ICACSIS 2014

International Conference on
Advanced Computer Science and Information System 2012
(ICACSIS 2014)
Hotel Ambhara, Jakarta
October 18th - 19th, 2014

View

Please enable **Javascript** on your browser to view all the page properly.

Contacts

ICACSIS Committee
Email: icacsis@cs.ui.ac.id
Phone: +62 21 786 3419 ext. 3225

Faculty of Computer Science, Universitas Indonesia
Kampus UI Depok
Indonesia - 16424
Phone: +62 21 786 3419
Fax: +62 21 786 3415
Email: humas@cs.ui.ac.id

International Conference on
Advanced Computer Science and Information System 2012
(ICACSIS 2014)
Hotel Ambhara, Jakarta
October 18th - 19th, 2014

Committees | Table of Contents | Author's Index | About This CD-ROM

Search

Committee

Honorary Chairs
A. Jain, Fellow IEEE, Michigan State University, US
T. Fukuda, Fellow IEEE, Nagoya-Meijo University, JP
M. Adriani, Universitas Indonesia, ID

General Chairs
E. K. Budiardjo, Universitas Indonesia, ID
D.I. Sensuse, Universitas Indonesia, ID
Z.A. Hasibuan, Universitas Indonesia, ID

Program Chairs
H.B. Santoso, Universitas Indonesia, ID
W. Jatmiko, Universitas Indonesia, ID
A. Buono, Institut Pertanian Bogor, ID
D.E. Herwindiati, Universitas Tarumanagara, ID

Section Chairs
K. Wastuwibowo, IEEE Indonesia Section, ID

Publication Chairs
A. Wibisono, Universitas Indonesia, ID

Program Committees
A. Azurat, Universitas Indonesia, ID
A. Fanar, Lembaga Ilmu Pengetahuan Indonesia, ID
A. Kistijantoro, Institut Teknologi Bandung, ID
A. Purwarianti, Institut Teknologi Bandung, ID
A. Nugroho, PTIK BPPT, ID
A. Srivihok, Kasetsart University, TH
A. Arifin, Institut Teknologi Sepuluh Nopember, ID
A.M. Arymurthy, Universitas Indonesia, ID
A.N. Hidayanto, Universitas Indonesia, ID
B. Wijaya, Universitas Indonesia, ID
B. Yuwono, Universitas Indonesia, ID

B. Hardian, Universitas Indonesia, ID
B. Purwandari, Universitas Indonesia, ID
B.A. Nazief, Universitas Indonesia, ID
B.H. Widjaja, Universitas Indonesia, ID
Denny, Universitas Indonesia, ID
D. Jana, Computer Society of India, IN
E. Gaura, Coventry University, UK
E. Seo, Sungkyunkwan University, KR
F. Gaol, IEEE Indonesia Section, ID
H. Manurung, Universitas Indonesia, ID
H. Suhartanto, Universitas Indonesia, ID
H. Sukoco, Institut Pertanian Bogor, ID
I. Budi, Universitas Indonesia, ID
I. Sitanggang, Institut Pertanian Bogor, ID
I. Wasito, Universitas Indonesia, ID
K. Sekiyama, Nagoya University, JP
L. Stefanus, Universitas Indonesia, ID
Marimin, Institut Pertanian Bogor, ID
M.T. Suarez, De La Salle University, PH
M. Fanany, Universitas Indonesia, ID
M. Kyas, Freie Universitat Berlin, DE
M. Nakajima, Nagoya University, JP
M. Widyanto, Universitas Indonesia, ID
M. Widjaja, PTIK BPPT, ID
N. Maulidevi, Institut Teknologi Bandung, ID
O. Sidek, Universiti Sains Malaysia, MY
O. Lawanto, Utah State University, US
P. Hitzler, Wright State University, US
P. Mursanto, Universitas Indonesia, ID
S. Bressan, National University of Singapore, SG
S. Kuswadi, Institut Teknologi Sepuluh Nopember, ID
S. Nomura, Nagaoka University of Technology, JP
S. Yazid, Universitas Indonesia, ID
T. Basaruddin, Universitas Indonesia, ID
T. Hardjono, Massachusetts Institute of Technology, US
T. Gunawan, Int. Islamic University Malaysia, MY
T.A. Masoem, Universitas Indonesia, ID
V. Allan, Utah State University, US
W. Chutimaskul, King Mokut's Univ. of Technology, TH
W. Molnar, Public Research Center Henri Tudor, LU
W. Nugroho, Universitas Indonesia, ID
W. Prasetya, Universiteit Utrecht, NL
W. Sediono, Int. Islamic University Malaysia, MY
W. Susilo, University of Wollongong, AU
W. Wibowo, Universitas Indonesia, ID
X. Li, The University of Queensland, AU
Y. Isal, Universitas Indonesia, ID
Y. Sucahyo, Universitas Indonesia, ID

ICACSIS 2014

International Conference on
Advanced Computer Science and Information System 2012
(ICACSIS 2014)
Hotel Ambhara, Jakarta
October 18th - 19th, 2014

Hotspot Clustering Using DBSCAN Algorithm and Shiny Web Framework

Karlina Khiyarin Nisa
Page(s): 144-147
Abstract | Full Text: PDF

Framework Model of Sustainable Supply Chain Risk for Dairy Agroindustry Based on Knowledge Base

Winnie Septiani
Page(s): 148-154
Abstract | Full Text: PDF

View: 1-25 | 26-50 | 51-75

ICACSIS 2014

International Conference on
Advanced Computer Science and Information System 2012
(ICACSIS 2014)
Hotel Ambhara, Jakarta
October 18th - 19th, 2014

A

B

C

D

E

F

G

H

I

J

K

L

M

N

P

R

S

T

V

W

Y

Z

ICACSIS 2014

International Conference on
Advanced Computer Science and Information System 2012
(ICACSIS 2014)
Hotel Ambhara, Jakarta
October 18th - 19th, 2014

View

Please enable **Javascript** on your browser to view all the page properly.

Contacts

ICACSIS Committee
Email: icacsis@cs.ui.ac.id
Phone: +62 21 786 3419 ext. 3225

Faculty of Computer Science, Universitas Indonesia
Kampus UI Depok
Indonesia - 16424
Phone: +62 21 786 3419
Fax: +62 21 786 3415
Email: humas@cs.ui.ac.id

IEEE INDONESIA SECTION

ICACSIS 2014

2014 International Conference on

Advanced Computer Science and Information Systems

(Proceedings)

Ambhara Hotel, Jakarta

October 18th-19th , 2014

Published by:

Faculty of Computer Science

Universitas Indonesia

Tourism Recommendation Based on Vector Space Model Using Composite Social Media Extraction

Husnul Khotimah, Taufik Djatna, and Yani Nurhadryani
Graduate School of Computer Science, Bogor Agricultural University
Email: husnulkhotimah@apps.ipb.ac.id, taufikdjatna@ipb.ac.id, yhadryani@ipb.ac.id

***Abstract*— Intentionally or not, social media users likely to share others recommendation about things, included tourism activities. In this paper we proposed a technique which was able to structure the joint recommendation of composite social media and extract them into knowledge about the tourist sites by deploying the vector space model. We included advice seeking technique to not only calculate recommendations obtained from the profile itself but also recommendations by social network users. This is a potential solution to handle sparsity problem that usually appears in conventional recommender systems. We further formulated an approach to normalize the unstructured text data of social media to obtain appropriate recommendation. We experimented the real world data from various source of social media in R language. We evaluated our result with Spearman's rank correlation and showed that our formulation has diversity recommendation with positive correlation to user's profile.**

Keywords--social media, vector space model, composite extraction

I. INTRODUCTION

Currently social media is one of tools for people to search for travel destination and information. The needs of online information for traveler, give an opportunity for tourism organization to enlarge their promotion through social media [1]. Social media has changed the way of traveler to planning their trip by combined information from various social media. For instance, we like to search comments or testimonies by others people in social media about any tourism site. Users assess and match it towards their preferences. They repeat these activities manually until they found the best tourism site that match with their preferences.

There were many tourism recommender systems (RSs) to support the selection process of tourism destination easily. In [2], [3] tourism RSs developed under static data which it represent characteristic of tourism site. In real world, the tourism characteristics does not compatible with static data which its value in crisp value (0 or 1). For instance in [3] was defined that travel goal is one of attribute in tourism context, it was divided to 9 value such as cultural experience, scenic/landscape, and education. One tourism site may have more than one value, and if one tourism site has one category such as cultural experience it doesn't mean its value to cultural experience is one and to other categories is zero. This method is not suitable with current user's behavior as we had mentioned in previous paragraph.

In this paper, we proposed a method for tourism RSs through social media extraction so it will not depend to static data that have high value of inflexibility. Information about tourism site was provided in several social media and dynamically grew as people share their experience through their comment, testimony, post, etc.

Our motivation was to overcome the drawback of static identifier for tourism site in RSs through text data in social media. There are some challenges when we need to utilize text data in social media. Data in social media is very unstructured, user are free to post with emoticon, word abbreviation, link, or any non-standardized text data. To clear these challenges, we presented a text mining method with some additional normalization process to obtain well-formed identifier of tourism site.

In this paper we did not use historical data about users and items as conventional RSs did. Our method provided recommendations based on user's posting in social media such as sparsity problem which potentially appear if users rarely make a post in their social media. This problem has potential to cause sparsity problem which in conventional RSs affected by sparse data rating [4], [5]. In order to avoid this problem, we utilized the power of the social network usage data that support data-network between users, such as in Facebook which are known as friends or in Twitter known as follower. Our assumption is the items that might be preferred by user's friend would directly influence user's choice. This assumption also known as advice-seeking [6]. We utilize users connection in social network to complete the process of this assumption.

We thus formulated the mechanism of RSs from the extraction process in user's social media and tourism

site's social media through text mining processes to be identifier for each object. Every tourism site has the set of term as their identifier. The occurrences of tourism site's identifier projected in vector space model with vector defined by the collection of term. User's identifier filtered based on tourism corpus and laid on tourism vector space model to calculate the proximity between user and each tourism site.

The objective of this paper was to formulate recommendation model based on composite social media and to extract unstructured composite social media text data in Indonesia language. The rest of this paper is organized as follows. In section II we briefly mention state of the art RSs based on social media. In section III we briefly present illustration how we formulate recommendation. In section IV we presented our proposed method and we conducted relevant experiment for our proposed method with real world data in section V. The conclusions are drawn in section VI.

II. Recommender Systems Based on Social Media

In social media, there are many aspects to provide recommendation process. Social media become people's need and influenced the data to keep on growing naturally and potentially to be resources for RSs. This power potentially provides RSs performance to be more dynamic but faced more challenges.

There are some studies to explore the benefit of social media data in RSs. In [7] presented two approaches for recommendation framework based on social media, there are interest-oriented and influenced-oriented which focused for content recommendation in social media. In [8] presented field experiment based on interview to demonstrate the benefit of social recommendation, trust-aware recommendation, and advice-seeking recommendation to improve the performance of RSs as it has similar mechanism with real world recommendation.

In practice, one person may have more than one social media, [9] presented recommendation with composite social media to acquire friend list and analyze friends who have impact in user decision making to generate personalized recommendation. The drawback of the conventional RSs in [9] still use static identifier that given by user manually. Most of research in object identifier extraction such as sentiment analysis and opinion mining are based on text mining method [10], [11], [12].

Based on this review, we caught a gap in social media data extraction for object identifier to infer in

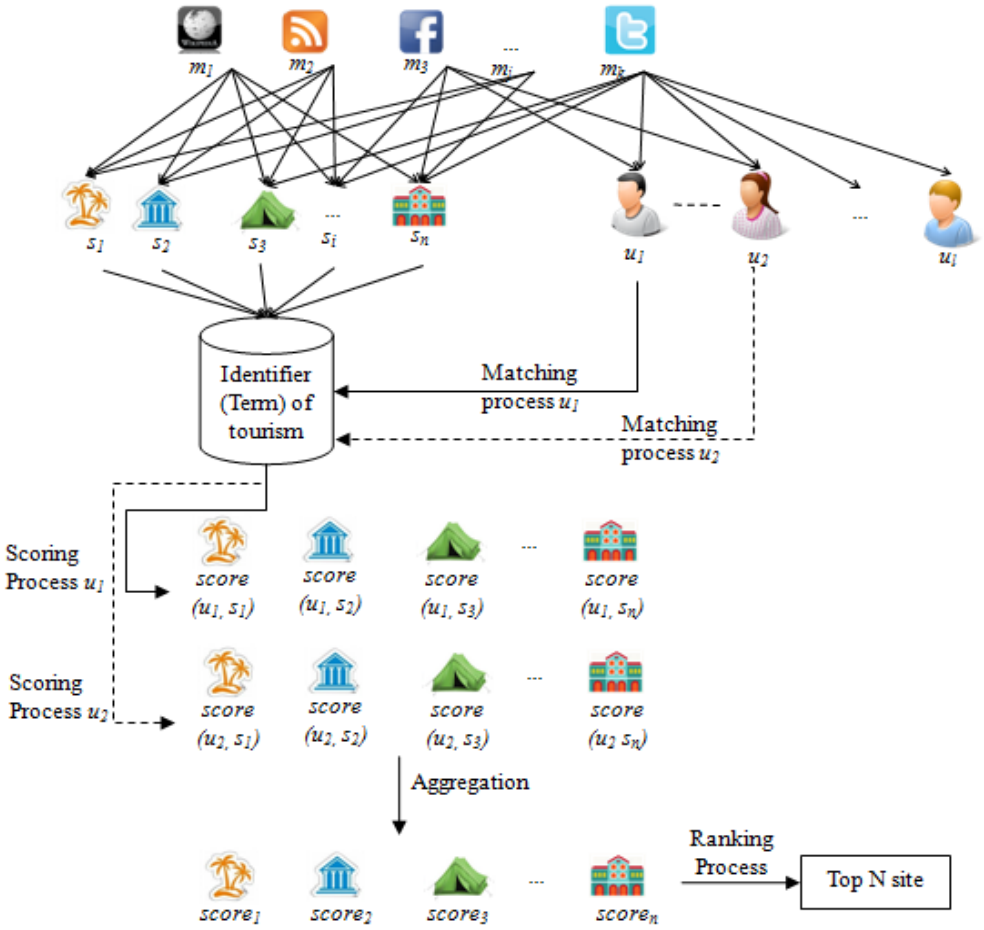

Fig. 1. Illustration of composite social media extraction and social recommender

process of recommendation. We further complete the recommendation by using relevant content from composite social media to enrich our findings.

III. Formulation of Recommendation based on Social Media Data

We briefly present our proposed method with illustration in Fig. 1. The information from various social media identified as characteristic or identifier for each tourism site. For each site, we extract information from various social media with text mining to emerge its characteristics. Then, the extraction result stored in database. The collection of extraction process from various tourism sites defined as tourism corpus.

In this research, we combined recommendation process with advice-seeking technique related to social recommender. We assumed the items that might be liked by other users who have strong relation with user also contribute to user's choice. Then, we identified user in the system who have strong relation with users who will receive recommendation, in this research we call those users as socialize users. We also projected socialize users into vector space model to calculate each user proximity with tourism site. Then, score for each tourism site aggregated based on level of trust (λ). For the last step, we rank tourism site based on score of aggregation function to generate top N recommendation.

IV. Proposed Approach

A. Tourism identifier Extraction

Nowadays information of tourism were supported by various social media. The collections of social media for each tourism site assumed as a corpus. In our approach, we defined identifier for each site by term occurrence in social media feature that based on text media.

Object identifiers for each site were generated from its corpus in Indonesia language with text mining method, there are tokenization, normalization, term compression, and term weighting. We normalized unstructured data on social media by following these step:

Step 1: remove punctuation and numbers.

Step 2: normalize based on words abbreviation. This is to to solve another challenges of data text in social media. When posting in social media, people like to use abbreviation of words. For instance, 'sepeda' (in English: bicycle), we can use abbreviation 'spd'. In this normalization process, we using abbreviation list in Indonesia Language from [13].

Step 3: stem all word based on Indonesia language [14].

Step 4: transform unstructured form of words using regular expression. Some challenges in social media text are people freely to post unstructured words like 'gunuuung' (in English: mountain) according to 'gunung'. Regular expression will transform the repeatedly letters into single letters. This is the formulation of regular expression transformation:

$$\begin{aligned} [a]^+ &\rightarrow [a] \\ [b]^+ &\rightarrow [b] \\ &\vdots \\ [z]^+ &\rightarrow [z] \end{aligned}$$

where $[a]^+$ means any words that consist repeatedly 'a' characters (more than one 'a' characters in a string) will be transformed into single 'a' characters.

Step 5: re-normalization process. In this normalization process, we use re-normalized words have been changed due to process in 1d, such as 'tanggal' (in English: date) have been changed to 'tangal' and we must re normalize to early form.

In order to avoid the term that generated abundant, we use term compression based on compression rate. The collection of identifier from all site were defined as tourism term. For each term will be calculated term frequency (*tf*) and then normalized sublinear *tf* scaling. We assign $tf_{t,s}$ depends on the number of occurrences of term *t* in site *s*. We use normalized sublinear *tf* scaling in [15] as follow :

$$wf_{t,s} = \begin{cases} 1+\log tf_{t,s} & \text{if } tf_{t,s} > 0 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

B. User Profiling

In this paper, the occurrence of tourism terms in user's social media content guide us to generate tourism site recommendation. First of all, user might have more than one account in various social media. The collection of user's social media can be seen as corpus for each user. Then, to get tourism term in user's corpus, the corpus were then will be proceed by text mining processes which are similar to the previous section of tourism identifier extraction.

C. Vector Space Model

The calculation of proximity between each tourism site user based on cosine similarity function in vector space model:

$$score(u_i, s_j) = \frac{\vec{v}(u_i) \cdot \vec{v}(s_j)}{\left|\vec{v}(u_i)\right|\left|\vec{v}(s_j)\right|} \quad (2)$$

Otherwise, in social media there is linkage between users that define connection between users. We identified advice-seeking process by the connection between users. We assumed if $user_1$ and $user_2$ were friend with each other in social media, $user_1$ will have contribution to influence $user_2$ recommendation, and vice versa. If we want to give recommendation to $user_1$ as main user, we must identified list of $user_1$'s friend, for example in this case we detect $user_2$ as $user_1$'s friend. We define notation for (f_1, $f_{2,\ldots\ldots}$ f_z) as the collection of main user's friends in our formulation.

We aggregated score from main user and main user's friends based on level of trust in range $0 \leq \lambda \leq 1$, the value of λ represent a weight of how we trust recommendation from our friends than we trust recommendation from our own profile. If we set the bigger value of λ , then we trust recommendation from our friend more than our own profile. We formulated equation that derived from function of weighted mean aggregation in [16] for each site (s_1, s_2,, s_j,...s_n) from this equation :

$$finalScore_j = (1-\lambda)\, score(u, s_j) + (\lambda)\frac{\sum_{i=1}^{z} score(f_i, s_j)}{z} \quad (3)$$

Then, we ranked tourism site based on the *final score* and filtered based on top-N.

We evaluated our formulation based on Spearman's rank correllation coefficient in [17] :

$$\rho = \frac{\sum (u-\bar{u})(v-\bar{v})}{n * stdev(u)\, stdev(v)} \quad (4)$$

The objective of our evaluation was to identify the effect of our assumption that the items might be liked by user's friend will influence user's choice. In Spearman's rank correlation coefficient, we compare ranking of recommendation between user's profile (without aggregation recommendation to user's friends) and with user's friend recommendation.

V. Experimental Result

In this section, we perform the result from our experiment with real world data and hypothetic data. We using R 3.1.0 software to retrieve data from social media and assist text mining process.

A. Experimental data

In this paper, we collect data about tourism site in Table 1 from various social media and compose the data based on tourism site. Then each tourism site take a role as documents and built a tourism corpus.

We retrieve data from 4 different source for each tourism site, there are Twitter search (tourism site's name as a query), Facebook page of tourism site (if any), Twitter account (if any), and Wikipedia webpage.

The data retrieval was assisted by R packages, there are RFacebook and twitter, and for Wikipedia data source firstly we saved html file and converted HTML to text by XML package. The usage of twitteR and RFacebook are we must get access token from API registration at https://developers.facebook.com/ and https://dev.twitter.com/.

B. Data Extraction

Firstly each source of data were processed independently, as data from facebook and twitter almost contain unstructured form but data from Wikipedia contain full structure form. Twitter data might be contain name of user, for example in Twitter post: '*RT @poo Taman Safari belajar keanekaragaman fauna*', we could detect '*@poo*' as name of user as '@' mark was at beginning of user name in twitter. In normalization process for Twitter data, we removed word with its formulation. In unstructured form of Facebook and Twitter data, we performed text mining with normalization process for abbreviation word. This process matched the words with dictionary of abbreviation and then replaced with word in the normal form. In [13] had been experimented with function of Levenshtein distance [18] to normalize abbreviation in Indonesian language, and the result showed matching process using dictionary was more accurate. In this experiment, we used dictionary from [13].

The challenges for text mining in Indonesian language is process of stemming. Porter algorithm [19] and Nazief & Adriani Algorithm [14] are two popular algorithm for stemming corpus in Indonesian language. The comparison of these two algorithm [20] showed Nazief & Adriani Algorithm was more accurate than Porter algorithm, although Porter algorithm faster than Nazief & Adriani algorithm. Then we implemented Nazief & Adriani algorithm in R environment with MySQL database to store word base of Indonesian language.

TABLE I
List of Tourism Site

Tourism Site (Tourism index)	Data Resources			
	Wikipedia	Query of Twitter Search	Facebook Account	Twitter Account
Bogor Botanical Garden (s_1)	√	kebun raya bogor	-	@kebunrayabogor
Safari Garden, Cisarua (s_2)	√	taman safari, cisarua	Taman Safari	@TSI_Bogor
Taman Mekarsari (s_3)	√	taman mekarsari	-	@TamanMekarsari
Kebun raya cibodas (s_4)	√	kebun raya cibodas	Kebun Raya Cibodas (KRC)	@KRCibodas19
Museum Fatahillah (s_5)	√	museum fatahillah	-	@Fatahillah_MSJ
Trans Studio Bandung (s_6)	√	transstudio bandung	-	@TransStudioBdng
Sea World (s_7)	√	sea world	-	@SEAWORLDANCOL
Monumen Nasional (s_8)	√	monumen nasional	Monumen Nasional - Monas	@Tugu_Monas
Taman Mini Indonesia Indah (s_9)	√	taman mini indonesia indah	-	@ilovetamanmini
Taman Impian Jaya Ancol (s_{10})	√	taman impian jaya ancol	-	@ancoltmnimpian

TABLE 2
EXAMPLE OF TERM-DOCUMENT MATRIX

Tourism Site	Term Frequency						
	air	alam	anak-anak	baca	buah	.....	wisata
Bogor Botanical Garden	1	2	0	1	0	.....	8
Safari Garden, Cisarua	1	5	0	2	3	.....	9
Taman Mekarsari	2	1	1	2	96	.....	34
Kebun raya cibodas	6	2	1	1	1	.....	39
Museum Fatahillah	8	1	2	1	10	.....	6
Trans Studio Bandung	0	3	1	1	1	.....	6
Sea World	12	6	4	1	2	.....	1
Monumen Nasional	3	0	0	1	5	.....	2
Taman Mini Indonesia Indah	6	2	4	1	1	.....	9
Taman Impian Jaya Ancol	4	4	1	1	7	.....	17

Stop word list in information retrieval depend on the context of its field. For instances, in the field of computer we adjust 'swim' as stop word, but not in the field of tourism. In our experiment, we just use 126 stop word contain conjunction such as 'yang', 'ke', 'pada'. In Fig. 2 we can see there are unrepresentative words for tourism context. We obtained 5018 term which contain 81% sparse term. Then, we reduce tourism term with value of compression rate 40% which means we filtered out term that were not appear in minimal 4 documents. Term compression reduce tourism term to 229 terms with 12% term. Table 2 is an example list for term document matrix.

C. *Recommendation Process*

We perform data acquisition from two users who have Twitter and Facebook account and they are friends in real world and in social network. $User_1$ is main user to given the recommendation and $user_2$ is friend to user1. In this process, we matched the occurrences of tourism term in user's social media post. Table 3 show term occurrences for $user_1$ and $user_2$ after user profiling process. $User_1$ like to travel to natural site and $user_2$ like to travel to historical site.

Then, we normalized term frequency with equation (1) then calculated cosine similarity for $user_1$ and $user_2$. Fig. 3 show illustration how vector space model projection for users and tourism site, tourism term take role as vector where cosine similarity defined the proximity between user and tourism site.

Fig. 2. Wordcloud for Wikipedia source data about Bogor Botanical Garden

In this experiment we used level of trust with λ=0.4 which show score for recommendation based on friend's profile. Table 4 is the result of our experiments, final score calculated by equation (3) and show the aggregated value of cosine similarity between $user_1$ and $user_2$. The result show how social recommender impact the result of recommendation for historical site by $user_1$. The top-3 recommendation are Sea world (s_7), Safari Garden (s_2), Monumen Nasional (s_8).

Then, we evaluated our result based on Spearman's rank correlation. The evaluation showed score of 0.78 which compared with ranking of recommendation based on user's social media content and our formulation (combine with friend's social media content). It showed that our recommendation has positive correlation with user's profile. Furthermore, our formulation has more diversity recommendation. If we generate a recommendation based on $user_1$ profile, the top-5 recommendation (s_7, s_2, s_4, s_1, s_3)

TABLE 3
TERM-DOCUMENT OF USER MATRIX

Tourism Term	$User_1$ (u_1)	$User_2$ (u_2)
air	2	7
alam	1	3
anak-anak	6	0
...	...	...
wisata	3	3

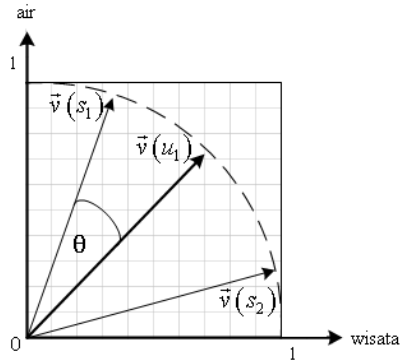

Fig. 3. Ilustration of proximity between user and tourism site in vector space model based on cosine similarity, $sim(s_1,u_1)=\cos\theta$

TABLE 4
CALCULATION FOR RECOMMENDATION SCORE

Tourism index (s_j)	Score (u_1,s_j)	Score (u_2,s_j)	Final Score	Rank based on u_1	Rank based on Final Score
s_1	0.8	0.79	0.80	4	4
s_2	0.86	0.75	0.82	2	2
s_3	0.78	0.7	0.75	5	6
s_4	0.82	0.73	0.78	3	7
s_5	0.67	0.84	0.74	9	8
s_6	0.76	0.83	0.79	6	5
s_7	0.87	0.79	0.84	1	1
s_8	0.73	0.92	0.81	7	3
s_9	0.71	0.77	0.73	8	9
s_{10}	0.65	0.8	0.71	10	10

related to natural tourism sites, but based on our formulation list of top-5 recommendation are more diverse (contain historical tourism sites), and user will have more experience with our recommendation. If user want obtain recommendation only based on their profile, we can adaptively set λ value related to user's needs on diversity recommendation. Thus briefly we have proved the power of composite social media content extraction to alleviate the tourism recommendations. However in further work to make our formulation more powerfull, we must construct integrated API from various social media and implement our formulation within real time process.

VI. CONCLUSION

We have proposed an approach for tourism recommendation based on composite social media data. This approach utilized the growth of tourism's data in social media to be an identifier which it represent tourism characteristic. Our recommendation adopted real world recommendation that our friend contribute in decision making process that provided by user's network in social media. The benefit of this adoption was to overcome sparsity problem that happened in conventional RSs. We solved the challenges to obtain well-formed identifier from unstructured text data in social media based on some normalization process in text mining method. The experiment with real world data show our formulation can adaptively implemented related to user's needs on diversity recommendation. Our approach potentially implemented in real time process to generate dynamic identifier of tourism sites.

REFERENCES

[1] Z. Xiang, and U. Gretzel, "Role of social media in online travel information search", *Tourism management*, vol 31, no. 2, pp. 179-188, 2010.

[2] N. Rosmawarni, T. Djatna, Y. Nurhadryani, "A mobile ecotourism recommendations systems using Cars-Context Aware Approaches", *Telkomnika*, vol 11, no 4, pp. 845-852, 2013.

[3] L. Baltrunas, B. Ludwig, S. Peer, and F. Ricci. "Context relevance assessment and exploitation in mobile recommender systems", *Personal and Ubiquitous Computing*, vol 16, no 5, pp 507-526, 2011.

[4] C. Desrosiers, and G. Karypis, "A comprehensive survey of neighborhood-based recommendation methods," in *Recommender Systems Handbook*, J. F. Ricci, L. Rokach, B. Shapira, and P.B. Kantor, Ed. New York: Springer, 2013, pp. 107–144.

[5] P. Symeonidis, D. Ntempos, and Y. Manolopoulos, *Recommender Systems for Location-based Social Network*, New York, Springer, 2014, pp. 7-20.

[6] P. Bonhard and M. Sasse, " 'Knowing me, knowing you' - using profiles and social networking to improve recommender systems", *BT Technology Journal*, vol. 24, no. 3, pp. 84-98, 2006.

[7] H. Ma, D. Zhou, C. Liu, M.R. Lyu, , and I.. King, "Recommender systems with social regularization", in *Proc 4th ACM Int. Conf. Web Search and Data Mining*, 2011, pp. 287-296.

[8] J. He and W. W. Chu, A social network-based recommender system (SNRS). Springer, 2010.

[9] J. Zhang, Y. Wang, and J. Vassileva, "Socconnect: A personalized social network aggregator and recommender", *Information Processing.& Management*, vol. 49, no. 3, pp. 721-737, 2013.

[10] J. Akaichi, Z. Dhouioui, and M. J. Lopez-Huertas Perez, "Text mining facebook status updates for sentiment classification", in System Theory, Control and Computing (ICSTCC), 2013 17th International Conference. IEEE, 2013, pp. 640–645.

[11] E. Martínez-Cámara, M. T. Martín-Valdivia, L. A. Ureñalópez, and A. R. Montejo-Ráez, "Sentiment analysis in twitter," Natural Language Engineering, vol. 20, no. 01, pp. 1–28, 2014.

[12] S. Stieglitz and L. Dang-Xuan, "Social media and political communication: a social media analytics framework," Social Network Analysis and Mining, vol. 3, no. 4, pp. 1277–1291, 2013.

[13] A. T. A. Aziz, "Sistem Pengklasifikasian Entitas Pada Pesan Twitter Menggunakan Ekspres Regular dan Naïve Bayes", Bachelor thesis, Dept. Computer Science, Bogor Agricultural University, Bogor, ID, 2013.

[14] M. Adriani, J. Asian, B.Nazief, S. M. Tahaghoghi, and H. E. Williams, "*Stemming Indonesian: A confix-stripping approach*", *ACM Transactions on Asian Language Information Processing (TALIP)*, vol. 6, no. 4, pp.1-33, 2007.

[15] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge: Cambridge university press, 2008, pp. 109-134.

[16] J. Marichal, "Aggregation functions for decision making," in *Decision Making Process :Concepts and Methods*, D. Bouyssou, D. Dubois, H. Prade, and M. Pirlot, Ed. Wiley, 2009, pp 673-722.

[17] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, "*Evaluating collaborative filtering recommender systems*," *ACM Transactions on Information Systems*, vol 22, no. 1, pp. 5-53, 2004.

[18] A.T. Freeman, L. C. Sherri, and M. A. Christopher, "Cross linguistic name matching in English and Arabic: a one to many mapping extension of the Levenshtein edit distance algorithm," in *Proceedings of the main conference on Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics*. Association for Computational Linguistics, 2006.

[19] Z.T. Fadhillah, A Study of Stemming Effect on Information Retrieval in Bahasa Indonesia,

[20] L. Agusta, Perbandingan algoritma stemming Porter dengan algoritma Nazief & Adriani untuk stemming dokumen teks bahasa indonesia, in *Konferensi Nasional Sistem dan Informatika*, Bali, 2009, pp. 196-201.