Characterizing Repeated Outbreak Patterns in Influenza Time Series

Miriam Nuño[1] , James Dunyak[2]

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

OBJECTIVE

In some influenza seasons, morbidity and mortality closely follow the expected seasonal variation. In these years, approaches such as Serfling's model and seasonal-based syndromic outbreak detectors, in use in EARS, work well. In other years, though, short but intense variations occur in addition to the longer term seasonal variation. These intense outbreaks, which are often multimodal, have important implications for both syndromic surveillance and influenza epidemiology. Unfortunately, they are both difficult to characterize and poorly understood.

In this paper, we apply techniques from time-frequency distribution theory to identify the temporal location, duration, and amplitude of intense outbreaks occurring in the presence of longer time scale variations.

BACKGROUND

Time-frequency analysis techniques allow us to decompose influenza time series to isolate and study these periods of multimodal outbreak interrupting the nearly periodic seasonal variation. Different time-frequency analysis methods tile the time-frequency plane in different ways. In our situation, the spectrogram or short-time periodogram, with uniform tiling in time and frequency, is the best match for the signals of interest.

METHODS

For an influenza case count time series $x(t)$, and windowing function $w(\tau)$ with finite support centered at $\tau = 0$, we calculate spectrogram [1] coefficients

$$A(t,\omega) = \left\| \int_{-\infty}^{\infty} x(t-\tau) w(\tau) e^{i\omega\tau} d\tau \right\|^2$$

The support (region of τ where $w(\tau) > 0$) describes the amount of time localization, which in turn controls the frequency resolution. Note that the spectrogram essentially takes a local Fourier transform, to identify energy localized in time and frequency. In practice, the integral is approximated for discrete time series with a Fast Fourier Transform. The time, frequency, and amplitude of peaks in the spectrogram characterize the multimodal outbreaks, while the phase of the integral provides more precise timing information.

RESULTS

This approach has been applied to several confidential syndromic data sets, including hospital emergency department visits coded for respiratory infection and ambulatory care visits coded for several syndromes. In those cases, Poisson regression was first applied to adjust for day-of-week variation. Here we illustrate the technique with U.S. pneumonia and influenza mortality time series from 1996-2006 (www.cdc.gov/mmwr). A Hamming window of length 20 weeks in applied

Figure 1 provides an example of the results. Note the intense oscillations in fall 2001 are clearly characterized in frequency and time in the spectrogram. The 2002/2003 season has both strong oscillations in the early fall, and separately, during the winter peak.

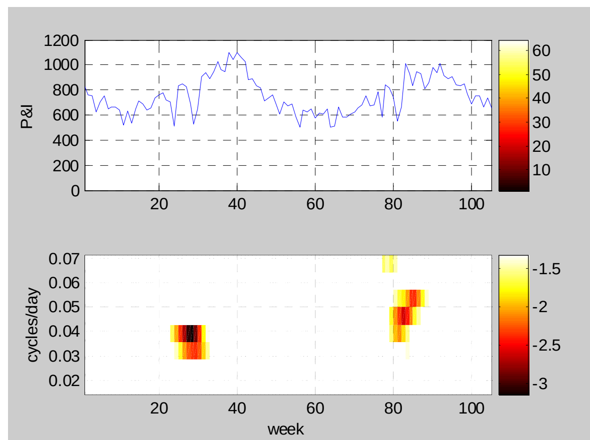

Fig. 1- (a) P&I time series from the 2001/2002 and 2002/2003 influenza seasons. The time scale is in weeks, starting at week 15 of 2001. (b) The resulting spectrogram $\log_{10}(A(t,\omega))$.

CONCLUSIONS

In this paper, we demonstrate how time-frequency analysis can be used to find and describe instabilities in syndromic and other healthcare time series. This work is part of our ongoing research to identify and characterize these instabilities, to improve both epidemiological understanding and baseline behavior for real-time syndromic surveillance.

REFERENCES

[1] Cohen, L. "Time-Frequency Distributions- a Review," Proc. of the IEEE, Volume 77, Issue 7, July 1989 Page(s):941 – 981.