

# New Kid on the Block: Exploring the Google+ Social Graph

Gabriel Magno

UFMG

Belo Horizonte, Brazil  
magno@dcc.ufmg.br

Giovanni Comarela

UFMG

Belo Horizonte, Brazil  
giovannicomarela@dcc.ufmg.br

Diego Saez-Trumper  
Universitat Pompeu Fabra  
Barcelona, Spain  
diego.saez@upf.edu

Meeyoung Cha  
KAIST  
Republic of Korea  
meeyoungcha@kaist.edu

Virgilio Almeida  
UFMG  
Belo Horizonte, Brazil  
virgilio@dcc.ufmg.br

## ABSTRACT

This paper presents a detailed analysis of the Google+ social network. We identify the key differences and similarities with other popular networks like Facebook and Twitter, in order to determine whether Google+ is a new paradigm or yet another social network. This work is based on large-scale crawls of over 27 million user profiles that represented nearly 50% of the entire network in 2011. We observe that the average path length between users is slightly higher than other networks, possibly because Google+ is a new system where relationships are still rapidly growing. Google+ shows a higher level of reciprocity than Twitter, which also has directed social links. The newly available “places lived” field could be used to study how users are distributed around the world and how aggressively the service has been adopted in different countries. We find that Google+ is popular in countries with relatively low Internet penetration rate. Based on the amount and types of information publicly shared in user profiles, we also find that the notion of privacy varies significantly across different cultures.

## Categories and Subject Descriptors

J.4 [Computer Applications]: Social and behavioral sciences Miscellaneous; H.3.5 [Online Information Services]: Web-based services

## General Terms

Human Factors, Measurement

## Keywords

Google+, Online Social Network, Geo-location

## 1. INTRODUCTION

Social networks are a global information infrastructure, where individuals bring their social relations online and share information,

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IMC'12, November 14–16, 2012, Boston, Massachusetts, USA.  
Copyright 2012 ACM 978-1-4503-1705-4/12/11 ...\$15.00.

photos, songs, videos, as well as ideas. Social networking sites like Facebook now reach 82% of the world’s Internet-using population or about 1.2 billion people in total according to comScore [11]. In fact social networking became the most popular online activity worldwide. Accordingly, a number of researchers have tried to understand user behaviors and characteristics of various online social networks, where Twitter and Facebook have been the two most popularly examined platforms [26, 7, 39, 3].

To compete in this field, Google has launched in June 2011 its own social networking service called Google+ (<https://plus.google.com/>). The platform was announced as a new generation of social network and included several new features, such as *circles* that allow users to share different content with different people and *hangouts* that let users to create video chatting session and invite up to nine people from their circles of friends to share the environment [22].

Since its launch, the Google+ social network has been adding new users at a rapid pace. In fact, it is known as the fastest growing network ever, reaching 20 million visitors in only 21 days [10]. The service has later reached 62 million registered users as of December 2011 [24] and a total of 250 million registered users of whom 150 million are active as of June 2012 [15].

Once Google+ has become a popular social media network, it is important to understand how it compares to other social network models. Typical questions follow. How are people connected on Google+? Who are the most popular users? How are users distributed worldwide? What is the impact of geography on the social relationships?

Furthermore, the rapid adoption rate of the service raises interesting questions about online privacy. One crucial question is on what the default privacy settings should be. Along these lines, it is worth examining how “closed” social networking sites are, compared to the “open” Internet. Google has positioned itself as promoter of the Internet openness against other social networking services that are often described “walled garden” [5] due to limited access to their internal web pages. Then, is Google+ different? How open is it and how does it impact user interactions?

To answer these questions we have crawled more than 27,556,390 user profiles and 575,141,097 relationship links among users, as well as any publicly available data about the users such as gender, geo-location, and relationship status. A relatively large number of users leave personal information publicly available for anyone to see. This kind of information allows us to analyze user behavior patterns and compare them to previous research results obtained for other social networks, e.g., Facebook and Twitter.

Based on the gathered data, we characterize the novel social network model provided by Google+ in depth, its user base, its geographical distribution, and compare its main characteristics with other social network services. Among various findings, some of the main results are summarized as follows:

1. Our analysis on the top users based on the circles list indicate that the majority of the top users (7 out of 20) are well-known individuals from information technology industry;
2. By looking into users who share their work or home contact information publicly (1% of all users), we observe that a large fraction of the users who share telephone numbers are male and single;
3. We find that users share strikingly different amounts of information to public in their profiles depending the country they live in;
4. By examining the social links between the users in relation to their countries, we observe that physical distance is crucial in the likelihood of forming a social link between two users;
5. The fraction of global and national links also vary according the countries, indicating the different patterns of usages of the Google+ service across different cultures.

As another contribution of this paper, we make the collected data available to the wider research community at <http://gplus.camps.dcc.ufmg.br>. Our Google+ data set could facilitate new projects in social computing and computer network research that need actual data for their experiments.

The rest of this paper is organized as follows. We begin with a detailed description of the Google+ platform and the data collection methodology in Section 2. In Section 3, we study the social graph of the Google+ users and describe in detail the public attributes available in Google+. Next, we analyze the impact of geographical distance and relationships in Google+ in Section 4. We examine the economic indicators of a country and their relationship to the adoption rate of Google+ within that country. We discuss related work and implications in Section 5 and Section 6. Finally, we summarize our findings and conclude in Section 7.

## 2. METHODOLOGY

We briefly describe key features of the Google+ service and the way we gathered data.

### 2.1 Platform Description

The Google+ service was released in June of 2011 [22]. In the first 90 days, the service has been on field trial and only those users who received an invitation could create an account. During this time, the network grew *virally* through social contacts. In September 20th, 2011, the service became publicly open and no invitation was required for a sign up [21]. These two different mechanisms of spreading would have attracted different kinds of users to Google+. For instance, users who joined through invitations are likely tech-savvy users who typically adopt new services early, compared to the users who join through open sign-up.

In Google+, users can manage their contact list through *circles*. Circles are labeled groups of friends, which allows a user to share or receive information with and from a specified subset of his contacts. For example, a user may manage “family”, “colleagues”, and “alumni” circles. When a user adds someone in one of his circles, he starts to receive updates from that person. This manual grouping of contacts alleviates some of the privacy problems that existed in

other “flat” social networks, where default privacy settings are set to maximize the visibility of users profile and only a small number of members change it [20]. There are two types of circles, namely in- and out-circles. While the latter represents the list of users that a given user has added to her circles, the former represents the list of other users who added that user to their circles.

Circle names and their user lists are private information that only the circle creator can see. A user can identify all the others who included the user in their circles (i.e., followers), because the user receives a notification when someone adds him to a circle. Similar to Twitter, people can add other users to their circles without confirmation. This is different from networks like Facebook, where all social links are reciprocal and both sides of the users should agree to own a social link. In the user profile page two lists are shown by default: The “Have user in circles” list, like a followers list, and “In user’s circles”, similar to the followees list in Twitter. The user has the option to set these lists as private.

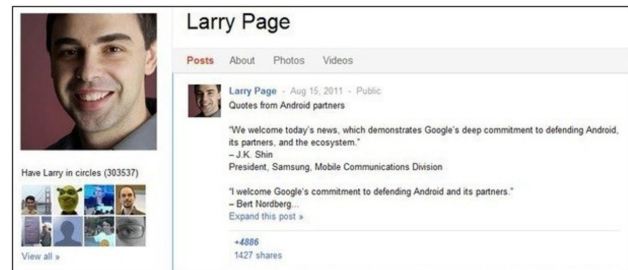


Figure 1: Google+ home page of Larry Page.

Google+ users can publish ideas (status), images, videos and any kind of URL. Whenever a user post something, she has the option to set the visibility of that content, choosing which of her circles can see it. On the other side, a user can choose from which circles she will receive content. Therefore, circles are the way to manage information flow in Google+. The continued information flow through circles is referred to as “stream” in the system.

There are several features that allow users to interact with others. User interactions are centered around content; users can comment, share (like retweet in Twitter), and click on the “+1” button (similar to Like button in Facebook) on a given content. When a user clicks on the “+1” button, she is publicly recommending that particular content to others and it will be saved in her “+1’s tab” similar to bookmark. There are other features such as photo albums (that allow users to upload, share and organize photos), hangout (a kind of collaborative video chat with friends), and games.

### 2.2 Data Collection

In order to collect user profiles in Google+, we implemented a breadth-first search (BFS) crawler in Python, considering both the public in-circles and out-circles lists (i.e. bidirectional BFS). We began our crawl with Mark Zuckerberg, the co-creator and chief executive of Facebook, because he was known to be one of the most popular users in Google+ at the time of data collection. Given that users are connected in social networks, our crawler although started from a single seed node soon reached other popular users in Google+. We could not repeat the crawl with randomly chosen seed nodes, because numeric user IDs were not supported at the time of data collection. Although the BFS technique is simple and efficient, it exhibits several well-known limitations such as the bias towards sampling high degree nodes, which may affect the degree distribution [18, 35].

The data collection process started on November 11th, 2011, and ended on December 27th, 2011. We used a total of 11 machines with different IP addresses to efficiently gather large amount of data. The profile information was retrieved by making HTTP requests to publicly available user profile pages. In total we crawled **27,556,390** profile pages, collecting public user information and its circles lists. With the social links of the users we have constructed a directed graph that has **35,114,957** nodes and **575,141,097** edges. As of the data collection date, we estimated that our data set represented 56% of all registered Google+ users[2].

There is a limit on the maximum number of users that could appear in any public circle, which is 10,000 users. Since the Google+ social graph was gathered in both directions (in-circles and out-circles), we were able to recover almost all “lost edges.” In order to estimate the fraction of missed links, we compared the number of users shown in their profile page with the actual number of edges we collected. In our dataset there are 915 users with more than 10,000 in-circles users, which should have 37,185,272 incoming edges according to their profile pages, while we found 27,600,503 links for those users in our graph. By dividing the difference of these numbers by the total number of edges, we estimate that 1.6% of the edges are lost because of the 10,000 limit on the circle list.

### 3. GRAPH ANALYSIS

In order to characterize social relationships of Google+ users, we define a social graph. The vertices of the social graph are Google+ users present in our dataset. A user  $v$  added by user  $u$  to her circles results in a directed edge from  $u$  to  $v$ . Therefore the social relations among Google+ users make a directed graph  $G(V, E)$ , where  $V$  represents the set of users and  $E$  is the set of directed edges  $(u, v)$ ,  $u, v \in V$ . Given the social graph construction, we analyze two types of properties: first on the node characteristics and then on the graph structure. The former capture the characteristics of Google+ users, as defined by the fields of the user profile, while the latter represents relationships between users.

**Table 1: Top 20 users ranked by in-degree**

Rank	Name	About
1	Larry Page	IT (Google)
2	Mark Zuckerberg	IT (Facebook)
3	Britney Spears	Musician
4	Snoop Dogg	Musician
5	Sergey Brin	IT (Google)
6	Tyra Banks	Model
7	Vic Gundotra	IT (Google)
8	Paris Hilton	Socialite
9	Richard Branson	Businessman (Virgin Group)
10	Dane Cook	Comedian
11	Jessi June	Model
12	Trey Ratcliff	Blogger
13	will.i.am	Musician
14	Felicia Day	Actor
15	Thomas Hawk	Blogger
16	Tom Anderson	IT (Myspace)
17	Pete Cashmore	IT (Mashable)
18	Guy Kawasaki	IT (Apple) & Writer
19	Wil Wheaton	Actor & Writer
20	Ron Garan	Astronaut (NASA)

### 3.1 Node Characteristics

To get a sense for what users expect from the Google+ service, we first examine who the most popular users are. Table 1 shows the top 20 users based on their in-degrees (i.e., how many circles these users are added to by others). The top list of Google+ is a mix of singers, bloggers, actors, and IT professionals. When we compare the list to that of Twitter in [26], only one user appears in both lists; singer Britney Spears ranked second as of mid 2009. The top list is particularly different from that of Twitter in that (1) we do not see any news media outlet like the New York Times and CNN, while (2) we see founders of large Internet-based companies like Google and Facebook. In fact 7 out of the 20 users are IT related, which is uncommon in other social networks.

Next focusing more on the average user, we examine what kinds of interactions they perform on the network. In general, users of social networking sites reveal different types of personal information in their profile, such as basic descriptors (e.g., gender, relationship status, cities lived), contact information (e.g., e-mail, phone number, address, Web site), personal interests (e.g., favorite TV shows, movies, books, quotes, music), education information (e.g. field of study, degree), work information (e.g., employer, position), etc.

Google+ users also publish information about themselves in their profiles. Some pieces of information are in “restricted fields”, where users have to choose among some options, while in “open fields” users can write anything they want. Only the fields “relationship”, “looking for”, and gender are restricted fields. The rest of the fields are open fields. In the field called, places lived, a user can write the name of any place she lived and the Google+ system automatically tries to mark the place on the map.

For all the fields, except for the name that is public by default, a user can control the privacy setting and set visibility of that field. There are five options: (1) *public*, which means open to anyone in the Internet, (2) *extend circles*, which means open to people that are in circles and people that are in the circles of those, (3) *your circles*, which means open to people in one’s circles, (4) *only you*, and (5) *custom*, which means a user can choose exactly which circles may view that field.

We have collected information about all the fields of users that were publicly accessible. In Table 2, we show the number and fraction of users that have made each type of information available.

**Table 2: Public attributes available in Google+**

Attribute	Available	%
Name	27,556,390	100.00
Gender	26,914,758	97.67
Education	7,471,191	27.11
Places lived	7,371,461	26.75
Employment	5,917,609	21.47
Phrase	4,075,132	14.79
Other profiles	3,713,546	13.48
Occupation	3,656,447	13.27
Contributor to	3,622,627	13.15
Introduction	2,149,191	7.80
Other names	1,210,760	4.39
Relationship	1,186,903	4.31
Braggin rights	1,074,964	3.90
Recommended links	1,001,349	3.63
Looking for	753,704	2.74
Work (contact)	60,434	0.22
Home (contact)	58,876	0.21

## 3.2 Privacy concerns

Studies on human behavior [38] show that individuals with profiles on social networking sites may have greater risk taking attitudes than those who do not. And sharing contact information, such as telephone numbers, may increase risks. As far as contact details are concerned, the work in [38] shows how many Facebook users disclose identity information in the form of contact details. The majority of the users in the sample used by the study publicly showed their e-mail address (64.1%). Only a few Facebook members published their mobile phone number (10.7%). Similarly, only a minority (10.7%) of the participants revealed their home address on Facebook.

Google+ allows their users to publish contact information in their profiles. Some users publicly share their work or home contact information. In our data set, a total of 72,736 users share telephone number in Google+, which represent 0.26% of the population. We call these users *tel-users* and because they represent a class of risk taking users we look into the details of the profile of these users.

In order to examine how much information *tel-users* share publicly compared to all users, we show the CCDF (Complementary Cumulative Distribution Function) of the number of fields in the profile shared for each user in Figure 2, removing the fields of Home and Work information from the contabilization. (The list of the fields available are given in Table 2.) As we can see, *tel-users* generally share more information in their profiles than other Google+ users, which confirm their risk taking attitude. For example, 10% of all Google+ users share more than six fields, while 66% of the *tel-users* do the same.

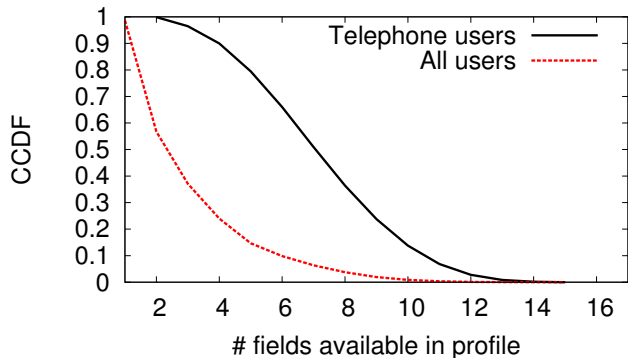


Figure 2: Number of fields shared by users in the profile

Concerning the information sharing behavior of Google+ users, table 3 displays the percentage of users who give information about gender, relationship, and location for all users and *tel-users*, considering only those users that had the field public. Among all users of the dataset, 68% are male and 31% are female. However, the difference is much higher when we consider *tel-users*; 86% are male and 11% female, indicating that female Google+ users are less likely to share phone numbers than male Google+ users. Similar to the observations confirmed in [16], more risk taking behaviors can be found for men and greater concern from women with regard to information provided on the Web.

What is particular about Google+ is that it asks users to provide a very detailed level of information about their relationship status as opposed to other social networks. The nine default options from which users can choose from are listed in the table. Conducting the same comparison of all users and *tel-users* over the relationship status, we find that user behaviors are similar between the two groups.

Table 3: Information shared by all users and *tel-users*

	All users	Tel-users
Total	27,556,390	72,736
Gender (N)	26,914,758	71,267
Male	67.65%	85.99%
Female	31.46%	11.26%
Other	0.89%	2.75%
Relationship (N)	1,186,903	29,068
Single	42.82%	57.24%
Married	26.59%	21.03%
In a relationship	19.80%	10.23%
It's complicated	3.16%	3.98%
Engaged	4.39%	2.98%
In an open relationship	1.26%	2.77%
Widowed	0.50%	0.58%
In a domestic partnership	1.08%	0.77%
In a civil union	0.39%	0.41%
Location (N)	6,621,644	45,676
United States	31.38%	8.92%
India	16.71%	31.90%
Brazil	5.76%	4.72%
United Kingdom	3.35%	2.19%
Canada	2.30%	1.52%
Other	40.50%	50.77%

However, those users who set their relationship status as “single”, “it’s complicated”, “in an open relationship”, “widowed”, and “in a civil union” were more likely to share their phone numbers publicly than others. In particular, we saw a high percentage of single users (57.24%) compared to all the users (42.82%). In contrast, only half of the users “in a relationship” shared their phone numbers.

The fraction of *tel-users* does not follow the rank of the top 10 countries in Figure 6. While the US take up 31.38% of all users, it counts for only 8.92% of those users who have made their phone numbers available in Google+. In contrast, India now becomes the most populated country based on the fraction of *tel-users* count (31.90%). The fraction of Indian users in the *tel-users* group is twice as big as in all other country users group.

While the different level at which users of a given country reveal their phone numbers is interesting, this may come as no surprise when we account for the fact that people’s perception of what is “private” is different. According to a report in [9], 65% of people in Germany find mobile phone number as personal, whereas only 28% of people in Romania think the same.

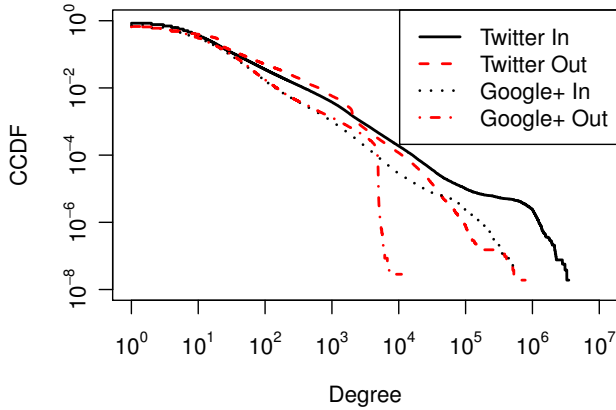
## 3.3 Structural characteristics

We next present characteristics of the Google+ social graph. For each network metric, we also show the results for the Twitter graph from other research [7] for comparison.

### 3.3.1 Degree Distribution

One of the most common structural measures analyzed in complex networks such as the Google+ social graph is the distribution of the number of the incoming and outgoing node connections or what is so called “degree.” Figure 3 shows the CCDF for the variables out-degree and in-degree of the Google+ social graph. We can see that these curves have approximately the shape of a Power Law distribution. The CCDF of a Power Law distribution is given by  $Cx^{-\alpha}$ ,  $x, \alpha, C > 0$ . If we compare the curves with Twitter, we observe similar patterns, although Google+ shows slightly lower degrees.

By using a simple statistical linear regression (in the log-log scale) we estimated the exponent  $\alpha$  that best models the data. We obtained  $\alpha = 1.3$  (with  $R^2 = 0.99$ ) for in-degree and  $\alpha = 1.2$  (with  $R^2 = 0.99$ ) for out-degree. Although both curves are similar, the out-degree curve drops sharply around 5000. We conjecture this is because Google maintains a policy that allows only some special users to outpass a specified threshold (unknown) and add more than 5000 friends to their circles.



**Figure 3: Degree Distributions**

The overall power-law trend shown in the degree distribution is characteristics of the human social networks. This means that a small fraction of the individuals have disproportionately large number of neighbors, while most users have a small number of neighbors. Examples of such hubs are listed in Table 1. For instance, Larry Page is listed in 3.7 million circles as of August 2012, while the majority are listed in no more than 10 circles. As studied in many other research, hubs play a central role in information propagation in social networks.

### 3.3.2 Reciprocity

In order to evaluate discrepancies between in- and out-degrees for a given node we use a simple metric called Relation Reciprocity ( $RR$ ) of a node  $u \in V$  as:

$$RR(u) = \frac{|OS(u) \cap IS(u)|}{|OS(u)|} \quad (1)$$

where  $OS(u)$  is the set of nodes (i.e. users) that have an incoming edge from  $u$  and  $IS(u)$  is the set of nodes with outgoing edges pointing to  $u$ .

Figure 4(a) shows the distribution of the Relation Reciprocity. This metric is able to effectively differentiate very popular users, such as celebrities and companies, with very low reciprocity from ordinary users, that have moderate to high  $RR$ .

The analysis of the relation reciprocity in circles links suggests a strong signature of the Google+ users. More than 60% of the users have  $RR$  higher than 0.6, which shows some sort of structural balance between the users of this new social networking service. This concept is related to the fraction of reciprocal relationships in a social graph. Based on our datasets, we calculate the percentage of global reciprocal relations and find 32% in Google+, compared to 22.1% reported for Twitter [26].

The high reciprocity rate in Google+ may be related to the scarcity of large media outlet profiles, since by the time of our collection the “pages” feature was new and not very popular. Media

outlet attract very large numbers of followers, but do not exhibit followees. Different kinds of online sharing services exhibit higher reciprocal relations, such as 68% for Flickr [8] and 84% for Yahoo! 360 [25].

### 3.3.3 Clustering Coefficient

Another common characteristic of social networks is a high average clustering coefficient ( $CC$ ). The  $CC$  of a node  $u$ , denoted by  $C(u)$ , is defined as the probability of any two of its neighbors (outgoing) being neighbors themselves [41]. This metric is associated to the number of triangles that contain a node  $u$ . For a directed graph, the maximum number of triangles connecting the  $|OS(u)|$  outgoing neighbors of  $u$  is  $|OS(u)|(|OS(u)| - 1)$ . Thus, the  $CC$  measures the ratio between actual triangles and their maximal value. During clustering coefficient analysis we only consider the nodes with  $|OS(u)| > 1$ , since this is a necessary condition for this computation.

Figure 4(b) shows the distribution for the  $CC$  of nodes in the social graph. We randomly sampled one million nodes and computed  $CC$  for each one of them. We can see that 40% of all users have a  $CC$  greater than 0.2. An approximate calculation, based on the results presented for Facebook in [39], allows us to estimate the clustering coefficient for part of Facebook population. In [39], we have that only users with degree smaller than 50 have an average  $CC$  greater than this value. However, these users represent less than 1% of the entire network [39], suggesting that Google+ has a higher average cluster coefficient than Facebook, which represent a more tightly connected network. Comparing with Twitter (as shown in Figure 4(b)), we can also see higher values of  $CC$  in Google+.

### 3.3.4 Strongly Connected Component

The study of the connected components of a social graph is a key factor to understand its structural properties. For example, if we know the WCCs (Weakly Connected Components) of a graph then we have information about the number of isolated nodes in the network as well if it has a giant component. Due to the data collection procedure used to crawl Google+ in this work, which is a bidirectional snow ball, the social graph  $G$  consists of only one WCC. As a consequence, we do not have information about the isolated nodes, but we know that we have collected a giant component of the graph (see Table 4).

In order to investigate the connectivity of  $G$  we decided to measure the number and size of all its Strongly Connected Components (SCC). A strongly connected component of a social graph (directed) is a subgraph, such that a node can be reached from any other node following edges between them. SCCs have an important role in directed social networks (like Google+) because they are central to information dissemination to the users that are part of the them. Graphs with large SCCs are amenable to quick information dissemination processes.

We identified 9,771,696 SCCs in  $G$ . To reach this number we used a procedure involving two Depth First Searches [12]. Figure 4(c) presents the CCDF of the size of all SCCs found in  $G$ . In this figure we can see that almost all of them are small. In fact, there is only one with more than 100 nodes, which is the SCC with 25,240,000, that means that  $G$  has a giant component and the graph we collected is highly connected.

### 3.3.5 Degrees of Separation

The degree of separation essentially describes the shortest possible routes between two nodes of our graph. Although the degree of separation has been commonly thought of in the social context, the concept has many applications in social networking such as infor-

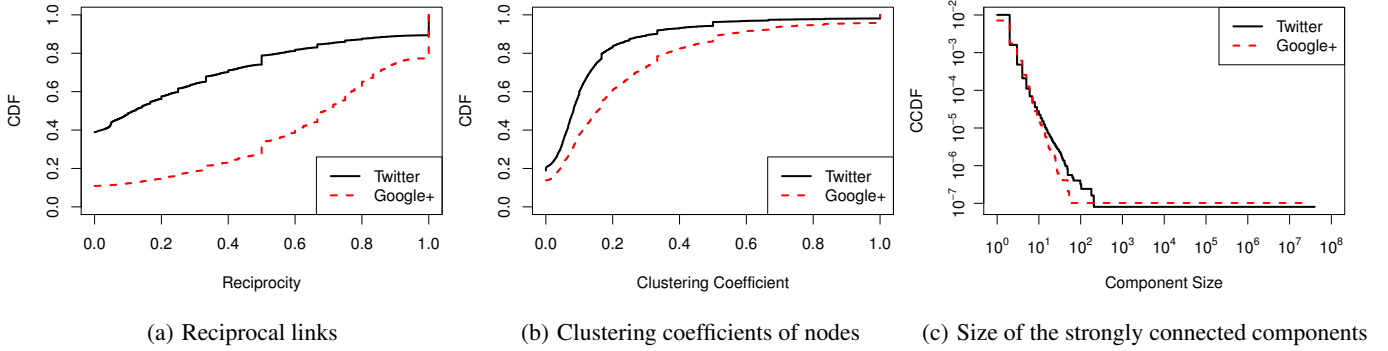


Figure 4: Distributions of various network properties in Google+

Table 4: Comparison of topological characteristics of Google+ and other online social networks

Network	Nodes	Edges	% Crawled	Path length	Reciprocity	Diameter	In-degree	Out-degree
Google+	35M	575M	56%	5.9	32%	19	16.4	16.4
Facebook	721M	62G	100%	4.7	100%	41	190.2	190.2
Twitter	41.7M	106M	100%	4.1	22%	18	28.19	29.34
Orkut	3M	223M	11%	4.3	100%	9	-	-

mation dissemination and friend recommendation [13]. We present an analysis of how many hops there are between two users in the Google+ social graph. In order to have the exact distribution we would need to compute the shortest path from all nodes to all nodes of the network. Due to the computational cost of this task we decided to use a random sampling procedure [1]. We sampled  $k$  different users and for each one of them we computed the shortest path to all others users in the network. We started with  $k = 2000$  and increased it until 10000, stopping in this value once there were no more changes in the distribution.

Figure 5 presents the final estimate of the path length distribution for two cases: the directed graph  $G$  and its undirected version. In the first case we can see that the most common value is 6 with an average of 5.9. In the second we have 5 as the most common and an average of 4.7. The graph  $G$  has a diameter of 19 and for its undirected version 13. This means that most users are only a few hops away from a random user, which has the important implica-

tion that information can spread quickly and widely throughout the network.

Although we have found the same mode of the well-known study of Milgram [31], it is important to remark that we are analyzing the public graph of Google+. So, adding back the edges omitted by users due to privacy constraints may reduce the average path length further. Comparing with other networks, we refer to Twitter with a mode of 4 and average of 4.12 [26], Facebook with an average of 4.74 hops [3] and a median of 6 for the MSN messenger network [27].

Table 4 summarizes the key structural features of Google+ and three others important OSNs in order to conclude this section. Statistics on other social networks are borrowed from [26, 3, 39, 32]. We can see some important differences, for example the Google+ social graph has a higher average path length and reciprocity. Its diameter is comparable to Twitter, but smaller than Facebook. Moreover, we can see that the number of friends (both in- and out-degrees) are much smaller when compared to Facebook.

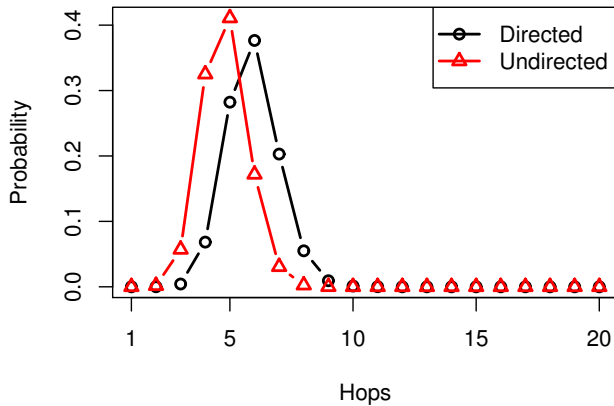


Figure 5: Estimated path length distribution

## 4. PATTERNS ACROSS GEO-LOCATIONS

Google+ users can list all the places they have lived at a field in their profile called “Places lived,” which is incorporated to the Google Map for visualization. Nearly 27% of the users in our dataset provide geo-location information. This feature is unique in Google+, for other social networks like Facebook only allows users to list their current location and, at most, their hometown information.

Using the places lived field, we analyzed how Google+ users are distributed around the world. For this, we first extracted the coordinates of the last location from the places lived field for each user and translated the coordinates into a valid country identifier. In this fashion, we were able to identify the country of 6,621,644 users.

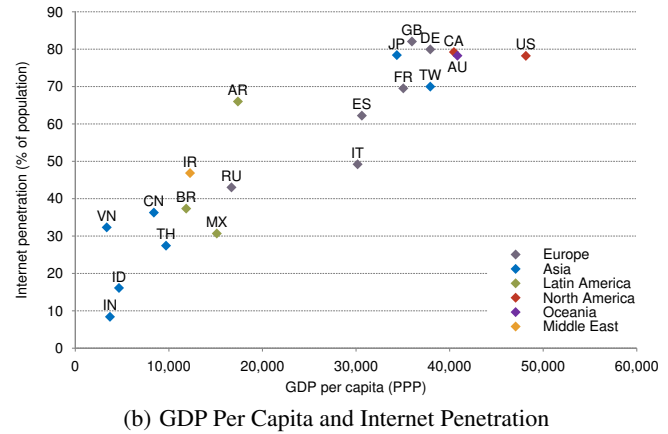
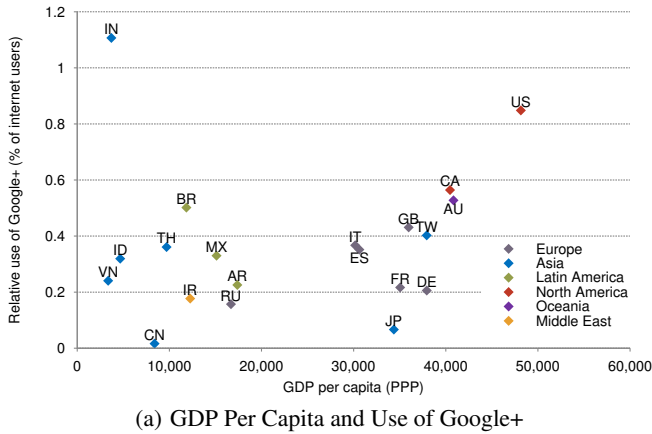


Figure 7: Usage of Internet and Google+ among countries

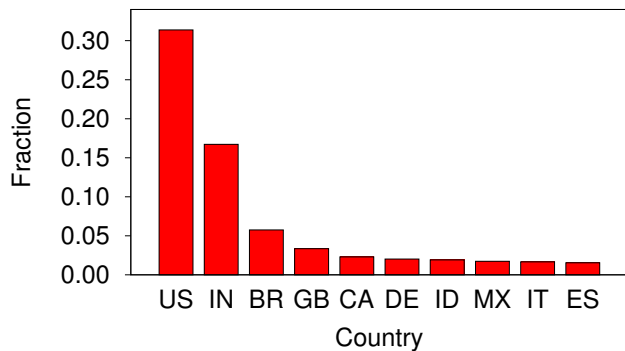


Figure 6: Top 10 countries with Google+ users

Figure 6 shows the top 10 countries in our dataset with their respective percentages of the registered users.<sup>1</sup> More than 30% of the users who share their location information are identified as living in the US. We observe Google+ is relatively popular in India and Brazil, which are also two of the countries with high presence of Google’s other social network, Orkut [19]. United Kingdom, Canada, and Germany also appear in the list, which are countries known to have high Internet Penetration Rate (IPR) or the percentage of Internet users out of the population of that country.<sup>2</sup> Interestingly, countries like Indonesia and Mexico appear in our top list, which are not part of the top countries based on the Internet penetration rate.

#### 4.1 Economics

Intrigued by the unique mix of countries based on Internet penetration rate, we further investigate which countries have high percentage of their Internet population on Google+. To do that we define *Google+ Penetration Rate*, that can be computed for each country  $C$  as follows:

$$GPR = \frac{\text{number of users in our dataset living in } C}{\text{Internet population of } C}. \quad (2)$$

<sup>1</sup>Country codes represent the following. US: United States; IN: India; BR: Brazil; GB: United Kingdom; CA: Canada; DE: Germany; ID: Indonesia; MX: Mexico; IT: Italy; ES: Spain.

<sup>2</sup>Statistics about the Internet penetration rate and population were obtained from <http://www.internetworldstats.com>

Note that our measure is meaningful only for the relative ranking of different countries, because our data is a sample taken from Google+ and in the sample only 27% of the users provide geo-location information.

Figure 7(a) shows the Google+ penetration rate for the top 20 countries. The top country in Google+ adoption now becomes India. We also see that countries like Taiwan and Thailand appear in the top ten list. For comparison, we show the Internet penetration rate of the same top 20 countries in Figure 7(b). The top five countries of Internet penetration are United Kingdom, Germany, Canada, Japan, and Australia. Both of the figures have Gross Domestic Product (GDP) per capita in the X-axis.

We make several observations. First, while there is a linear relationship between the GDP per capita of a country and its Internet penetration rate (identified as the data points clustered near a straight line in Figure 7(b)), we do not see the same trend in Google+ penetration rate. Countries with lower GDP per capita like Brazil, Mexico, and Thailand have equal footing in the penetration rate as with much wealthier countries such as United Kingdom, Australia, and Canada.

Second, certain countries showed a large gap between the Internet and Google+ penetration rate such as Japan, Russia, and China. In both of these countries, domestic social networks like Mixi in Japan, Odnoklassniki in Russia, and QQ in China are widely used. International social networking services like Facebook and Twitter are known to have little presence in these countries. Also in case of China, international social networking sites have been blocked [17].

Third, countries with lower Internet penetration like India and Brazil had a very high Google+ penetration rate. As we mentioned before, these two countries are known to favor Google’s other social network, Orkut. It is possible that users in India and Brazil are familiar with the Google product, hence are more likely to adopt the service quickly than other countries.

#### 4.2 User Occupation

Among various statistics we examined, the occupation-job title of the top users clearly distinguished Google+ from other well-known social networks, as we examined in Table 2. Interestingly, the top occupations also varied slightly across different countries. Table 5 shows the occupation-job title of the 10 most connected users in each of the top 10 countries, based on their in-degree (i.e., how many circles these users are added to by others). At a glance, the top list of Google+ is a mix of singers, bloggers, actors, and IT (i.e.,

**Table 5: Occupation-Job Title of the top users**

Country	Profession codes of the top-10 users*										Jaccard
United States	Co	Mu	IT	Mu	IT	Mu	Bu	IT	Mo	Ac	1.00
India	Mu	So	IT	Mu	Mo	Mo	IT	Bu	IT	Mu	0.57
Brazil	Co	TV	Jo	Wr	Ar	Bl	Bl	Co	Mu	Co	0.18
United Kingdom	Bu	Mu	IT	IT	Mu	Mu	IT	Mo	So	IT	0.57
Canada	IT	IT	Mu	Co	Bu	Ac	IT	Mu	Co	Ac	0.83
Germany	Bl	IT	IT	Jo	Bl	IT	Jo	Ec	Mu	Bl	0.22
Indonesia	Mu	IT	So	Mo	Mo	IT	Mu	Ec	Ph	Jo	0.30
Mexico	Mu	Mu	Mu	IT	Mu	Bl	Bl	Mu	Ac	Jo	0.33
Italy	Jo	Jo	IT	IT	Jo	IT	Jo	Mu	Mu	IT	0.29
Spain	Jo	Po	Po	IT	Mu	Mu	IT	Mu	Po	IT	0.25

\* Co: Comedian; Mu: Musician; IT: Information Technology Person; Bu: Businessman; Mo: Model; Ac: Actor; So: Socialite; TV: Television Host; Jo: Journalist; Bl: Blogger; Ec: Economist; Ar: Artist; Po: Politician; Ph: Photographer; Wr: Writer

Information Technology) professionals. When we compare the list to that of Twitter [26], the top list is particularly different in that we do not see any news media outlet like the New York Times and CNN, while we see founders of large Internet-based companies like Google and Facebook. In fact, seven out of the top 20 global users were IT related in Google+, which is uncommon in other social networks.

In the table, we also show the Jaccard index, used to compare the similarity and diversity of occupations in these country when compared to occupation-job titles in US. The top users in Canada have a very similar profile to that of the United States. Furthermore, the US, Canada, UK, and India share several top professions, which we may be due to the common British colonization. In contrast, Brazil, Italy, and Spain show a different set of celebrities and professions, and is worth noting that these three countries are Latin cultures, different from anglo-saxon cultures (US, CA, GB).

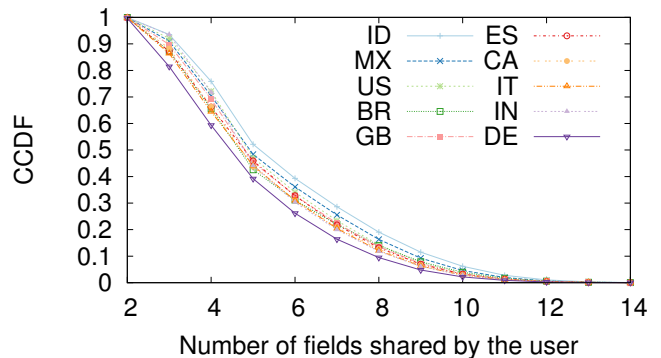
The top countries have very different kinds of popular users. IT professionals are popular in Google+. In Brazil, there are no famous IT related public figures, hence the list is dominated by comedians and bloggers. In Mexico, half of the top users are related to music. Italy is the country with more journalists among top users, 4 in total. Spain is the only country having Politicians in the top 10 user list. These lists suggest that each country has a different pattern of utilization of the information network provided by Google+, because the occupations of the top individuals represent what a typical user expect from Google+.

### 4.3 Openness

We next examine how these 10 countries differ in the notion of privacy, by looking at the number of different types of information publicly shared by users in their profiles (e.g., name, gender, education, occupation). As mentioned earlier, the name field is mandatory. Also, because of our methodology to utilize geo-location, all of the sample users studied in this section have shared “places lived” field. Therefore, the minimum number of shared fields is 2.

Figure 8 shows the CCDF of the number of fields users of the top 10 countries shares in their profiles. We present the X-axis in the range 2-14 for better visualization. We observe that, although the difference between the countries is not very pronounced, the ranking is slightly different.

Indonesia and Mexico share more information than other more popular countries like United States and United Kingdom. Germany is the most conservative when it comes to sharing personal information; it was the only country having less than 10% of the users sharing more than 12 fields and also the only country having less than 30% of the users sharing more than 10 fields.



**Figure 8: Number of fields in profiles in each country**

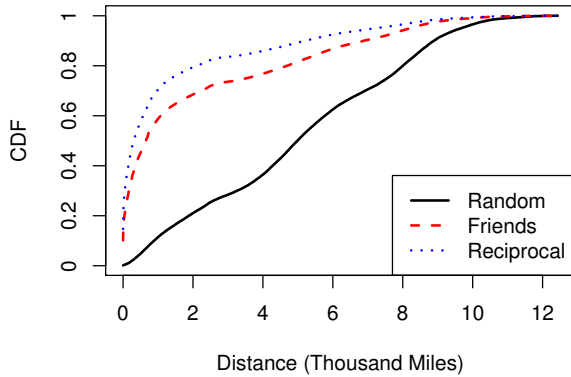
### 4.4 Average Path Miles

We now investigate the relationship between social network structure and geographical properties. We start by answering the following question: *is the geographical location of users an important factor in the formation of social links?* To understand if the distance has some influence on the formation of social links as described in circles, we estimated the physical distance of pairs of users in three cases: (1) every pair of socially connected users (approximately 60 million pairs), (2) pairs of reciprocally connected users (approximately 13 million pairs) and (3) randomly chosen pairs of users (20 million, not linked by a social relation). We then computed the physical distance between them. It is important to remark that we conducted this analysis only for users that share geo-location information, which represents 26.75% of the crawled Google+ network.

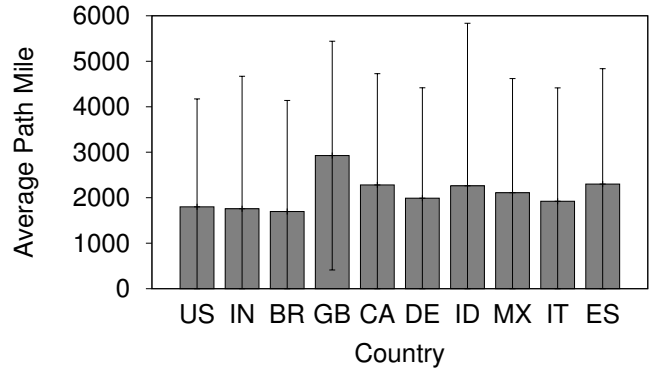
Figure 9(a) shows the cumulative distribution on the expected physical distance—which we call the *path mile*, similar to the notion of the path length—between pairs of circle friends and random user pairs in Google+. The friendship links in Google+ have higher geographical proximity than a random pairs of users. Nearly 58% of the users (friends) were separated by less than a thousand miles and 15% of them were separated by in fact 10 miles. This observation reinforces the high chance that the Google+ network largely capture the offline social relationships among users. As expected, users with symmetric links (reciprocal) live closer than those with asymmetrical links, indicating the influence of physical distance on the intensity of the relationship.

One natural question that arises from this result is whether it depends, or not, on the country size. For example, do geographically large countries like US have a higher average path mile than small





(a) Path Mile distribution of Google+ users



(b) Average path mile with standard deviation

**Figure 9: Path Mile distributions**

countries like Italy? If this is the case, large countries should have better investment in content distribution in order to minimize jitter and delay especially for the delivery of user generated videos. Figure 9(b) shows the average path miles along with the standard deviation error bar for the top 10 countries. Contrary to our expectation, there is no specific pattern relating the size of the country and its average path mile. One possible explanation could be that small countries have a considerable fraction of edges going outside the country. In fact, this result is discussed in next section.

#### 4.5 Social links across geography

The final question we ask is about the impact of country on friendship link formation. In particular, we ask: are users in the same country more likely to be friends in Google+ than users in different countries? To answer this question, we constructed a graph of countries, where each node is represented by one of the top 10 countries and the weight of each directed edge is given by the proportion of outgoing links from one country to another. Self-loop edges hence would represent the fraction of friendship links that bind two users in the same country.

Figure 10 shows the visualization of the links across the top 10 countries of Google+. The size of each node is normalized to represent the proportion of Google+ users in the associated country and the thickness of each edge is proportional to its weight. Edges with weight smaller than 0.01 were omitted to improve visibility. With this result we find that US has an important role in the overall landscape of Google+, as seen from the dominant influx of edges from most countries to the US. Moreover, the US is a node with low reciprocity, which means that there are a significant number of people of other places adding people in the US to their circles while those in the US, in general, prefer to form friendships among themselves.

Highly populated countries like Brazil, India, and Indonesia (and the US as already mentioned) tend to have a high weight in the self-loop edges. For the remaining six countries, the proportion of self-loops is much smaller. In particular, only 30% of the links are self-loops in United Kingdom and 33% in Canada. These two countries, as a result, have a large number of out-going edges to the US, which might be explained by geographical proximity and cultural similarity (e.g., sharing the same spoken language).

It is also worth noticing that in Figure 10, the countries that exhibit self-loop edges greater than 0.50 are those that do not have English as their first languages, which are Indonesia, India, Brazil, Italy. Perhaps because of its economical and technological leadership, the US also exhibits a high degree of self-loop edges. This indicates the language barrier in the set up of cross-national social relationships. Furthermore, this also means that the nature of language and geography will introduce interesting opportunity for growth strategies (e.g., launch of Google+ in a new non-English speaking country will likely show a similar organic growth pattern with many national links).

The average path mile discussed earlier could mean that content distribution in Google+ faces similar challenges for both small and large countries. In fact, smaller countries like United Kingdom may require more sophisticated measures to reduce delay in content delivery, as seen from its high average path mile. Furthermore, we see varying patterns of link formation across different countries. When it comes to building recommender systems, it may make sense to recommend domestic users and their content for those countries that have high degree of self-loop such as Brazil and India. However, it may be of more interest to the users to recommend foreign users and content to those in Germany and United Kingdom due to their low fraction of self-loops.

## 5. RELATED WORK

Characterization of social networks and user behavior is fundamental to the understanding and engineering of these services on the Internet. Many studies focus on the characterization of the most popular social network models, such as Facebook, Twitter, Orkut, Cyworld and others. Some of the important findings of these studies include establishing power law distributions for in- and out-degree, short average distance between pairs of users, a very large connected component, and a small number of extremely popular users. Thus, in the remainder of this section, we restrict our coverage of related work to studies that concentrate on characterization of other social network models.

Mislove et al. [32] studied graph theoretic properties of social networks, based on the friend network of Orkut, Flickr, LiveJournal, and YouTube. They confirmed the power-law, small-world, and scale-free properties of these social network services. Ahn et

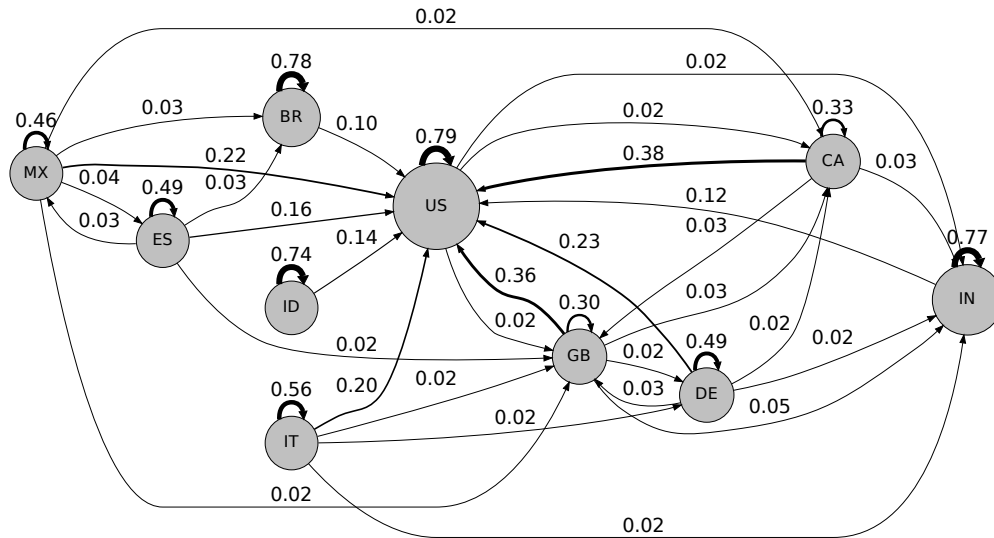


Figure 10: Link distribution across the top countries

al. [1] studied the network properties of Cyworld, a popular social networking service in South Korea. They compared the explicit friend relationship network with the implicit network created by messages exchanged on Cyworld’s guestbook. They found similarities in both networks: the in-degree and out-degree were close to each other and social interaction through the guestbook was highly reciprocal.

Recently, Ugander et al. [39, 3] used the complete Facebook dataset to study the social graph of Facebook. They show - among other things - that the degree of separation in that platform is 4.7, while we find that in Google+ it is 5.9. This difference may be explained by the fact that Google+ is a new platform at it should get denser in the future, as studied by [28] for different networks. Two recent references [26, 7] focus on the study of the Twitter graph. Other studies comparing different social network models were done by [32, 4]. In general, Google+ presents a combination of the characteristics of other networks, such as Facebook and Twitter.

When it comes to research on geo-location of users in online social networks, Liben-Nowell et al. [29] analyzed the geographical location of LiveJournal users and found a strong correlation between friendship and geographic proximity. This work confirms that most social links in the blog network are correlated with physical distance and only 33% of the friendships are independent of geography. We find a similar pattern in the friendship structure of Google+ in this paper. Recently, Scellato et al. [37] showed that there is a strong relationship between geographical distance and the probability of being friends in social networks. They discuss the implications of geo-location for social networking sites. Rodrigues et al. [36] investigate the word-of-mouth based content discovery by analyzing URLs in Twitter. They showed that propagation and physical proximity have correlation. Finally, Poblete et al. [34] studied a large amount of data gathered from Twitter and showed the various usages of the system depending across different countries.

Most existing work focus on several extremely popular social networks. Although it started much later in time, Google+ has grown rapidly as one of the top 10 popular social networks in the world. While there exist a large number of newspaper articles [24, 23, 40, 14] and blog articles [33, 6] on Google+, no research paper

has looked into its structure or user behavior in a research oriented way. This work provides a first look into the social graph and the geo-location patterns of Google+ users.

## 6. IMPLICATIONS

So far, we have made a series of observations about the service, network topology, and users of the Google+ social network based on large-scale data. In this section, we discuss the implications of these findings.

First, given that Google+ is a new social network, our first interest is to compare the topological structure of Google+ against other social networks. Compared to other social networks, Google+ cherishes openness in content sharing and is not a “walled garden” service like Facebook, where only the members can access content [30]. On the other hand, Google+ enforces a strong notion of friendship links by allowing users to manage different circles of friends. Our data analyses in Section 3 and Section 4 indicate that Google+ is in fact truly a social network, where the social links are correlated in geography reflecting offline friendship (i.e., friends are more likely to be located close), is far more reciprocal (i.e., bidirectional links), and have higher clustering coefficient (i.e., have triangle structures) compared to Twitter. The average path length is shown slightly longer than the other networks. As shown earlier, it is 5.9 in Google+ compared to 4.1–4.7 in other networks. A possible reason stem from the Google+ network is new and is still in the growing phase.

Second, observing the patterns of Google+ penetration worldwide can give insight into other new social networking service providers who would like to enter the market. While most new social network services typically starts their operation as a third party application or an adds-on service to the existing OSN services, Google+ is leading a full-fledged competition in the field. Therefore the pattern of how this new service is being adopted is important. While popular social networks like Facebook are known to have extremely high penetration rate of 50% or above [42], there is still room for a new social network service to become a hit in some countries. In particular, Google+ have been successfully adopted by countries with lower GDP per capita and this trend is important

because the Internet penetration rate of these countries are growing fast—meaning the user base could potentially grow more rapidly for Google+.

Third, our findings about the privacy concern of users indicate that users exhibit different privacy notions and expectations in Google+, based on geography. Such differences could be taken into account when trying to build a recommender system or run an advertisement campaign on top of Google+, for instance, the system could feature newly emerging musicians to users in Mexico, while recommend journalists to newly joining users in Italy. Also, running a political campaign on Google+ may not turn out successful for many countries, except for in Spain. Another example would be that marketers could build appealing profiles for companies by following the right level of privacy concerns in each country.

Third, based on the information about the circle list and the geography of users, we have examined how the social links are distributed across different countries. The resulting map in Figure 10 shows an interpreting landscape of user interactions. We find very different user behaviors in this case. Certain countries like Brazil, India, and Indonesia appear far more inward looking when forming social links, than those outward looking countries like United Kingdom and Canada. This means that based on the geographical location of where a user lives, her expectation towards finding a stronger local community in the network is different. We believe this kind of social network analysis allows us to study the collective and deviant behavior of particular demographics, which are increasingly considered important and useful both in research and practice.

## 7. CONCLUDING REMARKS

In this paper we study characteristics of the Google+ social graph. We present a comprehensive description of the platform, highlighting the main differences from other popular social network models. Our study is based on a large amount of data gathered encompassing 27 million user profiles and their connections to other users. With this dataset we analyze unique features of the Google+ demographics, especially on the gender, occupation, relationship status, and geo-location of users. We construct a graph representing the social relations of Google+ and analyze its structural properties, such as reciprocity, clustering coefficient, and node degree distribution. The Google+ social graph has a giant connected component that included 70% of the crawled users, which means that information can flow freely among all such users.

We also compute the average physical distance between two connected users. Exploiting the geo-location of users, we could see how aggressively Google+ has been adopted in different countries. We investigate relationships between economic indexes of countries and the adoption rate of Google+. We find that Google+ is popular in countries with relatively low Internet penetration rate. By examining the top users based on the circle link information, seven out of the top 20 users turned out to be in the information technology industry, a trend that is rather uncommon in other online social networks, where popular figures are media outlets, celebrities, and public figures. By looking into users who share their contact information publicly, we observe that a large fraction of the users are male and single.

There are several interesting directions for future research. First, we are interested in measuring the speed at which a new social network service grows and whether we can predict the phase transitions in the growth sparks (e.g., tipping point when a network suddenly shows a rapid growth or the point where the growth stabilizes and turns into a dormant phase). By collecting multiple snapshots of the Google+ topology, we hope to gain insight in

the dynamic changes in the internal structure of the social network over various adoption phases. Second, having seen the key differences of Google+ from other online social networks, we would like to understand how different privacy settings and openness impact the types of conversations and the patterns of content sharing in Google+. By gathering status updates of the most prolific users, we hope to understand what people expect to read and share from Google+ as opposed to other walled-garden or media-like social networks.

## Acknowledgements

We thank our shepherd Zhi-Li Zhang and the anonymous reviewers for their valuable comments. This research was partially funded by the Brazilian National Institute of Science and Technology for the Web (MCT/CNPq/INCT grant number 573871/2008-6). Gabriel Magno, Giovanni Comarela and Virgilio Almeida were partially supported by grants from CNPq, CAPES and FAPEMIG. Diego Saez-Trumper was supported by the HIPERGRAPH project (TIN2009-14560-C03-01) from the Spanish Economy and Competitiveness Ministry. Meeyoung Cha was supported by Basic Science Research Program through the National Research Foundation, Korea (2011-0012988).

## 8. REFERENCES

- [1] Y.-Y. Ahn, S. Han, H. Kwak, S. Moon, and H. Jeong. Analysis of topological characteristics of huge online social networking services. In *Proc. of ACM International World Wide Web Conference*, pages 835–844, 2007.
- [2] P. Allen. Google+ growth accelerating, passes 62 million users, adding 625,000 new users per day. prediction: 400 million users by end of 2012., December 2011. [http://bit.ly/gp\\_nusers](http://bit.ly/gp_nusers).
- [3] L. Backstrom, P. Boldi, M. Rosa, J. Ugander, and S. Vigna. Four degrees of separation. *CoRR*, abs/1111.4570, 2011.
- [4] F. Benevenuto, T. Rodrigues, M. Cha, and V. Almeida. Characterizing user behavior in online social networks. In *Proc. of ACM SIGCOMM Internet Measurement Conference*, pages 49–62, 2009.
- [5] S. Brin. Web freedom faces greatest threat ever, warns google's sergey brin., April 2012. <http://www.guardian.co.uk/technology/2012/apr/15/web-freedom-threat-google-brin>.
- [6] H. Carpenter. I already use facebook and twitter. why do i need google+?, July 2011. <http://tinyurl.com/7pwauo2>.
- [7] M. Cha, H. Haddadi, F. Benevenuto, and K. P. Gummadi. Measuring User Influence in Twitter: The Million Follower Fallacy. In *Proc. of AAAI International Conference on Weblogs and Social Media*, 2010.
- [8] M. Cha, A. Mislove, and K. P. Gummadi. A measurement-driven analysis of information propagation in the flickr social network. In *Proc. of ACM International World Wide Web Conference*, pages 721–730, 2009.
- [9] E. Comission. Attitudes on data protection and electronic identity in the european union, June 2011. <http://bit.ly/wkZhGR>.
- [10] comScore. Google+ Reaches 20 Million Visitors in 21 Days, July 2011. <http://tinyurl.com/3ox3lp9>.
- [11] comScore. It's a Social World: Top 10 Need-to-Knows About Social Networking and Where It's Headed, December 2011. <http://tinyurl.com/cfhqxke>.

- [12] T. H. Cormen, C. Stein, R. L. Rivest, and C. E. Leiserson. *Introduction to Algorithms*. McGraw-Hill Higher Education, 2nd edition, 2001.
- [13] D. Easley and J. Kleinberg. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, 2010.
- [14] E. Eldon. Larry page is super excited to announce that google+ has 90m users, January 2012. <http://tinyurl.com/82mkntq>.
- [15] engadget. Google+ has 250 million users, more mobile than desktop, June 2012. <http://tinyurl.com/d3ausse>.
- [16] J. Fogel and E. Nehmad. Internet social network communities: Risk taking, trust, and privacy concerns. *Computers in Human Behavior*, 25(1):153–160, 2009.
- [17] Freshtrax. Top 10 most popular websites in china, March 2010. <http://bit.ly/xgRIAq>.
- [18] M. Gjoka, M. Kurant, C. T. Butts, and A. Markopoulou. A walk in facebook: Uniform sampling of users in online social networks. *CoRR*, abs/0906.0060, 2009.
- [19] Google. Orkut demographic data, January 2012. <http://www.orkut.com/MembersAll>.
- [20] R. Gross and A. Acquisti. Information revelation and privacy in online social networks. In *Proc. of ACM CCS Workshop on Privacy in the Electronic Society*, pages 71–80, 2005.
- [21] V. Gundotra. Google+: 92, 93, 94, 95, 96, 97, 98, 99... 100., September 2011. [http://bit.ly/gp\\_noinvite](http://bit.ly/gp_noinvite).
- [22] V. Gundotra. Introducing the google+ project: Real-life sharing, rethought for the web, June 2011. [http://bit.ly/gp\\_released](http://bit.ly/gp_released).
- [23] J. Halliday. Google+ saw 67m visits in november, with 2.7m from uk, says comscore, January 2012. <http://tinyurl.com/7ydcnck>.
- [24] Q. Hardy. Google+ gains traction, researcher says, December 2011. <http://nyti.ms/wrEsOx>.
- [25] R. Kumar, J. Novak, and A. Tomkins. Structure and evolution of online social networks. In *Proc. of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 611–617, 2006.
- [26] H. Kwak, C. Lee, H. Park, and S. Moon. What is twitter, a social network or a news media? In *Proc. of ACM International World Wide Web Conference*, pages 591–600, 2010.
- [27] J. Leskovec and E. Horvitz. Planetary-scale views on a large instant-messaging network. In *Proc. of ACM International World Wide Web Conference*, pages 915–924, 2008.
- [28] J. Leskovec, J. Kleinberg, and C. Faloutsos. Graphs over time: densification laws, shrinking diameters and possible explanations. In *Proc. of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 177–187, 2005.
- [29] D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, and A. Tomkins. Geographic routing in social networks. *Proceedings of the National Academy of Sciences of the United States of America*, 102(33):11623–11628, 2005.
- [30] R. MacManus. Tim berners-lee calls facebook a walled garden - is that fair?, November 2010. [http://www.readwriteweb.com/archives/tim\\_berniers-lee\\_says\\_facebook\\_is\\_a\\_walled\\_garden.php](http://www.readwriteweb.com/archives/tim_berniers-lee_says_facebook_is_a_walled_garden.php).
- [31] S. Milgram. The small world problem. *Psychology Today*, 2:60–67, 1967.
- [32] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee. Measurement and Analysis of Online Social Networks. In *Proc. of ACM SIGCOMM Internet Measurement Conference*, 2007.
- [33] J. Mitchell. Google+ is going to mess up the internet, January 2012. <http://tinyurl.com/84x7o3a>.
- [34] B. Poblete, R. Garcia, M. Mendoza, and A. Jaimes. Do All Birds Tweet the Same? Characterizing Twitter Around the World. In *Proc. of ACM Conference on Information and Knowledge Management*, Glasgow, UK, 2011.
- [35] B. F. Ribeiro and D. F. Towsley. Estimating and sampling graphs with multidimensional random walks. *CoRR*, abs/1002.1751, 2010.
- [36] T. Rodrigues, F. Benevenuto, M. Cha, K. P. Gummadi, and V. Almeida. On Word-of-Mouth Web Based Discovery of the Web. In *Proc. of ACM SIGCOMM Internet Measurement Conference*, 2011.
- [37] S. Scellato, C. Mascolo, M. Musolesi, and V. Latora. Distance matters: geo-social metrics for online social networks. In *Proc. of ACM SIGCOMM Workshop on Social Networks*, Berkeley, CA, USA, 2010. USENIX Association.
- [38] T. Taraszow, E. Aristodemou, G. Shitta, Y. Laouris, and A. Arsoy. Disclosure of personal and contact information by young people in social networking sites: An analysis using Facebook profiles as an example. *International Journal of Media and Cultural Politics*, pages 81–101, 2010.
- [39] J. Ugander, B. Karrer, L. Backstrom, and C. Marlow. The anatomy of the facebook social graph. *CoRR*, abs/1111.4503, 2011.
- [40] T. Wasserman. Google+ most popular with male users, students, February 2012. <http://tinyurl.com/7kz2k3m>.
- [41] D. J. Watts and S. H. Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684):440–442, June 1998.
- [42] ZDNet. Top 10 countries in Facebook adoption, January 2012. <http://zd.net/xmA4rt>.