INSTITUTO DE INVESTIGACIONES ECONOMICAS Y SOCIALES DEL SUR

Documento de Trabajo Proyecto Unidad Ejecutora

"Inclusión social: innovaciones y políticas públicas. Un análisis regional"

Nº10

"Small World Topologies of Facebook Pages in Bahía Blanca"

Autores:

Emiliano Gutiérrez 1,2

Juan M.C. Larrosa 1,2

12/11/2019

[1] Departamento of Economics, Universidad Nacional del Sur.

[2] Instituto de Investigaciones Económicas y Sociales del Sur (IIESS- CONICET).

Abstract

Networks that exhibit large hubs are prone to show higher connectivity than random structures. This feature is also reflected in the presence of power law distribution in the degree distribution of nodes. We check for the presence of power law degree distributions in Facebook Page Networks (FPN). We analyze data collected from 325 Facebook fan pages from an Argentinian median city summarizing 94 thousand nodes and 6.5 million edges during a whole year. We found that whole network exhibits power law degree distributions but when we analyzed the subnetworks that compose it only a handful of marginal networks in terms of size and level of engagement. We find that main subnetworks exhibit small world properties such as average geodesic distance approximates and correlates to the natural logarithm of the graph range. We begin our contribution by surveying the topic of power law in degree distribution and small world in online networks. We estimate the slopes of the log-log degree distribution for each network for identifying free scale structures and statistically associate network structural properties with the presence of power law distributions. We present statistical results that associate structural parameters with the presence of power law distributions and small world effect shading light to new patterns for detecting the properties in networks.

1. Introduction

Social capital is inserted in Facebook sites (Brooks et al., 2014) and people use their personal sites for interacting with relatives, friends, acquaintances, and even unknown pairs. The rich diversity of interactions made this OSN a laboratory for testing diverse hypothesis. Whole new markets may emerge connecting users with firms on each commercial sector.

In this work, we sample different Facebook fan page from business geographically located in an urban area comprehending different socio-economic sectors and institutions. We look for understanding how the structure of these networks change across sectors.

Related contributions on Facebook has centered mainly on people (Backstrom et al., 2011; Ugander et al., 2011). Backstromet al. (2011) affirms “six degrees of separation” to be display within the Facebook chart, besides a significant neighborhood clustering. That is, Facebook, at the level of person clients, appears to be a small-world framework, and the degrees of separation would afterward be decreased to four in Ugander et al. (2011). Eikmeier and Gleich (2017) test spectral methods for detecting power law distributions in degree frequency distributions from several real networks finding that a significant percentage as following power law. Slattery et al. (2013) find that FPN formed by Likes follows power law distributions but Wohlgemuth and Matache (2014) find that power law and small coexisted in many Facebook groups. We find that the giant component of the FPN has free scale distribution in out-degree frequency and in-degree distributions but only a handful of subnetworks follow power law.

Power-laws and small world networks are a key component in any characterization of the networks gathered from the Internet and other large information sources (Blackmore, 2001). These include online social networks (OSN), forums, search engines, web crawls and recommender systems, among other examples. There are quite a few places that power-laws may arise in the description of these networks. For instance, the degree distributions of these networks are usually observed to have a power-law (Clauset et al. 2009) however some recent discussions (Holme, 2019). In this sense, the scale free form law has her root in the called “Zipf law”. This law was originally formulated in the context of analyze the frequency of words in natural language (Zipf, 1936) where it’s possible observe an inequal distribution. Many words have a small frequency, but a few words have a high occurrence.

In particular, the Facebook Page Network (FPN) is one of the primary avenues where firms and products engage. FPN provides a space where individuals may interact with various pieces of digital media, comment on publications by the business process outsourcing (BPO), and share certain items with friends within their own personal network. If a person would like to receive updates or display her preference or affiliation toward a particular BPO, one may do so by clicking *Like* on the BPO, which is then publicly displayed on the individual’s own page. Individuals may choose to interact with a BPO digitally for a variety of reasons: to genuinely support the BPO, access locked content, receive member benefits, derive entertainment value, acquire information, or signal things about themselves such as interests and associations. BPOs may choose to publicly interact with other BPOs for a variety of reasons as well: to reflect business associations, affiliate with a particular cause, or lure a potential response from the target BPO. Like some individuals BPOs may use this opportunity to enhance their public online appearance.

Networks are represented by graphs that connect vertices or nodes and links or edges. In the case of the FPN, a node is an entity, either an individual or a BPO, and a link is a directed Like between them. We also add links by commenting or posting. The total number of interactions is the degree of the node, and the degree distribution can be obtained by calculating the frequency of these degrees.

2. Network definitions

We can define a network as $G = (V, E)$, where *V* is a set of vertices $V = \{x_1, x_2, \dots, x_n\}$ with *N* vertices and *E* edges, when each vertex could be linked with other vertex. For instance. the pair $\{x_i, x_j\}$ is a link when the vertex x_i is connected with x_j.

Furthermore the connectivity degree for the vertex $x_j(k_i)$ is defined like the number of edges with the other vertices. If the parameter k_i = 0 then, the vertex is disconnected and it haven't any link with the other vertices. In this sense we could define the degree distribution like the distribution probability function $f(x) = P(k_i = x)$.

Thus, for evaluate of the degree to which vertices in a network tend to clustering together. Firstly, it's possible define the local clustering coefficient for each vertex. The local clustering is the proportion for the number of the links existing for the neighbors respect the all the possible links between them. The local clustering coefficient is defined such that:

$$C_i = \frac{|\{e_{j,s}\}|}{k_i(k_i - 1)}$$

Nevertheless for measure the clustering in a network, Watts and Strogatz (1998), propose a network overall level of clustering in a network, which is the average of the local clustering coefficients of all the vertices:

$$\bar{C} = \frac{1}{n} \sum_{i=1}^{N} C_i$$

Another metric very important is the average path length. This measure provides the average number of steps along the shortest paths for all possible pairs of network vertices (*d*), excluding the vertices with k = 0. Then, the average path length is

$$l = \frac{1}{n\,(n-1)} \sum_{i \neq h} e_{i,j}$$

Watts and Strogatz (1998) proposed a specifically classification for networks where the average path length scale with $ln\, n\, (l \approx \frac{log(n)}{\log(k)})$ and have high clustering coefficient than random networks. This type of networks is called "small-world networks". This network's topology could be found in social networks (Wohlgemuth and Matache, 2014), co-authorships (Goyal, Leij and Moraga-González, 2003, Ebadi, and Schiffauerova, 2015), terrorism (Badia I Dalmases, 2011, Lindelauf, Borm and Hamers, 2009), epidemic (Liu et al., 2015), among others.

In the other hand, Albert and Barabási (1999), consider the existence the scale free networks, where the probability a new node will be connected to vertex is determined for connectivity degree k_j

$$P(k_i) = \frac{k_i}{\sum_j k_j}$$

*A scale free networks are specific networks which the degree distribution obeys a power law (*Barabási, Pósfai, 2016)*.* In this sense, the probability a vertex has exactly *k* edges is

$$p_k = Lk^{-\gamma}$$

Given the cumulative probability is equal to 1 and *L* is a constant term, we could rewrite the previous expression as

$$L \sum_{k=1}^{\infty} k^{-\gamma} = 1$$

Hence,

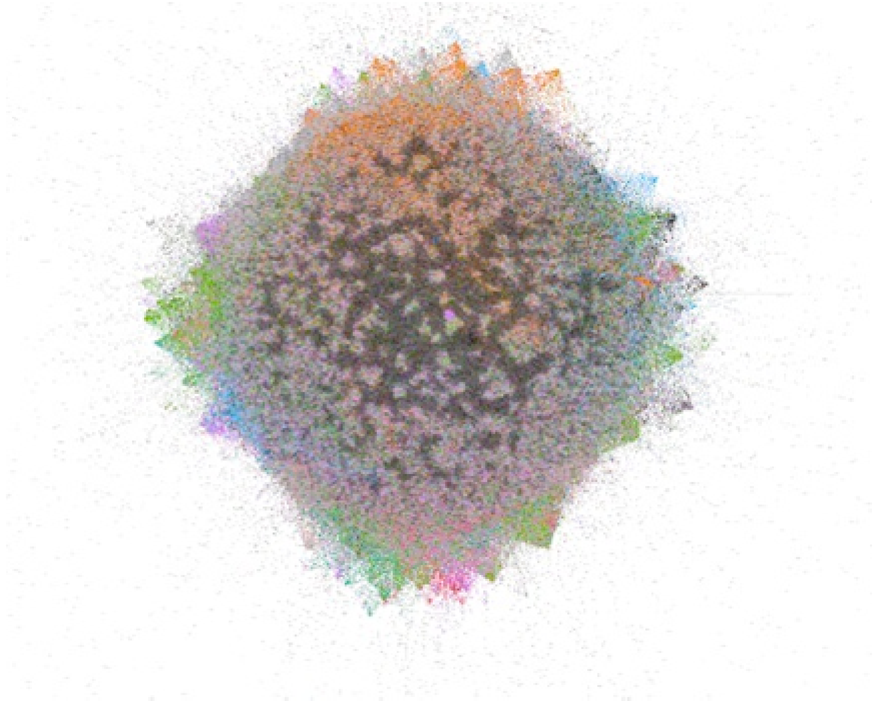
$$L = \frac{1}{\sum_{k=1}^{\infty} k^{-\gamma}} = \frac{1}{\zeta}$$

Where ζ is known as the Riemann-zeta function. Thus, is possible redefine the probability like as log-log function

$$\log p_k = -\gamma \log k + \log L$$

And if the γ value is between 2 and 3, it is possible to characterize a network as scale free *(*Barabási and Pósfai, 2016).

Figure 1. The entire scrapped network



.

Additionally, another contributions find scale free networks in business (Rendón de la Torre et al., 2019), disease sexual transmission (Schneeberger et al. 2004), social networks (Sadri et al, 2017), among others.

However, an interesting discussion is presented in Broido and Clauset (2019), about if the scale free networks are often in the real world. This research studied nearly 1000 real networks (social, biological, technological, transportation, and information networks) and found only a minimum percentage present a strongest evidence about the scale free topology.

3. Data Analysis

We obtain data from Facebook fan pages with almost all business with physical store in the city of Bahia Blanca (Argentina). We scraped 325 networks that comprehends 93,692 nodes with 6,638,433 edges (including self-loops) by using the NodeXL software (Smith et al, 2010). This was made prior to the 2018 Cambridge Analytic scandal so data represents the capture of entire FPN. The completed network that includes all the 325 subnetworks is presented in Figure 1. Networks are Facebook fan pages captured along the year 2015 (from 1/1/2015 to 12/31/2015). Network seems highly dense but represents the interaction of an entire year. We capture a significant part of different digital of an entire city with a total population of 335,000 inhabitants. Table 1 shows descriptive statistics of main structural metrics from all subnetworks. We observe and average size of almost 590 nodes (*N*) but with a great standard deviation (s.d.). Edges (*G*) are on 25 thousands on average, again with great s.d. Geometric means exhibit a more accurate picture given the assumed non-normal data distribution with average range of 155 and average edges of almost 1,463. Reciprocity indexes are low (RVPR and RER) with an index higher in edge reciprocity. Isolated nodes (SVCC) are negligible while the greatest component has 866 nodes (MVCC) and 27,289 edges (MECC). Average diameter is small, less than four steps away (MGD) and short average geodesic distance (AGD) are indications of small world phenomenon. Density is also low (Dens) but that's not implies much about the presence of clusters. Assortativity indexes indicate dissortative networks on average (IIDA, IODA, OIDA, and OODA) especially in the case of output-input degree assortativity, i.e. nodes with higher out-degree tend to avoid connecting to nodes with higher in-degree. Average degree (AG) and betweenness centralization (Bet) are high, depicting networks with hubs centralizing information distribution. All indications of community detection are high: Modularity (Mod) and Clustering Coefficient (Clus), Watts-Strogatz Clustering Coefficient (WSCC), and Network Clustering Coefficient (Transitivity) (NCC (T)) consistently reveal groups in all subnetworks, with low s.d. in all of these metrics.

Finally, female (pF) participation is overwhelming ranging to .64 on average, while male participation reaches only a quarter on average in the whole sample. Business (pB) and Other (pO) range .4 and .3 on average, respectively.

Table 1. Descriptive statistics of network structural metrics

Code	Mean	Standard deviation	Mode	Median	Geometric Mean	Kurtosis	Skewness	Max	Min
N	587.3	1347	14	175	154.81	109	8.83	18887	4
G	24710	61513	5	1580	1463.13	38	5.38	614145	3
g'	7757	19469	0	741		25.82	4.63	162478	0
TE	32936	75380	16	2579	2477	36.24	5.02	776623	4
SL	168.3	560.6	1.00	25.00		99.02	8.74	7497	0
RVPR	0.11	0.10	0.00	0.07		4.45	1.94	0.57	0
RER	0.18	0.14	0.00	0.14		1.83	1.37	0.73	0
CC	8.40	99.31	1.00	1.00	1.64	319.11	17.83	1781	1
SVCC	1.17	4.40	0.00	0.00		75.41	7.88	52.00	0
MVCC	866	6042.6	4.00	174	154	307.41	17.36	107641	4
MECC	27289	545452	15	2355	2098	17.20	3.45	469890	4
MGD	3.41	1.38	4.00	4.00	3.13	1.72	0.84	10.00	1
AGD	1.84	0.38	1.00	1.88	1.79	0.90	-0.19	3.17	0.75
Dens	0.12	0.12	0.25	0.09	0.08	10.87	2.65	1.00	0
IIDA	-0.04	0.19	-0.08	-0.07		1.88	0.67	0.68	-0.71
IODA	-0.14	0.14	0.00	-0.12		3.12	-0.33	0.55	-0.71
OIDA	-0.26	0.16	-0.50	-0.24		2.25	0.20	0.62	-0.71
OODA	-0.03	0.26	0.00	-0.07		-0.02	0.27	0.61	-0.71
AD	52.43	79.02	5.00	25.40	23.29	58.19	5.80	973.87	0.4
Bet	0.24	0.15	0.00	0.23		-0.59	0.36	0.67	0
Mod	0.29	0.15	0.00	0.28		0.81	0.57	0.92	-0.01
Clust	13.98	99.26	3.00	6.00	6.28	317.57	17.75	1785	1.00
WSCC	0.51	0.08	0.50	0.51		11.09	-1.60	0.78	0
NCC	0.31	0.13	0.50	0.29		0.15	0.38	0.75	0

NACC	0.83	0.19	0.00	0.90	8.65	-2.80	1.00	0
***M* (*pM*)**	0.26	0.16	0.25	0.24	0.67	0.82	0.82	0
***F* (pF)**	0.68	0.19	0.50	0.71	0.93	-0.94	1.00	0
***B* (*pB*)**	0.04	0.07	0.00	0.01	20.24	4.04	0.57	0
***I* (*pI*)**	0.00	0.04	0.00	0.00	204.85	14.00	0.62	0
***O* (*pO*)**	0.03	0.07	0.00	0.00	33.21	5.20	0.67	0

See definitions and descriptions in Table 2 in the Appendix.

We identified the following sectors among fan page networks with common socioeconomic drivers: Restaurant, Bakeries, Ice cream shops, Bars and Pubs, Pizza shops, Schools, Bookstores, Universities, Institutions, Theaters, Gyms, Hotels, Drugstores, Wardrobes, Workshops, Entertainment, Sport clubs, Butcheries, Hardware stores, Petty shops, Supermarkets, Professional services, Music stores, News, Radios, Real estate, Discos, Barber shops, and Churches. This way, each of the detected network belongs to one of these categories. Figure 2 depicts the sample composition of the FPN.

For these predefined clusters we estimate structural metrics and compare them for obtaining a taxonomy of relevant proprieties. They all share internal similarities and external dissimilarities that allow to easily identify how clusters are formed.

Figure 2. FPN type composition

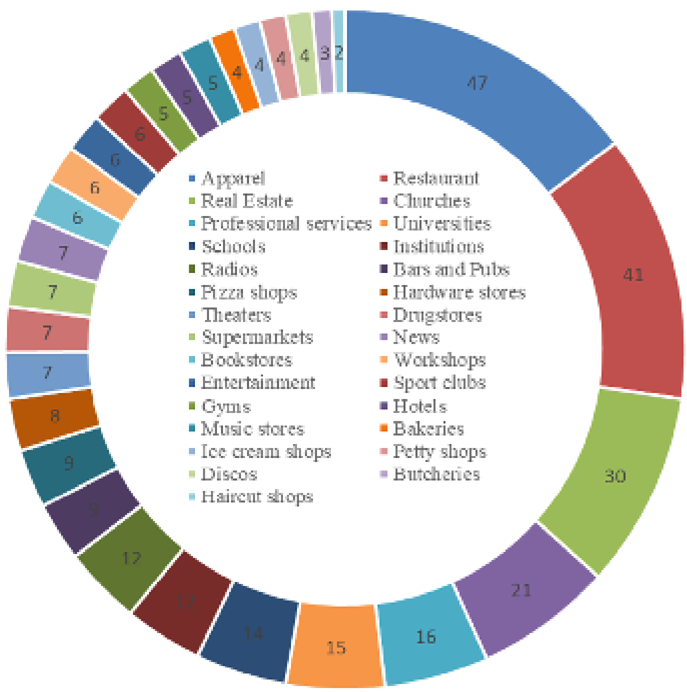
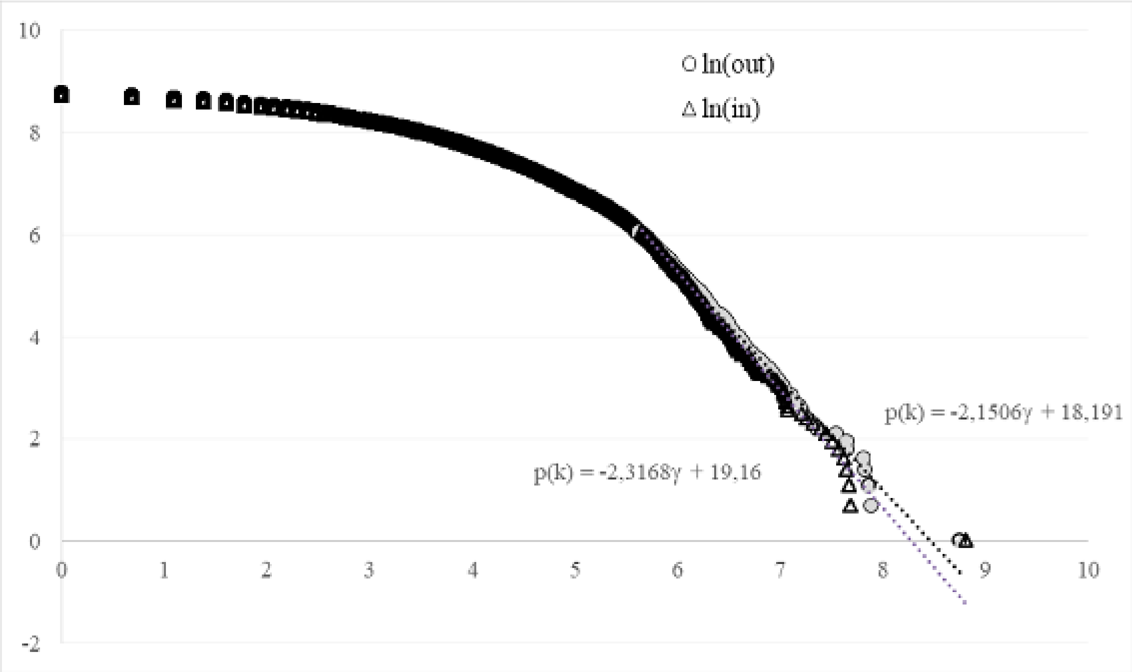


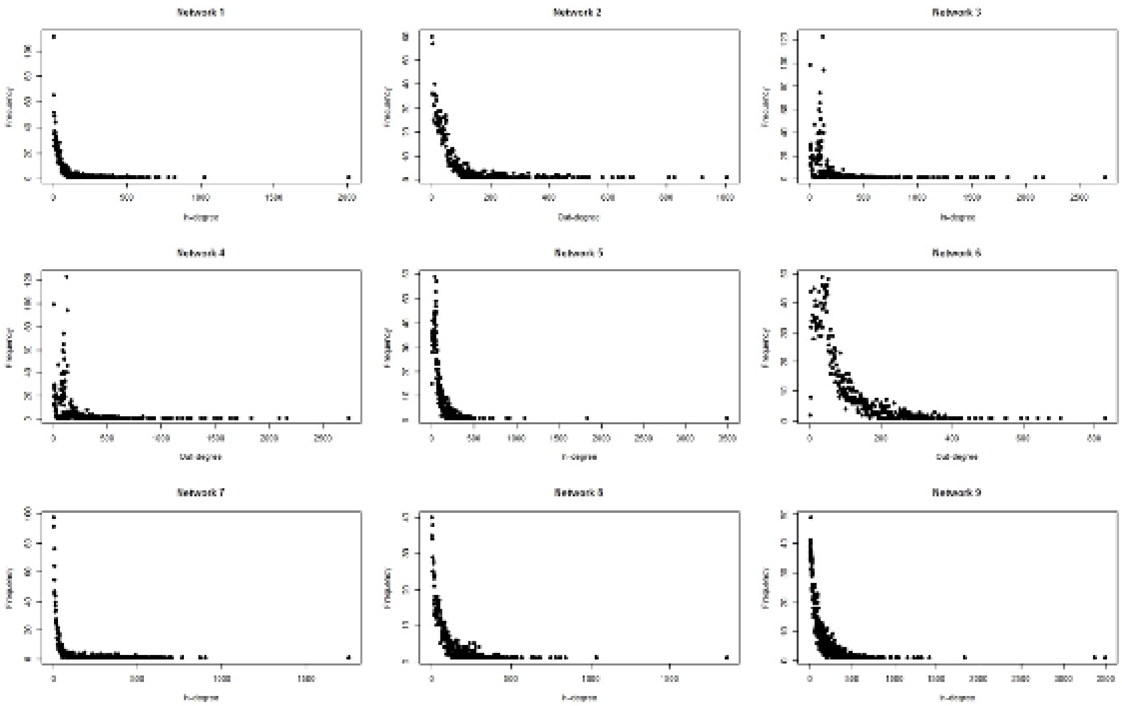
Figure 3 presents the in-degree -In(in)- and out-degree -In(out)- distributions of the complete FPN. When observing Figure 2 we detect a classical pattern of free scale network degree distribution. The γ exponent presents values of 2.3168 and 2.1506, respectively. So both distributions present evidence of free scale architecture at the highest level of aggregation.

Figure 3. Out-degree and in-degree frequency distribution and log-log representation inserted



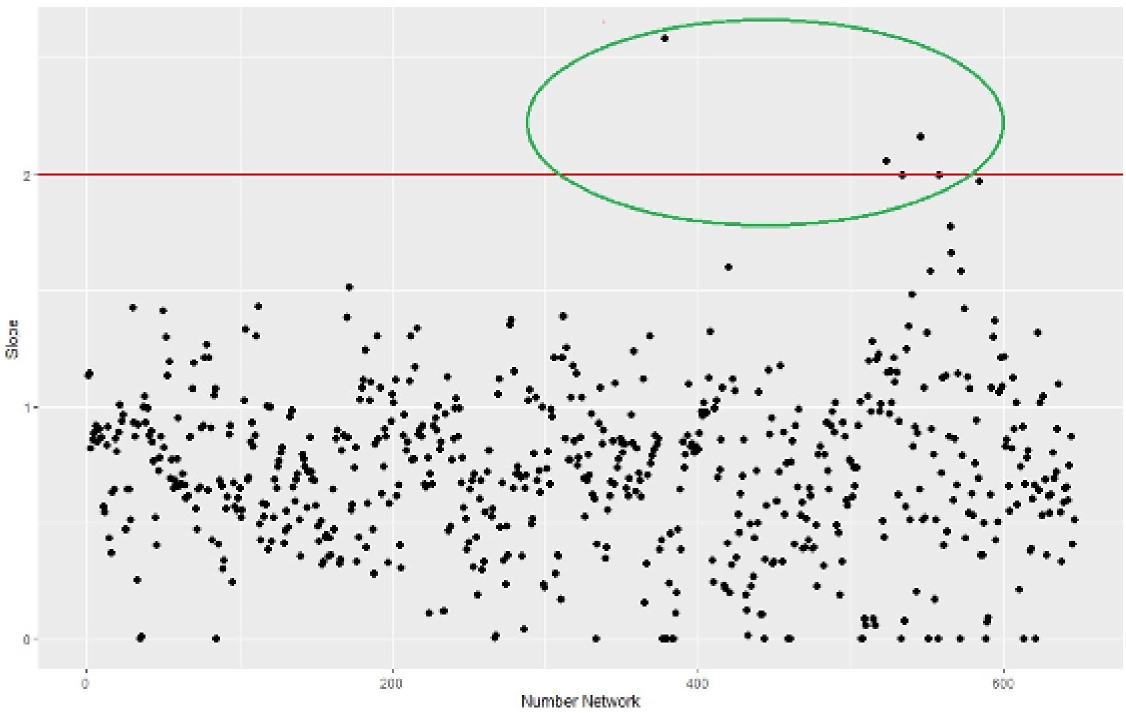
We also find intriguing initial evidence of potential free scale presence in diverse individual networks as presented in the Figure 4. Pictures depict degree frequency distributions of individual selected networks. At a glance, they seem to show evidence of the presence of power law in degree distribution but when estimating the slope of the log-log representation the different values of γ were lower than expected values. We test all subnetworks adjusting a linear regression to the slope of the log-log representation. Values of γ are presented in Figure 5.

Figure 4. Examples of degree distributions of selected networks



On the other hand, a few cases when we detected a γ value corresponding a scale-free topology are only present in very small subnetworks. One possible explanation for these results is linked to in these networks exist only an engaged node who post continuously but he receives a sporadic answers from isolated nodes. Values are presented in Figure where in an ellipse are remarked those few subnetworks that show slope close to free scale degree frequency distribution ($\gamma \geq 2$). However, these subnetworks deserve particular clarification. They are 5 subnetworks with a handful of nodes and only one has more than a thousand nodes. They belong to the Real Estate and Mechanical cluster of subnetworks with almost negligible engagement. They scale free networks because they have the majority of non-engaged participants, passive to the engagement of the fan page owner. So, as star networks, they show few highly connected nodes to a huge amount of poorly connected counterparts.

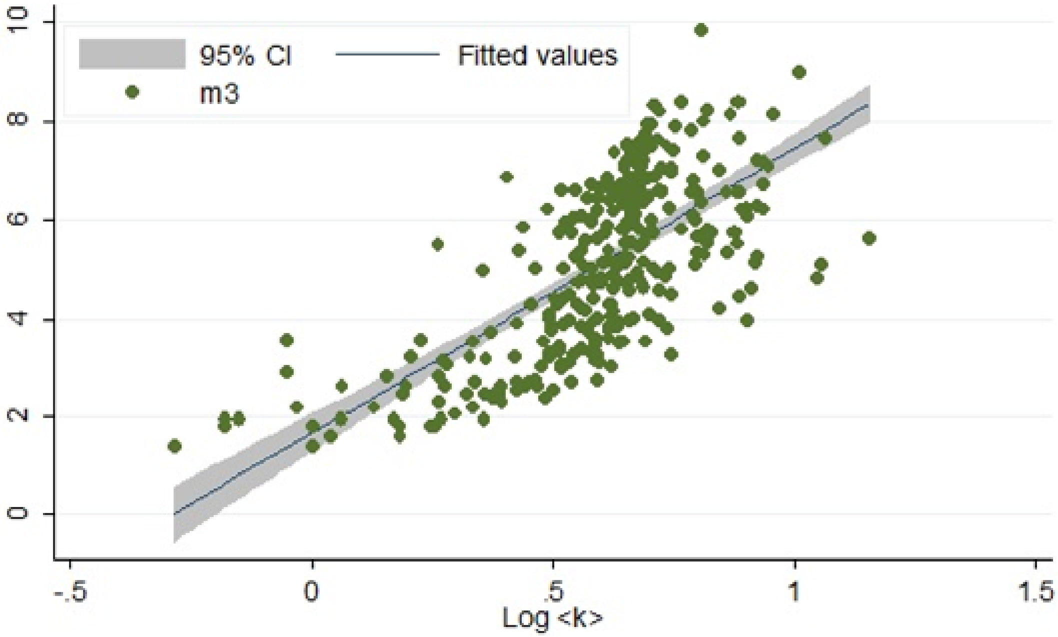
Figure 5. Scatter plot of γ for in-degree and out-degree frequency distributions of all subnetworks (encircled free scale subnetworks)



3.1 Small world

We focus in the degree distributions by estimating how close $\langle k \rangle$ approximates to ln N. By estimating the linear regression between *k* and ln *N* we check who data points lie along the best fit curve. The information is presented in the following Figure 6 where it can be pointed out that values in the confidence interval might be interpreted as statistically approximating the values of *k* to the ln *N*, then verifying network that approximates the small world phenomenon. In any case, the correlation between both variables is significant, i.e. values of *k* tend to be in the ln *N* neighborhood.

Figure 6. Fitted values and confidence interval of ln *N* and <k>



The potential explanation for this observation is that subnetwork in FPN tend to attract highly engaged people and participants. That's gives the more engaged nodes are not that distant from low engaged ones. Interaction is promoted via advertising and promotions as well as intentional FPN administrator policies for incrementing participation. Scale free architectures require, on the other hand, an enormous difference between more engaged hubs to low engaged marginal nodes and that is something hardly observed in a 2.0 environment.

4. Conclusions

We explore 325 FPB from a sample geographically located in Argentina summarizing thousands of nodes and relationships. We initially explore, as demonstrated by many other contributions, the topologies expecting to find free scale degree distributions as the more present topology in the sample. Instead, we detected only a handful of FPN with such topologies (including the in-degree and out-degree distribution of the complete FPB sample). A posterior analysis of the free scale networks show that these subnetworks were marginal in terms of the values of their metrics. However, by studying structural metrics of the subnetworks we find patterns that associate topologies to small world phenomenon. Specifically, average geodesic distance is positively correlated to the logarithm of the network range.

This way, we support the findings of Wohlgemuth and Matache (2014) and Caci et al. (2012), where Facebook networks (in different variants such as FPN or groups) exhibit small world properties and, in some particular cases or at the highest aggregate level, exhibit also power law distributions in degree frequencies. Again, for highly engaged actors in the network, it is difficult for a specific to scale far away from the rest in a sizes that would represent a scale free topology.

However, we need recognize the important temporal constraint: it is almost impossible to collect data after the 2018 Cambridge Analytics scandal. This event affects the possibility of continue this research along the time with the very same quality of data.

References

Backstrom, L., P. Boldi, M. Rosa, J. Ugander, and S. Vigna (2011), "Four Degrees of Separation." arxiv.org/abs/1111.4570

Barabási, A.-L. and R. Albert (1999), "Emergence of Scaling in Random Networks", *Science* 286 (5439): 509-512. DOI: 10.1126/science.286.5439.509

Barabási, A.-L. and M. Pósfai (2016), *Network Science*, Cambridge University Press.

Bastian M., S. Heymann, and M. Jacomy (2009). "Gephi: an open source software for exploring and manipulating networks". International AAAI Conference on Weblogs and Social Media.

Batagelj, V. and A. Mrvar (1998), "Pajek—a program for large network analysis". *Connections* 21(2): 47–57.

Blackmore, S. (2001), "Evolution and memes: the human brain as a selective imitation device". *Cybernetic Systems* 32:225–255

Brooks, B., B. Hogan, N. Ellison, C. Lampe, and J. Vitak (2014), "Assesing structural correlates to social capital in Facebook ego", *Social Networks* 38: 1-15. https://doi.org/10.1016/j.socnet.2014.01.002

Broido, A. and A. Clauset (2019), "Scale-free networks are rare", *Nature Communications* 10 (1017). https://www.nature.com/articles/s41467-019-08746-5

Caci, B., M. Cardaci, and M.E. Tabacchi (2012), "Facebook as a Small World: a topological hypothesis", *Social Network Analysis and Mining* 2: 163-167. DOI 10.1007/s13278-011-0042-8

Clauset, A., C. Shalizi, and M. Newman (2009). "Power-Law Distributions in Empirical Data". *SIAM Review*, 51(4): 661-703. Retrieved from http://www.jstor.org/stable/25662336

Ebadi, A. and A. Schiffauerova (2015), "How to Receive More Funding for Your Research? Get Connected to the Right People!", *PLOS One*. https://doi.org/10.1371/journal.pone.0133061

Eikmeier, N. and D.F. Gleich (2017), "Revisiting Power-law Distributions in Spectra of Real World Networks", *KDD '17 Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*: 817-826. DOI: 10.1145/3097983.3098128

Goyal, S., M.J. van der Leij and J.L. Moraga-Gonzalez (2006), "Economics: An Emerging Small World," *Journal of Political Economy*, University of Chicago Press, vol. 114(2): 403-432.

Holme, P. (2019), "Rare and everywhere: Perspectives on scale-free networks", *Nature Communications* 10, 1016. Retrieved from https://www.nature.com/articles/s41467-019-09038-8

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rendón de la Torre, S., L. Kalda, R. Kitt, and J. Engelbrecht, (2019) “Detecting overlapping community structure: Estonian network of payments”, *Proceedings of the Estonian Academy of Sciences* 68(1):79. DOI: 10.3176/proc.2019.1.08

Sadri, A.M., S. Hasan, S.V. Ukkusuri and M. Cebrian (2017), “Understanding Information Spreading in Social Media during Hurricane Sandy: User Activity and Network Properties”, https://arxiv.org/abs/1706.03019

Schneeberger, A, Mercer CH, Gregson SA, Ferguson NM, Nyamukapa CA, Anderson RM, Johnson AM, Garnett GP (2004), “Scale-free networks and sexually transmitted diseases: a description of observed patterns of sexual contacts in Britain and Zimbabwe”. *Sexually Transmitted Diseases* 31 (6): 380-7 DOI: 10.1097/00007435-200406000-00012

Slattery, R.E., R.R. McHardy, and R. Bairathi (2013), “On the Topology of the Facebook Page Network”, Proceedings from Engagement Science: Theory and Methods: Insights from the Emergence API. ArXiv: 1307.02.2189v1

Smith, M., Ceni A., Milic-Frayling, N., Shneiderman, B., Mendes Rodrigues, E., Leskovec, J., Dunne, C., (2010). *NodeXL: a free and open network overview, discovery and exploration add-in for Excel 2007/2010/2013/2016*, from the Social Media Research Foundation: https://www.smrfoundation.org

Ugander, J., B. Karrer, L. Backstrom, and C. Marlow (2011), “The Anatomy of the Facebook Social Graph.” arxiv.org/abs/1111.4503.

Watts, D.J. and S.H. Strogatz (1998). “Collective dynamics of 'small-world' networks.” *Nature* 393 (6684): 440-2. doi:10.1038/30918.

Wickham, H. (2016), *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wohlgemuth, J. and M. T. Matache (2014), “Small-world properties of Facebook group networks”. *Complex Systems* 23(3): 197-225.

Zipf, G. (1936), *The Psychobiology of Language*. London: Routledge;

Appendix

Metrics were calculated by using freely available software such as Pajek (Batagelj and Mrvar, 1998), NodeXL Basic (Smith et al., 2010) and Gephi (Bastian et al., 2009). We also use R packages for graphics (R Core Team, 2019; Wickham, 2016).

Table 2. Codes, names and definitions

Code	Metric and variable	Description
N	Vertices	Range of the graph
G	Unique Edges	Non-repeated edges
g^l	Edges With Duplicates	Repeated interaction edges
TE	Total Edges	Sum of two previous edge classification
SL	Self-Loops	Edges connecting vertices with themselves.
RVPR	Reciprocated Vertex Pair Ratio	The number of vertex pairs that have edges in both directions divided by the number of vertex pairs that are connected by any edge.
RER	Reciprocated Edge Ratio	The number of edges that are reciprocated divided by the total number of edges.
CC	Connected Components	The number of connected components in the graph
SVCC	Single-Vertex Connected Components	The number of connected components that have only one vertex.
MVCC	Maximum Vertices in a Connected Component	The number of vertices in the connected component that has the most vertices.
MECC	Maximum Edges in a Connected Component	The number of edges in the connected component that has the most edges.
MGD	Maximum Geodesic Distance (Diameter)	Largest path between more distant pair of vertices
AGD	Average Geodesic Distance	Average path between every pair of vertices
Dens	Graph Density	This is a ratio that compares the number of edges in the graph with the maximum number of edges the graph would have if all the vertices were connected to each other.
IIDA	Input-Input Degree Assortativity	Indicates the correlation between node in-degree to its connections in-degree.
IODA	Input-Output Degree Assortativity	Indicates the correlation between node in-degree to its connections out-degree.
OIDA	Output-Input Degree Assortativity	Indicates the correlation between node out-degree

		to its connections in-degree.
OODA	Output-Output Degree Assortativity	Indicates the correlation between node out-degree to its connections out-degree.
AD	Average degree	Nodes' weighted average degree
Bet	Network betweenness centralization	Nodes' weighted average betweenness
Mod	Modularity	Graph modularity
Clust	Clusters	Based on modularity, the number of identified clusters
WSCC	Watts-Strogatz Clustering Coefficient	Clustering coefficient based on Watts-Strogatz model
NCC (T)	Network Clustering Coefficient (Transitivity)	Clustering coefficient based on transitivity
NACC	Network All Closeness Centralization	Closeness centralization
M (*pM*)	Male (Percentage of Male)	Number (percentage) of males
F (pF)	Female (Percentage of Female)	Number (percentage) of females
B (*pB*)	Business (Percentage of Business)	Number (percentage) of business
I (*pI*)	Institution (Percentage of Institution)	Number (percentage) of institutions
O (*pO*)	Other (Percentage of Other)	Number (percentage) of other nodes