



博士论文答辩报告

基于知识迁移的社交媒体事件检测方法研究

博士生：黄威靖

导师：王腾蛟教授

北京大学信息科学技术学院数据库实验室

2018-06-05



1. 引言

2. 基于知识迁移的社交媒体事件检测通用框架 K2Social

3. 基于知识库结构的主题抽取与迁移方法

4. 用户兴趣建模与知识迁移方法

5. 基于知识迁移的领域事件检测通用方法

6. 总结与展望

Motivation

Many Web applications need the **accurate event detection** technique on microblog stream, including:

1. disaster response [Sakaki,WWW 2010]
2. breaking news report¹

But detecting events on twitter stream accurately is still challenging.

¹<http://www.theverge.com/2016/12/1/13804542/reuters-algorithm-breaking-news-twitter>

1. 引言

2. 基于知识迁移的社交媒体事件检测通用框架 K2Social

3. 基于知识库结构的主题抽取与迁移方法

4. 用户兴趣建模与知识迁移方法

5. 基于知识迁移的领域事件检测通用方法

6. 总结与展望

Motivation

Many Web applications need the **accurate event detection** technique on microblog stream, including:

1. disaster response [Sakaki, WWW 2010]
2. breaking news report²

But detecting events on twitter stream accurately is still challenging.

²<http://www.theverge.com/2016/12/1/13804542/reuters-algorithm-breaking-news-twitter>

1. 引言

2. 基于知识迁移的社交媒体事件检测通用框架 K2Social

3. 基于知识库结构的主题抽取与迁移方法

4. 用户兴趣建模与知识迁移方法

5. 基于知识迁移的领域事件检测通用方法

6. 总结与展望

Motivation

Many Web applications need the **accurate event detection** technique on microblog stream, including:

1. disaster response [Sakaki, WWW 2010]
2. breaking news report³

But detecting events on twitter stream accurately is still challenging.

³<http://www.theverge.com/2016/12/1/13804542/reuters-algorithm-breaking-news-twitter>

1. 引言

2. 基于知识迁移的社交媒体事件检测通用框架 K2Social

3. 基于知识库结构的主题抽取与迁移方法

4. 用户兴趣建模与知识迁移方法

5. 基于知识迁移的领域事件检测通用方法

6. 总结与展望

Motivation

Many Web applications need the **accurate event detection** technique on microblog stream, including:

1. disaster response [Sakaki,WWW 2010]
2. breaking news report⁴

But detecting events on twitter stream accurately is still challenging.

⁴<http://www.theverge.com/2016/12/1/13804542/reuters-algorithm-breaking-news-twitter>

1. 引言

2. 基于知识迁移的社交媒体事件检测通用框架 K2Social

3. 基于知识库结构的主题抽取与迁移方法

4. 用户兴趣建模与知识迁移方法

5. 基于知识迁移的领域事件检测通用方法

6. 总结与展望

Motivation

Many Web applications need the **accurate event detection** technique on microblog stream, including:

1. disaster response [Sakaki, WWW 2010]
2. breaking news report⁵

But detecting events on twitter stream accurately is still challenging.

⁵<http://www.theverge.com/2016/12/1/13804542/reuters-algorithm-breaking-news-twitter>

1. 引言

2. 基于知识迁移的社交媒体事件检测通用框架 K2Social

3. 基于知识库结构的主题抽取与迁移方法

4. 用户兴趣建模与知识迁移方法

5. 基于知识迁移的领域事件检测通用方法

6. 总结与展望

Motivation

Many Web applications need the **accurate event detection** technique on microblog stream, including:

1. disaster response [Sakaki,WWW 2010]
2. breaking news report⁶

But detecting events on twitter stream accurately is still challenging.

⁶<http://www.theverge.com/2016/12/1/13804542/reuters-algorithm-breaking-news-twitter>

论文发表情况 (1/3)

1. **Weijing Huang**, Tengjiao Wang, Wei Chen, Siyuan Jiang, Kam-Fai Wong, PhraseCTM: Correlated Topic Modeling on Phrases within Markov Random Fields, ACL 2018, accepted, to appear. (CCF A 类)
2. **Weijing Huang**, Tengjiao Wang, Wei Chen, Yazhou Wang, Category-Level Transfer Learning from Knowledge Base to Microblog Stream for Accurate Event Detection, DASFAA 2017. (EI index number: 20174404323179, CCF B 类)
3. **Weijing Huang**, Wei Chen, Tengjiao Wang, Shibo Tao. TaxoPhrase: Exploring Knowledge Base via Joint Learning of Taxonomy and Topical Phrases, the 2nd Open Knowledge Base and Question Answering Workshop at SIGIR 2017.
4. **Weijing Huang**, Wei Chen, Tengjiao Wang, Shibo Tao. Efficient Topic Modeling on Phrases via Sparsity, the 29th IEEE International Conference on Tools with Artificial Intelligence (ICTAI) 2017, accepted, to appear.
5. **Weijing Huang**, Wei Chen, Lamei Zhang, Tengjiao Wang, An Efficient Online Event Detection Method for Microblogs via User Modeling, the 18th Asia Pacific Web Conference (APWeb) 2016. (EI index number: 20164102880250)

论文发表情况 (2/3)

6. Ruhui Wang, **Weijing Huang**, Wei Chen, Tengjiao Wang, Kai Lei. ASEM: Mining Aspects and Sentiment of Events from Microblog, the 24th ACM International Conference on Information and Knowledge Management (CIKM) 2015.
7. Yue Wang, **Weijing Huang**, Lang Zong, Tengjiao Wang, Dongqing Yang: Influence maximization with limit cost in social network. SCIENCE CHINA Information Sciences 56(7): 1-14 (2013)
8. Yue Wang, **Weijing Huang**, Wei Chen, Tengjiao Wang, Dongqing Yang. Informed Prediction with Incremental Core-based Friend Cycle Discovering, the 12th International Conference on Web-Age Information Management (WAIM) 2011.
9. 张腊梅, **黄威靖**, 陈薇, 王腾蛟, 雷凯. EMTM: 微博中与主题相关的专家挖掘方法. 计算机研究与发展. 2016 年 (第 53 卷) (萨师煊优秀学生论文奖).
10. 吴良, **黄威靖**, 陈薇, 王腾蛟, 雷凯, 刘月琴. ACT-LDA: 集成话题、社区和影响力分析的概率模型. 计算机科学与探索. 2013 年第 7 卷第 8 期 (718-728).

论文发表情况 (3/3)

11. Shibo Tao, Xiaorong Wang, **Weijing Huang**, Wei Chen, Tengjiao Wang, Kai Lei, From Citation Network to Study Map: A Novel Model to Reorganize Academic Literatures, Big Scholar workshop at WWW 2017.
12. Wei Chen, Lang Zong, **Weijing Huang**, Gaoyan Ou, Yue Wang, Dongqing Yang, An Empirical Study of Massively Parallel Bayesian Networks Learning for Sentiment Extraction from Unstructured Text, the 13th Asia-Pacific Web Conference (APWeb) 2011.
13. Xilian Li, Wei Chen, Tengjiao Wang, **Weijing Huang**, Target-specific Convolutional Bi-directional LSTM Neural Network for Political Ideology Analysis, the Asia Pacific Web and Web-Age Information Management Joint Conference on Web and Big Data (APWeb-WAIM) 2017.
14. Xiao Zhang, Xiaorong Wang, Wei Chen, Jie Tao, **Weijing Huang**, Tengjiao Wang, A Taxi Gap Prediction Method via Double Ensemble Gradient Boosting Decision Tree, the 2nd IEEE International Conference on Intelligent Data and Security 2017.

参与科研项目情况

硕博连读期间参与的项目：

1. 非结构化数据管理系统北大部分——国家“核高基”重大专项 (2010ZX01042-002-002-02)
2. 海量 Web 数据结构化内容提取与集成及大型示范应用——国家 863 计划课题 (2012AA011002)
3. 微博用户社区及主题时序方法研究——中国信息安全测评中心合作项目
4. 大数据驱动的航天航空装备创新研发与应用示范——十三五国家重点研发计划课题

硕博连读期间参与的专利：

1. 基于交互式文档聚类的信息检索方法及系统, **黄威靖**, 于倩, 陈薇, 王腾蛟, 杨冬青, CN103514183A.
2. Web 社会网络核心用户信息交互演化分析方法, 王悦, **黄威靖**, 陈薇, 王腾蛟, 杨冬青, CN102637182A.

感谢导师王腾蛟教授对我的悉心指导！

感谢所有关心、支持我的人！

谢谢各位老师！