

exchange or simply relay information. To quantify the level of balance in the connectivity of individual nodes, Figure 9(a) plots the summary distribution of the ratio of followers to friends (using boxplots) for different group of users based on their number of followers in our most recent snapshot (LCC-Nov). This figure demonstrates that only low degree nodes (with less than 100 followers) exhibit some balance between their number of followers and friends. Otherwise, the number of friends among G+ users grows much slower than the number of followers.

We calculate the percentage of bidirectional relationships for a node u , called $BR(u)$, as expressed in Equation 1 where $Friend(u)$ and $Follower(u)$ represent the set of friends and followers for u , respectively. In essence, $BR(u)$ is simply the ratio of the total number of bidirectional relationships over the total number of unique relationships for user u .

$$BR(u) = \frac{Friend(u) \cap Follower(u)}{Friend(u) \cup Follower(u)} \quad (1)$$

Figure 9(b) presents the summary distribution of $BR(u)$ for different groups of G+ users in LCC the based on their number of followers using the LCC-Nov snapshot. The results for other recent LCC snapshots are very similar. As expected, popular users ($> 10k$ followers) have a very small percentage of bidirectional relationships. As the number of followers decreases, the fraction of bidirectional relationships slowly increases until it reaches around 40% for low-degree users ($< 1K$ followers). In short, even low degree users that maintain a balanced connectivity, do not reciprocate more than 40% of their relationships. Our inspection of 5% of LCC users who reciprocate more than 90% of their edges revealed that 90% of them maintain less than 3 friends/followers and less than 5% of them have any public posts. These results collectively suggest that G+ users reciprocate a small fraction of their relationships which is often done by very low degree users with no activity.

Clustering Coefficient: Figure 10 depicts the summary distribution of the undirected version of the clustering coefficient (CC) among G+ users in different LCC snapshots. This figure clearly illustrates that during the roughly one year period from Dec 2011 to Nov 2012, the CC among the bottom 90% of users remained below 0.5 and continuously decreases. On the other hand, the CC for the top 10% of users has been very stable. In essence, *the G+ structure has become less clustered as new users joined the LCC over the one year period.* A similar trend in cluster coefficient has been recently reported for a popular Chinese OSN [49] that indicates that such an evolution in the CC might be driven by underlying social forces rather than features of the OSNs. We also notice that the distribution of the CC among G+ users exhibits only minor changes between Aug and Nov 2012 which is another sign of stability in the connectivity features of G+ network. Compared to Twitter network where the CC is less than 0.3 for 90% of users, G+ is still more clustered. Furthermore, using the approximation presented in [39], we conclude that just 1% of the nodes in a complete Facebook snapshot collected in May 2011 [46] have a CC larger than 0.2 in comparison with the 16% and 30% in Twitter and G+ (using LCC-Nov snapshot). In summary, as the population of G+ has grown, its connectivity has become less clustered but it is still the most clustered network compared to Twitter and Facebook.

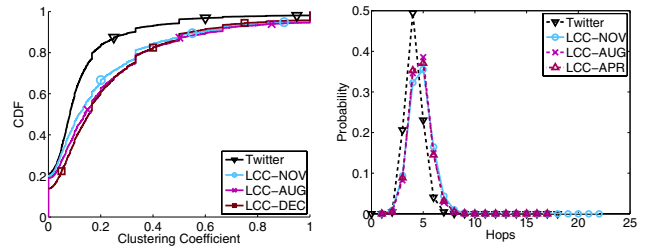


Figure 10: Clustering Coefficient

Figure 11: Average Path Length

	LCC-Nov	FB	Twitter
Path Length (Avg)	4.7	4.7	4.1
Path Length (Mode)	5	5	4
Eff. Diameter	6	-	4.8
Diameter	22	41	18

Table 6: Summary of path length and diameter characteristics for G+, Facebook and Twitter

Path Length: Figure 11 plots the probability distribution function for the pairwise path length between nodes in different LCC snapshots for G+ and a snapshot of Twitter (TW-Con). We observe that roughly 97-99% of the pairwise paths between G+ users are between 2 to 7 hops long and roughly 68-74% of them are 4 or 5 hops. The diameter of the G+ graph has increased from 17 hops (in April) to 22 hops (in November of 2012). The two visibly detectable changes in this feature of the G+ graph as a result of its growth are: a small decrease in typical path length (from April to November) and the increase of its diameter in the same period. Table 6 summarizes the average and mode path length, the diameter and the efficient diameter [38] (*i.e.*, 90 percentile of pairwise path lengths) for the G+ network (using LCC-Nov), Twitter (using TW-Con) and a Facebook snapshot from [23]. We observe that G+ and Facebook exhibit similar average (and mode) path length but Facebook has a longer diameter. One explanation is the fact that the size of Facebook network is roughly one order of magnitude larger than G+ LCC. Twitter has the shortest average and mode path length and diameter among the three. We conjecture that this difference is due to the lack of restriction in the maximum number of friends that leads to many shortcuts in the network as Twitter users connect to a larger number of friends.

Relating User Activity & Connectivity: We also analyzed the correlation between the connectivity and activity of individual users in the LCC. Our results reveal a strong positive correlation between the popularity of a user (*i.e.*, number of followers) and the user's post rate. The post rate of individual users exhibit a weaker correlation with the number of friends. Further details on all of our analysis can be found in our related technical report [31].

In summary, our analysis on the evolution of LCC connectivity led to the following key findings: (i) As the size of LCC significantly increased over the past year, all connectivity features of LCC have initially evolved but have become rather stable in recent months despite its continued growth. (ii) Only low degree and non-active users may reciprocate a moderate fraction of their relationships. (iii) Many key features of connectivity for G+ network (e.g., degree distribution, fraction of bidirectional relationships) have striking similarity with the Twitter network and very different from

the Facebook network. These connectivity features collectively suggest that G+ is primarily used for message propagation similar to Twitter rather than pairwise users interactions similar to Facebook.

7. RELATED WORK

OSN characterization: The importance of OSNs has motivated researchers to characterize different aspects of the most popular OSNs. The graph properties of Facebook [46, 23], Twitter [37, 26] and other popular OSNs [42] have been carefully analyzed. Note that all these studies use a single snapshot of the system to conduct their analysis, instead we analyze the evolution of the G+ graph over a period of one year. In addition, some other works leverage passive (e.g., click streams) [24, 45] or active [48, 32] measurements to analyze the user activity in different popular OSNs. These papers are of different nature than ours since they use smaller datasets to analyze the behaviour of individual users. Instead, we use a much larger dataset to analyze the evolution of the aggregate public activity along time as well as the skewness of the contribution overall activity across users in G+. Finally, few works have also analyzed users' information sharing through their public attributes in OSNs such as Facebook [41].

Evolution of OSN properties: Previous studies have separately studied the evolution of the relative size of the network elements for specific OSNs (Flickr and Yahoo 360) [36], the growth of an OSN and the evolution of its graph properties [40, 22, 49, 28, 29, 43] or the evolution of the interactions between users [34] and users' availability [25]. In this paper, instead of looking at a specific aspect, we perform a comprehensive analysis to study the evolution of different key aspects of G+ namely, the system growth, the representative of the different network elements, the LCC connectivity and activity properties and the level of information sharing.

Google+ Characterization: G+ has recently attracted the attention of the research community. Mango et al. [39] use a BFS-based crawler to retrieve a snapshot of the G+ LCC between Nov and Dec 2011. They analyze the graph properties, the public information shared by users and the geographical characteristics and geolocation patterns of G+. Schiberg et al. [44] leverage Google's site-maps to gather G+ user IDs and then crawl these users' information. In particular, they study the growth of the system and users connectivity over a period of one and a half months between Sep and Oct 2011. Unfortunately, as acknowledged by the authors the described technique was anymore available after Oct 2011. Furthermore, the authors also analyze the level of public information sharing and the geographical properties of users and links in the system. Finally, Gong et al. [30] use a BFS-based crawler to obtain several snapshots of the G+ LCC in its first 100 days of existence. Using this dataset the authors study the evolution of the main graph properties of G+ LCC in its early stage. Our work presents a broader focus than these previous works since in addition to the graph topology and the information sharing we also analyze (for first time) the evolution of both the public activity and the representativeness of the different network elements. Furthermore, our study of the graph topology evolution considers a 1 year window between Dec 2011 and Nov 2012 when the network is significantly larger and presents important differences to its early status that is the focus of the previous works. In another interesting, but less related

work, Kairam et al. [35] use the complete information for more than 60K G+ users (provided by G+ administrators) and a survey including answers from 300 users to understand the selective sharing in G+. Their results show that public activity represents 1/3 of the G+ activity and that an important fraction of users make public posts frequently. Finally, other papers have studied the video telephony system of G+ [47], the public circles feature [27] and the collaborative privacy management approaches [33].

8. CONCLUSION

This paper examines the key features of G+ network and their evolution during the first year of G+ operation. We conduct large scale measurement on G+ and collect some of the largest public datasets on any OSN to date to characterize connectivity, activity and information sharing across G+ users along with their evolution over a one year period. We develop an efficient technique to collect random samples of G+ user. This in turn enables us to determine the relative size of key components (*i.e.*LCC, partitions, singletons) of G+ network.

We show that while the size of LCC component of G+ has grown at a high rate (200K user per day), the relative size of LCC has decreased with time. Our investigations reveal that a significant fraction of new G+ users appear to be implicitly added by Google while they register for other Google services. Furthermore, the main connectivity features of LCC have become relatively stable in recent months which suggests that the G+ network has reached a steady state. We show that these stable connectivity features of LCC component of G+ have a striking similarity with Twitter but are very different from Facebook. This similarity indicates that users use G+ for message propagation similar to Twitter rather than pairwise user interaction like Facebook. In terms of user activity, even LCC users are not actively engaged in G+ network. The contribution of user activity in terms of posting is skewed among LCC users (*i.e.*10% of users are responsible for 80% of posts) and user reactions to activities is an order of magnitude more skewed (*i.e.*1% of users generate 80% of reactions to all posts). Our findings collectively demonstrate that in the current OSN marketplace with two dominant players, namely Facebook and Twitter, a new OSN such as G+ might be able to attract a rather significant number of users to become part of the network (*i.e.*connect to its LCC). However, it is much more challenging to get these users meaningfully engaged in the system.

9. ACKNOWLEDGEMENTS

The authors would like to thank anonymous reviewers for their valuable feedback as well as Meeyoung Cha, Sue Moon, Diego Saez and Haewoon Kwak for sharing information from their datasets with us. This work has been partially supported by the European Union through the FP7 TREND (257740) and eCOUSIN (318398) Projects and the TWIRL (ITEA2-Call 5-10029) Project, the Spanish Government through the MINECO eeCONTENT Project (TEC2011-29688-C02-02) and the MECD Jose Castillejo Grant (JC2011-0353), the Regional Government of Madrid through the MEDIANET project (S-2009/TIC-1468), the Social Networks Chair of Institut-Mines Telecom SudParis and the National Science Foundation under Grant IIS-0917381.

10. REFERENCES

- [1] <http://google-plus.com/category/statistics/>.
- [2] http://online.wsj.com/article/SB10001424052970204653604577249341403742390.html?mod=WSJ_hp_LEFTTopStories.
- [3] <http://en.wikipedia.org/wiki/Google+>.
- [4] <https://plus.google.com/107117483540235115863/posts/YUniwagZuKZ>.
- [5] *Announcement of New Feature for Google+*. <http://www.workinghomeguide.com/9918/google-new-features-hashtag-auto-complete-text-to-photos-and-video-status>.
- [6] *Announcement of New Feature to add text to photos in Google+*. <https://plus.google.com/107814003053721091970/posts/D7gfxe4bU7o>.
- [7] *Buzz shut down announcement*. <https://support.google.com/mail/bin/answer.py?hl=en&answer=1698228>.
- [8] *Flow Chart for Google+ Sharing*. <http://googleplushowto.com/2011/07/will-user-a-see-my-post-in-google-plus/>.
- [9] *Genealogy Data: Frequently Occurring Surnames from Census 2000*. US Census Bureau. <http://www.census.gov/genealogy/www/data/2000surnames/index.html>.
- [10] *Google+ (maximum number of contacts in your circles)*. <http://support.google.com/plus/bin/answer.py?hl=en&answer=1733011>.
- [11] *Google+ official learn more site*. <http://www.google.com/intl/en+/learnmore/>.
- [12] *Google+ official support site*. <http://support.google.com/plus/>.
- [13] *Google+ Pages*. <http://www.google.com/+business/>.
- [14] *Google+ Pages announcement*. <http://googleblog.blogspot.com/2011/11/google-pages-connect-with-all-things.html>.
- [15] *Google Privacy Policy*. <http://www.google.com/policies/privacy/>.
- [16] *Google+ Redesign*. <http://mashable.com/2012/04/11/google-plus-redesign/>.
- [17] *Google Registration*. *Arstechnica*. <http://arstechnica.com/gadgets/2012/01/google-doubles-plus-membership-with-brute-force-signup-process/>.
- [18] *Google Registration*. *Blogspot*. <http://googlesystem.blogspot.com/2012/01/new-google-accounts-require-gmail-and.html>.
- [19] *Google wave shut down announcement*. <http://support.google.com/bin/answer.py?hl=en&answer=1083134>.
- [20] *Orkut Official Site*. <http://www.orkut.com>.
- [21] *Orkut Statistics*. *Wikipedia*. <http://en.wikipedia.org/wiki/Orkut>.
- [22] Y.Y. Ahn, S. Han, H. Kwak, S. Moon, and H. Jeong. Analysis of Topological Characteristics of Huge Online Social Networking Services. In *WWW*, 2007.
- [23] Lars Backstrom, Paolo Boldi, Marco Rosa, Johan Ugander, and Sebastiano Vigna. Four Degrees of Separation. *CoRR*, abs/1111.4570, 2011.
- [24] F. Benevenuto, T. Rodrigues, M. Cha, and V. Almeida. Characterizing User Behavior in Online Social Networks. In *ACM IMC*, 2009.
- [25] A. Boutet, A.M. Kermarrec, E. Le Merrer, and A. Van Kempen. On the Impact of Users Availability in OSNs. In *ACM SNS*, 2012.
- [26] M. Cha, H. Haddadi, F. Benevenuto, and K.P. Gummadi. Measuring User Influence in Twitter: The Million Follower Fallacy. In *AAAI ICWSM*, 2010.
- [27] L. Fang, A. Fabrikant, and K. LeFevre. Look Who I Found: Understanding the Effects of Sharing Curated Friend Groups. In *ACM WebSci*, 2012.
- [28] S. Gaito, M. Zignani, G.P. Rossi, A. Sala, X. Wang, H. Zheng, and B.Y. Zhao. On the Bursty Evolution of Online Social Networks. In *ACM KDD HotSocial Workshop*, 2012.
- [29] S. Garg, T. Gupta, N. Carlsson, and A. Mahanti. Evolution of an Online Social Aggregation Network: an Empirical Study. In *ACM IMC*, 2009.
- [30] N.Z. Gong, W. Xu, L. Huang, P. Mittal, E. Stefanov, V. Sekar, and D. Song. Evolution of Attribute-augmented Social Networks: Measurements, Modeling, and Implications Using Google+. In *ACM IMC*, 2012.
- [31] R. Gonzalez, R. Cuevas, R. Motamedi, R. Rejaie, and A. Cuevas. Google+ or Google-?. Dissecting the Evolution of the New OSN in its First Year. Technical report available at: <http://www.it.uc3m.es/~rcuevas/techreports/g+TR2012.pdf>, Universidad Carlos III de Madrid, 2012.
- [32] L. Gyarmati and T.A. Trinh. Measuring User Behavior in Online Social Networks. *Network*, *IEEE*, 24(5):26–31, 2010.
- [33] H. Hu, G.J. Ahn, and J. Jorgensen. Enabling Collaborative Data Sharing in Google+. In *IEEE Globecom*, 2012.
- [34] Jing Jiang, Christo Wilson, Xiao Wang, Peng Huang, Wenpeng Sha, Yafei Dai, and Ben Y. Zhao. Understanding Latent Interactions in Online Social Networks. In *ACM IMC*, 2010.
- [35] S. Kairam, M. Brzozowski, D. Huffaker, and E. Chi. Talking in Circles: Selective Sharing in Google+. In *ACM CHI*, 2012.
- [36] R. Kumar, J. Novak, and A. Andtomkins. Structure and Evolution of Online Social Networks. In *ACM KDD*, 2006.
- [37] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. What is Twitter, a Social Network or a News Media? In *WWW*, 2010.
- [38] J. Leskovec, J. Kleinberg, and C. Faloutsos. Graphs Over Time: Densification Laws, Shrinking Diameters and Possible Explanations. In *ACM SIGKDD*, 2005.
- [39] Gabriel Magno, Giovanni Comarella, Diego Saez-Trumper, Meeyoung Cha, and Virgilio Almeida. New kid on the block: Exploring the google+ social graph. In *ACM IMC*, 2012.
- [40] A. Mislove, H.S. Koppula, K.P. Gummadi, P. Druschel, and B. Bhattacharjee. Growth of the Flickr Social Network. In *WOSN*, 2008.
- [41] A. Mislove, B. Viswanath, K.P. Gummadi, and P. Druschel. You Are Who You Know: Inferring User Profiles in Online Social Networks. In *ACM WSDM*, 2010.
- [42] Alan Mislove, Massimiliano Marcon, Krishna P. Gummadi, Peter Druschel, and Bobby Bhattacharjee. Measurement and Analysis of Online Social Networks. In *ACM IMC*, 2007.
- [43] R. Rejaie, M. Torkjazi, M. Valafar, and W. Willinger. Sizing Up Online Social Networks. *IEEE Network*, 24(5):32–37, Sept-Oct 2010.
- [44] D. Schiöberg, F. Schneider, H. Schiöberg, S. Schmid, S. Uhlig, and A. Feldmann. Tracing the Birth of an OSN: Social Graph and Profile Analysis in Google. In *ACM WebSci*, 2012.
- [45] F. Schneider, A. Feldmann, B. Krishnamurthy, and W. Willinger. Understanding Online Social Network Usage from a Network Perspective. In *ACM IMC*, 2009.
- [46] Johan Ugander, Brian Karrer, Lars Backstrom, and Cameron Marlow. The Anatomy of the Facebook Social Graph. *CoRR*, abs/1111.4503, 2011.
- [47] Y. Xu, C. Yu, J. Li, and Y. Liu. Video Telephony for End-consumers: Measurement Study of Google+, iChat, and Skype. In *ACM IMC*, 2012.
- [48] Z. Xu, Y. Zhang, Y. Wu, and Q. Yang. Modeling User Posting Behavior on Social Media. In *ACM SIGIR*, 2012.
- [49] X. Zhao, A. Sala, C. Wilson, X. Wang, S. Gaito, H. Zheng, and B.Y. Zhao. Multi-scale Dynamics in a Massive Online Social Network. In *ACM IMC*, 2012.