Enabling Syndromic Surveillance in Pakistan

Ross Maciejewski*[1], Shehzad Afzal[2], Adam J. Fairfield[1], Arif Ghafoor[2], David S. Ebert[2], Naeem Ayyaz[4] and Maaz Ahmed[3]

[1]Computer Science, Arizona State University, Tempe, AZ, USA; [2]Purdue University, West Lafayette, IN, USA; [3]King Edward Medical University, Lahore, Pakistan; [4]University of Engineering and Technology, Lahore, Pakistan

Objective

This work presents our first steps in developing a Global Real-time Infectious Disease Surveillance System (GRIDDS) employing robust and novel in-fectious disease epidemiology models with real-time inference and pre/exercise planning capabilities for Lahore, Pakistan. The objective of this work is to address the infectious disease surveillance challenges (specific to developing countries such as Pakistan) and develop a collaborative capability for monitoring and managing outbreaks of natural or manmade infectious diseases in Pakistan.

Methods

Utilizing our partner hospitals in the Lahore, Punjab area, we have begun developing a theoretical model of patient hospital visits with respect to diseases and syndromes within Pakistan. Our first thrust has focused on the collection, categorization and cleansing of data based on expert knowledge from our partnering institutions in Pakistan. Data consists of a patient's home address and chief complaint which is then categorized into syndromes. Home addresses are geocoded utilizing the Google API with a resultant 72% accuracy. Unknown geolocations are aggregated only at the hospital level. Using this cleaned data, we employ methods similar to our previous work [1] on syndromic surveillance for early disease detection. Currently, we have collected over 600,000 patient records over 1.5 years.
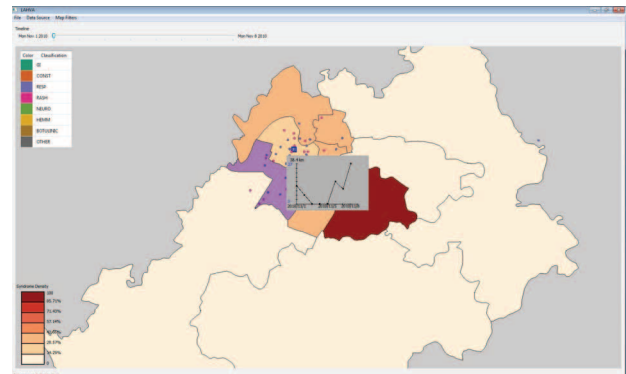
We employ the use of choropleth maps, isopleth maps utilizing kernel density estimation of patient addresses, traditional control chart methods such as exponentially weighted moving averages (EWMA), and a non-parametric time series analysis approach (seasonal trend decomposition using loess smoothing (STL) [2]) which requires only 90 days of historical data to be put into operation. The time series models are deployed as part of a real-time surveillance system in which temporal anomalies over regions can be analyzed and disease outbreaks reported.

Results

Figure 1 illustrates our visual analytics toolkit in operation. Here we see the location of our partner hospital in the Lahore region. The hospital coverage is in the most populous location of the city, providing data as a sentinel site for the overall health of the city. Currently, our system employs the use of interactive filters and linked isopleth or choropleth maps with time series analysis on mouse over.

Conclusions

Currently our research has focused on one partner location within the city of Lahore. Our ongoing work is focusing on the adoption of such a system to other regions of the country and the development of disease spread simulations (particularly Dengue Fever) utilizing baseline data collected by our partners. We plan to integrate these models into our visual analytics system for real-time planning and simulation.

Keywords

syndromic surveillance; visual analytics; pakistan

References

[1] Maciejewski, R., et al.., "Forecasting Hotspot – A Predictive Analytics Approach," IEEE Transactions on Visualization and Computer Graphics, 17(4): 440-453, 2011.

[2] Hafen, R. P., et al., "Syndromic surveillance: STL modeling, visualizing and monitoring disease counts," BMC Medical Informatics and Decision Making, 2009.

*Ross Maciejewski
E-mail: rmacieje@asu.edu