

K-TEAMS / Computer-Supported Collaborative Knowledge Construction Research Laboratory

User-Level Opinion Propagation Analysis in Discussion Forum Threads

Dumitru-Clementin Cercel¹ & Ştefan Trăuşan-Matu^{1, 2}

¹University POLITEHNICA of Bucharest ²Romanian Academy Research Institute for Artificial Intelligence

16th International Conference on Artificial Intelligence: Methodology, Systems, Applications (AIMSA) Varna, September 11-13, 2014

Why Opinion Propagation in Online Forum Threads?

- Online forums are very popular among Internet users
- A mixture of positive and negative opinions
- There are multiple applications in the real world

Opinion Concept

- The definition given in the Oxford Dictionary
 - "A view or judgement formed about something, not necessarily based on fact or knowledge"
 - "A statement of advice by an expert on a professional matter"
- Computational perspective (Ding et al., 2008)
 - Target entity
 - Holder
 - Sentiment
 - Timestamp

Related Work *vs.* Our Original Contribution

Related Work: Drawbacks

- The Voter Model, The Sznajd Model, The Deffuant Model, The Hegselmann-Krause Model
- (-) The propagation of opinions in the networks (e.g. the Watts-Strogatz network model or the Barabási-Albert network model)
- (-) Opinions are numerical values, randomly assigned to individuals
- (-) The interactions over time between neighboring individuals are also randomly established

Our Original Contribution: Opinion Propagation in Forum Threads

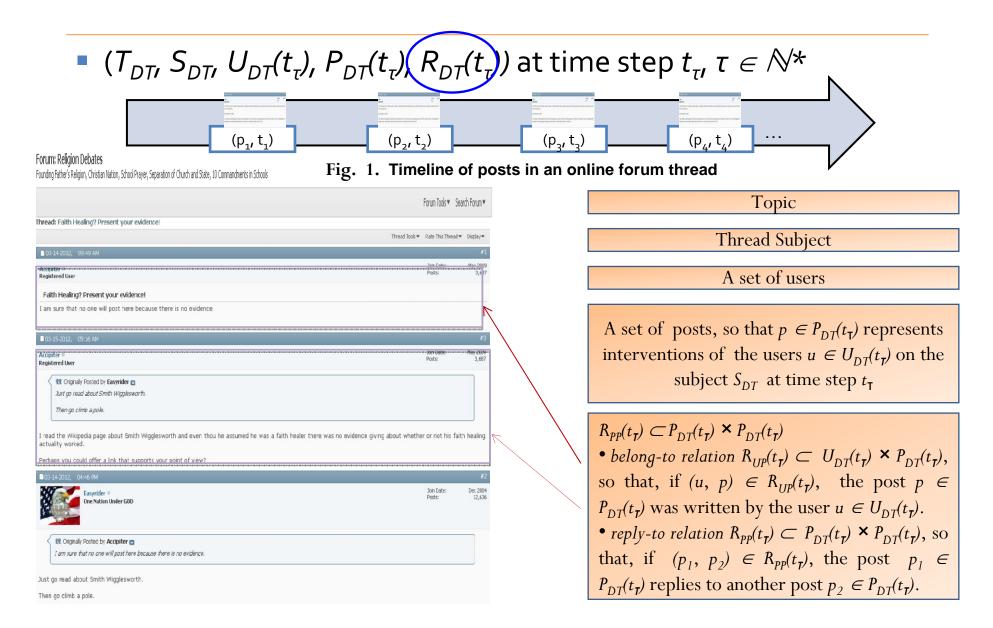
- Our method consists in determining whether, until a given time step, the users in discussion are in agreement or continue to have different or even contrary opinions.
 - (+) We take into account the opinions written by users
 - (+) We use the real-world's online forum threads

• $(T_{DT}, S_{DT}, U_{DT}(t_{\tau}), P_{DT}(t_{\tau}))$	R_{DT}	r(t ₁)) at time step $t_{\tau}, \tau \in \mathbb{N}^*$
(p ₁ , t ₁)	t ₂)		$(p_3, t_3) $
Forum: Religion Debates Founding Father's Religion, Ciristian Nation, School Prayer, Separation of Church and State, 10 Commandments in Schools Fig. 1. Time	eline	of po	osts in an online forum thread
	Forum Tools▼ Se	earch Forum 🔻	Торіс
Thread: Faith Healing? Present your evidence!	s 💌 Rate This Thread	• Nierlau•	
■ 09-14-2012, 09-149 AN		#1	
Accipiter o Registered User	Join Date: Posts:	May 2009 3,687	
Faith Healing? Present your evidence!			
I am sure that no one will post here because there is no evidence.			
■ 03-15-2012, C5:16 AM		#3	
Accipiter o Registered User	Join Date: Posts:	May 2009 3,687	
R Originaly Possed by Easymider Xet ao read about Smith Wizzlesworth.			
Then go climb a pole.			
I read the Wikipedia page about Smith Wigglesworth and even thou he assumed he was a faith healer there was no evidence giving about whe actuality worked.	ther or not his fai	ith healing	
Perhaps you could offer a link that supports your point of view?			
■ 03:14:2012, 04:46 PM Easyrider ○ Dne Nation Under GOD	Join Date: Posts:	#2 Det 2004 12,636	
R Originally Posted by Accipiter I am sure that no one will post here because there is no evidence.			
Just go read about Smith Wigglesworth.			
Then go climb a pole.			

• $(T_{DT}, S_{DT}, U_{DT}, (t_{\tau}), F$	$P_{DT}(t_{\tau}), R_{DT}(t_{\tau}))$ at	time step $t_{\tau}, \tau \in \mathbb{N}^*$
Forum: Religion Debates	(p_2, t_2) ig. 1. Timeline of posts in	$p_3, t_3)$ (p_4, t_4)
Founding Father's Religion, Christian Nation, School Prayer, Separation of Church and State, 10 Commandments in Schools	Forum Tools * Search Forum *	Topic
Thread: Faith Healing? Present your evidence!		
	Thread Tools * Date This Thread * Display *	Thread Subject
09-14-2012, 09-19 AM Accipiter •	#1 .3on Date: May 2009 Posts: 3,667	
Registered User Faith Healing? Present your evidence!	i vanus againi	
Tam sure that no one will post here because there is no evidence.		
■ 03-15-2012, 05:16 4M	#3	
Accipitar 0 Registred User	Join Date: May 2009 Posts: 3,687	
R Originally Posted by Easynider Xust go read about Smith Wigglesworth.		
Then go climb a pole.		
I read the Wikipedia page about Smith Wigglesworth and even thou he assumed he was a faith healer there was actuality worked.	no evidence giving about whether or not his faith healing	
Perhaps you could offer a link that supports your point of view?		
■03.14-2012, 04×6 PM	#2 Join Dete: Dec 2004	
Easyrider © One Nation Under GOD	Posts: 12,636	
R Criginally Posted by Accipiter I am sure that no one will post here because there is no evidence.		
Just go read about Smith Wigglesworth.		
Then go climb a pole.		

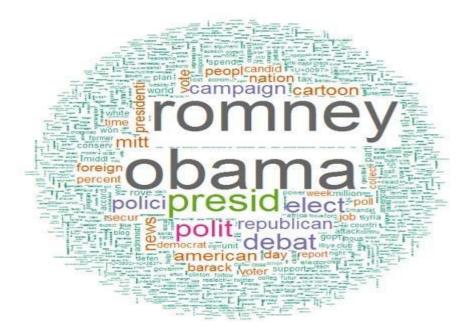
• $(T_{DT}, S_{DT}, U_{DT}(t_{T}), P_{DT})$	$(t_{\tau}), R_{DT}(t_{\tau}))$ at	time step $t_{\tau}, \tau \in \mathbb{N}^*$
Forum: Religion Debates Founding Father's Religion, Christian Nation, School Prayer, Separation of Church and State, 10 Commandments in Schools Fig.		(p_3, t_3) (p_4, t_4) an online forum thread
	Forum Tools 🔻 Search Forum 💌	Торіс
Thread: Faith Healing? Present your evidence		
	Thread Tools 💌 Rate This Thread 🐨 Display 🕶	Thread Subject
Accipiter • Accipiter •	#1 Jon Date: May 2009	
Registered User	Posts: 3,687	A set of users
Faith Healing? Present your evidence! I am sure that no one will post here because there is no evidence.		
rain sure mac no one will post nele occasse there is no evidence.		
■ 03-15-2012. (5:16 4M	#3	
Accipiter o Registered User	Join Date: May 2009 Posts: 3,687	
Image: State		
I read the Wikipedia page about Smith Wigglesworth and even thou he assumed he was a faith healer there was no evidence actuality worked.	giving about whether or not his faith healing	
Perhaps you could offer a link that supports your point of view?	- C4	
■ 03-14-2012, 04:46 PM Easyrider © One Nation Under G0D	3on Dete: Dec 2004 Posts: 12,436	
R Originally Posted by Accipiter a I am sure that no one will post here because there is no evidence,		
Just go read about Smith Wigglesworth. Then go climb a pole.		

 (T_{DT}, S_{DT}, U_{DT}(t_τ), P_{DT} 	(t ₇), R _{DT} ((t_{τ})) at tim	ne step $t_{\tau}, \tau \in \mathbb{N}^*$
Forum: Religion Debates Founding Father's Religion, Christian Nation, School Prayer, Separation of Church and State, 10 Commandments in Schools Figure Church and State, 10 Commandments in Schools	(p_2, t_2) 1. Timeline of	(p ₃ , t ₃) posts in an on	line forum thread
	Forum Tools 🔻 Search F	orum 🔻	Торіс
Thread: Faith Healing? Present your evidence! 03-14-2012, 09-19-AM	Thread Tools 💌 Rate This Thread 💌 Dis	play *	Thread Subject
Octaveller of Registered liser	Jain Date: Mi Posts:	v 2009 3,687	A set of users
Faith Healing? Present your evidence! I am sure that no one will post here because there is no evidence. 03-15-2012, 05:16 AM Arcipiter © Registered User It or ignally Posted by Easyrider © Just go read about Smith Wigglesworth.	Jon Date: Ma Posts:	13 2009 3,667	A set of posts, so that $p \in P_{DT}(t_T)$ represents interventions of the users $u \in U_{DT}(t_T)$ on the subject S_{DT} at time step t_T
Then go climb a pole. I read the Wikipedia page about Smith Wigglesworth and even thou he assumed he was a faith healer there was no evidence actuality worked. Perhaps you could offer a link that supports your point of yew?	giving about whether or not his faith he	aling L	
O3:14:2012, D4:46 FM Easynther © Done Nation Under GOD		#2 x 2004 12,436	
Rt Originally Posted by Accipiter I am sure that no one will post here because there is no evidence.			
Just go read about Smith Wigglesworth. Then go climb a pole.			



Concepts

- Noun Term Vocabulary
 - All distinct noun terms on which the users $U_{DT}(t_{\tau})$ expressed their opinions and that are semantically related to one of the noun terms by which the subject S_{DT} is written
 - $V^{d}_{DT}(t_{\tau}) = \{n_{1}, n_{2}, ..., n_{d}\}$ denotes a *d*-dimensional vocabulary of noun terms at time step $t_{\tau}, \tau \in \mathbb{N}^{*}$



Concepts

Discrete Opinion Space

- The opinions of each user $u \in U_{DT}(t_{\tau})$ on the noun terms from the vocabulary $V_{DT}^{d}(t_{\tau})$ can be represented by a vector in a *d*-dimensional discrete opinion space $OS_{DT}^{d} = \{-1, 0, +1\}^{d}$
- $o_{DT}^{d}(t_{\tau}) = [o_1 \ o_2 \ \dots \ o_d]^T$ denotes an opinion vector at time step $t_{\tau}, \tau \in \mathbb{N}^*$, in the *d*-dimensional discrete opinion space OS_{DT}^{d}
 - The opinion entries o_k can take one of the following sentiment scores: -1, 0, or +1

Observation!

- If the user $u \in U_{DT}(t_{\tau})$ does not express his opinion on the noun term $s_k \in V_{DT}^d(t_{\tau})$ until time step $t_{\tau}, \tau \in \mathbb{N}^*$, then we consider the value 0 for the entry $o_k \in o_{DT}^d(t_{\tau})$ (t_{τ})
- If until time step t_{τ} , $\tau \in \mathbb{N}^*$, the user $u \in U_{DT}(t_{\tau})$ gives more opinions on a noun term, then only his last opinion is taken into consideration

Concepts

- Term-User Opinion Matrix
 - We construct a $d \times n$ term-user matrix $A_{T-U}(t_{\tau}) = [A_1(t_{\tau}) A_2(t_{\tau}) \dots A_n(t_{\tau})]$ at time step $t_{\tau}, \tau \in \mathbb{N}^*$
 - *n* denotes the number of users in the set $U_{DT}(t_{\tau})$
 - Each column $A_u(t_{\tau}) = [a_{1,\mu}(t_{\tau}) \ a_{2,\mu}(t_{\tau}) \ \dots \ a_{d,\mu}(t_{\tau})]^T$ corresponds to a user $u \in U_{DT}(t_{\tau})$ and denotes the *d*-dimensional opinion vector of the user $u \in U_{DT}(t_{\tau})$ in the discrete opinion space OS^d_{DT}
- User-User Similarity Matrix
 - We construct a $n \times n$ user-user similarity matrix $B_{U-U}(t_{\tau})$ at time step t_{τ} , $\tau \in \mathbb{N}^*$
 - The entry of the row k^{th} and of the column h^{th} of the matrix $B_{U-U}(t_{\tau})$ is denoted by $b_{k,k}(t_{\tau})$ and represents the similarity between users k and $u \in U_{DT}(t_{\tau})$ from the perspective of the opinion vector expressed by these users

•
$$b_{k,h}(t_{\tau}) = sim(A_k(t_{\tau})) = [a_{1,k}(t_{\tau}) a_{2,k}(t_{\tau}) \dots a_{d,k}(t_{\tau})]^T, A_h(t_{\tau}) = [a_{1,h}(t_{\tau}) a_{2,h}(t_{\tau})]^T$$

$$\dots a_{d,h}(t_{\tau})]^{T}) = \frac{\sum_{i=1}^{d} a_{i,k}(t_{\tau}) a_{i,h}(t_{\tau})}{\sqrt{\sum_{i=1}^{d} \left(a_{i,k}(t_{\tau})\right)^{2}} \sqrt{\sum_{i=1}^{d} \left(a_{i,h}(t_{\tau})\right)^{2}}}$$

Problem Formalization

- The Problem of User-Level Opinion Propagation in Online Forum Threads:
 - Input: Given, at time step t_{τ} , $\tau \in \mathbb{N}^*$, a subset of users $U'_{DT}(t_{\tau}) \subset U_{DT}(t_{\tau})$ who have similar opinion vectors for any time step t_i , $t_i \ge t_{\tau}$, (i.e. $b_{u2u1} \le \varepsilon_1$, u_2 , $u_1 \in U'_{DT}(t_{\tau})$) and who initiated the opinion propagation process at time step t_{τ}
 - **Goal**: A user $u \in U_{DT}(t_j) \setminus U'_{DT}(t_{\tau})$ is considered to be influenced, at time step t_j , $t_j > t_{\tau}$ by the opinion propagation only if the following condition is met:

$$\pi \ b_{u3,u}(t_{\tau}) \leq b_{u1,u}(t_{j}) \ \bowtie \ \mathcal{E}_{2}$$

where $u_3 \in U_{DT}(t_i) \setminus U'_{DT}(t_r)$, $u_1 \in U'_{DT}(t_i)$ and the parameters ε ε can beThe opinion vector of the user uIly set.The opinion vector of the user ushould be dissimilar to the opinion $u_1 \in U'_{DT}(t_i)$ should be similar to the opinionvectors of the users $U_{DT}(t_i) \setminus U'_{DT}(t_T)$ $u_1 \in U'_{DT}(t_i)$ vectors of the user u

We study the case in which $t_{\tau} = t_1$ and the set $U'_{DT}(t_1) = \{u_1\}$

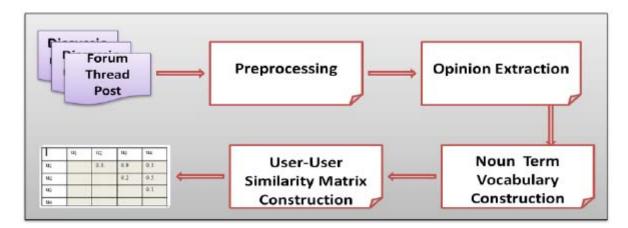


Fig. 2. General architecture of the user-level opinion propagation analysis in online forum threads

Step 1: Preprocessing

• tokenization, part-of-speech tagging, syntactic parsing, and coreference resolution

Step 2: Opinion Extraction

- The Stanford Dependencies:
 - binary semantic relations in a sentence between a *governor* and a *dependent* (abbreviated_relation_name(governor, dependent))



- "dobj", "nsubj", "amod", "acomp", "advmod", "xcomp ", "neg"
- For example, "nsubj" (nominal subject): the governor is any adjective, and the dependent is a noun term
 - For the sentence "The movie is interesting" the Stanford dependencies are: det (movie, The), nsubj (interesting, movie), cop (interesting, is), root (ROOT, interesting)

Step 3: Noun Term Vocabulary Construction

Algo	rithm 1: Construction of Noun-Term Vocabulary	Algorithm 2: Mining Dependency Relations from a Forum Thread Input: $P_{DT}(t_r) = \{p_i\}_{i \in \mathbb{N}}$ - set of posts in the forum thread at time step t_i .	
Inpu	t : $P_{DT}(t_{\tau}) = \{p_{\tau}\}_{\tau \in \mathbb{N}}$ - set of posts in the forum thread at time step t_{τ} ;	Output: $\Omega = \{(h_m, d_n)\}_{m,n \in \mathbb{N}}$ - pairs (noun_term, opinion_word)	
Inpu	t : $S_{DT} = \{W_k\}_{k \in \mathbb{N}}$ – the subject of the forum thread;	1: $\Omega \leftarrow \emptyset$ 2: for each post p_i in $P_{DT}(t_e)$ do	
Outp	ut: $V^{d}_{DT}(t_{\tau})$ – noun-term vocabulary at time step t_{τ} ;	 R(p_i) ← Parse(p_i) // the set {rel_{ij}(h_{i,k}, d_{i,l})}_{j,k,l∈N} of dependency relat 4: 	ions from
1:	$\Omega \leftarrow MiningDependencyRelations (P_{DT}(t_{t})) // \Omega = \{(h_{m}, d_{n})\}_{m,n \in \mathbb{N}}$	5: for each dependency relation $rel_{ij}(h_{ijk}, d_{ij})$ in $R(p_i)$ do	post p _i ;
2:	for each word wk in S _{DT} do	6: if $(rel_{i,j} = 'dobj' and checkVerb(h_{i,k}) and checkNoun(d_{i,l}) then 7: \Omega \leftarrow \Omega \cup (h_{i,k}, d_{i,l})$	
3:	if $(checkNoun(w_k) \text{ and } length(w_k) > 3)$ then	8: end if	
4:	$w_k \leftarrow \text{lemmatization}(w_k)$	9: if $(rel_{i,j} = 'nsubj' and checkAdjective(h_{i,k}) and checkNoun(d_{i,l}) then 10: \Omega \leftarrow \Omega \cup (h_{i,k}, d_{i,l})$	
		10: $\Omega \leftarrow \Omega \cup (h_{i,k}, d_{i,l})$ 11: end if	
5:	$w_k \leftarrow lowercase(w_k)$	12: if $(rel_{i,j} = 'amod' and checkAdjective(h_{i,l}) and checkNoun(d_{i,k}) then$	
6:	else	13: $\Omega \leftarrow \Omega \cup (d_{i,l}, h_{i,k})$	
7:	$S_{DT} \leftarrow S_{DT} \setminus W_k$	14: end if	
		15: if (rel _{i,j} = 'advmod' and checkVerb(h _{i,k}) and there is rel _{i,jj} (h _{i,kk} , d _{i,ll}) so	
8:	end if	16: $\operatorname{rel}_{i,jj} = \operatorname{insubj}^{i}$ and $h_{i,k} = h_{i,jj}$	i,kk then
9:	end for	$\begin{array}{rcl} 17: & \Omega \leftarrow \Omega \cup (d_{i,l}, d_{i,l}) \\ 18: & \text{end if} \end{array}$	
		19: if $(rel_{i,i} = 'acomp' and checkVerb(h_{i,k}) and there is rel_{i,ii}(h_{i,kk}, d_{i,il}) so t$	hat
10:	for each pair (h_m, d_n) in Ω do	20: $rel_{ij} = rcmod$ and $h_{ik} = r$	
11:	$sim \leftarrow \sum_{w_k \in S_{DT}} sim_{w_k}(h_m, w_k)$	21: $\Omega \leftarrow \Omega \cup (d_{i,l}, h_{i,kk})$	
12:	if sim $!= 0$ and $h_m \notin V_{DT}^d(t_r)$ then	22: end if	
		23: if (rel _{i,j} = 'xcomp' and checkVerb(h _{i,k}) and there is rel _{i,jj} (h _{i,kk} , d _{i,ll}) so t	
13:	$V^{d}_{DT}(t_{\tau}) \leftarrow V^{d}_{DT}(t_{\tau}) \cup (h_{m})$	24: rel _{i,jj} = 'ccomp' and $h_{i,k}$ = 25: $\Omega \leftarrow \Omega \cup (h_{i,kk}, d_{i,l})$	d _{ill} then
14:	end if	25: $S_2 \leftarrow S_2 \cup (h_{i,kk}, a_{i,1})$ 26: end if	
		27: end for	
15:	end for	28: end for	

Step 3: Noun Term Vocabulary Construction

Algorithm 1: Construction of Noun-Term Vocabulary	
Input: $P_{DT}(t_{\tau}) = \{p_{\tau}\}_{\tau \in \mathbb{N}}$ - set of posts in the forum thread at time step t_{τ} ;	
Input: $S_{DT} = \{W_k\}_{k \in \mathbb{N}}$ – the subject of the forum thread;	
Output: $V^{d}_{DT}(t_{\tau})$ – noun-term vocabulary at time step t_{τ} ; 1: $\Omega \leftarrow MiningDependencyRelations (P_{DT}(t_{\tau})) // \Omega = \{(h_m, d_n)\}_{m,n \in \mathbb{N}}$ 2: for each word w_k in S_{DT} do	The extraction of pairs (noun_term, opinion_word) by using the dependency relations considered for opinion mining
3: if $(checkNoun(w_k) \text{ and } length(w_k) > 3)$ then 4: $w_k \leftarrow lemmatization(w_k)$ 5: $w_k \leftarrow lowercase(w_k)$	
$ \begin{array}{l} \textbf{6:} \textbf{else} \\ \textbf{7:} S_{DT} \leftarrow S_{DT} \setminus w_k \\ \textbf{8:} \textbf{end if} \end{array} $	The identification of the nouns terms in the subject of the forum thread
9: end for 10: for each pair (h_m, d_n) in Ω do	
11: $\operatorname{sim} \leftarrow \sum_{w_k \in S_{DT}} \operatorname{sim}_{wu}(h_m, w_k)$ 12: if $\operatorname{sim} != 0$ and $h_m \notin V^d_{DT}(t_t)$ then 13: $V^d_{DT}(t_t) \leftarrow V^d_{DT}(t_t) \cup (h_m)$ 14: end if	We remove the noun terms that are not relevant to the subject of the forum thread
15: end for	

Opinion Lexicon	Opinion Lexicon Dimension	Sentiment Categories	Characteristics
SentiWordNet 3.0 (based on WordNet 3.0)	155,287 words from 117,659 synsets	Each word has a score ranging between 0 and 1 for each positive, negative, or neutral sentiment category.	 It is semi-automatically generated. It distinguishes between parts of speech of the opinion words.
Micro-WNOp (based on WordNet 2.0)	1,960 words from 1,105 synsets	Each word has a score ranging between 0 and 1 for each positive, negative, or neutral sentiment category.	 It is manually generated. It distinguishes between parts of speech of the opinion words.
MPQA Subjectivity Lexicon	8,221 words	Each word is in one of the following sentiment categories: positive, negative, both (positive and negative), or neutral.	 It is semi-automatically generated. It distinguishes between parts of speech of the opinion words.
Bing Liu's Opinion Lexicon (based on WordNet 2.0)	6,786 words	Each word is in one of the following sentiment categories: positive or negative.	 It is semi-automatically generated. It does not distinguish between parts of speech of the opinion words.

Opinion Lexicon	Opinion Lexicon	Sentiment Categories	Characteristics
	Dimension		
SentiWordNet 3.0	155,287 words	Each word has a score	• It is semi-automatically
(based on WordNet	from 117,659	ranging between 0 and 1 for	generated.
3.0)	synsets	each positive, negative, or	• It distinguishes between parts
		neutral sentiment category.	of speech of the opinion words.
Micro-WNOp	1,960 words from	Each word has a score	• It is manually generated.
(based on WordNet	1,105 synsets	ranging between 0 and 1 for	• It distinguishes between parts
2.0)		each positive, negative, or	of speech of the opinion
		neutral sentiment category.	words.
MPQA Subjectivity	8,221 words	Each word is in one of the	• It is semi-automatically
Lexicon		following sentiment	generated.
		categories: positive, negative,	• It distinguishes between
		both (positive and negative),	parts of speech of the
		or neutral.	opinion words.
Bing Liu's Opinion	6,786 words	Each word is in one of the	• It is semi-automatically
Lexicon		following sentiment	generated.
(based on WordNet		categories: positive or	• It does not distinguish
2.0)		negative.	between parts of speech of
			the opinion words.

Opinion Lexicon	Opinion Lexicon Dimension	Sentiment Categories	Characteristics
SentiWordNet 3.0 (based on WordNet 3.0)	155,287 words from 117,659 synsets	Each word has a score ranging between 0 and 1 for each positive, negative, or neutral sentiment category.	 It is semi-automatically generated. It distinguishes between parts of speech of the opinion words.
Micro-WNOp (based on WordNet 2.0)	1,960 words from 1,105 synsets	Each word has a score ranging between 0 and 1 for each positive, negative, or neutral sentiment category.	 It is manually generated. It distinguishes between parts of speech of the opinion words.
MPQA Subjectivity Lexicon	8,221 words	Each word is in one of the following sentiment categories: positive, negative, both (positive and negative), or neutral.	 It is semi-automatically generated. It distinguishes between parts of speech of the opinion words.
Bing Liu's Opinion Lexicon (based on WordNet 2.0)	6,786 words	Each word is in one of the following sentiment categories: positive or negative.	 It is semi-automatically generated. It does not distinguish between parts of speech of the opinion words.

Opinion Lexicon	Opinion Lexicon Dimension	Sentiment Categories	Characteristics
SentiWordNet 3.0 (based on WordNet 3.0)	155,287 words from 117,659 synsets	Each word has a score ranging between 0 and 1 for each positive, negative, or neutral sentiment category.	 It is semi-automatically generated. It distinguishes between parts of speech of the opinion words.
Micro-WNOp (based on WordNet 2.0)	1,960 words from 1,105 synsets	Each word has a score ranging between 0 and 1 for each positive, negative, or neutral sentiment category.	 It is manually generated. It distinguishes between parts of speech of the opinion words.
MPQA Subjectivity Lexicon	8,221 words	Each word is in one of the following sentiment categories: positive, negative, both (positive and negative), or neutral.	 It is semi-automatically generated. It distinguishes between parts of speech of the opinion words.
Bing Liu's Opinion Lexicon (based on WordNet 2.0)	6,786 words	Each word is in one of the following sentiment categories: positive or negative.	 It is semi-automatically generated. It does not distinguish between parts of speech of the opinion words.

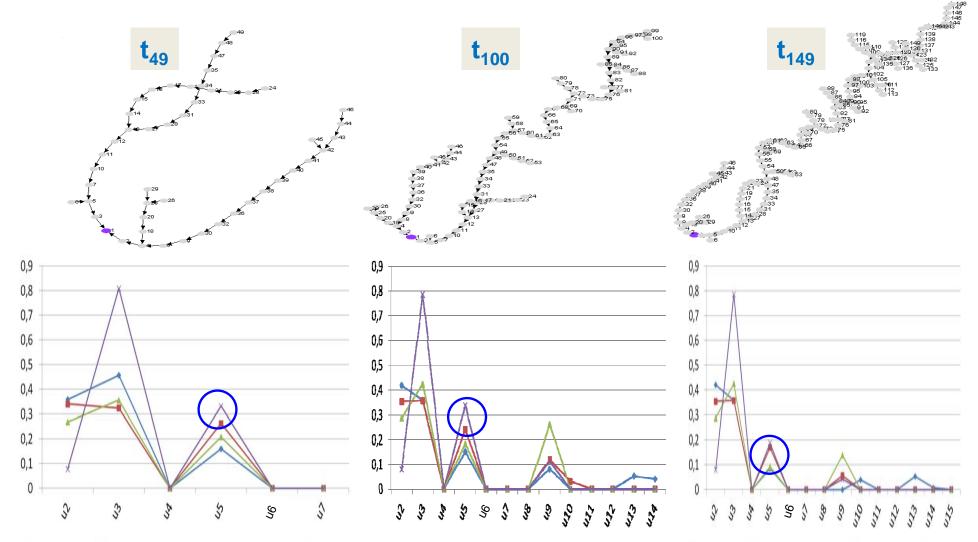
Algorithm 3: Opinion Word Sentiment Mining using SentiWordNet / Micro-WNOp	Algorithm 4: Opinion Word Sentiment Mining using Bing Liu's opinion lexicon
Input: w - opinion word;	Input: w – opinion word; Output: a value $c_{1}(1, \pm 1)$ – indicates the continuent of animien word w where: 1 denotes a
Input: pos – part of speech for opinion word w;	Output: a value \in {-1, +1} – indicates the sentiment of opinion word w, where: -1 denotes a negative sentiment and +1 denotes a positive sentiment;
Output: a value $\in \{-1, 0, +1\}$ – indicates the sentiment of opinion word w, where: -1 denotes a	
	2: return 0
negative sentiment, 0 denotes a neutral sentiment, and +1 denotes a positive sentiment;	3: else
 if !FindOpinionWord(w, pos) then 	4: if w ∈ GetPostiveList() then
2: return 0	5: return +1
3: end if	6: end if 7: if m. a. Cathlanding list() then
4: S ← GetSenses(w, pos)	7: if w ∈ GetNegativeList() then 8: return -1
5: for each $s \in S$ do	9: end if
	10: end if
6: difference ← GetPositiveScore(w, s) - GetNegativeScore(w, s)	
7: if difference = 0 and GetPositiveScore(w, s) != 0 then	Algorithm 5: Opinion Word Sentiment Mining using the MPQA subjectivity lexicon Input: w - opinion word;
8: continue	Input: pos – part of speech for opinion word w;
9: end if	Output: a value \in {-1, 0, +1} – indicates the sentiment of opinion word w, where: -1 denotes a
 if difference = 0 and GetPositiveScore(w, s) = 0 then 	negative sentiment, 0 denotes a neutral sentiment, and +1 denotes a positive sentiment; 1: if !FindOpinionWord(w, pos) then
11: return 0	2: return 0
	3: else
12: end if	 4: polarity ← GetPolarity(w) 5: if polarity = "positive" then
13: if difference > 0 then	6: return +1
14: return +1	7: end if 8: if polarity = "negative" then
15: end if	9: return -1
16: if difference < 0 then	10: end if
17: return -1	11: if polarity = "neutral" then 12: return 0
18: end if	13: end if
	14: if polarity = "both" then
19: end for	15: return +1 16: end if
20: return +1	17: end if

Dataset

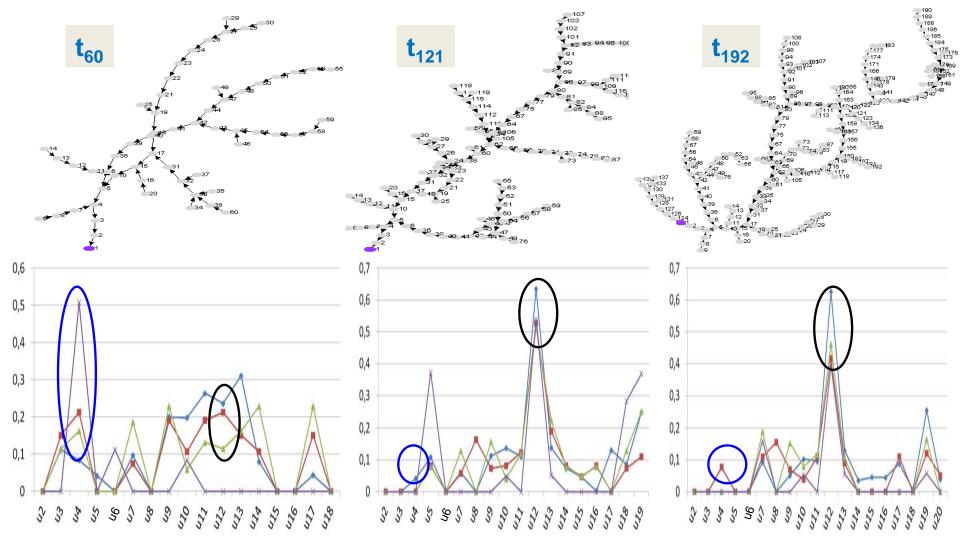
- We use the Internet Argument Corpus (IAC)
 - http://www.4forums.com
 - A dataset freely available
 - Each discussion thread is saved in the JSON format

Forum Thread Identifier	Forum Thread Subject	Time Steps (t _j)	Number of Users (U _{DT} (t _j))
Forum Thread 1	"proof of God's existence"	t49	7
		t ₁₀₀	14
		t ₁₄₉	15
Forum Thread 2	"Atheists - America's greatest threat?"	t ₆₀	18
		t ₁₂₁	19
		t ₁₉₂	20

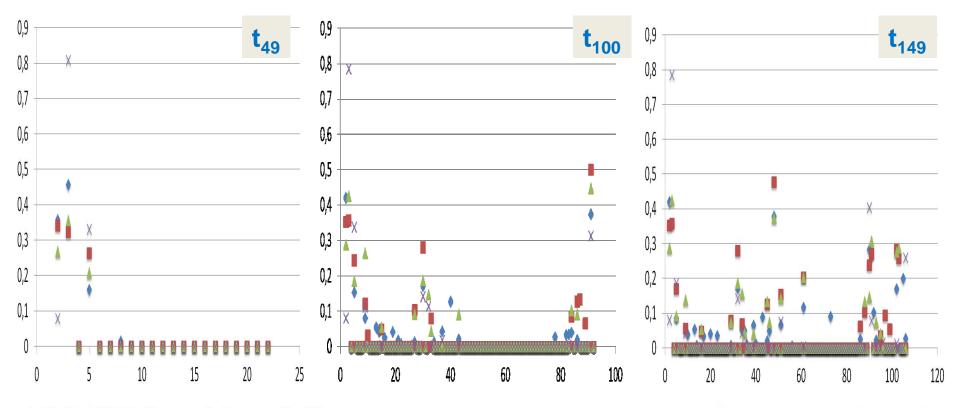
Table 1. Statistics on the experimental corpus.



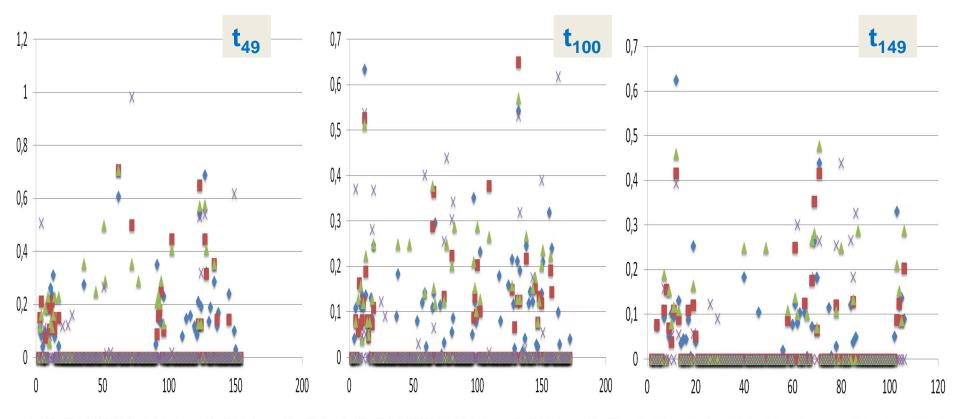
SentiWordNet - MPQA Subjectivity Lexicon - Bing Liu's Lexicon - MicroWNOp - SentiWordNet - MPQA Subjectivity Lexicon - Bing Liu's Lexicon - MicroWNOp - SentiWordNet - MPQA Subjectivity Lexicon - Bing Liu's Lexicon -



--SentiWordNet --MPQA Subjectivity Lexicon --Bing Liu's Lexicon --MicroWNOp --SentiWordNet --



◆ SentiWordNet ■ MPQA Subjectivity Lexicon ▲ Bing Liu's Lexicon ★ MicroWNOp SentiWordNet ■ MPQA Subjectivity Lexicon ★ Bing Liu's Lexicon ★ MicroWNOp
Fig. 5. The values of the user-user similarity matrix B_{U-U}



◆ SentiWordNet ■ MPQA Subjectivity Lexicon ▲ Bing Liu's Lexicon ★ MicroWNOr ◆ SentiWordNet ■ MPQA Subjectivity Lexicon ▲ Bing Liu's Lexicon ★ MicroWNOp

Fig. 6. The values of the user-user similarity matrix B_{U-U}

Conclusions

- More research on information propagation, but little on opinion propagation in social media
- It is useful to solve the opinion propagation problem because of its multiple applications in the real world
- We proposed a user-level opinion propagation analysis method in online forum threads, by combining opinion mining and natural language processing techniques
- The results of our method for the opinion propagation problem depend on the opinion mining techniques we use

Future Work

- Application of our opinion propagation analysis method to other online forum threads
- Analysis of users' behavior in the opinion propagation process in online forum threads
- Study of the opinion propagation problem in different types of social media, such as blogs and online social networks (e.g. Twitter), in order to detect and describe the opinion propagation process in these media

Thank You!

