Incorporating Water Security into Syndromic Surveillance

S. Babin MD PhD[1], H. Burkom PhD[1], Z. Mnatsakanyan PhD[1], L. Ramac-Thomas PhD[1], M. Thompson PhD[1], R. Wojcik MS[1], J. Lombardo MS[1], K. Clayton[2]

[1]*Johns Hopkins University Applied Physics Laboratory*, [2]*US Environmental Protection Agency*

OBJECTIVE

The objective of this paper is to illustrate a technique for combining water quality and population-based health data to monitor for water-borne disease outbreaks.

BACKGROUND

Although rare in the US, the CDC reports 13-14 drinking-water-related disease outbreaks per year, affecting an average of about 1000 people [e.g., 1]. The US EPA has determined that the distribution system is the most vulnerable component of a drinking water system. Recognizing this vulnerability, water utilities are increasingly measuring disinfectant levels and other parameters in their distribution systems. The US EPA is sponsoring an initiative to fuse this distribution system water quality (WQ) data with health data to improve surveillance by providing an assessment of the likelihood of the occurrence of a waterborne disease outbreak. This fused analysis capability will be available via a prototype water security (WS) module within a population-based public health syndromic surveillance system.

METHODS

The WS module applies Bayesian Belief Network (BBN) technology [2] using historical data and expert knowledge from public health professionals, physicians, and water engineers. Three types of BBNs are used: a Health BBN, a WQ BBN, and a Fusion BBN. The Health BBN is applied to population health data (e.g., emergency department chief complaints) to seek anomalous levels of gastrointestinal-related (GI) illness potentially caused by waterborne pathogens. The WQ BBN combines algorithm outputs for various water parameters described below to test for anomalies at the zone level. WQ data are provided by the participating public water utilities and are analyzed in the context of sampling frequencies and their respective distribution zones (distribution system sectors that have similar water characteristics depending on the hydrostatic pressure and distance from the treatment plant). A modified version of the Canary algorithm suite [3] is used for analysis of individual water sensor data (e.g., disinfectant levels). These algorithms compare recent parameter values with baseline values derived from nearby and historically similar sensors identified with a divisive hierarchical clustering algorithm. Finally, a Fusion BBN is applied to the outputs of both the Health and WQ BBNs to assess the degree of belief that degraded WQ is or is not associated with illnesses. The final output is presented as a summary alert webpage showing the outputs of the three BBNs over the previous eight days. Webpage alert hyperlinks are built in for maps, time series and other details.

RESULTS

Early evaluations have focused on efficient fusion of the WQ and health data. Figure 1 shows the chosen structure for the WQ BBN. The WQ spatial cluster analysis yielded 5-7 water characteristic regions. These clusters enabled the inclusion of sensors with intermittent data and provided a stable baseline. As an example of the algorithm detection performance achieved for WQ BBN input with this stabilization, the measured probability of detection for transient signals was 0.69 (std. dev. = 0.10) at a background alert rate of 1 every 190 days.

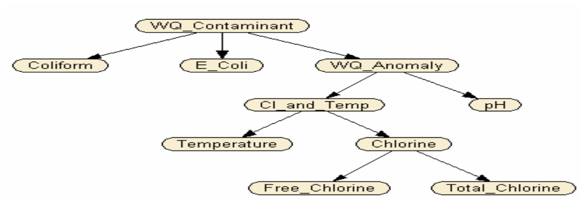

Figure 1 – Example of nodes for the Water Quality BBN.

CONCLUSIONS

Using WQ data for health monitoring must account for operational characteristics of the distribution system while restricting attention to sub-syndromes characteristic of waterborne disease. The nested BBN approach allows fusion of water information with dissimilar population health data to encapsulate both engineering and epidemiological expertise.

ACKNOWLEDGMENTS

The authors would like to thank Sean McKenna and David Hart at Sandia National Laboratories for sharing an early version of the CANARY Event Detection Software. The expert knowledge of the Washington Suburban Sanitary Commission, Milwaukee Water Works, and the Milwaukee, Montgomery County, and Prince Georges County Health Departments is gratefully appreciated.

REFERENCES

[1] Liang JL et al, Surveillance for waterborne disease and outbreaks associated with drinking water and water not intended for drinking – United States, 2003-2004. MMWR 55(SS-12), 31-68. Dec. 2006.

[2] Pearl, Judea. "Fusion, propagation, and structuring in belief networks." Artificial Intelligence 29(3):241–288, 1986.

[3] McKenna SA, Wilson M, Klise KA, On change detection in water quality data, JAWWA, in press, 2007.

Further Information:
Steven Babin, steven.babin@jhuapl.edu