Trajectory Networks and Their Topological Changes Induced by Geographical Infiltration

Luciano da Fontoura Costa

Institute of Physics at São Carlos, University of São Paulo,

P.O. Box 369, São Carlos, São Paulo, 13560-970 Brazil

(Dated: 14th March 2008)

In this article we investigate the topological changes undergone by trajectory networks as a consequence of progressive geographical infiltration. Trajectory networks, a type of knitted network, are obtained by establishing paths between geographically distributed nodes while following an associated vector field. For instance, the nodes could correspond to neurons along the cortical surface and the vector field could correspond to the gradient of neurotrophic factors, or the nodes could represent towns while the vector fields would be given by economical and/or geographical gradients. Therefore trajectory networks are natural models of a large number of geographical structures. The geographical infiltrations correspond to the addition of new local connections between nearby existing nodes. As such, these infiltrations could be related to several real-world processes such as contaminations, diseases, attacks, parasites, etc. The way in which progressive geographical infiltrations affect trajectory networks is investigated in terms of the degree, clustering coefficient, size of the largest component and the lengths of the existing chains measured along the infiltrations. It is shown that the maximum infiltration distance plays a critical role in the intensity of the induced topological changes. For large enough values of this parameter, the chains intrinsic to the trajectory networks undergo a collapse which is shown not to be related to the percolation of the network also implied by the infiltrations. (Copyright Luciano da F. Costa, 2008)

PACS numbers: 89.75.Hc, 89.75.Fb, 89.75.-k

'No one remembers what need or command or desire drove Zenobia's founders to give their city this form, ..., which has perhaps grown through successive superimpositions from the first, now undecipherable plan.' (I. Calvino, Inivisible Cities

I. INTRODUCTION

Graphs and complex networks can be classified into two major categories: geographical and emphnongeographical. The former type of networks is characterized by the fact that each of their nodes has a welldefined spatial position, expressible by respective coordinates. Contrariwise, the nodes of non-geographical networks do not have specific positions, or if they have we do not know what they are. Several real-world networks are geographical, including power distribution (e.g. [1]), tourism (e.g. [2]), transportation (e.g. [3]), biological networks (e.g. bone structure [4] and gene expression expression [5, 6]), amongst many others. The connectivity of geographical networks is often, but not always, affected or even defined by the proximity and spatial adjacencies between the nodes, in the sense that nodes which are neighbors or close one another tend to have larger chances of being connected. Several models of geographical networks have been proposed in the literature (e.g. [7-10]). A new family of networks, namely the *knitted networks*, was proposed recently [11, 12] to include all networks defined and composed by paths, i.e. sequences of edges without repetition of nodes.

In this article, we expand the family of knitted net-

works by incorporating structures generated by trajectories defining paths following a given vector field. More specifically, a set of nodes is distributed within a given domain (a 2D space in this article, but the extension to higher dimensions is immediate), one node is chosen as origin, and the respective trajectory (line of force) is obtained while the nodes which are closer than a given maximum distance to the current point of the trajectory are sequentially incorporated into the path. This procedure is repeated several times, yielding a network with connections aligned to the vector field. In other words, the paths correspond to approximations of the solutions of the dynamical system represented by the vector field. Figure 1 illustrates two trajectory networks obtained from the vector fields $\vec{\phi}(x, y) = (y, x)$ and $\vec{\phi}(x, y) = (y, -x)$ (b).

Trajectory networks are important because they represent a natural putative model for several real-world structures and phenomena including the establishment of neuronal connections under the influence of neurotrophic fields (e.g. [13–15]), the growth of transportation systems under geographical and economical influences (e.g. 'every path leads to Rome'), the growth of trees and roots under influence of trophic factors [16], the development of channel-based systems such as bone structure and the vascular system, amongst many other important systems.

The focus of attention in the current work is to investigate how the topology of trajectory networks, a geographical type of knitted network, is affected as the consequence of progressive *geographical infiltration*. By geographical infiltration (hence infiltration for short), it is meant any process which tend to interconnect pairs of nodes which are spatially close one another. Several



FIG. 1: Trajectory networks obtained for the fields $\vec{\phi}(x,y) = (y,x)$ (a) and $\vec{\phi}(x,y) = (y,-x)$ (b).

real-world systems are exposed to this type of alterations, such as the appearance of cracks along channels, the establishment of new local routes between towns and cities, contaminations between vessels of fibers, gallery building by parasites, intentional attacks, internal spreading of diseases, to cite just a few cases. In the current work, the infiltration process is simulated by selecting nodes at random and connecting this node to all other nodes which are closer than a maximum distance D_p . Therefore, the adopted infiltration corresponds to the progressive incorporation of *tufts* of local connectivity.

The effects of progressive infiltration on the topology of trajectory networks is here investigated by quantifying the degree, clustering coefficient, size of the largest component, as well as the number and length of the chains present in the network. The characterization of chains as an important category of motifs in networks was reported recently [17]. As a matter of fact, several realworld networks incorporate several chains in their structure in ways which are intrinsically specific to their dynamics and organization. The trajectories networks are possibly the first theoretical model of complex networks which naturally incorporates a large number of chains. These motifs are a consequence of the linking of spatially distributed nodes along the trajectories defined by the given vector fields. Indeed, the chains appearing in several of the real-world situations which can be modeled by trajectory networks represent an important structural feature in the sense of providing relatively independent (i.e. with few interconnections) routes between possible destinations. Therefore, it becomes particularly important to characterize the structure of trajectory networks before and after infiltration by considering the number and length of the existing chains. Interestingly, the effect of infiltrations can be either bad or good, depending on each specific system. For instance, the incorporation of additional local routes is in principle beneficial for transportation and communication systems. On the other hand, the addition of local connections in biological networks (e.g. bone or neuronal networks) may have catastrophic consequences. Observe that in the latter situation the main purpose of the chains/fibers is actually to provide mutual isolation. In both cases, the quantification of the effects of the infiltration over the topology of the respective networks can provide valuable information to be interpreted from the perspective of each problem.

This article starts by presenting the basic concepts including the generation of trajectory networks and the geographical infiltrations — and follows by describing the experiments and discussing the respectively obtained results.

II. BASIC CONCEPTS

A complex network is a graph exhibiting a particularly intricate structure. The connectivity of a undirected, unweighted network can be completely represented in terms of the respective adjacency matrix K, such that each interconnection between two nodes i and j implies K(i, j) = K(j, i) = 1, with K(i, j) = K(j, i) = 0 being otherwise imposed. The *immediate neighbors* of a node iare those nodes which receive an edge from i. The *degree* of a node i is equal to the number of its immediate neighbors. Two nodes are said to be *adjacent* if they share an edge; two edges are adjacent if they share one node. A sequence of adjacent edges is a *walk*. A *path* is a walk which never repeats a node or edge. The length of a walk (or path) is equal to the respective number of involved edges. The *clustering coefficient* of nodes i is calculated by dividing the number of interconnections between its immediate neighbors and the maximum possible number of connections which could be established between those neighbors.

A connected component of a network is a subgraph such that each of its nodes can be reached from any of its other nodes [31]. A chain is a subgraph of a network such as that each of its nodes has degree 1 or 2 and not additional nodes of degree 1 or 2 are connected to it [17]. The *length* of a chain is given by its number of edges. Two measurements which can be used to characterize the chains in a given network include the number of such chains and average and standard deviation of their respective lengths. Chains are naturally related to paths along the network.

III. TRAJECTORY NETWORKS

A family of networks, namely the knitted complex networks, was introduced recently [11, 12] incorporating all networks organized around the concept of *paths*. Two main types of knitted networks were initially identified: path-transformed and path-regular. The former subcategory of knitted complex network is obtained by performing the start-path transformation [11] on a given network (star and path connectivities can be understood as duals, e.g. through the line-graph transformation). Therefore, networks with power-law distribution of path lengths can be obtained by star-path transforming Barabási-Albert networks [18]. The second type of knitted complex networks, namely the path-regular networks, is particularly simple and involves starting with a set of N isolated nodes and performing several paths encompassing all nodes. Path-regular networks have been found to exhibit marked similar properties between different configurations or nodes in the same configuration (e.g. [12, 19]). An even more regular version of the path-regular network, with all nodes exhibiting identical degrees, was later reported in [20, 21].

Geographical networks are characterized by the fact that each of their nodes has a well-defined spatial position. Geographical networks represent an important category of complex networks because several real-world structures are inherently embedded into 2D or 3D spaces, and their connectivities are strongly affected by proximity and spatial adjacency. Given a set of spatially distributed nodes embedded in a continuous space to which a vector field is associated, it is possible to obtain geographical networks whose connections are a consequence not only of the proximity between nodes, but also of the orientations implied by the respectively associated vector field. Several real-world can be thought as involving a geographical distribution of nodes and associated vector fields. For instance, the neurons along the cortical surface can be represented as a set of geographically distributed nodes, while their connections are established to a great extent as a consequence of neurotrophic fields (e.g. electrical or chemical gradients). Systems of streets, roads and highways can also be understood as involving a set of spatially distributed nodes (the intersections between routes), with the interconnections being established in terms of the spatial proximity between nodes as well as geographical and economical fields (e.g. the trend to connect to a big city, to avoid a geographical obstacle or to follow level-sets of height). Several other natural and human-made complex systems can be modeled by trajectory networks. Trajectory networks are related to gradient networks (e.g. [22-24]), field interactions [5, 6, 25], as well as dynamical systems (e.g. [26, 27]). In the present work, we understand trajectory networks as a particular case of knitted networks.

The trajectory networks considered in the present article are obtained as follows. First, a two-dimensional workspace of size $L \times L$ is defined, and a vector field $\phi(x,y)$ is associated to it. For simplicity's sake we assume that $-L/2 \leq x, y \leq L/2$. All networks considered henceforth in this work are obtained for the vector field $\phi(x,y) = (y,x)$. N points are distributed along this space with uniform probability. A total of N_p trajectories are then performed while obtaining each network. A starting point is randomly selected, and the respective line of force (always parallel to the vector field) is calculated by using the Euler leapfrog numerical method (e.g. [28]). At each current time, if a new node is found at a distance not exceeding D_p , that node is connected to the previous node, and so on. As it is clear from the example of trajectory network shown in Figure 1, the combination of proximity and orientation constraints while performing the connections yield networks incorporating several chains, which closely follow the vector field orientation. Different degrees of interconnectivity between and along the chains can be obtained by varying the total number of points and the parameter D_p . Observe that the number of chains is reduced for larger values of D_p/N . Once all trajectories are performed, the isolated points can be removed (as adopted henceforth) or not (allowing further connections).

IV. GEOGRAPHICAL INFILTRATIONS

Given a geographical network, several types of perturbations of its structure can arise as a specific consequence of its geographical nature, in the sense that nodes which are spatially closer may interfere one another. For instance, in a neuronal system, unwanted connections may appear between nearby neurons as a consequence of diseases. In transportation systems, it is only too natural to incorporate new local connections to the network. Several other types of geographical interferences are possible, including those arising as a consequence of contaminations, attacks, infiltrations, amongst many other. In this work we incorporate progressive infiltrations to a given network geographical network by selecting one of its nodes and connecting to it all other nodes which are not further than a maximum distance D_i .

V. RESULTS AND DISCUSSION

A set of 30 trajectory networks was obtained for the field $\vec{\phi}(x,y) = (y,x)$. A total of 1000 nodes was initially distributed within a squre region of side L = 100 centered at (0,0), and $N_p = 100$ trajectories were numerically calculated. Starting from a randomly chosen node, each node at a maximum distance $D_p = 2$ from the current growing extremity of each trajectory was successively connected. An example of obtained trajectory network is shown in Figure 1. Each of the 30 networks underwent progressive infiltrations assuming $D_i = 5$ and $D_i = 10$. Figure 2 shows four stages (100, 200, 300 and 400) along the successive infiltrations for $D_i = 5$. Examples of the results of infiltrations with $D_i = 10$ are depicted in figure 3.

In order to characterize the alterations in the topology of the trajectory networks as they underwent progressive infiltrations, a set of measurements (e.g. [29]) was taken along the process. These measurements included the average and standard deviation of the node degree, clustering coefficient, size of the largest connected component, and chain lengths along successive infiltration stages. Only chains longer than 3 edges were considered in the respective measurements. These chains were identified by starting from each of the network nodes with degree 1 or 2 and following along both sides (in case of degree 2) until the respective extremities of the chains (nodes with degree 1 or larger than 2) were found (each detected chain was removed from the network in order to accelerate the processing of the remaining nodes). The results obtained for $D_i = 5$ and $D_i = 10$ are shown in Figure 2 and 3, respectively. Figure 6 and 7 show the above measurements for all the 30 considered networks.

It is clear from Figures 4 to 7 that, as could be expected, the degree and clustering coefficient both increased as a consequence of the addition of the infiltration tufts. Both such increases are sublinear, with a steeper decrease in the rate of clustering coefficient increase observed for $D_i = 10$ (Fig. 7). The relative sizes of the maximum connected components suffer an abrupt transition before the 160 first infiltrations (most of the transitions take place before that value) for both settings of D_i , but is more abrupt for $D_i = 10$. This change is related to the percolation of the chains in the original network. Another relatively abrupt change is observed for the path lengths, most of which stabilizing themselves at a value near 6 for $D_i = 5$ and 4 for $D_i = 10$. The interval from the start of the infiltrations until the average length of the chains stabilizes (as observed above) is called the *period of collapse* of the chains. Very few networks remained with large average chain lengths larger after 200 infiltrations. This confirms the fact, evident from Figure 8, that the tuft infiltrations tend to quickly eliminate most of the long chains in the trajectory networks (the chain collapse). For larger values of D_i , after the collapse of the chains, the vector field influence on the network connectivity can be hardly distinguished by visual inspection, such as in Figures 3(b-d). It is important to keep in mind that the fact that small values of D_i tend to imply little effect over the chain structure of the trajectory networks is ultimately related to the number N of initial nodes and the maximal distance D_p considered for chaining the nodes during the construction of the networks.

The two involved critical phenomena, namely the percolation of the networks and the collapse of the chains, were investigated further in order to search for possible relationship between their respective onsets. In order to do so, transition points along the successive infiltrations were identified automatically. These points, respectively T_p and T_c , correspond to the first occurrence of the value 1 for the relative size of the largest connected component and the first occurrence of the average chain length which is smaller or equal than 5, respectively. Figure 8 shows the respectively obtained distribution of T_p and T_c obtained for the 30 realizations of networks with $D_i = 10$. It is clear from this figure that the two critical phenomena taking place in the considered trajectory networks seem to be largely independent, in the sense that no correlation has been observed between their critical values. Interestingly, as shown in Figure 8, the collapse of the chains can take place before the respective percolation.

VI. CONCLUDING REMARKS

Geographical networks represent an important category of complex networks because of their natural potential for modeling a large number of real-world and human-made complex structures and systems. At the same time, the category of complex networks build up by paths, namely the *knitted networks*, constitutes an important superclass of complex structures because of their intrinsic association with the concept of paths (as opposited to star connectivity) and random walk dynamics (e.g. [11, 12, 19]). In this work, trajectory networks have been understood to belong to the supercategory of knitted networks as a consequence of the fact that these structures are obtained by performing paths (trajectories) along the nodes. In this sense, trajectory networks become a special case where the paths tend to follow an associated vector field. Our main interest in the present work, however, consisted in investigating how the topology of trajectory networks changed as a consequence of geographical infiltrations. While several types of attacks and perturbations have been considered and investigated in complex network research, relatively lesser attention has been focused on perturbations intrinsically related to geographical constraints, especially the adjacency and proximity between nodes. Yet, several important real-



FIG. 2: The network in Fig. 1 after 100 (a), 200 (b), 300 (c) and 400 (d) infiltrations with $D_i = 5$.

world and human-made systems are prone to this type of perturbations, ranging from the onset of unwanted neuronal connections as a consequence of diseases to the incorporation of new local routes to transportation systems.

The main contributions reported in this article are listed and reviewed in the following:

Trajectory networks as a special case of knitted complex networks: We have enlarged the family of knitted complex networks through the incorporation of trajectory networks. This type of geographical knitted network corresponds to an interesting case where the connectivity is the consequence of both the proximity between nodes and the orientation of the underlying vector field. New type of perturbation of network structure: We considered, possibly for the first time, perturbations (or 'attacks') to geographical networks which depend on the proximity between the spatially distributed nodes. We focused attention on 'tuft' infiltrations, where a node iis randomly chosen and all other nodes which are closer than a maximum distance D_i are connected to node i. This type of topological change can be related to several real-world effects such as unwanted neuronal tangles as a consequence of diseases, establishment of local connections in transportation networks, contaminations, and attacks.

Identification of drastic variation of the effects of the infiltrations: The progressive infiltration of a trajectory



FIG. 3: The network in Fig. 1 after 100 (a), 200 (b), 300 (c) and 400 (d) infiltrations with $D_i = 10$.

network was investigated in a systematic manner, considering 30 realizations of networks obtained for the same configuration with respect to the vector field $\vec{\phi}(x, y) = (y, x)$. The changes in the networks topology was monitored by taking several measurements including the degree, clustering coefficient, size of the largest connected component, as well as the particularly relevant lengths of the existing chains. The latter measurements are especially important because the trajectory networks are inherently composed by chains. While the degree and clustering coefficients underwent relatively smooth increases, the size of the largest component and average chain lengths were subjected to relatively abrupt variations related to the percolation of the network (in the case of the largest connected component) and to the collapse of the chain structure (in the case of the average chain lengths). The value of D_i was found to be have great influence on such topological changes induced by the infiltrations, with values much larger than D_p implying particularly intense changes, especially regarding the chain structure. After the collapse of the chains, the effect of the original vector field on the network connectivity could hardly be discerned. Such findings are particularly important for a large number of real-world structures underlain by trajectory networks and geographical infiltrations.

Independence of percolation and collapse of chains: The progressive infiltration of trajectory networks involves two critical phenomena: its percolation and the collapse of its chain structure. Interestingly, no clear re-



FIG. 4: Measurements of degree, clustering coefficient, size of the largest connected component and chain lengths in terms of the number of infiltrations (identified as 'time') with $D_i = 5$ for a network obtained for the vector field $\vec{\phi}(x, y) = (y, x)$.

lationship between these phenomena has been identified by considering the critical times T_p and T_c . This implies that the collapse of the chains can not be predicted from the percolation of the respective network, and vice-versa. As a matter of fact, it has also been observed that the collapse of the chains can take place before the percolation of the respective network.

The several possibilities of future work include but are not limited to the following:

Other types of vector fields: It would be interesting to investigate how the patterns of topological changes observed in this work extends to trajectory networks obtained by considering other vector fields, as well as other configurations of the involved parameters.

Orthogonal infiltrations: In this work we focused attention on tuft infiltrations. It would be interesting to study the topological changes of trajectory networks with respect of other types of geographical perturbations, such as connecting points according to proximity and orientations orthogonal to the vector field (possibly also through trajectories).

Infiltration by increasing distances: While the infiltrations implemented in this article consisted in selecting nodes followed by tuft interconnection, it would be particularly interesting to investigate the topological alterations of trajectory networks while all pairs of nodes are joined according to successive distances. Such a type of infiltration is guaranteed to completely eliminate the chains after a critical interval.

Application to real-world networks: It would be interesting to quantify the alterations of real-world networks expressible by trajectory networks, including transportation networks, power distribution, communications, tourism and neuronal systems.

Application to Image and Shape Analysis: The analysis of images containing objects and shapes has remained a great challenge (e.g. [25, 30]). It would be particularly interesting to consider the application of the concepts and methods reported in the current work to such problems. More specifically, trajectories can be obtained in gray-level images by considering their respective gradient fields. So, by distributing points through the image and interconnecting them while taking in to account trajectories driven by the gradient fields, it is possible to obtain respective network representations incorporating a great deal of the intrinsic geometric features. Shapes



FIG. 5: Measurements of degree, clustering coefficient, size of the largest connected component and chain lengths in terms of the number of infiltrations (identified as 'time') with $D_i = 10$ for a network obtained for the vector field $\vec{\phi}(x, y) = (y, x)$.

represented by their contour can also be mapped into trajectory networks by considering vector fields induced by their borders (e.g. electrical or distance fields). The topological properties of the respective measurements are expected to provide valuable features for image and shape analysis and classification. Signatures obtained by considering the evolution of several measurements of the soobtained networks as the consequence of geographical infiltration can provide additional features for visual characterization and classification.

Acknowledgments

Luciano da F. Costa thanks CNPq (308231/03-1) and FAPESP (05/00587-5) for sponsorship.

- R. Albert, I. Albert, and G. L. Nakarado, Phys. Rev. E 69, 025103 (2004).
- [2] R. Baggio, Phys. A **379**, 727 (2007).
- [3] M. Barthelemy and A. Flammini, J. Stat. Mech. p. 07002 (2006).
- [4] L. da F. Costa, M. P. Viana, and M. E. Beletti, Appl. Phys. Letts. 88, 033903 (2006).
- [5] L. da F. Costa, B. A. N. Travencolo, A. Azeredo, M. E. Beletti, G. B. Muller, D. Rasskin-Gutman, G. Sternik, M. Ibanes, and J. C. I. Belmonte, Appl. Phys. Letts. 86, 3901 (2005).
- [6] L. Diambra and L. da F. Costa, Bioinformatics 21, 3846 (2005).
- [7] M. Kaiser and C. C. Hilgetag, Phys. Rev. E 69, 036103 (2004).
- [8] L. da F. Costa, Intl. J. Mod. Phys. C 16, 1001 (2005).
- [9] M. T. Gastner and M. E. J. Newman, Eur. Phys. J. B 49, 247 (2006).
- [10] C. P. Warren, L. M. Sander, and I. M. Sokolov, Phys. Rev. E 66, 056105 (2002).
- [11] L. da F. Costa (2007), arXiv:0711.1271.
- [12] L. da F. Costa (2007), arXiv:0711.2736.



FIG. 6: Averages of degree, clustering coefficient, size of the largest connected component and chain lengths in terms of the number of infiltrations (identified as 'time') with $D_i = 5$ for each of the 30 networks obtained for the vector field $\vec{\phi}(x, y) = (y, x)$.

- [13] E. J. Huang and L. F. Reichardt, Ann. Rev. Neurosc. 24, 677 (2001).
- [14] M. V. Sofroniew, C. L. Howe, and W. C. Mobley, Ann. Rev. Neurosc. 24, 1217 (2001).
- [15] L. R. Squire, F. E. Bloom, S. K. McConnell, J. L. Roberts, N. S. Spitzer, and M. J. Zigmond, *Fundamental Neuroscience* (Academic Press, 2003).
- [16] P. J. Gregory, Plant Roots: Growth, activity and interactions with the soil (Wiley-Blackwell, 2006).
- [17] P. R. V. Boas, F. A. Rodrigues, G. Travieso, and L. da F. Costa, Phys. Rev. E 77, 026106 (2008).
- [18] R. Albert and A. L. Barabási, Rev. Mod. Phys. 74, 47 (2002).
- [19] L. da F. Costa (2007), arXiv:0712.0415.
- [20] L. da F. Costa (2008), arXiv:0802.0421.
- [21] L. da F. Costa (2008), http://hal.archivesouvertes.fr/hal-00258893/fr/.
- [22] Z. Toroczkai, B. Kozma, K. E. Bassler, N. W. Hengartner, and G. Korniss (2004), cond-mat/0408262.

- [23] Z. Toroczkai and K. E. Bassler, Nature 428, 716 (2004).
- [24] L. da F. Costa and G. Travieso, Intl. J. Mod. Phys. C 16, 1097 (2005).
- [25] L. da F. Costa (2006), arXiv:physics/0603025.
- [26] E. J. Huang and L. F. Reichardt, Europhysics News 36, 218 (2005).
- [27] E. P. Borges, D. O. Cajueiro, and R. F. S. Andrade, Eur. Phys. J. B 58, 469 (2007).
- [28] J. H. Mathews, Numerical Methods for Mathematics, Science and Engineering (Prentice Hall, 1987).
- [29] L. da F. Costa, F. A. Rodrigues, G. Travieso, and P. R. V. Boas, Advs. in Phys. 56, 167 (2007).
- [30] L. da F. Costa and R. M. Cesar, Shape Analysis and Classification: Theory and Practice (CRC Press, 2001).
- [31] Often a connected component is understood to be maximal, in the sense of incorporating all mutually reachable nodes in that component.



FIG. 7: Averages of degree, clustering coefficient, size of the largest connected component and chain lengths in terms of the number of infiltrations (identified as 'time') with $D_i = 10$ for each of the 30 networks obtained for the vector field $\vec{\phi}(x, y) = (y, x)$.



FIG. 8: Scatterplot of the percolation (T_p) and collapse (T_c) critical times for $D_i = 10$.