On-line Argumentative Rationality Analysis by case of TCM Debate at Tianya Forum

Xijin Tang[1,2]
[1]Institute of Systems Science
[1]Academy of Mathematics and Systems Science, Chinese Academy of Sciences
Beijing, 100190, China

Can Wang[1]
[2] School of Economics and Management
University of Chinese Academy of Sciences
Beijing, 100049 , China

Abstract—**Nowadays Internet-based social media platforms provide rooms for wider discussions toward a variety of topics, from normal living affairs to national policies, from gossips of celebrities to fight between square dancers and young basketball fans, from folk remedies to scientific stories, etc. All those emerging comments often reflect communities concerns towards those topics to some extent. Some topics may arouse hot discussions or long lasting debates. Instead as many studies conducted to analyze the on-line behavior patterns, such as general replies patterns fit power law, etc., this paper concentrates on debates styles, i.e. by what kind of ways those on-line debates go on, especially toward those bi-polarization debates. With Paul Graham's disagreement hierarchy and LDA topic models, the debate styles are analyzed toward the on-line traditional Chinese medicines debates in the famous Chinese BBS, Tianya Forum. Such kind of analysis aims to expose the rationality about the on-line debates and then may be helpful to conduct intervention for quality debates.**

Keywords- On-line debates; opinion mining; disagreement hierarchy; LDA, traditional Chinese medicine

I. INTRODUCTION

The Internet technologies facilitate people to express their personal opinions at the cyberspace. BBS, blogs, microblogs, and a variety of social media platforms not only record people's emerging wild ideas, but also accelerate and expand those thoughts or opinions dissemination which often inspires more involvements and brings potential impacts, especially those on-line discussions about a variety of topics across all the fields. Wide communities participation, insightful ideas for different groups of people, disclosure of unknown stories, stances and arguments, especially the on-line discussions and debates not only show the emerging and accumulating of public opinions, but illustrate the potentiality of those public opinions toward policy making. In this paper we refer those on-line debates as on-line free discussion on one topic, such as abortion, political elections, etc. and always with two polarized opinions. Lots of studies have been conducted toward on-line debates. Among two main streams of literature in this domain, one is to distinguish subjective expressions from factual information [1, 2], another is to detect the text polarity, positive or negative; both to label the sentiments or judgments toward the concerned topic, e.g. agree or disagree, etc. Those studies heavily focused on feature selection [3-5] and classifiers optimization [6] for performance improvement. Most of those studies belong to opinion mining with limitations toward the practical debate analysis. Usually opinion mining and sentiment analysis are used as synonyms, by applying data mining and natural language processing (NLP) techniques to process textual information [7], while sentiments cannot truly represent stances [8]. Besides, corpora are important for opinion mining. Many studies used users' comments or news as corpora. Those studies mostly rely on existing lexicons, or generated lexicons by seed words [9], mainly based on experiences. While on-line debates are more diverse, conversational and highly contextualized; one word may have opposite meanings under different contexts. Some studies focused on automatically determining the stance of a participant, such as Wiebe's group from University of Pittsburgh [10, 11], Anand et al. from University of California Santa Cruz [12-14], and Tikves et al. from Arizona State University on profiling Islamic organizations' ideology and activity patterns [15,16], limited researches are seen on illustrating how people express their different perspectives towards the concerned topics along the unstructured on-line debate. By reviewing the up-to-date stance analysis research, Wang and Tang then tried stance analysis toward on-line debate on traditional Chinese medicine (TCM) [17]. The debates over TCM in China last long and lead to 2-polarization; thus even the attitudes on TCM are used as one question in the Zuobiao Survey (i.e. China political compass survey) [18].

As depicted by [17], one hot post on TCM lasting one year from one of the most influential Chinese BBS, Tianya Forum, was selected to explore machine learning to determine the replies' stances about TCM. Two ways to select features were tested for SVM classifiers, while using logistic regression to select domain feature words outperformed using adjectives, adverbs, verbs and nouns as features. Furthermore, logic regression was conducted to select top discriminating technical terms and human names for both "preserving TCM" and "abolishing TCM" stances respectively to illustrate specific arguments from each side during the debate. Moreover, 10 topics were generated for both camps by latent Dirichlet allocation (LDA) respectively to explicitly indicate that the emphases of the two camps were different during the debate. The “preserving TCM” stance holders concern the motivations of the opposite camp, the effectiveness of the TCM, etc. The “abolishing TCM” stance holders doubt the scientific nature and the rationality of TCM, introduce the modern medicine, and condemn the illegal medical practice relevant to TCM.

978-1-5386-2623-8/17 $31.00 © 2017 IEEE
DOI 10.1109/KICSS.2017.32

All of above addressed stance analysis studies focus on contents. Due to anonymity and informality, emerging new Internet languages (such as a variety of acronyms) and oral utterances are flooded with lots of war of words happen with the on-line debates. To some extent, various diatribes show that verbal violence is a normal state of the on-line debates as no obligation is concerned. With no strict regulations, rational debates may be very difficult to be achieved, while still lots of useful information are available. Thus it is interesting to explore how to extract rational or meaningful contents for better understanding and how to achieve more rational discussions so as to show a comprehensive scenario toward the concerned topic.

This paper goes beyond previous work and concentrates on debates styles, i.e. by what kind of ways those on-line debates go on still with the on-line TCM debates as illustrations. Paul Graham's disagreement hierarchy is adopted to show how much the debate is rational by their arguments based on those topics generated by LDA models. The rest of the paper is organized as follows. Section II describes the basic figures of the highlighted TCM posts at Tianya Club and their generated topics by LDA from two camps. Section III analyzes those topics from both camps to illustrate how rational the on-line debate is based on Paul Graham's disagreement hierarchy. After constructing the on-line debate network based on replying among the participants, the key players and their distributions among those generated topics from two camps are discussed in Session IV to show the key players' rational performance within the debates. Conclusions are presented in Section V.

II. TOPIC MODELING ON 2 CAMPS AT THE ON-LINE TCM DEBATES

As addressed in [17], the TCM debate is a typical polarized debate with two camps, "preserving TCM" and "abolishing TCM" and the everlasting TCM debates in daily life also exist at the social media platforms. There are many threads on TCM topic with replies more than 5000 at Tianya Forum. As the main disputes toward TCM have not changed great during the past years, this paper continues to study the highlighted 3 posts as listed in Table I.

TABLE I. HOT POSTS ABOUT TCM AT TIANYA FORUM [17]

Thread ID	*Reply #*	*Participants*	*Start Day*	*End Day*
2822432	117318	4890	2012-10-16	2013-11-29
2121178	36592	5522	2011-03-21	2015-01-24[a]
2317943	33547	6067	2011-11-12	2015-01-24[a]

a. Data were captured to that day

A. Labeling the Stance of Posts

Firstly, we label the replies by user IDs' stances. For simplification, 603 participants (authors) with more than 10 replies are selected to be manually labeled with results listed in Table II. Authors may join different threads, thus we have 179 authors on "abolishing TCM" and 417 on "preserving TCM" among those labeled authors. The stance of one post is its author's stance. In TCM debates, few people change stance. Then, we have 75153 posts on "preserving TCM" and 59414 posts on "abolishing TCM". Next we generate topics based on those posts from 2 camps to see what kind of perspectives each camp exposes during the debates.

TABLE II. STANCE LABELING TOWARD HOT TCM POSTS AT TIANYA FORUM

Post label	*Authors #*	*Labeled authors*	*Preserving TCM #*	*Abolishing TCM #*	*No Stance*
zy-1	4890	282	199	82	1
zy-2	5522	288	186	94	8
zy-3	6067	259	173	84	2

B. Generating Two Camps' Topics of TCM Debate

Latent Dirichlet allocation (LDA) is used to generate topics for each camp. LDA is a generative statistical model which treats documents as bags of words generated by one or more topics [19]. In LDA, each document may be viewed as a mixture of various topics where each document is considered to have a set of topics that are assigned to it via LDA.

Here is the procedure to generate the topics from replies on either "preserving TCM" or "abolishing TCM" stance.

1) Remove replies with fewer than 10 Chinese characters.

2) Filter out urls.

3) Segment words with Rwordseg package[1]. One TCM terminology dictionary with 28428 TCM technical terms from Sougou Cell dictionary[2] and all participants' Tianya IDs are selected as reserved words.

4) Remove stop words (such as "oh") from the bag of words and words with only one character.

5) Calculate perplexity of the topic models for each camp. Based on computation results, we select to generate 30 topics for each camp at each thread.

6) Generate 30 topics for each camp using the "topicmodels" package[3] in R.

Due to limited spaces, only part of results are listed in Tables III and IV for illustrations with no original Chinese words.

TABLE III. PART OF 30 GENERATED TOPICS FROM "ABOLISHING TCM" CAMP OF THREAD "ZY-1"

Topic label	*Topic focus*	*Topic label*	*Topic focus*	*Topic label*	*Topic focus*
Ta-1	Case I	Ta-11	TCM Toxic side-effects II	Ta-20	Ridicules on TCM supporters
Ta-2	Artemisinin	Ta-12	Meridian-acupoints I	Ta-21	TCM is out of date
Ta-3	Ridicules on TCM logic	Ta-15	*Yin Yang* and five elements	Ta-22	Science discovery
Ta-4	TCM disciples	Ta-16	TCM therapeutic methods	Ta-23	TCM supporter's logic in debates
Ta-6	TCM Toxic side-effects I	Ta-17	TCM and Modern medicine	Ta-26	Western medicine popularity
Ta-7	Deception of TCM I	Ta-18	Virusology	Ta-27	Meridian-acupoints II

[1] http://cran.r-project.org/web/packages/Rwordseg/

[2] http://pinyin.sogou.com/dict/detail/index/20664

[3] http://cran.r-project.org/web/packages/topicmodels/

Topic label	*Topic focus*	*Topic label*	*Topic focus*	*Topic label*	*Topic focus*
Ta-8	Deception of TCM II	Ta-19	TCM education system	Ta-30	TCM supporters normal attitudes

TABLE IV. PART OF 30 GENERATED TOPICS FROM "PRESERVING TCM" CAMP OF THREAD "ZY-1"

Topic label	*Topic focus*	*Topic label*	*Topic focus*	*Topic label*	*Topic focus*
Tp-1	TCM current situation	Tp-9	Ridicules on against TCM people II	Tp-19	Side-effects of western medicines
Tp-2	Mission of "abolishing TCM"	Tp-10	Ridicules on against TCM people III	Tp-22	Insults on TCM opponents III
Tp-3	Disadvantages of western medicine	Tp-12	Insults on TCM opponents I	Tp-22	Science discovery
Tp-4	TCM prescriptions	Tp-14	Science, TCM and philosophy	Tp-23	Insults on TCM opponents IV
Tp-5	TCM going abroad I	Tp-15	Insults on TCM opponents II	Tp-25	Artemisinin
Tp-6	TCM theory	Tp-16	Diseases and TCM treatments	Tp-26	Mission of TCM opponents'
Tp-7	Side-effects of antibiotics	Tp-17	TCM opponents' logic	Tp-29	CASE IV
Tp-8	Ridicules on against TCM people I	Tp-18	TCM successful case	Tp-30	TCM opponents are liars

From all those topics by 3 threads relevant to 2 camps, it can be summarized that main argumentative topics from the "preserving TCM" camp during the on-line debate indicate the side-effects of western medicine, emphasize TCM's functions, prescriptions, and successful cases, while also concerning the motivations of the opposite camp and insulting the opponents. The arguments from the "abolishing TCM" camp cover explanation of western medicine, doubts toward the scientific nature and rationality of TCM, illustrations of ineffectiveness of TCM, toxic and side-effects of TCM, condemns on the illegal medical practice relevant to TCM together with lots of ridicules. Next we go further to observe the rationality of those arguments.

III. DISAGREEMENT HIERARCHY OF TCM DEBATES

The normative debates always follow the general rules. Kunz and Rittel (1970) proposed an argumentation scheme *issue based information system* (IBIS) as a way to support coordination and planning of political decision process [20] and later Conklin and Begeman (1988) developed an IBIS based computerized tool [21], which gradually evolved into QuestMap and support so-called Dialog Mapping for group discussions process [22]. After entering into Web Era, Klein and Iandol (2008) reported a study using Collaboratorium, with the same function as QuestMap [23]. They argued that the open-source/peer production (OSPP) technology enabled large scale distributed participation but was not capable of collaborative deliberation, since the coverage of a topic was created bottom-up and then generally unsystematic. That kind of technology was more time-based, while collaborative deliberation required logic-based postings. As a matter of fact, traditional argumentative support tools help to obtain possible structures of the unstructured problems while sacrifice freedom of wild thinking and then may lead to loss of novel ideas, the typical disadvantages of consensus built top-down. While those on-line discussions at BBS are always happened bottom-up and difficult to be logically all through the debate especially emerging participation with anonymity. Thus general IBIS scheme may not be practicable to structure the on-line debates. Here we apply Graham's disagreement hierarchy to analyze the rationality of the 2-polarization debate on TCM.

A. Disagreement hierarchy by Graham

Paul Graham thought "agreeing tends to motivate people less than disagreeing", and thus disagreeing may expand another's territory. Even it is easier to tell the difference between mere name-calling and a carefully reasoned refutation, Graham proposed a disagreement hierarchy (DH) to put names on the intermediate stages, as shown by Figure 1 with simple explanations for each level, as the lowest level is DH0: name-calling and the highest level is DH6: refuting the central point. Here we simply regard that rationality is decreasing from the highest to the lowest in this hierarchy. Next we analyze the rationality at the TCM on-line debate.

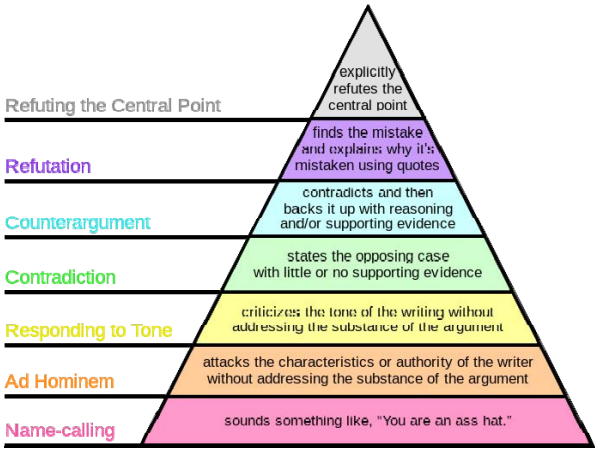

Figure 1. Graham's disagreement hierarchy [24]

B. Rationality of on-line TCM debate

We observe the rationality of debate from the topics generated by two camps from three threads using DH. As not all the topics are listed due to space limitation, here just some illustrations with labeled topics for easy understanding.

DH0: Name-calling is the lowest form of disagreement, such as Ta1-(3, 20) in Table II and Tp1-(8, 9, 12, 15, 23) in Table III. More topics not listed in Table II are within this level.

DH1: "An ad hominem attack is not quite as weak as mere name-calling", such as Ta1-4 and Tp1-(2,10, 22, 30).

DH2: Responding to tone, such as Ta1-(23,30) and Tp1-17.

DH3: Contradiction. From this stage, finally responses are got to what is said, rather than how or by whom. "The lowest form of response to an argument is simply to state the opposing case, with little or no supporting evidence". As replies with fewer than 10 Chinese characters have been removed, no such topics are generated in our case.

DH4: Counterargument means contradiction plus reasoning or evidence. Actually the majority of topics belong to this form, such as Ta1-(1, 6, 7, 8, 16, 17, 19, 21) and Tp1-(1, 4, 5, 6, 16, 18).

DH5: Refutation, which is the most convincing form of disagreement. Due to the pyramid hierarchy, the higher the fewer instances, such as Ta1-2 and Tp1-(14, 15).

DH6: Refuting the central point, which is the most powerful form of disagreement, such as Ta1-(12, 15, 18, 22, 27) and Tp1-(3, 7, 19).

With the topic distributions generated by LDA, we summarize the probabilities of those topics at each DH and acquire the DH distributions for each thread, as shown in Figures 2 and 3.

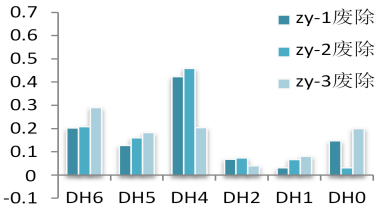

Figure 2. DH distributions at "abolishing TCM" camp of three threads

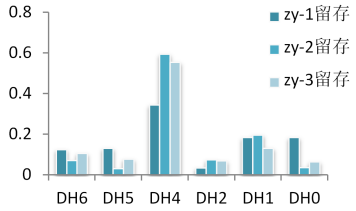

Figure 3. DH distributions at "preserving TCM" camp of three threads

We count the topics into each level for 2 camps of the 3 threads and find that there are more topics from "abolishing TCM" within DH4 to DH6 (23 for zy-1 and 25 for zy-2) than those from "preserving TCM" (19 for zy-1 and 21 for zy-2) for two threads. Only for Thread zy-3, topics from two camps are balanced. Such a study is to illustrate "abolishing TCM" camp expressing their opinions more rational than their opponent.

IV. KEY PLAYERS' PERFORMANCE ALONG THE DEBATES

During the on-line debate happened at BBS, the replying relationships between the participants in one thread construct a debate network. The vertices (nodes) of the network are authors and the edges (links) indicate an author comments on at least one previous message from another author. The network is directed and there are at most two links between author *i* and author *j* (*i*→*j*, *j*→*i*). In the real world, two participants may communicate many times while the frequencies are not used as the weights of the edges in this paper.

A. Detecting Key Players from on-line Debate Network

Wang and Tang have constructed three reply networks based on three threads, where reply network for Thread "zy-1" (Network 1) has 4890 authors, reply network for Thread "zy-2" (Network 2) has 5514 authors and reply network for Thread "zy-3" (Network 3) has 6065 authors [25] and conducted network structure characteristics analysis together with structure balance analysis. Here we study key players from different measures within the network. Table V lists key players by in-degree, out-degree, betweenness, Page Rank and cutpoint at the reply network of Thread "zy-1".

TABLE V. TOP 15 KEY PLAYERS THREAD "ZY-1" REPLAYING NETWORK

User ID	*In-degree*	*Out-degree*	*Between-ness*	*Rank*	*PageRank*	*Rank*	*Cut-point*
施正义	854	1240	2.15E+06	1	0.1296	1	Yes
恶海捕蟹	198	372	3.07E+05	3	0.0177	2	Yes
djm1000418873	125	36	7.91E+04	4	0.0094	3	Yes
活济公2012	108	158	1.30E+05	2	0.0090	4	Yes
来这里看这里一时	99	266	1.63E+05	6	0.0074	8	Yes
sanbenwu	92	170	8.78E+04	7	0.0063	12	Yes
木口林是马扁子	91	245	1.47E+05	11	0.0064	9	Yes
南洋一游子	88	74	1.20E+05	17	0.0092	10	Yes
davy1002011	84	111	6.39E+04	10	0.0063	7	Yes
枫林隐士	64	192	9.22E+04	22	0.0041	17	Yes
jjsquid	63	172	8.10E+04	29	0.0049	14	Yes
dragon200934	62	0	0.00E+00	16	0.0070	5	Yes
1984 不再天真	58	147	8.95E+04	28	0.0050	34	No
反中医为己任	53	67	3.24E+04	41	0.0031	16	Yes
全真教义杰道人	50	83	3.96E+04	33	0.0038	32	Yes

It can be seen that even ID " djm1000418873" ranks among top 3 by in-degree, betweenness, PageRank and is a cutpoint, the low outdegree shows he did not participate the debate actively, only his posts may be referred by others. Based on all 5 measures, we select "施正义", "恶海捕蟹", "活济公 2012", "来这里看这里一时", "sanbenwu", "木口林是马扁子", "南洋一游子" and "davy1002011" as the key players at Thread "zy-1" debate, where "施正义", "恶海捕蟹", "来这里看这里一时", "木口林是马扁子" and "南洋一游子" hold stance on "abolishing TCM" while "活济公 2012", "sanbenwu" and "davy1002011" are TCM supporters. Those 8 players produce 53,242 replies, 45.38% of all replies at Thread "zy-1", while replies from "活济公 2012" amount to 22,784, up to 19.42% of all replies.

All through those key players detected from three reply networks of three threads, "zy-1", "zy-2" and "zy-3", there are 5, 5, and 5 players from "abolishing TCM" in each thread respectively and 3, 1 and zero players from "preserving TCM" camp. Thus we say key players from "abolishing TCM" camp play more important roles along the debates.

B. *Debate style of key players by DH*

As previously explained, we get the DH distributions at each camp of three threads based on the topic distributions generated by LDA. Here we go further to extract the distributions of key players at different topics and different DHs. Still with Thread "zy-1", Figure 4 shows the DH distributions of 5 key players from "abolishing TCM" camp. It can be seen that "恶海捕蟹" values highest at DH6 while "施正义" and "木口林是马扁子" values higher at DH0.

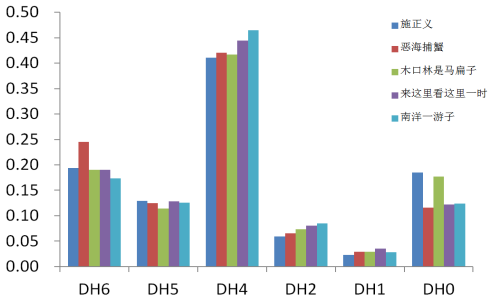

Figure 4. The DH distributions of 5 key players' topics from "abolishing TCM" camp at Thread "zy-1"

Due to space limitation, we do not list the distributions of each player from each camp along those 30 generated topics of all three threads. Table VI is a glimpse of 5 key players' top 3 topics and their corresponding DH levels at Thread "zy-1".

TABLE VI. THE DISTRIBUTIONS OF TOP 3 TOPICS AND RELEVANT DH LEVES OF 5 KEY PLAYERS FROM "ABOLISHING TCM" CAMP AT THREAD "ZY-1"

User ID	*Top 1 topic*	*Prob.*	*DH level*	*Top 2 topic*	*Prob.*	*DH level*	*Top 3 topic*	*Prob.*	*DH level*
木口林是马扁子	Ta1-3	0.0855	0	Ta1-21	0.0613	4	Ta1-30	0.0432	2
施正义	Ta1-20	0.0602	0	Ta1-24	0.0568	0	Ta1-26	0.0492	6
恶海捕蟹	Ta1-27	0.0504	6	Ta1-22	0.0497	6	Ta1-15	0.0456	6
来这里看这里一时	Ta1-30	0.0480	2	Ta1-6	0.0428	4	Ta1-11	0.0387	4
南洋一游子	Ta1-19	0.0535	4	Ta1-23	0.0456	2	Ta1-17	0.0450	4

As a matter of fact, 5 players' performance at the other two threads can also be acquired and then the DH differences of each player at different threads can be analyzed. At different debates, the responses of each player may spread across different DH levels.

Table VII lists the top 3 topics and the corresponding DH levels of the 3 key players from "preserving TCM" camp at Thread "zy-1". We know that the refutations by the 3 key players from "preserving TCM" are by form of DH4 and below, some at DH0 level.

TABLE VII. THE DISTRIBUTIONS OF TOP 3 TOPICS AND RELEVANT DH LEVES OF 3 KEY PLAYERS FROM "PRESERVING TCM" CAMP AT THREAD "ZY-1"

User ID	*Top 1 topic*	*Prob.*	*DH level*	*Top 2 topic*	*Prob.*	*DH level*	*Top 3 topic*	*Prob.*	*DH level*
活济公2012	Tp1-29	0.0881	4	Tp1-22	0.0837	1	Tp1-16	0.0795	4
davy1002011	Tp1-16	0.0795	4	Tp1-10	0.0429	1	Tp1-9	0.0403	0
Sanbenwu	Tp1-12	0.1496	0	Tp1-15	0.0654	0	Tp-17	0.0571	2

V. CONCLUSIONS

Nowadays on-line free discussions and debates by presenting rich pictures with wide participation, insightful ideas from different communities, disclosure of unknown stories, arguments and evidences, etc. not only show the emerging and accumulating of public opinions, but illustrate the potentiality of those public opinions toward policy making. How to extract the useful information and arguments while excluding the low-level conversations is a challenging task. This paper uses a typical 2-polarization on-line debate on TCM for some trials.

The studies are undertaken from three perspectives. Topic modeling is conducted to generate main topics from all replies of each camp from each thread, to show how each camp, "abolishing TCM" or "preserving TCM" expresses their standpoints during the debate. Then we adopt a disagreement hierarchy to further illustrate the rationality of those topics. It is seen that different camp demonstrates different rational performance during the on-line free debate. Moreover, based on emerging reply network, we detect key players and go further to acquire the distributions of topics and DH forms of each player from different camps to display the rational performance of those key players during the debates. Such kind of studies help to accumulate the reasons and evidences for the arguments during the debates. Moreover, it may also helpful to expose the potential ways to intervene the ongoing on-line debate via key players' behaviors.

Habernal and Gurevych are exploring to provide convincing argument from Web argumentation [26]. By their ideas, it is worth going further not only analyze the rationality of debate ways, but also the weights of evidences within the free debates.

ACKNOWLEDGMENT

This work is supported by the National Key Research and Development Program of China under grant 2016YFB1000902 and the Natural Science Foundation of China (Nos. 61473284 and 71371107).

REFERENCES

[1] E. Riloff, "Automatically Generating Extraction Patterns from Untagged Text", in the proceedings of 13th National Conference on Artificial Intelligence, Portland, 1996, pp.1044-1049.

[2] E. Riloff, and J. Wiebe, "Learning Extraction Patterns for Subjective Expressions", in the proceedings of Conference on Empirical Methods in Natural Language Processing, Sapporo, 2003, pp.105-112.

[3] H. Cui, V. Mittal, and M. Datar, "Comparative Experiments on Sentiment Classification for Online Product Reviews", in the proceedings of the 21st National Conference on Artificial Intelligence, Boston, 2006, pp.61-80.

[4] V. Ng, S. Dasgupta and S. M. N. Arifin, "Examining the Role of Linguistic Knowledge Sources in the Automatic Identification and Classification of Reviews", in the proceedings of the International Conference on Computational Linguistics and Meeting of the Association for Computational Linguistics, Sydney, 2006, pp.381-393.

[5] M. Gamon, "Sentiment Classification on Customer Feedback Data: Noisy Data, Large Feature Vectors, and the Role of Linguistic Analysis", in the proceedings of the 23rd International Conference on Computational Linguistics, Beijing, 2010, pp.841-847.

[6] B. Pang, L. Lee and S. Vaithyanathan, "Thumbs Up?: Sentiment Classification Using Machine Learning Techniques", in the proceedings of the Conference on Empirical Methods in Natural Language Processing, Philadelphia, 2009, pp.79-86.

[7] B. Liu, "Opinion Mining and Sentiment Analysis", Foundations & Trends in Information Retrieval, vol. 2, issues 1-2, 1-135, 2008.

[8] Y. L. Zhao and X. J. Tang, "In-depth Analysis of Online Hot Discussion about TCM", in the proceedings of the 15th International Symposium on Knowledge and Systems Sciences, JAIST Press, Sapporo, 2014, pp.275-283.

[9] P. D. Turney and M. L. Littman, "Measuring praise and criticism: Inference of semantic orientation from association", ACM Transactions on Information Systems, vol. 21, issue 4, pp315-346, 2003

[10] W. H. Lin, T. Wilson, and J. Wiebe, "Which Side Are You on? Identifying Perspectives at the Document and Sentence Llevels", in the proceedings of the 10th Conference on Computational Natural Language Learning, New York, 2006, pp.109-116.

[11] S. Somasundaran and J. Wiebe, "Recognizing stances in ideological on-line debates", in the proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, Los Angeles, 2010, pp.116-124.

[12] R. Abbott, M. Walker, P. Anand, et al., "How Can You Say Such Things?!?: Recognizing Disagreement in Informal Political Argument", in the Proceedings of the Workshop on Language in Social Media, Portland, 2011, pp.2-11.

[13] P. Anand, M. Walker, R. Abbott, et al., "Cats Rule and Dogs Drool!: Classifying Stance in Online Debate", in the Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, Portland, 2011, pp.1-9.

[14] M. A. Walker, P. Anand, R. Abbott, et al., "That is Your Evidence?: Classifying Stance in Online Political Debate", Decision Support Systems, vol. 53, issue 4, pp.719-729, 2012.

[15] S. Tikves, S. Gokalp, M. Temkit, et al., "Perspective Analysis for Online Debates", in the Proceedings of the Interna-tional Conference on Advances in Social Networks Analysis and Mining, Istanbul, 2012, pp.898-905.

[16] S. Tikves, S. Banerjee, H. Temkit, et al., "A System for Ranking Organizations Using Social Scale Analysis", Social Networks Analysis & Mining, vol.3, issue 3, pp.313-328, 2013.

[17] C. Wang and X. J. Tang, "Stance Analysis for Debates on Traditional Chinese Medicine at Tianya Forum", in H.T. Nguyen and V. Snasel (Eds.): CSoNet 2016 (Ho Chi Minh City), LNCS 9795, Springer, 2016, pp. 321-332.

[18] J. Pan and Y. Xu, "China's Ideological Spectrum", The Journal of Politics. Available at http://dx.doi.org/10.2139/ssrn.2593377, 2017, Forthcoming

[19] D. M. Blei, A. Y. Ng and M. I. Jordan, "Latent Dirichlet allocation", Journal of Machine Learning Research, vol. 3, pp.993-1022, 2003.

[20] W. Kunz and H. W. J. Rittel, "Issues as elements of information systems" (Working Paper). Berkeley: Institute of Urban and Regional Development, University of California, Berkeley, 1970, https://www.cc.gatech.edu/~ellendo/rittel/rittel-issues.pdf

[21] E. J. Conklin and M. L. Begeman, "gIBIS: a hypertext tool for exploratory policy discussion", ACM Transactions on Information Systems, vol.6, no.4, pp.303–331, 1988.

[22] E. J. Conklin, A. Selvin, S. B. Shum and M. Sierhuis, "Facilitated hypertext for collective sensemaking: 15 years on from IBIS", in the Proceeedings of the 12th ACM Conference on Hypertext & Hypermedia, Arbus, Demark, August 14–18, 2001, pp. 123–124.

[23] M. Klein, and L. Iandol, "Supporting collaborative deliberation using a large-scale argumentation system: the MIT Collaboratorium", in the proceedings of the 11th symposium on directions and implications of advanced computing, Berkeley, 26–29 June 2008, pp 5–12.

[24] P. Graham, How to disagree, http://www.paulgraham.com/disagree.html, 2008

[25] C. Wang and X. J. Tang, "The Online Debate Networks Analysis: A Case Study of Debates at Tianya Forum", in J. Chen et al. (eds.): KSS 2016 (Kobe, Nov 4-5, 2016) CCIS 660, Springer, 2016, pp. 140–150.

[26] I. Habernal, and I. Gurevych, "What makes a convincing argument? Empirical analysis and detecting attributes of convincingness in Web argumentation", in the Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing (EMNLP2016), Austin, Texas, November 1-5, 2016, pp 1214–1223.