Computing and Informatics, Vol. 33, 2014, 591–608

SOCIAL DATA VISUALIZATION SYSTEM FOR UNDERSTANDING DIFFUSION PATTERNS ON TWITTER: A CASE STUDY ON KOREAN ENTERPRISES

Dosam Hwang

Department of Computer Engineering, Yeungnam University
Gyeongsan, Korea 712-749
e-mail: `dshwang@yu.ac.kr`

Jai E. Jung*

Department of Computer Engineering, Chung-Ang University
Seoul, Korea
&
Department of Information System, Universiti Malaya
Kuala Lumpur, Malaysia
e-mail: `j2jung@gmail.com`

Seungbo Park

Institute of Media Content, Dankook University
Yongin, Korea, 448-701
e-mail: `molaal@naver.com`

Hien T. Nguyen

Faculty of Information Technology, Ton Duc Thang University
Ho Chi Minh City, Vietnam
e-mail: `hien@tdt.edu.vn`

Abstract. Online social media have been playing an important role of creating and diffusing information to many users. It means the users can get cognitive influence to the other users. Thus, it is important to understand how the information can be diffused by interactions among users through online social media. In this paper, we design a social media monitoring system (called "TweetPulse") which can analyze and show meaningful diffusion patterns (DP) among the users. Particularly, TweetPulse focuses on visualizing information diffusion in Twitter, given a certain time duration. Also, this work has investigated the relationships 1) between DP and event detecting, 2) between DP and emotional words, and 3) between DP and the number of followers of the users. Thereby, to understand the continuous patterns of the information diffusion, we propose two different types of analytic methods, which are 1) macroscopic approach and 2) microscopic approach. For evaluating the proposed method, we have collected and preprocessed the dataset during about 4 months (14 March 2012 to 12 July 2012). As a conclusion, TweetPulse has helped users to easily understand DP from a large scale dataset streaming through Twitter.

Keywords: Information visualization, diffusion patterns, marketing strategies, Twitter

1 INTRODUCTION

Online social media (e.g., MySpace, Twitter, and Facebook) have been regarded as an important channel where information can be efficiently diffused to other users. This information diffusion can be accelerated (in opposite, decelerated), according to a number of situations. For example, the information about tsunami and earthquakes has been significantly distributed to the people in the certain areas. Thus, information diffusion is an important social phenomenon that we have to recognize and understand.

These social networking systems have supported collaborative tagging service and RSS-based technologies for users, and also they have become the factors which give important influences on diffusing the information through online social system [1, 2, 20]. For example, in Twitter, retweet (in short, RT) function has the effect of another information diffusion which is related to own opinion; also, hash tagging can attract more attention for creating and pulling attention from others [3, 4].

In fact, we have already received a lot of information from traditional media, e.g., television or radio. However, with the advent of new communication technology and internet development, users are able to collect new information more easily. Unlike the existing one-sided mass media, the current social media can support users to generate information and share with each other more widely [5, 21, 22].

[*] Corresponding author

Users can communicate with friends via online social network services like Twitter and Facebook. Moreover, they can exchange various opinions with strangers, and receive news from celebrity [6]. It is no longer a surprising social process that social network services bring news quicker than the traditional media. Actually, the subway in Seoul had been stopped on the rush time because of accident. The people who were in the subway uploaded the news on SNS and these tweets were RT to other users; this prevented rush time delays. There are ever more people who are contacting on SNS with the commuting time.

1.1 Backgrounds on Online Social Networking Services

Social networking service (SNS) is an efficient tool for sharing tremendous amount of information and disseminating information quickly [7]. In particular, Twitter can only represent one-side social relationships (called following and follower), while other online social networking services can represent two-side social relationships. In order to establish online relationships, it does not require the consent of the opponents. It means that social relationships in Twitter can be formed more easily than in other SNS. Next, Twitter supplies Twitter API to access a variety of information, and developers can easily implement third party applications by collecting from Twitter. Also, Twitter can supply re-transmission feature of the tweets, called retweet (RT). If a user creates a tweet, it will be delivered to the followers of the user, and the followers who received tweet can also retweet the same tweets to their followers for various purposes. In this case, the tweet can deliver the information like word-of-mouth process (from one user to his/her followers, and again, from his/her followers to their followers). Because there is no limit of the number of retweeting, more and more people can get the tweet. In terms of usability, it is convenient to retweet to their followers by clicking a button one time for rapid information diffusion on Twitter [8, 9, 10]. Thus, in this study, we study online information diffusion on Twitter by finding the pattern of the spread of information from Twitter and also the causes of this pattern. For this purpose, we have developed a software system (called TweetPulse) to show the relationships of diffusion of information with emotional language on Twitter.

1.2 Outline of This Paper

The outline of this paper is as follows. In Section 2, we introduce the related work of SNS diffusion and trend, and describe the information diffusion on the SNS. Section 3 discusses TweetPulse and its diffusion speed and diffusion convergence rate. Section 4 addresses the detecting events, and how long is the diffusion continued. In Section 5, we present the simulation to show the impact of events in the information diffusion and what information is propagated speedily. Finally, the conclusion of this work is given in Section 6.

2 RELATED WORK

2.1 Information Diffusion and Trend Analysis with SNS

SNS is a service to support users to build a human network on the web. As society advances, it is urged to express personal information and feelings grows stronger. Thus, the way of building social relationships changed to make friends on the Internet and websites like SNS is able to develop. Now more and more users use SNS to support new information. Because the SNS's information speed is very high, concise, and accurate, it is very advantageous for people in modern society [11, 12]. Unlike the traditional media's one-sided nature the SNS is based on a variety of digital media and thus it enables the contents, consumption and production by the user; thus, it opens a new chapter of society and culture. SNS can provide the opportunity for the formation of new connections through a variety of ways to communicate, various discussions on social issues on SNS, and leads the citizens to a diversity of social participation. SNS has characteristics such as speed, personality, information openness and easiness and can be classified into eight types depending on the capability of providing, such as profile-based, business-based, blog-based, vertical, collaborative, communication-centric, topic based, and micro blogging [13, 18, 23].

2.2 Information Diffusion on the SNS

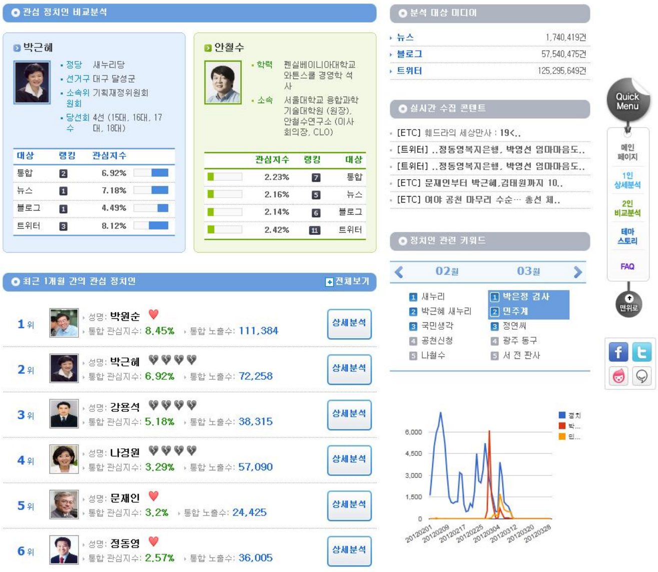
Just a few years ago, we find the information necessary for academic or business on portal site. However, we are able to quickly and easily share various kinds of information created in real time through SNS with several members who form communities [16, 14]. With SNS explosive diffusion, information sharing activity among SNS users is increasing [15]. With the development of the SNS, the website shows changes of information diffusion on the SNS. Especially it has become popular, as shown in Figure 1 which shows politician interest index, the preference, keyword changes like TrueStory.

Also, the Pulse-K is a social media monitoring and analysis service developed by Conan technology. As shown in Figure 2, Pulse-K[1] provides media and emotional on the strip real-time which keywords to monitor, and users are able to see how to change the reaction of the people on those keywords.

Figure 3 shows a social search tool which can provide original search results from Twitter and Blog; it also can check the maximum diffusion message. Social metrics insight[2] provides crisis management and issues detected by services through sensitivity word analysis to help companies. SNS such as Twitter or Blog help also companies to communicate with consumers to promote new products.

[1] http://www.pulsek.com/

[2] http://insight.some.co.kr/



관심 정치인 비교분석

박근혜
- 정당 새누리당
- 선거구 대구 달성군
- 소속위원회 기획재정위원회
- 당선회수 4선 (15대, 16대, 17대, 18대)

대상	랭킹	관심지수
통합	2	6.92%
뉴스	1	7.18%
블로그	1	4.49%
트위터	3	8.12%

안철수
- 학력 펜실베이니아대학교 와튼스쿨 경영학 석사
- 소속 서울대학교 융합과학기술대학원 (원장), 안철수연구소 (이사회의장, CLO)

관심지수	랭킹	대상
2.23%	7	통합
2.16%	5	뉴스
2.14%	6	블로그
2.42%	11	트위터

최근 1개월 간의 관심 정치인 전체보기

1위	성명: 박원순	통합 관심지수: 8.45%	통합 노출수: 111,384
2위	성명: 박근혜	통합 관심지수: 6.92%	통합 노출수: 72,258
3위	성명: 강용석	통합 관심지수: 5.18%	통합 노출수: 38,315
4위	성명: 나경원	통합 관심지수: 3.29%	통합 노출수: 57,090
5위	성명: 문재인	통합 관심지수: 3.2%	통합 노출수: 24,425
6위	성명: 정동영	통합 관심지수: 2.57%	통합 노출수: 36,005

분석 대상 미디어

뉴스	1,740,419건
블로그	57,540,475건
트위터	125,295,649건

실시간 수집 콘텐트
- [ETC] 훼드라의 세상만사 : 19<..
- [트위터] ..정동영복지은행, 박영선 엄마마음도..
- [트위터] ..정동영복지은행, 박영선 엄마마음도..
- [ETC] 문재인부터 박근혜,김태원까지 10..
- [ETC] 여야 공천 마무리 수순… 총선 체..

정치인 관련 키워드

02월	03월
1 새누리	1 박은정 검사
2 박근혜 새누리	2 민주계
3 국민생각	3 정연씨
4 공천신청	4 광주 동구
5 나철수	5 서 전 판사

Figure 1. Snapshot of diffusion of politicians related keywords in True Story

3 MODELING SOCIAL PULSE

The information can be collected from Twitter diffusion. The aggregated information can be visualized on the chart so that users can easily understand the diffusion patterns (increasing or decreasing) on TweetPulse. Once a set of information of each user is projected from Twitter, it can be easily interpreted as a discrete signal which is a sequence of events over time. More importantly, we can set a time

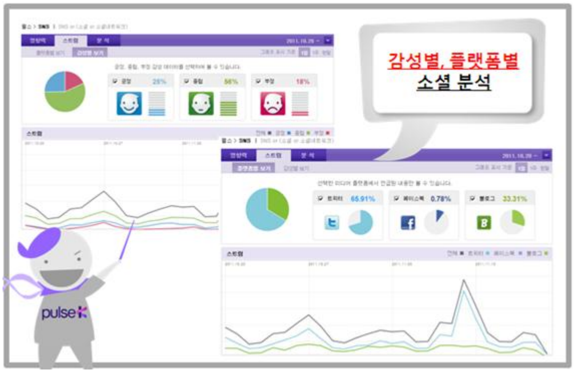
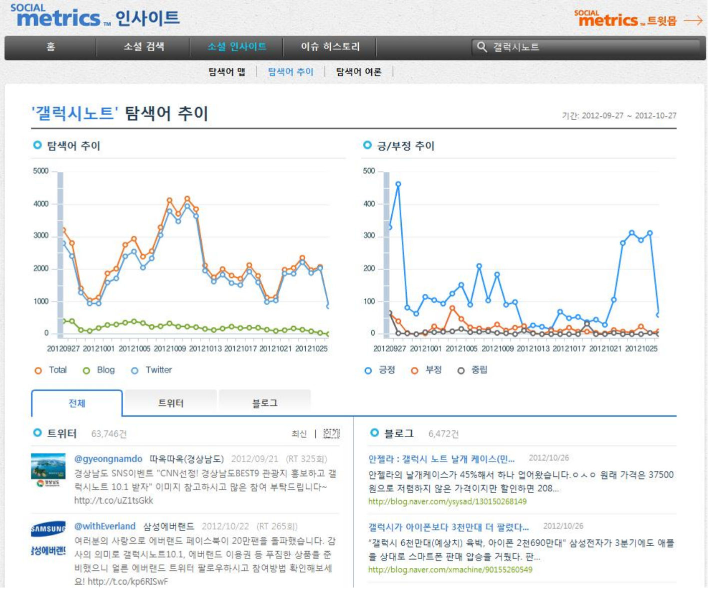


Figure 2. Snapshot of Pulse-K services



@gyeongnamdo 따옥따옥(경상남도) 2012/09/21 (RT 325회)
경상남도 SNS이벤트 "CNN선정! 경상남도BEST9 관광지 홍보하고 갤럭시노트 10.1 받자" 이미지 참고하시고 많은 참여 부탁드립니다~ http://t.co/uZ1tsGkk

@withEverland 삼성에버랜드 2012/10/22 (RT 265회)
여러분의 사랑으로 에버랜드 페이스북이 20만팬을 돌파했습니다. 감사의 의미로 갤럭시노트10.1, 에버랜드 이용권 등 푸짐한 상품을 준비했으니 얼른 에버랜드 트위터 팔로우하시고 참여방법 확인해보세요! http://t.co/kp6RISwF

안젤라 : 갤럭시 노트 날개 케이스(민... 2012/10/26
안젤라의 날개케이스가 45%해서 하나 업어왔습니다.ㅇㅅㅇ 원래 가격은 37500원으로 저렴하지 않은 가격이지만 할인하면 208...
http://blog.naver.com/ysysad/130150268149

갤럭시가 아이폰보다 3천만대 더 팔렸다... 2012/10/26
"갤럭시 6천만대(예상치) 육박, 아이폰 2천690만대" 삼성전자가 3분기에도 애플을 상대로 스마트폰 판매 압승을 거뒀다. 판...
http://blog.naver.com/xmachine/90155260549

Figure 3. Snapshot of social metrics insight

window (of size δ), since we need to understand how the information is diffused in Twitter [17, 19].

Assume that $t \in T$, a TweetPulse is composed of two parts; a time window $w_{\tau_t} = [\tau_t, \tau_t + \delta]$, and a pitch (a height of the pulse). Especially, the pitch can be explained as either the number of users who have applied the same information or the number of information. Thus, it can be formalized by

$$P_t = \{\langle w_{\tau_t}, \pi_{\tau_t} \rangle | \tau_t \in \tau\} \tag{1}$$

where π_{τ_t} is the pitch of a time window and can be computed by

$$\pi_{\tau_t} = \Big| \{ r_k | u_i \times t_j \times r_k \times \tau_{t_j} \in F_\tau, t_j = t, \tau_{t_j} \in [\tau_t, \tau_t + \delta] \} \Big| \tag{2}$$

where $u_i \in U$ and $r_k \in R$. More interesting work is to find relationships between given particular information. For example, if two tags are supported by more resources at the same time, we determine two sets of information have greater correlation than others.

3.1 Diffusion Speed

We can comprehend the speed of information diffusion, i.e., how quickly a tweet is propagated to other users in certain time duration. Since a tweet t has been firstly used at τ_t^0, we can measure the pitch of the TweetPulse π_{τ_t}. Given a tweet t and its TweetPulse P_t, the speed of this diffusion can be measured by

$$S_t = \max_{P_t} \frac{\pi_{\tau_t}}{\delta} \tag{3}$$

where δ indicates a size of the time window.

The speed simply indicates the maximum diffusion power of the corresponding tweet. We can measure the speed, when the TweetPulse of the tag shows the highest pitch. In other words, if we can construct a cumulative curve of the TweetPulse, we can easily find the speed at the steepest slope.

3.2 Convergence Rate of Diffusion

We can also measure the convergence rate of the propagated tweet, i.e., how quickly the tweet is spread to most of users. Given a tag t and its TweetPulse P_t, the convergence rate of its propagation can be measured by a temporal duration

$$C_t = |\tau_t^\Omega - \tau_t^0| \tag{4}$$

where τ_t^Ω is the time ending the TweetPulse (i.e., $\pi_{\tau_t^\Omega} = 0$). It means that after τ_t^Ω, there is no meaningful TweetPulse. Similar to the diffusion speed, convergence rate is also an important indicator for measuring diffusion power.

4 DETECTING AND ANALYZING EVENTS

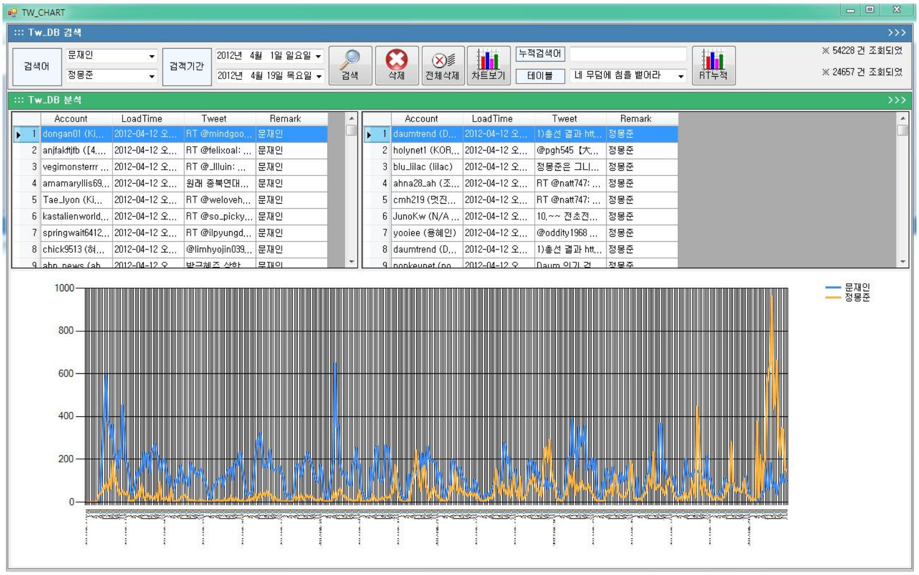
4.1 Event Detection

If at any time a large rate of diffusion is shown on the chart, maybe this is due to any key causes (social issues or events), and we will call it events detecting. Figure 4 is a graph showing the comparison of TweetPulse for change of counts searched by two keywords (Politicians) from 12 April 2012 to 29 April 2012.

We can see in some time-zones that diffusion speed is faster than in other time-zones for each keyword, and we are able to surmise at the time-zones which showed the highest diffusion speed that the time-zones had some social issue or events. On 12 April 2012, relationship with a certain politician "P_1" has social issue with tweets, e.g., "$tweet_1$", and the tweet has influenced diffusion speed of the keyword of P_1. On 29 April 2012, the social event which "$tweet_2$" has influenced the diffusion speed of the keyword of another politician "P_2". As this views a big diffusion speed at the time-zones, it will probably have events detecting about the keyword.

Also, in Figure 5, there is a time-zone when the diffusion speed is maximum on 12 April 2012. It means at that time there were some events relationships with the keyword of 'movie1'.

Thus, Figure 6 shows the snapshot before the time-zone when the tweet was propagated at maximum diffusion speed. An actress appears in movie "M_1" and



TW_CHART

::: Tw_DB 검색

검색어: 문재인 / 정몽준

검객기간: 2012년 4월 1일 일요일 – 2012년 4월 19일 목요일

누적검색어:

테이블: 네 무덤에 침을 뱉어라

※ 54228 건 조회되었

※ 24657 건 조회되었

::: Tw_DB 분석

	Account	LoadTime	Tweet	Remark
1	dongan01 (Ki...	2012-04-12 오...	RT @mindgoo...	문재인
2	anjfakftjfb ([4....	2012-04-12 오...	RT @felixoal: ...	문재인
3	vegimonsterrr ...	2012-04-12 오...	RT @_Illuin: ...	문재인
4	amamaryllis69,..	2012-04-12 오...	원래 종북연대...	문재인
5	Tae_lyon (Ki...	2012-04-12 오...	RT @weloveh...	문재인
6	kastalienworld...	2012-04-12 오...	RT @so_picky...	문재인
7	springwait6412...	2012-04-12 오...	RT @ilpyungd...	문재인
8	chick9513 (허...	2012-04-12 오...	@limhyojin039...	문재인
9	ahn_news (ah	2012-04-12 오	[illegible]	문재인

	Account	LoadTime	Tweet	Remark
1	daumtrend (D...	2012-04-12 오...	1)총선 결과 htt...	정몽준
2	holynet1 (KOR...	2012-04-12 오...	@pgh545 【大...	정몽준
3	blu_lilac (lilac)	2012-04-12 오...	정몽준은 그니...	정몽준
4	ahna28_ah (조...	2012-04-12 오...	RT @natt747: ...	정몽준
5	cmh219 (멋진...	2012-04-12 오...	RT @natt747: ...	정몽준
6	JunoKw (N/A ...	2012-04-12 오...	10.~~ 전초전...	정몽준
7	yooiee (용혜인)	2012-04-12 오...	@oddity1968 ...	정몽준
8	daumtrend (D...	2012-04-12 오...	1)총선 결과 htt...	정몽준
9	ponkeynet (po	2012-04-12 오	Daum [illegible]	정몽준

Figure 4. Detecting events on TweetPulse by politicians

writes "$tweet_3$" on the Twitter. This tweet includes the keyword of "M_1", and the users who follow the actress RT this message, so on 12 April 2012 the keyword of "M_1" has propagated fast.

4.2 Continuous Time of the Information Diffusion

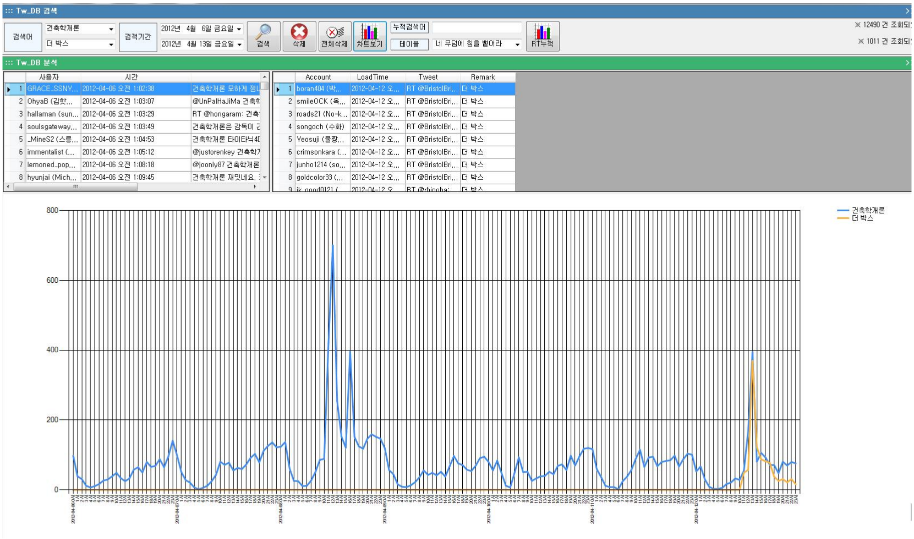
Figure 7 shows highest diffusion speed in some time-zones. At 9 AM on 16 April, there is a high diffusion speed compared with other time-zones of keyword "P_2" related to a certain politician. Also, at 11 AM on 21 April, a keyword of another politician "P_3" has a higher diffusion speed than other time-zones.

In this paper, we have also studied diffusion continuance time when an event has happened to one keyword. According to how long a tweet has been retweeted, we can check it using two types of analytics:

1. macroscopic and
2. microscopic analytic methods.

4.2.1 Macroscopic Analytic Method

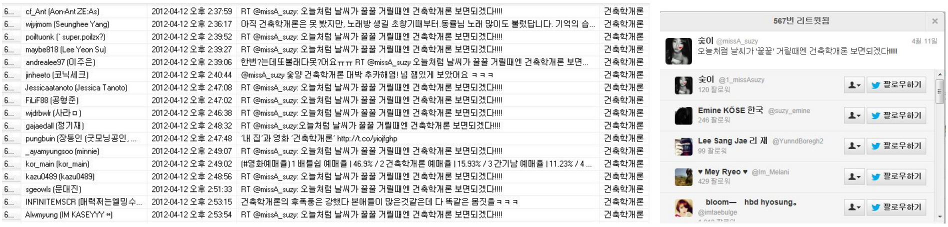
First we can check it on the monitor of events detecting on TweetPulse. We can call it *macroscopic* analytic method. For example, in Figure 7, we can see that "Ahn" keyword at the time-zone of 16 April 2012 (9 AM) has an event detected, also "Kim" keywords have event detected at 21 April 2012 (1 AM) and the diffusion



	사용자	시간	
1	GRACE_SSNY...	2012-04-06 오전 1:02:38	건축학개론 묘하게 잼니
2	OhyaB (김핫...	2012-04-06 오전 1:03:07	@UnPalHaJiMa 건축학
3	hallaman (sun,...	2012-04-06 오전 1:03:29	RT @hongaram: 건축
4	soulsgateway...	2012-04-06 오전 1:03:49	건축학개론은 감독이 乙
5	_MineS2 (스릉...	2012-04-06 오전 1:04:53	건축학개론 타이타닉40
6	immentalist (...	2012-04-06 오전 1:05:12	@justorenkey 건축학기
7	lemoned_pop...	2012-04-06 오전 1:08:18	@joonly87 건축학개론
8	hyunjai (Mich...	2012-04-06 오전 1:09:45	건축학개론 재밌네요.

	Account	LoadTime	Tweet	Remark
1	boran404 (박...	2012-04-12 오...	RT @BristolBri...	더 박스
2	smileOCK (옥...	2012-04-12 오...	RT @BristolBri...	더 박스
3	roads21 (No-k...	2012-04-12 오...	RT @BristolBri...	더 박스
4	songoch (수화)	2012-04-12 오...	RT @BristolBri...	더 박스
5	Yeosuji (물짱...	2012-04-12 오...	RT @BristolBri...	더 박스
6	crimsonkara (...	2012-04-12 오...	RT @BristolBri...	더 박스
7	junho1214 (so...	2012-04-12 오...	RT @BristolBri...	더 박스
8	goldcolor33 (...	2012-04-12 오...	RT @BristolBri...	더 박스
9	ik_good0121 (	2012-04-12 오	RT @rhinoba:	더 박스

Figure 5. Detecting events on TweetPulse by movie title



6...	cf_Ant (Aon-Ant ZE:As)	2012-04-12 오후 2:37:59	RT @missA_suzy: 오늘처럼 날씨가 꿀꿀 거릴때엔 건축학개론 보면되겠다!!!!	건축학개론
6...	wjyjmom (Seunghee Yang)	2012-04-12 오후 2:36:17	아직 건축학개론은 못 봤지만, 노래방 생길 초창기때부터.동률님 노래 많이도 불렀답니다. 기억의 습...	건축학개론
6...	poiltuonk (` super.poilzx?)	2012-04-12 오후 2:39:52	RT @missA_suzy: 오늘처럼 날씨가 꿀꿀 거릴때엔 건축학개론 보면되겠다!!!!	건축학개론
6...	maybe818 (Lee Yeon Su)	2012-04-12 오후 2:39:27	RT @missA_suzy: 오늘처럼 날씨가 꿀꿀 거릴때엔 건축학개론 보면되겠다!!!!	건축학개론
6...	andrealee97 (이주은)	2012-04-12 오후 2:39:06	한번?는데또볼래다못?어요ㅠㅠ RT @missA_suzy 오늘처럼 날씨가 꿀꿀 거릴때엔 건축학개론 보면...	건축학개론
6...	jinheeto (코닉세크)	2012-04-12 오후 2:40:44	@missA_suzy 숯양 건축학개론 대박 추카해염! 넘 잼있게 보았어요 ㅋㅋㅋ	건축학개론
6...	Jessicaatanoto (Jessica Tanoto)	2012-04-12 오후 2:47:08	RT @missA_suzy: 오늘처럼 날씨가 꿀꿀 거릴때엔 건축학개론 보면되겠다!!!!	건축학개론
6...	FiLiF88 (공형준)	2012-04-12 오후 2:47:02	RT @missA_suzy: 오늘처럼 날씨가 꿀꿀 거릴때엔 건축학개론 보면되겠다!!!!	건축학개론
6...	wjdrbwlr (사라ㅁ)	2012-04-12 오후 2:46:38	RT @missA_suzy: 오늘처럼 날씨가 꿀꿀 거릴때엔 건축학개론 보면되겠다!!!!	건축학개론
6...	gajaedall (정기재)	2012-04-12 오후 2:48:32	RT @missA_suzy:오늘처럼 날씨가 꿀꿀 거릴때엔 건축학개론 보면되겠다!!!!	건축학개론
6...	pungbuin (장동인 (굿모닝공인, ...	2012-04-12 오후 2:47:48	'내 집'과 영화 '건축학개론' http://t.co/yiojfghp	건축학개론
6...	_ayamyungsoo (minnie)	2012-04-12 오후 2:49:07	RT @missA_suzy: 오늘처럼 날씨가 꿀꿀 거릴때엔 건축학개론 보면되겠다!!!!	건축학개론
6...	kor_main (kor_main)	2012-04-12 오후 2:49:02	(#영화예매율) 1 배틀쉽 예매율	46.9% / 2 건축학개론 예매율	15.93% / 3 간기남 예매율	11.23% / 4 ...	건축학개론
6...	kazu0489 (kazu0489)	2012-04-12 오후 2:48:56	RT @missA_suzy: 오늘처럼 날씨가 꿀꿀 거릴때엔 건축학개론 보면되겠다!!!!	건축학개론
6...	sgeowls (문대진)	2012-04-12 오후 2:51:33	RT @missA_suzy: 오늘처럼 날씨가 꿀꿀 거릴때엔 건축학개론 보면되겠다!!!!	건축학개론
6...	INFINITEMSCR (매력쩌는엘밍수...	2012-04-12 오후 2:53:15	건축학개론의 후폭풍은 강했다 본애들이 많은것같은데 다 똑같은 몸짓을ㅋㅋㅋ	건축학개론
6...	Alwmyung (IM KASEYYY ∞)	2012-04-12 오후 2:53:54	RT @missA_suzy: 오늘처럼 날씨가 꿀꿀 거릴때엔 건축학개론 보면되겠다!!!!	건축학개론

567번 리트윗됨

숯이 @missA_suzy 4월 11일
오늘처럼 날씨가 '꿀꿀' 거릴때엔 건축학개론 보면되겠다!!!!

숯이 @1_missAsuzy
120 팔로워

Emine KÖSE 한국 @suzy_emine
246 팔로워

Lee Sang Jae 리 재 @YunndBoregh2
99 팔로워

♥ Mey Ryeo ♥ @Im_Melani
429 팔로워

bloom— hbd hyosung。 @imtaebulge

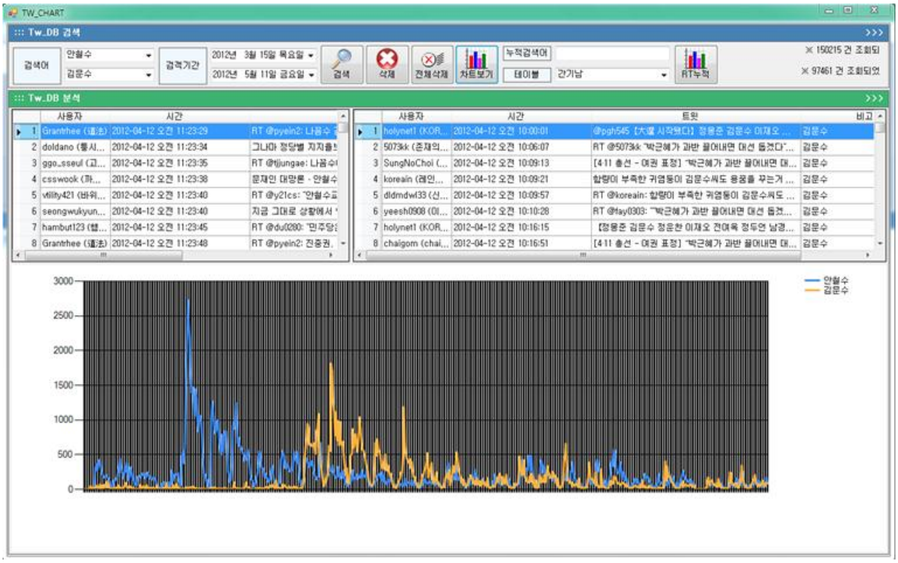
Figure 6. DB data of "M_1"

phenomenon lasts for 3 to 5 days; but maybe there are other events at the same time. For example, one event happened on 21 April 2012, and it just continued for a day, and another event happened on the next day.

4.2.2 Microscopic Analytic Method

To discover how long the diffusion phenomenon holds details, we have to use the microscopic analytic method. By help of this method, we will check the continuance time of the information diffusion. Thus, we can see it on Tweet cumulative analysis program. There will be many tweets related with a keyword to be uploaded every day. We study the emergence of some important events, so we just select the ones with over 80 tweets propagated.

As shown in Figure 8, on 16 April, the important event which has caused significant diffusion is "*tweet*$_4$". Compared in Figure 8, we cannot see the "*tweet*$_4$" for



	사용자	시간	
1	Grantrhee (道法)	2012-04-12 오전 11:23:29	RT @pyein2: 나꼼수
2	doldano (돌시...	2012-04-12 오전 11:23:34	그나마 정당별 지지율
3	ggo_sseul (고...	2012-04-12 오전 11:23:35	RT @tjjungae: 나꼼수
4	csswook (파...	2012-04-12 오전 11:23:38	문재인 대망론 - 안철수
5	vtility421 (바위...	2012-04-12 오전 11:23:40	RT @y21cs: "안철수교
6	seongwukyun...	2012-04-12 오전 11:23:40	지금 그대로 상황에서
7	hambut123 (햄...	2012-04-12 오전 11:23:45	RT @du0280: "민주당
8	Grantrhee (道法)	2012-04-12 오전 11:23:48	RT @pyein2: 진중권,

	사용자	시간	트윗	비고
1	holynet1 (KOR...	2012-04-12 오전 10:00:01	@pgh545 【大選 시작됐다】 정몽준 김문수 이재오 ...	김문수
2	5073kk (존재의...	2012-04-12 오전 10:06:07	RT @5073kk "박근혜가 과반 끌어내면 대선 등겠다"...	김문수
3	SungNoChoi (...	2012-04-12 오전 10:09:13	[4·11 총선 - 여권 표정] "박근혜가 과반 끌어내면 대...	김문수
4	koreain (레인...	2012-04-12 오전 10:09:21	함량이 부족한 귀염둥이 김문수씨도 용꿈을 꾸는거 ...	김문수
5	dldmdwl33 (신...	2012-04-12 오전 10:09:57	RT @koreain: 함량이 부족한 귀염둥이 김문수씨도 ...	김문수
6	yeesh0908 (이...	2012-04-12 오전 10:10:28	RT @lay0303: "박근혜가 과반 끌어내면 대선 등겠...	김문수
7	holynet1 (KOR...	2012-04-12 오전 10:16:15	【정몽준 김문수 정운찬 이재오 전여옥 정두언 남경...	김문수
8	chaigom (chai,...	2012-04-12 오전 10:16:51	[4·11 총선 - 여권 표정] "박근혜가 과반 끌어내면 대...	김문수

Figure 7. Detecting events on TweetPulse by politicians

another two days except the 16th. However, we can see the tweet whose "$tweet_4$" lasted for 3 days.

Next we want to compare the tweets of "Kim". As shown in Figure 9, on 21 April, the important event which caused tweet diffusion is "$tweet_5$". Figure 9 shows that "$tweet_5$" has lasted for 2 days, and on 23 April, the tweet "comment" has been diffused more quickly than other tweets. It means not only one event happened, but also several events happened together. Also, most of the events have been just diffused during 1 to 2 days, while only a few events have lasted for 3 to 5 days.

5 EXPERIMENTATION

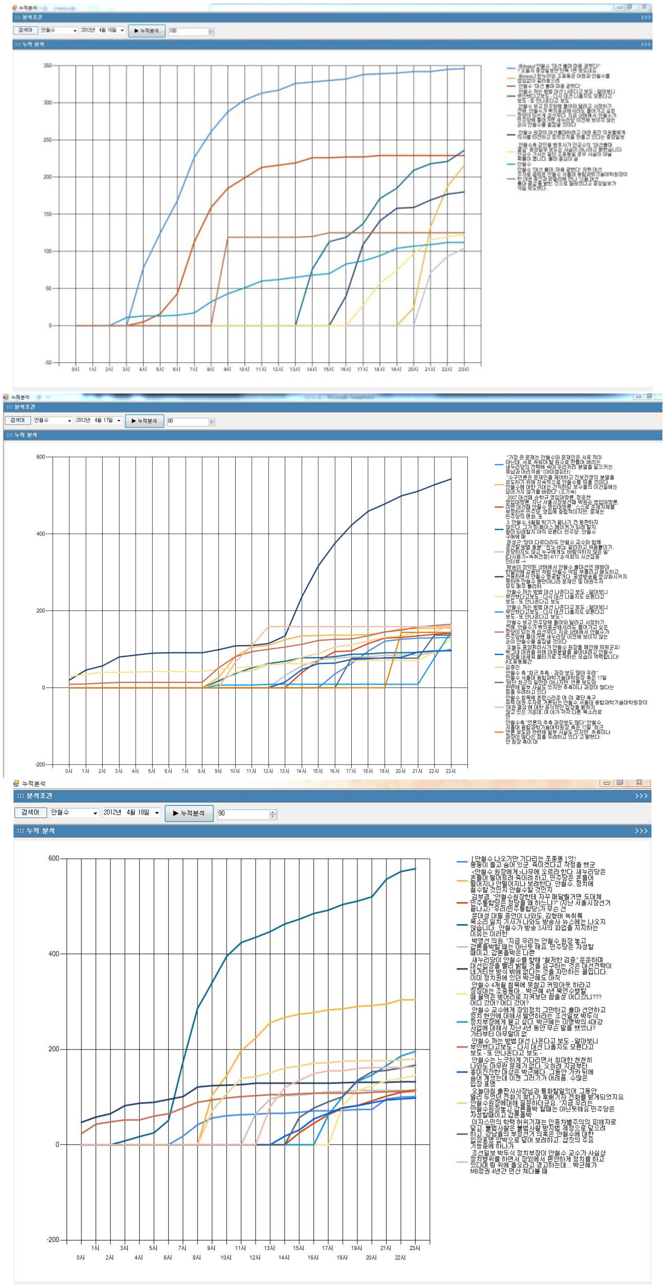
5.1 System Architecture

To understand the TweetPulse, the simulation system is made up of two software modules:

1. data collection function and
2. analysis tool for understanding diffusion phenomenon.

The analysis tool for understanding diffusion phenomenon is also composed of two parts; the first one is Tweet counter analysis program and the second is Tweet cumulative analysis program.

As shown in Figure 10, the data collecting program divides the table for each keyword to collect data in the database by Tweet API. Tweet counter analysis



누적분석

::: 분석조건

검색어 안철수 2012년 4월 16일 ▶ 누적분석 100

::: 누적 분석

@dogsul 안철수 "대선 출마 마음 굳혔다" ? 오늘자 중앙일보만 단독 1면 보도네요.

@oisoo2 한누리와 조중동은 야권과 안철수를 분임없이 갈라놓으려

안철수 "대선 출마 마음 굳혔다"

안철수 까는 방법 대선 나온다고 보도 - 알아보니 부인했다고보도 - 다시 대선 나올지도 모른다고 보도 - 또 안나온다고 보도

안철수 보고 민주당에 들어와 달라고 사정하기 전에, 안철수가 백의종군해서라도 들어가고 싶은 정당이 되는게 급선무다. 지금 상태에서 안철수가 민주당에 들어가면 새누리당 이전에 보이지 않는 손이 안철수를 흠집낼 것이다

안철수 원장이 대선출마하려고 야권 중진 의원들에게 의사를 타진하고 정치조직을 만들고 있다는 중앙일보

안철수측 강인철 변호사가 안교수의 "대선출마 결심" 중앙일보 보도는 사실이 아니라고 밝혔습니다. 안교수 기사는 일단 조중동일 경우 사실이 아닐 확률이 큽니다. 출마 결심이 설

안철수

안철수 "대선 출마, 마음 굳혔다" 유력 대선 주자로 꼽히온 안철수 서울대 융합과학기술대학원장이 한 야권 중진과 비밀리에 만나 '12월 대선 출마 결심'을 밝힌 것으로 알려졌다고 중앙일보가 16일 보도했다.

누적분석

::: 분석조건

검색어 안철수 2012년 4월 17일 ▶ 누적분석 80

::: 누적 분석

"가장 큰 문제는 안철수와 문재인은 서로 적이 아닌데, 서로 싸워야 할 원수로 만들어 버리는 새누리당의 전략에 속아 우리끼리 분열을 일으키는 못남과 어리석음" (아이엠피터)

"수구언론은 문재인을 제어하고 진보진영의 분열을 유도하기 위해 지속적으로 안철수를 띄울 것이다. 안철수에 대한 기대는 간직하되 보수들의 이간질에는 넘어가지 않기를 바란다" (조기숙)

'2007 대선때 손학규 영입대망론, 정운찬 영입대망론, 지난 서울시장보선때 박원순 영입대망론, 이번 대선때 안철수 영입대망론'. 스스로 존재자체를 부정하는 민주당. 영입에 중립적이지만, 문제는 민주당의 변화. 또

3. 안철수, 6월말 학기가 끝나기 전 등판하지 않는다. 그가 킹/페이스 메이커가 되려 할지, 왕이 되려할지 아직 모른다. 민주당, 안철수 구애에 매

문성근 "당이 다르더라도 안철수 교수와 함께 경선할 방법 충분" "친노-비노 갈라치고 싸움붙이기, 온당하지도 않고 누구에게도 바람직하지 않은 일" [다시듣기+녹취전문] 4/17 손석희의 시선집중 인터뷰 →

방송이 장악된 상태에서 안철수 출마선언 해봐야 한달만에 김용민 처럼 안철수 약점 부풀리고 매도하고 선동하면서 안철수 맹공할거다. 공영방송을 정상화시키지 못하면 안철수 뿐만아니라 문재인 및 야권주자 모두 매우 불리하

안철수 까는 방법 대선 나온다고 보도 - 알아보니 부인했다고보도 - 다시 대선 나올지도 모른다고 보도 - 또 안나온다고 보도

안철수 까는 방법 대선 나온다고 보도 - 알아보니 부인했다고보도 - 다시 대선 나올지도 모른다고 보도 - 또 안나온다고 보도 -

안철수 보고 민주당에 들어와 달라고 사정하기 전에, 안철수가 백의종군해서라도 들어가고 싶은 정당이 되는게 급선무다. 지금 상태에서 안철수가 민주당에 들어가면 새누리당 이전에 보이지 않는 손이 안철수를 흠집낼 것이다

오늘도 중앙찌라시가 안철수 원장을 메인에 띄웠군요! 박그네 대권을 위해 야권분열을 끌어내려고 안철수 원장을 내세워 물타기로 조작하는 모습이 역력합니다! #조중동폐간

김종인

안철수 측 "최근 추측 · 과장 보도 많아 우려" 안철수 서울대 융합과학기술대학원장 측은 17일 "비단 최근의 일만은 아니지만, 언론 보도와 관련해 일부 사실도 있지만 추측이나 과장이 많다는 점을 우려하고 있다

안철수 침묵에 혼란스러운 여·야, 결단 촉구 유력 대권 주자로 거론되는 안철수 서울대 융합과학기술대학원장이 '대권 결심'에 대한 공식적인 입장을 밝히지 않고 있는 가운데, 여·야가 각각 다른 목소리로 안

안철수측 "언론의 추측 과장보도 많다" 안철수 서울대 융합과학기술대학원장 측은 17일 "최근 언론 보도와 관련해 일부 사실도 있지만, 추측이나 과장이 많다는 점을 우려하고 있다"고 밝혔다. 안 원장 측이 대

누적분석

::: 분석조건

검색어 안철수 2012년 4월 18일 ▶ 누적분석 90

::: 누적 분석

[안철수 나오기만 기다리는 쪼중똥] 앗! 몽둥이 들고 숨어 있군. 죽이겠다고 작정을 했군

<안철수 원장에게>나무에 오르라 한다. 새누리당은 흔들어 떨어트려 죽이려 하고, 민주당은 흔들어 떨어지나 안떨어지나 보려한다. 안철수, 정치에 철수할 것인지 안철수할 것인지

김부겸, "안철수원장한테 자꾸 매달릴거면 도대체 민주통합당은 정당을 왜 하느냐?" (지난 서울시장선거 끝나고) "우리(민주통합당)가 무슨 선

문대성 대필 증언이 나와도, 김형태 녹취록 목소리 일치 기사가 나와도 방송사 뉴스에는 나오지 않습니다. 안철수가 방송 3사의 파업을 지지하는 이유는 이러한

박영선 의원, "지금 우리는 안철수 원장 놓고 갑론을박할 때는 아닌듯 해요. 민주당은 자성할 때이고, 갑론을박은 나쁜

새누리당이 안철수를 향해 "철저한 검증" 운운하며 대선입장을 빨리 밝힐 것을 요구하는 것은 대선전략이 네거티브 방식 밖에 없다는 것을 자인하는 꼴입니다. 이미 정치권에 있던 박근혜도 아직

안철수 4개월 침묵에 못참고 커밍아웃 하라고 징징대는 조중동아... 박근혜 4년 묵언수행할 때 꿀먹은 벙어리로 지켜보던 참을성 어디갔니??? 어디 갔어? 어디 갔어?

안철수 교수에게 장외정치 그만하고 출마 선언하고 정치 현안에 대해서 발언하라는 조선일보 박두식 정치부장에게 묻고 싶다. 박근혜는 이명박의 4대강 사업에 대해서 지난 4년 동안 무슨 말을 했었나? 가타부타 아무말이 없

안철수 까는 방법 대선 나온다고 보도 - 알아보니 부인했다고보도 - 다시 대선 나올지도 모른다고 보도 - 또 안나온다고 보도 -

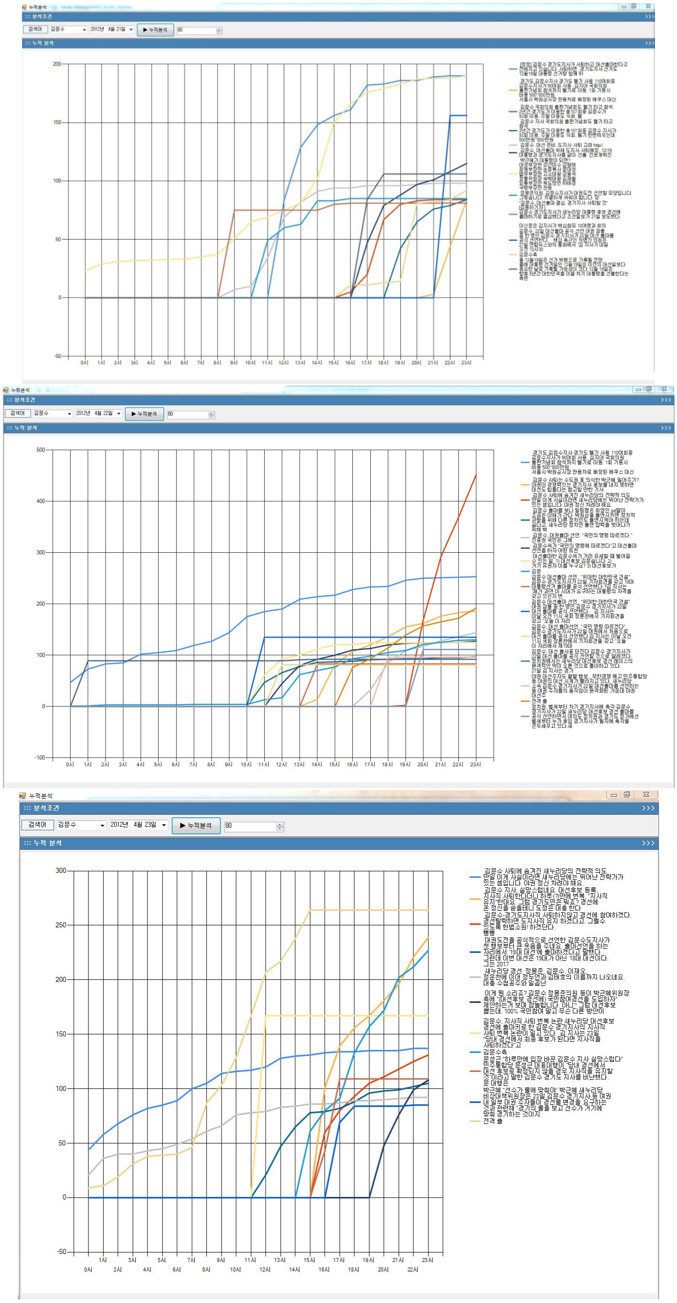
안철수는 느긋하게 기다리면서 최대한 천천히 나와도 아무런 문제가 없다. 오히려 지금부터 흥미진진한 대상은 박근혜다. 그동안 가카 뒤에 숨어 계셨는데 이젠 그러기가 어려움. 수많은 입장 표명

오늘아침 출판사사장님과 통화할일있어 그동안 멀리 두었던 전화기 찾다가 후배기자 전화를 받게되었지요 안철수원장에대해 질문하더군요. "지금 우리는 안철수원장놓고 갑론을박 할때는 아닌듯해요 민주당은 자성할때이고 갑론을박

이자스민의 학력 허위기재는 인종차별주의의 피해자로 덮고, 불법사찰은 불법사찰 방지법 제정으로 덮으려 하고, 강남을의 부정선거 의혹은 안철수에 대한 입장표명 압박으로 덮어 보려하고. 삽질의 주요 기능중에 하나가

조선일보 박두식 정치부장이 안철수 교수가 사실상 정치행위를 하면서 장외에서 편안하게 정치를 하고 있다며 링 위에 올오라고 경고하는데... 박근혜가 MB정권 4년간 먼산 쳐다볼 때

Figure 8. Tweet cumulative analysis of "Ahn" on a) 16 April, b) 17 April, and c) 18 April



김문수 2012년 4월 21일 80

[정정] 김문수 경기도지사가 사퇴하고 대선출마한다고 전해지고 있습니다. 사퇴하면, 경기도지사 선거도 12월19일 대통령 선거랑 함께 하

경기도 김문수지사 경기도 헬기 사용 110여회중 김문수지사가 90여회 사용. 심지어 국회의원 출판기념회 참석까지 헬기로 이동. 1회 기동시 비용 500~800만원. 서울시 박원순시장 관용차로 배정된 에쿠스 대신

김문수 국회의원 출판기념회도 헬기 타고 참석. 2년간 경기도가 이용한 총107회중 김문수가 93회 이용, 주말 이용도 16회, 헬

김문수 지사 국회의원 출판기념회도 헬기 타고 참석 2년간 경기도가 이용한 총107회중 김문수 지사가 93회 이용, 주말 이용도 16회, 헬기 한번띄우는데 500만원~800만원

김문수, 대선 준비..도지사 사퇴 고려 http//

김문수, 대선출마 위해 도지사 사퇴예정.. 12/19, 대통령과 경기도지사를 같이 선출. 진보개혁진

박근혜가 대통령이 되면?
여성부장관 공간미수 김현태
문체부장관 논문복사 문대성
법무부장관 고소대왕 강용석
[illegible]
[illegible]
국방부장관 관등

정몽준의원, 김문수지사가 대권도전 선언할 모양입니다. 그렇습니다. 치열하게 싸워야 합니다. 당

"김문수, 대선출마 결심, 경기지사 사퇴할 것" [김동하기자] 김문수 경기도지사가 새누리당 대통령 후보 경선에 출마하기로 결심했다고 조선일보가 21일 보도했다.

이신문은 김지사가 핵심참모 10여명과 회의
김문수, 22일 대선출마 공식 선언 여권 잠룡 중 한 명인 김문수 경기지사가 22일 대선 출마를 공식 선언한다. 핵심 측근인 차명진 의원은 21일 연합뉴스와의 통화에서 "김 지사가 내일 오전 10시30

김문수측

올 12월19일은 선거 빅뱅으로 기록될 전망 올해 대통령 선거일인 12월19일은 이전의 대선일보다 중요한 날로 기록될 가능성이 크다. 12월 19일은 향후 5년간 대한민국을 이끌 차기 대통령을 선출한다는 측면

김문수 2012년 4월 22일 80

경기도 김문수지사 경기도 헬기 사용 110여회중 김문수지사가 90여회 사용. 심지어 국회의원 출판기념회 참석까지 헬기로 이동. 1회 기동시 비용 500~800만원. 서울시 박원순시장 관용차로 배정된 에쿠스 대신

김문수 사퇴는 수도권 표 의식한 박근혜 밀어주기? 야권이 경쟁력있는 경기지사 후보를 내지 못하면 대선도 힘들다는 참고할 만한 기사.

김문수 사퇴에 숨겨진 새누리당의 전략적 의도. 만일 이게 사실이라면 새누리당에는 뛰어난 전략가가 있는 셈입니다. 야권 정신 차려야 해요.

김문수 출마를 보니 힐링캠프 최영인 pd말이 조금은 이해가 간다. 박원순을 출연시키면 정치적 균형을 위해 다른 정치인도 출연시켜야 하는데 싫다고. 새누리당 정치인 출연 압력을 벗어나기 위해 박

김문수, 대권출마 선언. "국민의 명령 따르겠다." 진중권 국민은 그에

김문수씨가 "국민의 명령에 따르겠다"고 대선출마 선언을 하자 어떤 트친

대선출마한 김문수씨가 거리 유세할 때 벌어질 수 있는 일. 1) 대선후보 김문습니다. 2) 거기 유권자 이름 누구요? 3) 대선후보가

김문

김문수 대선출마 선언.. "위대한 대한민국 건설" 김문수 경기도지사가 22일 기자회견을 갖고 18대 대통령선거 출마를 공식 선언했다.?김 지사는 "제가 과연 이 시대가 요구하는 대통령의 자격을 갖고 있는지 번

김문수 대선출마 선언.. "위대한 대한민국 건설" 여권 잠룡 중 한 명인 김문수 경기지사가 22일 대선 출마를 공식 선언했다. 김 지사는 이날 오전 11시 국회 정론관에서 기자회견을 갖고 "오늘 이 자리

김문수, 대선 출마선언.."국민 명령 따르겠다" 김문수 경기도지사가 22일 여권에서 처음으로 대선 출마를 공식 선언했다.김 지사는 이날 오전 11시 국회 정론관에서 기자회견을 갖고 "오늘 이 자리에서 제19대

김문수, 대선 출사표 던진다 김문수 경기지사가 22일 대선 출마를 공식 선언할 것으로 알려졌다. 정치권에서는 새누리당 대선후보 경선 레이스의 본격적인 막이 오른 것으로 풀이하고 있다. 21일 김 지사는 경기

야권 대선주자도 활발 행보.. 무한경쟁 예고 민주통합당 등 야권의 대선 시계가 빨라지고 있다. 새누리당 소속 김문수 경기지사가 22일 대선출마를 선언하는 등 여권 주자들의 움직임이 본격화된 가운데 야권 대선주

전격 출

정치권, 벌써부터 차기 경기지사에 촉각 김문수 경기지사가 22일 새누리당 대선후보 경선 출마를 공식 선언하면서 여의도 정치권과 경기도 정가에선 벌써부터 누가 후임 경기지사가 될지에 촉각을 곤두세우고 있다.새

김문수 2012년 4월 23일 80

김문수 사퇴에 숨겨진 새누리당의 전략적 의도. 만일 이게 사실이라면 새누리당에는 뛰어난 전략가가 있는 셈입니다. 야권 정신 차려야 해요.

김문수 지사, 실망스럽네요. 대선후보 등록, 지사직 사퇴한다더니 하루(?)만에 번복, "지사직 유지"한대요. 그럼 경기도민은 뭐죠? 경선에 온 정신을 쏟을테니 도정은 대충 한다

김문수-경기도지사직 사퇴하지않고 경선에 참여하겠다. 경선탈락하면 도지사직 유지 하겠다고. 그럴수 있도록 헌법소원! 하겠단다. 뻥뻥

대권도전을 공식적으로 선언한 김문수도지사가 첫 행보부터 큰 웃음을 주네요. 출마선언을 하는 자리에서 '19대 대선'에 출마하겠다고 말했다. 그런데 이번 대선은 19대가 아닌 18대 대선이다. 그는 2017

새누리당 경선. 정몽준, 김문수, 이재오, 정운찬에 이어 정두언과 김태호의 이름까지 나오네요. 대충 수첩공주와 일곱난

이게 뭔 소리죠? 김문수 정몽준의원 등이 박근혜위원장 측에 "(대선후보 경선에) 국민참여경선을 도입하자" 제안하는거 보며 깜놀합니다. 아니~ 그럼 대선후보 뽑는데, 100% 국민참여 말고 무슨 다른 방안이

김문수, 지사직 사퇴 번복 논란 새누리당 대선후보 경선에 출마키로 한 김문수 경기지사의 지사직 사퇴 번복 논란이 일고 있다. 김 지사는 23일 "당내 경선에서 최종 후보가 된다면 지사직을 사퇴하겠다"고

김문수측

문성근 "하루만에 입장 바꾼 김문수 지사 실망스럽다" 민주통합당 문성근 대표대행이 "당내 경선에서 대선 후보로 확정되지 않을 경우 지사직을 유지할 것"이라고 말한 김문수 경기도 지사를 비난했다. 문 대행은

박근혜 "선수가 룰에 맞춰야" 박근혜 새누리당 비상대책위원장은 23일 김문수 경기지사 등 여권 내 일부 대권 주자들이 경선룰 변경을 요구하는 것과 관련해 "경기의 룰을 보고 선수가 거기에 맞춰 경기하는 것이지

전격 출

Figure 9. Tweet cumulative analysis of “Kim” on a) 21 April, b) 22 April, and c) 23 April

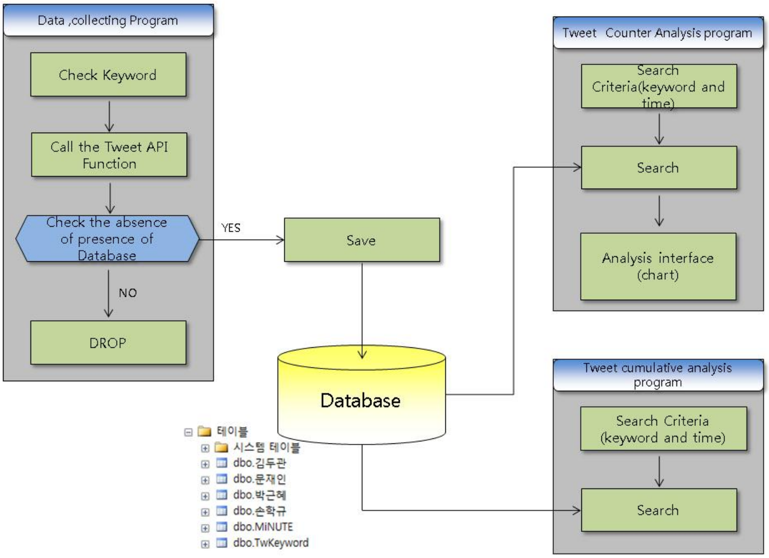

Data ,collecting Program
Check Keyword
Call the Tweet API Function
Check the absence of presence of Database
YES
NO
DROP
Save
Database
테이블
시스템 테이블
dbo.김두관
dbo.문재인
dbo.박근혜
dbo.손학규
dbo.MiNUTE
dbo.TwKeyword
Tweet Counter Analysis program
Search Criteria(keyword and time)
Search
Analysis interface (chart)
Tweet cumulative analysis program
Search Criteria (keyword and time)
Search

Figure 10. System architecture

program shows the tweet counter on the time during which the tweet information was collected in the database. Tweet cumulative analysis program shows the tweet information which was collected in the database by accumulating during one day.

5.2 Data Collection

For experimentation, we have targeted 8 politicians and 16 movie titles for collecting datasets and the keywords. In this study, we have collected data during 4 months (14 March 2012 to 12 July 2012). Thus, we have collected about 4 000 000 tweets. With this, we have constructed a vast database, and were able to search for 4 months data to cover the flaw which search tweet just for 7 days in Twitter.

5.3 Experimental Environment

In this paper, we have collected data for 4 months (14 March 2012 to 12 July 2012). We have developed a monitoring program, as shown in Figure 11.

This system is developed by using C# and MS-SQL. It can save the tweets which were searched using 8 keywords at special interval simultaneously. The most superior point in this system is that it can be run and collect data in PC without using the web page. The advantages of this program are in the ability to speculate what event has happened at the time-zones when a large diffusion speed is shown. Also, we can guess at the time-zone when there is a large diffusion speed that some events happened. More importantly, tweet cumulative analysis program is a monitor

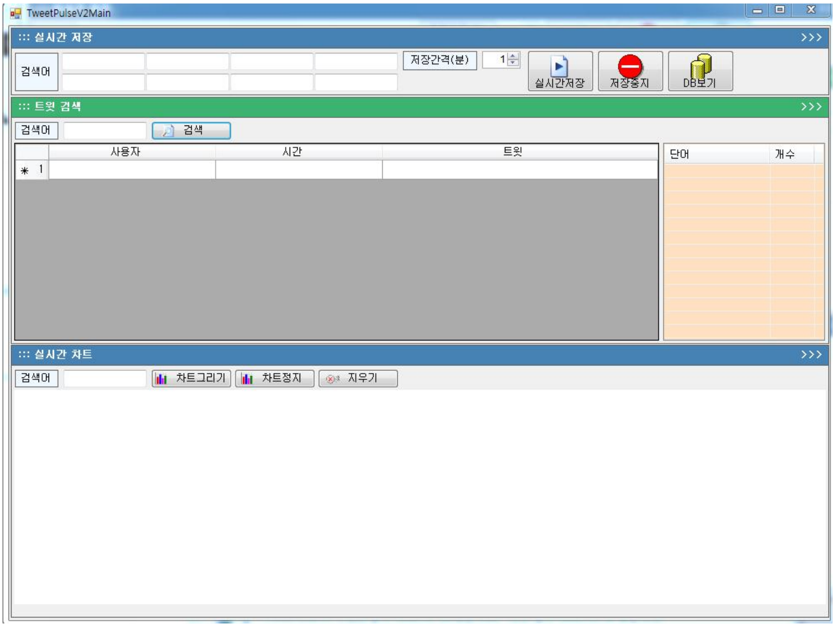


Figure 11. Snapshot on GUI of Tweet collection software

program made to study the relationship of diffusion phenomena and perceptual language of tweets. The program is also used for microscopic analysis continuance time of the information diffusion.

Data cleansing such as de-duplication is an important issue for understanding information diffusion patterns. But, in these systems, the meaning of the diffusion of information is not accurate. Tweet search on 100 tweets per second is used in the data collecting program. Every 1 second returns 100 tweets, thus there will be collect duplication. Due to this problem, the process of duplication removal is necessary. Given a data which want to save t, the data can be formalized by

$$t = \langle RT, U, T, S, sURL \rangle. \tag{5}$$

As shown in the Figure 12, before saving the data in DB which collect 100 data every 1 second to check user, upload-time, and tweet in DB, when $(U_1 = U_2) \wedge (S_1 = S_2) \wedge (T_1 = T_2)$, in other words $t_1 = t_2$, it means the data is already saved in database, so it can be dropped. In opposite, if $t_1 \neq t_2$, the data will be saved in DB.

Figure 13 is a chart which cumulates tweets for more than 30 days. The tweets which view significant diffusion speed include emotional words. The tweets which include emotional words diffused faster and more abundantly.

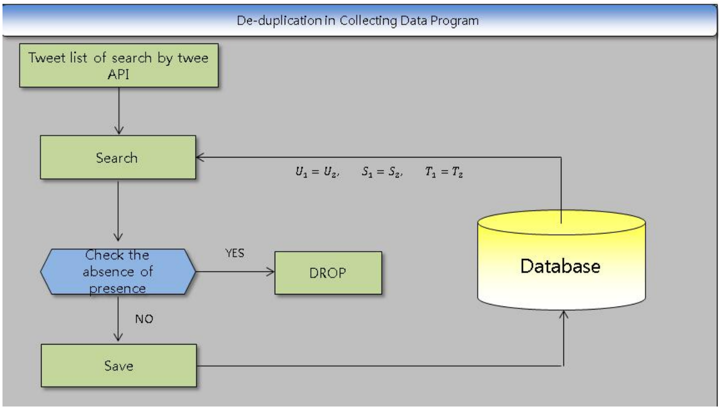
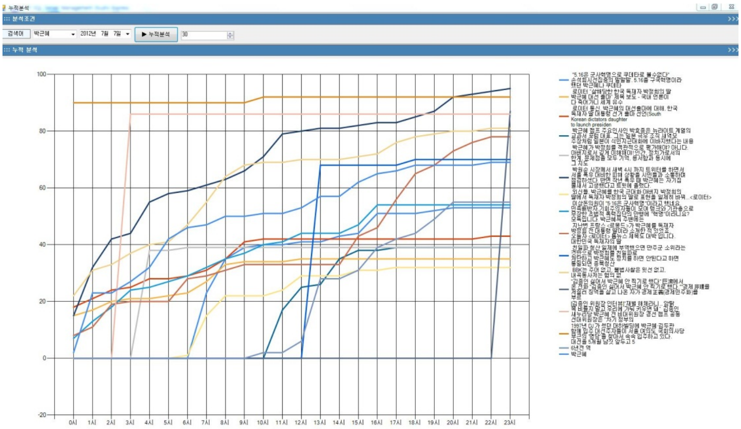

De-duplication in Collecting Data Program
Tweet list of search by twee API
Search
$U_1 = U_2$, $S_1 = S_2$, $T_1 = T_2$
Database
Check the absence of presence
YES
DROP
NO
Save

Figure 12. System architecture of the duplication removal process



검색어 박근혜 2012년 7월 7일 ▶ 누적분석 30

- "5.16은 군사혁명으로 쿠데타로 볼수없다" 손석희시선집중의 말말말. 5.16를 구국혁명이라 했던 박근혜다 쿠데타
- 로이터 "살해당한 한국 독재자 박정희의 딸 박근혜 대선 출마" 제목 보도 - 국내 언론이 다 죽어가니 세계 유수
- 로이터 통신, 박근혜의 대선출마에 대해, 한국 독재자 딸 대통령 선거 출마 선언(South Korean dictators daughter to launch presiden
- 박근혜 캠프 주요인사인 박효종은 뉴라이트 계열의 교과서 포럼 대표. 그는 일본 극우 조직 새역모 주장처럼 일본이 식민지근대화에 이바지했다는 내용
- 박근혜가 박정희를 객관적으로 평가해야? 아니다. 아버지로서 깊게 이해해야! 인간, 정치가로서의 한계, 문제점을 모두 기억, 용서함과 동시에 그 지도
- 박원순 시장께서 새벽 4시 까지 트위터를 하면서 서울 폭우 대비한 피해 상황을 시민들과 소통하며 점검하셨다. 반면 작년 폭우 때 박근혜는 자기집 물새서 고생했다고 트윗에 올렸다.
- 외신들, 박근혜를 한국 근대화 아버지 박정희의 딸에서 독재자 박정희의 딸로 표현을 일제히 바꿔...<로이터>
- 이상돈의원이 "5.16은 군사혁명"이라고 했네요. 민족배반자,기회주의자들이 모여 탱크와 기관총으로 무장한 초법적 폭력집단의 만행에 "혁명"이라니요? 모독입니다. 박근혜씨 주변에는
- 지난번 프랑스 <르몽드>가 박근혜를 독재자 박정희 전 대통령 딸이라 소개한 적 있었죠. 오늘자 <로이터> 통뉴스 제목도 대박 입니다. 대한민국 독재자의 딸
- 친일파 청산 일제에 부역했으면 만주군 소위라는 것만으로 박정희를 친일파로 처단하고 박근혜도 정치를 하면 안된다고 하면 통일되면 종북청산
- BBK는 주어 없고, 불법사찰은 윗선 없고, 내곡동사저는 혐의 없
- "김종인 싫어서 박근혜 안 찍기로 했다" 巨濟에서 온 전화 "김종인 싫어서 박근혜 안 찍기로 했다." "경제 非理를 저질러 징역을 살고 나온 자가 경제 正義(경제민주화)를 부르
- [김종인 위원장 인터뷰]"재벌 해체라니.. 암탉 목 비틀지 말고 우리에 가둬 키우면 돼" 김종인 새누리당 박근혜 전 비대위원장 경선 캠프 공동 선대위원장은 "자기 정부의
- 1997년 DJ가 썼던 대하빌딩에 박근혜·김두관 함께 입주 대선주자들이 서울 여의도 국회의사당 부근의 '명당'을 찾아서 속속 입주하고 있다. 대선을 5개월 남짓 앞두고 5
- 6년전 역
- 박근혜

Figure 13. Tweet cumulative analysis program

6 CONCLUDING REMARKS AND FUTURE WORK

In this paper, we designed and implemented a model which can analyze the diffusion of information efficiently by finding the TweetPulse, and showed the rate of diffusion using visual interface. First of all, the relationship between information diffusion and some events keyword were analyzed, and then the number of followers of the user who write a tweet and its effect on information diffusion. In our next researches, we shall try to find what feelings are added when a tweet is diffused.

Acknowledgement

This work was supported by the Yeungnam University research grants in 2013.

REFERENCES

[1] Yang, J.—Counts, S.: Comparing Information Diffusion Structure in Weblogs and Microblogs. In: W. W. Cohen, S. Gosling (Eds.): Proceedings of the 4th International Conference on Weblogs and Social Media (ICWSM 2010), May 23–26, AAAI Press, Washington, DC, USA 2010.

[2] Pham, X. H.—Jung, J. J.—Hwang, D.: Beating Social Pulse: Understanding Information Propagation via Online Social Tagging Systems. Journal of Universal Computer Science, Vol. 18, 2012, No. 8, pp. 1022–1031.

[3] Cha, M.—Mislove, A.—Gummadi, K. P.: A Measurement-Driven Analysis of Information Propagation in the Flickr Social Network. In: J. Quemada, G. Léon, Y. S. Maarek, W. Nejdl (Eds.): Proceedings of the 18th International Conference on World Wide Web (WWW 2009), Madrid, Spain, April 20–24, ACM 2009, pp. 721–730.

[4] Jung, J. J.: CONTEXTGRID: A Contextual Mashup-Based Collaborative Browsing System. Information Systems Frontiers, Vol. 14, 2012, No. 4, pp. 953–961.

[5] Lenhart, A.: Adults and Social Network Websites.

[6] Hu, M.—Liu, B.: Mining and Summarizing Customer Reviews. In: W. Kim, R. Kohavi, J. Gehrke, W. DuMouchel (Eds.): Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 22–25, ACM, Seattle, Washington, USA 2004, pp. 168–177.

[7] Jung, J. J.: Attribute Selection-Based Recommendation Framework for Short-Head User Group: An Empirical Study by Movielens and IMDb. Expert Systems with Applications, Vol. 39, 2012, No. 4, pp. 4049–4054.

[8] Siganos, G.—Faloutsos, M.—Faloutsos, P.—Faloutsos, C.: Power Laws and the As-Level Internet Topology. IEEE/ACM Transactions on Networking, Vol. 11, 2003, No. 4, pp. 514–524.

[9] Schneider, A.—Jackson, R.—Baum, N.: Social Media Networking: Facebook and Twitter. The Journal of Medical Practice Management, Vol. 26, 2010, No. 3, pp. 156–157.

[10] Jung, J. J.: Evolutionary Approach for Semantic-Based Query Sampling in Large-Scale Information Sources. Information Sciences, Vol. 182, 2012, pp. 30–39.

[11] Boyd, D.—Golder, S.—Lotan, G.: Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter. In: Proceedings of the 43rd Hawaii International Conference on Systems Science (HICSS-43), IEEE Computer Society, Koloa, Kauai, HI, USA 2010, pp. 1–10.

[12] Krishnamurthy, B.—Willinger, W.—Gill, P.—Arlitt, M. F.: A Socratic Method for Validation of Measurement-Based Networking Research. Computer Communications, Vol. 34, 2011, No. 1, pp. 43–53.

[13] Vaithyanathan, S.: The Value of Social Media Data in Enterprise Applications. In: K. S. Candan, Y. Chen, R. T. Snodgrass, L. Gravano, A. Fuxman (Eds.): Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD 2012), May 20–24, ACM, Scottsdale, AZ, USA 2012, pp. 755–756.

[14] Jung, K.—Heo, W.—Chen, W.: IRIE: Scalable and Robust Influence Maximization in Social Networks. In: Proceedings of the 12th IEEE International Conference on Data Mining (ICDM 2012), Brussels, Belgium 2012.

[15] Jung, J. J.: Semantic Optimization of Query Transformation in a Large-Scale Peer-to-Peer Network. Neurocomputing, Vol. 88, 2012, pp. 36–41.

[16] Valler, N.—Prakash, B. A.—Tong, H.—Faloutsos, M.—Faloutsos, C.: Epidemic Spread in Mobile Ad Hoc Networks: Determining the Tipping Point. In: J. Domingo-Pascual, P. Manzoni, S. Palazzo, A. Pont, C. M. Scoglio (Eds.): Proceedings of the 10th International IFIP TC 6 Networking Conference (NETWORKING 2011), May 9–13, Vol. 6640 of Lecture Notes in Computer Science, Springer 2011, pp. 266–280.

[17] Jung, J. J.: Boosting Social Collaborations Based on Contextual Synchronization: An Empirical Study. Expert Systems with Applications, Vol. 38, 2011, No. 5, pp. 4809–4815.

[18] Ellison, N. B.—Steinfield, C.—Lampe, C.: The Benefits of Facebook "Friends": Social Capital and College Students' Use of Online Social Network Sites. Journal of Computer-Mediated Communication, Vol. 12, 2007, No. 4, pp. 1143–1168.

[19] Jung, J. J.: Discovering Community of Lingual Practice for Matching Multilingual Tags from Folksonomies. Computer Journal, Vol. 55, 2012, No. 3, pp. 337–346.

[20] Vossen, G.: Big Data as the New Enabler in Business and Other Intelligence. Vietnam Journal of Computer Science, Vol. 1, 2014, No. 1, pp. 3–14.

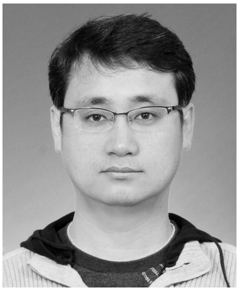
[21] Pham, X. H.—Nguyen, T. T.—Jung, J. J.—Nguyen, N. T.: ⟨A, V⟩-SPEAR: A New Method for Expert Based Recommendation Systems. Cybernetics and Systems, Vol. 45, 2014, No. 2, pp. 165–179.

[22] Nguyen, D. T.—Jung, J. J.: Privacy-Preserving Discovery of Topic-Based Events from Social Sensor Signals: An Experimental Study on Twitter. Scientific World Journal, Vol. 2014, 2014, Article ID 204785.

[23] Jung, J. J.: Semantic Wiki-Based Knowledge Management System by Interleaving Ontology Mapping Tool. International Journal on Software Engineering and Knowledge Engineering, Vol. 23, 2013, No. 1, pp. 51–63.

Dosam Hwang received his Ph. D. degree in natural language processing at Kyoto University. In 1987, he was chosen the best researcher at Korea Institute of Science and Technology. Since 2005 he has been the Professor at Yeungnam University in Korea. Currently he is working on natural language processing, machine translation, ontology, semantic web, and information retrieval. He has served for a number of international conferences and a technical committee for ISO.

Jai E. Jung is a corresponding author of this paper. He is an Associate Professor in Chung-Ang University, Korea, since September 2014. Before joining Chung-Ang University, he was a faculty member in Yeungnam University from 2007. He was a postdoctoral researcher in INRIA Rhône-Alpes, France in 2006, and a visiting scientist in Fraunhofer Institut (FIRST) in Berlin, Germany in 2004. He received the B. Eng. in computer science and mechanical engineering from Inha University in 1999. He received M. Sc. and Ph. D. degrees in computer and information engineering from Inha University in 2002 and 2005, respectively. His research topics include knowledge engineering on social networks by using machine learning, semantic Web mining, and ambient intelligence.



Seungbo Park is a research professor in Kyung Hee University, Korea. He received the B. Sc. and the M. Sc. degrees in Electrical Engineering from Inha University, Korea, in 1995 and 1997. He received the Ph. D. degree in Computer Information Engineering from Inha University, Korea. His research interests include video story analyzing, semantic contents, video knowledge representation, social network analysis, and AI. He worked at Daewoo Electronics as an engineering researcher.

Hien T. Nguyen has served as the Dean, Faculty of Information Technology, Ton Duc Thang University, Vietnam, since 2012. He received Ph. D. in computer science, M. Sc. in computer science, and B. Eng. in computer engineering from Ho Chi Minh City University of Technology, Vietnam, in 2011, 2005, and 2002, respectively. His research interests include job shop scheduling, information extraction, natural language processing, web/text mining, semantic web, and social network analysis and mining.