Correlation - between Social Relation Value and Search Pattern in Social Networks

GunWoo Park
Information System Management
Group
Republic of Korea Army
Gyeryong-si, Republic of Korea
kndugunwoo@gmail.com

Abstract— **Social Network Service (SNS) is a popular and powerful Web service to connect and/or find friends, and the tendency of a user's interest often affects his/her friends who have similar interests. We can 'Collective Intelligence and Social Big Data' through the SNS. If we can track users' preferences in certain boundaries in terms of Web searching, we can we can improve search efficiency and reliability in view of users. In this paper, we analyze correlation - between Social Relation Value and Search Pattern of web users to improve search efficiency and. Social Relation Value indicates degree of relation on latent characteristic among web users who exist in Social Networks. We calculate Social Relation Value in granting association between 'M' topics and 'L' Attributes that are applicable to 'N' Factors(Similarity, Adjacency, Etc.). And then, we analyze the Correlation - between Social Relation Value and Search Pattern by comparing the search result of each web users correspond to input 'Query' based on Social Relation Value calculated according to topics. Consequently, we find that Search Pattern is similar to that of web users who have high Social Relation Value. Namely, web users who have similar characteristic and interact with each other with activity in (Online) Social Networks have high Correlation. Thus, we are able to improve search efficiency and reliability when the correlation is applied to search.**

Keywords—Social Network, Social Network Analysis, Social Big Data, Correlation, Relation Value, Search Pattern, Query Log, Click History, Collective Intelligence

I. INTRODUCTION

It's getting difficult to improve efficiency and reliability of search results. It is because of two reasons- the explosive increase of web information and the needs of the web users. Most of web users only look at the search results from 5th [1] through 20th [2], and if there is no information that they want to find, they tend to immediately go over next search. Such like Search Pattern depends on search intent could be obtained by using 'Query Log(or Click History)' and 'User Interaction',. Etc. But it is difficult to understand the search intent by them and also they are limited because of privacy. Also, user query is very short [2] and unspecified [3], and improving search efficiency and reliability is very hard because every web users have different search intent on same query. However, despite of such ambiguity, the best indicator of web user's search intent is basically 'Query' itself [5].

In this paper, we look for a solution from web users who input query. This idea is started with a concept that if we use common Search Pattern following Social Relation Value of web users exist in Social Networks, efficient search could be possible.

Social Network is composed of members with the center of individual. Social Network is the 'Community' that network formed by friends, an alumni association, and a sort of club members in the Off-Line is moved to On-Line. Also it could be understood as a connection which has been artificially set to exchange the information among web users. How to generate, dismantle or maintain connection of members in Social Networks is core approaching method of Social Network Analysis [6]. It is, while web 2.0 is a network between person and person, person and information, information and information, Social Network is between person and person who is a part of huge web environment. Such Social Networks can analyze connection intensity based on various attributes and logics that express about individual-centered direct-connection or indirect-connection among web users. Therefore web user's satisfaction could be improved if Social Network, which is formed by web users who have similar latent information among them, is applied to web search.

In this paper, we analyze the mutual relation between Social Relation Value and Search Pattern to improve search efficiency. Social Relation Value is a value of similarity about latent characteristic or tendency between 'I' and other web users in Social Networks. Social Relation Value is calculated by giving each 'M' topics' association to 'N' factors(Similarity, Access, Intimacy, Adjacent, Rewardingness,. etc.) that are able to affect mutual human network {('M' Topics) X ('L' Attributes) Association Matrix is generated}. The reason why we calculate Social Relation Value according to topics is the difference of interest and search intent among web users. On each topics, ranking with 'I' could be changed by topics. We show this "**Fig. 1**". For example, in terms of 'Arts', active interaction and information sharing will be occurred with active between 'I' and 'user A'. But when applying topic 'Sports', on or low interaction and information sharing will be occurred since Social Relationship is none. That is, even though web users who have much similarities with other web users in general field, they can have low or no association with each other on other specific fields. So varying Social Relation Value according to topics is need.

Arts		Sports		Business		Education	
Ranking	User	Ranking	User	Ranking	User	Ranking	User
1	A	1	C	1	L	1	Y
2	B	2	M	2	S	2	Z
3	C	3	L	3	Y	3	C
4	X	4	S	4	Q	4	Q
5	M	5	W	5	M	5	M
6	L	6	J	6	J	6	L
7	S	7	O	7	G	7	T
8	Y	8	Y	8	W	8	P
9	Z	9	Z	9	Z	9	S
10	Q	10	V	10	C	10	J
...	...	...	...	...	...	...	...

Fig. 1. Reranking Social Realtion Value according to Topics

For this, first, we build Social Network based on ordinary Social Relation Value calculated according to the association of latent attributes among web users who have direct or indirect connections. And then, we reconstruct a Social Network according to varying Social Relation Value according to topics. Next, on the basis of calculated Social Value according to topics, we make an experiment for comparing Search Pattern among web users. Finally, we analyze correlation - between Social Relation Value and Search Pattern. As a result, we find the fact that Search Pattern is similar to that among web users who have High-Social Relation Value. Based on this fact, we can improve search efficiency and reliability in Personalized / Social Search. This is, on each topic, if web users exchange and share their information with other web users who have High-Social Relation Value according to topics, they can get more accurate and High-Relative Information.

II. Related Work

Recently, much studies of improving web search have been continued, and they could be classified into 2 large fields : Personalized Search and Contextual Search.

First, research on personalization was caused by a study result that a person has more interest on specified topic and that tendency affect to search result. To do Personalized Search, many experiments have been proceeded by applying users' topic preferences into search result based on specific algorithm [7] and reflect the preference of interaction between choice on search result and user query record [8]. But using personal information needs much time and storage space, and reflecting accurate weight on each topic is still difficult things even the user's topic preference is found [8]. It is because user's query pattern is entangled with complicated Relationship between user query and topic search intent. Contextual Search is the method to find such problem through context. Such research includes query context understanding, query context and previous record, synthesis of user reaction [9] or consensus with query, recognition and a method using contextual learning effect followed by document vectors [10, 11]

On the other hand, it is common nowadays to share individual daily life and information and chat with intimates by using SNS(Social Network Service). Focused on that point, online community services based on relationship among web users are becoming active. So, on the basis of social network generated by individual-centered network, many research and services are in process by using active interaction among users in community.

A. Social Network

Human Network which is a kind of structure how people make relationships, is recently introduced as a concept of Social Network especially on online network. Social Network could be defined as a social structure consisted of individual who is made up of at least one interdependent relationship ,and a group. There are typical SNS(Social Network Service) on the web : Friendster [1] , Orkut [2] , Facebook[3] and Cyworld[4] from Korea.

Research on Social Network could be classified to 3 fields. First is a research on improving search efficiency about members or club by using the information of connectivity among users [12, 13]. Second is comparative analysis about situation between Actual Society and Social Network [14, 15, 23, 24, 25, 26, 27, 28, 29, 30, 31]. Final one is a field about efficiency of Social Network composition and Network Security [16, 17]. Recently, in the field of sociology, communication engineering, and economy, research about SNA(Social Network Analysis) which is one of social network, Value calculating algorithm using various internal attributes are ongoing. And study on applying previous situation of social network to web environment is in progress in the field of information science. Method to approach social network analysis is separated into 2 ways : Positional Approach, which is measuring the location in the whole connection network and effect of it, and Relational Approach, which is focused on direct relationship among connection networks. Adjacency Matrix expresses whether relationship between member(i, j) is exist or not using '1' and '0', and it is a basic form of completed network [18]. Usually, 'Cell' of matrix express relationship between i and j. For example, in **"Fig.2"**, the value between 'A' and 'G' is '1'. That means there exist relationship between 'A' and 'G'(for example, in case that 'A' select the 'G' as a friend) and they are expressed by both side arrow.

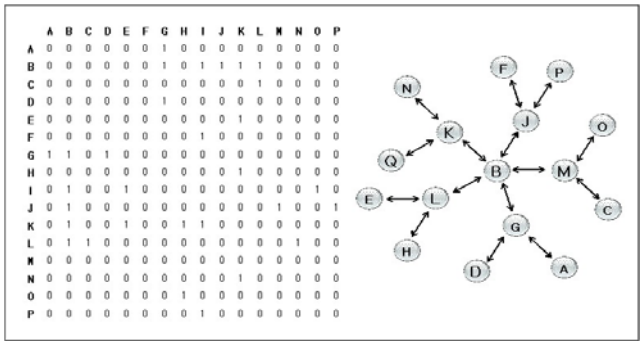


Fig. 2. Adjacency Matrix and Graph

B. Social Network Analysis

There are so many techniques(Connection Union, Centrality, Structural Equivalence, Density, Network Range, Prestige, Cliques, Clustering, Multi-Dimensional Scaling,. Etc.) to analyze Social Network. We present just Connection Union, Centrality and Structural Equivalence.

[1] www.friendster.com

[2] www.orkut.com

[3] www.facebook.com

[4] www.cyworld.com

- Connection Union : There are examples of typical search topics about connection union - who and who become friends, what clique people form, and so on. For instance, if 'N' people want to become a clique on friend network, there are some requirements - all of them are friends each other and they are in an independent group, not a part of large one. Such defined clique has high level of solidarity, collaboration, shared identity, information exchange, and possibility of doing collective action. Another typical concept is the path length. It is a minimum level to connect two nodes. In other words, it means how many steps are needed to connect two nodes. The numerical formula of connection union is expressed below as (1).

$$P_{ij} = 1 - \frac{d_{ij} - 1}{d_{\max}} \quad (1)$$

  • ij : user (i,j)

  • d(max) : Maximum path length in network

- Centrality is one of the most frequently used term related with authority and influence. In most of experiential analysis, an high-centrality individual has special social or economical position. Also, in terms of organization, higher centrality makes high company outcome or high survival rate. Link degree, which means number of connected nodes, is a value calculating centrality. In the graph which has direction, when connections are heading to the node, it is called in-centrality. And the opposite case, it is called out-centrality. The numerical formula of centrality is expressed below as (2).

$$outdegree_{ik} = \sum_{j=1}^{N} Z_{ijk} = Z_{ik}, \quad indegree_{ik} = \sum_{j=1}^{N} Z_{jik} = Z_{ik} \quad (2)$$

  • Z_{ijk} : Relationship from i to j in the k network

  • outdegree : Number of outbound relation from i to ($j_0, j_1, .. j_n$)

  • indegree : Number of inbound relation from ($j_0, j_1, .. j_n$) to i

- Structural Equivalence : Structural Equivalence means how similar form of relationship performers make among them. If performer A and B aren't connected directly but connected with others, they occupy same position. The numerical formula of structural equivalence is expressed below as (3). Like formula (3), if 'i' and 'j' is connected to all q's with same strength(relation from row to column) and if the q's are connected to (i,j) with same strength(relation from column to row), it is defined that distance between (i,j) is '0' and they are in the structural same position.

$$d_{ij} = \left(\sum_{q} (Z_{iq} - Z_{jq})^2 + \sum_{q} (Z_{qi} - Z_{qj})^2\right)^{1/2} \quad (3)$$

Network data is generally shown as binary type(expresses the relationship as '0'- none, and '1'- exist), so structural equivalence could be estimated by similarity-difference Value designed for such data type.

Users are participate in generating a Social Network, and link strength increases as their activity are associated each other. Furthermore it could be seen that a Social Network is evolving and constructing the form of actual society as time goes on. Using such characteristic, research on improving searching efficiency applying social annotation among users and collaborative filtering method, is under active progress.

III. Correlation between Social Realtion Value and Search Pattern

Main feature of a Social Network is that there is various connection due to members' internal attributes, and users use a Social Network with at least one reason. But users who are connected with simple relationship don't have same interest on every topic, and their satisfaction about search result on same query is also different. It is because when users have completely different interest or intention even if they input same query. **"Fig.3"**, is made by web user 'A' who works for IT company. User 'A' inputted query 'Human Brain' to find the image that he/she'd like to insert into his/her technical report, and select good result with 'O' mark and others 'X'.

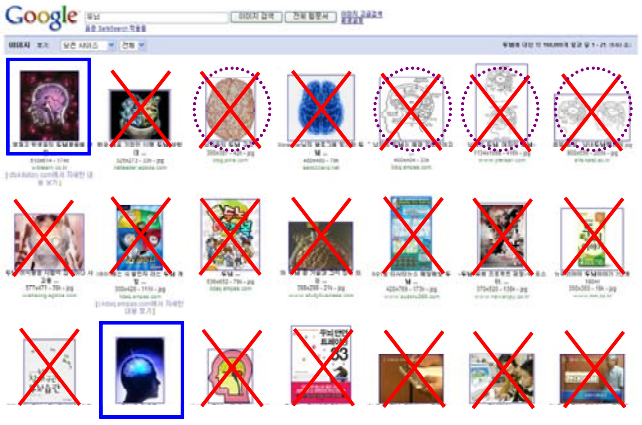


Fig. 3. Search Result about Query 'Brain'

As shown at **"Fig.3"**, web user 'A' satisfied with only 2 search results. If other web user who works for medical field searches image for writing an essay, he will select black dotted line result. If at that time search result is satisfiable, he/she can recommend his/her search results when other web users who have similar occupation input same query with same purpose. Therefore, if one of attributes correspond to Relationship Value like 'occupation', is similar with other web users, they can easily find the information as they want.

Like this, if ever-changing interest or search intent could be analyzed in Social Networks and applied to search, it can help improve search efficiency and reliability. For this, correlation analysis between Social Relation Value and Search Pattern is very important. If web users who have similar characteristic and tendency have same Search Pattern, it could be used effectively to search in Social Networks' environment. Namely, the fact that Search Pattern among web users who have High-Social Relation Value is very meaningful element because it is able to applied to improvement of search efficiency and reliability.

A. Calculation of Relation Value

First of all, to calculate Social Relation Value, factors and attributes which can characterize each individual are need to be defined and selected. Also because web users' interest is different, various topics which correspond to their interest should be selected. And then, we can calculate Social Relation Value by algorithm giving association between attributes and topics.

1) Factors and Attributes Selection

It is very important to determine how to calculate Social Relation Value according to topics. To calculate Social Relation Value among web users, 'N' Factors(Similarity, Access, Intimacy, Adjacent, Rewardingness,. etc.) are used. Each factor has detail attributes, which is personal profile and SNS refers. Such 'N' Factors are separated into total 'L' Attributes including representative SNS site Friendster, Orkut, Facebook and Cyworld's general attributes and other additional attributes.

2) Topics Selection

Social Relation Value is changed according to topics that web users have interest in Social Networks. For Example, we assume that there exist a Social Network and can see Social Relation Value based on Topic 'Arts' and 'Sports'. Structure of two networks and features(for example, link strength, existence of connection, relation,. Etc) among web users will be different according to topics. In other words, although web users have same interest and characteristic generally, their interest and characteristic will be changed according to topics. Due to such characteristic, 'M' different topics are selected to calculate Social Value. After that, on each subject, Social Relation Value among 'I' and other web users is calculated.

B. Matrix of Association among Attributes and Topics

"Fig.4" shows {'L'(Attributes) x 'M'(Topics)} Association Matrix. Among the Attributes corresponding to each web users' Factors, Attributes which are affected by topics should be defined first to calculate web users' Social Relation Value on the Social Network. That is, if Attributes are affected by a certain topic, association is given to that attributes.

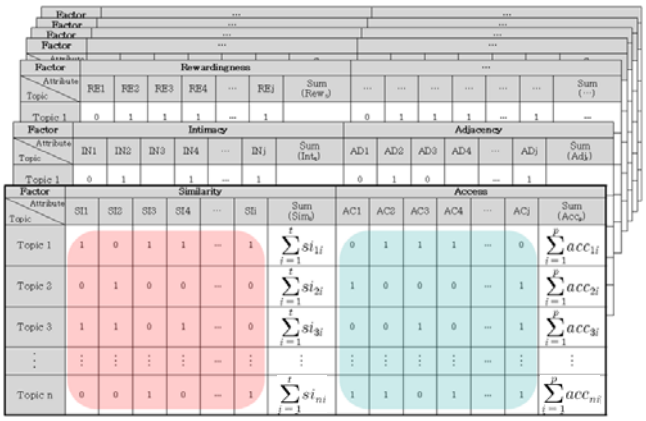


Fig. 4. 'L' X 'M' Association Matrix(topics_attributes)_(Topic Sensitive)

C. Algorithm for Calculating Relation Value

Network analysis is expressed as binary type(indicate the relationship as 0-none, and 1-yes), sometimes structural equivalence could be estimated by similarity-difference value designed for such data type. **"Fig.5"** is an example for calculating similarity value, shown by cross table.

user 1 : 0 0 0 0 1 1 1 1 1 1
user 2 : 0 0 0 1 1 1 1 1 1 1
user 3 : 0 0 0 0 0 0 0 1 1 1

item 1	item 2	
	exist	not exist
exist	a	c
not exist	b	d

Fig. 5. Division into two parts data for Calculate a Similarity Characteristic

Based on {a,b,c,d} 4 items, calculate all binary data. In above example, in case of 'user 1' and 'user 2', a=6,b=1,c=0,d=3. And case of 'user 2' and 'user 3', a=3,b=0,c=4,d=3. In this case, simple matching Similarity Value analysis is used for method to analyze network. That is, as shown on formula (4), it is measured using number which is matched with all numbers. In terms of this measurement, case of corresponding with both 'exist' and 'not exist' are concluded in numerator.

$$Similarity(X, Y) = \frac{a+d}{a+b+c+d} \qquad (4)$$

Referring to formula (4), we suggest the algorithm to calculate Relation Value among web users according to topics.

We can extract the Row Vector related to Topic and Attributes from Association Matrix, $T_i = [1, 0, 0, 0, 0, \dots, 0]$, where the value '1' appears just only in the i^{th} column, $A_{j\_topic_i} = [1, 1, 0, 0, 1, \dots, 1]$, where the value '1' appears in the j^{th} column. The numerical formula of Social Relation Value among web users at $Topic_i$ is expressed below as (5).

$$SRV_{me\_user_r} = \alpha \cdot \frac{t_i \cdot \sum_{j=1}^{l_1} si_j}{SI_i} + \beta \cdot \frac{t_i \cdot \sum_{j=1}^{l_2} acc_j}{ACC_i} + \dots\dots + \gamma \cdot \frac{t_i \cdot \sum_{j=1}^{l_k} inti_j}{INTI_i} \qquad (5)$$

- $\alpha + \beta + \dots + \gamma = 1$, α, β, γ : Balance Factors
- $A_j = \{si_1, si_2, \dots, si_{lp}, acc_1, acc_2, \dots, acc_{lq}, inti_1, inti_2, \dots, inti_{lr}\}$, for each i = 1, 2, 3, … m
- si_j, acc_j, $inti_j$: Attributes that coincide with web users in case of applying Row Vector of Association Matrix
- SI_i, ACC_i, $INTI_i$: Sum of Attributes which have association about Topics

IV. EXPERIMENT AND EVALUATION

This section take comparison experiment about web users' search satisfaction to verify whether there exist correlation between Social Relation Value and Search Pattern according to topics or not. In other word, this experiment is to verify the fact that if web users have High-Social Relation Value

then their Search Pattern is similar, too. The process of experiment and evaluation is as follows.

- Step1 : The members of the experiment and data set composition
- Step2 : Calculate the relation value according to topics(① ODP-based 15 topics selection, ① 2 factors and 11 attributes selection, ③ 15x11 Association Matrix giving association among each topic and attributes)
- Step3 : Analysis on search pattern(① Representative 'query' selection according to topics, ② Search by 'www.google.com', ③ Gather search result of web users, ④ Analyze satisfaction of search result(search pattern) based on relation value)
- Step4 : Analysis correlation between relation value and search pattern(① Search pattern among the similar web users who have a high relation value, ② Compare the graph of relation value with search pattern

A. *Data Set*

We made a Social Network using members of SNS(friendester, orkut, facebook, cyworld,. Etc.) that we have account. **"Fig.67"** shows Social Networks which is used for experiment and evaluation. Total 1,000 members were selected as members of the experiments. 250 members are directly connected with author(1-Hop Distance), and the others are indirectly connected with author(2-Hop Distance). And we made 1,000 different Social Networks with each web user in the center. To obtain reliability, personal profiles of each member are composed of people who live in various country(for example, Korean, Japanese, Thailander, Chinese, American, Canadian, Austrian, Spanish,. Etc.). And also, data set related to profiles(Gender, Relationship, Birth, City, Language, Interest,. Etc.) of members are very diversify.

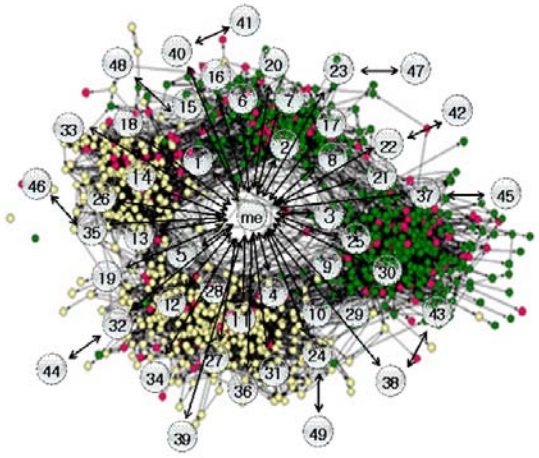


Fig. 6. Social Network for Experiment

B. *Calculation of Social Relation Value*

1) *Factors and Attributes Selection*

We limit 'N' Factors and 'L' Attributes to '2' and '11' referring to 'Orkut' for experiment. We select Similarity and Access on '2' Factors. And also we select mandatory input item of ordinary field on '11' Attributes. In case of Orkut, profile which is equal to detail attributes, is classified into '5' parts : Common, Friendship, Address, Occupation, Personal Information. It identifies web users through e-mail address based on Name, Gender, and Nation. If web users want to various service including finding friends service, they have to input their profile about other information section. Important information is being provided at common part, other '4' part(Friendship, Address, Occupation, Personal Information) is being automatically amended according to common part. User Characteristic, an element to calculate user Social Relation Value, is belong to Common Part which is mandatory for all web users. **"Table 1"** shows '11' Attributes.

TABLE I. 11 ATTRIBUTES

Attributes	Content
Gender	Male / Female
Relationship	Solo / Married / Dating
Birth Date	1920 ~ 1990
City	Direct Input
Language	Basic One
High School	Direct Input
College	Direct Input
Company, Organization	Direct Input
Field of Interest	Can choose from Friends, Partner for leisure activities, Business Network, Date
Registering Friends	Mutual Friends, or Non-Registered
Coincidence status about joined Group	Status among web users

2) *Topics Selection*

We limit 'M' Topics to '15' Topics for experiment. '15' different topics are from Orkut ODP [5] (Open Directory Project), which classified all subjects into '15' communities.

ODP is most comprehensive web directory to obtain ideal search result, and **"Fig.7"** shows its structure. Thus, it could be considered that all topics are included in one of '15' directories. Each directory divided into classes like **"Table 2"**.

[5] www.dmoz.org

Fig. 7. 15 Directory of ODP

TABLE II. 15 TOPICS

Topics	Classification
Arts	Science/History, Culture Life/Local Community, Arts/Entertainment, Music
Business	Alumni/alma mater, Business, Company
Computers	Computer/Internet
Games	Game
Health	Health/Well Being/Exercise, Food/Beverage /Wine
Home	Family/Home, Pet/Animal, Food/Beverage/ Wine, Hobby/Handicraft
Kids and Teens	Game, Pet/Animal, Music
News	Health/Well Being/Exercise, Science/History, Arts/Entertainment, Government/Politics
Recreation	Health/Well Being/Exercise, Recreation/ Sports, Travel, Hobby/Handicraft
Reference	Business, Travel, School/Education
Regional	Nation/Region, City/village, Travel
Science	Science/History
Shopping	Health/Well Being/Exercise, City/village, Pet/Animal, Food/Beverage/Wine, Cars, Fashion/Beauty
Society	Nation/Region, City/Village, Romance/Date, Culture Life/Local Community, Religion/Faith
Sports	Health/Well Being/Exercise

3) Calculation of Social Relation Value

We limit 'L' X 'M' Association Matrix size to '11' X '15' for experiment. **"Fig.8"** shows '11' X '15' Association Matrix.

Factor	Similarity										Access		
Attribute / Topic	Gender	Relationship	Birth Date	City	Language	High School	College	Company	Interest	Sum (SI)	Registering Friend	Join Group	Sum (AI)
Arts	1	0	0	1	1	1	1	1	1	7	1	1	2
Business	0	0	1	1	0	1	1	1	1	6	1	1	2
Computer	1	1	0	1	1	0	0	0	1	5	1	1	2
Game	1	0	1	1	0	0	0	0	1	4	1	1	2
Health	1	0	1	1	0	0	0	1	1	5	1	1	2
Home	1	1	1	1	1	0	0	1	1	7	1	1	2
Kids & Teens	1	1	1	1	1	0	0	0	1	6	1	1	2
News	0	0	0	1	1	1	1	1	1	6	1	1	2
Recreation	0	1	1	1	0	1	1	1	1	7	1	1	2
Reference	0	1	1	1	1	1	1	1	1	8	1	1	2
Regional	1	0	0	1	1	0	0	1	1	5	1	1	2
Science	1	0	1	1	1	1	1	1	1	8	1	1	2
Shopping	1	1	1	1	0	0	0	1	1	6	1	1	2
Society	1	0	1	1	1	0	0	1	1	6	1	1	2
Sports	1	1	1	1	0	0	0	0	1	5	1	1	2

Fig. 8. '11' X '15' Association Matrix$_{(topics_attributes)_(Topic\ Sensitive)}$

a) Assigning Association between Factors and Attributes(Assign positive value : '1')

Accessibility could be calculated from whether members' connection exists or not. Borgatti said that club joining and friend registration is a kind of accessibility(access) [21]. Also many sociologists and psychologists studied link and interaction among people in a virtual space and social phenomenon. Of course this definition varies according to time and location, so this paper gave connection based on empirical elements and academic research advanced until now. Namely, we assign the Attributes {Gender, Relationship, Birth Date, City, Language, High School, College, Company, Interest} to Factor {Similarity} and the Attributes {Registering Friend, Join Group} to Factor {Access}.

b) Assigning Association between Attributes and Topics(Assign positive value : '1')

Hagel said, "Members of virtual community have a goal to share their interest or expertise about certain topic widely by interaction among them [19]." It could be said that user attributes like interest field, friend registration, joining a group is related to topics as previously stated.

Also, the fact "people who have similar experience in community aim at generating meaningful relationships based on their own experience" means that there is a link among people like alumni or living nearby about certain topic.

M.K.Smith mentioned that gender, human race, and age is the category which distinguish the main differences, and its characteristic appears in their own culture and society [20]. That is, 'Gender', 'Human Race', and 'Age' is related with culture or society like 'Arts', 'Home', 'Regional', 'Science', 'Society' category.

'1' is positive count item which gives certain value, and '0' is negative count item. Let's take a look at connection between attribute 'Gender' and each topic on **"Fig.8"**. 'Gender' is a related attribute which affects to topic 'Arts' because it was given '1' by 'Arts'. But in terms of 'Business', it was given '0'. So 'Gender' doesn't have relationship with 'Business'.

4) Algorithm for Calculating Relation Value

The method to calculate Relation Value about 2 Factors among topic-specified web users is as expression (6).

$$SRV_{me\_user_t} = \alpha \cdot \frac{t_i \cdot \sum_{j=1}^{l_1} si_j}{SI_i} + \beta \cdot \frac{t_i \cdot \sum_{j=1}^{l_2} acc_j}{ACC_i} \quad (6)$$

Relation Value of j couple web users for certain topic Ti is normalized sum of association and accessibility. At this time, constant k is a balancing factor of 2 Factors Similarity and Access.

5) Calculation of Relation Value

After generation of a Social Network with 'I' at the center, Relation Value among web users by topics was calculated using algorithm and results are as **"Fig.9"**.

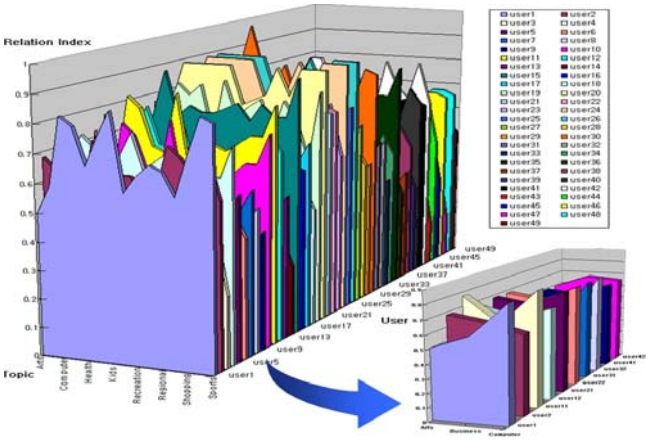


Fig. 9. Social Relation Value of whole web users

Like **"Fig.9"**, Social Relation Value among web users is changed according to topics. Namely, although web users, who have Low-Social Relation Value in a certain topic, their Social Relation Value could be high in other topics. So Social Relation Value need to be calculated according to topics.

C. Analysis of Search Pattern

1) Query Selection by Topics

5 query items per topic was selected according to user's high preference items from the ODP's 15 directories. 75 query is like **"Table 3"**.

Google[6] search engine was used to finding search result, and page was limited within 3 pages. According to recent study on web user's search range, 88% of web users using search engine look just under 3 pages [22]. So, when there is no information within 3 page, other query was input in this experiment.

TABLE III. '5' QUERY EACH 15 TOPICS

Topics	Query
Arts	Song, Theater, Character, Literature, Design
Business	Communication, Tour, Travel, Auction, Part time job
Computers	Internet, Online Game, Chatting, Computer, Finding Friends
Games	Video, Simulation, Internet Game, Game, Play
Health	Hospital, Oriental Medicine, Mental Health, Diet, Food
Home	Marriage Information, Infant Care, Recipe, Moving, Pregnancy
Kids and Teens	Infant Health, Kids Movie, Fairy Tale, Kids Journal, Education
News	Local Newspaper, Broadcasting, Internet broadcasting, Broadcasting Center, Forecast
Recreation	Free Gift, Celebrity, Movie, Online Bookstore, News
Reference	Study Abroad Agency, Foreign Language, Online Study, Public Office Examinations, Kindergarten
Regional	Travel Information, Fishing, Cars, Pet, Book
Science	Science, Social Science, Science Journal, Education, Environment
Society	Religion, Law Information, Politics, Issues, Social Movement
Shopping	Flower, Fashion, Food, Shopping, Home Appliances
Sports	Sports, Exercise, Ball Game, Sports Match, Exercise for Health

2) Analysis of Search Result

In this section, experiment focused on verifying this fact : Like **"Fig.11"**, web users who is similar to my characteristic and tendency, in other words, High-Social Relation Value user by each topic, has similar Search Pattern. To obtain results, coincidence rate of the members of experiment was used. Coincidence rate is a chance to select the same page with 'I' when certain experimental query was entered. In every 5 topics, it was made to select 3 satisfied web pages and coincidence rate is able to calculated on this way.

For example, let's calculate coincidence rate of search between 'I' and 'user 1' about topic 'Arts'. Input 'Song', which is selected from topic 'Arts' into google search engine, and confirm web pages until 3 page. If the number of searched 'url' and 'snipet' is 'n', number from 1st ranked 'url' and 'snipet' to 'n'th items, and make a Set. In state of S={1,2,3,4,....n}, if 'I' select 1st, 2nd, 3rd 'url' and 'snipet' and 'user 1' selects 1st, 5th, 7th, coincidence rate is {(1/3)*100}. Like this, calculate coincidence rate between 'I' and whole 49 members' selection.

"Fig.10" expresses two facts : Selection by web users according to search result and search result about each representative query from 'Arts', 'Business' and 'Computer'. Like this method, coincidence rate for '5' queries from each '15' topics, total '75' queries was confirmed by each web users.

Section		User					
topic	query	me	user1	user2	user3	···	user49
Arts	song	1,3,9	1,4,10	1,5,9	1,2,5	···	1,2,3
	theater	1,4,13	2,5,13	1,13,26	2,5,10	···	3,5,7
	character	1,13,16	1,13,20	13,16,25	1,7,9	···	1,12,13
	literature	3,6,8	3,6,9	1,3,10	1,3,7	···	3,5,7
	design	2,4,7	2,5,7	2,4,5	1,2,5	···	1,5,8
Business	communication	4,7,16	4,5,9	1,4,7	1,2,4	···	3,4,7
	travel	2,5,9	2,5,7	5,6,9	1,3,5		2,5,7
	bank	10,11,13	2,10,11	1,10,11	1,2,3	···	2,4,6
	auction	1,2,7	1,3,4	1,2,3	1,4,5		2,5,7
	arbeit	1,4,7	1,5,7	1,3,9	1,2,3	···	1,2,4
Computer	internet	2,4,10	1,2,4	2,3,17	1,3,5		1,2,5
	onlinegame	4,8,17	2,4,8	1,7,8	1,2,4	···	3,4,6
	chatting	1,3,4	1,2,3	2,3,8	2,3,6		1,3,4
	computer	8,12,21	2,8,12	1,6,8	1,2,8	···	1,7,8
	computergame	3,4,14	2,3,4	3,4,13	1,2,3		2,45

Fig. 10. Selection of Search Result

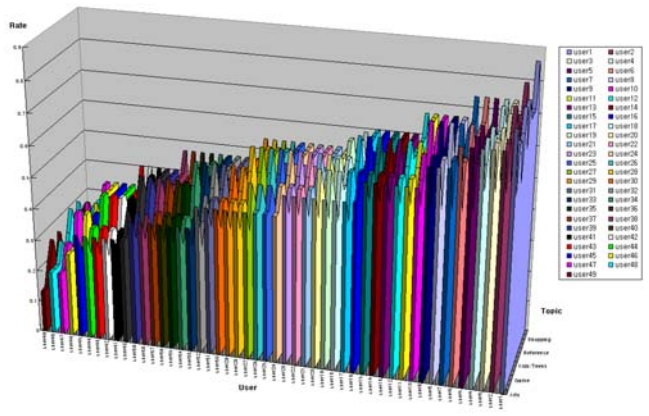


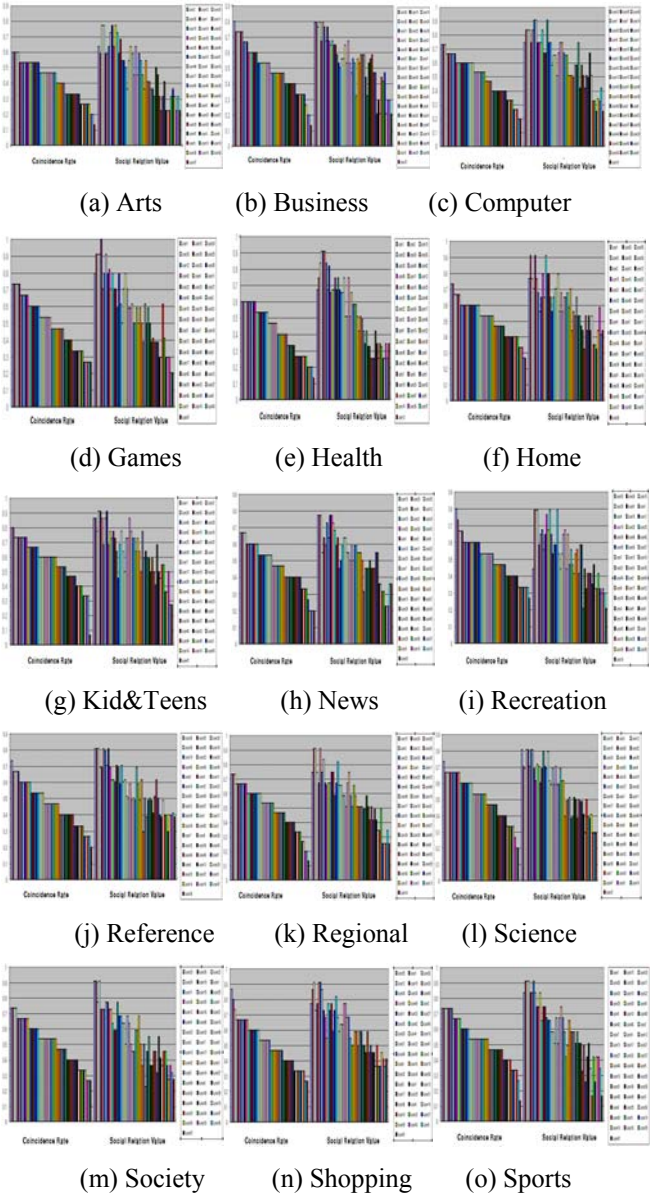
Fig. 11. Coincidence Rate of Search according to Topics

[6] www.google.co.kr

"Fig.11" expresses search coincidence rate between 'I' and other web users. Like **"Fig.11"**, Coincidence Rate related to Pattern among web users is changed according to topics, too. Therefore we can know that web users' Search Pattern is changed according to web users' interest or inclination.

D. Analysis of Correlation between Social Relation Value and Search Pattern

This section is the last phase of paper. In this section, Relation Value and web users' Search Pattern is compared so as to verify whether mutual correlation is exist or not. If there is High-Coincidence among high Relation Value web users, they have mutual correlation.



(a) Arts (b) Business (c) Computer

(d) Games (e) Health (f) Home

(g) Kid&Teens (h) News (i) Recreation

(j) Reference (k) Regional (l) Science

(m) Society (n) Shopping (o) Sports

Fig. 12. Correlation between Social Relation Value and Coincidence Rate of Search

"Fig.12" shows correlation between Social Relation Value and Coincidence Rate of Search among web users. As you can see from **"Fig.12"**, preferences of search result(or Coincidence Rate of Search) is similar according to Social Relation Value among web users classified by topics. From this, you can find a fact that web users who have High-Social Relation Value, in other words, who have strong link strength by topics, have similar Search Pattern. So we can confirm the possibility to improve search efficiency and reliability in Personalized / Social Search based on the fact : among web users who have High-Social Relation Value, that is, who have strong link strength according to personal Attributes have similar Search Pattern.

V. CONCLUSION AND FUTURE WORK

In this we analyze Correlation between Social Relation Value and Search Pattern in a Social Network.

For this, we construct a Social Network, and then, we reconstruct Social Networks based on Social Relation Value calculated using characteristic and tendency among web users according to topics. After that, we conduct comparison experiment on web users' search result satisfaction by queries. As a result of this experiment, we can confirm a result that among web users who have High-Social Relation Value, that is, who have strong link strength according to personal attributes have similar Search Pattern. This fact is very important and meaningful when it is applied to Social Networks, because if web user obtain Search Pattern(including usage query, click through data,. Etc.) from other user who have High-Relation Value and similar characteristics and tendency, he/she could get accurate information without any unnecessary time wasting. That is, such fact is applied to search algorithm, it could be possible to improve search efficiency and reliability in personalized search.

In the future, additional research on how to give association into relationship between information and user attributes, and topic and user's attributes, will be necessary. Also it will be needed to search how to improve searching efficiency and reliability by analyzing web users' behavior pattern and user's attributes through an implementation of the idea in a Personalized/Social Search.

VI. REFERENCES

[1] K. Tanaka. : Usage of Web Knowledge and User Behaviors for Web Search. In Proceedings of KJDB 2007, 2-19, 2007

[2] B. J. Jansen, A. Spink and T Saracevic. : Real life, real web users, real needs : A study and analysis of user queries on the Web. Information Processing and Management, 36(2) : 207-227, 2000

[3] J. R. Wen, J. Y. Nie and H. J. Zhang. : Clustering user queries of a Search Engine. In Proceedings of the 10th International World Wide Web Conference, 162-168, 2001

[4] J. Carroll and M. Rosson.: The paradox of the active user. Interfacing Thought: Cognitive Aspects of Human-Computer Interaction, 80-111, 1987.

[5] R. Kraft and C. C. Chang : Search with Context. In Proceedings of the 11th International World Wide Web conference, 477-486, 2002.

[6] JiSu Kim. : Human Network, Policy of Information and Communication, a volume 16, an issue 16, 2004

[7] T. H. Haveliwalla.: Topic-Sensitive PageRank., In Proceedings of the 11th International World Wide Web Conference, 517-526, 2002

[8] F. Qiu and J. Cho. : Automatic Identification of User Interest For Personalized Search. In Proceedings of the 15th International World Wide Web Conference, 727-736, 2006.

[9] A. Broder.: A Taxonomy of Web Search. SIGIR Forum. 36(2), 3-10, 2002

[10] U. Lee, Z. Liu and J. Cho.: Automatic Identification of User Goals in Web Search. In Proceedings of the 14th International World Wide Web Conference, 391-400, 2005

[11] D. E. Rose and D. Levinson.: Understanding User Goals in Web Search. In Proceedings of the 13th International World Wide Web Conference, 13-19, 2004

[12] M. V. Vieira, B. M. Fonseca, R. Damazio, P. B. Golgher, D. de Castro Reis and B. Ribeiro-Neto.: Efficient Search Ranking in Social Networks. In Proceedings of the 16th ACM Conference on Information and Knowledge Management (CIKM) '07, 563-572, 2007

[13] E. Spertus, M. Sahami and O. Buyukkokten.: Evaluating Similarity Measures: A Large-Scale Study in the Orkut Social Newtork. In Proceedings of the 11th ACM SIGKDD international conference on Knowledge Discovery in Data mining, 678-684, 2005

[14] L.A. Adamic, O. Buyukkokten and E. Adar.: A social network caught in the Web. First Monday, 8(6), 2003

[15] D. Fono and K. Raynes-Goldie.: Hyperfriendship and Beyond: Friends and Social Norms on LiveJournal. In Internet Research Annual Volume 4: AIOR Conference, 91-103, 2006

[16] C. Dwyer: Digital Relationships in the 'MySpace' Generation: Results From a Qualitative Study. In Proceedings of the 40th Annual Hawaii International Conference on System Sciences (HICSS), 19-28, 2007

[17] C. Dwyer, S. R. Hiltz and K. Passerini.: Trust and A comparison of Facebook and MySpace. In Proceedings of the 13th Americas Conference on Information Systems (AMCIS), 2007

[18] YongHak Kim. : Social Network Analysis, ParkYoung Company, 2003

[19] Hagel, Jhon and Arthur G. Amstrong. : Net Gain ; Expending Markets through Virtual Communities, HBS Press, 1997

[20] Marc A. Smith. : Communities in cyberspace, 11. 1998.

[21] S. P. Borgatti and R. Cross.: A Relational View of Information Seeking and Learning in Social Networks. Management Science, 49(4): 432-445, 2003

[22] IProspect, Search Engine User Behavior, 2006. http://iprospect.com/WhitePaper_2006_Search EngineUser Behavior.pdf

[23] Y.A. Kim, G.W. Park. : Topic-Driven SocialRank: Personalized search result ranking by identifying similar, credible users in a social network, Knowledge-Based Systems Volume 54, December 2013, Pages 230-242

[24] GunWoo Park , SooJin Lee , SangHoon Lee, To Enhance Web Search Based on Topic Sensitive_Social Relationship Ranking Algorithm in Social Networks, Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology, p.445-448, September 15-18, 2009

[25] Young-an Kim and Gun-woo Park. : Topic Sensitive_Social Relation Rank Algorithm for Efficient Social Search, The Journal of Korean Institute of Communications and Information Sciences, Vol. 38, No. 5, pp. 385~393, 2013.

[26] G.W. Park, J.W. Oh, S.H. Lee. : Personalized and Social Search by Finding User Similarity based on Social Networks. The KIPS Transactions, Part D, Vol.16 No.5, 2009.

[27] ValeriaOrso, Tuukka Ruotsalo, Jukka Leino, Luciano Gamberini, Giulio Jacucci. : Overlaying social information: The effects on users' search and information-selection behavior, Information Processing & Management, Volume 53, Issue 6, November 2017, Pages 1269-1286

[28] Omair Shafiq Reda Alhajj John G.Rokne : On personalizing Web search using social network analysis, Information Sciences, Volume 314, 1 September 2015, Pages 55-76

[29] Vicenç Torra, Termeh Shafie, Julián Salas. : Data Protection for Online Social Networks andP-Stability for Graphs, IEEE Transactions on Emerging Topics in Computing (Volume: 4, Issue: 3, July-Sept. 2016)

[30] Kyoungsoo Bok, Jongtae Lim, Minje Ahn and Jaesoo Yoo. : A Social Search Scheme Considering User Preferences and Popularities in Mobile Environments, KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS VOL. 10, NO. 2, Feb. 2016

[31] Shubham Goel Ravinder Kumar. : Folksonomy-based user profile enrichment using clustering and community recommended tags in multiple levels, Neurocomputing Available online 7 August 2018