to explore specific signal types. In our example, we use the Haar wavelet to detect localized sustained changes in signal level characteristic of an outbreak.

$$\psi_{\text{Haar}}(t) = \begin{cases} 1 & 0 \le t \le 1 \\ -1 & -1 \le t \le 0 \\ 0 & \text{otherwise} \end{cases}$$

Wavelet based time-scale representations of the signals do offer promise, but substantial interpretation difficulties remain [1]. The key to addressing this problem is the development of rigorous statistical detection techniques.

METHODS

Our approach provides a highly nonparametric statistical detection test, based on the concept of an incoherent signal [1]. Incoherent signals result, for example, from driving linear filters with white noise. These signals are at the heart of AR, ARMA, moving window averages, spectral methods, and many other popular syndromic surveillance approaches to background modeling. The energy in temporally localized outbreaks, in contrast, requires phase dependency to focus energy in time. Thus we seek localized coherent structures in the signal.

This statistical technique is based on one of the oldest methods in nonparametric statistics: the development of a randomized reference distribution. To build our randomized reference distribution, we use phase-randomized versions of the syndromic time series measurement. See [2] for details. Retrospective detection is performed in a fixed window, while real-time detection applies a sliding window.

RESULTS AND CONCLUSIONS

We have applied this approach to identify, with high statistical confidence, localized outbreaks in several confidential syndromic data sets, including hospital emergency department visits coded for respiratory infection and ambulatory care visits coded for several syndromes. Here we illustrate the technique with U.S. pneumonia and influenza death rates from 1996-2006 (www.cdc.gov/mmwr). Figure 1(a) shows the time series, while 1(b) shows the resulting wavelet coefficients. Note the intriguing but difficult to interpret two-dimensional signal representation. Figure 1(c) applies the coherent structure detector [2], with p-values describing the probability a false alarm at a specified scale anywhere in the time window. Note the strong detections of localized outbreaks, indicated by small p-values.

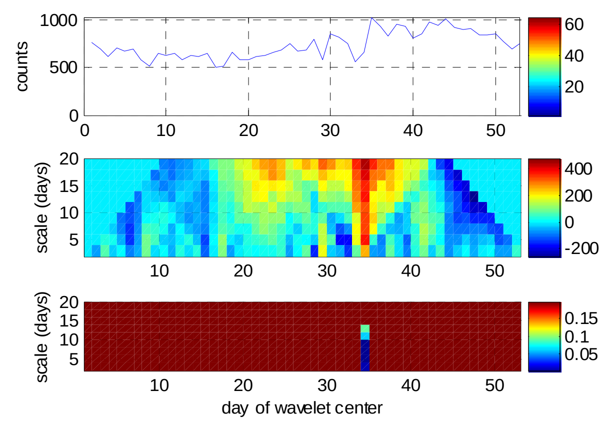

Figure 1 (a) 2005 time series (b) wavelet transform (c) p-values

REFERENCES

[l] J. Dunyak, X. Gilliam, R. Peterson and D. Smith, “Coherent Gust Detection by Wavelet Transform,” J. Wind Engineering and Industrial Aeradynamics, 77/78 p. 467-478, 1998.

[2] X. Gilliam, “Statistical techniques for identification of coherent structures by wavelet analysis,” PHD Dissertation, Texas Tech University, 1998. 77/78 p. 467-478, 1998.