Scientific
Research
Publishing

Examining the Effects of Digital Social Networks on New Physical Human Interactions and Social Networks: A Validation of Dunbar's Numbers

Fergus J. D. Keatinge

Department of Geography, University of Florida, Gainesville, FL, USA
Email: fergus@ufl.edu

Abstract

The digital social network of a user, who had undergone two radical locational changes, was analyzed to assess if digital social networks were influencing the ability of the user to create new physical social bonds in regards to proximal distance of existing social interactions and if new physical social networks conformed to Dunbar's theorem of the social network size limit. The social network data (users equating to nodes and physical friendships to links) was implemented into the network analysis software "Gephi". Standard network measures were assessed on the three digital sub-networks with the user removed from the calculations. Two separation algorithms were assessed on the network data, the Force Atlas algorithm and the Fruchterman Reingold algorithm. The results contradicted with existing research indicating that existing digital social networks did not have an effect on the creation of new social bonds after a radical locational change. The creation of new physical social networks conformed in part to Dunbar's network size limit theorem and existing social links did not affect the user's ability to create new social partnerships.

Keywords

Network Analysis, Facebook, Gephi, Force Atlas, Fruchterman Reingold

1. Introduction

"Facebook" is a digital social networking service created in 2004 by Mark Zuckerberg [1]. A user creates a dig-

ital profile and then has the ability to create links with other user profiles, in essence create a digital link between friends. A user also has the ability to perform many other actions, such as exchanging messages with other users (both public and private), posting status updates, sharing photos/videos/web links and importantly displaying common-interests by either joining user groups or 'liking' a common-interests page. Facebook has over 1.3 billion users, with an estimated 49.5% of Americans having a user Facebook profile as of 2011 [2]. In the modern day social network, digital services such as Facebook are now an integral part of the social system [3]. The ability to communicate with a member of your designated friendship group both personal conversations (private messages) and broadcast messages (status updates, picture sharing etc.), means friendships are unconstrained by physical distances (as the network is digital). With these factors in mind, there arises the issue of dependence on digital social networks, and how these dependences are affecting physical social networks, *i.e.* the want/need/ ability for forming new social network connections in the real world.

In 1992 and 1993, Dunbar examined the limit and typical social network size of humans and primates. Dunbar's [4] [5] research found that the maximum limit of a human social network size was between 100 and 200 individuals, due to the limit of the Neocortex size in human brains. While studies have been undertaken to examine the implications of digital social networks of physical human social networks, such as the research of Goncalves *et al*. [6], little research has been placed into these implications using geospatial distance as a defining feature. This research aims to analyze the digital social network of a subject (test human) that has undergone the radial relocation across the earth multiple times, to examine if digital social networks are effecting the want and ability of the subject to create new links within a physical human social network. In simplified terms, what effect is Facebook having on this subjects social network within new environments located at a distance (+100 km) from the previous base of the social network, and does this conform with Dunbar's [4] network limit size theorem.

This research is of both economic and social significance. Socially, the results of this investigation will contribute to ongoing the research in the implications of digital social networks in regards to the ratification or refutation of Dunbar's [4] numbers. Recent investigations have focused on the analysis of social networks from a single point, whereas this investigation has focused on social network analysis using geospatial distance as an influencing factor. The understanding of this problem can help improve the size, the quality and the speed of creation of new physical social networks improving a subject's life both physically and emotionally [7]. This research also has economic implications for various reasons. The better understanding of digital social networks allows for the improvement movement of information, skills and resources. The improved movement and logistics of these attributes allows for a better efficiency and improvement to all. Improved social network analysis also has many other economic benefits [8]. According to software engineering company [9], social network analysis has vastly improved customer targeting, brand awareness, the detection of health care fraud and many others. Improvements in these areas helps develop the free market to which global economies are based.

Taking these factors into consideration, the research question/study objective of this investigation is: What effects are digital social networks having on new physical human interactions/bonds in regards to the proximal distance of existing social interactions? Do these results confirm to the network limit size theorem proposed by Dunbar [4] [5].

The research objective suggests that the null and alternative hypothesis should be:

Ho: Digital social networks are not having an effect on new physical human interactions/connections in regards to distance.

Ha: Digital social networks are having an effect on new physical human interactions/connections in regards to distance.

2. Data Description and Network Construction

The data used in this network was obtained from the digital social network Facebook. In its simplest form, the data comprises nodes and links, nodes being Facebook user profiles "friends" and links being a digital friendship connection between two Facebook user profiles. A link is formed by a node (user profile) manually confirming the link between, by accepting a friendship request. This is the only method of link formation, unlike other digital social networks like Twitter, where a user can create a link without the confirmation of the user in question. This gives Facebook a closer relationship to a real human social network, as a physical confirmation is required before a link is formed.

The data was acquired from Facebook via the application Netvizz v1.01. This retrieves a user's Facebook

friend list (the nodes) and the mutual connections between nodes within this network. The data was formatted into a matrix database GDF file similar to that of a coma separated file (CSV), also allowing for the support of attribute data of both the nodes and edges, making this format ideal of digital social network analysis. The user profile under analysis was created 09/11/2006 and contains 619 nodes (friends) and 8805 edges (mutual friendship connections). The radical relocations occurred twice, firstly August 2012, just under 6 years after the creation of the profile) and secondly August 2014, just less than 8 years after the profile was created. The network data itself removes the user profile edges from the data, *i.e.* the analysis excludes all the links/ties that include Ego (the profile account holder) as to not falsely skew the statistical network analysis results.

The digital social network data acquired for this investigation is the data of the author's (as of this point known as the user) personal Facebook. The data in question is ideal in regards to fulfilling the study objective of this research (the effects of digital social networks on real social networks in regards to geospatial distance), as the user in question has undergone multiple radical geo-relocations at various periods from the creation of the user Facebook profile. The separate geo-relocations were to areas where the total number of existing social network links, both physical and digital, was equal to zero. That is, the user moved to locations where they did not have previous social connections with a single person of any kind. Additionally the data contains zero fake user profiles (bots/spam), a common factor in today's digital social networks [10]. Each node is the user profile of that specific person, known to the user in the physical world. These factors make the digital social network data of this user Facebook profile ideal for this investigation.

3. Methodology of Network Analysis

The digital social network data was implemented into the network analysis and projection software Gephi (version 0.8.2—development tool). A series of network measures and projections were undertaken to analyze both the total network and sub-networks to meet the study objectives of this research. As mentioned previous, a physical confirmation needs to be undertaken for a link between two nodes to occur, this therefore implies that the network is undirected, as the confirmation implies a mutual agreement of friendship. Therefore all network measures are performed on the undirected network.

Initial network measures were run using the Gephi network analysis tools. These include: Network density measures (average degree, total/in-degree/out-degree distribution plots), the clustering coefficient metrics (average clustering coefficient, total triads and the distribution), modularity (size distribution and number of communities, centrality measures (diameter, average path length, betweenness/closeness/eccentricity distribution reports) and number of connected components (weakly-undirected network). The calculation of these standard network measures allows for both the analysis of the network itself but also gives meaning to projection algorithms when used to label and identify nodes and links (using both color and size).

The main objective of this research is to analyze the effect of digital social networks on physical social interactions in regards to distance. Therefore the identification of communities within the network is the key factor in the analysis. Using the modularity (the fraction of the edges that fall within the given groups minus the expected number of edges if distribution of edges occurred randomly) of each node as a method of community detection nodes can be colored in regards to groups.

With clusters/communities detected and identified separation algorithms were implemented to project nodes into their designated communities, for further node analysis. Two separate algorithms were applied to the network data; the Force Atlas and the Fruchterman Reingold algorithm. The Force Atlas algorithm applies inertia, repulsion, attraction and gravitational (location attraction) forces to nodes and links to give motion to the network to project nodes into communities with edges similar lengths. The Fruchterman Reingold [11] algorithm is similar in nature to the Force Atlas (force-directed layout algorithm) but differs as a maximum allowable area (circular) is imposed on the projection, as a defined by the number of nodes. Upon completion of the algorithms, the communities identified were split into their geographic designation and further analyzed against time (denoted by the user's total time within the location). The communities against time period calculations are then compared, to analyze the effect (if any) digital social networks are having on physical social interactions in relation to time.

4. Results

Table 1 displays the results of various selected network measures. As a mandate for the data acquired for this

Table 1. A table displaying the results of selected network measures.

| Network Measure	Result
Density	0.046
Average Weighted Degree	14.225
Average Clustering Coefficient	0.596
Total Triads	88750
Modularity	0.609
Communities/Clusters	21
Average Path Length	4.469
Connected Components	13

investigation, is the movement of user to areas where zero previous network connections exist, a low network density of 0.046 is expected. 619 nodes gives the number of potential connections to be 191271, a vast number of connections compared with the actual number of social friendship links (8805), further network measures displayed in **Table 1** are discussed further below (Discussion and Conclusions section). **Figure 1** (below) displays an initial projection of digital social network data (top) and this initial projection with the modularity (bottom) used to identify clusters (with these separate clusters colored differently). **Figure 2** (below) displays the final result of the separation algorithms applied to the social network data (Force Atlas—top, Fruchterman Reingold—bottom), animations of these algorithms (**Figure S1** and **Figure S2**) can be found in the Supporting Figures section.

Figure 3 (below) gives the results of different network measures used to alter the size and color of the nodes, three of the images have previously been run with the Fruchterman Reingold separation algorithm, whereas The bottom right image has been projected using the Force Atlas algorithm. Image 1 (top left) is a display of degree density (high density = larger node, darker color). Image 2 (top right) is a representation of node Eigenvector Centrality score (higher importance = larger node, darker color). Image 3 (bottom left) is a representation of Eccentricity Centrality (node size—number of links, node color eccentricity score). Image 4 (bottom right) is a representation of sex assigned to each node (blue male, pink female, green unidentified).

Table 2 contains the results of various network measures run on network partitions indicated by the 3 geo-locations as specified by the user (UK, Canada and USA). The forced partitioning algorithms confirmed these separate clusters/communities. Simple division into categories by time is not a suitable option, due to the ability of the user/users social network to create links at any time, regardless of geolocation. The last row of **Table 2**. gives an average number of friends created per day, this is calculated by dividing the number of nodes (friends) within these sub-networks, by the total number of days that sub-network has existed. Time of sub-networks is not concurrent, as the subject has left the geo-locational base of the sub-network. Therefore once radial movement has occurred; the membership of the sub-network is closed. Comparisons between results will be discussed further below (Discussion and Conclusions section).

5. Discussion and Conclusions

The results of the network analysis, both of the total network itself and sub-networks, provides interesting results into the effects of digital social networks (in this case Facebook) on physical human social networks when the user incurs a radial geo-locational movement. As **Table 2** shows, the user creates new extensive physical social networks when relocates to a new location meaning leading to two possible conclusions. Either digital social networks have little to effect on the user's ability/want to create new friendships or aid the user in new friendship creation (link creation) by previous familiarity with a previously unknown person by association. In any case, the data suggests that digital social networks are having a neutral or a positive impact, not a negative impact. This analysis confirms Dunbar's [4] [5] network limit size theorem, as the number of nodes is close to his predicted figure. The data suggests that there should be a difference in the quality of the sub-social networks. The average weighted degree difference and the average Clustering Coefficient for the UK sub-network,
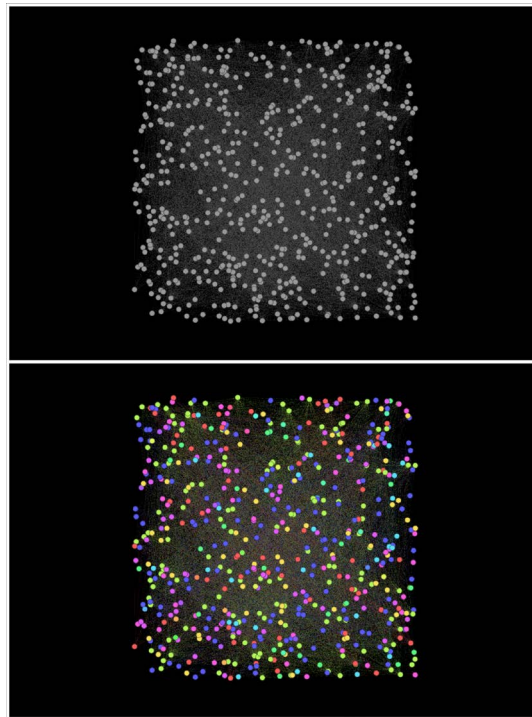
Figure 1. Two images showing the initial display of the network with no layout alterations. Facebook network data visualized using the software Gephi V0.8.2. The bottom image has assigned color to each nodes using modularity as a method of node community detection.
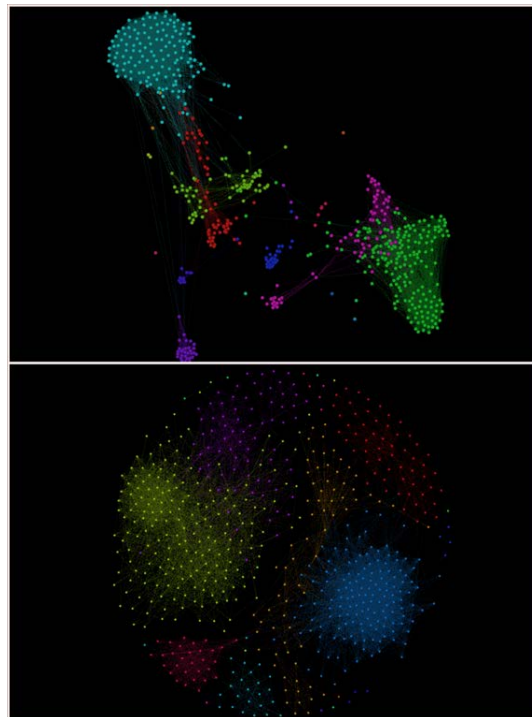
Figure 2. Two images showing the results of the separation algorithms. The top image displays results of the Force Atlas algorithm and the bottom, Fruchterman Reingold. Facebook network data displayed after implementation through various separation algorithms. Node cluster colors differ between each image.
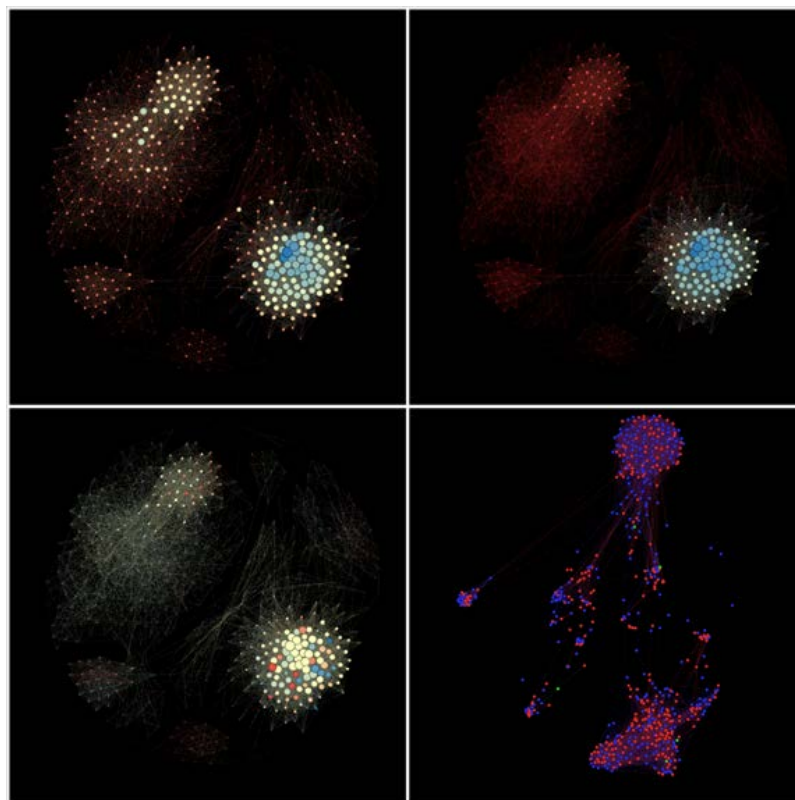
Figure 3. Four images displaying the results of node alteration to represent different network measures. Facebook network data implemented through separation algorithms, with node color and size representative of different network measures.

Table 2. A table displaying the results of selected network measures on nodes and links that had been partitioned into geographic communities (UK, Canada and USA).

Network Measure	UK	Canada	USA	Network Total
Network Duration (Days)	2190	730	96	3016
Nodes	314	261	44	619
Links	5446	3211	134	8805
Density	0.125	0.097	0.142	0.046
Average Weighted Degree	36.725	24.891	6.091	14.225
Average Clustering Coefficient	0.698	0.464	0.744	0.596
Modularity	0.332	0.436	0.232	0.609
Diameter / Radius	7/4	6/3	3/0	12/0
Average Path Length	2.807	2.355	1.643	4.469
Connected Components	1	1	13	13
Friendship Rate (Network Duration/Number of Nodes)	0.143	1.135	0.458	0.205

are much higher than the Canadian sub-network. This means that on a network level, each node shares more links with other nodes, and on a social network level, the user's friends are more likely to share a social connection.

Price [12] has suggested that digital social networks are having a negative effect on physical human social networks by a user resorting to online communication with old links rather than pursuing new social friendships

(nodes and links), but the results from this investigation prove otherwise in a certain context. One theory can be that both the results of this investigation and Price [12] are correct, where the analysis differs is that the user from this investigation is relocated to an area where no previous social links exist and the distance from the previous network (UK sub-network) is extremely far (2700 miles/4345 kilometers). As the results from the network analysis indicate, both the UK and Canadian sub-networks are fully matured and thriving. Therefore it is proposed that digital social networks can have an effect on physical human social networks but distance between sub-networks is a very important factor. This theory makes logical sense considering the following two examples. The relocation to an area 1 hour away (by car—100 miles) from the current physical social network base will not require the user to create a new sub-network of social links. Due to the short distance, the user can easily maintain contact with previous nodes and have physical social interactions due to the short distance (both users traveling 30 minutes towards each other needs little effort). Relocation by the user to an area 10 hours away (by plane—2000 miles) will mean the ease of physical contact with previous nodes is removed (due to high cost/time), therefore the user is forced to create a new social sub-network to avoid loneliness. These logical examples agree with the theory that digital social networks are not having an effect on physical social networks but only over a certain distance. Further research will be required to calculate this distance, which more than likely varies from person to person depending on social characteristics (introvert/extrovert etc).

According to Backstrom *et al*. [13], a social network requires a certain amount of time to develop. Therefore comparisons with the results gained from the USA sub-network will be premature, as the network is still in its infancy. But the information from the analysis of this network is useful in regards to the processes that start when a new sub-network is formed. The data indicates that the creation of a social network can follow a Logistic Sigmoid curve when friendship number is plotted over time. When considered as the logical shape, the graph begins with initial friendship creations. This is currently highlighted by the number of individual weakly connected components within the USA sub-network. Friendships are initially created through randomized social interactions, meaning a minimal chance of these initial friendships sharing a link. As time progresses nodes and links exponentially increase as friends of friends are incorporated into the user's social network, then this rate flattens out as Dunbar's [4] numbers begin to take effect, as displaying with the mature UK and Canadian sub-networks.

The forced separation algorithms also provide very useful information when considering what effect digital social networks are having. The results from **Table 2** indicate that both the UK and Canadian sub-network are one connected component, when analyzed together it is also found to be one connected component. The forced separation algorithms allowed for the identification of the single link that formed a connection between the two sub-networks (exchange student from UK). Without the separation algorithms, the location of this link would have been close to impossible to find (visually) given the vast amounts of nodes and links. The visualization of data forms an important part of network analysis as it provides context to individuals without understanding of network analysis concepts (such as centrality measures).

In conclusion, the results of this investigation suggest that digital social networks are having either a positive or no effect on physical social networks in relation to distance. It is theorized that due to conflicting results of investigations, there is a possibility that digital social networks such as Facebook may affect physical social networks but only when relocation distance is small. In addition the results of the analysis of the separate sub-networks confirm Dunbar's [4] network limit size theorem.

There are several options to be considered for future work. This investigation is focused on an egocentric network *i.e.* the individual digital social network of the investigator; performing identical analysis on multiple digital social networks, which have also undergone radical relocation and comparing the results of the standard network measures, is the next logical step in the analysis of effect of digital social networks on physical social networks.

6. Supporting Information

Figure S1. An animation displaying the Force Atlas separation algorithm.
https://www.youtube.com/watch?v=ibdqTyZYSvc
Figure S2. An animation displaying the Fruchterman Reingold separation algorithm.
https://www.youtube.com/watch?v=v1ZZFPmuAWo

Acknowledgements

References

[1]   Phillips, S. (2005) A Brief History of Facebook. The Guardian.
       http://www.theguardian.com/technology/2007/jul/25/media.newmedia. Accessed 11 November 2014

[2]   Miniwatts Marketing Group (MMG) (2014) Internet World Stats, Estimated USA Internet Users.
       http://www.internetworldstats.com/unitedstates.htm

[3]   The Nation (2012) Social Media Now A Part of Everyday Life. Nation Multimedia 24 January 2012.
       http://www.nationmultimedia.com/technology/social-media-now-part-of-every-day-life.html

[4]   Dunbar, R. (1992) Neocortex Size as a Constraint on Group Size in Primates. *Journal of Human Evolution*, **22**. 469-493. http://dx.doi.org/10.1016/0047-2484(92)90081-J

[5]   Dunbar, R. (1993) Co-Evolution of Neocortex Size, Group Size and Language in Humans. *Behavioral and Brain Sciences*, **16**, 1-27. http://dx.doi.org/10.1017/S0140525X00032325

[6]   Goncalves, B., Perra, N. and Vespignani, A. (2011) Modeling Users' Activity on Twitter Networks: Validation of Dunbar's Number. *PLoS ONE*, **6**, e22656. http://dx.doi.org/10.1371/journal.pone.0022656

[7]   Kadushin, C. (2012) Understanding Social Networks: Theories, Concepts and Findings. Oxford University Press, New York.

[8]   Albarra, A. (2013) The Social Media Industries. Routledge Publishers, New York.

[9]   Bowes, P. (2012) Sphere of Influence—The importance of Social Network Analysis.
       http://www.pb.com/docs/US/Software/Industry-Pages/Telecommunications/churn-management/PDFs/Sphere-of-Influence-Importance-Social-Network-Analysis_WP.pdf

[10]   Wang, A.H. (2010) Detecting Spam Bots in Online Social Networking Sites: A Machine Learning Approach. *Lecture Notes in Computer Science*, **6166**, 335-342. http://dx.doi.org/10.1007/978-3-642-13739-6_25

[11]   Fruchterman, J. and Reingold, M. (1991) Graph Drawing by Force-Directed Placement. *Software*: *Practice and Experience*, **21**, 1129-1164. http://dx.doi.org/10.1002/spe.4380211102

[12]   Price, E. (2011) Does Facebook Affect How You Interact with People in the "Real World".
       http://www.technobuffalo.com/2011/06/08/does-facebook-affect-how-you-interact-with-people-in-the-real-world-infographic

[13]   Backstrom, L., Huttenlocher, D., Kleinberg, J. and Lan, X.Y. (2006) Group Formation in Large Social Networks: Membership, Growth and Evolution. Database Applications. KDD06.
       http://www.cs.cornell.edu/~lars/kdd06-comm.pdf