Distance Matters: An Exploratory Analysis of the Linguistic Features of Flickr Photo Tag Metadata in Relation to Impression Management

Syed Ishtiaque Ahmed
Department of Information Science
Cornell University
301 College Avenue
Ithaca, NY, US
sa738@cornell.edu

Shion Guha
Department of Information Science
Cornell University
301 College Avenue
Ithaca, NY, US
sg648@cornell.edu

ABSTRACT

Tags are words that users add to shared multimedia contents as metadata to facilitate better categorization and improved sharing experiences. With the burgeoning growth of shared images and videos over online social networks, a huge number of tags is being populated everyday in public or shared databases. While one major reason for tagging a photo or a video incorporates the functional needs for the organization of that shared object, people also use tags as a medium of communication for conveying their emotions to their family, friends, and other contacts. The diversity in the linguistic features of these tags demonstrates some interesting patterns that reflect different facets of human nature in managing their online impression to their social peers. This paper investigates how some linguistic features of tags associated with the Flickr photos change with the distance between the user's home location and the location where the photo is taken. In our exploratory analysis "affective" and "relative" words and their multiplicative interaction show correlations with this distance. These initial findings help us to have a better understanding of online social phenomena related to the expression of emotions and sharing information. At the same time, this might have some indirect implications to understand the insight of impression management in online communities.

Categories and Subject Descriptors

H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing – Dictionaries, Indexing methods, Linguistic processing.

General Terms

Measurement, Human Factors, Languages.

DBSocial '12 Scottsdale, AZ USA

Keywords

Geotag, Flickr, Linguistic Analysis, Impression Management

1. INTRODUCTION

Flickr is an online photo and video sharing network owned by Yahoo! Incorporation [9]. It was launched in February 2004. Although we observe an increasing interest among Flickr users in sharing videos, we limit our interest in Flickr photos in this study. According to Wikipedia, Yahoo reported in June 2011 that Flickr had a total of 51 million registered members and 80 million unique visitors. In August 2011 the site reported that it was hosting more than 6 billion images and this number continues to grow steadily according to reporting sources [2]. For the ease in retrieving photos owned by a user and for communicating to the audience, tagging photos and videos with different keywords or arbitrary texts was allowed in the Flickr users. In August 2006, Flickr allowed the users to geotag the photos, and it opened up the opportunity to add the spatial information (such as: longitude, latitude, etc.) to a photo. Since then, a big portion of Flickr photos has been geotagged. With the invention of smart mobile devices with embedded GPS unit and camera, it has become very easy for the users to take photos and share those to social networks with geotags. As a result, a huge number of geotagged photos are being added regularly to Flickr.

Before going into detail, we define two classes of words, which will be frequently used in this paper. This classification has been made by the linguistic features of the words and is widely used for linguistic analyses. (1) Relative words are the words, which are related to space, time and motion (Example: day, walk, with). (2) Affect words are the words that are related to emotion. (Example: joy, love, sad). LIWC dictionary has a detailed list of different classes of words [15]. Now we come back to photo-tagging behavior. Two main reasons for tagging photos are: (1) to classify the photos for helping search engines, and (2) to make the photos easily retrievable and understandable to the users' friends, families, different other communities or to the public [4, 17]. Nonetheless, users demonstrate a high level of diversity in their behavior of tagging photos, which has been discussed in the section 2 of this paper. However, from the social psychological literature [12, 7, 6, 11, 10, 18, 5], we understand that the diversity in tagging is a consequence of the users' self-representations inside their online communi-

Figure 1: Two publicly available Flickr photos of Eiffel Tower, Paris, France. This photo on the left was uploaded by a user whose hometown is San Diego, California, United States. The tags she made were: "Paris","France", "Europe", "TourEiffel", "Eiffel Tower", "Travel 2009". The photo on the right was uploaded by a user whose hometown is Paris, France. The tags she made were: "French", "night", "city", "lights", "at", "winter", "event", "street", "trade", "leisure", "iPhone", "Eiffel", "entertainment", "iphonography", etc.

ties. Due to the presence of an invisible audience in the mind of the users, the tags were targeted to portray themselves before their social peers. Assuming the fact that most of the social peers of a person live in, or near to, that person's hometown, we expect that the diversity of the tags should follow a certain pattern with the change of distance between where the photo is taken and where the hometown is. From this point of view, some interesting tagging phenomena can be explained. For example, in most of the cases, when people upload the photos of their cars, the photos are taken near their living places. Similarly, they love to tag about time, space, and motion when they are around their home. Since the presence of invisible audience (consisted of family and friends) inside their mind is already aware of the information of the bigger things near their hometowns, they provide less "information" while tagging a "local" photo. Figure 1 illustrates how two users (one from California, United States of America and one from Paris, France) tagged their photos of Eiffel Tower differently. We can safely assume that the friends and family of a person living in Paris should be aware of the basic information of the Eiffel Tower. Hence, it is unusual for a local citizen of Paris to tag a photo of Eiffel tower with only the basic information of the tower. As a result, the tags the local citizen of Paris makes on a photo of Eiffel Tower include less informational words. Rather, those contain more "emotional" words in it. On the other hand, when a visitor from California, United States, visits Paris, she might be interested in introducing her friends and family (who are expected not to be citizens of Paris) to the basic information of Eiffel Tower. In this way, the social psychological theories help us to understand why people use more "relative" and "affect" words while describing the photos near their hometowns. Our exploratory study with a set of photos demonstrated how people tagged photos in their local space with more "relative" and "affect" words. We found how the use of these tags decreased with the increase of distance between one's hometown and the location where the photo was taken. We call this distance Photo Distance (DP) and we will use this notion in the rest of this paper. Although this particular phenomenon is also related to the users' photo taking behavior, we focus on this issue from the linguistic viewpoint of tags; since we can also categorize the photos using these linguistics features.

From social psychology literature, we came up with a set of hypotheses and tested those against a dataset of publicly available geotagged Flickr photos taken in different times over 2011. We found negative correlations in the amount of affect and relativity and positive correlations in their multiplicative interactions while considering them against DP. We present the mathematical model for each of these relationships based on the statistics we have. Furthermore, we try to find the rationales behind these patterns.

2. RELATED WORK

Despite having a lot of tags on the shared contents over Internet, not much research has been found in literature regarding tagging until Golder and Huberman's work on "collaborative tagging systems" [12]. They addressed the problems regarding the information dynamics found with collaborative tagging. They pointed out how polysemy (using a word that has more than one meanings) and synonymy (using different words to mean the same thing) in the tag database decrease the precision and recall of tagging systems. Following their work, Marlow et al proposed a model for the tagging systems that could handle those problems effectively [16]. Cuttuto et al. devised a stochastic model for collaborative tagging system which explained how people's tagging behavior followed simple activity rule despite complex cognition process when the contribution of other users are exposed to them [8]. Although their works contributed to the organization of the shared contents by tags and thus improved the probability of a better search, those could tell a little about how tagging behavior was related to the human factors.

Ames et al. investigated users' motivations behind tagging photos in mobile and online media through a qualitative study [4]. In their studies, besides the usual web interface of Flickr, they used a mobile phone application, called "ZoneTag", which helped the participants to upload photos directly to their Flickr accounts and to tag those. Based on their experiment, they classified users' motivations behind tagging the photos into two broad categories: self and social. The self-motivations included the factors that would help the users to make the images more useful to themselves. On the other hand, the social motivation incorporated the factors that would help the users to make the image more meaningful to their family, friends, and online communities. For both of these two categories the functions of the tags were either the organization of the shared images or facilitating the communication in the future. For example, a better organizational tag could offer better searching results in the future both for themselves and for the people in their social networks. On the other hand, a better communicative tag would help them to reminiscing the past in the future and help their family and friends to perceive their feelings

associated with the content of the images.

Since people try to express themselves through these tags, the theory of Self-perception is also related to the social factors of tagging [7, 6]. In this way, tags are also considered as a tool for impression management to users' social peers [11, 10]. This explains why users might be motivated to put more emotional words while describing a place in their locality and put more information while describing a place far from their hometowns. Hence, the presence of invisible social peers are not only present in the users' mind while tagging a photo, but also plays an important role to contribute to the linguistic features of the tags. So, a photo shared over Flickr is not only a photo, but also a self-portrait of the user through which she presents herself to her social peers. Similarly, the tags are not only some words for categorizing a photo, but also an important tool for managing impressions in online communities. This leads us to understand how self-identity, self- categorization, and self-presence happen in virtual environments like online social networks [18, 5]. People categorize their identities inside their online communities by following a particular pattern of expressing themselves, which they figure out by their own observations and perceptions. These self-perceptions are reflected in the images they share and the tags they made on those. At the same time, these tags allow them to create their self-identities among their peers by which they express their own opinions and create their own spaces in the online virtual communities like Flickr.

The findings of [4] support the social psychological theory on the dependency of human behavior upon the presence of others [3]. This presence can be actual, imagined or implied. Lee et al argued that perceived social presence was found to have a positive effect on tagging in del.icio.us, a bookmark managing system in which tagging is used extensively [13]. Similarly, while talking about the social motivation of tagging, the authors of [4] implicitly assumed that there was an invisible audience always present in a photographer's mind. That audience had been mainly composed of the people from the photographer's social surroundings including family and friends. Hence, the tags made by the photographers often had the objectives like making themselves identifiable to their friends and family or sharing their feelings with them. Nov et al made further investigation on the similar topic, but by narrowing down their interest to the Flickr photos[17]. In their studies they addressed the question, how these motivations are reflected in the tags generated by the users on Flickr. They found that the levels of the Self and Public motivations, together with social presence indicators, are positively correlated with tagging level (number of tags); while *Family and Friends* motivations are not significantly correlated with tagging.

Although there has been a number of studies related to collaborative tagging behaviors [14, 16, 12], very few literature were found to study the interplay between the tagging behavior and the spatial information of the image. Lee et al showed that there was a strong correlation between human generated tags and geotags [2]. They claimed that the user-generated tags on different images of geographically similar places are also similar. They also showed, this relationship followed the power-rule, which strengthened their actual claim of the similarity measurement. Although their paper could establish the relationship between these two types of tags, they did not consider the locality effect in tagging behavior. From their model, we can infer that if A and B are two geographically dissimilar places, and if the sets of tags made at those places are T_A and T_B respectively, then there should sufficient dissimilarity between T_A and T_B. While this model helped us to understand the geographic effect on tagging behavior, the question of how the tags made by a local person are different from those made by a foreigner still remained unaddressed. Wu et al first showed that while tagging places on a map of a particular area, the local people were interested in tagging places which were different from the places the foreigners were interested to tag [19]. But how this differences effect in the linguistic features of the contents of the tags has not yet been studied to our best knowledge. Moreover, this effect has not yet been studied for the tags on photos either.

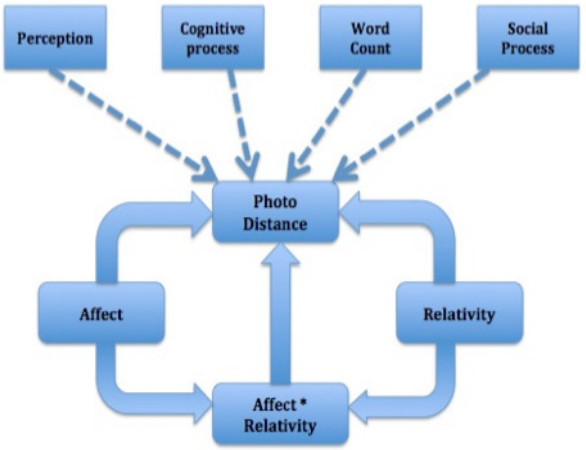

Figure 2: Summary of the research model

3. RESEARCH MODEL

Based on the theories regarding the changes in linguistic patterns of tags with the Photo Distance (D_P), we expect to see evidence that the D_P is influenced by the independent variables. We make the following hypotheses according to our expectation:

H1: D_P will be negatively correlated with the amount of affect contained in the photo tags.

H2: D_P will be negatively correlated with the amount of relative words contained in the photo tags.

H3: D_P will be positively correlated with the interaction effect between tags and relativity contained in photo tags.

Figure 2 summarizes our research model in general. Moreover, we included the perception, social, cognitive content of the tags in addition to the total word counts of every tag associated with a given image as a control variable.

4. METHOD

Data: Using the APIs, we extracted the geotagged photos of the first two days of the every moth of 2011. Due to the limitation of the usage of the APIs we could only get about four thousand photos with each query. Hence the initial database of photos was made with about forty eight thousand. Next we discarded the photos for which the uploades' home locations were not available. This left us

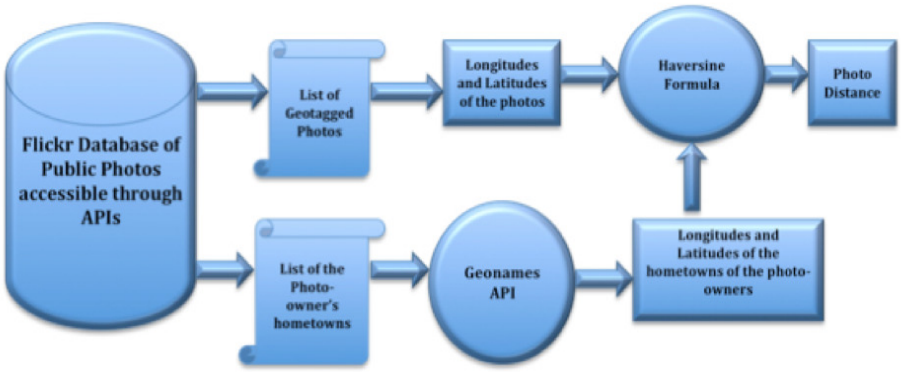

Figure 3: Acquisition of Photos and Determination of Photo Distances

with 22,957 photos from 3,198 unique users, where we got some texts in the "location" field of their profile. It is worthmentioning that the users are prompted to fill up the field for their location when they open their Flickr account. We used Geonames API to get the longitude and latitude values of the home locations of the users. In our study we found that a good number of entries in the "location" field of the users' could not be identified, which we considered to be usual in case of social networks. Later, we computed the distance (in kilometer) between the photo location and home location for each of the photos using "Haversine" formula [1]. Using the Flickr APIs we extracted the tags of those photos that are made by the users themselves. We used the software program Linguistic Inquiry and Word Counts (LIWC) for analyzing the linguistic features of these tags. LIWC analyzed the tags of each of the photos and returned a set values corresponding to their linguistic features. Among others, these linguistic features include word and parts of speech counts, social process, affective process, and cognitive process.

LIWC: Linguistic Inquiry and Word Count (LIWC), a computerized text analysis program that outputs the percentage of words in a given text that fall into one or more of over 80 linguistic (e.g., first-person singular pronouns, conjunctions), psychological (e.g., anger, achievement), and topical (e.g., leisure, money) categories [15]. It builds on previous research establishing strong links between linguistic patterns and personality or psychological state, but makes possible far more detailed results than did hand counts. This is widely used as standard linguistic analysis tool.

Geonames API: The GeoNames geographical database covers all countries and contains over eight million place names that are available for download free of charge. A number of free APIs are available that allows us to get the longitude and latitude values of a place. (www.geonames.org)

5. ANALYSIS

After obtaining the cleaned dataset, we looked at the descriptive statistics of all the response and predictor variables, which are presented in Figure 4. Next, we looked at the distributions of the response and predictor variables to ensure that they meet the conditions of normality for a multiple linear regression model. Our response variable, distance, followed a skewed distribution and was the primary variable

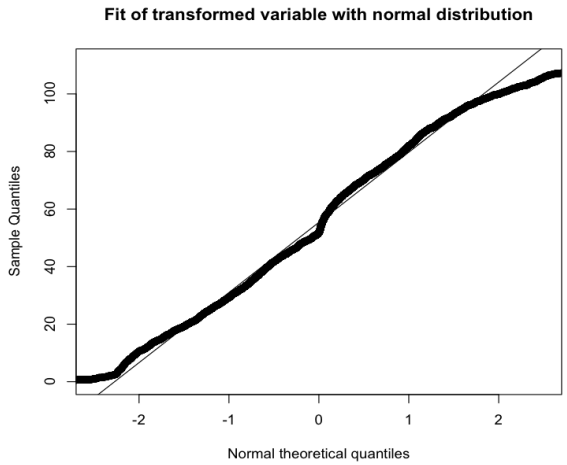

Figure 5: Fit of transformed variable with normal distribution

of concern. Therefore, after considering a range of transformations, a square root transformation on distance as the response variable was decided upon by the research team. The fit of this transformed variable to the theoretical normal distribution having similar parameters can be seen in Figure 4.

The Anderson-Darling (A=0.499, p = 0.22) and Kolmogorov-Smirnov test statistics (D = 0.031, p = 0.22) also support this hypothesis at the 95 confidence level. Figure 6. represents the summary statistics for this variable. Therefore, we can proceed to construct our analytic model with this transformed variable as our dependent variable. A multiple linear regression model with the preceding variable as a response variable, 3 predictor variables and 4 control variables (obtained from LIWC analysis) was conducted as the initial model and the results presented in Figure 8. It was seen that the following theoretical model seemed to be the best in terms of explaining the most variance in the given data.

$$y = \beta_0 + \beta_1 * affect + \beta_2 * relative + \beta_3 * affect * relative + e \quad (1)$$

Parameters	Variables							
	Dependent	Independent			Control			
	Distance (km)	Affect	Relativity	Affect * Relativity	Perception	Cognitive	Word Count	Social
Minimum	0^	0	0	0	0	0	0	0
Maximum	12360	100	100	2500	100	100	119	100
Mean	3659	1.314	5.47	12.68	3.345	1.096	7.621	1.846
Median	2681	0.898	4.888	10.533	2.821	0.734	7.03	1.266
Standard Deviation	2781.837	2.296	3.201	8.541	2.849	2.06	3.259	3.185

Figure 4: Descriptive statistics of independent, dependent and control variables ($\wedge$ distances $<$ 1km are rounded down to zero)

Parameters	
Minimum	0
Maximum	111.20
Mean	55.37
Median	51.78
Standard Deviation	24.37

Figure 6: Descriptive statistics of transformed dependent variable (distance)

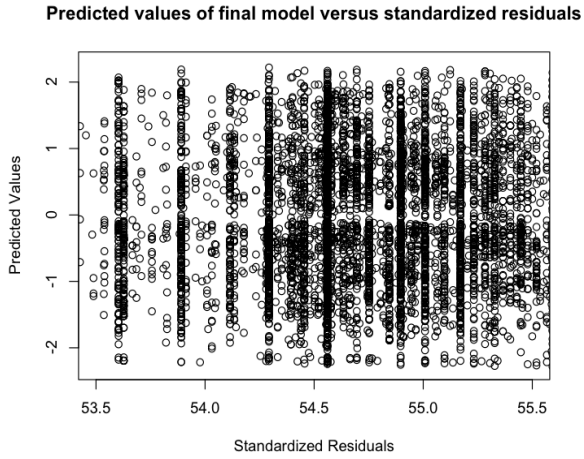

Figure 7: Predicted values of the final model versus standardized residuals

6. RESULT

Our final regression model seems to suggest that affect, relative and their interaction affects the variances in the given data. The final model fitted equation was found to be:

$$distance = 3718.7147 - 18.3245 * affect - 8.4009 * relativity + 0.8276 * affect * relative + e \quad (2)$$

We found that both affect ($|t| = 4.462$, $p < 0.001$) and relative components ($|t| = 4.486$, $p < 0.001$) in Flickr image tags decreased quadratically with a unit increase in distance, which is consistent with the assertions in **H1** and **H2**. LIWC decomposes affect further into positive and negative emotions. However, the correlation between these two is - 0.3 ($p<0.001$) and the correlation between affect and positive emotions is 0.78 ($p < 0.001$). This suggests that any affective component in tags is mainly driven by positive emotional content which may be a reflection of certain impression management strategies by the users which will discussed in the succeeding section. The interaction effect between these two variables increased quadratically ($|t| = 2.684$, $p < 0.01$) with a unit increase in distance. It should be noted here that the correlation between affect and relative is 0.1 ($p <0.01$). Therefore, pure correlation is not driving this interaction. This is consistent with **H3**. We did not find any sort of statistical significance with any of the control variables.

7. LIMITATIONS AND FUTURE WORK

We only analyzed 22, 957 public photos uploaded in 2011, which do not represent all the Flickr data. The home location field in the profiles of the users might not be completely accurate. The LIWC values that we used in our studies might not reflect to the actual meaning of the tags that the users made. We are also interested in looking at the semantic content of tags from such image data. Moreover, future work in this area should focus on incorporating additional demographic variables like age and gender for a much more detailed model.

8. CONCLUSION

We have shown that affective and relative components, as well as their interaction have a significant impact on this distance. We have constructed possible hypotheses from the theories and the result of our analysis supported our hypotheses. This is an exploratory study of how linguistic features obtained from the tags of an image shared over Flickr may have an impact of the content of the image.

Initial Model			Final Model		
Variables	**Parameter Estimates**	**Standard Error**	**Variables**	**Parameter Estimates**	**Standard Error**
Word Count	- 0.021	0.016	Affect	-18.325***	4.107
Social	0.016	0.016	Relativity	-8.401***	1.872
Affect	- 0.159***	0.037	Affect*Relativity	0.828**	0.308
Cognitive	- 0.019	0.039	F statistic		12.8***
Perception	0.018	0.02	Adjusted R-squared		0.63
Relativity	- 0.075***	0.017	*** p < 0.001 ** p < 0.01,		
Affect*Relativity	0.009***	0.003			
F statistic		6.378***			
Adjusted R-squared		0.54			

Figure 8: Parameter estimates and fit statistics of initial and final models

9. REFERENCES

[1] *http : //en.wikipedia.org/wiki/haversine_formula.*

[2] Flickr - advertising solution. *Yahoo! After June 2011*, September 2011.

[3] G. Allport. In the handbook of social psychology. Addison-Wesley Publishing Company, 1968.

[4] M. Ames and M. Naaman. Why we tag: Motivations for annotationin mobile and online media. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pages 971–980. ACM, 2007.

[5] J. Bailenson, J. Blascovich, A. Beall, and J. M. Loomis. Proxemic behavior as a function of behavioral realism within immersive virtual environments. *Presence: Teleoperators and Virtual Environments*, 10:583–589, 2001.

[6] D. J. Bem. *Self-Perception Theory, In L. Berkowitz (Ed.), Advances in Experimental Social Psychology (Vol. 6, pp.1-62)*. New York: Academic Press.

[7] D. J. Bem. Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, 74:183–200, 1967.

[8] C. Cattuto, V. Loreto, and L. Pietronero. Semiotic dynamics and collaborative tagging. *Proceedings of National Academy of Sciences*, 104:1461–1464, 2007.

[9] Flickr. *http : //www.flickr.com.*

[10] E. Goffman. Embarrassment and social interaction. *The American Journal of Sociology*, 62(3):264–271, 1956.

[11] E. Goffman. The presentation of self in everyday life. New York: Doubleday, 1959.

[12] S. Golder and B. Huberman. The structure of collaborative tagging systems. *Journal of Information Science*, 10(4):198–208, 2006.

[13] K. Lee. What goes around comes around: An analysis of del.icio.us as social space. In *Proceedings of Computer Supported Collaborative Work (CSCW)*. ACM, 2006.

[14] K. Ling, G. Beenen, P. Ludford, X. Wang, K. Chang, X. Li, D. Cosley, D. Frankowski, L. Terveen, A. Rashid, P. Resnick, and R. Kraut. Using social psychology to motivate contributions to online communities. *Journal of Computer-Mediated Communication*, 10:4, 2005.

[15] LIWC. http://www.liwc.ne.

[16] C. Marlow, M. Naaman, M. Davis, and D. Boyd. Tagging paper, taxonomy, flickr, academic article, to read. In *Proceedings of Hypertext*. ACM, 2006.

[17] O. Nov, M. Naaman, and C. Ye. What drives content tagging: The case of photos on flickr. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI)*, pages 1097–1100. ACM, 2008.

[18] R. Ratan and B. S. Hasler. Exploring self-presence in collaborative virtual teams. *PsychNology Journal*, 8(1):11–31, 2010.

[19] S. Wu, S. Liu, D. Cosley, and M. Macy. Mining collective local knowledge from google mymaps. In *Proceedings of the 20th International Conference companion on World Wide Web (WWW)*. ACM, 2011.