Eur. Phys. J. B 49, 513–522 (2006)
DOI: 10.1140/epjb/e2006-00085-1

THE EUROPEAN PHYSICAL JOURNAL B

Topological patterns in street networks of self-organized urban settlements

J. Buhl[1,a], J. Gautrais[1], N. Reeves[2], R.V. Solé[3], S. Valverde[3], P. Kuntz[4], and G. Theraulaz[1]

[1] Centre de Recherches sur la Cognition Animale, CNRS UMR 5169, Université Paul Sabatier, 118 route de Narbonne, 31062 Toulouse Cedex 4, France
[2] Laboratoire NXI GESTATIO, Département de Design, Université du Québec à Montréal, Case Postale 8888, Succursale Centre-Ville, Montréal (Québec), H3C 3P8, Canada
[3] ICREA-Complex Systems Lab,Universitat Pompeu Fabra, Dr Aiguader 80, 08003 Barcelona, Spain
[4] Laboratoire d'Informatique Nantes-Atlantique, Université de Nantes, 2 rue de la Houssinière, BP 92208, 44322 Nantes Cedex 03, France

Received 7 September 2005 / Received in final form 5 January 2006
Published online 31 March 2006 –

Abstract. Many urban settlements result from a spatially distributed, decentralized building process. Here we analyze the topological patterns of organization of a large collection of such settlements using the approach of complex networks. The global efficiency (based on the inverse of shortest-path lengths), robustness to disconnections and cost (in terms of length) of these graphs is studied and their possible origins analyzed. A wide range of patterns is found, from tree-like settlements (highly vulnerable to random failures) to meshed urban patterns. The latter are shown to be more robust and efficient.

PACS. 89.75.Hc Networks and genealogical trees – 89.40.Bb Land transportation – 89.65.Lm Urban planning and construction

1 Introduction

Human activities take place on a spatial matrix largely defined by three types of transportation networks. These involve matter (streets, roads, highways, railways, airport networks), energy (the power grid) and information (Internet, telephone networks). One of the oldest human-made webs is the city. The network of streets in a city is describable in terms of a grid of linked nodes defining a more or less ordered pattern. Although urban structures are not limited to a two-dimensional domain, cities are essentially planar structures. The architecture of streets and roads crossing the urban structure defines a basic template that strongly constrains the further development of other webs (such as the power grid or communication nets). Cities are usually considered as organized hierarchically into neighbourhoods, whose rank and spatial extent depend on a number of traits, including the economic function which they offer to the surrounding population. Hierarchical structures would be organized around specialized centres serving larger areas, whereas those of local needs serve smaller ones.

The growth dynamics of cities has changed over time, and modern towns experience rapid, often uncontrolled patterns of growth that escape from a top-down, planned scheme [1]. This happens at many different scales: today, in developing countries, besides the exponential growth of the main urban centers, we attend the simultaneous growth of thousands of slums, shantytowns and squatter settlements which accounts for one half of all urbanization processes in the world. Their growth follows a pattern often described as "organic", marked by many distributed decisions, and often characterized by a large degree of complexity [2]. At the scale of their global shape, many cities expand today largely through this second mechanism and recent studies reveal that they actually display fractal-like patterns not very different from those displayed by natural systems [3,4]. In this context, recent studies on morphological patterns of large urban structures suggest that some universality classes might be at work [5]. Specifically, Carvalho and Penn studied the statistical patterns of street maps in a large collection of cities, looking for regularities derived from rank-order curves using data collapse. Fractal patterns were reported and their features were suggested to be the result of distinct Levy processes.

A complementary analysis can be performed by looking at smaller scales in urban zones involving distributed (instead of planned) decisions. At a lower scale, whether in parts of a large modern city or in smaller settlements, decentralized growth of street networks can also be observed throughout the whole history of the city. Many of these networks do not result from a planning process where addresses are attributed prior to any construction, but rather

[a] *Current address*: Department of Zoology, University of Oxford, Tinbergen Building, Oxford OX1 3PS, UK
e-mail: buhl@cict.fr

emerge or evolve in an incremental way through a physical process involving local aggregation rules [6]. It follows that in many cases, the resulting topologies are complex and deviate from simple regular patterns such as square-grids. The common traits of such street networks are largely unknown and their quantitative description is lacking.

When dealing with the architecture of city networks (including the type of networks considered here but also others such as road networks) two different approaches can be considered. One is to explore the dynamic patterns that can arise on top of a pre-defined web. In this approximation the key question would be how the traffic of matter can be optimally distributed so that congestion (efficiency) is properly minimized (optimized). As discussed by some authors [7–9], two key requirements need to be fulfilled, namely avoid large detours and reduce the cost, which are assumed proportional to the length of the paths. Using Boltzmann and Darwinian strategies, optimal solutions can be found [10]. Such methods deal with global energy functions, and require solving a class of frustrated optimization problems. Not surprisingly, the optimal networks are somewhere located between two extreme situations. A different approach, which we take here, is to explore the topological organization of street networks as a static object. Such an approach is of value in those cases where no evidence for a planned building is at work (and thus no global optimization is involved). In this case, the network is the result of local decisions performed by a distributed set of individuals who made their decisions based on multiple constraints, not necessarily associated to global detour lengths or efficient traffic. As shown below, real street networks are also located somewhere between two well-known types of graphs which are used here as null models for comparison.

In this paper, we make use of graph theory as a powerful tool to characterize and compare the topology of street networks of non-planned settlements and to provide insights about their evolution and functional properties. The paper is organized as follows: in Section 2 the data set of networks used here is described. In Section 3, we quantify topological patterns in these networks. In Section 4, the efficiency of the webs is analyzed and the network robustness studied in Section 5. A general discussion is presented in Section 6.

2 Street networks

A number of theoretical studies have been exploring some relevant patterns present in city maps and road networks. Some of these studies involve considering streets themselves as the units of a so called information city network [11,12]. Thus, two streets are linked if they intersect each other. In this way, topological patterns are used to characterize information patterns in city structure or navigation along road networks [13]. Here we concentrate our analysis in both topological and geometrical patterns displayed by self-organized street networks.

Any of these street networks (SNS) can be described by an embedded planar graph [14] $G = (V, E)$ where $V = \{(v_i, x_i, y_i),\ (i = 1, \ldots, n)\}$ is the set of n nodes

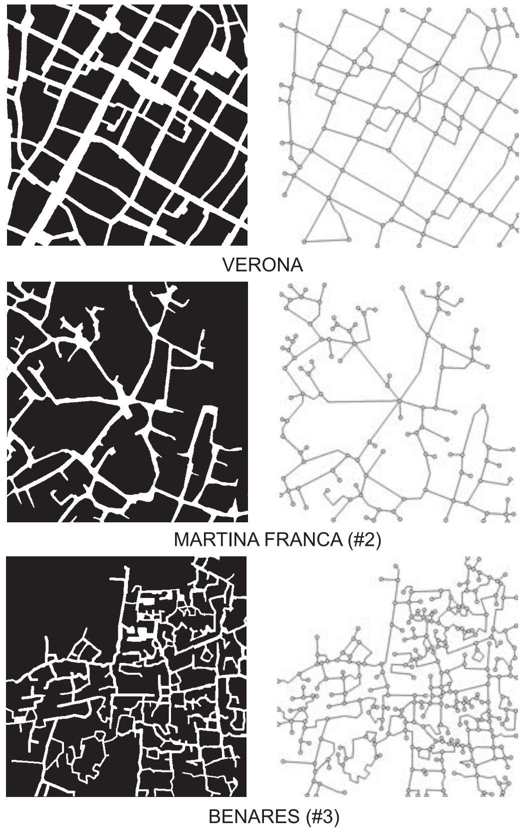

Fig. 1. Binary maps (left side) and the graph representation (right side) of 3 examples of street networks. On the binary maps, the white space correspond to public space. The graphs were established from the binary maps by considering the intersections and dead-ends of streets as the nodes, and the segments of streets joining the nodes as the edges.

characterized by their (x, y) position and diameter, and $E = \{(v_i, v_j)\}$ the set of m edges/connections between nodes and characterized by their length d_{ij}. The edges correspond to sub-sections of streets connecting two nodes. The nodes correspond to the squares, the intersections between streets and their dead ends. Here we study a large set of 41 non-planned settlements. These networks correspond to urban environments that evolved without global supervision through at least part of their development. The selected networks correspond mainly to settlements originated from Europe, Africa, Central America, India and one shantytown (El Agustino, Lima, Peru).

Street networks where selected among a database of 300 maps maintained by the NXI GESTATIO laboratory. The samples analyzed here correspond to the largest zones possible that were not directly in contact with strong heterogeneities, such as rivers or interface with other types of urbanization. The samples were square delimited, in order to introduce an equivalent artificial limit for all samples following a standard introduced by Caminos et al. [15]. The edges and nodes correspond exclusively to public spaces. Private spaces such as courtyards, gardens or private paths were excluded. We thus established a database of selected binary maps where public space was indicated in white and private spaces in black (Fig. 1). Table 1

Table 1. Basic characteristics of the networks: settlement name and number of the sample, country, number of nodes n, mean degree $\langle k \rangle$; average path length l and its value l_{rand} in a corresponding random graph (see Ref. [35]); parameter ξ in an exponential decay of the degree (k) distribution tail $P(k) \sim e^{-k/\xi}$ (for $k \geq 3$) and assortativity coefficient Γ (* indicates that Γ is significantly different from 0) .

Settlement	Country	n	$\langle k \rangle$	l	l_{rand}	ξ	Γ
ALGER CASBAH	ALGERIA	132	2.17	9.50	6.32	0.468	–0.207*
AMIENS (XIXth)	FRANCE	159	2.81	8.43	4.91	0.692	–0.047
ARLES (#1)	FRANCE	110	2.58	6.97	4.96	0.668	0.103
ARLES (#2)	FRANCE	99	2.16	7.61	5.96	0.506	–0.144
AT SJEN	ALGERIA	173	2.24	9.47	6.38	0.518	–0.165*
BENARES (#1)	INDIA	339	2.48	13.48	6.40	0.365	–0.135*
BENARES (#3)	INDIA	280	2.33	13.60	6.67	0.447	–0.153*
BEZIERS	FRANCE	90	2.33	6.99	5.31	0.610	–0.147
BONHEIDEN (#1)	BELGIUM	191	2.34	10.51	6.19	0.497	–0.116*
BONHEIDEN (#2)	BELGIUM	213	2.26	10.80	6.56	0.719	–0.116*
BONHEIDEN (#3)	BELGIUM	237	2.24	13.61	6.79	0.478	–0.183*
CAHORS	FRANCE	143	2.56	8.49	5.28	0.582	–0.108
CORDOBA (#1)	SPAIN	78	2.21	6.67	5.51	0.535	–0.230*
CORDOBA (#2)	SPAIN	66	2.42	5.73	4.73	0.756	–0.277*
CORDOBA (#3)	SPAIN	82	2.61	6.62	4.59	0.501	–0.097
CORDOBA (#4)	SPAIN	65	2.31	5.80	4.99	0.572	–0.138
DAX	FRANCE	55	2.62	4.85	4.16	0.692	–0.178
ECIJA (#1)	SPAIN	70	2.31	6.86	5.06	–	–0.139
ECIJA (#2)	SPAIN	69	2.32	6.25	5.03	0.550	–0.202*
EL AGUSTINO	PERU	118	2.02	11.99	6.80	–	–0.241*
GHARDAIA (#1)	ALGERIA	170	2.07	11.24	7.06	–	–0.207*
GHARDAIA (#2)	ALGERIA	177	2.18	9.84	6.64	–	–0.253*
GOSLAR	GERMANY	63	2.41	5.73	4.70	0.796	–0.079
HOMS (#1)	SYRIA	149	2.79	7.67	4.87	0.721	0.040
HOMS (#2)	SYRIA	137	2.34	7.93	5.80	0.542	–0.126
LUBECK	GERMANY	60	2.80	5.12	3.98	0.976	–0.149
MADRIGAL	SPAIN	63	2.60	5.38	4.33	0.647	–0.074
MARTINA FRANCA (#1)	ITALY	115	2.28	7.82	5.76	–	–0.122
MARTINA FRANCA (#2)	ITALY	116	2.14	7.89	6.26	0.746	–0.230*
PUTIGNANO	ITALY	133	2.20	8.06	6.22	0.469	–0.189*
QAZVIN (#1)	ALGERIA	101	2.79	6.70	4.49	0.851	–0.037
QAZVIN (#2)	ALGERIA	104	2.75	6.90	4.59	0.832	0.061
ROME	ITALY	108	2.76	6.80	4.61	0.695	–0.099
ROSTOCK	GERMANY	52	2.73	4.58	3.93	0.747	–0.035
SABVIZAR	IRAN	91	2.04	7.83	6.31	–	–0.209*
SIENE	ITALY	61	2.39	5.40	4.71	0.760	–0.176
SOEST (#1)	GERMANY	58	2.52	5.49	4.40	0.673	–0.123
SOEST (#2)	GERMANY	45	2.44	5.04	4.26	0.600	–0.176
UDINE	ITALY	87	2.87	5.92	4.23	0.823	–0.039
VERONA	ITALY	93	2.86	6.19	4.31	0.782	–0.019
ZARAGOZA	SPAIN	106	2.43	7.17	5.24	0.585	–0.170*

summarizes the main characteristics of these networks. They are relatively small, ranging from 45 to 339 nodes and all defining a single connected component (i.e. there is a path through the graph connecting every two nodes).

The standard comparisons with theoretical models (such as the standard, Erdös-Renyi random graph [16]) often taken as references in the studies of Small-World and Scale-Free networks lose their relevance in the framework of planar graphs [17]. Unfortunately, there are few general analytic results (such as path lengths or degree distributions) on random planar graphs [18,19], and they essentially consist in asymptotic results difficult to apply on the finite size networks considered in this paper. Here, we will focus on a comparison with two extremal models, the Minimal Spanning Tree (MST, [20–22]) and the Greedy Triangulation (GT). Like the settlements analyzed here, they are planar graphs. Given the set of nodes V in the corresponding city network, the MST defines the shortest tree which connects all the nodes into a single connected component. We used the Kruskal's algorithm [21] to build MST, that is, edges are sorted in increasing order of their length, and added to the graph following this sorting order and only if they do not introduce a cycle. A triangulation defines a planar subdivision whose bounded faces are triangles and whose vertices are points of V [23]. There are several ways for computing the triangulation, the GT being a very simple case of triangulation of a planar point set [24]: it can be constructed by sorting all edges by increasing order of their length, and, browsing the list of edges in this order, inserting each edge only if it does not produce any intersections with any other edge already added.

3 Topological patterns

3.1 Degree and degree correlations

The characteristic feature of the degree distributions is that they are single-scaled (Fig. 2), i.e. they follow a fast decay that can be approximated by an exponential decay $P(k) \sim e^{-k/\xi}$. Table 2 shows the parameter values ξ in the networks where the tails were long enough. Degree correlations were estimated by calculating the "assortativity coefficient" Γ as proposed by Newman [25]. It is defined as:

$$\Gamma = \frac{c\sum_i j_i k_i - \left[c\sum_i \frac{1}{2}(j_i + k_i)\right]^2}{c\sum_i \frac{1}{2}(j_i^2 + k_i^2) - \left[c\sum_i \frac{1}{2}(j_i + k_i)\right]^2} \tag{1}$$

where j_i and k_i are the degrees of the vertices at the ends of the ith edge, with $i = 1, \ldots, m$, $c = 1/m$ and m being the number of edges. When $\Gamma = 0$, nodes are connected independently from their neighbor degree. $\Gamma > 0$ indicate an assortative network where nodes with a given degree connect preferentially with nodes having similar degrees, while $\Gamma < 0$ corresponds to disassortative networks where nodes connect preferentially with nodes having a different degree to them. Most street networks appear to be disassortative, among which 17 are significant, and three networks have a positive but non-significant assortativity coefficient (Tab. 1).

3.2 Structure of cycles and path system

Street networks show no or little clustering (i.e. a low probability of triangles). This is expected, given the tendency of streets to be parallel, at least on local scales. Intersections of pairs of parallel streets will tend to enhance the presence of cycles with four elements, as expected from a mesh-like system. Since the SNS are planar graphs, we can easily quantify the amount of cycles in these networks: using Euler's formula, the number of faces f (excluding the infinite face) associated with any planar graph is

$$f = m - n + 1, \tag{2}$$

where m is the number of edges and n the number of nodes. As the number of edges m of a planar graph with n nodes is bounded by $3n - 6$, the number of faces is

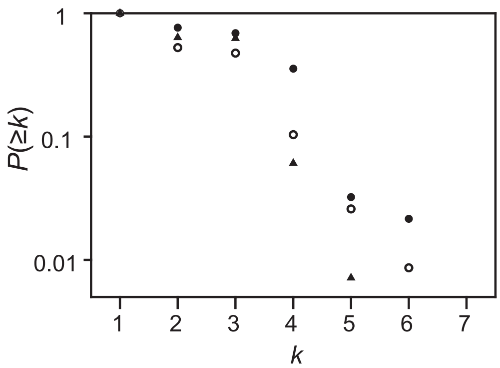

Fig. 2. Cumulative degree distribution of 3 examples of street networks: Verona (closed-circles), Martina Franca #2 (open-circles) and Benares #3 (triangles). The distributions are single-scaled, i.e. they decrease very fast from $k = 3$. The distributions tails (for $k \geq 3$) can be approximated by an exponential decay.

bounded by $2n - 5$. We can thus compute a normalized "meshedness coefficient":

$$M = \frac{m - n + 1}{2n - 5}, \tag{3}$$

where M can vary from zero (tree structure) to one (complete planar graph, which is a triangulation).

None of the networks analyzed here is acyclic, but the meshedness M shows considerable variability (Tab. 2) from $M = 0.009$ (El Agustino), which corresponds to a structure very close to a tree (only 2 cycles in El Agustino), to $M = 0.211$ (Qazvin #2, Iran), which corresponds to almost grid-like structures. In these networks, M appears to be independent of the network size n ($r = -0.179$, $N = 41$, $p = 0.263$).

4 Network efficiency

The characteristics of the path system can be obtained by analyzing the shortest paths between all pairs of nodes. Street networks are spatially extended webs, it is thus possible to work with two types of distance measures for the shortest paths: topological and geometrical. Topological path length is computed from the number of nodes the path is going through; while the geometrical path length is the sum of the lengths of all edges the path is going through. Recently, Latora and Marchiori [26] proposed the so-called "average efficiency":

$$E(G) = \frac{1}{n(n-1)} \sum_{v_i \neq v_j \in V} \frac{1}{d^*_{ij}}, \tag{4}$$

where d^*_{ij} corresponds to the shortest path between the nodes i and j. For a given graph G, the so-called global efficiency E_{glob} is determined by computing the ratio between the average efficiency measure $E(G)$ for all paths in the graph G,and $E(K_n)$, the average efficiency for all paths in the complete graph of order n (it possesses the

Table 2. Relative geometric and topologic global efficiency $E^*_{glob,G}$ and $E^*_{glob,T}$ respectively; Meshedness coefficient M; robustness to random removal of nodes f_R and selective removal of highest degrees f_S, relative robustness f^*_R and relative length L^*. Settlements are sorted in the decreasing order of their $E^*_{glob,G}$ value. Note that the two lower values of $E^*_{glob,G}$ are negative, which indicates that these networks have lower efficiencies than would have the MST.

Settlement	$E^*_{glob,G}$	$E^*_{glob,T}$	M	f_R	f_S	f^*_R	L^*
AMIENS (XIXth)	0.625	0.447	0.176	0.283	0.157	0.550	0.235
ROSTOCK	0.621	0.462	0.182	0.327	0.173	0.598	0.228
VERONE	0.619	0.508	0.177	0.301	0.161	0.569	0.237
LUBECK	0.614	0.492	0.209	0.317	0.167	0.600	0.219
SOEST (#1)	0.580	0.389	0.126	0.276	0.172	0.488	0.150
QAZVIN (#1)	0.559	0.478	0.193	0.287	0.168	0.548	0.232
UDINE	0.553	0.502	0.188	0.299	0.161	0.579	0.242
ARLES (#1)	0.540	0.450	0.095	0.273	0.145	0.517	0.187
QAZVIN (#2)	0.536	0.474	0.211	0.269	0.115	0.515	0.245
HOMS (#1)	0.533	0.510	0.201	0.275	0.148	0.555	0.238
DAX	0.525	0.461	0.165	0.309	0.145	0.553	0.204
ROME	0.496	0.485	0.164	0.287	0.176	0.569	0.237
CORDOBA (#3)	0.480	0.378	0.157	0.244	0.122	0.409	0.184
MADRIGAL	0.450	0.445	0.157	0.286	0.127	0.533	0.216
SOEST (#2)	0.447	0.358	0.106	0.267	0.133	0.447	0.157
CAHORS	0.442	0.419	0.146	0.245	0.126	0.487	0.212
SIENE	0.413	0.390	0.086	0.279	0.098	0.480	0.139
CORDOBA (#2)	0.412	0.373	0.118	0.258	0.136	0.423	0.163
ZARAGOZA	0.401	0.408	0.123	0.245	0.085	0.464	0.155
BONHEIDEN (#1)	0.392	0.355	0.100	0.162	0.094	0.293	0.116
MARTINA F. (#1)	0.390	0.369	0.059	0.217	0.070	0.393	0.138
PUTIGNANO	0.382	0.412	0.058	0.195	0.090	0.381	0.129
BONHEIDEN (#2)	0.362	0.369	0.074	0.183	0.033	0.333	0.117
BENARES (#1)	0.350	0.382	0.130	0.177	0.121	0.350	0.192
ECIJA (#1)	0.342	0.313	0.088	0.243	0.114	0.415	0.122
GOSLAR	0.321	0.384	0.091	0.254	0.079	0.424	0.156
AT SJEN	0.309	0.364	0.063	0.191	0.075	0.350	0.147
GHARDAIA (#2)	0.307	0.340	0.046	0.186	0.068	0.347	0.117
BONHEIDEN (#3)	0.305	0.278	0.060	0.139	0.051	0.237	0.101
ECIJA (#2)	0.301	0.383	0.075	0.232	0.116	0.380	0.142
BEZIERS	0.291	0.377	0.079	0.222	0.078	0.376	0.163
CORDOBA (#4	0.288	0.391	0.088	0.246	0.092	0.446	0.177
ARLES (#2)	0.273	0.344	0.058	0.182	0.081	0.311	0.110
ALGIERS CASBAH	0.258	0.299	0.061	0.159	0.076	0.266	0.120
CORDOBA (#1)	0.252	0.325	0.075	0.218	0.090	0.350	0.111
MARTINA F. (#2)	0.251	0.385	0.037	0.181	0.052	0.317	0.108
GHARDAIA (#1)	0.245	0.272	0.027	0.141	0.088	0.218	0.068
BENARES (#3)	0.188	0.306	0.090	0.150	0.075	0.267	0.174
HOMS (#2)	0.134	0.415	0.093	0.219	0.088	0.403	0.169
EL AGUSTINO	-0.055	0.165	0.009	0.085	0.034	0.096	0.090
SABVIZAR	-0.119	0.273	0.034	0.143	0.055	0.144	0.104

same vertices as G, but with all the $n(n-1)/2$ possible edges):

$$E_{glob} = E(G)/E(K_n). \tag{5}$$

We can compute two different measures of global efficiency, $E_{glob,G}$ and $E_{glob,T}$, whether we use geometrical or topological distances respectively.

The geometric efficiency $E_{glob,G}$ also varies much from low values (minimal $E_{glob,G} = 0.4$ in El Agustino) to values close to the one obtained in a complete graph (maximal $E_{glob,G} = 0.837$ in Rostock, Germany; see Tab. 2). $E_{glob,G}$ does not vary with the network size n in GT networks (Fig. 3a), while it decreases with the increase of n in MST networks. All the SNS display intermediate values of $E_{glob,G}$ comprised between the one of MST and GT networks, with the exception of two settlements (El Agustino and Sabvizar) that have even lower values for $E_{glob,G}$ than the corresponding MST. In the SNS, $E_{glob,G}$ decreases quite similarly to MSTs with size n, but with a lot of variability between networks of similar sizes. This variability may be due to the network meshedness. We compute the relative value of the geometric global efficiency,

$$E^*_{glob,G} = \frac{E^S_{glob,G} - E^{MST}_{glob,G}}{E^{GT}_{glob,G} - E^{MST}_{glob,G}}, \tag{6}$$

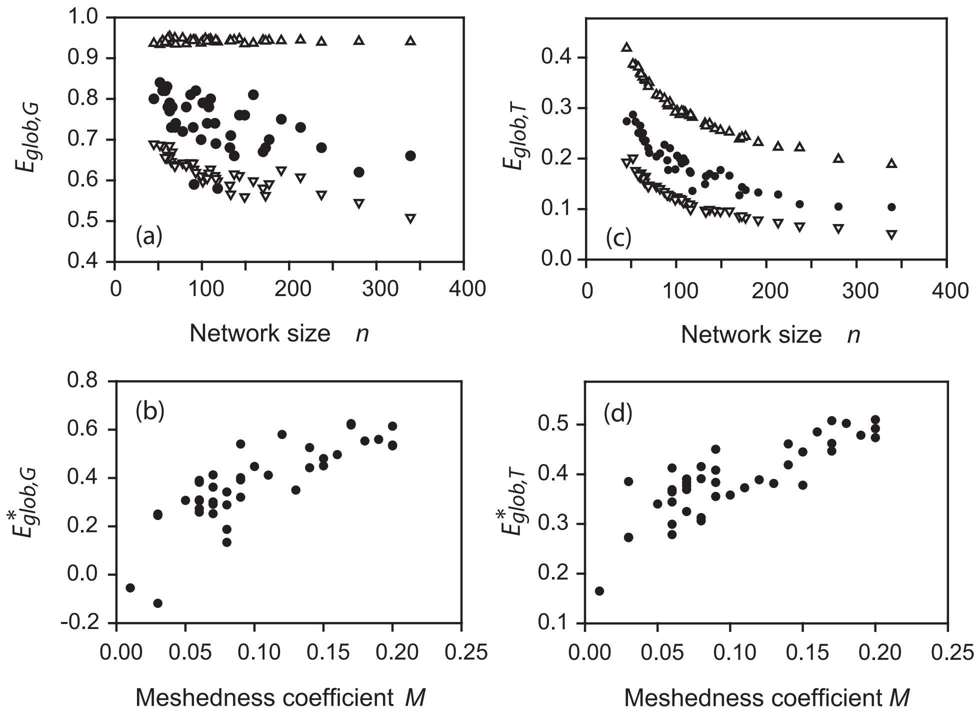

Fig. 3. Relation between geometric (a), topologic (c) global efficiency $E_{glob,G}$ and$E_{glob,T}$ and network size n. $E_{glob,G}$ is independent of n in GT networks (upward triangles), while it decrease with an increase in n for both settlement networks (circles) and MST networks (downward triangles). $E_{glob,T}$decrease with an increasing network size n for all networks. However, much variability is present in both $E_{glob,G}$ and $E_{glob,T}$. The relative efficiency, both geometrical $E^*_{glob,G}$ (c) and topological $E^*_{glob,T}$ (d) is positively correlated with network meshedness M.

where $E^S_{glob,G}$, $E^{MST}_{glob,G}$ and $E^{GT}_{glob,G}$ is the geometric global efficiency in the SNS, MST and GT networks respectively. Indeed, there is a clear correlation between $E^*_{glob,G}$ and $M(r = 0.82; N = 41, p < 0.001$, Fig. 3b), while there is no significant correlation between $E^*_{glob,G}$ and n ($r = -0.236$, $N = 41$, $p = 0.137$). Settlements that are more meshed have more efficient path systems at the geometric level.

The topological global efficiency $E_{glob,T}$ also shows much variability ranging from 0.104 to 0.287 in the SNS. It decreases with an increase in the network size n for SNS, GT and MST networks (Fig. 3c). However, $E_{glob,T}$ also exhibits a certain variability for similar network size n in the SNS. As for geometric efficiency, this variability may be due to the network meshedness. We thus compute the relative value of $E^*_{glob,T}$ normalized between the one observed in the MST and in the GT networks:

$$E^*_{glob,T} = \frac{E^S_{glob,T} - E^{MST}_{glob,T}}{E^{GT}_{glob,T} - E^{MST}_{glob,T}}, \tag{7}$$

where $E^S_{glob,T}$, $E^{MST}_{glob,T}$ and $E^{GT}_{glob,T}$ is the topologic global efficiency in the SNS, MST and GT networks respectively.

Again, there is a clear correlation between meshedness M and $E^*_{glob,T}$ ($r = 0.83$, $N = 41$, $p < 0.001$; Fig. 3d) while there is no significant correlation between $E^*_{glob,T}$ and n ($r = -0.267$, $N = 41$, $p = 0.09$). Therefore, in the SNS, both higher topologic and geometric path system efficiencies are reached by increasing the meshedness of the network.

5 Network robustness

Beyond the efficiency associated to a given network topology, an additional and complementary approach is the analysis of fragility against random failures. The nature of such failures depends on the particular kind of network considered: it might consist in mutations in gene networks, failures of routers in the Internet or species loss in an ecosystem. For urban networks, failures can consist in streets interrupted by landslides, floods, crumbled buildings, or by more temporary events such as demonstrations. The *robustness* of a network is measured by studying how it becomes fragmented as an increasing fraction of nodes is removed. The network fragmentation is usually measured as the fraction of nodes contained by the largest connected component. This node removal can take place either randomly or in decreasing order of their degree (selective removal). In homogeneous random graphs, the fragmentation of the network is similar under random and selective removal of nodes [27,28]. Several real networks have been reported to deviate clearly from this prediction of random graph theory and to exhibit a high resilience to random removal and high vulnerability to selective removal of nodes [27–30]. This property was first proposed as a unique feature of scale-free networks [27]. However several food-webs that are not scale-free networks also exhibit this property, and it has thus been conjectured that it could come from a more general feature of the degree distribution, such as its degree of asymmetry [31] or degree correlations [25].

Here we explore the effects of node removal in SNS and compare them with both MST and GT networks.

For each experiment, we have determined the evolution of the relative size S of the largest connected component, i.e. the fraction of nodes contained in it, with the fraction f of disconnected nodes (1000 runs for each network) under a random or a selective removal of nodes. We define the random and selective *robustness* (f_R and f_S, respectively) as the values of f required for S to reach the value $S = 0.5$ in each type of removal.

The fragmentation shows a common feature in all SNS (Fig. 4a): The decrease of S is clearly slower when the nodes are disconnected randomly than when higher degrees are disconnected first. This property is very different from what is observed in a classical random graph, where the decrease of S is similar under both random and selective removal of nodes. Indeed, the fragmentation of SNS is also quite different from the one observed in the corresponding MST and GT networks, where most of the fragmentation is similar under random and selective removal of nodes (Fig. 4b).

However, the speed at which fragmentation occurs shows a great variability among SNS, between $f_R = 0.085$ (El Agustino) to $f_R = 0.327$ (Rostock). There is a tendency for f_R to decrease with network size n, as it does in the MST networks, while in the GT networks f_R is independent from n (Fig. 4c). However, in the SNS, there still exists a strong variability for networks of equivalent size n.

We compute the relative robustness f_R^* between MST and GT networks defined as:

$$f_R^* = \frac{f_R^S - f_R^{MST}}{f_R^{GT} - f_R^{MST}}, \tag{8}$$

where f_R^S, f_R^{MST} and f_R^{GT} correspond to the random robustness in the settlement, the MST and the GT networks respectively. The relative robustness f_R^* is significantly correlated with several topological characteristics of the network: it is positively correlated with meshedness M ($r = 0.843$, $N = 41$, $p < 0.001$, Fig. 5a). This further supports the idea that the cycles may be a key feature for the network properties of robustness to disconnections. There is also a significant positive correlation between f_R^* and network assortativity Γ ($r = 0.624$, $N = 41$, $p < 0.001$; Fig. 5b), which adds weight to the conjecture of Newman [25] which states that disassortative networks may be more vulnerable to random disconnections.

The parameter ξ in the degree distribution tail, which reflects the skewedness of the tail, is also significantly positively correlated with f_R^* ($r = 0.685$, $N = 35$, $p < 0.001$, Fig. 5c). The more heterogeneous is the degree distribution, and the more robust to disconnection is the network, which gives support to the conjecture of [31]

A striking feature in SNS is that $E_{glob,G}$ and f_r are strongly correlated ($r = 0.915$, $N = 41$, $p < 0.001$). In our sample of settlements, the relation between path system efficiency and network robustness appears to be linear (Fig. 6). We have shown that this increase is clearly correlated with an increase in meshedness. However, increasing

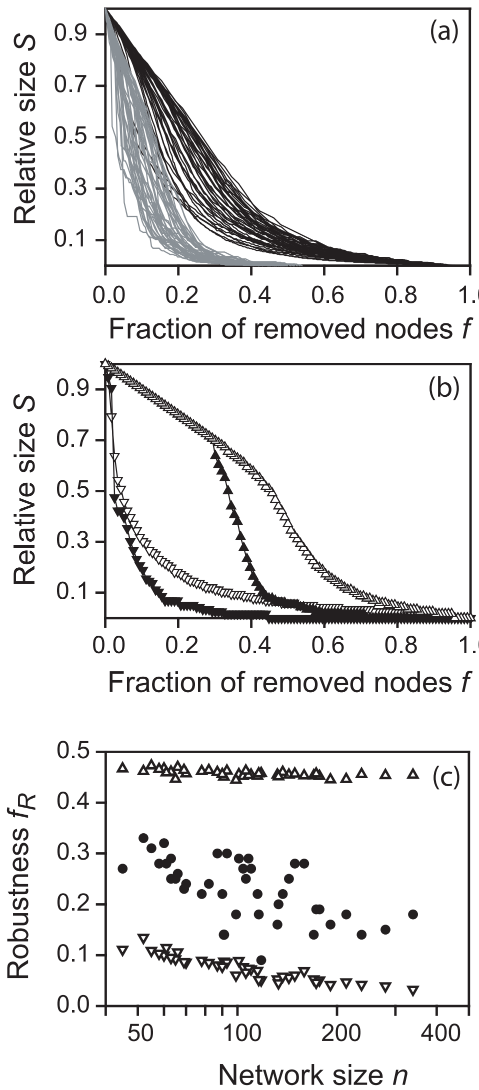

Fig. 4. (a) Settlement network fragmentation, measured by the relative size of the largest component S, under random (black curves) and selective (grey curves) removal of a fraction of nodes f. (b) Fragmentation of the MST (downward triangles) and triangulated network (upward triangles) corresponding to the Martina Franca #2 network under random (open symbols) and selective (closed symbols) removal of nodes. (c) Relation between the robustness to random node removal f_R and the network size n.

M means developing longer networks. We thus introduce a normalized measure of length, where the lower bound is represented by the length of the MST while the upper bound corresponds to the GT:

$$L^* = \frac{L_S - L_{MST}}{L_{GT} - L_{MST}}, \tag{9}$$

where L_S, L_{MST} and L_{GT} correspond respectively to the total length of edges in the SNS, the corresponding MST and GT networks.

The relative length of SNS varies from values close to the MST length (minimal observed $L^* = 0.068$ in Ghardaia #1) to values that did not exceed the third of

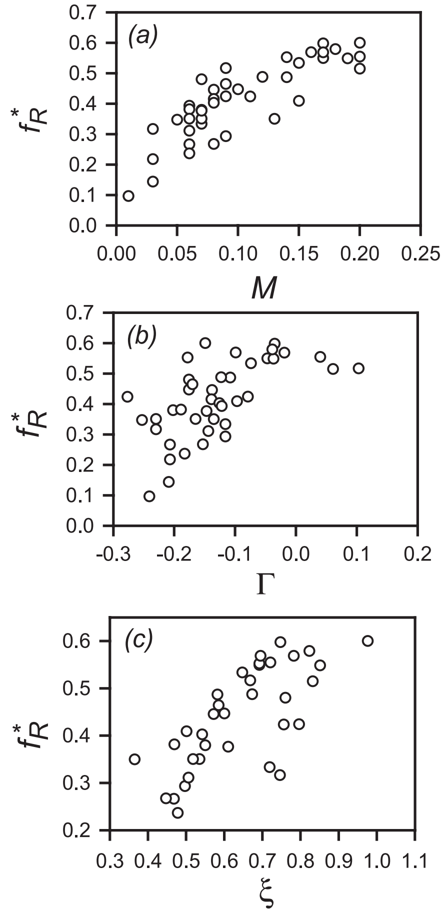

Fig. 5. Topological correlation with the relative robustness f_R^*. Correlation with the meshedness coefficient M (a), the assortativity coefficient Γ (b) and the skewedness of the degree distrubtion ξ (c).

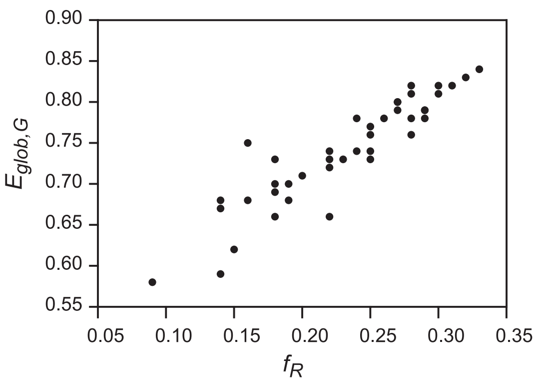

Fig. 6. Relation between robustness f_R and geometric global efficiency $E_{glob,G}$.

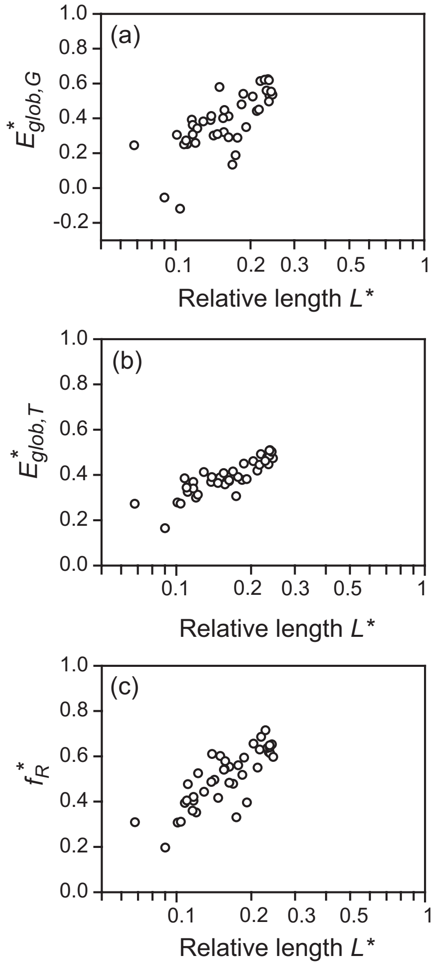

Fig. 7. Relation between the relative length L^* and the relative geometric global efficiency $E^*_{glob,G}$ (a), the relative topological global efficiency $E^*_{glob,T}$ (b), and the relative robustness f_R^* (c), The down-left corner (toward coordinates 0, 0) corresponds to the length, robustness and efficiency of the MST while the upper left corner (coordinates 1, 1) correspond to these characteristics for GT networks.

the triangulation length (maximal observed $L^* = 0.245$ in Qazvin #2). However, as relative global efficiency $E^*_{glob,G}$, $E^*_{glob,T}$ and relative robustness f_R^* increase faster than L^* (Fig. 7), high levels of path system efficiency and robustness can be reached with only a slight increase in L^*. From the relation that we observe, it seems that most of the gain in efficiency and robustness is achieved in the initial increase in length departing from a MST.

6 Discussion

Urban centers are one the first complex webs constructed by human societies. They involve several layers of complexity and provide the matrix for economic and social evolution. One of the most obvious levels of description deals with their static architecture, as revealed by their street maps. In some cases, such maps are highly ordered (such as in Manhattan or Barcelona cores) but many others (such as those analyzed here) clearly depart from the grid-like picture. In urban settlements, particularly at some small scales, these networks result from decentralized growth and are closer to other patterns observed in nature, such as ant tunneling networks [17]. In this paper we provide the first analysis of a large sample of self-organized SNS, which complements the recent analysis by Carvalho and Penn involving large-scale urban space organization [5]. The basic observations reported here are summarized as follows:

(1) Street networks are planar graphs and such planarity strongly constrains their heterogeneity, which has been shown to be rather limited (they exhibit single-scale distributions of links and are sparse, with $\langle k \rangle \approx 2.5$).
(2) The analysis of correlations reveals that most of these networks are diassortative and thus nodes with large degree tend to be unconnected among them.
(3) By means of Euler's formula, we defined a measure of meshedness, M, that quantifies the relative amount of faces in comparison with the two extremes that are trees and triangulations. Far from being ordered square meshes, many of the SNS that we have studied are closer from tree-like structures, with the denser SNS hardly reaching half of the cycles of a square grid. Our measure shows that such measures as the clustering coefficient (and any variants measuring short cycles only) would not allow capturing the characteristic presence of a limited amount of large cycles, though our results show that the presence of few of these cycles may have strong consequences on several network properties.
(4) Network efficiencies have been characterized using the Latora-Marchiori measure [26,32,33]. Though this measure deals with the length of shortest-paths between nodes rather than with actual trips and flows taking place on these paths, our results bring interesting insights about the relation between several network characteristics and their consequences on the path system, which may influence strongly the traffic and flows taking place on these networks. Indeed, efficiency displayed much variation among different networks, from values as low as trees to values close to triangulations, and correlated well with M. Thus the presence of multiple paths and the reduction of detours in SNS seem to be mainly achieved by a slight increase in meshedness.
(5) As shown for ecological systems exhibiting limited heterogeneity in degree distributions [30,31], the SNS exhibit fragility under selective removal. Given their spatially extended character and decentralized origins, the removal of high-k nodes has an important impact on network reliability. Conversely, removal of low-k nodes (the most common) has little effects. These effects of node removal are not exhibited by MST and GT networks, which behave similarly for both types of perturbation. The correlations between the level of robustness in a SNS and its meshedness suggest that these properties of the SNS may be related to their cycle structure.

The local character of evolutionary rules implicit in the development of SNS has shaped them through their history. The patterns revealed by our study indicate that self-organized SNS display special features not shared by other standard random graphs. They differ from greedy triangulations and spanning trees. At one extreme, we have SNS corresponding to an assembling of tortuous tree-like structures. From this extreme, there exists a whole range of SNS that are characterized by a progressive and parallel increase in meshedness and heterogeneity in the degree distributions. These progressive changes in the SNS topology lead to a web that is robust against random failure and yet fragile under removal of key nodes.

Remarkably, the street networks studied here share most of these structural properties with at least one other class of planar networks in nature, namely the ant tunnelling networks [17]. In particular, both types of networks exhibit a similar relationship between cost and efficiency: the efficiency increases sharply as soon as a few cycles are introduced, while it saturates before the number of cycles reaches values of grid-like patterns. The same type of relationship has also been recently observed in another set of street networks studied by Cardillo et al. [34]. Their sample spanned networks with higher relative costs than the one studied here, and the authors were able to estimate that the saturation of the relative efficiency around a value of 0.8 for relative costs higher than 0.25.

It can be noted that some of the street networks studied here exhibit at least one feature that was never observed in tunnelling networks: diassortativity. This property seems to be associated with street patterns standing between tree-like networks and mesh networks. In these networks, diassortativity might be explained by the existence of hierarchical like tree structures partially merged.

What are the factors associated with these variations in the SNS topology? One can speculate that the poor global efficiency of several SNS may be associated with a different use of the path system (e.g. different transportation modes) or different distribution of trips among the network, which may compensate the presence of so strong detours. A possible correlation between a particular network topology and the actual use of these SNS networks remains to be studied.

How does these factors relate to the network of relationships and collaborations between the agents that built the network, in a particular context? Understanding the link between them may bring new insights about the evolution of the urban networks, and more generally of complex networks whose growth occurs in planar constraints.

The authors would like to thank V. Fourcassié, C. Jost and J.L. Deneubourg for many helpful suggestions and comments. J.B. was supported by a doctoral grant from the French Ministry of Scientific Research. This work has been supported by the "Programme Cognitique" from the French Ministry of Scientific Research (J.B., J.G., G.T., N.R., P.K.), by the grant BFM2001-2154 (RVS, SV) and the Santa Fe Institute (RVS).

References

1. M. Batty, P. Longley, *Fractal cities: a geometry of form and function* (Academic Press, London, San Diego, 1994)
2. S. Kostov, *The city shaped: Urban patterns and meaning through history* (Thames and Hudson, London, 1991)
3. H.A. Makse, J.S. de Andrade, M. Batty, S. Havlin, H.E. Stanley, Phys. Rev. E **58**, 7054 (1998)
4. S. Manrubia, D.H. Zanette, R.V. Solé, Fractals **7**, 1 (1999)
5. R. Carvalho, A. Penn, Physica A **332**, 539 (2004)

6. These non-globally-planned urban morphologies were defined and regrouped by N. Reeves under the acronym Mesap for "Morphologies Evolutives Sans Adressage Préalable" (Evolving morphologies without pre-defined addresses)
7. F. Schweitzer, W. Ebeling, H. Rosé, O. Weiss, Evol. Comp. **5**, 419 (1998)
8. M.T. Gastner, M.E.J. Newman, e-print `arXiv:cond-mat/0409702`
9. M.T. Gastner, M.E.J. Newman, e-print `arXiv:cond-mat/0407680`
10. F. Schweitzer, *Brownian Agents and Active Particles* (Springer, Berlin, 2001)
11. M. Rosvall, A. Trusina, P. Minnhagen, K. Sneppen, Phys. Rev. Lett. **94**, 028701 (2005)
12. B. Jiang, C. Claramunt, Env. Planning B **31**, 151 (2004)
13. V. Kalapala, V. Sanwalani, A. Clauset, C. Moore, e-print `arXiv:physics/0510198`
14. T. Nishizeki, N. Chiba, *Planar Graphs; Theory and Algorithms* (North-Holland, Amsterdam, 1988)
15. H. Caminos, J. Turner, J. Steffian, *Urban Dwelling Environments: An Elementary Survey of Settlements for the Study of Design Determinants* (MIT Press, Cambridge, Massachussetts, 1969)
16. B. Bollobas, *Random graphs,* 2nd edn. (Cambridge University Press, Cambridge, 2002)
17. J. Buhl, J. Gautrais, R.V. Solé, P. Kuntz, S. Valverde, J.L. Deneubourg, G. Theraulaz, Eur. Phys. J. B **42**, 123 (2004)
18. A. Denise, M. Vasconcellos, D.J.A. Welsh, Congressus Numerantium **113**, 61 (1996)
19. D. Osthus, H.J. Promel, A. Taraz, J. Comb. Theory B **88**, 119 (2003)
20. E.W. Dijkstra, Numer. Math. **1**, 269 (1959)
21. J.B. Kruskal, Proc. Amer. Math. Soc. **2**, 48 (1956)
22. D. Cheriton, R.E. Tarjan, SIAM J. Computing **5**, 724 (1976)
23. M. de Berg, M. van Kreveld, M. Overmars, O. Schwarzkopf, *Computational geometry*, 2nd rev. edn. (Springer, Berlin, 2000)
24. C. Levcopoulos, A. Lingas, Algorithmica **2** (1987)
25. M.E.J. Newman, Phys. Rev. Lett. **89**, 208701 (2002)
26. V. Latora, M. Marchiori, Phys. Rev. Lett. **87**, 198701 (2001)
27. R. Albert, H. Jeong, A.L. Barabasi, Nature **406**, 378 (2000)
28. P. Holme, B.J. Kim, C.N. Yoon, S.K. Han, Phys. Rev. E **65**, 056109 (2002)
29. H. Jeong, S.P. Mason, A.L. Barabasi, Z.N. Oltvai, Nature **411**, 41 (2001)
30. R.V. Solé, J.M. Montoya, Proc. R. Soc. Lond. B **268**, 2039 (2001)
31. J.A. Dunne, R.J. Williams, N.D. Martinez, Ecol. Lett. **5**, 558 (2002)
32. V. Latora, M. Marchiori, Physica A **314**, 109 (2002)
33. V. Latora, M. Marchiori, Eur. Phys. J. B **32**, 249 (2003)
34. A. Cardillo, S. Scellato, V. Latora, S. Porta, e-print `arXiv:physics/0510162`
35. D.J. Watts, *Small Worlds: The Dynamics of Networks Between Order and Randomness* (Princeton University Press, Princeton, 1999)