

Estimating Purposes of Users in Social Networking Service Public Contents

ンヴンギ バシリサ[†] 岩井原瑞穂[‡]

[†] 早稲田大学大学院情報生産システム研究科

〒808-0135 福岡県北九州市若松区ひびきの 2-7

E-mail: [†] basilisam@akane.waseda.jp, [‡] iwaiharu@waseda.jp

Abstract: We call as publicly active users (PAUs) those who actively post contents to publicly-visible locations in social networking services (SNSs). In this paper, we explore posting purposes of PAUs (open-post purposes) through inspecting their public contents. We sample PAUs from five public groups of different interests and one set of general users whose interests are unknown. We can safely assume that each PAU is sharing the group's interest by posting to the interest group. We manually classify posting purposes based on types of activities into 19 purpose items, and five high-level categories of open-post purposes: promoting self, promoting own activities, promoting opinion, propagating quotes, and propagating news. We further evaluate for independence of these five OP-purposes through statistical analysis, which results into promoting self, propagating news quotes, and propagating sociality information. We find that PAUs have the following five purpose items: posting group or fellow photos (83% PAUs), propagating positive social news (83% PAUs), posting self photos (77% PAUs), propagating positive quotes (65% PAUs), and posting personal details (60% PAUs). These findings suggest that PAUs are relying more on posting about their close people and propagating news (lifestyle, technology, entertainment) than either posting self photos or sharing photos containing their favorite messages. However, correlation analysis reveals no relation between PAUs posting photos of themselves and their close people, and propagating of information (news, quotes) from SNS users. We also introduce post-frequency level of 0 to 2 to represent how much degree a PAU exhibits a particular purpose item.

Keyword: Publicly active users, user behavior analysis, purpose analysis, SNS user contents.

1. INTRODUCTION

Uses and gratifications approach is commonly used to investigate motivations of SNS users [3,4]. To assess SNS uses and gratifications, surveys are usually conducted because most of uses are privately performed and gratifications are latent. Publicly active users (PAUs for short) publicly publish contents such as video, photos, text, and links to articles via SNS. Upon observing contents, SNS users often find that PAUs publish their photos to promote their identity. Also, PAUs share photos containing words to propagate quotes or news. These few examples give an insight that PAUs are publishing contents because they have purposes to do so. In [2], three high-level purposes of 2012 US presidential tweets are identified: oppose, favor, and other.

In this paper, we explore posting purposes (open-post purposes or OP-purposes for short) of PAUs by inspecting their SNS public contents. From identified OP-purposes, we construct a purpose model by classifying purpose items based on types of activities. After that, we refine the purpose model by statistical analysis. Contrary to [2], we are identifying multiple purposes on one post. Also, we focus on what are typical purposes, which drive a PAU to

post openly.

Identifying OP-purposes and revealing predictors of these purposes are useful for SNS providers to classify users and their posts based on their latent purposes.

2. PURPOSE MODEL CONSTRUCTION

2.1 Data Collection

We use two methods to sample user IDs. In the first method, we generate random user IDs and filter out user IDs that are invalid. We term users with random IDs as *general users* and we assume that their interests are unknown. In the second method, we sample user IDs from public groups. We infer users' interest from the group that they are participating. We focus in public group(s) on which users share one of the four topics, namely business, politics, animals, and music. Therefore, we use the SNS search function to search for each topic in order to find the public group(s). If one or more public group are returned, we roughly observe the contents of each group for activeness and homogeneity to judge whether the group's contents reflects the group's interest. Such observation is necessary because we safely assume that users belonging to a group tend to share the common interest of the group.

Activeness means the group contains members who are actively or continuously participating to discussions. *Homogeneity* means the contents of the group are mostly related to the group name. If no public group is returned, we look for similar topics(s) and again search until we find a public group, which contents are reflecting the group's interest.

By the above procedure, we find five public groups as seen in Table 1 and collect user IDs from these groups (column "Collected user IDs" in Table 1). These user IDs are collected from groups' posts and visible comments that are published from newest (April 2015) to oldest (March 2015). *Visible comments* are comments, which do not require a user to click a link in order to view them. Then, we extract user IDs from those posts and comments and remove duplicate user IDs.

2.2 Selecting PAUs

To identify OP-purposes consistently, it is necessary to select users who are active in publicly posting contents via SNS. Therefore, we apply conditions to the collected user samples in Table 2 in order to select publicly active users (PAUs). We set the following conditions for a user to be a PAU: i) post frequency of the PAU. We set the criteria such that the user should have at least one post per month in year 2015 (January - May), the user should have at least one post in year 2014, and the user posted at least 15 posts between January 2014 and May 2015. ii) The language used on PAU's posts should be English. English language is essential to enable collaboration of raters with different ethnicity when analyzing PAUs' contents. Nevertheless, we select a PAU who has a few posts written in other languages when translation option is available.

Table 1 shows the counts of user IDs and PAU IDs that are sampled from five public groups and one set of general users. We process 2467 user IDs and 350 PAU IDs matched our criteria on validity. We observe that proportion of general PAUs is smaller compared with that of PAUs from public groups (see Table 1). Upon processing random IDs for general PAUs, we found the following: i) most of sampled general users have not satisfied the conditions to be PAUs, and ii) some IDs are for public pages and groups.

Table 1. Data set summary

Source of data	Group's interest	Collected user IDs	Processed user IDs ^a	PAUs IDs (%) ^b
General	Unknown	6993	1658	100 (6.0)
Public group names				

Hiking With Dogs	Pet	952	115	46 (40.0)
Jazzmasters&Jaguars	Music	565	249	49 (19.7)
Constitutional Patriots	Political-enthusiast	175	172	72 (41.9)
Like For Like Promote Your Business	Trade	544	250	74 (29.6)
Promote My Business	Trade	517	23	9 (39.1)
All	5	9746	2467	350 (14.2)
Note: ^a Number of users whose pages are checked for meeting the conditions to be PAUs. ^b Proportion of processed users in percentage that are PAUs.				

2.3 Constructing the Purpose Model

In this section, we explain about how we construct the purpose model from purpose items based on types of activities. Then, we will explain about how the purpose model is refined using statistical approach. Table 2 shows the description of PAU sets that are used in this section.

2.3.1 Constructing Purpose Model Manually

First we discuss listing up candidate open-post purposes (OP-purposes). We adopt the following steps: (1) Identifying candidate OP-purposes from public contents, (2) filtering out OP-purposes that are rarely observed, (3) refining candidate OP-purpose documented in Step 1 to remove duplicated or similar OP-purposes, and (4) clustering of similar purpose items into high-level categories.

(1) Identifying candidate OP-purposes from public contents. We identify OP-purposes of first 185 PAUs in Table 2. For each PAU, we write an outline of any OP-purposes that can be identified from his/her publicly observed contents, dated between January 2014 and May 2015. There exist posts that are difficult to interpret their respective OP-purposes. For example, posts with words such as "Good night!!", "Hello friends!", "Noooo", or posts with pictures of places (parks, beaches, etc.), and objects (buildings, sculpture, etc.). These posts are therefore not counted.

(2) Filtering out OP-purposes that are rarely observed. This step is necessary for statistical analysis, to avoid OP-purposes having only a few PAUs. We remove OP-purposes in Step 1 that are identified in not greater than 20 PAUs.

(3) Refining candidate OP-purposes documented in Step 1 to remove duplicated or similar OP-purposes. We manually classify OP-purposes based on types of activities into 19 purpose items. As an example, Table 3 shows 2 out of 19 purpose items.

(4) Clustering similar purpose items into high-level

categories. We identified the following five high-level categories: promoting self, promoting opinion, promoting professional activities, propagating quotes, and propagating news. *Promoting self* is associated with a PAU's intention to expose his/her personal life via SNS within the contexts of themselves, families (including pets), friends, non-professional activities, and possessions. *Promoting own activities* is associated with a PAU's intention to expose his/her social and professional activity commitments via SNS. *Promoting opinion* is associated with a PAU's intention to express his/her thoughts on news or encountered situations via SNS. *Propagating quotes* is associated with a PAU's intention to share words from other sources, which reflect his/her taste. *Propagating news* is associated with a PAU's intention to share information about events that happens in a society.

The hierarchical structure of purposes is shown in Figure 1, whose purpose symbols are defined in Table 4. In Figure 1, a *promoting purpose* is a purpose of a PAU posting for apparently promoting or advertising his/her own agenda. On the other hand, a *non-promoting purpose* is the complement of the promoting purpose such that promotional/advertising purpose is not explicit. Public postings can convey promoting purposes in various forms, either implicitly or explicitly. Implicit promoting purpose can be exemplified as propagating others' posts and news that match with the PAU's purpose, as well as, a PAU sharing a snapshot, which contains details of upcoming music performance of his/her favorite musician. However, in this study we label a post with the promoting purpose when the PAU publicly posts contents that are directly linked to him/her, or share such contents from other sources (public page, website, other SNSs). One public post can be labeled with multiple purposes. For example, a PAU posts his/her comments on a shared news article. In this case, the post is labeled as both promoting (explicitly expressing thoughts on news) and non-promoting (propagating news). However, if the PAU shares another's post or content but not giving his/her own comment, we regard that the PAU's intension is implicit and we label the post as non-promoting only.

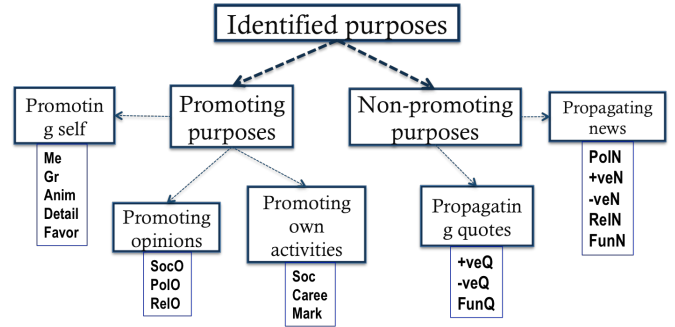


Figure 1. Purpose model constructed from identified purpose items

Table 2 Summary of research sets and their descriptions

PAU sets	Sec(s) ^a	Description
185	2.3.1	First ^b PAUs out of 350 PAUs.
200	2.3.2; Steps 2-3	PAUs who are randomly selected out of 350 PAUs. In Step 2, rater1 crosscheck public contents of the first 100 PAUs, while rater 2 crosscheck that of the remaining 100 PAUs.
150	2.3.2; Step 3	These are remaining PAUs after selecting 200 PAUs above.
350		These are all PAUs that match our criteria for validity in Section 2.2.
Note : ^a Sec(s); Section(s). ^b 54 are general PAUs, 46 from "Hiking With Dogs", 41 from "Jazzmasters&Jaguars", 18 from "Constitutional Patriots", 17 from "Like For Like Promote Your Business", and 9 from "Promote My Business"		

Table 3. Purpose identification table

Symbol	Purpose item	PAUs' observed content that constitute purpose items
Me	Posting self photos	Photos of PAU either at home or visiting several places (restaurant, mall, beach, holiday, etc.)
+veN	Propagating positive social news	Social news (lifestyle, technology, entertainment, sports) shared by PAU to provide useful self-education, awareness, or enthusiasm.
Note: To identify PAU's photos, at least one photo should have description about PAU. News is shared from other users, other SNSs, or websites as articles, videos, photos, and links.		

2.3.2 Post-frequency Levels on Purpose Items

Now we introduce the *post-frequency level* of 0 to 2 to represent how much degree a PAU exhibits a particular purpose item, as follows:

(1) Assigning the post-frequency level of 0 to 2 to a PAU on each purpose item, as follows: i) level 0 if public contents that constitute the purpose item are not observed, ii) level 1 if public contents that constitute the purpose item are observed between 1 to 5 times, and iii) level 2 if public contents that constitute the purpose item are

observed at least 6 times. If no more than five posts are observed per PAU, then only levels 0 and 1 are used.

(2) Crosschecking by raters. Here, *raters* are colleagues (not authors of this paper) with expertise in the social network field. We use two raters, for assuring accuracy of our initial inspection of public contents. Raters are also instructed to give comments where necessary. Table 2 shows details of PAUs that are used in this step, and crosschecked by raters. Using the table (*purpose item table* for short) that describes guidelines of each purpose item as a guideline and post-frequency levels in Step 1, we select 200 PAUs and first inspect each PAU's public contents that match with each purpose item. After inspection, we record post-frequency level of each purpose item for each PAU. Then, raters are given instructions to repeat the procedure above. We calculate Cohen's Kappa coefficient to measure agreement on recording post-frequency levels of purpose items between the first author and each rater. We find that agreement between raters and the first author is above moderate.

(3) After collecting feedbacks from raters, we inspect again public contents of 200 PAUs (in Step 2), which have disagreement between the first author and raters, and resolve into unique post-frequency levels. We also inspect public contents of 150 PAUs (Table 3) and resolve the purpose item post-frequency levels in the same manner. In total, our data set consists of 350 PAUs and 19 purpose items.

2.3.3 Purpose analysis

Table 4 shows the results of Section 2.3.2 on determining post-frequency levels of 0 – 2 for each purpose item over 350 PAUs. We observe that the five most frequent purpose items (with lowest frequencies at level 0) are posting group or fellow photos (Gr), posting self photos (Me), propagating positive social news (+veN), propagating positive quotes (+veQ), and posting personal details (Detail). These purpose items also have high frequencies at level 2 in comparison with other purpose items. These tendencies suggest that PAUs are relying more on posting about their close people and propagating news (lifestyle, technology, entertainment) than either posting self photos or sharing photos containing their favorite messages (quotes). Also, as Detail has 129 PAUs at level 2, writing details about their personal lives via SNS is popular among PAUs.

Furthermore, we conduct correlation analysis on the top 5 purpose items and find that Me is positively correlated

with Gr ($r = .393, p < .01$) and Detail ($r = .250, p < .01$), Gr is positively correlated with Detail ($r = .336, p < .01$), and +veN is positively correlated with Detail ($r = .198, p < .01$) and +veQ ($r = .286, p < .01$). From these results, we can say that it is not surprising to find that PAUs who are actively posting photos of themselves, are also posting about their close people and writing details about their personal lives via SNS. However, it is surprising to find that there is no relationship between PAUs posting photos of themselves and their close people, and their tendency of propagating news (technology, entertainment, lifestyle) from SNS users. Indicating that, PAUs who are actively propagating news are not necessarily more/less active in posting photos that are directly linked to them.

In addition, we perform descriptive statistics on number of purpose items having post-frequency levels of 0 to 2 for each PAU. We find that PAUs are likely to have more purpose items with the post-frequency level of 0 (Mean = 11.1, SD = 3.1) than those with post-frequency levels of 1 (Mean = 4.4, SD = 2.4) and 2 (Mean = 3.5, SD = 1.7).

Table 4. Distribution of PAUs' post-frequency levels for each purpose item (N=350)

Symb ol	Purpose item	Leve l 0	Leve l 1	Lev el 2
Me	Posting self photos	82	141	127
Gr	Posting about close people	60	107	183
Anim	Posting pet photos	198	90	62
Detail	Posting personal details	141	80	129
Favor	Requesting favor	265	80	5
Soc	Promoting social activities	307	33	10
Caree	Promoting career activities	276	21	53
Mark	Promoting marketing activities	209	61	80
SocO	Promoting social opinion	187	118	45
RelO	Promoting religious opinion	306	32	12
PolO	Promoting political opinion	289	37	24
+veQ	Propagating positive quotes	122	129	99
-veQ	Propagating negative quotes	204	103	43
FunQ	Propagating funny quotes	241	95	14
PolN	Propagating political news	254	49	47
+veN	Propagating positive social news	60	117	173
-veN	Propagating negative social news	165	113	72
RelN	Propagating religious news	269	58	23
FunN	Propagating funny news	243	77	30

2.3.4 Refining the Purpose Model through Statistical Analysis

Now we evaluate independence of the five high-level OP-purposes in Figure 1 through factor analysis, to refine the model. We adopt from [1] a benchmark of item loading greater than 0.5 to identify a purpose item as a marker for the factor. After examining factor loadings, we retain three factors (Table 5) corresponding to three high-level OP-purposes: promoting self, propagating news quotes, and propagating sociality information, shown in Figure 2.

Promoting self is associated with a PAU’s intention to expose his/her personal life without posting self photos via SNS. An item posting self photos is discarded due to its low factor loading. *Propagating sociality information* is associated with a PAUs’ intention to share words that are either encouraging, educating, or entertaining. *Propagating news quotes* is associated with a PAU’s intention to share news (political, religious, violence, etc.) and his/her words that either support or oppose people’s opinions. Three purpose items promoting social opinion, promoting political opinion, and promoting religious opinion loaded on propagating news quotes factor. However, we discard them because they are classified under promoting purpose, and high-level OP-purpose promoting opinion in Figure 1. Joinson [1] also discarded item(s) with acceptable factor loading(s) that do not represent a factor.

If we adopt our refined purpose model, the following purpose items will be excluded due to not fit to be in the three factors above: promoting social opinion (SocO), promoting religious opinion (RelO), promoting political opinion (PolO), promoting social activities (Soc), promoting career activities (Caree), promoting marketing activities (Mark), and posting self photos (Me). However, these purpose items are clearly visible in postings and explaining PAUs behaviors well. So we retain these purpose items and adopt the original purpose model in Figure 1, and we study statistical correlations between these purpose items and other observable features of PAUs.

Table 5. Summary statistics and factor analysis of OP-purposes

Symbol	Purpose items and scales ¹	Mean	SD	Factor loadings ²		
				Propagating news quotes	Promoting self	Propagating sociality information
	Propagating news quotes scale (Alpha = 0.78)	0.50	.541			
PolN	Propagating political news	0.36	.558	.809	-.228	.041
RelN	Propagating religious news	0.63	.671	.650	-.164	.120
-veQ	Propagating negative quotes	1.13	.704	.596	-.150	.432
-veN	Propagating negative social news	0.45	.639	.547	-.136	.453
	Promoting self scale (Alpha = 0.62)	0.80	.503			
Detail	Posting personal details	0.97	.874	-.035	.746	.191
Favor	Asking favor	0.26	.469	-.017	.670	.050
Gr	Posting about close people	1.35	.757	-.295	.567	.112
Anim	Posting pet photos	0.61	.770	-.048	.564	.031
	Propagating sociality information scale (Alpha = 0.56)	0.75	.453			
FunQ	Propagating funny quotes	0.35	.556	-.039	.127	.687
FunN	Propagating funny news	0.39	.641	-.061	.003	.629
+veQ	Propagating positive quotes	0.93	.793	.122	-.087	.605
+veN	Propagating positive social news	1.32	.750	.082	.139	.574

Note: SD, Standard deviation. ¹Principal component factor analysis with Varimax rotation, explaining about 41.4% of the variance. ²Purpose item scales are 0 = contents not observed, 1 = contents observed between 1 and 5 times, and 2 = contents observed at least 6 times.

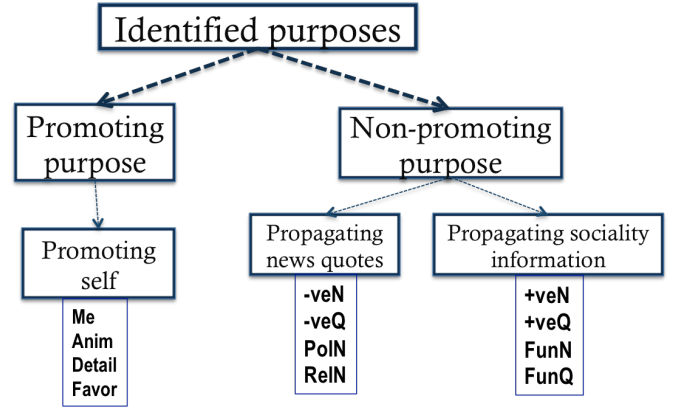


Figure 2. Purpose model constructed through factor analysis

3. OBSERVABLE FEATURES

In this section, we will discuss the observable features of PAUs, which are numerical indicators extracted from PAU’s public posts and profile items. We investigate possibilities of utilizing observable features as predictors of his/her purposes of posting contents in public. In the following, we denote by *OF* the set of observable features, where *X* is divided into non-overlapping subsets *PI* (profile items), *AI* (activity indicators), *GI* (group interests), and *PPT* (profile photo types).

3.1 Profile Items

Profile items are information elements in a PAU’s profile page. In this paper, we consider Gender, ProfilePhoto, and Weighted non-contact openness (Non-contact openness) as profile items *PI*.

We encode Gender as follows: Gender equals 0, 1, or system missing, if the PAU is male, female, or undisclosed, respectively. In our data set, there are 158 males, 141 females, and 51 PAUs whose gender information are not shown in their public profile pages. For ProfilePhoto, we set ProfilePhoto = 1 if the PAU discloses his/her profile photo to the public, and 0 otherwise.

Contact profile items are those that strongly identify user identities, such as MobilePhone, Websites, BirthdayFull, BirthdayPartial (year omitted), Email, and ScreenNames. For security reasons, very few users open such contact profile items. We define that *NG* is the set of the profile items where contact profile items, Gender, and ProfilePhoto are excluded.

We utilize the following weighted non-contact openness *NC* [3] as a normalized feature for openness of non-contact profile items, ranging between 0 and 1:

$$x_i = P_{NC} = \frac{1}{|NG|*N} \sum_{k \in NG} |R_k| y_k. \quad (1)$$

Here, y_k , $k = \{1, 2, 3, \dots, K\}$ is such that $y_k = 1$, if the k -th non-contact profile item is disclosed to the public and 0 otherwise, N is the total number of PAUs, and R_k equals the number of PAUs who opened the k -th non-contact profile item. The feature NC gives weighting such that a profile item has more weight if more users disclose it. This weighting gives us benefits such as prediction accuracy improvement, variable reduction, and portability between SNSs.

Figure 3 shows the non-contact profile items plus ProfilePhoto, and their *disclosure rates*, which is the percentage of the PAUs that open the item. We note that Figure 3 shows the profile items having disclosure rates over 10 percent and used in computing the weighted non-contact openness NC . We exclude items whose disclosure rates are no greater than 10 percent, because such profile items are sporadic and hard to be utilized to explain purpose items.

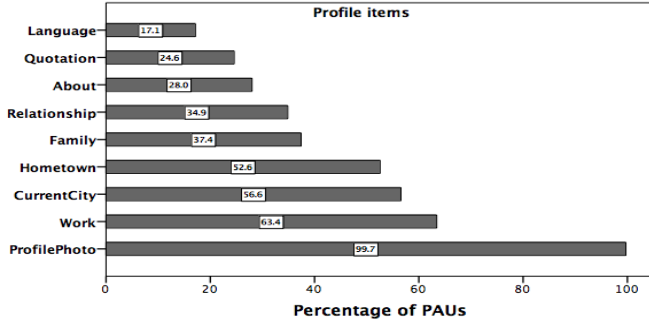


Figure 3. Disclosure rates non-contact profile items and ProfilePhoto

3.2 Activity Indicators

An *activity indicator* is a numerical feature on how much degree a user is active on a specified aspect. We consider the following nine activity indicators that are observable from their public pages: Friends, Followers, Likes, Groups, Music, Photos, ProfilePhotos, CoverPhotos, and Posts.

Table 6 shows their description, disclosure rates, mean, maximum, and minimum values. We observe high variability (range between minimum and maximum value) in these activity indicators. The cases of minimum value of 0 of an activity indicator can be interpreted as: i) the PAU hides the activity indicator value from being publicly viewed (e.g. Friends), ii) the PAU does not post regarding

the activity indicator (e.g. Music), and iii) SNS users do not interact with the PAU regarding the activity indicator (e.g. Followers). Due to high variability, we convert activity indicators into their natural logarithmic forms.

We denote by $AI \subseteq OF$ the set of the observable features consisting of the activity indicators. Values of AI are continuous.

Table 6. Description and descriptive statistics of the activity indicators

Activity indicator	Details	Disclosure rates (%)	Mean	Minimum	Maximum
Friends	Number of SNS users that are directly connected with PAU	31	260.3	0	5000
Followers	Number of SNS users that follow PAU's SNS page	28	116.9	0	20000
Likes	Number of public pages that PAU click a 'Like' button	60	429.2	0	8098
Groups	Number of publicly visible groups that PAU joined	84	59.4	0	1527
Music	Number of music related public pages that PAU click a 'Like' button	56	53.0	0	1738
Photos	Number of photos uploaded by PAU excluding profile and cover photos.	99	939.6	0	13079
ProfilePhotos	Number of profile photos that are uploaded by a PAU	99	71.4	0	1146
CoverPhotos	Number of cover photos that are uploaded by PAU	95	36.5	0	814
Posts	Number of posts that we collect (Jan. 2014 – Feb. 2015).	100	275.7	15	910

3.3 Group Interests

As we discuss in Section 2.1, we assume that a PAU participating to an interest group share the interest of the group. We utilize the group membership information as an observable feature. Let $GI \subseteq OF$ denote the set of group-interest features such that a binary feature takes value 1 if the PAU is participating to its corresponding public group, otherwise the feature takes value zero. We note that we only check whether each PAU belongs to the group or not, and do not examine his/her postings in the interest group. As shown in Table 1, our dataset samples PAUs from five public groups, so we construct four features for these groups, namely Pet, Jazz, Trade, and Political-enthusiast.

3.4 Types of profile photos

Finally, we introduce observable features on what types of profile photos are presented in PAUs' system-generated profile photo albums. In Facebook, users can upload their profile photo as their icons, and past profile photos are accessible as an album. Based on the tendencies we observe on user profile photos, we classify profile photos into the following three types: profile photo with human face(s), profile photo containing pet face(s), and profile photo with non-human/pet (objects).

We denote by $PPT \subseteq OF$ the set of features based on the above three profile photo types. Each feature x in PPT has value $x = 1$ on a PAU, if the PAU shows the corresponding photo type in one or more profile photos in the first five profile photos, otherwise $x = 0$.

3.4.1 Challenges of observing one profile photo

Relying on profile photos of PAUs that are currently displayed to observe the above types has several challenges:

(i) PAUs tend to have at least one profile photo, and their profile photo albums contain past profile photos. In our data set, 96% of PAUs have uploaded at least two profile photos, and as shown in Table 10, the maximum number of one user's uploaded profile photos is 1146. To identify profile photo types, we examine the first five profile photos of each PAU.

(ii) PAUs tend to upload several different types of profile photos in their profile photo albums. For example, one PAU has profile photo of him/herself, with a pet, and with other people. We record frequencies of each profile photo type in each album.

(iii) Profile photos that are uploaded by PAUs tend to have a variety of styles such as face visibility (none, partial, complete), type of photos (colored photos, black and white, altered colors, have texts), subjects (human(s), pet(s), human and pet, objects), and facial expression (serious, action, posing, normal smile, making a face). Due to these varieties of styles, annotators (the first author and two raters) independently examine for types of profile photos for each PAU. By recent image recognition, human face(s) and pet face(s) in photos can be automatically detected. However, to recognize objects in profile photos and to be more accurate in identifying human and pet face(s), manual detection of profile photo types is necessary.

3.4.2 Manual detection of profile photo types

The following steps are carried out to manually identifying the profile photo types of PAUs:

(1) Raters receive links of the 350 PAUs that are described in Section 3.1, and are instructed to record the frequency of each of profile photo type of each PAU. Raters are also given the following additional instructions: i) If the PAU's profile photo contains more than one human, then raters should regard the type as profile photo with human face(s), when at least one human face in that photo is clearly identified. Also, if the PAU's profile photo contains more than one pet, then raters should regard the type as profile photo containing pet face(s), when at least one pet face in that photo is clearly identified. ii) If a human or pet face(s) in the PAU's profile photo is not clearly identified (backside, half face, taken in the dark

place, blurred photo), then raters should ignore that profile photo.

In this work, we identify one profile photo type per each photo. For photos with multiple profile photo types, raters are given the following instructions: i) if a PAU's profile photo has background of human/pet and covered with texts, we give priority to the texts and raters should regard the type as profile photo with objects. However, if the face(s) are clearly visible and not covered with texts, then we give priority to those face(s) that are identified. ii) If the PAU's profile photo contains both human and pet faces, then we give priority to pet face(s) and raters should regard the type as profile photo containing pet face(s). The first author also performs the tasks in this step.

(2) Determining the frequency level of each profile photo type. The frequency level is given as the frequencies of the type found in the first five photos. If there exists a disagreement in the frequencies between the annotators, we adopt the smallest frequency as profile photo type's frequency level.

(3) Computing agreement statistics. Table 11 shows *inter-annotator agreements (IAA)*, the number of times in percentage that annotators record the same frequency of the profile photo type (Step 1). Also, Table 11 shows the probability that annotators agree on the frequency level of each profile photo type. IAA for the type profile photo with human face(s) is surprisingly lower compared with that of other types of profile photos. That is probably because a profile photo might contain a small-sized face, people took photo in dark place, or people put shades (flag) on photo, which make it difficult for all annotators to identify human face(s) in that photo. Also wearing costume tends to change the appearance of a human face in the profile photo. High probability values in Table 11 indicate that the annotators mostly agree on the frequency levels of each profile photo type of PAUs.

(4) Converting the frequency level of each profile photo type in Step 2 to a binary feature. For each PAU, if the frequency level of the type is greater than 0, then the feature has value 1, otherwise 0.

Table 7. Agreement statistics: inter-annotator agreement and probability of recording true frequency

Type of profile photo	IAA	Probability of agreement with true value		
		Annot ator 1	Annotat or 2	Annot ator 3
Containing pet face(s)	0.89	0.96	0.95	0.97
With human face(s)	0.64	0.87	0.90	0.83
With objects	0.75	0.91	0.93	0.88

3.4.3 Profile photo type results

Table 8 shows the results of Section 3.4.2, Step 2 on determining frequency levels of 0 to 5 for each profile photo type. Number of PAUs is reduced (from N=350 to N=346) because links of four PAUs are no longer available when annotators are identifying profile photo types. Profile photo with human face(s) has the lowest frequency level of 0, indicating that majority of PAUs have that profile photo type. In Table 12, the distributions of profile photo containing pet face(s) and profile photo with objects are positively skewed indicating that these profile photo types are rarely found in PAUs' profile photo albums.

We further investigate the distribution of number of profile photo types of PAUs and find that about 45% of PAUs have at least two types of profile photos in their first five profile photos. Such findings reveal that PAUs are commonly changing profile photo types; therefore, we should not adopt a model, which assumes one PAU has a single profile photo type.

Table 8. Distribution of PAUs' frequency levels of each profile photo type (N=346)

Profile photo type	0	1	2	3	4	5
Profile photo containing pet face(s)	279	31	15	11	8	2
Profile photo with human face(s)	70	56	55	52	39	74
Profile photo with objects	191	64	34	24	14	19

4. CONCLUSION and FUTURE WORK

We explored posting purposes of PAUs through inspecting their SNS public contents. We manually classified posting purposes based on types of activities into 19 purpose items and five high-level OP-purposes. Through statistical analysis, these five OP-purposes were refined into three high-level OP-purposes. However, the refined purpose model excluded purpose items that are useful in application such as determining sentiments and emotions of PAUs, among others. Also, we discussed about observable features that will be used for estimating levels of the five OP-purposes for a PAU. We introduced a method of detecting several types of profile photos of PAUs in their albums. We found that 45% of PAUs have at least two out of three types of profile photos. In addition, we also introduced post-frequency level of 0 to 2 to represent how much degree a PAU exhibits a particular purpose item.

As future work, we plan on estimating OP-purposes from

observable features (profile items, activity indicators, photos, and groups' interests).

5. REFERENCES

1. Joinson, A.: 'Looking at', 'Looking up', or 'Keeping up with' People? Motives and Uses of Facebook. In: Proceedings of SIGCHI conference on Human Factors in Computing Systems, pp. 1027-1036. New-York: ACM Press, 2008.
2. Mohammad, S.M., Zhu, X., Kiritchenko, S., Martin, J.: Sentiment, Emotion, Purpose, and Style in Electoral Tweets. Information Processing and Management, vol. 51, pp. 480-499, 2015.
3. Mvungi, B., Iwaihara, M.: Associations between Privacy, Risk Awareness, and Interactive Motivations of Social Networking Service Users, and Motivation Prediction from Observable Features. Computers in Human Behavior, vol. 44, pp. 20-34, 2015.
4. Spiliotopoulos, T., Oakley, I.: Understanding Motivations for Facebook Use: Usage Metrics, Network Structure, and Privacy. In: Proc. the SIGCHI Conference. 3287-3296, Paris, France, 2013.