Learning Specific Detectors of Adverse Events in Multivariate Time Series

Josep Roure, Artur Dubrawski, Jeff Schneider

The Auton Lab, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 15213

OBJECTIVE

This paper describes how powerful detectors of adverse events manifested in multivariate series of bio-surveillance data can be learned using only a few labeled instances of such events.

BACKGROUND

The context of the work presented here is rapid detection of statistically significant emerging patterns of adverse events in data related to food-and agriculture-safety collected by the U.S. Department of Agriculture. The particular data under consideration includes records of daily counts of condemned and healthy cattle, counts of positive and negative microbial tests of food samples, and counts of passed and failed sanitary inspections of slaughter houses. Effective monitoring of those streams of heterogeneous data is instrumental in early detection of adverse food events and in their subsequent mitigation.

METHOD

We use temporal scan [2] as a basic detection tool. It slides a fixed-width time window along a time series and compares the positive and negative counts inside of it against the aggregated counts observed outside, and applies either Chi-square or Fisher's exact test of significance of the obtained contingency table. In our approach, temporal scan is being applied individually to each of the available data-streams and the resulting series of p-values are then combined using Fisher's method of p-values aggregation [3]. Detectors based on Fisher's method benefit multivariate analysis by being able to raise an alert even if none of the component signals is critical, but if some of them are near critical. They are non-specific because they target any departure of the combined series from normal, and they are not tailored to any specific scenario.

	FP when DtD fixed to			DtD when FP fixed to			
	2	2.5	3	2	3	4	5
A	65.65	33.9	18.6	5.22	4.49	4.04	3.9
B	114.48	79.81	45.15	7.45	6.4	4.98	4.84
C	62.15	45.92	35.1	5.04	4.83	4.66	4.55
F	35.76	16.35	7.2	3.36	3.25	3.19	3.14
F+	na	8.73	3.24	3.05	3.01	2.97	2.93
F+FP	9.27	4.18	0.67	2.84	2.64	2.52	2.42
F+TP1	na	2.36	1.22	2.78	2.39	2.3	2.22
F+TP5	na	2.31	0.82	2.7	2.28	2.18	na
F+TP10	na	2.04	0.94	2.59	2.26	na	na
L	2.32	1.43	0.13	2.18	1.81	1.66	1.56

Table 1 – Number of Potential False Positives (FP) and Detection Delay in days (DtD) for different detection methods: temporal scan for individual streams (A, B, C), Fisher's aggregation method (F), hand-crafted specific filter (F+), Fisher's combined with filters learned from data using labeled false positives only (F+FP), the same using FP and one, five, and ten labeled outbreaks (F+TP1, FP+TP5, FP+TP10), and classification based specific detector (L).

Knowing particular signatures of events of interest, one can design more powerful detectors to specifically target them. Sometimes, such dedicated detectors have to be hand-crafted using domain expertise, if the amount of available training data is insufficient to support automated learning [4]. Such situations happen very often in practice of bio-surveillance because collecting labeled data is costly and, more importantly, because instances of real adverse events are – luckily – rare. Therefore, there is a need for machine learning techniques which would allow for training specific detectors even if the number of identified positive examples in data is small.

RESULTS

Experimental results summarized in Table 1 indicate that Fisher's aggregation substantially improves detection power (timeliness as well as alert frequency) over the results obtained using individual streams. Hand-crafted specific detector slightly outperforms Fisher's method, but the best results were obtained with a classifier [5] trained using sample evidence (100 labeled outbreaks). Much more practical alternative is based on kernel density model [6] trained with labels on false positive detections produced by Fisher's method. It beats the hand-crafted detector, but looses against the classifier trained on rich dataset. Incremental improvement can be obtained by combining into the training data labels on false positives with one, then five and eventually ten labeled instances of true positives.

CONCLUSION

It is possible to use machine learning to construct powerful specific detectors of adverse events in multivariate time series data, even if the number of available labeled examples of such events is very small.

ACKNOWLEDGEMENTS This work was supported by the Centers of Disease Control (award R01-PH 000028) and based upon work that was supported by the National Science Foundation (grant IIS-0325581)

REFERENCES

[1] P.S. Mead, L. Slutsker, V. Dietz, L.F. McCaig, J.S. Bresee, C. Shapiro, P.M. Griffin, and R.V. Tauxe. Food-related illness and death in the U.S. Emerging Infectious Diseases, 5(5), 1999.

[2] J. Roure, A. Dubrawski, and J. Schneider. A study into detection of bio-events in multiple streams of surveillance data. In Biosurveillance 2007, LNCS Vol. 4506, Springer-Verlag, 2007.

[3] L. Wasserman. All of Statistics. Springer, 2004.

[4] A. Dubrawski, K. Elenberg, A. Moore, M. Sabhnani, Monitoring Food Safety by Detecting Patterns in Consumer Complaints. AAAI/IAAI'06, Boston, USA, July 2006.

[5] L. Breiman. Random forests. Machine Learning, 45(1):, 2001.

[6] A. Gray, D. Lee, and A.W. Moore. Dual-tree fast gauss transforms. Neural Information Processing Systems 2005.