Category-Level Transfer Learning from Knowledge Base to Microblog Stream for Accurate Event Detection

Weijing Huang, Tengjiao Wang, Wei Chen, Yazhou Wang

School of Electronics Engineering and Computer Science, Peking University

@DASFAA 2017, Suzhou, China

Motivation

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

1. public opinion analysis [Chen, SIGIR 2013]
2. public security [Li, ICDE 2012], [Imran, WWW 2014]
3. disaster response [Sakaki,WWW2010]
4. breaking news report[1]

But detecting events on twitter stream accurately is still challenging.

[1] `http://www.theverge.com/2016/12/1/13804542/reuters-algorithm-breaking-news-twitter`

Challenges (1/2)

According to [Huang, WWW 2016], the challenges include,

1. fast changing
2. high noise
3. short length

And, we found another key factor,

1. Small events with fewer tweets ➔ Hard to trade off between precision and recall.

Challenges (2/2)

Exploratory study on the *Edinburgh twitter corpus*: 11/29 events contain less than 50 tweets.

Table: Statistics of labeled events.

Event	Date	Event Size
S&P downgrade US credit rating	05/08/2011	656
Atlantis shuttle lands	21/07/2011	595
US increases debt ceiling	25/07/2011	485
Plane with Russian hocky team Lokomotiv crashes	07/09/2011	286
Amy Winehouse dies	23/07/2011	283
Gunman opens fire in youth camp in Norway	23/07/2011	260
Earthquake in Virginia	24/08/2011	246
First victim of London riots dies	09/08/2011	174
Explosion in French nuclear plant in Marcoule	12/09/2011	135
Google announces plans to bury Motorola Mobility	15/08/2011	127
NASA announces there might be water on Mars	04/08/2011	124
Car bomb explodes in Oslo, Norway	22/07/2011	114
...	...	...
Indian and Bangladesh sign a border pact	06/09/2011	25
Flight 4896 crash	13/07/2011	21
First aritficial organ transplant	12/07/2011	18
three men die in riots in england	10/08/2011	16
rebels capture interational tripoli airport	21/08/2011	13

11 Small events with fewer tweets

How about existing methods? (1/2)

Event detection methods *without extra information*, such as

1. clustering articles
   - LSH[Petrovic, NAACL 2010]
   - need to set threshold to determine whether new article represents a new event.
2. analyzing word frequencies
   - EDCoW[Weng, ICWSM 2011]
   - treat the word as the basic unit in analysis, without regarding polysemy words (words have different meanings, e.g. "apple")
3. finding bursty topics via topic modeling
   - TimeUserLDA[Diao, ACL 2012], BurstyBTM[Yan, AAAI 2015]
   - detects the "large" events but may ignore the "small" ones.

How about existing methods? (2/2)

Event detection methods *by leveraging extra information*

1. typical one: Twevent[Li, CIKM 2012]
   - divides the tweet into segments according to the Microsoft Web N-Gram service and Wikpedia
   - detects the bursty segments and cluster these segments into candidate events
   - still have to trade off between precision and recall

An Example

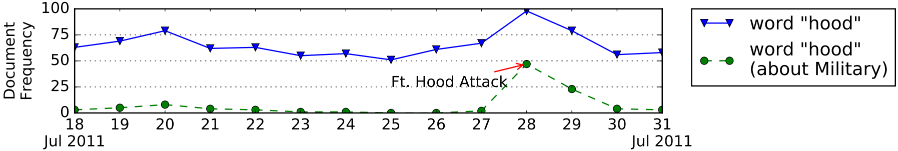

Figure: The comparison of the time series between the raw word *hood* and the *Military* related word *hood*, computed on the *Edinburgh twitter corpus*. The rise of document frequency on July 28th, 2011 is corresponding to the event mentioned in `https://en.wikipedia.org/wiki/Fort_Hood#2011_attack_plot`.

The insights on the example

Overview of our method

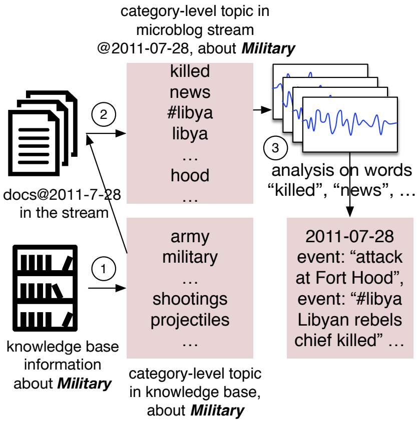

Figure: TRANSDETECTOR's processing flow, taking *Military* related events in microblogs as an example.

TransDetector: Phrase 1

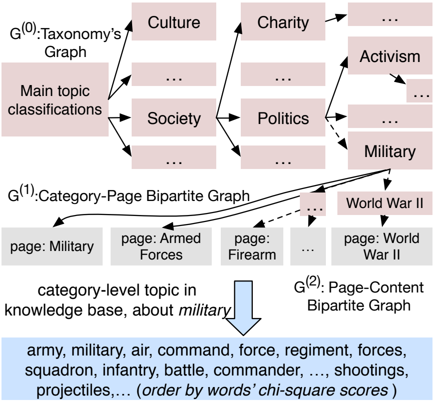

Figure: Extracting Category-Level Topics in Knowledge Base via its three fold hierarchical structure, taking *Military* as an example.

TransDetector: Phrase 2

TRANSDETECTOR: Phrase 3

Experimental Results

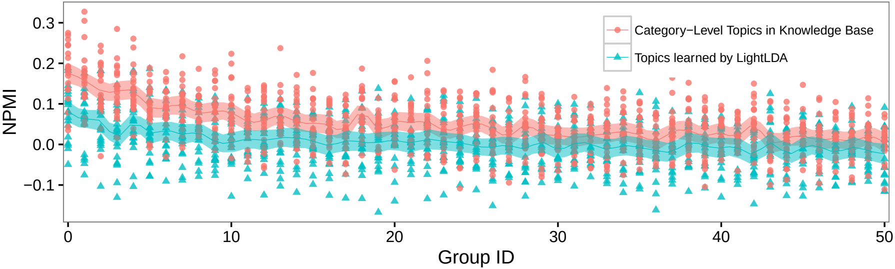

Figure: More topics are compared at the NPMI metrics between our method and LightLDA

some

Table: The comparison on the topic coherence(NPMI) between our method and LightLDA, taking *Aviation* as an example. (NPMI is computed on a group of ten words. $\sim$ stands for the top five words.)

Category-Level Topics extracted from Wikipedia by TRANSDETECTOR				Topics Learned from Wikipedia by LightLDA			
GID	#words*	words	NPMI	GID	#words*	words	NPMI
-	1-5	aircraft air airport flight airline	-	-	1-5	engine aircraft car air power	-
0	1-5, 6-10	$\sim$, airlines aviation flying pilot squadron	0.113	0	1-5, 6-10	$\sim$, design flight model production speed	0.112
1	1-5, 11-15	$\sim$, flights pilots raf airways fighter	0.155	1	1-5, 11-15	$\sim$, system vehicle cars engines mm	0.062
2	1-5, 16-20	$\sim$, boeing runway force crashed flew	0.092	2	1-5, 16-20	$\sim$, fuel vehicles designed models type	0.072
3	1-5, 21-25	$\sim$, airfield landing passengers plane aerial	0.179	3	1-5, 21-25	$\sim$, version front produced rear electric	0.035
4	1-5, 26-30	$\sim$, bomber radar wing bombers crash	0.137	4	1-5, 26-30	$\sim$, space control motor standard development	0.085
5	1-5, 31-35	$\sim$, airbus airports operations jet helicopter	0.189	5	1-5, 31-35	$\sim$, film range light using available	-0.002
6	1-5, 36-40	$\sim$, squadrons base flown havilland crew	0.088	6	1-5, 36-40	$\sim$, wing powered wheel weight launch	0.087
7	1-5, 41-45	$\sim$, combat luftwaffe aerodrome carrier fokker	0.159	7	1-5, 41-45	$\sim$, developed low test ford cylinder	0.007
8	1-5, 46-50	$\sim$, planes fly engine takeoff fleet	0.186	8	1-5, 46-50	$\sim$, equipment side pilot hp aviation	0.091
9	1-5, 51-55	$\sim$, fuselage helicopters aviator naval aero	0.157	9	1-5, 51-55	$\sim$, systems us sold body drive	-0.051
10	1-5, 56-60	$\sim$, glider command training balloon faa	0.166	10	1-5, 56-60	$\sim$, gear introduced class safety seat	0.069
...	...	...	...	...	...	...	...
18	1-5, 96-100	$\sim$, scheduled carriers military curtiss biplane	0.131	18	1-5, 96-100	$\sim$, transmission special replaced limited different	0.059
19	1-5, 101-105	$\sim$, accident engines iaf albatross rcaf	0.068	19	1-5, 101-105	$\sim$, features machine nuclear even unit	0.011

Experiment Settings

Experimental Results

Table: Category-Level Topics extracted from knowledge base and the corresponding topics on microblog stream learned from CTrans-LDA. The words in **bold** font are newly learned on the microblog stream by the transfer learning.

Aviation		*Health*		*Middle East*		*Military*		*Mobile Phones*	
Knowledge Base	Microblog Stream	Knowledge Base	Microblog Stream	Knowledge Base	Microblog Stream	Knowledge Base	Microblog Stream	Knowledge Base	Microblog Stream
aircraft	air	health	weight	al	**#syria**	army	killed	android	iphone
air	plane	patients	loss	israel	**#bahrain**	military	news	mobile	apple
airport	flight	medical	diet	iran	people	air	**#libya**	nokia	android
flight	time	disease	health	arab	israel	command	libya	ios	app
airline	airlines	treatment	cancer	israeli	police	force	rebels	phone	ipad
airlines	news	hospital	lose	egypt	**#libya**	regiment	people	samsung	samsung
aviation	boat	patient	fat	egyptian	#egypt	forces	police	game	mobile
flying	airport	clinical	tips	ibn	news	squadron	war	app	blackberry
pilot	force	symptoms	treatment	jerusalem	**#israel**	infantry	libyan	iphone	tablet
squadron	fly	cancer	body	syria	world	battle	attack	htc	apps

Experimental Results

Table: Overall Performance on Event Detection

Method	Number of Events to be Evaluated	Recall@ Benchmark1	Precision@ Benchmark2	Recall@ Benchmark2	F@ Benchmark2	DERate[a] (Duplicate Event Rate)@ Benchmark2
LSH	500	0.704	0.788	0.651	0.713	0.348
TimeUserLDA	100	0.370	0.790	0.177	0.289	0.114
Twevent	375	0.741	0.808	0.658	0.725	0.142
EDCoW	349	0.556	0.748	0.511	0.607	0.226
BurstyBTM	200	0.667	0.825	0.384	0.497	**0.079**
TransDetector	457	**0.889**	**0.912**	**0.876**	**0.894**	0.170

[a] DERate = (the number of duplicate events) / (the total number of detected realistic events)

Experimental Results

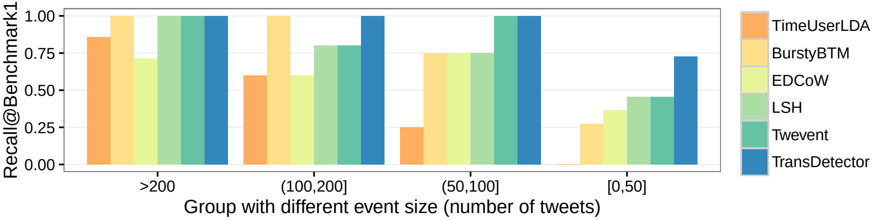

Figure: The relation between the recall and the event size

Experimental Results

Table: Events about *military* detected by systems between 2011-07-22 and 2011-07-28

Date	Event key words	Representative event tweet	Number of event tweet	Methods[a]					
				L	TU	TW	E	B	TD
7/22/11	Norway, Oslo, attacks, bombing	Terror Attacks Devastate Norway: A bomb ripped through government offices in Oslo and a gunman... http://dlvr.it/cLbk8	557	✓	✓	✓	✓	✓	✓
7/23/11	Gunman, rink	Gunman Kills Self, 5 Others at Texas Roller Rink http://dlvr.it/cLcTH	43	-	-	✓	✓	-	✓
7/26/11	Kandahar, mayor, suicide, attack	TELEGRAPH]: Kandahar mayor killed by Afghan suicide bomber: The mayor of Kandahar, the biggest city in south _	47	✓	-	✓	✓	-	✓
7/28/11	Ft., Hood, attack	Possible Ft. Hood Attack Thwarted http://t.co/BSJ33hk	52	-	-	-	-	-	✓
7/28/11	Libyan, rebel, gunned	Libyan rebel chief gunned down in Benghazi http://sns.mx/prfvy1	44	-	-	-	-	-	✓

[a] L=LSH, TU=TimeUserLDA, TW=Twevent, E=EDCoW, B=BurstyBTM, TD=TransDetector.

Thanks!

Q&A

[1] This slide and more data are available at http://q-r.to/bajx8I