Detecting Conversations Going Viral

Evaluation of time-aware change detection algorithms

Matt Chapman
matthew.chapman@student.uva.nl

May 1, 2017, 4 pages

Supervisor: Evangelos Kanoulas, Universiteit van Amsterdam
Host organisation: Buzzcapture International, http://www.buzzcapture.com

Universiteit van Amsterdam
Faculteit der Natuurwetenschappen, Wiskunde en Informatica
Master Software Engineering
http://www.software-engineering-amsterdam.nl

Contents

Abstract

Bibliography

[1] Leman Akoglu and Christos Faloutsos. "Anomaly, event, and fraud detection in large network datasets". In: *Proceedings of the sixth ACM international conference on Web search and data mining - WSDM '13* (2013), p. 773. DOI: 10.1145/2433396.2433496. URL: http://dl.acm.org/citation.cfm?id=2433396.2433496.

[2] Foteini Alvanaki et al. "EnBlogue: emergent topic detection in Web 2.0 streams". In: *Proc. ACM SIGMOD International Conference on Management of Data* (2011), pp. 1271–1274. ISSN: 07308078. DOI: 10.1145/1989323.1989473. URL: http://doi.acm.org/10.1145/1989323.1989473%7B%5C%%7D5Cnpapers2://publication/uuid/44E009D1-8574-49C1-A0AC-C2B247FE9043.

[3] M Basseville and Igor V Nikiforov. *Detection of Abrupt Changes: Theory and Application*. 1993. ISBN: 0-13-126780-9. DOI: 10.1016/0967-0661(94)90196-1. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.77.6896%7B%5C&%7Drep=rep1%7B%5C&%7Dtype=pdf.

[4] S. Bersimis, S. Psarakis, and J. Panaretos. "Multivariate statistical process control charts: An overview". In: *Quality and Reliability Engineering International* 23.5 (2007), pp. 517–543. ISSN: 07488017. DOI: 10.1002/qre.829.

[5] Cody Buntain, Christopher Natoli, and Miroslav Zivkovic. "A Brief Comparison of Algorithms for Detecting Change Points in Data". In: *Supercomputing*. 2014. URL: https://github.com/cbuntain/ChangePointDetection.

[6] Tamraparni Dasu et al. "Change (detection) you can believe in: Finding distributional shifts in data streams". In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 5772 LCNS (2009), pp. 21–34. ISSN: 03029743. DOI: 10.1007/978-3-642-03915-7_3.

[7] Frédéric Desobry, Manuel Davy, and Christian Doncarli. "An online Kernel change detection algorithm". In: *IEEE Transactions on Signal Processing* 53.8 (2005), pp. 2961–2974. ISSN: 1053587X. DOI: 10.1109/TSP.2005.851098.

[8] Allen B. Downey. "A novel changepoint detection algorithm". In: *Applied Microbiology and Biotechnology* (2008), pp. 1–11. arXiv: 0812.1237. URL: http://arxiv.org/abs/0812.1237.

[9] Tom Fawcett and Foster Provost. "Activity monitoring: Noticing interesting changes in behavior". In: *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining* 1.212 (1999), pp. 53–62. ISSN: 09258574. DOI: 10.1016/j.ecoleng.2010.11.031. URL: http://portal.acm.org/citation.cfm?id=312195.

[10] Jeremy Ginsberg et al. "Detecting influenza epidemics using search engine query data." In: *Nature* 457.7232 (2009), pp. 1012–4. ISSN: 1476-4687. DOI: 10.1038/nature07634. URL: http://www.ncbi.nlm.nih.gov/pubmed/19020500.

[11] Yoshinobu Kawahara and Masashi Sugiyama. "Change-point detection in time-series data by direct density-ratio estimation". In: *Proceedings of the 2009 SIAM International Conference on Data Mining* (2009), pp. 389–400.

[12] Daniel Kifer, Shai Ben-david, and Johannes Gehrke. "Detecting Change in Data Streams". In: *Proceedings of the 30th VLDB Conference* (2004), pp. 180–191. ISSN: 00394564. DOI: 10.1016/0378-4371(94)90421-9. arXiv: 9310008 [cond-mat].

[13] Martin Kulldorff et al. “A space-time permutation scan statistic for disease outbreak detection”. In: *PLoS Medicine* 2.3 (2005), pp. 0216–0224. ISSN: 15491277. DOI: 10.1371/journal.pmed.0020059.

[14] Tze Leung Lai and Jerry Zhaolin Shan. “Efficient recursive algorithms for detection of abrupt changes in signals and control systems”. In: *IEEE Transactions on Automatic Control* 44.5 (1999), pp. 952–966. ISSN: 00189286. DOI: 10.1109/9.763211.

[15] Calle Madrid et al. “Variance Changes Detection in Multivariate Time Series”. 2004.

[16] David S. Matteson and Nicholas A. James. “A nonparametric approach for multiple change point analysis of multivariate data”. In: *Submitted* 14853 (2012), pp. 1–29. ISSN: 0162-1459. DOI: 10.1080/01621459.2013.849605. arXiv: 1306.4933.

[17] Anita M Pelecanos, Peter a Ryan, and Michelle L Gatton. “Outbreak detection algorithms for seasonal disease data: a case study using Ross River virus disease.” In: *BMC medical informatics and decision making* 10.1 (2010), p. 74. ISSN: 1472-6947. DOI: 10.1186/1472-6947-10-74. URL: http://www.biomedcentral.com/1472-6947/10/74.

[18] Abdulhakim A. Qahtan et al. “A PCA-Based Change Detection Framework for Multidimensional Data Streams”. In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD ’15* (2015), pp. 935–944. DOI: 10.1145/2783258.2783359. URL: http://dl.acm.org/citation.cfm?id=2783258.2783359.

[19] D. Siegmund and E. S. Venkatraman. “Using the generalized likelihood ratio statistic for sequential detection of a change-point”. In: *The Annals of Statistics* 23.1 (1995), pp. 255–271. ISSN: 0090-5364. DOI: 10.1214/aos/1176324466.

[20] Alexander G Tartakovsky and Boris L Rozovskii. “A Nonparametric Multichart CUSUM Test for Rapid Intrusion Detection”. In: *Proceedings of Joint Statistical Meetings* (2005), pp. 7–11.

[21] Alexander G. Tartakovsky et al. “Detection of intrusions in information systems by sequential change-point methods”. In: *Statistical Methodology* 3.3 (2006), pp. 252–293. ISSN: 15723127. DOI: 10.1016/j.stamet.2005.05.003.

[22] Dang-Hoan Tran, Mohamed Medhat Gaber, and Kai-Uwe Sattler. “Change Detection in Streaming Data in the Era of Big Data: Models and Issues”. In: *ACM SIGKDD Explorations Newsletter - Special issue on big data* 1 (2014), pp. 30–38. ISSN: 1931-0145. DOI: 10.1145/2674026.2674031.

[23] Alan S. Willsky and Harold L. Jones. “A Generalized Likelihood Ratio Approach to the Detection and Estimation of Jumps in Linear Systems”. In: *IEEE Transactions on Automatic Control* 21.1 (1976), pp. 108–112. ISSN: 15582523. DOI: 10.1109/TAC.1976.1101146.

[24] Yi Xu et al. “Pattern change discovery between high dimensional data sets”. In: *Proceedings of the 20th ACM international conference on Information and knowledge management - CIKM ’11* (2011), p. 1097. DOI: 10.1145/2063576.2063735. URL: http://dl.acm.org/citation.cfm?doid=2063576.2063735.