Predicting Facebook Users' Interactive Motivations from Observable Factors

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, ‡iwaihara@waseda.jp

Abstract: Motivations to use social networking service (SNS) can be classified into inward, outward, and neutral. Inward motive is to interact with existing SNS friends, outward motive is to acquire new friends via the SNS, and neutral motive is the motive which cannot classify whether a user has inward or outward motive. Predicting users' interactive (inward, outward, and neutral) motivations of using Facebook is important for its users, providers, application developers, and marketers for target advertising and friendship recommendation. We have conducted a survey via a questionnaire, to study relationship between motivations and observable factors that are visible from the public. In this paper, we discuss constructing prediction models; based on binary logistic regressions, to predict motivations from observable factors such as profile photo, number of friends, contact openness, non-contact openness, gender, and age. Results show that our prediction models (inward and outward) have high accuracy and c statistics, indicating that the models can predict latent motives well and have good discrimination power. We also report that our prediction models perform better than a random selection model. Furthermore, using dimensional reduction technique important predictors for optimum prediction models were identified.

Keyword: Accuracy, c-statistics, dimensional reduction, f1-score, Facebook, kappa, motivations of using SNS.

1. INTRODUCTION

It is known that not a small fraction of Facebook users are disclosing private information to a large group of people in spite of their risk awareness[6]. This is partly due to their motivation(s) behind, for example, communicating with new friends[6], which cause users to compromise their long-term risk awareness, and communicating for short-term benefits.

Motivations of using SNSs (motivations for short), particularly Facebook; have been widely studied[1,2,5,7,8]. However, all previous works related to motivations relied sorely on users' responses from questionnaires because motivations are latent and subjective factors. The survey process can be very laborious in terms of time and cost, especially when sample data of diverse demography is required. In an SNS a user's profile data can be observed by other SNS user(s), however reason(s) behind the user information disclosure is quiet unknown to other SNS users. If motivations can be detected from observed user's profile, then it will be known whether the user unintentionally or intentionally publish his/her data. Using constructed prediction model, it is possible to predict user's motivations provided that a data set consists of users' response on motivations and behaviors is available for training purpose.

In this paper, we discuss estimating motives of using SNSs through prediction models. The models estimate likelihood of having inward, neutral, and outward motives from only observable factors. Here, observable factors refer to user-generated data visible from the public level. We focus on interactive motives where lurking is not considered, because a user with lurking motives do not require any information in his/her profile page. While a user with interactive motives require disclosing identifying information in order to facilitate communications. We classify motives into three groups, which are *inward*, *outward*, and *neutral motives*. *Inward motive* is to interact with existing SNS friends. *Outward motive* is to acquire a new friend via the SNS. *Neutral motive* is a motive toward activities a user performs on the SNS, where the activities do not indicate explicit distinction between new and existing friends. In this paper, we will focus on inward and outward motives. With our prediction models, researchers with a data set that contain only observable factors can estimate motives for further detailed analysis. SNS users (providers, end users, marketers, and application developers) can use the models to improve users experience, targeted advertising and proper friendship recommendations.

Contribution: To the best of our knowledge, we are the first to derive prediction models that estimate motives of using an SNS from observable factors. The models we

propose in this paper have several advantages: i) The models are simple and intuitive, only requiring user's SNS profile data that are visible from the public, ii) using dimensional reduction technique optimum predictor variables were identified with only little degradation in the performance of the models and, iii) the models are flexible to the changes in requiring data sets and they can be applied to different SNSs that have similar attributes. In this paper, due to space constraint, we will discuss the first two advantages.

The rest of the paper is organized as follows. In Section 2, we review related work. In Section 3, describe our prediction problem. The construction of our prediction models are presented in Section 4. In Section 5, we will present our experiment results. In Section 6, we will present a summary in the form of discussion, and Section 7 is a conclusion and future work.

2. RELATED WORK

Motivations of using SNSs have been studied as a separate variable[2,5] or related to other factors[1,7,8] beyond our work. [2,5] found that, Facebook users use the site more to communicate with their friends than to acquire new friends. In our previous work[6], we investigated factors causing private information disclosure. We found that motivation to acquire new friends is among factors that cause private information to be open, even in spite of the risk awareness. Other studies about motivations were related to factors such as trust and Internet privacy concern[1], Facebook features[7], demography, time spent on Facebook, Facebook usage and personal network metric[8].

In spite of the abundance of researches related to motivations as reviewed in the above, the current approach of using surveys to obtain users' motivations responses has several challenges: i) the questionnaire-based approach suffers from low response rate i.e. the ratio of users who responded to total number of users who received invitation request to participate in a survey. Compensation is therefore awarded to users in order to motivate them to participate in the survey. ii) Analysis technique such as logistic regression requires at least 10 users per predictor variable. The more the number of predictor variables, the more the minimum number of required users and therefore the more duration is required to obtain enough responses. iii) Since most of SNSs adopt privacy control, only permitted contents are visible from the public. By incorporating only observable factors in our prediction models, we can analyze a vast collection of SNS data that are accessible from the public. The requiring factors of our prediction models are easily translatable between different SNSs, which enable comparison of user motivations between SNSs.

3. PREDICTION MODEL METHODOLODY and PREDICTOR VARIABLES

Binary logistic regression is one of prediction methodologies that are widely used in clinical fields to predict the outcome of a future patient (predicted probability) given the patient characteristics[3], and also used in psychological analysis. In this paper, we adopt binary logistic regression to find whether an SNS user is {high, medium, low} in {outward, inward} motives, through combination of observable factors. We focus on finding which combinations of factors are high in accuracy and practically useful.

We set a threshold termed as a *scope threshold* for each profile attribute of a user. If a user opens his/her attribute value beyond the scope threshold, we regard that he/she *opened* this attribute. Facebook has five levels of *disclosure scopes* which are 1) only me, 2) selected individuals, 3) friends 4) friends of friends, and 5) everyone. For our data, we set its *scope threshold* at between 3) friends and 4) friends of friends. This threshold is intended to illuminate whether the user has an intention to reveal information to friends of friends and beyond, or just to share with his/her direct friends.

We model profiles of SNS users as follows. An SNS user j has a profile page consisting of items x_i, $i=\{1,2,3,\ldots,I\}$, and each item x_i has a non-negative value, representing openness of x_i. For most of items, we use binary number {0,1} as values except for items like number of friends and age where we use continuous values. Depending on user j's privacy settings, each item is either visible from the public (value 1) or not (value 0). Not that the latter case includes the situation where user j only shares the items with friends and friends of friends, or simply does not supply item x_i to the SNS. However it is difficult to separate these non-public cases and therefore we only focus on whether user j opens item x_i to the public or not. Let $X = [x_1,x_2,\ldots,x_I]$ be the vector of items values of user j. Let $Z\in\{0,1\}$be one of the six dependent variables on whether the user is {high, low} in {outward, neutral, inward}motive. Given observable factors X_j of user j, we model the predicted probability $P_j(Z=1|X_j)$ of user j's motive satisfying $Z=1$ as follows:

$$P_j = \frac{1}{1+e^{-(\beta_0+\sum_{i=1}^{I}\beta_i x_i)}} \quad (1)$$

where β_0 is a constant factor and β_i is a coefficient of predictor variable x_i. Equation 1 is derived from the following logistic regression formula

$$\ln\left(\frac{P_j}{1-P_j}\right) = \beta_0 + \sum_{i=1}^{i=I}\beta_i x_i. \quad (2)$$

The constant factor and coefficients of predictor variables are computed by using maximum likelihood estimation, where both predictor variables and motive values are required as a training data set.

We conducted a survey on Facebook users with a questionnaire[6], in which users' motives were asked among other questions. The question consists of 16 items, each anchored using five point Likert scale (from 1= strongly disagree to 5= strongly agree). We label user *j's* response *r* on a motive *m* based on the Likert scale as follows: (a) *j* is low on *m* when *r* is 1 or 2, (b) *j* is medium on *m* when *r* is 3, and (c) *j* is high on *m* when *r* is 4 or 5.

Applying a dimension reduction technique called *factor analysis*, we further grouped these motives into three independent motives (inward, neutral, and outward) as shown in Table 1. Inward, neutral, and outward motives of user *j* are calculated as the mean value of the element motives in each motivation. Due to space constraint, we will present results for inward and outward motives only. Later we discuss thresholds to determine whether user *j* is either high, medium, or low in these motives.

Motivations of using SNS items and scales[2]	Mean	SD	Factor loadings[1]		
			Neutral	Outward	Inward
Neutral scale (Cronbach's alpha =0.86)	**3.03**	**0.89**			
Commented on photo/posts	3.13	1.10	.833		
Changing current status	2.87	1.15	.815		
Updating or editing profile	2.88	1.15	.787		
Posting on wall	3.11	1.22	.707		
All my friends have account	2.64	1.22	.605		
Entertainment (pass time)	3.54	1.15	.510		
Outward scale (Cronbach's alpha =0.82)	**2.94**	**0.91**			
Creating group/groups	2.72	1.20		.789	
Have direct encounter with Facebook friend	2.38	1.15		.782	
Meeting new people and make new friends	3.20	1.29		.762	
Finding about, organizing or joining event	3.02	1.14		.694	
Interacting with group/groups	3.38	1.17		.521	
Inward scale (Cronbach's alpha = 0.77)	**4.03**	**0.81**			
Reconnect with people you have lost contact with	4.20	.974			.878
Maintaining relationship with friends	4.05	.958			.816
Contacting people far from you cheaply	3.98	1.13			.665
Keeping in touch with old friends	3.90	1.15			.629

Notes: [1]Principal component factor analysis with varimax rotation, cut off point is 0.500 [2]Individual items ranged from 1=strongly disagree to 5=strongly agree,

Tab.1. Summary statistics and factor analysis results for motives.

Binary logistic regression requires that a dependent variable *Z* (motives) to be dichotomized. Therefore, we set thresholds corresponding to several top *Y*% users following descending order of the mean of each motive, where Y= 10 + 5g, g ={0,1,2, …..,17}. If a user is among top *Y*% users, we regard the user as having the motive, otherwise the user has no motive. Then we select two thresholds for each motive, which we use to build two classifiers C1 and C2. Classifier C1 separates {inward, outward}-high and {inward, outward}-medium, while classifier C2 separates {inward, outward}-medium and {inward, outward}-low. The criteria for selecting two thresholds will be discussed in Section 5.

We use the following observable factors as predictor variables to predict motives: Age, Gender (we encode female = 0 and male =1), number of friends in logarithmic form (LogFriends), ProfilePhoto, Contact information which is the number of open contact attributes of a user, i.e. $k \in \{1,2,3,\ldots,m\}$where m is the number of non-contact attributes and k is openness, and Non-contact openness which is the number of open non-contact attributes of a user, i.e. $k \in \{1,2,3,\ldots,n\}$where n is the number of non-contact attributes and k is openness. We use the following for Contact openness attributes: Address, IM, Email, and MobilePhone. We use the following for Non-contact openness attributes: RelationshipStatus, Family, CurrentLocation, Hometown, Work, BirthdayFull, and Birthday Partial. However, any profile attribute(s), which fall(s) under contact or not contact information category, can be included in Contact and Non-contact openness, respectively.

4. PREDICTION MODEL CONSTRUCTION

4.1 Data Collection Method

A total of 276 participants with age ranging from 20 to 59, from 17 current countries of residence agreed to participate in our survey. Females were 134 while males were 142. As compensation, participants received a complimentary copy of possible social network risks and recommendations for safe practice in their email. Participants were recruited in six ways: invitation posted on the Foreign Student Division Board and on an international house board, mailing lists, visiting close members who have a Facebook account to their houses and giving them a questionnaire to participate, posting an invitation on several Facebook group walls and giving links to members who were interested in participating, recruited by a Japanese company, and chatting with online Facebook members and inviting them to participate by giving a link to a survey for those who showed interest. The link was opened for three weeks.

4.2 Survey Design

The survey questionnaire we prepared contained fourteen questions, which were FB related questions (motiva-

tions, disclosure scopes of some attributes, and the total number of friends). The survey also contained the set of demographic questions (gender, age, status, and current country of residence). The survey was prepared by using Microsoft Word to create a paper version and was also hosted on GoogleDocs. The survey is available in English and Japanese versions upon request from the authors.

To ensure content validity, pre-tested scales were relied on where possible. All items in questions were anchored using a Likert scale except for questions related to age, current country of residents, and the total number of friends. We relied on[2,4,5,9] to prepare items for motive construction. We modified some items where needed. For the purpose of this study, we only considered interactive motivations and activities without lurking.

5. EXPERIMENTS

This section discusses about the performance tests results used to determine the appropriate classifiers C1 and C2 for each motive. Each performance measure will be explained in the next paragraph.

We used the following performance tests to assess thresholds in each motive. *Accuracy* is the proportion of users that are correctly predicted. When any of our prediction models is applied to a particular user, accuracy gives the expected probability that the motive will be correctly predicted. *C-Statistics*[3] is a discrimination measure, which refers to the ability of the prediction model to distinguish correctly the two classes of outcomes. In binary logistic regression, this measure corresponds to the area under the receiver operating characteristic (ROC) curve. *F1 score* is the performance measure for finding portion of objects out of the large set. *Kappa* is a statistical measure that shows our prediction models are better than random selection. If kappa is near zero, it means that our prediction model is not much improving from random selection. If kappa is close to one, the prediction is almost perfect. *Diff accuracy* is accuracy with predictor variables minus accuracy with constant value only.

Top Users(%)	Mean cut point	Recall	Presicion	F1 score	Accuracy	C-ststistics	Kappa	Diff accuracy
50	4.25	0.572	0.53	0.55	0.531	0.605	0.062	2.9
60	4	0.91	0.648	0.757	0.647	0.621	0.265	4.3
65	4	0.95	0.679	0.792	0.673	0.633	0.281	1.8
70	3.75	0.964	0.733	0.832	0.727	0.667	0.35	2.2
75	3.75	0.976	0.777	0.865	0.771	0.682	0.389	1.8
80	3.5	0.982	0.828	0.899	0.822	0.711	0.443	1.8
85	3.5	0.996	0.867	0.927	0.865	0.729	0.471	1
90	3.25	0.996	0.905	0.948	0.902	0.714	0.456	-0.3

Tab.2. Performance tests results for inward motive

Top Users(%)	Mean cut point	Recall	Presicion	F1 score	Accuracy	C-statistics	Kappa	Diff accuracy
10	4	0.071	1	0.132	0.905	0.772	0.472	0.7
20	3.8	0.107	0.667	0.184	0.807	0.702	0.397	1.1
25	3.6	0.13	0.529	0.209	0.753	0.706	0.341	0.4
30	3.4	0.217	0.529	0.308	0.705	0.714	0.298	0.7
40	3.2	0.378	0.6	0.464	0.647	0.713	0.265	5.1
50	3	0.623	0.628	0.625	0.625	0.658	0.25	12.3
60	2.8	0.855	0.651	0.739	0.636	0.661	0.242	3.2
65	2.6	0.922	0.695	0.792	0.684	0.67	0.305	2.9
70	2.4	0.964	0.73	0.831	0.724	0.689	0.342	1.9
80	2.2	0.977	0.812	0.887	0.8	0.675	0.375	-0.4
90	1.8	0.996	0.912	0.952	0.909	0.695	0.494	0.4

Tab.3. Performance tests results for outward motive

Table 2 and Table 3 show the performance results for the inward and outward motives, respectively. We can observe that irrespective of the distributions of users for each motive, performance results tend to increase when at least top50% users have a motive. The cut points for inward motive are higher than outward motive, because the distribution of inward motive is skewed to the right indicating many users are high in inward motive. While the distribution of outward motive has a single peak and symmetric. We can also observe from Table 1 that the mean of inward motive is 4.03, while mean of outward motive is 2.94.

We used the following conditions to select appropriate classifiers for each motive. For both classifiers in each motive, we set lower limits of 0.6, 0.6, and 0.3 to accuracy, c-statistics, and kappa, respectively, to ensure prediction qualities. For classifier C1 in each motive, we selected a cut point for top Y% users based on the following conditions: 1) the mean value corresponding to top Y% users should be above 3.6, 2) the value of F1 score should not be equal to zero (see Table 2 and Table 3). For classifier C2 in each motive, we selected a cut point based on the following conditions: 1) the mean value correspond to top Y% users should be below 3.6, 2) the value for F1 score should not be equal to zero. Classifier C2 aims to predict users with low motives, therefore we computed F1 score for bottom Y% users (bottom 40% users and below). The results are shown in Figure 1.

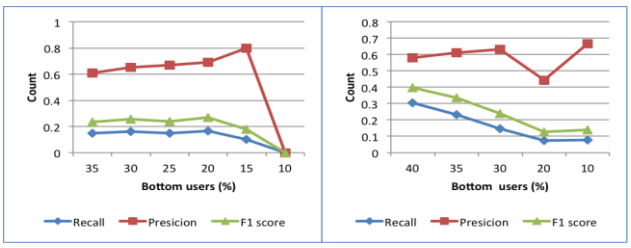

Fig.1. F1 score for left: inward motive and right: outward motive

From the above conditions on classifiers, we found that top 70% and top 80% users with corresponding inward motive mean values of 3.75 and 3.5 are appropriate cut

points for classifiers C1 and C2, respectively. For the outward motive, we found that top 20% and top 70% users with corresponding mean values of 3.8 and 2.4 are appropriate cut points for classifiers C1 and C2, respectively.

Furthermore, we performed dimensional test for each selected prediction model by removing variables, and observe changes in performance results. From this dimensional reduction test, an optimum prediction model with minimum number of predictors, causing little or no degradation in performance measures, is obtained. A prediction model with minimum predictors is desirable, because the model is applicable to a wider range of data sets where certain attributes may be unavailable.

In Tables 4-7, each row shows the number of removed predictor variable(s), predictor variables included in prediction model, and performance results. The results reveal that when predictor variable(s) are removed in ascending order of Wald's (without replacement), the first few models indicate certain degradations from the prediction model with no attribute removed (Row in Tables 4-7 with the number of removed predictor = 0). Then the performance results tend to either slightly increase or decrease with subsequent removals of predictors. When the performance results start to decrease, we include the predictor variable that causes reduction in subsequent analysis. Therefore the prediction models with the minimum number of predictors (the dotted circle in Tables 4-7) include two types of predictors; significant predictor(s) and included predictor(s) that avoid performance degradation.

No. of removed predictors	Predictor variables included in prediction model	Recall	Precision	F1 score	Accuracy	C-statistics	Kappa	Diff accuracy
Inward high motive prediction model (cut point = top 70% users) with less-influential predictor removed in ascending order of Walds								
0	ProfilePhoto, LogFriends, Non-contact openness, Contact openness, Age, Gender	0.964	0.733	0.832	0.727	0.667	0.350	2.2
1	ProfilePhoto, LogFriends, Non-contact openness, Contact openness, Age	0.964	0.733	0.832	0.727	0.677	0.350	2.2
2	ProfilePhoto, LogFriends, Non-contact openness, Contact openness,	0.974	0.730	0.834	0.727	0.656	0.350	2.2
3	ProfilePhoto, LogFriends, Non-contact openness	0.964	0.730	0.831	0.724	0.638	0.342	1.9
	Diff accuracy = Accuracy with predictor variables – Accuracy with constant only							

Tab.4. Performance results for inward high prediction model when attribute(s) are removed

No. of removed predictors	Predictor variables included in prediction model	Precision	Recall	F1 score	Accuracy	C-Statistics	Kappa	Diff accuracy
Inward low motive prediction model (cut point = top 80% users) with less-influential predictor removed in ascending order of Walds								
0	LogFriends, ProfilePhoto, Non-contact openness, Contact openness, Age, Gender	0.692	0.167	0.269	0.822	0.711	0.443	1.8
1	LogFriends, ProfilePhoto, Non-contact openness, Contact openness, Age	0.692	0.167	0.269	0.822	0.711	0.443	1.8
2	LogFriends, ProfilePhoto, Non-contact openness, Contact openness	0.625	0.162	0.257	0.811	0.648	0.409	0.7
2	LogFriends, ProfilePhoto, Non-contact openness, Age	0.692	0.167	0.269	0.822	0.640	0.443	1.8
3	LogFriends, ProfilePhoto, Age	0.625	0.093	0.162	0.811	0.617	0.409	0.7
3	LogFriends, Non-contact openness, Age	0.778	0.130	0.223	0.822	0.580	0.443	1.8
	Diff accuracy = Accuracy with predictor variables – Accuracy with constant only							

Tab.5. Performance results for inward low prediction model when attribute(s) are removed

No. of removed predictors	Predictor variables included in prediction model	Precision	Recall	F1 score	Accuracy	C-Statistics	Kappa	Diff accuracy
Outward high motive prediction model (cut point = top 20% users) with less-influential predictor removed in ascending order of Walds								
0	LogFriends, Age, Contact openness, Non contact openness, ProfilePhoto, Gender	0.667	0.107	0.184	0.807	0.702	0.397	1.1
1	LogFriends, Age, Contact openness, Non contact openness, ProfilePhoto	0.667	0.107	0.184	0.807	0.702	0.397	1.1
2	LogFriends, Age, Contact openness, Non contact openness	0.667	0.107	0.184	0.807	0.702	0.397	1.1
3	LogFriends, Age, Contact openness	0.750	0.107	0.187	0.811	0.693	0.55	1.5
4	LogFriends, Age	0.800	0.071	0.130	0.807	0.659	0.397	1.1
4	LogFriends, Contact openness	0.750	0.054	0.101	0.804	0.643	0.388	0.8
	Diff accuracy = Accuracy with predictor variables – Accuracy with constant only							

Tab.6. Performance results for outward high prediction model when attribute(s) are removed

No. of removed predictors	Predictor variables included in prediction model	Precision	Recall	F1 score	Accuracy	C-Statistics	Kappa	Diff accuracy
Outward motive prediction model (cut point = top 70% users) with less-influential predictor removed in ascending order of Walds								
0	LogFriends, Age, ProfilePhoto, Contact openness, Non contact openness, Gender	0.632	0.148	0.240	0.724	0.689	0.342	1.9
1	LogFriends, Age, ProfilePhoto, Contact openness, Non contact openness	0.722	0.160	0.262	0.735	0.683	0.369	3.0
2	LogFriends, Age, ProfilePhoto, Contact openness	0.684	0.160	0.259	0.731	0.676	0.360	2.6
3	LogFriends, Age, ProfilePhoto	0.611	0.136	0.222	0.720	0.671	0.333	1.5
4	LogFriends, Age, Contact openness	0.667	0.148	0.242	0.727	0.657	0.350	2.2
5	LogFriends, Age, Contact openness, Non contact openness	0.684	0.160	0.259	0.731	0.676	0.360	2.6
	Diff accuracy = Accuracy with predictor variables – Accuracy with constant only							

Tab.7. Performance results for outward low prediction model when attribute(s) are removed

Below, we present the optimum prediction models selected in Tables 4-7, where the four prediction models consist of the combinations of {high, low} in {inward, outward} motives.

$$P_{InHigh} = \frac{1}{1+e^{1.193-a}}\ , \qquad (3)$$

where a = $.713x_{Logfriends}$ + $.944x_{ProfilePhoto}$ - $.363x_{Contact}$ - $.180x_{NonContact}$

$$P_{InLow} = \frac{1}{1+e^{2.392-1.403x_{LogFriends}-.019x_{Age}+.095x_{NonContact}}}, \qquad (4)$$

$$P_{OutHigh} = \frac{1}{1+e^{8.197-.95x_{Age}-1.498x_{LogFriends}}}\ , \qquad (5)$$

$$P_{OutLow} = \frac{1}{1+e^{4.645-c}}\ , \qquad (6)$$

where

$c=1.256x_{LogFriends}+.075x_{Age}+.006x_{NonContact}-.347x_{Contact}$.

6. DISCUSSIONS

We used Facebook as a case study to find prediction models that predict latent motives of using the SNS from observable factors. Our prediction models can be applicable to other SNSs having the following features: (a) the SNS has similar distribution of motives as Facebook users, and (b) the SNS has access controls for profile attributes (especially for profile attributes, to apply the optimum prediction model) that can be specified by privacy settings. With our prediction models, it is also possible to predict a user motives across multiple SNSs.

The attribute for the number of friends appears as a significant predictor in both high and low motives. With-

out the number of friends as an observable factor, it is difficult to predict a user's motive. Age appear as significant factor to predict users with outward motive and inward low motive (equations 4-6). However, users with outward high motive (equation 4) have less information disclosed in their profile page than users with outward low motive. Some of predictor variables (Contact openness and Non-contact openness) for low outward motive are similar to high inward motive, showing that users, who are reluctant to communicate with new friends, use the SNS mostly to communicate with known friends. ProfilePhoto and Contact openness are identifying attributes which predict a user with inward high motive. Users with inward high motive who disclose identifying attributes probably want to be easily identified by people they know.

We selected inward, outward, and neutral motives for our study, because these motives were highly mentioned on previous researches related to motivations and users' response was high on these motives. However, this work can be extended by adding other interactive motivations when services offered by an SNS increases.

7. CONCLUSION and FUTURE WORK

We investigated four prediction models that predict motives to use SNSs, classified as {inward high, inward low, outward high, and outward low} from observable factors that can be viewed by the public. The models were developed based on user responses from the questionnaire.

Our performance results show that motives could be predicted from observable factors with high accuracies of 0.724 to 0.822. Also c-statistics results show that our models can distinguish users with motives (either high or low) or not. Kappa results show that our prediction models perform better than the random selection. We further performed dimensional reduction test our appropriate prediction models. We found that removing variable(s) causes little or no degradation in performance results. We explored variables that are in optimum prediction models with minimum numbers of predictors.

Our future work includes investigating prediction models for neutral motives. We also plan to use our model on other data sets for further analysis.

8. REFERENCES

1. Dwyer, C., Hiltz, S. R, Passerini, K.: Trust and Privacy Concern within Social Networking Sites: A comparison of Facebook and MySpace. Proceedings of AMCIS 2007, Keystone, Co. Retrieved November 15, 2008, from http://csis.pace.edu/~dwyer/research/DwyerAMCIS2007.pdf ,2007.
2. Ellison, N. B., Streinfield, 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, no. 4, 2007.
3. Giancristofaro, R., Salmaso, L.: Model Performance Analysis and Model Validation. STATISTICAL, anno LXIII, n.2, 2003.
4. Joinson, A.: 'Looking at', 'looking up' or 'Keeping up with' People? Motives and Uses of Facebook. In: Proceeding of the SIGCHI conference on Human Factors in Computing Systems, pp.1027-1036. New York: ACM Press, 2008.
5. Lampe, C., Ellison, N., Streinfield, C.: A Face (book) in the Crowd: Social Searching vs. Social Browsing. In: Proceeding of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work, pp. 167-170. New York: ACM Press, 2006.
6. Mvungi, B., Iwaihara, M.: Factors that Influence Social Networking Service Private Information Disclosure at Diverse Openness and Scopes, Proc. SocInfo2013, LNCS 8238, pp. 119-128, Japan, Nov.2013.
7. Smock, A., Ellison, N., Lampe, C., Wohn, D.: Facebook as a toolkit: A Uses and Gratification to Unbundling Feature Use. Computer in Human Behavior, vol. 27, issue 6, pp. 2322-2329, 2011.
8. Spiliotopoulos, T., Oakley, I.: Understanding Motivations for Facebook Use: Usage Metrics, Network Structure, and Privacy. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 3287-3296, Paris, France, 2013.
9. Subrahmanyam, K., Reich, S., Waechter, N., Espinoza, G.: Online and Offline Social Networks: Use of Social Networking Sites by Emerging Adults. Journal of Applied Developmental Psychology, vol. 29, issue 6, pp. 420-433, 2008.