Feature Characterization in Some Weighted Directed Networks

Pedro J. ZUFIRIA[1,*] and Colleen MCMILLON[2]

[1]Information Processing and Telecommunications Center (IPTC), ETSI Telecomunicación, Universidad Politécnica de Madrid, Spain

[2]Treelogic, Pozuelo de Alarcón, Spain

*Corresponding author

Abstract. The paper studies the time evolution of weighted directed networks by evaluating some of their fundamental features along time. The main concepts are illustrated by analyzing the evolution of a network which represents the call flow between the towers of a telephone operator in Senegal. Different time scales (hour, day, week, month) are considered to compute the evolution of features in order to characterize properties which may follow cyclic behaviors. A joint characterization of the features is also performed via a paired correlation analysis which also shows to be sensitive to time scales, suggesting different seasonal behaviors in the communication patterns.

Introduction

In the last two decades, many complex systems have being studied using network models [1,2,3], since these models allow for the characterization of some structural properties which become fundamental for understanding the behavior of the systems. These similar structural properties usually appear in social networks [4], biology networks [5], technological networks [6] or information networks [7]. Since these networks are frequently very large, some of their features must be analyzed from a statistical perspective; in this framework, random graph models have been successfully employed for characterizing such networks.

Starting from the seminal Erdós-Rényi (ER) model [8] which served as a baseline (i.e., the simplest null model) for comparative purposes, several more sophisticated models have been proposed to explain the behavior of complex networks (e.g., [3,9]). All these models can be put into the formal framework of a probability space in the so called Random Network Models (RNMs) [10]. This formulation allows for a rigorous and unified approach to the problem of analyzing the properties of different network models [11,12]. In this paper, we illustrate the use of RNMs to analyze the relationship between several features in some weighted directed networks which evolve in time.

The paper is organized as follows. Section 2 presents the basic concepts related to Random Network Models and the analysis of features or properties defined on networks. Section 3 illustrates the main ideas on a real network obtained from the communication flows between phone towers during a given time period in Senegal. Finally, concluding remarks are presented in Section 4.

Random Network Models. Time Evolution and Joint Characterization of Features

Many complex networks can be modeled in a probabilistic framework which gathers some of their fundamental properties. Given a Random Network Model (RNM) [10,11], any quantifiable property or feature f that can be defined on a sample network (such as the number of links, diameter, number of triangles, connectivity, assortativity, reciprocity, closeness centrality, betweenness centrality, node degrees, etc.) can be characterized via a corresponding random variable F as defined in [11]. Such random variable or vector F is denoted as *Random Network Model Feature*, and provides some (indirect) information concerning the structure and properties of the RNM.

The evolution of networks can be characterized via the corresponding analysis of the associated time evolving RNMs or Random Network Processes (RNPs) [10,12]. At the same time, RNPs can be studied by analyzing the time evolution of some of the above mentioned features. The joint characterization of several time series of features becomes a challenge specially when the networks we have to analyze may be sensitive to the time scale employed in the analysis.

Analysis of a Real Network

In this section we illustrate the above mentioned analysis procedure by characterizing some weighted directed networks derived from the data provided in the second Orange D4D Challenge [13]. Precisely, the Challenge Data Set 1 provided the total flows of both sms and telephone calls during every hour of the year 2013, between 1666 towers distributed all over Senegal. Such Data Set allows the straightforward construction of several weighted directed networks by simply aggregating the desired flow (sms or calls) over a selected period of time.

Since the networks that can be constructed every hour were too many, as a first approach only the evolution along a month (April) was considered for such fine hourly resolution. The rest of network sequences were constructed aggregating the flows at daily, weekly and monthly scales, and they were considered over the whole year range. For each network a set of 8 features was computed: density (related to number and weight of links), diameter, transitivity (related to clustering coefficient), assortativity, reciprocity, closeness centrality, betweenness centrality, degree distribution. Note that some vector features are difficult to analyze and represent in a plot (specially along time); hence, the entropy of such vector features becomes an appropriate scalar summary or simplification easy to handle and represent, still preserving some useful information [11].

Figures 1, 2 and 3 represent the time evolution of the density, diameter, transitivity, assortativity, reciprocity, closeness entropy, betweenness entropy and degree distribution entropy at monthly, weekly and daily scales respectively. The comparison among the figures illustrates how the comparative time evolution of the features strongly depends on the selected time scale. This result suggests the existence of different types of seasonal behaviors in the communication patterns.
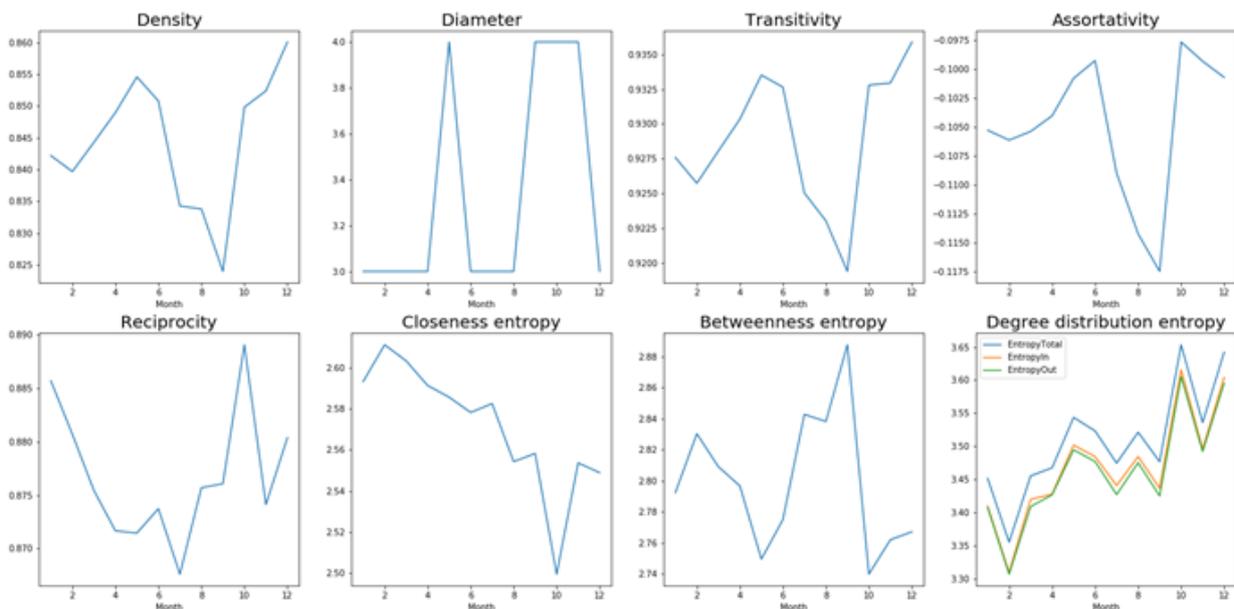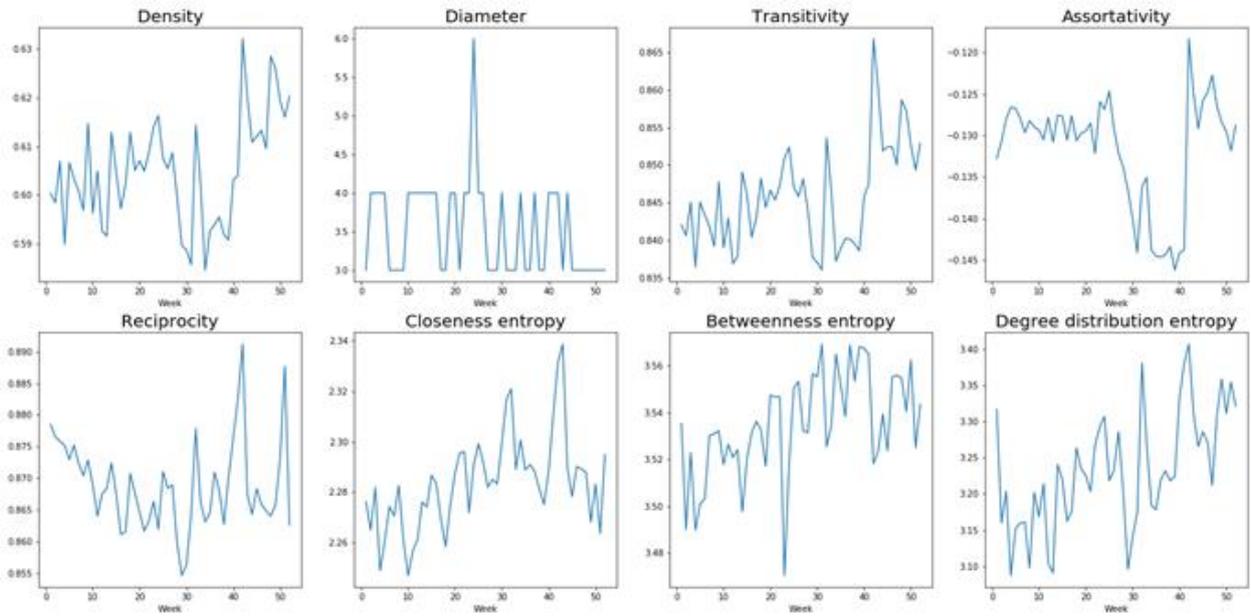

Figure 1, Monthly evolution of features,

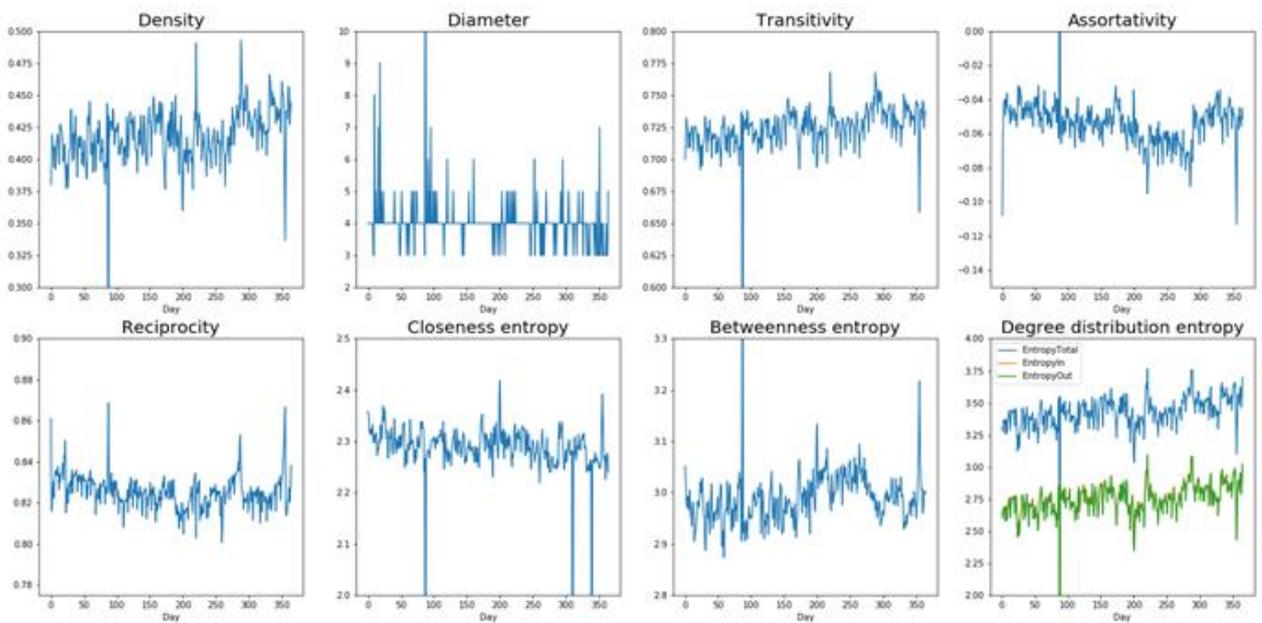Figure 2. Weekly evolution of features.

Figure 3. Daily evolution of features.

Figure 4 represents the paired correlations between the density, diameter, transitivity, assortativity, reciprocity, closeness entropy, betweenness entropy and degree distribution entropy at monthly, weekly, daily and hourly scales.
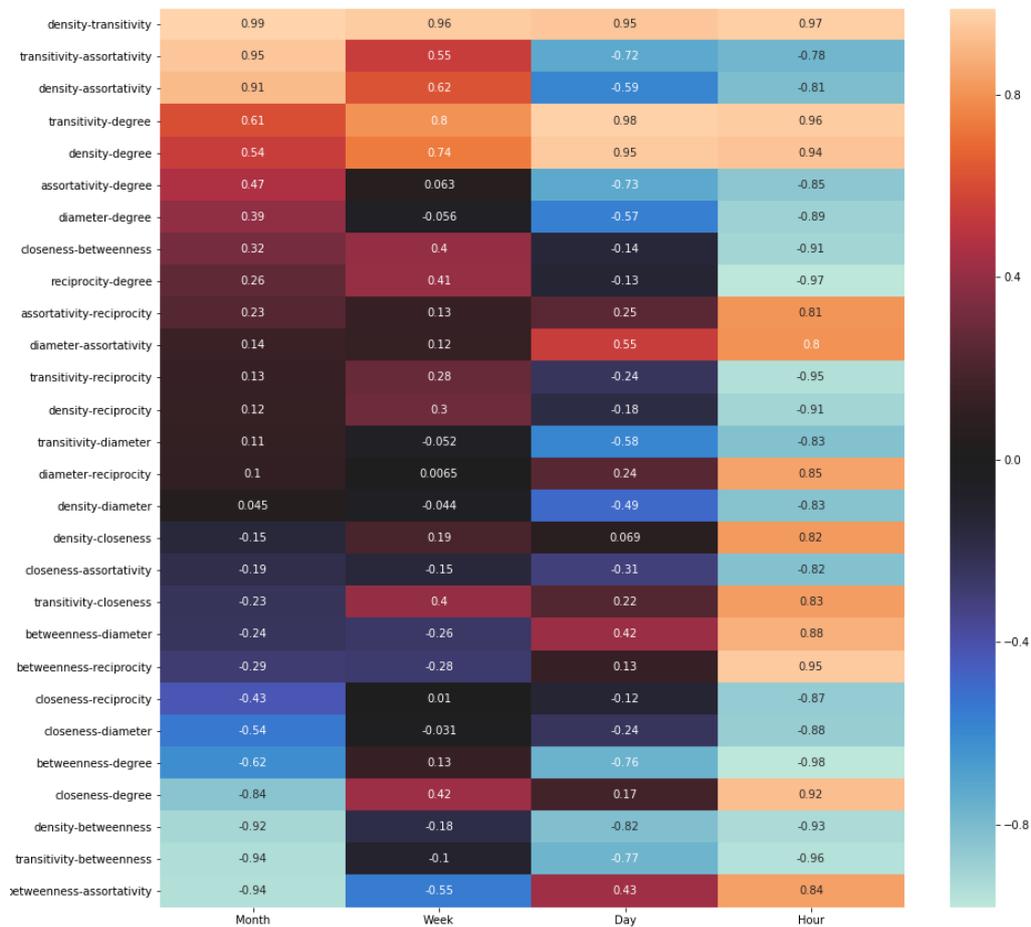
Figure 4. Correlations between features at montly, weekly, daily and hourly scales.

The correlations between the different properties (e.g., transitivity-assortativity, closeness-degree) do change in a very drastic way when we move along the different time scale resolutions. This fact suggests the existence of different seasonal behaviors in the country communication patterns.

Concluding Remarks and Further Research

The analysis of network features along time has proven to be useful to characterize the communications patterns, detecting the existence of seasonal behaviors due to the sensitivity to times scales. Further research is being performed to assess the relevance of the features and the appropriate time scales for detecting specific patterns. In addition, the relevance of some network properties is being tested by comparing them against some null models (i.e., RNMs based on minimum number or simplest assumptions).

Acknowledgement

References

[1] Albert-László Barabási and Jennifer Frangos. *Linked: the new science of networks science of networks*. Basic Books, 2014.

[2] Mark JE Newman. Networks. Oxford University Press, 2018.

[3] Duncan J Watts and Steven H Strogatz. Collective dynamics of small world networks. *Nature*, 393(6684):440–442, 1998.

[4] J-P Onnela, Jari Saramäki, Jorkki Hyvönen, Gy˙orgy Szabó, David Lazer, Kimmo Kaski, János Kertész, and A-L Barabási. Structure and tie strengths in mobile communication networks. *Proceedings of the National Academy of Sciences*, 104(18):7332–7336, 2007.

[5] Hawoong Jeong, áalint Tombor, Réka Albert, Zoltan N Oltvai, and A-L Barabási. The large-scale organization of metabolic networks. *Nature*, 407(6804):651–654, 2000.

[6] Michalis Faloutsos, Petros Faloutsos, and Christos Faloutsos. On powerlaw relationships of the internet topology. In *ACM SIGCOMM Computer Communication Review*, volume 29, pages 251–262. ACM, 1999.

[7] Sidney Redner. How popular is your paper? an empirical study of the citation distribution. *The European Physical Journal B-Condensed Matter and Complex Systems*, 4(2):131–134, 1998.

[8] Paul Erdös and Alfréd Rényi. On random graphs. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.

[9] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

[10] Pedro J. Zufiria and Iker Barriales-Valbuena. Evolution models for dynamic networks. In 38th International Conference on Telecommunications and Signal Processing (TSP), pages 252–256. IEEE, 2015.

[11] Pedro J. Zufiria and Iker Barriales-Valbuena. Entropy characterization of random network models. Entropy, 19(7):321, 2017.

[12] Pedro J. Zufiria and Iker Barriales-Valbuena. Analysis of basic features in dynamic network models. Entropy, 20(9):681, 2018.

[13] de Montjoye YA, Smoreda Z, Trinquart R, Ziemlicki C, Blondel VD. D4D-Senegal: the second mobile phone data for development challenge. arXiv preprint arXiv:14074885. 2014.