

Information Trustworthiness

AAAI 2013 Tutorial

Jeffrey Pasternack[‡] Dan Roth[†] V.G.Vinod Vydiswaran[†]

[†] Department of Computer Science, University of Illinois, Urbana, IL 61801

[‡] Facebook Inc., Menlo Park, CA 94025

Tutorial Syllabus

Date: Monday, July 15th, 2013

Time: 9:00am - 1:00pm

Decision makers and citizens are heavily influenced these days by information they obtain from online resources – from news portals, online encyclopedias, blogs and forums, product websites and reviews, etc. However, the lack of control over what gets published online can often lead to dissemination of unreliable and misleading information. In such a scenario, how can one verify if a claim is true or determine which sources are trustworthy?

In this tutorial, the instructors systematically consider various approaches that address this problem. This tutorial is aimed at AI researchers who are interested in future research on trustworthiness of information. The goal of the tutorial is to present the audience with an exhaustive survey of relevant research on this topic and to outline future directions of research that may interest the AI community. No specific background knowledge is assumed of the audience.

Tutorial outline

The intended duration of the tutorial is four hours. The detailed outline and syllabus of the tutorial is given below.

1. Introduction & Motivation [15 min]

- Real-life examples showing need to verify claims and study information trustworthiness across different domains

2. Trustworthiness approaches based solely on source features [45 min]

- Graph based approaches: PageRank [3], topic-sensitive PageRank [7], propagation of trust and distrust [6]
- Medical source reliability using feature-based classification approaches [14]
- Text-centric approaches to check claims, using community knowledge [17]
- Using reputation of sources [2, 1]

3 Fact-finders: Trustworthiness approaches based on sources and claims [30 min]

- Voting, “Sums”
- TruthFinder [20]
- “Investment” and variants [10].
- Probabilistic models [21, 22]

4. Extensions to basic models, adding structure [60 min]

- Incorporating prior knowledge into models, Generalized fact-finders [12, 11]
- Incorporating context: Evidence-based Trust Framework [16]
- Handling source dependence [5, 4]
- Latent credibility analysis [13]

5. Presenting trustworthy information [30 min]

- External factors influence human perception of trust [15, 8].
- Presenting credible information on controversial issues [18, 19, 9]

6. Conclusion and Future Research Directions [30 min]

Tutorial Instructors

Jeff Pasternack is a research scientist at Facebook. His dissertation work with advisor Dan Roth focused on computing the trustworthiness of information via both non-probabilistic models such as *fact-finders* which iteratively evaluate the believability of claims given the trustworthiness of their sources (and vice versa) and, more recently, generative *Latent Credibility Analysis* models that, by describing how a source “decides” to assert a claim, provide a principled and highly effective method of assessing trust; making efficient use of prior knowledge has been of substantial interest in both approaches. Earlier work includes metrics for measuring the collective trustworthiness of sets of claims (such as news articles), information extraction, transliteration, and constrained learning.

Dan Roth is a Professor in the Department of Computer Science at the University of Illinois at Urbana-Champaign and the Beckman Institute of Advanced Science and Technology, and a University of Illinois Scholar. He is a fellow of AAAI and the ACM. Roth has published broadly in machine learning, natural language processing, knowledge representation and reasoning and received several paper, teaching and research awards. He has developed several machine learning based natural language processing systems that are widely used in the computational linguistics community and in industry and has presented invited talks and tutorials in several major conferences.

V.G.Vinod Vydiswaran is a post-doctoral research associate at the Information Trust Institute at the University of Illinois at Urbana-Champaign. His dissertation research focused on modeling and predicting trustworthiness on online textual information, and was advised by Prof.Dan Roth and Prof.ChengXiang Zhai. As part of his research, he addressed various challenges of computing trustworthiness of online textual information, viz. algorithmic, data-driven, and user-centric. His earlier work has included developing a textual entailment system and applying entailment techniques to relation extraction and information retrieval. His research interests include text informatics, natural language processing, machine learning, and information extraction.

References

- [1] B Thomas Adler, Krishnendu Chatterjee, Luca De Alfaro, Marco Faella, Ian Pye, and Vishwanath Raman. Assigning Trust to Wikipedia Content. *Computer Engineering*, 2008.
- [2] B. Thomas Adler and Luca de Alfaro. A Content-driven Reputation System for the Wikipedia. In *Proceedings of the 16th Intl. Conf. on World Wide Web (WWW)*, pages 261–270, 2007.
- [3] Sergey Brin and Lawrence Page. The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer Networks*, 30(1–7):107–117, 1998.
- [4] Xin Luna Dong, Laure Berti-Equille, Yifan Hu, and Divesh Srivastava. Global Detection of Complex Copying Relationships Between Sources. *Proc. of VLDB Endowment (PVLDB)*, 3(1):1358–1369, 2010.
- [5] Xin Luna Dong, Laure Berti-Equille, and Divesh Srivastava. Truth Discovery and Copying Detection in a Dynamic World. *Proc. of VLDB Endowment (PVLDB)*, 2(1):562–573, 2009.
- [6] R. Guha, Ravi Kumar, Prabhakar Raghavan, and Andrew Tomkins. Propagation of trust and distrust. In *Proceedings of the 13th International Conference on World Wide Web (WWW)*, pages 403–412, 2004.
- [7] Taher H. Haveliwala. Topic-sensitive PageRank. In *Proc. of 11th Intl. Conf. on World Wide Web (WWW)*, pages 517–526, 2002.
- [8] Robert R Hoffman, Machine Cognition, John D Lee, and Wisconsin Madison. The Dynamics of Trust in Cyberdomains. *IEEE Intelligent Systems*, pages 5–11, 2009.
- [9] Jeff Pasternack and Dan Roth. Comprehensive trust metrics for information networks. In *Proc. of the Army Science Conference (ASC)*, Orlando, Florida, December 2010.
- [10] Jeff Pasternack and Dan Roth. Knowing what to believe (when you already know something). In *Proc. of Intl. Conf. on Computational Linguistics (COLING)*, pages 877–885, 2010.
- [11] Jeff Pasternack and Dan Roth. Generalized fact-finding. In *Proc. of the International World Wide Web Conference (WWW)*, 2011.
- [12] Jeff Pasternack and Dan Roth. Making better informed trust decisions with generalized fact-finding. In *Proc. of the International Joint Conference on Artificial Intelligence (IJCAI)*, pages 2324–2329, 2011.
- [13] Jeff Pasternack and Dan Roth. Latent credibility analysis. In *Proc. of the International World Wide Web Conference (WWW)*, 2013.
- [14] Parikshit Sondhi, V.G.Vinod Vydiswaran, and ChengXiang Zhai. Reliability Prediction of Webpages in the Medical Domain. In *Proceedings of the 34th European Conference on Information Retrieval (ECIR)*, 2012.
- [15] S Tseng and B J Fogg. Credibility and computing technology. *Communications of the ACM*, 42(5):39–44, 1999.
- [16] V.G.Vinod Vydiswaran, ChengXiang Zhai, and Dan Roth. Content-driven Trust Propagation Framework. In *Proceedings of the 17th SIGKDD Conference on Knowledge Discovery and Data Mining (KDD)*, pages 974–982, 2011.
- [17] V.G.Vinod Vydiswaran, ChengXiang Zhai, and Dan Roth. Gauging the Internet Doctor: Ranking Medical Claims based on Community Knowledge. In *Proceedings of the KDD Workshop on Data Mining for Medicine and Healthcare*, pages 42–51, 2011.
- [18] V.G.Vinod Vydiswaran, ChengXiang Zhai, Dan Roth, and Peter Pirolli. BiasTrust: Teaching biased users about controversial topics. In *Proceedings of the 21st ACM International Conference on Information and Knowledge Management (CIKM)*, pages 1205–1209, 2012.

- [19] V.G.Vinod Vydiswaran, ChengXiang Zhai, Dan Roth, and Peter Pirolli. Unbiased Learning of Controversial Topics. In *Proceedings of the 75th Annual Meeting of the American Society for Information Science and Technology (ASIS&T)*, pages 291.1–291.4, 2012.
- [20] Xiaoxin Yin, Jiawei Han, and Philip S. Yu. Truth Discovery with Multiple Conflicting Information Providers on the Web. *IEEE Transactions on Knowledge and Data Engineering*, 20(6):796–808, 2008.
- [21] Bo Zhao and Jiawei Han. A probabilistic model for estimating real-valued truth from conflicting sources. In *Proc. of 10th Intl. Workshop on Quality in Databases, in conjunction with VLDB 2012 (QDB'12)*, pages 1–7, 2012.
- [22] Bo Zhao, Benjamin I. P. Rubinstein, Jim Gemmell, and Jiawei Han. A bayesian approach to discovering truth from conflicting sources for data integration. *Proc. of VLDB Endowment (PVLDB)*, 5(6):550–561, 2012.