Are Circles Communities? A Comparative Analysis of Selective Sharing in Google+

Steffen Brauer, Thomas C. Schmidt
Internet Technologies Research Group, Dep. Informatik
Hamburg University of Applied Science, Hamburg, Germany
steffen.brauer@haw-hamburg.de, t.schmidt@ieee.org

***Abstract*—The audience of shared content in Social Media is often hard to determine. To protect users from over-sharing, several services provide a feature for grouping contacts. Communities, interest groups, and circles are common examples. In this work, we investigate the structural properties of the circles in Google+ in comparison to the well-known communities. Based on several data sets and scoring functions, we search for the specific characteristics of circles. Our findings indicate that circles indeed form a special substructure that clearly differs from community groups. While the internal connectivity of circles and communities appear fairly similar, circles admit a much enhanced intensity of external relations. Circles resemble communities to which a large number of external links have been added. Selective sharing in circles is thus less confined.**

Keywords—Social media analysis; social graph structure; group formation; privacy;

I. INTRODUCTION

In Social Media applications like Blogs, Online Social Networks (OSNs), or Micro Feeds, users can communicate and share content in groups. While the audience of direct communication and private messages is explicitly known, content sharing and status messaging are commonly distributed via implicit replication on the platform. A possible large audience of a post is often not considered by the user, but may have significant impact on his personal life. The so called over-sharing of content can be avoided using features for selective sharing [1], which are provided by several Social Media services. The concept of selective sharing was designed to support a context-dependent publication behaviour. In different contexts like work, family, or friends, a person can act differently according to appropriate norms and accepted conventions. These so called facets [2] or foci [3] describe different social aspects in the life of a person and provide the theoretical background of selective sharing features.

Released only in 2011, Google+ is a latecomer in the world of Online Social Networks. One of its characteristic features are the circles. Circles allow users to group their contacts in different containers provided by the platform, and manage their different social facets. Preconfigured circles are friends, family and acquaintances, as well as all following users. It is also possible to rename the default circles or add customized ones. The range of users, who can read a post, is limited to members of the target circle that was selected when the post was created. Besides selective sharing, the Google+ circles also enable selective reception of content streams generated by the contacts in each circle. While the circles are private to the user per default, there is the option to share circles. Another interesting feature of Google+ is directed edges, which are created by adding users to circles. Outgoing edges are named 'In your circles' and incoming edges 'In circles of others'. These directed links in Google+ make it a hybrid network between typical OSNs like Facebook and Micro-blogging services like Twitter [4].

We are interested in understanding the group structures within OSNs [5], [6]. In this work, we investigate the characteristics of circles within the Google+ graph and compare to well-known group structures like communities. We want to find out how particular these circles are, and try to identify a clear signature of their structure if available.

A circle is actively created and shaped by its owner, possibly based on a personal perspective. In contrast, communities or groups follow a subscribe model and are open for joining by different users who share a common attribute. Given such difference in composition, it is natural to question the similarity of the results. Using a variety of empirical data sets, we explore the following details:

1) Are circles pronounced structures in the social graph, well distinguished from the underlying network?
2) Resemble circles communities or other group structures of classical social networks, or do circles show distinct characteristics?

To answer these question, we analyze the data sets using different scoring functions. We base our results on four scoring functions from the field of community detection that focus on specific characteristics of network communities. Our findings indicate that circles do admit a very pronounced structure that differs from other groups by its connectivity with the remaining social graph. While traditional communities are rather closed groups with few relations to the outside, circles in Google+ are densely connected internally *and* externally. This makes them appear like communities with many additional transit links.

The remainder of this paper is structured as follows. We continue with reviewing related work on Google+ and its circle feature. In Section III, we discuss the theoretical differences between circles and traditional communities. The following Section IV characterizes the Google+ and LiveJournal data sets that we use in our evaluation. In Section V, we evaluate and discuss the questions stated above. We conclude in Section VI and give an outlook.

II. Related Work

Since its release, Google+ has attracted research. Schiöberg et al. [4] traced the social graph in the beta and initial public phase. They found that the asymmetric relations in Google+ lead to a hybrid form composed of classic social network like Facebook and of Social Media services like Twitter.

Gonzales et al. [7] likewise performed several crawls of the Google+ graph in a one year period. Their extensive analysis captures the connectivity of the network, user activity and information sharing. The authors conclude that Google+ is a broadcast social media system with a small group of very active users. These users create the largest part of the visible activity and attract other users to join. Our work backs these first-hand observations by a quantitative structural analysis.

Magno et al. [8] also crawled Google+ in its creation phase. While the authors evaluated metrics typical for social network analysis, they also included reciprocity. This measure describes describes the correlation between in- and out-degrees for a given vertex or set of vertices. Besides the network structure, the authors focus on demographic properties and investigate the geo-location of users.

Aside from the typical social network features, circles are an interesting field for analyzing social behaviour. Kairam et al. [1] focus on the selective sharing aspect and how it is used in Google+. They apply the theory of facets of a person's life. Each facet covers a specific group of personal contacts from daily life. While some people share the same information with all people they know, some people want to share targeted information just with some people. Kairam et al. show in their work, that circles on Goolge+ correspond to these natural groupings.

McCauley and Leskovec [9] propose a model for automatically discovering circles in a given ego-network. The ego-network of a user covers all vertices he is connected to and all edges between these vertices. Based on the ego-network of a user, they formulate the circle detection as a clustering problem that is applied to a user's ego-network. The authors aim at modelling properties with respect to the assumed characteristics of circles: i) vertices in a circle share a common property or aspect, ii) each circle is formed by a different aspect and iii) circles can overlap meaning that strong circles can be within weaker ones. These characteristics define circles as some kind of network community, and we want to evaluate how far this agreement holds.

Fang et al. [10] investigate the impact of the circle-sharing feature for the network growth and structure of the Google+ social network. According to their clustering, shared circles can be categorized into two main groups: communities, which have a high link density and reciprocity, as well as celebrities, which have a low in-circle density, low reciprocity but a very high in-degree. These two categories show the main motivation of users to share their circles: they want to share group of users, which may have a common attribute or represent a facet of their life, or they want to share a group of popular people.

In a methodological study of common concern, Yang and Leskovec [11] evaluate community metrics on data sets, which include ground-truth communities. The authors map ground-truth communities to crawled network communities, which are explicitly labelled in the provided evaluation data. They use 230 different social, collaboration and information networks to test 13 scoring functions, which characterize how well a set of vertices is connected. As one of their major findings, the scoring functions correlate and can be grouped in four subsets based on the community characteristics they measure. We base our work on this categorization of scoring functions and use three data sets which the authors provide.

III. Circles vs. Communities

Circles are a core features of Google+. While other social networks like Facebook support groups for selective sharing, Google+ forces the user to put new contacts in circles. It thus inverts the application logic of communities. Whereas groups in traditional OSNs form optional overlays of the social graph, circles are mandatory sub-structures of a user network in Google+. Some users try to avoid this categorization by putting all contacts in one circle, but the majority actively adopts this perspective change when building social contacts [12].

The mapping from contacts to users is only visible to the creator of the circle, which makes quantitative evaluations on circles difficult from the public perspective. The only way to extract circles from Google+ without asking users to manually publish their circles is to search for posts with shared circles. Users are able to share circles they created by others. While reasons for sharing a circle are manifold, this feature creates an opportunity of accessing the structure of circles. The limitation of shared circles is that they may be created only for sharing and do not represent an actual facet of the creator. Fang et al. [10] found that there are two main categories of shared-circles. Circles, that cover communities. They have a very high density and a high reciprocity with the circle owner. And there are circles which cover very popular users. The popularity is determined by the users in-degree.

Communities can be found in different types of networks like social, biological, or information networks. They all refer to a common attribute of the vertices that centers connectivity. A general definition of communities in networks is that a community is a set of vertices with many connections within the set and just few connections from the community to the remaining network [11], [13].

Like classical communities, circles are created around a common attribute. Unlike communities, circles directly relate to the creating user and may only carry meaning for its creator. Thus a community can be regarded as a member-initiated rendezvous, whereas a circle more closely resembles a replication channel of the initiator. The underlying nature of sharing is pull-based for communities, but push-based for circles. Another distincton of circles in comparison to classical communities is that only vertices from the ego-network of the creator can be added to the circle.

IV. Data Description

In the remainder of the paper, we refer to the social graph as the directed graph $G(V, E)$, where V are the user profiles and E the relations between them. A vertex $v \in V$ has an ID to identify it. Other available attributes are not considered, as we are only interested in evaluating the structural properties of the social graph. Based on the directed graph $G(V, E)$, we

TABLE I. NOMENCLATURE

Notation	Description
n	Number of vertices in the graph
m	Number of edges in the graph
f	Scoring function
C	Circle/ community
n_C	Number of vertices in C
m_C	Number of edges in C
c_C	Number of edges at the boundary of C
$d(v)$	Degree of vertex v

define n as the number of vertices in the graph $n = |V|$ and m as the number of edges $m = |E|$. A scoring function f is applied to a circle C of n_C vertices in C, $n_C = |C|$, m_C edges in C and c_C edges on the boundary of C. $d(v)$ is the degree of a vertex v.

A. Google+

Even if the contacts of a user are publicly available and there are shared circles, it is difficult to load the data from Google+, because the official API does not provide calls for loading relationships. We use the only available data set provided by McCauley and Leskovec [9], which includes circles. It covers 133 ego-networks of users, who share at least two circles. While the authors also use Facebook and Twitter for their evaluations, they crawl a Google+ data set. Since circles are per default private to users, they had to select users, who shared at least two circles, and then crawled their ego-network. Even though the data set only consists of ego-networks, by joining all ego-networks a large connected component is formed with 107.614 vertices and 13.673.453 edges.

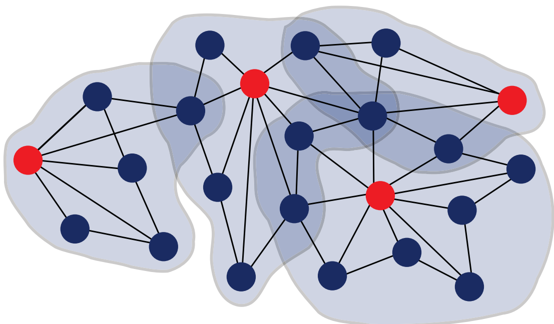


Fig. 1. Schematic visualization of the data set from McAuley and Leskovec: Red vertices are the owners of the ego-networks, the vertices in the overlapping ego-networks are the bridges between the ego-networks.

Data, which is generated with this method, is biased towards too dense components, connected only by a few edges between each other. Figure 1 visualises the nature of the data set. It consists of ego-networks (light blue areas) that include the owners (red vertex). Owners are indirectly connected to each other via blue vertices from their ego-networks. Because several vertices appear in more than one ego-network, the joint graph of all ego-networks is a fully connected component. We find that 93,5 % of the ego-network in the data set overlap. This means that they share at least one common vertex. The overlap of ego-networks is shown in Figure 2. More than 55.000 vertices are only in one, and around 14.000 in two networks. Comparing to this high count of vertices, there are just a few vertices, which are members of more than 50 ego-networks. These vertices have a high impact on the connectivity of the data set.

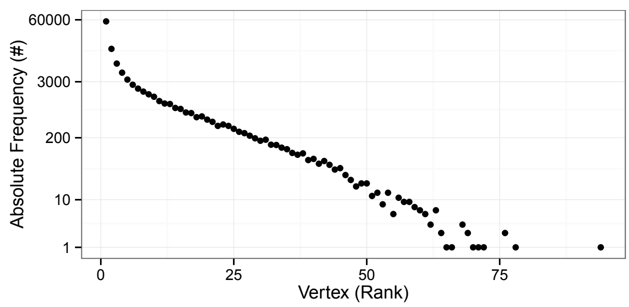


Fig. 2. Log plot of vertex membership count in ego-networks.

Before we start with the core evaluations, we want to characterize the data set by applying common structural features of Online Social Networks: Degree Distribution, Clustering Coefficient and Node Separation [14].

1) Degree Distribution: The degree $d(v)$ of a vertex v in a graph G is defined as the number of connected edges. Previous research on Online Social Networks indicated that the distribution of the degree often follows a power-law distribution. Magno et al. [8] even claimed a power-law distribution in their Google+ data set. Gonzales et al. [7] also found a power-law distribution in a snapshot of the largest connected component they crawl. To investigate the degree distribution in the McAuley and Leskovec [9] dataset, we use a set of methods introduced by Clauset et al. [15]. The authors stress that determining a power-law distribution by simply comparing plots is insufficient. Previously, Van Mieghem [16] had disclosed several prominent claims of power laws in the Internet as mistaken. Following the method, we create models for a power-law, exponential and log-normal distribution and than check which fits best to the degrees in the used data set using the log likelihood ratio.

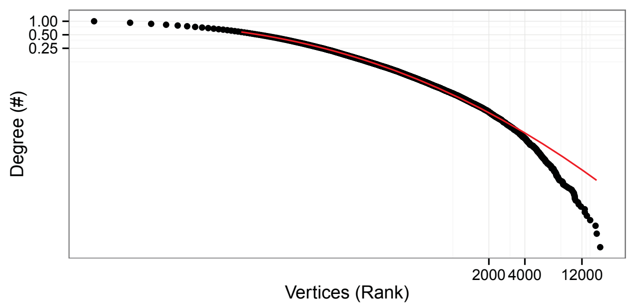


Fig. 3. Log-log plot of the in-degree distribution of the Google+ data set. The red solid line shows the fit of the log-normal distribution.

When following the proposed method, we cannot match a power-law distribution to the degree distribution of the Google+ data set. Rather we find an approximate fit of a log-normal distribution for the in-degree of the vertices (see Figure 3). The latter corresponds to an exponential of a Normal random process and would indicate that inviting membership into circles is rather a multiplicative than an additive process.

2) Clustering Coefficient: Besides the degree distribution, the Clustering Coefficient (CC) is a common measure on OSN. It tests for the local clustering structure of vertices, i.e., whether neighbours of a vertex tend to connect to each other as well [14]. This property is formalized by counting the number of triangles a vertex is part of in relation to the maximum number of triangles a vertex could possibly be part of [8].

The distribution of the clustering coefficient for our Google+ data set is displayed in Figure 4. It admits a smooth, almost symmetric shape with an average of 0.4901. Clustering thus seems a rather random phenomenon in the data without any pronounced tendency. This indicates that users in Google+ interconnect rather independent of a common circle membership. These findings are in contrast to the results of previous measurements. Magno et al. observe asymmetry with a low average coefficient of ≈ 2.5. Gong et al. [17] continuously evaluated the CC during the creation phase of Google+. The highest value they observed at the very beginning of Google+ was about 0.32. These previous results rather indicate a layout of many small stars with neighboring users that do not connect.

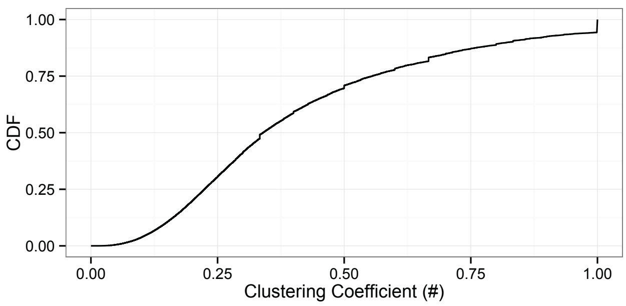


Fig. 4. CDF of the Clustering Coefficient of the Google+ data set.

3) Node Separation: As a third characteristic of OSNs, we measure the Node Separation. Node Separation relates to the initial observation from social network research, known as the small world problem by Milgram 1967 [18]. The diameter of a network, which is the length of the longest shortest path between all vertices, remains surprisingly short in the small world. While we observe a diameter of 13, Gong et al. [17] find a value of 6 in their data set, and Magno et al. [8] even see 19. We also evaluate the average path length of the dataset with 3.32 (Magno et al. 5.9).

Comparing the statistics from our evaluation with the results obtained by Magno et al. [8], it becomes visible that the data set we use is significantly smaller and better connected. Table II presents an overview of the two data sets. Magno et al. crawled a graph that has 326 times more vertices than ours, but the edge count is only enhanced by 40. The Magno data set only has 42 times more edges than ours. The higher connectivity visible in our evaluation corresponds to a divergent average degree of the vertices in the two networks. While the Magno network has an average of 16 edges per vertex, the vertices in our data set have an average degree of 127. This difference is caused by the crawling methods. While Magno et al. use a breath-first-search, the McAuley and Leskovec data set is crawled around predefined users and their ego-networks. It is important to consider the high connectivity of our data set in the evaluation, because we investigate the structure of circles within the network, focusing on their internal connectivity. Metrics applied to the circles, which are based on the vertex edge ratio, will produce high values, because of the overall high density in the network as compared to the result of the same metric in other Google+ data sets.

TABLE II. STATISTICAL COMPARISON OF THE MCAULEY AND LESKOVEC [9] AND MAGNO ET AL. [8] DATA SETS. (ASP: AVERAGE SHORTEST PATH)

Metric	Magno et al.	McAuley and Leskovec
Vertices	35.114.957	107.614
Edges	575.141.097	13.673.453
Diameter	19	13
ASP	5.9	3.32
Degree		
distribution (in)	power-law $\alpha = 1.3$	log-normal
distribution (out)	power-law $\alpha = 1.2$	-
average degree (in)	16.4	127
average degree (out)	16.4	189

B. Other Data Sets

For comparison, we use several other social network graphs in addition to the Google+ data set. To extend the analysis of circles, we add a Twitter data set to our evaluation. Like Google+, Twitter consists of a directed graph created by users start following the post of others. The data set is also provided by McAuley and Leskovec [9] and forms a graph with 81,306 vertices and 1,768,149 edges. The 100 included communities are created by a selective sharing feature called *lists* in Twitter.

To compare these circle-like structures with classical communities, we also use data sets of two traditional OSNs. The first data is crawled from the LiveJournal OSN by Yang and Leskovec [11]. It includes 3,997,962 vertices and 34,681,189 edges. The communities within the network are explicitly labelled as interest-based groups. The same type of communities are included in the second traditional OSN data obtained from Orkut. The Orkut data set has been crawled by Mislove et al. [19] and provided by [11]. It subsumes 3,072,441 vertices and 117,185,083 edges. We restrict our evaluation to the top 5000 communities of these two graphs ranked by their size, to limit calculation complexity. A summary of all data sets is shown in Table III.

The major difference between the circle-type and community data is their edge type. While Google+ and Twitter use directed edges to represent a link between its users, LiveJournal and Orkut use undirected links. This may have an impact on our evaluations, because a bidirectional relation between two vertices is represented by one edge in undirected and by two in directed graphs. Directed communities could score higher than undirected ones, when edge counts are used in the scoring functions. A fully connected set of vertices in a directed graphs has twice the number of edges than a fully connected set of vertices in an undirected graph. To test for this impact on our evaluation, we preformed the scoring functions on the Twitter and Google+ and on a undirected representation of these networks with bidirectional edges combined to one.

TABLE III. COMPARISON OF THE EVALUATED DATA SETS

Graphs	Google+	Twitter	LiveJournal	Orkut
Vertices	107,614	81,306	3,997,962	3,072,441
Edges	13,673,453	1,768,149	34,681,189	117,185,083
Type	directed	directed	undirected	undirected
Structure	Circles	Circles	Communities	Communities
# Communities	468	100	5000	5000

The results show a minimal deviation of about 2,38 % and do not have an impact on the results of our overall evaluation.

V. EVALUATION

The goal of our evaluation is to investigate how far the characteristics of classic network communities coincide with circles in Google+. Here the first question was, whether circles form pronounced structures in the social graph. The second question asked whether these structures resemble classic communities in typical social networks. To quantify the particularity of communities, we use scoring functions from the field of community detection. These functions score sets of vertices with extremal values, whenever they match the concept of a community.

a) Internal Connectivity—Average Degree: A community should consist of a set of highly interconnected vertices. A scoring function that focuses on the internal connectivity is the Average Degree [20]. In the definition provided by Yang and Leskovec [11]

$$f(C) = \frac{2m_C}{n_C}, \tag{1}$$

the mean degree of the vertices in C is calculated by twice the number of edges within the community $2m_C$ to obtain the link contacts at each vertex, divided by the total number of vertices n_C in C. Values of this function depend on the density of the underlying social graph.

b) External Connectivity—Ratio Cut: The second characteristic of a community is its separation from the remaining network. That means, the vertices in the set share just a few links with its embedding graph. The Ratio Cut function is based on the edges which are at the border of a community c_C, divided by the balancing product of the number of vertices within C and the complement of C [13].

$$f(C) = \frac{c_C}{n_C(n - n_C)} \tag{2}$$

By counting bordering edges, only, the Ratio Cut remains independent of the connectivity within the community—and a measure opposed to the Average Degree.

c) Combined Internal and External Connectivity—Conductance: While Average Degree and Ratio Cut only focus on a single aspect of the relation between the community and network, the Conductance metric covers both aspects. This is achieved by measuring the number of edges pointing outside the community in relation to the total degree of the vertices within the community [21].

$$f(C) = \frac{c_C}{2m_C + c_C} \tag{3}$$

In a sense, the Conductance scoring function is able to capture the common intuition of a community. A well pronounced community will have a low score, due to many internal edges and significantly fewer inter-connects to the 'outside world'. It is noteworthy that by evaluating edge ratios, the Conductance corrects for the density of the underlying graph.

d) Modularity: A very popular scoring function is Modularity introduced by Newman [22]. It uses a so called null model to verify whether the given network has a community structure or not [13]. The model we use was proposed by Newman and Girvon [23] and consists of a randomized graph, which has the same degree sequence as the original graph. We use the algorithm proposed by Viger and Latapy [24] to generate the random graph. Based on the null model, the Modularity is positive, whenever the number of edges inside the community is higher than the expectation value of edges in the null model.

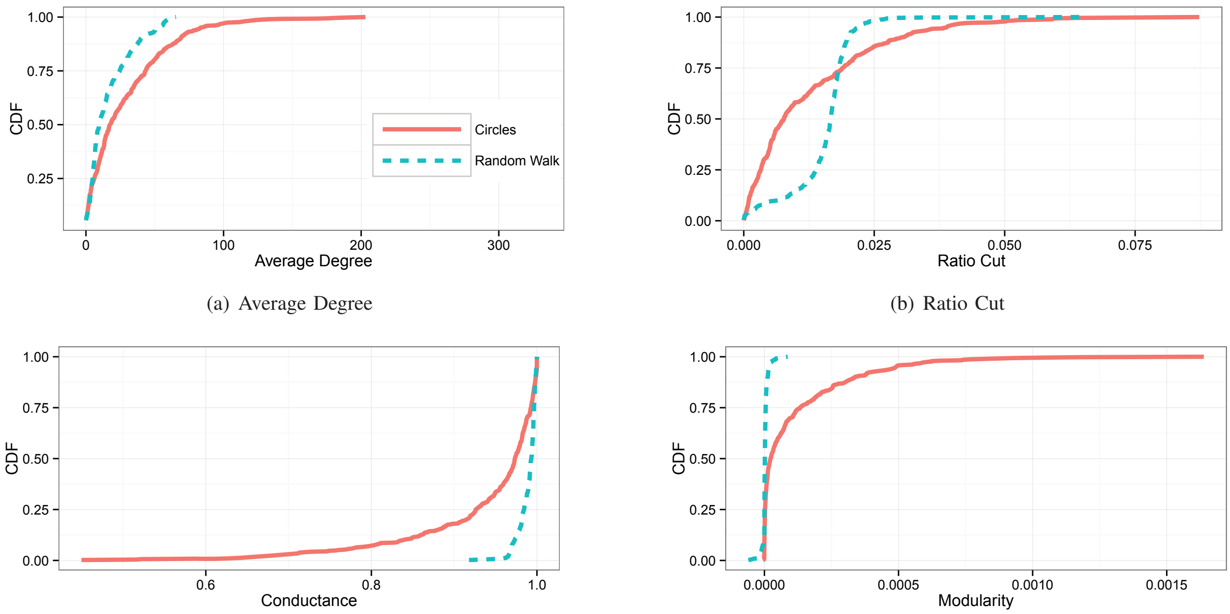
$$f(C) = \frac{1}{2m}(m_C - E(m_C)) \tag{4}$$

A. The Structures of Circles within the Social Graph

We now want to answer the question, whether circles form pronounced structures in the social graph. We apply the four scoring functions described above to the circles available in the Google+ data set and to randomly selected sets from the graph with the same size as the circles. The functions scores should clearly separate circles from random sets, whenever community structures are present. We sample the random sets by preforming random walks on the graph. Starting from a randomly selected vertex, the walk continues by selecting neighbors at random until sufficiently many vertices are found. The walks is restarted, whenever no new neighbour is available. While generating an unbiased selection of the sub-graph [25], the method of performing random walks produces a widely connected, representative network.

Figure 5 comparatively displays the CDFs for the circle data and the randomly traversed sets. All four functions clearly differentiate circles from the random sets. Average Degree and Conductance are separated only quantitatively for the circles and the random sets, but attain distributions of similar shape. While circles score higher on Average Degree, the overall connectivity measured by the Conductance is significantly lower. Without surprise, the selection of random paths in a connected social graph tends to produce a more even, fairly flat network.

The scores for Ratio Cut (Fig. 5(b)) produce two different distributions for the circles and the random sets. While the scores show a smooth rise, the random sets peak around 0.02. This peak reflects the average ratio between connected vertices and edges in between them, since Ratio Cut scores the edges on the boundary and the included vertices. It is also worth mentioning that the score for more than 70% of the circles is lower than for the random sets. Circles are thus better separated from the remaining network than an average subset. The clearest differentiation between the random sets and the circles is visible from the Modularity function (Fig. 5(d)). Scores for the random sets are very low, which indicates a good coincidence of our selection from random walks with the random null model of Fortunato [13]. More than 50 % of the circles show a significant deviation from the expected results given by the null model.



(a) Average Degree

(b) Ratio Cut

(c) Conductance

(d) Modularity

Fig. 5. Group Metrics applied to Circles in the dataset (red solid) and group metrics applied to a selection of vertices obtained from a random walk (blue dashed).

All distributions of circles admit smooth long tails. This shows, that there is always a small but finite probability of the circles for significantly different graph properties. A justification for the tails could be found in the work of Fang et al. [10], who claimed that there are two categories of shared-circles in Google+. Communities, which are highly connected and celebrities, which have a low connectivity but very popular members (i.e., star-like structures). Their first category of circles produces dominant scores, while a few circle of, e.g., pop singers have a low internal connectivity and score low.

To intermediately sum up the results from applying the scoring functions to circles and random sets, we could show that our tests for community characteristics did produce significantly diverging results. So we can conclude that circles form pronounced structures which differ from the nature of the entire graph.

B. Circles vs. Classical Communities

In the following, we want to find out how similar the structures of circles approach the communities that are observed in traditional social networks. We use the data sets from LiveJournal and Orkut networks (Section IV-B) as community samples and compare to circle-type structures using the data from Google+ and also from Twitter. We chose to compare to additional Twitter data to avoid a bias inherited from the individual measurement on Google+.

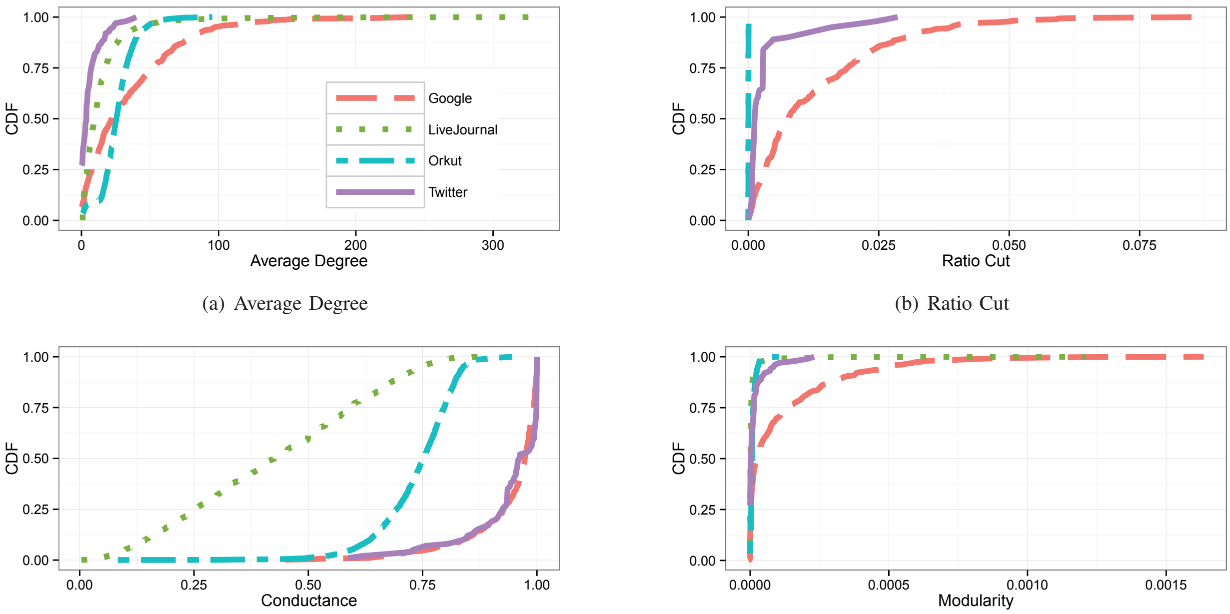
Figure 6(a) shows the CDF of the four social network graphs for the Average Degree function. While there is a large difference in the maximum score in the data sets (Twitter: 40, LiveJournal: 332), no significant difference in the shape of the distribution of the community types can be found. Google+ admits the lowest Average Degree.

The CDF of the Ratio Cut function displayed in Figure 6(b) attains a divergent picture for external connectivity. While LiveJournal and Orkut – the two original community networks – have vanishing relative frequencies for linking to the outside,the networks with circles score visibly higher (Twitter has a mean of 6 and Google+ 34) than for the classical communities. Connectivity beyond the group structure is thus omnipresent in circles, but utmost uncommon in communities.

A reason for this phenomenon lies in the perspective of the creators, who may not be connected with the entire community, they want to share. This is supported by the findings of Fang et al. [10], who found that sharing a circle leads to a densification of community circles, because missing members of the community can create connections to user, they did not connect yet. This phenomenon goes along with the overall high density within an ego-network, which also lead to poor separation between circles and the remaining network.

The Conductance function, which scores densely on the internal and sparsely for externally connected communities with a low value, shows the most striking difference between the data sets. While LiveJournal almost attains a uniform distribution, and Orkut still has 50 % of the communities below 0.75, almost all circles ($\approx$ 90 %) show a Conductance close to 1 (> 0.9). These results report on a remarkably high balance between internal and external link density in the group formation of Google+ and Twitter, which is caused by the same phenomenon observed for the Ratio Cut.

The CDFs for the Modulartiy scoring function, plotted in Figure 6(d), all show a similarly high increase. While Twitter, LiveJournal and Orkut nearly overlap, Google+ shows a smoother curve and a higher maximum value. This implies that the Google+ circles show a higher connectivity than the



(a) Average Degree

(b) Ratio Cut

(c) Conductance

(d) Modularity

Fig. 6. CDFs of the results from the scoring function applied to the communities in the four data sets.

other data sets in comparison to their null model. Thus the Google+ graph has the highest connectivity, which is also shown by the high scores of the Ratio Cut function. It also has to be mentioned that the overall scores for Orkut are very low compared to the other data sets, which is also visible for Twitter.

Summing up the results of this evaluation, we could show that circles internally attain the same structure as classical communities. However, the separation of circles from the remaining network is significantly reduced. The organisation of ego-networks in Circles thus produces much less confined group structures than communities, which are shaped by actively joining users. While this may appear surprising at first sight, an explicit consideration of context may open a more appropriate view. Circles represent the context of its creator[1], while community members actively opt for its context. It may be just more likely that users who join a community have a shared perspective centered around its subject.

As a consequence, members of circles should regard selective sharing more thoughtful than community members. The latter are part of much clearer determined context group.

VI. Conclusion and Outlook

By introducing the concept of circles, Google+ enables users to transfer the different facets of their social live to the management of their online contacts and share information selectively. While in traditional communities users join on their own will, the circles are created by a peer user from his own ego-network. In this work, we tried to explore the effects of these different building mechanisms on the social networks and on the processes of selective sharing therein.

Based on a Google+ data set with shared circles, we characterized the structures of circles that are embedded in the Google+ social graph. We started with classification of the used data set by comparing it characteristics to other Google+ data sets. Comparing two social networks of circular structures with two data sets that are built from traditional communities, we could show that i) circles form pronounced community-like structures in Google+, and ii) circles attain an individual structural signature. In particular, circles are significantly less separated from the remaining network than classical communities. Selective sharing in Google+ is thus more diffusive and less confined.

In our future work, we will extend our research on group structures from a global to an ego-centred view. While only publicly shared circles were available for this evaluation, private circles could provide a deeper insight into the user-centric deployment of circles.

[1]In a sense, placing a user in a Circles resembles the placement of a hyperlink. For hyperlinks, context is well known to split into two, the context of departure and that of arrival [26].

Acknowledgment

We would like to thank the anonymous reviewers for their very thoughtful and encouraging comments. This work is funded by the German Federal Ministry of Education and Research (BMBF) within the project Mindstone, see http://mindstone.hylos.org.

References

[1] S. Kairam, M. Brzozowski, D. Huffaker, and E. Chi, "Talking in circles: Selective sharing in google+," in *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*. ACM, 2012, pp. 1065–1074.

[2] S. D. Farnham and E. F. Churchill, "Faceted identity, faceted lives: social and technical issues with being yourself online," in *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, ser. CSCW '11. New York, NY, USA: ACM, 2011, pp. 359–368.

[3] S. L. Feld, "The focused organization of social ties," *American journal of sociology*, pp. 1015–1035, 1981.

[4] D. Schiöberg, S. Schmid, F. Schneider, S. Uhlig, H. Schiöberg, and A. Feldmann, "Tracing the birth of an osn: Social graph and profile analysis in google+," in *Proceedings of the 3rd Annual ACM Web Science Conference*. ACM, 2012, pp. 265–274.

[5] H. Roreger and T. C. Schmidt, "Socialize Online Learning: Why we should Integrate Learning Content Management with Online Social Networks," in *Proc. of IEEE Intern. Conf. on Pervasive Computing and Communication (PerCom), Workshop PerEL.* Piscataway, NJ, USA: IEEE Press, March 2012, pp. 685–690.

[6] S. Brauer and T. C. Schmidt, "Group Formation in eLearning-enabled Online Social Networks," in *Proc. of the International Conference Interactive Computer aided Learning (ICL'12)*, M. E. Auer, Ed. Piscataway, NJ, USA: IEEEPress, Sep. 2012.

[7] 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," in *Proceedings of the 22Nd International Conference on World Wide Web*, ser. WWW '13. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee, 2013, pp. 483–494.

[8] G. Magno, G. Comarela, D. Saez-Trumper, M. Cha, and V. Almeida, "New kid on the block: exploring the google+ social graph," in *Proceedings of the 2012 ACM conference on Internet measurement conference*, ser. IMC '12. New York, NY, USA: ACM, 2012, pp. 159–170.

[9] J. McAuley and J. Leskovec, "Learning to discover social circles in ego networks," in *Advances in Neural Information Processing Systems 25*, 2012, pp. 548–556.

[10] L. Fang, A. Fabrikant, and K. LeFevre, "Look Who I Found: Understanding the Effects of Sharing Curated Friend Groups," in *Proceedings of the 3rd Annual ACM Web Science Conference*, ser. WebSci '12. New York, NY, USA: ACM, 2012, pp. 95–104.

[11] J. Yang and J. Leskovec, "Defining and Evaluating Network Communities based on Ground-truth," in *Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics*, ser. MDS '12. New York, NY, USA: ACM, 2012, pp. 3:1–3:8.

[12] J. Watson, A. Besmer, and H. R. Lipford, "+Your Circles: Sharing Behavior on Google+," in *Proceedings of the Eighth Symposium on Usable Privacy and Security*, ser. SOUPS '12. New York, NY, USA: ACM, 2012, pp. 12:1–12:9.

[13] S. Fortunato, "Community detection in graphs," *Physics Reports*, vol. 486, no. 3-5, pp. 75–174, 2010.

[14] A. Sala, L. Cao, C. Wilson, R. Zablit, H. Zheng, and B. Y. Zhao, "Measurement-calibrated graph models for social network experiments," in *Proceedings of the 19th international conference on World wide web*, ser. WWW '10. New York, NY, USA: ACM, 2010, pp. 861–870.

[15] A. Clauset, C. R. Shalizi, and M. E. Newman, "Power-law distributions in empirical data," *SIAM review*, vol. 51, no. 4, pp. 661–703, 2009.

[16] P. Van Mieghem, *Performance Analysis of Communications Networks and Systems*. Cambridge, New York: Cambridge University Press, 2006.

[17] N. Z. Gong, W. Xu, L. Huang, P. Mittal, E. Stefanov, V. Sekar, and D. Song, "Evolution of Social-Attribute Networks: Measurements, Modeling, and Implications using Google+," in *Proceedings of the 2012 ACM conference on Internet measurement conference*. ACM, 2012, pp. 131–144.

[18] S. Milgram, "The small world problem," *Psychology today*, vol. 2, no. 1, pp. 60–67, 1967.

[19] A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee, "Measurement and Analysis of Online Social Networks," in *Proceedings of the 5th ACM/Usenix Internet Measurement Conference (IMC'07)*, San Diego, CA, October 2007.

[20] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi, "Defining and identifying communities in networks," *PNAS*, vol. 101, no. 9, 2004.

[21] J. Shi and J. Malik, "Normalized cuts and image segmentation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 8, pp. 888–905, 2000.

[22] M. E. Newman, "Modularity and community structure in networks," *Proceedings of the National Academy of Sciences*, vol. 103, no. 23, pp. 8577–8582, 2006.

[23] M. E. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Physical review E*, vol. 69, no. 2, p. 026113, 2004.

[24] F. Viger and M. Latapy, "Efficient and simple generation of random simple connected graphs with prescribed degree sequence," in *Computing and Combinatorics*. Springer, 2005, pp. 440–449.

[25] S. Lu, J. Kang, W. Gong, and D. Towsley, "Complex Network Comparison Using Random Walks," in *23rd Intern. World Wide Web Conf. WWW '14 (Companion Volume)*. New York, NY, USA: ACM, 2014, pp. 727–730.

[26] G. P. Landow, "The rhetoric of hypermedia: Some rules for authors," *Journ. of Comp. in Higher Education*, vol. 1, no. 1, pp. 39–64, 1989.