An Examination of the Behaviour of Young and Older Users of Facebook

Darren Quinn, Liming Chen, and Maurice Mulvenna

School of Computing and Mathematics, University of Ulster,
Jordanstown. BT37 0QB, Northern Ireland
quinn-d15@email.ulster.ac.uk,
{l.chen,md.mulvenna}@ulster.ac.uk

Abstract. Increased use in recent years of social networking applications such as Facebook has resulted in a rapid growth of user generated content. As a communication tool for social engagement, the online patterns and characteristics of individuals and user types can now be better understood by monitoring user behaviour, activity frequency and modality of use. The aim of this paper is to present an analysis of the behaviour and characteristics of young and older user groups in an online social network. Our findings show that correlations exist between user behaviour and age, based upon activities as evidenced within the data. Visualisation of full user activity has further helped our understanding of user behaviour by identifying interaction patterns, which may be used for the monitoring of social engagement.

Keywords: Interaction Analysis, Social Networking, Social Computing.

1   Introduction

The growth and adoption of Internet technologies has increased markedly over the past decade. From 2000 to 2011, growth in Internet usage has witnessed an increase of 353% and as of March 2011, 58% of Europe's 816 million population were Internet users [1]. As a subsequent development of the Internet, online social networks emerged through a desire by users to share data, experiences, events and emotions. As a communication tool they performed a major role in the increased adoption rates of Internet technologies. As such a diverse range of social networking sites have been developed, from generic social networks such as Facebook[1], to specialist content sharing sites such as the photo sharing network of flickr[2]. The impact of social networking has been acclaimed as rendering a new global footprint [2], due to its profound social and economic impact, on both business and individual users. With increased access to web services, and ever more innovative developments for end users, high profile social networking applications such as Facebook and Twitter[3] are being routinely applied as a core communication approach for users desiring to share

---

[1] www.facebook.com
[2] www.flickr.com
[3] www.twitter.com

information on a range of topics. An increased use of social networking sites has led to a rise in research interest, particularly within the domain of Social Network Analysis (SNA), a research area with a traditional role of analysis and visualisation of networks [3], discovering details such as a networks structure, or the strength and cohesiveness of community structures [4].

More recently within SNA a trend is emerging which is extending traditional concepts towards a more contemporary approach of Social Network Interaction Analysis (SNIA). As an example, Facebook is currently the most popular Social Networking Site with an estimated 700 million unique monthly visitors [5]. The aim of SNIA is to understand the "*individual*" and their use of functions in these vast networks. Through a variety of techniques such as web crawlers or bespoke programs, SNIA seeks to discover the behavioural patterns of each individual by analysing their user generated content. Contemporary studies have been interested in aspects such as user profiles, communication patterns and interactions in online networks [6], [7], [8]. As an approach it not only allows for the analysis of users and groups etc, but may also be applied to develop a greater understanding of the behavioural patterns and related characteristics of each user.

1.1   Aim

The aim of our current research is to harness and mine user generated content, through data collection approaches applied to Facebook, exploring the behaviour and characteristics of two distinct user groups, young (15-30) and older (50+) users, comprised of equal numbers of males and females. This study is made up in two parts; firstly we investigated user activity, defined as the posting of comments or replies to comments. As a study, its aim is to define the individual interaction metrics of any given user's activity. These measures are applied to compare interaction rates of *a user*, and across two *age categories*. Secondly we explore activity visualisation, to disclose interaction patterns. Illustrating user interaction data may help to understand previously undisclosed patterns, detailing the: *who, what* and *when* of a user's social network engagement, illustrating patterns such as network joining or the activity evolution of a user, defining their high and low periods of user activity.

The remainder of the paper is organised as follows. Section 2 discusses the concept, exploring in detail the chosen social network structure, defining interactions metrics and analysis methods applied in the study. Section 3 describes the user data and the data collection process, with results of activity frequencies and activity visualisation presented. Section 4 discusses the patterns that can be observed from the results. Section 5 summarises current work, discussing future work and limitations.

2   Concept and Methodology

The aim of the research is to develop an understanding into the individual activity of users. The research is aimed at determining individual metrics, and also to establish if correlations are identifiable between age groups and interaction frequency. Studies are

designed with the primary objectives of: quantifying user activity, visualising user engagement, identifying behavioural patterns and characteristics of disparate user groups, based on real world observations in an online social network. Metrics are subsequently applied to categorise users based on measures of activity, based upon frequency of posting comments and replies to comments. The primary motivation for investigating young and older online users is that it helps to establish if different behaviours (e.g., high or low activity patterns) are present and identifiable. Secondly the research may help to determine if these behaviours are attributable to the age of online social networks users. Future investigations will focus on establishing if indicators in the behaviour of people in online social networks relate to user well being (e.g., low activity frequency corresponding to low feelings of well being).

2.1   Methodology

"*Facebook is a social utility that connects people with friends and others who work, study and live around them. People use Facebook to keep up with friends, upload an unlimited number of photos, post links and videos and learn more about the people they meet*" [9]. Facebook was selected to investigate user's online social network engagement for the following reasons: prominence as the leading social network, user volume, data accessibility and a range of interaction features. Essentially Facebook users engage through their '*Wall*', a facility which controls user content, such as for the posting of comments. As users contribute to their wall a chronicle of interactions is amassed. User's security options are self-imposed and dictate the accessibility of data by other Facebook network users. Privacy settings are set and controlled by the user, and information can be disclosed to: '*everyone*', '*friends of friends*' or '*friends only*'. We class information as publically available if a user's privacy settings are set to everyone. It is only these profiles with which the study was concerned.

As is dependent on the information provided by the user, the age of a user is not always explicitly given. However, provided certain information is available it then becomes possible to determine a user's age (within a close proximity), in one or a combination of three ways, either by: date of birth, school leaving year or university leaving year. Interactions were defined within the context of this study as being any measurable user activity occurring on the user wall. A full range of potential Facebook interactions were assessed for their inclusion within the scope of the current study. However, the remit of the current study is to assess only the user's personal contribution to their network, therefore all non-user elements (e.g., non-user comments) were omitted. Due to the subjective nature of particular wall features our study concerned itself solely on user comments and user replies as core application components to provide a measurement of user contribution. It is viewed that later studies may provide an opportunity to provide a broader scoped analysis. As a specific example, the following comment scenario is provided; a user posts a comment and another user posts a reply in response to this comment, both interactions are recorded. However, only the first interaction will be of interest in the scope of this study as this is the only interaction attributable to the user in question. Corresponding metadata of date and time was also captured for each user element.

2.2   User Frequency

User metrics were derived using the following approach. Activity frequency was calculated as a measure of each individual's activity count, which was then divided over the length duration of available user data. In reality this means extraction of the first and last activity dates of any particular user, resulting in a value for the total number of active days. This value was then divided by the number of activities to obtain an overall frequency metric for the entire active period. Activity frequency calculation:

$$af = n / (x1 - x2) \hspace{3cm} (1)$$

Where; activity frequency (af), first date stamp (x1), last date stamp (x2), total number of user activities (n).

3   Data Collection and Trends

This section describes the data collection process, and provides specific details of the user data obtained for young and older users. The process was repeated in both instances altering only for the setting of profile ages. A user profile was created and using the '*find classmates*' tool, a seeding point of local universities was applied. Based on the defined profile attributes (age, gender, etc), a listing of potential friends were identified and returned by Facebook. User suitability for inclusion in the study was then determined on the two following issues; firstly a user's wall must be publically viewable; secondly age of the user must be identifiable from the stated user profile. Profiles of each suitable user's '*friends*' were then assessed using the same approach. Process was repeated until a sufficient amount for each user group and gender was acquired. Data was acquired over two months from December 2010, and subjected to a three stages of analysis process. (1) User walls were manually extracted through generated source code, containing all visible wall data. (2) A bespoke parsing program was developed for the identification and extraction of required user interaction data. (3) Comparison evaluations were applied for each user group and gender set. In total 500 full user profiles were obtained (125 young males, 125 young females, 125 older males and 125 older females). User data consisted of the entire wall history for each user, with walls exhausted until posts were no longer available. User age data was determined on a users '*info*' page, and extracted in conjunction with wall data.

3.1   Activity Frequency

Based on initial analysis, user groups were categorised into the following five activity groupings: '**G1**': Zero activity, '**G2**': Greater than 0 and less than 5 days, '**G3**': Greater or equal to 5 days and less than 10 days. '**G4**': Greater or equal to 10 days and less than 20 days. '**G5**': Greater than 20 days. These classifications are designed to represent increasing activity ranges, from G1 to G5. The first group 'G1' is the lowest possible activity value of zero, applied when no user activity was recorded on the user wall. As detailed the majority of users can then be subsequently categorised within

the remaining groupings of G2 to G5. The user comment frequency for both the older and younger users (Fig.1), and also the user reply frequencies for older and younger users (Fig.2) have been illustrated below. Further detailed is the activity frequency comparison table of young and older users (Table 1).
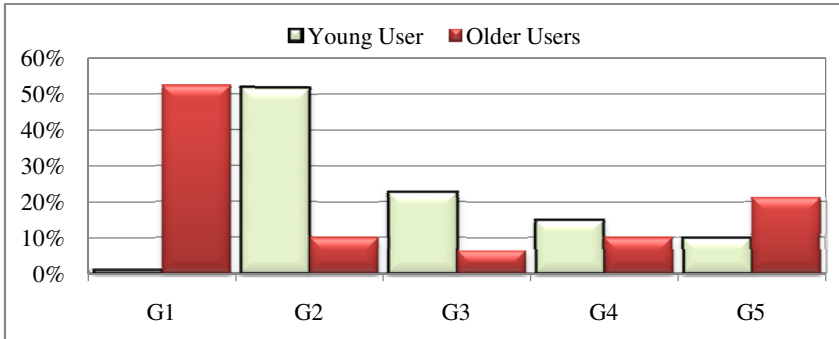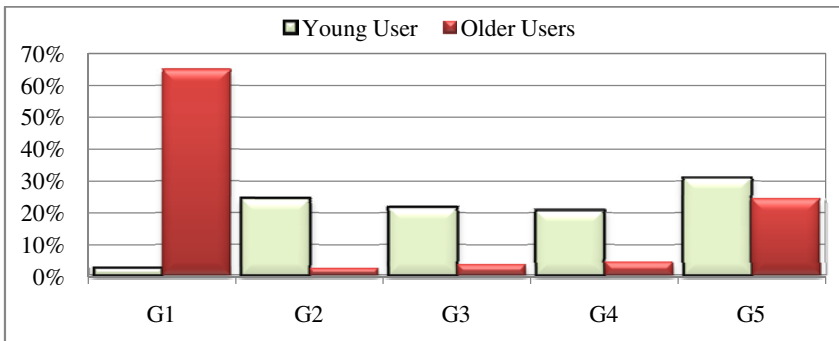

Fig. 1. User Comment Frequency

Fig. 2. User Reply Frequency

Table 1. Activity Frequency Table

Group	Comment Frequency (%)			Reply Frequency (%)	
	Younger	Older		Younger	Older
G1: Zero days	1 %	52 %		2 %	65 %
G2: 0 - 4 days	52 %	10 %		24 %	2 %
G3: 5 - 9 days	23 %	6 %		21 %	4 %
G4: 10 - 19 days	15 %	10 %		21 %	4 %
G5: Over 20 days	10 %	21 %		31 %	24 %

3.2   Activity Visualisation

In order to visualise user activity a '*representative user*' was selected (Fig.3). Analysis of activities for both groups indicates that younger users generate greater volumes of

activity, therefore the younger cohort was chosen. Of the active groups, statistically 'G2' (Frequency values 0-5) was shown to be the most representative for younger users. In order to isolate and visualise a user, an individual was selected which contained the value closest to the average number of days value. The average number of active user days for younger users was 368 days. To ensure objectivity the foremost G2 young user to be found either side of the average number of days was selected. In this instance (whilst ensuring user privacy) we can reveal a number of user elements; the user was male, active for 370 days, retaining a frequency for their user comments of 2.03 and for their replies 3.06. As shown below (Fig.3):
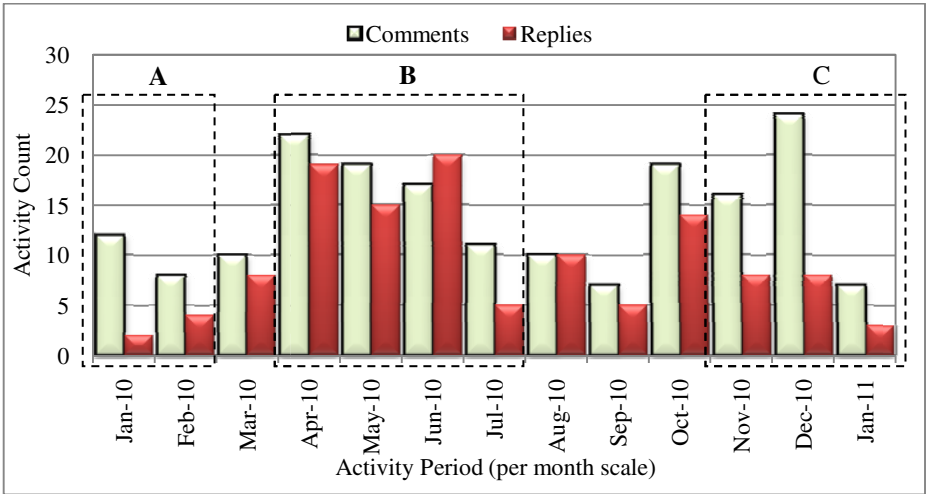

Fig. 3. User Activity Chart

4   Analysis and Discussion

Analysis of comment activity frequencies discloses the following. Relating specifically to older users, the most dominant grouping was G1, containing 52% of all older users activity frequency values. As a group it identifies users whose walls contain no user activity. The second largest grouping was 'G5', identifying all user activity frequencies greater than 20, with 21%. A decreasing trend further continued through groups G4 and G3 with 10% and 6% respectively. 'G2' observed 10%. Within younger users notably only 1% of user values were contained within 'G1'. The largest younger user grouping was 'G2' retaining 52% of all younger frequency values. Younger user values declined sharply throughout the remainder of the groupings with; G3 containing 23%, G4 15% and G5 10%. Comparisons of the two groups show clear disparities are identifiable between the activity frequencies of young (Fig.4a) and older groups (Fig.4b). It can be shown that the majority of young users engage frequently, in contrast to older users. Older user engagement can now be said to be much less frequent with markedly lower levels of activity.
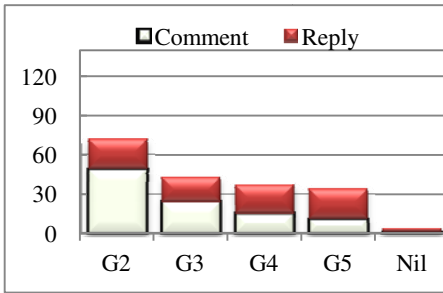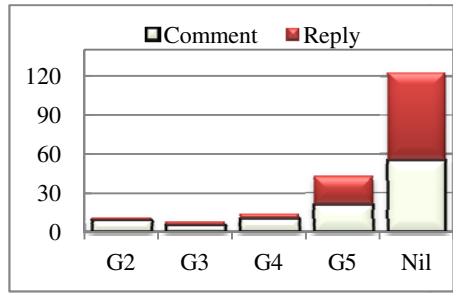
Fig. 4a. Younger Frequencies

Fig. 4b. Older Frequencies

As illustrated (Fig.3) in plotting user activity patterns are identifiable providing and understanding of a users habits. A total of 303 activities were observed made up of 182 user comments and 121 user replies. Within Fig.3 a series of groupings have been highlighted (A-C). Band A illustrates early social network engagement were a cluster of low volume recurring activity was recorded. A period of consistent activity can be seen over approximately the first 16 weeks from January 2010 to April 2010, were the majority of activity focused on user comments. Band B is identifiable as a period of heightened and concerted activity, were a significant increase in activity is visible, with a nucleus of activity between May 2010 and June 2010. Band C is identified as current behaviour in the final quarter of the year, with consistent activity. Spiking patterns are observed in this period resulting in increased user volumes. Given the length of engagement (370 days), it is viewed activity is non random and that patterns are evident. Notable however is the observation of low or non engagement periods, such as that of late July 2010 or early November 2010.

5   Conclusion

This study makes a number of contributions to the knowledge of online user interactions and user behaviour. This paper has shown that in the online social network of facebook.com the activity frequency at which younger and older users engage is varied. By harvesting user generated content, contained within a user's wall profile, the frequency at which a user engages can be determined. The development of activity frequency metrics now quantifies the frequency at which a user contributes to a social network. As is shown, such approaches can enable users to be classified based on their activity. Subsequently it is possible for the interaction rates to be compared across different users and against different user types in age groups. These results show that younger users can be classified as high frequency users, producing large volumes of online user data. Older users have been shown to engage less with a more infrequent online social network engagement. Through the plotting of a user's online history, results show that the behavioural patterns of a user can be observed, particularly in regard to high, low and non engagement.

6   Summary of Results, Future Work and Limitations

In determining the activity frequency of a user's data, and its subsequent visualisation, it has helped further the understanding of user's behaviour, detailing when and how users engage, identifying patterns of early, mid and current engagement. Importantly it defines high and low periods of activity. As with just one user example, it may be possible to hypothesize as to the reasoning of activity patterns. They are observed throughout the spring and early summer months as having a concerted period of activity. One tentative hypothesis is that behavioural patterns may have been stimulated by the reaction to increased real world social activities (holidays, etc.). Although conjecture for this particular user, patterns may realistically be applied to specific user types for monitoring social engagement, as early warnings relating to increasing social isolation of vulnerable individuals or groups. Online social networks have been shown as a means to track a user's social engagement levels, which may be applied in epidemiology for pattern anomaly detection.

It is envisioned that our future work in this area will extend current research, establishing if correlations exist between online activity and user well being. Limitations of current work are concerned with user data, specifically that the study is limited to only users that defined privacy settings to 'everyone'. It is hoped that a tool or process can be developed to overcome this issue, whilst acquiring a larger data set.

References

1. Miniwatts Marketing Group, Internet World Stats. Internet (March 2011), `http://www.internetworldstats.com/stats4.htm` (accessed July 12, 2011)
2. Charron, C., Favier, J., Li, C.: Social Computing, `http://www.forrester.com` (accessed April 02, 2010)
3. Lewis, K., Kaufman, K., Gonzalez, Christakis, M.: Tastes, ties and time: A new social network dataset using facebook.com. Social Networks 30(4), 330–342 (2008)
4. Porter, M.A., Onnela, J.P., Mucha, P.J.: Communities in networks. Notices of the American Mathmatical Society 56(9), 1082–1097 (2009)
5. eBizMBA. Top 15 most popular Social Networking Sites. Internet, `http://www.ebizmba.com/articles/social-networking-websites` (accessed July 12, 2011)
6. Massari, L.: Analysis of MySpace user profiles. Information Systems Frontiers 12(4), 357–359 (2010)
7. Huang, L., Xia, L.: User Character and Communication Pattern Detecting on Social Network Site. In: International Conference on Engineering Computation, ICEC 2009, May 2-3, pp. 261–264 (2009)
8. Wilson, C., Boe, B., Sala, A., Puttaswamy, K.P.N., Zhao, B.Y.: User interactions in social networks and their implications. In: Proceedings of the 4th ACM European Conference on Computer Systems (Eurosys 2009), pp. 205–218 (2009)
9. Facebook. Facebook Press Room, About Facebook. Internet, `http://www.facebook.com/press.php` (July 12, 2011)