Detecting Conversations Going Viral

Evaluation of time-aware change detection algorithms

Matt Chapman

matthew.chapman@student.uva.nl

May 1, 2017, 4 pages

Supervisor: One does not simply do a thesis without \supervisor!
Host organisation: Buzzcapture International, http://www.buzzcapture.com

Contents

Abstract

Bibliography

[1] L. Akoglu and C. Faloutsos. Anomaly, event, and fraud detection in large network datasets. *Proceedings of the sixth ACM international conference on Web search and data mining - WSDM '13*, page 773, 2013.

[2] F. Alvanaki, S. Michel, K. Ramamritham, and G. Weikum. EnBlogue: emergent topic detection in Web 2.0 streams. *Proc. ACM SIGMOD International Conference on Management of Data*, pages 1271–1274, 2011.

[3] M. Basseville and I. V. Nikiforov. *Detection of Abrupt Changes: Theory and Application*. 1993.

[4] S. Bersimis, S. Psarakis, and J. Panaretos. Multivariate statistical process control charts: An overview. *Quality and Reliability Engineering International*, 23(5):517–543, 2007.

[5] C. Buntain and C. Natoli. A Brief Comparison of Algorithms for Detecting Change Points in Data.

[6] T. Dasu, S. Krishnan, D. Lin, S. Venkatasubramanian, and K. Yi. Change (detection) you can believe in: Finding distributional shifts in data streams. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 5772 LCNS:21–34, 2009.

[7] F. Desobry, M. Davy, and C. Doncarli. An online Kernel change detection algorithm. *IEEE Transactions on Signal Processing*, 53(8):2961–2974, 2005.

[8] A. B. Downey. A novel changepoint detection algorithm. *Applied Microbiology and Biotechnology*, pages 1–11, 2008.

[9] T. Fawcett and F. Provost. Activity monitoring: Noticing interesting changes in behavior. *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, 1(212):53–62, 1999.

[10] J. Ginsberg, M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, and L. Brilliant. Detecting influenza epidemics using search engine query data. *Nature*, 457(7232):1012–4, 2009.

[11] Y. Kawahara and M. Sugiyama. Change-point detection in time-series data by direct density-ratio estimation. *Proceedings of the 2009 SIAM International Conference on Data Mining*, pages 389–400, 2009.

[12] D. Kifer, S. Ben-david, and J. Gehrke. Detecting Change in Data Streams. *Proceedings of the 30th VLDB Conference*, pages 180–191, 2004.

[13] M. Kulldorff, R. Heffernan, J. Hartman, R. Assun????o, and F. Mostashari. A space-time permutation scan statistic for disease outbreak detection. *PLoS Medicine*, 2(3):0216–0224, 2005.

[14] T. L. Lai and J. Z. Shan. Efficient recursive algorithms for detection of abrupt changes in signals and control systems. *IEEE Transactions on Automatic Control*, 44(5):952–966, 1999.

[15] C. Madrid, D. D. Estadística, U. Carlos, and I. I. I. D. Madrid. Variance Changes Detection in Multivariate Time Series. 2004.

[16] D. S. Matteson and N. A. James. A nonparametric approach for multiple change point analysis of multivariate data. *Submitted*, 14853:1–29, 2012.

[17] A. M. Pelecanos, P. a. Ryan, and M. L. Gatton. Outbreak detection algorithms for seasonal disease data: a case study using Ross River virus disease. *BMC medical informatics and decision making*, 10(1):74, 2010.

[18] A. A. Qahtan, B. Alharbi, S. Wang, and X. Zhang. A PCA-Based Change Detection Framework for Multidimensional Data Streams. *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '15*, pages 935–944, 2015.

[19] D. Siegmund and E. S. Venkatraman. Using the generalized likelihood ratio statistic for sequential detection of a change-point. *The Annals of Statistics*, 23(1):255–271, 1995.

[20] A. G. Tartakovsky and B. L. Rozovskii. A Nonparametric Multichart CUSUM Test for Rapid Intrusion Detection. *Proceedings of Joint Statistical Meetings*, pages 7–11, 2005.

[21] A. G. Tartakovsky, B. L. Rozovskii, R. B. Blažek, and H. Kim. Detection of intrusions in information systems by sequential change-point methods. *Statistical Methodology*, 3(3):252–293, 2006.

[22] D.-H. Tran, M. M. Gaber, and K.-U. Sattler. Change Detection in Streaming Data in the Era of Big Data: Models and Issues. *ACM SIGKDD Explorations Newsletter - Special issue on big data*, (1):30–38, 2014.

[23] A. S. Willsky and H. L. Jones. A Generalized Likelihood Ratio Approach to the Detection and Estimation of Jumps in Linear Systems. *IEEE Transactions on Automatic Control*, 21(1):108–112, 1976.

[24] Y. Xu, Z. Zhang, P. Yu, and B. Long. Pattern change discovery between high dimensional data sets. *Proceedings of the 20th ACM international conference on Information and knowledge management - CIKM '11*, page 1097, 2011.