





























**Table 9** Corresponding relationship between the event and the real event detected by the Enron network**表 9** Enron 网络检测出的事件与真实事件对应关系

事件编号	发生时间	具体事件描述
事件 1	2001 年 10 月 16 日	安然发布 2001 年第二季财务报表
事件 2	2001 年 10 月 22 日	美国证券交易委员会要求提交交易内容
事件 3	2001 年 10 月 31 日	美国证券交易委员会开始对安然进行正式调查,次日安然抵押了部分资产
事件 4	2001 年 11 月 8 日	安然承认做了假账,次日迪诺基公司宣布准备收购安然
事件 5	2001 年 11 月 28 日	标准普尔将安然调低至“垃圾债券”,30 日,安然股价跌至 0.26 美元
事件 6	2001 年 12 月 2 日	安然申请破产保护

综上,通过对 VAST 和 Enron 真实网络上事件检测分析,证明了基于 AUC 和基于 mAUC 的 FDA 在单个事件发生的网络和多个事件发生的网络上都能进行有效的事件检测,并且可以对不同事件对网络的影响进行定量评估.基于 AUC 的 FDA 对网络波动更加敏感,容易受到非事件引起的网络波动的干扰.基于 mAUC 的 FDA 由于考虑了节点的微观演化,能够更精准地检测出事件引起的波动,避免无效波动的干扰.

#### 4 结论及展望

为了检测网络事件,量化事件对社会网络演化产生的影响,本文提出了一种混合指标群智能方法 IndexEvent:利用最佳权重算法 OWA 来确定当前时段网络的最佳混合指标;然后,通过基于 AUC 或基于 mAUC 的网络波动检测算法 FDA 检测事件.为了验证 IndexEvent 方法的有效性,本文基于 WS 小世界网、BA 无标度网络、VAST 和 Enron 真实网络进行了大量实验探讨,并得出以下结论:

- (1) 对于特定真实网络,在最佳混合指标为某个单位指标权重为 1、其他指标权重均为 0 的特殊情况下,最佳混合指标和这个权重为 1 的独立相似性指标的链路预测精度一样.其他情况均可以通过 OWA 找到一个链路预测精度优于独立相似性指标的混合指标;并且,真实网络的演化机制越复杂,通过 OWA 得到最佳混合指标的优势越明显.
- (2) 基于 AUC 的 FDA 对网络演化波动更加敏感,非事件引起的正常网络波动容易对事件检测结果造成干扰.基于 mAUC 的 FDA 由于考虑了节点的微观演化,能够更好地避免正常网络波动的干扰,精准地检测出事件引起的网络异常波动.
- (3) 无论是在 WS 小世界网络和 BA 无标度网络的实例,还是在 VAST 和 Enron 的真实网络,IndexEvent 方法检测出的事件与真实事件有很好的匹配关系,证明了 IndexEvent 的准确性高,具有很好的实用性.进一步的研究仍然需要在以下两个方面继续:
  - (1) 尝试利用链路预测去确定真实网络中不同的网络演化机制所占的具体比重;
  - (2) 进一步探讨各种相似性指标及最大似然估计指标之间的关系.

**致谢** 我们真诚地向对本文的工作给予支持和建议的审稿人、主编、编辑、同行、老师和同学表示感谢.

#### References:

- [1] Reuter T, Bielefeld U, Papadopoulos S, Petkos G, Vries CD, Mezaris V. Social event detection at mediaeval 2013: Challenges, datasets, and evaluation. In: Proc. of the MediaEval Workshop. 2013. 18–19.
- [2] Hidalgo CA, Rodriguez-Sickert C. The dynamics of a mobile phone network. Physica A Statistical Mechanics & Its Applications, 2008,387(12):3017–3024. [doi: 10.1016/j.physa.2008.01.073]
- [3] Watts DJ, Strogatz SH. Collective dynamics of “small-world” networks. Nature, 1998,393(6684):440–442. [doi: 10.1038/30918]
- [4] Zhang Z, Wu B. Pfaffian orientations and perfect matchings of scale-free networks. Theoretical Computer Science, 2015,570(C): 55–69. [doi: 10.1016/j.tcs.2014.12.024]
- [5] Washio T, Motoda H. State of the art of graph-based data mining. AcmSigkdd Explorations Newsletter Homepage, 2003,15(1): 59–68. [doi: 10.1145/959242.959249]
- [6] Papadimitriou P, Dasdan A, Garcia-Molina H. Web graph similarity for anomaly detection. Journal of Internet Services & Applications, 2010,1(1):19–30. [doi: 10.1007/s13174-010-0003-x]

- [7] Mcculloh IA, Carley KM, Mcculloh IA, Carley KM. Social network change detection. Technical Report, CMU-ISR-08-116, Institute for Software Research, School of Computer Science, Carnegie Mellon University, 2008. [doi: 10.2139/ssrn.2726799]
- [8] Wan X, Milios E, Kalyaniwalla N, Janssen J. Link-Based event detection in email communication networks. In: Proc. of the SAC ACM Symp. on Applied Computing. 2009. [doi: 10.1145/1529282.1529618]
- [9] Taskar B, Abbeel P, Koller D. Discriminative probabilistic models for relational data. Eprint Arxiv, 2012. 485–492.
- [10] Aaron C, Christopher M, Newman MEJ. Hierarchical structure and the prediction of missing links in networks. *Nature*, 2008,453(7191):98–101. [doi: 10.1038/nature06830]
- [11] Liben-Nowell D, Kleinberg J. The link-prediction problem for social networks. *Journal of the American Society for Information Science and Technology*, 2007,58(7):1019–1031. [doi: 10.1002/asi.20591]
- [12] Kossinets G, Watts DJ. Reports empirical analysis of an evolving social network. *Science*, 2006,311(5757):88–90. [doi: 10.1126/science.1116869]
- [13] Sarukkai RR. Link prediction and path analysis using Markov chains. *Computer Networks*, 2000,33(1):377–386. [doi: 10.1016/S1389-1286(00)00044-X]
- [14] Zhu J, Hong J, Hughes JG. Using Markov chains for link prediction in adaptive Web sites. In: Proc. of the ACM SigWeb Hypertext. 2002. 60–73.
- [15] Clauset A, Moore C, Newman MEJ. Hierarchical structure and the prediction of missing links in networks. *Nature*, 2008,453(7191):98–101. [doi: 10.1038/nature06830]
- [16] Guimera R, Sales-Pardo M. Missing and spurious interactions and the reconstruction of complex networks. *Proc. of the National Academy of Sciences*, 2010,106(52):22073–22078. [doi: 10.1073/pnas.0908366106]
- [17] Liben-Nowell D, Kleinberg J. The link-prediction problem for social networks. *Journal of the American Society for Information Science & Technology*, 2007,58(7):1019–1031. [doi: 10.1002/asi.20591]
- [18] Rapoport A. Spread of information through a population with socio-structural bias: I. Assumption of transitivity. *The Bulletin of Mathematical Biophysics*, 1953,15(4):523–533. [doi: 10.1007/BF02476440]
- [19] Jaccard P. Etude Comparative de la distribution florale dans une portion des Alpes et du Jura. *Bulletin de La Societe Vaudoise Des Sciences Naturelles*, 1990,37(142):547–579.
- [20] Adamic LA, Adar E. Friends and neighbors on the Web. *Social Networks*, 2003,25(3):211–230. [doi: 10.1016/S0378-8733(03)00009-1]
- [21] Barabasi AAR. Emergence of scaling in random networks. *Science*, 1999,286(5439):509–512. [doi: 10.1126/science.286.5439.509]
- [22] Lichtenwalter RN, Lussier JT, Chawla NV. New perspectives and methods in link prediction. In: Proc. of the ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. Washington, 2010. 243–252. [doi: 10.1145/1835804.1835837]
- [23] Symeonidis P, Iakovidou N, Mantas N, Manolopoulos Y. From biological to social networks: Link prediction based on multi-way spectral clustering. *Data and Knowledge Engineering*, 2013,87(4):226–242. [doi: 10.1016/j.datak.2013.05.008]
- [24] Huang Z, Li X, Chen H. Link prediction approach to collaborative filtering. In: Proc. of the Joint Conf. on Digital Libraries. ACM Press, 2005. 7–11. [doi: 10.1145/1065385.1065415]
- [25] Rao J, Wu B, Dong YX. Parallel link prediction in complex network using MapReduce. *Ruan Jian Xue Bao/Journal of Software*, 2012,23(12):3175–3186 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4206.htm> [doi: 10.3724/SP.J.1001.2012.04206]
- [26] Salton G, McGill MH. Introduction to Modern Information Retrieval. New York: McGraw-H Hill, 1983.
- [27] Sørensen T. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Biol Skr*, 1948,5:1–34.
- [28] Ravasz E, Somera AL, Mongru DA, Oltvai ZN, Barabási AL. Hierarchical organization of modularity in metabolic networks. *Science*, 2002,297(5586):1551–1555. [doi: 10.1126/science.1073374]
- [29] Leicht EA, Holme P, Newman MEJ. Vertex similarity in networks. *Physical Review E*, 2006,73(2):026120. [doi: 10.1103/PhysRevE.73.026120]
- [30] Noble CC, Cook DJ. Graph-Based Anomaly Detection. In: Proc. of the ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining. Washington, 2003. 631–636. [doi: 10.1145/956750.956831]
- [31] Priebe CE, Conroy JM, Marchette DJ. Scan statistics on enron graphs. *Computational & Mathematical Organization Theory*, 2005, 11(3):229–247. [doi: 10.1007/s10588-005-5378-z]
- [32] Wu B, Wang B, Yang SQ. Framework for tracking the event-based evolution in social networks. *Ruan Jian Xue Bao/Journal of Software*, 2011,22(7):1488–1502 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3841.htm> [doi: 10.3724/SP.J.1001.2011.03841]
- [33] Baruah RD, Angelov P. Evolving social network analysis: A case study on mobile phone data. In: Proc. of the 2012 IEEE Conf. on Evolving and Adaptive Intelligent Systems (EAIS). IEEE, 2012. 114–120. [doi: 10.1109/EAIS.2012.6232815]

- [34] Qiao SJ, Tang CJ, Peng J, Liu W, Wen FL, Qiu JT. Mining key members of crime networks based on personality trait simulation email analysis system. Chinese Journal of Computer, 2008,31(10):1795-1803 (in Chinese with English abstract). [doi: 10.3321/j.issn:0254-4164.2008.10.014]
- [35] Kashima H, Abe N. A parameterized probabilistic model of network evolution for supervised link prediction. Trans. of the Japanese Society for Artificial Intelligence, 2006,22(2):340-349. [doi: 10.1109/ICDM.2006.8]
- [36] Hanley JA, Mcneil BJ. The meaning and use of the area under a receiver operating characteristic (roc) curve. Radiology, 1982, 143(1):29-36. [doi: 10.1148/radiology.143.1.7063747]
- [37] Herlocker JL, Konstan JA, Terveen LG, Riedl JT. Evaluating collaborative filtering recommender systems. ACM Trans. on Information Systems, 2004,22(1):5-53. [doi: 10.1145/963770.963772]
- [38] Zhou T, Ren J, Medo M, Zhang YC. Bipartite network projection and personal recommendation. Physical Review E (Statistical Nonlinear and Soft Matter Physics), 2007,76(4):70-80.
- [39] Bouwmeester D, Ekert A, Zeilinger A. The physics of quantum information. In: Proc. of the Quantum Cryptography Quantum Teleportation Quantum Computation. 2010.
- [40] Tang D, Cai Y, Zhao J, Xue Y. A quantum-behaved particle swarm optimization with memetic algorithm and memory for continuous non-linear large scale problems. Information Sciences, 2014,289(24):162-189. [doi: 10.1016/j.ins.2014.08.030]
- [41] Schutte JF, Groenwold A. A study of global optimization using particle swarms. Journal of Global Optimization, 2004,31(31):93-108. [doi: 10.1007/s10898-003-6454-x]
- [42] <http://www.cs.cmu.edu/~enron/2012/8/1>
- [43] Liu HK, Lü LY, Zhou T. Uncovering the network evolution mechanism by link prediction. Scientia Sinica PhysMechAstron, 2011, 41(7):816-823 (in Chinese with English abstract). [doi: 10.1360/132010-922]
- [44] <http://www.cs.umd.edu/hcil/VASTchallenge08/2012/10/8>

#### 附中文参考文献:

- [25] 饶君,吴斌,东昱晓.MapReduce 环境下的并行复杂网络链路预测.软件学报,2012,23(12):3175-3186. <http://www.jos.org.cn/1000-9825/4206.htm> [doi: 10.3724/SP.J.1001.2012.04206]
- [32] 吴斌,王柏,杨胜琦.基于事件的社会网络演化分析框架.软件学报,2011,22(7):1488-1502. <http://www.jos.org.cn/1000-9825/3841.htm> [doi: 10.3724/SP.J.1001.2011.03841]
- [34] 乔少杰,唐常杰,彭京,等.基于个性特征仿真邮件分析系统挖掘犯罪网络核心.计算机学报,2008,31(10):1795-1803. [doi: 10.3321/j.issn:0254-4164.2008.10.014]
- [43] 刘宏颀,吕琳媛,周涛.利用链路预测推断网络演化机制.中国科学:物理学力学天文学,2011,41(7):816-823. [doi: 10.1360/132010-922]



胡文斌(1976—),男,湖北武汉人,博士,副教授,博士生导师,主要研究领域为复杂网络,人工智能,调度优化.



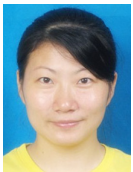
邱振宇(1992—),男,硕士生,主要研究领域为复杂网络,社会网络分析.



王欢(1989—),男,博士生,主要研究领域为社会网络分析,数据挖掘.



肖雷(1992—),男,硕士生,主要研究领域为社会网络分析,数据挖掘.



严丽平(1980—),女,博士生,主要研究领域为社会网络分析,数据挖掘



杜博(1983—),男,博士,教授,博士生导师,CCF 专业会员,主要研究领域为复杂网络,社会网络分析.