Physica A 515 (2019) 641–648

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Connecting network science and information theory

Henrique F. de Arruda [a,*], Filipi N. Silva [b,c], Cesar H. Comin [b], Diego R. Amancio [a,c], Luciano da F. Costa [b]

[a] *Institute of Mathematics and Computer Science, University of São Paulo, São Carlos, SP, Brazil*
[b] *São Carlos Institute of Physics, University of São Paulo, São Carlos, SP, Brazil*
[c] *School of Informatics, Computing, and Engineering, Indiana University, Bloomington, IN, USA*

ARTICLE INFO

Article history:
Received 13 July 2018
Received in revised form 28 September 2018
Available online 9 October 2018

Keywords:
Information theory
Complex networks
Compression
Data compression
Random walks
And network dynamics

ABSTRACT

A framework integrating information theory and network science is proposed. By incorporating and integrating concepts such as complexity, coding, topological projections and network dynamics, the proposed network-based framework paves the way not only to extending traditional information science, but also to modeling, characterizing and analyzing a broad class of real-world problems, from language communication to DNA coding. Basically, an original network is supposed to be transmitted, with or without compaction, through a sequence of symbols or time-series obtained by sampling its topology by some network dynamics, such as random walks. We show that the degree of compression is ultimately related to the ability to predict the frequency of symbols based on the topology of the original network and the adopted dynamics. The potential of the proposed approach is illustrated with respect to the efficiency of transmitting several types of topologies by using a variety of random walks. Several interesting results are obtained, including the behavior of the Barabási–Albert model oscillating between high and low performance depending on the considered dynamics, and the distinct performances obtained for two geographical models.

1. Introduction

Much effort in science and technology has been focused on the study of information theory [1] and network science [2,3], two seemingly independent realms. In information theory, probabilities are assigned to symbols and used to derive important results, such as minimum bandwidth and minimal sampling rates. On the other hand, in *network science* [2], focus is given to understanding the intricate topology of complex networks, and its interaction with various types of dynamics. Interestingly, these two different perspectives – broadly related to time series compaction and studies of topology/dynamics complexity – can be shown to ultimately be closely interrelated. For instance, information theory has been used to define causal relationships between nodes [4,5], characterize networks according to their compressibility [6], define topological similarity [7], map time-series to networks [8], reveal community structure [9] quantify the diversity of ecological networks [10], and characterize network dynamics [11].

Because of the fundamental importance of information science in the development of a large number of fields – including communications [12], computing [13] and even neurosciences [14] – any extension of the concepts in that area has the potential for significant theoretical and applied impact. This is the main motivation and general objective of the present work, which is developed with special attention given to including the systems generating the time series commonly studied in

* Corresponding author.
E-mail address: h.f.arruda@gmail.com (H.F. de Arruda).

https://doi.org/10.1016/j.physa.2018.10.005
0378-4371/

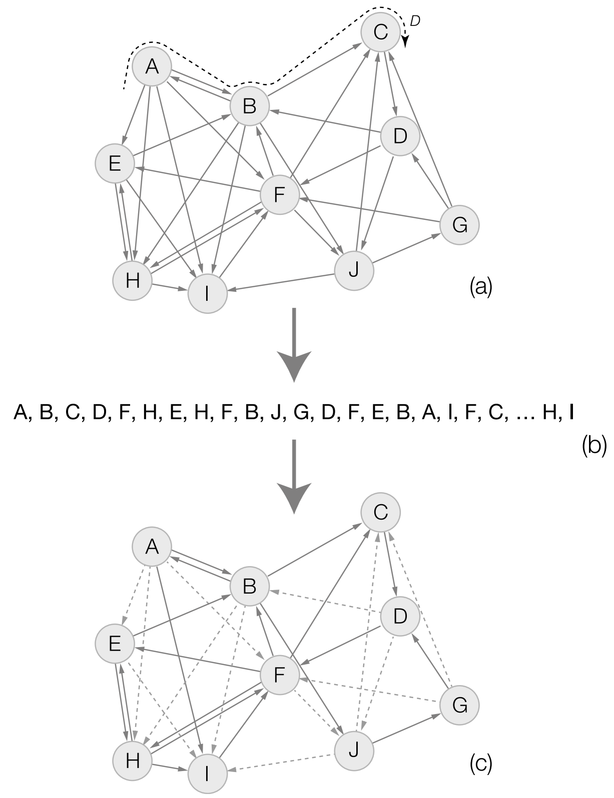

Fig. 1. Integrating network science and information theory. A network (a) is sampled by a given dynamics *D* and a respective sequence of symbols (b) is obtained and transmitted. The receiver can then try to recover the original network (c).

information theory into a more complete and integrated framework. More specifically, we represent the generating systems as complex networks, which immediately implies the critical question regarding how the topology of these generating networks influences the compaction of the produced time series. To study and better understand this interrelationship provides the main specific objective of the present work. The basic idea, illustrated in Fig. 1, is to understand sequences of symbols or time series as projections of an original network, e.g., obtained by some sampling dynamics (such as random walks), transmitted, and then reconstructed to some accuracy.

Underlying such an approach is the hypothesis that every time series or sequence of symbols is produced by some discrete system, which can be represented as a complex network. So, these generated series and sequences inherit, to a great extent, the properties of the generating networks. The second underlying hypothesis is that the interaction between such complex network systems takes place through communication channels, which necessarily have limited bandwidth. So, it becomes important to use methods for more effective/robust transmission of the networks, such as by using compression. At the same time, these sequences are a byproduct of the interaction between the topology of the original networks and the unfolding dynamics. Therefore, the proposed approach reinstates information theory from looking only at the time series to the network level.

Several important problems can be naturally conceptualized and represented according to the proposed framework. These include: the transformation from thoughts into language, encompassing the whole of literature in the process, arts, routing in transportation and computing systems, distributed computing, teaching and planning of syllabuses, economics and financial indices, genes and proteins, scientific modeling, and even WWW surfing and the flow of consciousness. An interesting aspect shared by all such cases is that the linearization of a higher dimensional structure (a network) derives from imposed constraints such as finite bandwidth channel, storage systems, etc. By the way, intermediate representations with dimensions higher than one (time series) are also possible and naturally incorporated in the proposed theory. In addition, the efficiency of coding by projections and respective transmission are related to the *complexity* of the original network, therefore emphasizing another critical issue shared by information theory and network science.

The basic construction in Fig. 1 can be immediately used to model more complex systems, such as opinion dynamics, as illustrated in Fig. 2. Here, each of the large nodes in Fig. 2 correspond to an agent transmitting its beliefs, originally represented by the networks inside the nodes, through time series along a network with a particular topology.

Optimization of transmission can be achieved by compacting the time-series. In information theory, this is typically achieved by using the frequency of symbols as a means to derive optimal code words. For instance, Huffman coding [15]

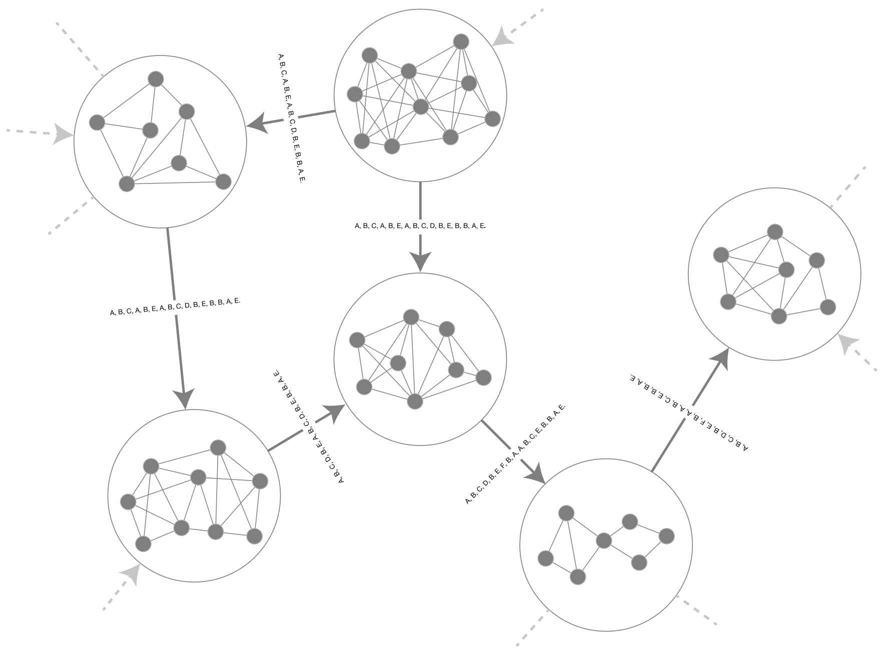

Fig. 2. The integration of network science and information theory can be used to address systems involving several interactions of the type represented in Fig. 1, with applications in ares such as opinion spreading.

provides a means to achieve lossless, optimal symbol-by-symbol coding of time series, considering the probability of each symbol. In the herein proposed framework, optimization involves accurate prediction of symbols by considering the topology of the network and the probing dynamics, instead of only the sequence of symbols. For instance, in the case of symbols taken individually from an undirected graph by using a traditional random walk, it is known that the frequency of each symbol can be fully predicted from the respective node degrees [16]. The precision of such a prediction can be expressed in terms of the Pearson correlation coefficient between the frequency of visits and degrees (or other topological property). Because of its ability to characterize how much the dynamics is affected by the topological features of the network, such a measurement is henceforth called *steering coefficient* S of the topology over the dynamics. All in all, the coding efficiency, given a specific network and dynamics, will probably depend on the value of the steering coefficient, which can vary among different network topologies and dynamics.

The present work illustrates and explores the potential of the proposed framework with respect to synthetic networks, allowing the consideration of several topologies and different sampling dynamics. So, we can investigate interesting questions such as: (i) how does the symbol prediction from graphs impact the transmission?; (ii) how do the topological features of different graph models affect the performance?; (iii) how do the performances of the considered types of dynamics compare one another?

This paper is organized as follows. Section 2 presents the proposed framework. Section 3 presents the obtained results and discussions. Finally, Section 4 provides the conclusions and perspectives for future studies.

2. Methodology

Here we proposed a framework for understanding time series as being generated by some dynamics taking place in a complex network. In such an approach, the time series can be used as a one-dimensional projection of the network that can be used to transmit information stored on it. Such a time series can also be used to reconstruct a model of the original network. This procedure is illustrated in Fig. 1. Note that previous studies employed random walk dynamics to generate time series to infer structural properties of networks [17,18].

Our methodology starts by obtaining time series of symbols for each considered network. More specifically, the topology of each network is probed according to each of the considered random walks dynamics. A respective time series comprising the sequence of visited nodes is obtained and, at each time instant, a reconstruction of the original network is generated (as illustrated in Fig. 1(a) and (b)). It can be easily achieved by starting with a disconnected set of symbols (nodes) and adding each new received edge, defined by a pair of following symbols, into this network. This reconstruction depends on: (i) the time series size; (ii) the considered dynamics; and (iii) the network topology.

An important issue regards how effectively each of the network models can be recovered from a respectively generated time series, in the presence or not of compression. In particular, the problem of measuring the transmission efficiency of networks is described in the following section (Section 2.1). This is followed by the description of the adopted complex network models (Section 2.2), random walk dynamics (Section 2.3), and employed compression technique (Section 2.4).

2.1. Measuring the efficiency of transmission

The efficiency of the transmission can be quantified in terms of the time series length (number of steps) required for reconstruction of 90% of the original network (measured in terms of number of edges). This critical time is henceforth referred to as T_{90}. The effect of compressing the time series by using the frequency of visits to nodes predicted by the respective degrees, referred to as T_{90}^C, is also considered, giving rise to another series of experiments. Furthermore, we also calculated the long term transmission ratio, defined as

$$R_L = \frac{T_L^C}{T_L}, \tag{1}$$

where T_L is a sufficiently large time (a total of one million symbols was used). In principle, a combination of topology and dynamics that allows large compression ratio should lead to faster reconstruction.

In order to quantify the efficiency of a given dynamics to characterize how much the dynamics are affected by the topological features, we compute the Pearson coefficient between the topological and dynamical measurements. More specifically, we considered the frequencies of visits and the network node degrees, and this measurement is henceforth called *steering coefficient S*. Furthermore, steering coefficient is employed in a twofold manner, by considering the steering coefficient attained after exploring 90% of the network nodes (S_{90}), and the long-term steering coefficient values (S_L), which is computed after a sufficiently large time of exploration (a total of one million steps). Note that S_{90} and S_L are inherently related to T_{90} and T_L, respectively.

2.2. Adopted dynamics

In order to investigate the proposed framework, we adopted four random walk dynamics, used to generate respective sequences of symbols. The considered dynamics are: the random walk (RW) [19], a variation in which the transition probabilities are preferential to nodes with higher degree (RWD) [20], another variation in which the inverse of the node degree is considered (RWID) [20], and the true self-avoiding random walk (TSAW) [21,22].

In a traditional RW dynamics, the next node to be taken is selected uniformly among its neighbors. In a degree-biased random walk, the probability p_{ij} that the agent goes from node i to j depends on the degree of each neighbor. Here we consider a dependence of the form

$$p_{ij} = \frac{k_j^\alpha}{\sum_{l \in \Gamma_i} k_l^\alpha}, \tag{2}$$

where Γ_i is the set of nodes connected to node i. When $\alpha = 1$ we have the RWD dynamics, while $\alpha = -1$ results in the RWID case. Note that the RW dynamics is obtained when $\alpha = 0$. An interesting property of the RW dynamics is that, on undirected networks, as the generated sequence increases, the frequencies of visits to nodes – as inferred from the current number of times each node has been visited – become directly proportional to the degree [16]. For the degree-biased case, the steady state probabilities of a degree-biased random walk is [23]

$$P_i = \frac{\sum_{j \in \Gamma_i} k_j^\alpha k_i^\alpha}{\sum_h \sum_{l \in \Gamma_h} k_l^\alpha k_h^\alpha}. \tag{3}$$

If the degrees of the neighbors of node i can be approximated as the average degree of the network (which happens, for instance, for narrow degree distributions), the numerator of Eq. (3) can be written as $\sum_{j \in \Gamma_i} k_j^\alpha \approx \langle k \rangle^\alpha k_i$, leading to

$$P_i \approx \frac{\langle k \rangle^\alpha k_i k_i^\alpha}{\langle k \rangle^\alpha \sum_h k_h k_h^\alpha} = \frac{k_i^{\alpha+1}}{N \langle k^{\alpha+1} \rangle}. \tag{4}$$

This means that the probabilities become related to the degree taken to $\alpha + 1$.

In the TSAW, the memory of the path already taken by the agent is kept and considered for determining its next step. Here, we choose the frequency of edges in contrast to the frequency of nodes [22,24]. In this way, edges already visited many times are avoided. Thus, the probability p_{ij} that the agent moves through a certain edge e at its immediate neighborhood Γ_e is

$$p_{ij} = \frac{\gamma^{-f_e}}{\sum_{j \in \Gamma_e} \gamma^{-f_j}}, \tag{5}$$

where γ is a parameter of the dynamics and f_e is the frequency of visits to edge e. Note that self-avoiding behavior is achieved for $\gamma > 1$. The TSAW dynamics can be expected to be usually faster to cover a network, compared to the RW since unvisited connections are prioritized [22]. For large number of iterations, this dynamics tends to behave like diffusion, i.e. similarly to the RW dynamics [21]. We adopt $\gamma = 2$.

2.3. Complex network models

Six network models were used to investigate the proposed framework, namely the Erdős–Rényi (ER) [25], Barabási–Albert (BA) [26], Watts–Strogatz (WS) [27], Waxman (WAX) [28], random geometric (GEO) [29] and Knitted (KN) [30] models. The latter model is used only on experiments involving directed networks. The ER model generates small-world networks having a binomial degree distribution [2], meaning that all nodes in the network have similar degree. In contrast, networks generated by the BA model have a power-law degree distribution [26], implying some nodes, called hubs [3], possessing relatively large degree. The WS model can be used to generate networks having the small-world property, while also exhibiting large clustering coefficient values [27]. We adopt a variation where one starts with a lattice network and edges are rewired with probability p. We consider two rewiring probability values for the WS model: $p = 0.01$ (WS1) and $p = 0.005$ (WS2). In the GEO model, nodes are randomly placed, with uniform probability, in a two-dimensional space and pairs of nodes are connected if their distance is smaller than a given value. The WAX model begins with the same node placement procedure as in the GEO model, but pairs of nodes are connected according to a probability decaying exponentially with the distance between the nodes. Networks generated by the GEO and WAX models tend to have large diameter. KN networks are formed by treading paths. Initially, a set of unconnected nodes is created. Then, a sequence of distinct nodes is randomly selected and visited until a stop condition is met. Adjacent nodes in the sequence are connected through respective directed links. The process can be repeated many times until a desired average degree is reached. KN networks can be used for modeling co-occurrence networks in texts and other real-world situations involving sequential, uninterrupted visits to nodes [31–34].

All the above models, except KN networks, correspond to undirected networks. We also considered directed networks. A network can be made directed by assigning directions to the edges. First, we defined a reciprocity parameter r, which is the probability of a bidirectional edge. For each original edge, a random number n in the range [0, 1] was generated with uniform probability. If $n \leq r$, the original edge was split into two edges (in and out); otherwise, a single direction was randomly selected with equal probability. In addition, only the largest strongly connected component of the network was considered, so as to avoid the random walker to become trapped. We set $r = 0.6$ so as to preserve the original network size while accounting for substantial directionality.

2.4. Huffman algorithm

In digital media, a message can be encoded as a set of organized symbols, which are stored by using a fixed number of bits. For example, texts are formed of symbols which can be represented as characters of 8 bits. In order to store or transmit messages in an effective way, several lossless compression algorithms have been proposed in the literature [35].

The *Huffman code* is a particular data compression algorithm, based on information theory [1]. This method generates a dictionary of bit sequences employed to represent each symbol in a message. The compression is achieved by associating shorter bit sequences to more frequent symbols and longer sequences to symbols that appears more rarely in the message. To do so , the Huffman algorithm uses a binary tree, whose leaves represent symbols. Starting from the root, every edge is associated to a bit. Typically, left and right children are associated to the bits 0 and 1, respectively. The symbol code associated to each leaf node is then obtained by concatenating, from root to leaves, all edge values from the root. Such structure is used as a dictionary to encode the original message. In a similar fashion, the dictionary is used to decode the message.

3. Results and discussion

In this section we present the results of the transmission efficiency obtained for different models and random walk dynamics. We also compare the obtained results for the models with those obtained from text and word association networks.

3.1. Theoretical networks

For the analysis conducted for the network models, a total of 30 simulations were performed for each combination of model and dynamics. Furthermore, 30 network realizations were used for each network model. The considered networks had approximately 1000 nodes and average degree near 8.

Fig. 3 shows a parallel coordinates plot obtained for the undirected network models. The overall best transmission times were obtained for the TSAW dynamics. The BA topology implied a substantially higher value of T_{90} (number of steps required for reconstruction of 90% of the original network) for the RWD dynamics (Fig. 3(a)). This is probably a consequence of the fact that, in this dynamics, the moving agent tends to alternate between hubs, overlooking nodes with small degree. As shown in Fig. 3(b), similar results were obtained when considering compressed times, T_{90}^C, although in this case, the values of T_{90}^C for the RWD dynamics in the BA model are not as prominent as those obtained for T_{90}. The results for the WAX and GEO models differed significantly, with the former being transmitted much more effectively. This result is surprising because both these models share a geographical nature, in the sense that nodes that are spatially close one another tend to be connected.

Also shown in Fig. 3 is the steering coefficient (S_{90}). Similar values were obtained for the RW, RWD and TSAW dynamics. In the RWID case, the degree is not a good predictor of the frequency of visits, as indicated in Eq. (4), which leads to low

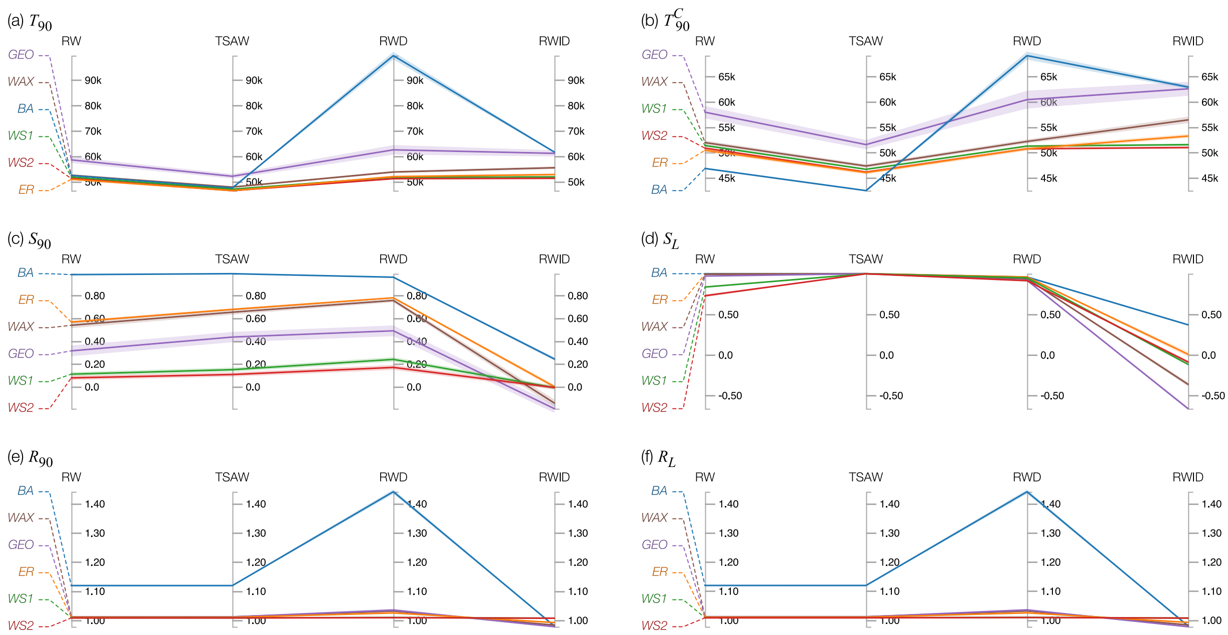

Fig. 3. Transmission and compaction average values and standard deviations obtained for *undirected* networks for diverse types of random walks. (a) Time to transmit 90% of the network (T_{90}); (b) Time to transmit 90% of the network with Huffman compression (T^C_{90}); (c) Steering coefficient attained after exploring 90% of the network (S_{90}); (d) Steering coefficient for longer term exploration (S_L); (e) Compression ratio for 90% exploration (R_{90}), and (f) Compression ratio obtained after a large number of iterations (R_L).

S_{90}. The WS1 and WS2 models usually led to low steering coefficient values. This probably happens because most nodes in these networks have the same degree, and therefore they also possess similar frequency of visits. The BA model always resulted in the largest steering coefficient. Such an effect is possibly a consequence of the power-law nature of the degree distribution, in which nodes with small degree tend to be scarcely visited while the opposite happens for nodes with large degree. A more in-depth discussion about the encoding efficiency is presented in Section S1 of the supplementary material, where we compare the compression achieved when estimating the symbol probabilities using the nodes degrees with those achieved when using individual messages to estimate the probabilities. Remarkably, in most cases we found that predicting the symbol probabilities from the network (instead of from the sequence of symbols) yielded better compaction.

The long term steering coefficient values (S_L) are shown in Fig. 3(d). The TSAW dynamics led to the highest S_L values, followed closely by the RWD. A variety of behaviors were observed for RWID, all of them yielding S_L values smaller than those obtained for the other dynamics.

Fig. 3(e,f) shows the compression ratio for 90% of network recovery, R_{90}, and for long term exploration, R_L, obtained for the several network models and dynamics. Similar results were obtained for most cases. Interestingly, the GEO and WAX, which had produced substantially different compression and transmission times in the previous experiment, implied similar compression ratios.

The experimental results for the directed cases are similar to the results shown in Fig. 3 (more details are presented in Section S2 and Fig. S3 in the supplementary material). In this experiment we have the inclusion of the KN model (intrinsically directed), which always led to low reconstruction times in all cases. The better transmission obtained for the KN networks is possibly related to the process of network construction, which can be understood as a kind of random walk. This type of network has two key aspects: (i) for every node, the inward degree (k_{in}) is equal to the outward degree (k_{out}) and (ii) the reciprocity is smaller than the other considered networks. In Section S3 of the supplementary material we compare the transmission times of KN networks with those obtained for networks generated by a directed ER model (thus having reciprocity close to zero) and a configuration model [2] having the constraint $k_{\text{in}} = k_{\text{out}}$. The results indicate that these two aspects are responsible for the better transmission of KN networks. Furthermore, the results were again similar for T_{90} and T^C_{90}, with the exception of the BA model (see Fig. S3 in the supplementary material). This model has large T_{90} for the RWD dynamics. Overall, the values of S_{90} showed similar trends as in the undirected case. The long term steering coefficients, S_L, resulted smaller than for the undirected cases, with exception of the BA model, which was similar to that case.

The compression ratios R_{90} and R_L obtained for the directed networks are shown in Fig. S3(e,f) (please refer to Section S2 in the supplementary material). These results are generally similar to those obtained for undirected networks.

Transmission and compaction average values and standard deviations obtained for *directed* networks for diverse types of random walks. (a) Time to transmit 90% of the network (T_{90}); (b) Time to transmit 90% of the network with Huffman compression (T^C_{90}); (c) Steering coefficient attained after exploring 90% of the network (S_{90}); (d) Steering coefficient for

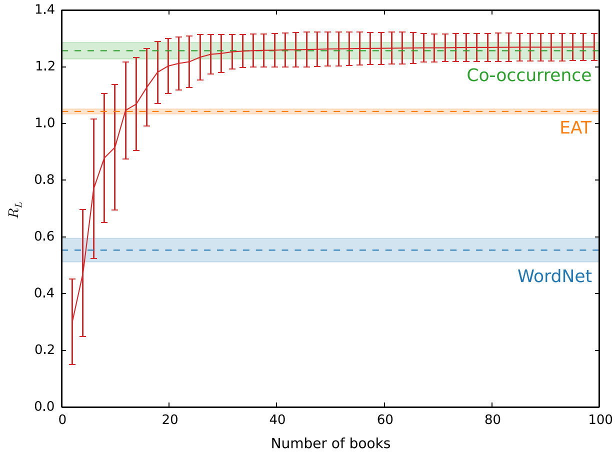

Fig. 4. Comparison between the average compression ratio (R_L) computed from word statistics of real books and the average R_L from network topology (degree). The dashed lines represent the average R_L by considering the network topologies. The red curve includes the average values and standard deviations in terms of the number of books used for word frequency estimation.

longer term exploration (S_L); (e) Compression ratio for 90% exploration (R_{90}), and (f) Compression ratio obtained after a large number of iterations (R_L).

3.2. Real-world networks

In this section we compare the compression ratio (R_L) of texts by using dictionaries from books and network degrees. All texts were pre-processed [36,37] so that the words were lemmatized, and the stop-words (words with low semantic content) and punctuation were removed. When deriving statistics only from books, a total of 100 books were considered, and for each book b a set of $1 \leq N \leq 99$ randomly chosen books, not including b, was used to create the Huffman code. In order to calculate R_L, we employed 14-bit symbols to represent the words. When using the network degree for compression, the following datasets were used: the Edinburgh Associative Thesaurus (EAT) [38], WordNet [39], and a co-occurrence network [36] created from the 100 considered books.

The obtained R_L values are shown in Fig. 4. When considering the networks for compression, the best R_L was obtained when the co-occurrence network was employed. The EAT network led to a lower compression ratio, while a larger text size ($R_L < 1$) was implied by WordNet. The compression ratio calculated from book statistics is, on the average, slightly better than the previous network approach when 40 or more books were considered.

4. Conclusions

The areas of information theory and complex networks have developed in a mostly independent way. However, as argued in the present work, these two areas present several shared and complementary elements which, when integrated, can be used to model, characterize and analyze a broad range of important real-world problems ranging from spoken/written language to DNA sequences. A formal framework, leading to a potentially new investigations, has been reported, involving the transmission of an original network, by using a sampling dynamics such as random walks, which produces a sequence of symbols or time series that can be used by a receiver to reconstruct the original network. More effective transmission demands compaction of the time series which, we argue, is directly related to the topology of the original network. We also show that the critical issue of compaction is directly related to one of the central paradigms in network science, namely the relationship between topology and dynamics, more specifically regarding the ability to predict the frequency of symbols from the very topology of the original network. Interestingly, the quality of such a prediction depends on the interplay between the topology of the original network and the adopted dynamics.

In addition to proposing the systematic integration between network science and information theory, we also illustrated a typical problem that can be tackled in this area, namely the efficiency of transmission of several types of networks by using different kinds of random walks. A number of interesting results has been reported. First, we confirmed that different network topologies and dynamics can lead, irrespectively of compaction, to rather distinct performances. Interestingly, the BA model exhibited a markedly distinct behavior, oscillating between the best and worst performances, depending on the probing random walk. On the other hand, the KN model, in almost all cases, led to the best performance in the case of directed networks. In addition, the two adopted geographical networks, namely WAX and GEO, despite their seemingly analogous spatial organization, yielded rather different results in the case of undirected networks. It is particularly interesting to observe that the BA model, which is topologically very complex (non-uniform degree distribution), led to the best overall compaction in most cases, except the RWID. This is because the power law degree distribution implies in asymmetric distribution of frequency of symbols, and therefore, more effective Huffman coding.

Some interesting promising applications include the modeling of opinion spreading, syllabuses planning, and language evolution. Particularly, in the case of opinion spreading, the proposed framework can be employed in a new direction that considers a network as the full message to be transmitted. More specifically, the message is transmitted as a sequence of symbols from a spreader (source) to a receiver (target) and is translated into a network. This approach allows the simulation of different dynamics, such as failures in the transmission (missing symbols), which are reflected in incomplete transmitted networks. Also, it would be interesting to consider noisy transmission, as well as higher order statistical coding of symbols. Regarding network topology, it would be particularly promising to investigate how modular structure can impact the transmission. Other types of dynamics can be also considered, especially those related to neuronal signal propagation.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. The authors also acknowledge financial support from FAPESP (2016/19069-9, 2015/08003-4, 2015/18942-8, 2014/20830-0 and 2011/50761-2), CNPq (grant no. 307333/2013-2) and NAP-PRP-USP.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.physa.2018.10.005.

References

[1] T.M. Cover, J.A. Thomas, Elements of Information Theory, in: Wiley Series in Telecommunications and Signal Processing, Wiley-Interscience, 2006.
[2] M. Newman, Networks: An Introduction, Oxford University Press, Inc., New York, NY, USA, 2010.
[3] A.-L. Barabási, Network Science, Cambridge University Press, 2016.
[4] J. Borge-Holthoefer, N. Perra, B. Gonçalves, S. González-Bailón, A. Arenas, Y. Moreno, A. Vespignani, The dynamics of information-driven coordination phenomena: A transfer entropy analysis, Sci. Adv. 2 (4) (2016) e1501158.
[5] J. Sun, D. Taylor, E.M. Bollt, Causal network inference by optimal causation entropy, SIAM J. Appl. Dyn. Syst. 14 (1) (2015) 73–106.
[6] S.E. Ahnert, Generalised power graph compression reveals dominant relationship patterns in complex networks, Sci. Rep. 4 (2014).
[7] M. De Domenico, J. Biamonte, Spectral entropies as information-theoretic tools for complex network comparison, Phys. Rev. X 6 (4) (2016) 041062.
[8] L. Lacasa, V. Nicosia, V. Latora, Network structure of multivariate time series, Sci. Rep. 5 (2015) 15508.
[9] M. Rosvall, C.T. Bergstrom, Maps of random walks on complex networks reveal community structure, Proc. Natl. Acad. Sci. 105 (4) (2008) 1118–1123.
[10] R. Ulanowicz, Quantitative methods for ecological network analysis and its application to coastal ecosystems, Treatise Estuar. Coast. Sci. 3 (2011) 5–57.
[11] M. Andjelković, N. Gupte, B. Tadić, Hidden geometry of traffic jamming, Phys. Rev. E 91 (5) (2015) 052817.
[12] X.-B. Chen, Y. Su, G. Xu, Y. Sun, Y.-X. Yang, Quantum state secure transmission in network communications, Inform. Sci. 276 (2014) 363–376.
[13] Q. Wang, Z. Gong, An application of fuzzy hypergraphs and hypergraphs in granular computing, Inform. Sci. 429 (2018) 296–314.
[14] J.P. Sutton, Neuroscience and computing algorithms, Inform. Sci. 84 (3) (1995) 199–208.
[15] T.H. Cormen, C. Stein, R.L. Rivest, C.E. Leiserson, Introduction to Algorithms, second ed., McGraw-Hill Higher Education, 2001.
[16] L.F. Costa, O. Sporns, L. Antiqueira, M.G.V. Nunes, O.N. Oliveira Jr, Correlations between structure and random walk dynamics in directed complex networks, Appl. Phys. Lett. 91 (2007) 054107.
[17] L. Lacasa, I.P. Mariño, J. Miguez, V. Nicosia, É. Roldán, A. Lisica, S.W. Grill, J. Gómez-Gardeñes, Multiplex decomposition of non-markovian dynamics and the hidden layer reconstruction problem, Phys. Rev. X 8 (3) (2018) 031038.
[18] L. Lacasa, J. Gómez-Gardenes, Correlation dimension of complex networks, Phys. Rev. Lett. 110 (16) (2013) 168703.
[19] L. Lovász, Random Walks on Graphs: A Survey, in: Combinatorics, Paul Erdos is Eighty, vol. 1, Janos Bolyai Mathematical Society, Hungary, 1993.
[20] M. Bonaventura, V. Nicosia, V. Latora, Characteristic times of biased random walks on complex networks, Phys. Rev. E 89 (1) (2014) 012803.
[21] D.J. Amit, G. Parisi, L. Peliti, Asymptotic behavior of the true self-avoiding walk, Phys. Rev. B 27 (3) (1983) 1635.
[22] Y. Kim, S. Park, S.-H. Yook, Network exploration using true self-avoiding walks, Phys. Rev. E 94 (4) (2016) 042309.
[23] J. Gómez-Gardeñes, V. Latora, Entropy rate of diffusion processes on complex networks, Phys. Rev. E 78 (2008) 065102.
[24] H.F. de Arruda, F.N. Silva, L.d.F. Costa, D.R. Amancio, Knowledge acquisition: A complex networks approach, Inform. Sci. 421 (2017) 154–166.
[25] P. Erdős, A. Rényi, On the evolution of random graphs, Publ. Math. Inst. Hungar. Acad. Sci. 5 (1960) 17–61.
[26] A.-L. Barabási, R. Albert, Emergence of scaling in random networks, Science 286 (5439) (1999) 509–512.
[27] D. Watts, S. Strogatz, Collective dynamics of 'small-world'networks, Nature 393 (6684) (1998) 440–442.
[28] B.M. Waxman, Routing of multipoint connections, IEEE J. Sel. Areas Commun. 6 (9) (1988) 1617–1622.
[29] J. Dall, M. Christensen, Random geometric graphs, Phys. Rev. E 66 (1) (2002) 016121.
[30] L.F. Costa, Knitted complex networks, arXiv:abs:0711.2736.
[31] A. Cohen, W. Hersh, C. Dubay, K. Spackman, Using co-occurrence network structure to extract synonymous gene and protein names from medline abstracts, BMC Bioinformatics 6 (1) (2005) 103.
[32] H. Liu, J. Cong, Language clustering with word co-occurrence networks based on parallel texts, Chin. Sci. Bull. 58 (10) (2013) 1139–1144.
[33] D.R. Amancio, Probing the topological properties of complex networks modeling short written texts, PLoS One 10 (2) (2015) e0118394.
[34] D.R. Amancio, L.F. Costa, Unveiling the relationship between complex networks metrics and word senses, Europhys. Lett. 98 (1) (2012) 18002.
[35] D. Salomon, G. Motta, Handbook of Data Compression, fifth ed., Springer Publishing Company, Incorporated, 2009.
[36] D.R. Amancio, Comparing the topological properties of real and artificially generated scientific manuscripts, Scientometrics 105 (3) (2015) 1763–1779.
[37] F.N. Silva, D.R. Amancio, M. Bardosova, L.d.F. Costa, O.N. Oliveira Jr, Using network science and text analytics to produce surveys in a scientific topic, J. Informetrics 10 (2) (2016) 487–502.
[38] G.R. Kiss, C. Armstrong, R. Milroy, J. Piper, An associative thesaurus of english and its computer analysis, Comput. Lit. Stud. (1973) 153–165.
[39] G.A. Miller, Wordnet: a lexical database for english, Commun. ACM 38 (11) (1995) 39–41.