cases	doc_1		doc_2		decision	id
	authors	• Hao Peng • Jianxin Li • Yangqiu Song • Renyu Yang • R. Ranjan • Philip S. Yu • Lifang He	**authors**	• Hao Peng • Jianxin Li • Yangqiu Song • Renyu Yang • Rajiv Ranjan • Philip S. Yu • Lifang He	DUPLICATES	229
	title	Streaming Social Event Detection and Evolution Discovery in Heterogeneous Information Networks	**title**	Streaming Social Event Detection and Evolution Discovery in Heterogeneous Information Networks		
	publication_date	2021-04-02 00:00:00	**publication_date**	2021-04-02 02:13:10+00:00		
	source	SupportedSources.SEMANTIC_SCHOLAR	**source**	SupportedSources.ARXIV		
	journal	ACM Transactions on Knowledge Discovery from Data (TKDD)	**journal**	None		
	volume	15	**volume**			
	doi	10.1145/3447585	**doi**			
	urls	• https://www.semanticscholar.org/paper/e767fb1a7227660a8ea8138dfdf2b73cb3f1bf24	**urls**	• http://arxiv.org/pdf/2104.00853v1 • http://arxiv.org/abs/2104.00853v1 • http://arxiv.org/pdf/2104.00853v1		
	id	id2285382340562254883	**id**	id-854271328413357332		
	abstract	Events are happening in real world and real time, which can be planned and organized for occasions, such as social gatherings, festival celebrations, influential meetings, or sports activities. Social media platforms generate a lot of real-time text information regarding public events with different topics. However, mining social events is challenging because events typically exhibit heterogeneous texture and metadata are often ambiguous. In this article, we first design a novel event-based meta-schema to characterize the semantic relatedness of social events and then build an event-based heterogeneous information network (HIN) integrating information from external knowledge base. Second, we propose a novel Pairwise Popularity Graph Convolutional Network, named as PP-GCN, based on weighted meta-path instance similarity and textual semantic representation as inputs, to perform fine-grained social event categorization and learn the optimal weights of meta-paths in different tasks. Third, we propose a streaming social event detection and evolution discovery framework for HINs based on meta-path similarity search, historical information about meta-paths, and heterogeneous DBSCAN clustering method. Comprehensive experiments on real-world streaming social text data are conducted to compare various social event detection and evolution discovery algorithms. Experimental results demonstrate that our proposed framework outperforms other alternative social event detection and evolution discovery techniques.	**abstract**	Events are happening in real-world and real-time, which can be planned and organized for occasions, such as social gatherings, festival celebrations, influential meetings or sports activities. Social media platforms generate a lot of real-time text information regarding public events with different topics. However, mining social events is challenging because events typically exhibit heterogeneous texture and metadata are often ambiguous. In this paper, we first design a novel event-based meta-schema to characterize the semantic relatedness of social events and then build an event-based heterogeneous information network (HIN) integrating information from external knowledge base. Second, we propose a novel Pairwise Popularity Graph Convolutional Network, named as PP-GCN, based on weighted meta-path instance similarity and textual semantic representation as inputs, to perform fine-grained social event categorization and learn the optimal weights of meta-paths in different tasks. Third, we propose a streaming social event detection and evolution discovery framework for HINs based on meta-path similarity search, historical information about meta-paths, and heterogeneous DBSCAN clustering method. Comprehensive experiments on real-world streaming social text data are conducted to compare various social event detection and evolution discovery algorithms. Experimental results demonstrate that our proposed framework outperforms other alternative social event detection and evolution discovery techniques.		
	versions		**versions**			