Structure Entropy for Urban Street Networks

Jérôme Benoit and Saif Eddin Jabari

New York University Abu Dhabi, Abu Dhabi, United Arab Emirates

jerome.benoit@nyu.edu

Objective and General Approach

Our primary aim is to find patterns among Urban Street Networks (USN) to gather them in categories.
Our approach envisages USN as evolving social structures subject to an entropic equilibrium similar to the one effectively observed among cities of a same country with respect to their sizes [1].

Overview

While our *raw material* are data extracted from data/map comprehensive archive networks as provided for instance by the OpenStreetMap (OSM) project, our investigation is based on well-known properties among USN that are sketched in a preliminary section.
The entropy is computed within the framework of physics information theory [2]. Accordingly, the concrete structures of USN must be structured as abstract algebraic ordering structures known as Galois lattices in order to apply the involved formalism: it appears that the Galois lattices for USN take a very intuitive form. As a result, the structure entropy for USN can be easily expressed: an explicit formula is given.
Then, we apply the entropic equilibrium model employed in [1]. Hypothesizing a crude asymptotic agent based model along the spirit of the city model [1] allows us to predict two power-law distributions: a pure power-law distribution for *"named streets"* as already reported in literature, and a generalized power-law for street junctions.
We believe that the ratio of the two scaling exponents may allow us to categorize USN.

Geometrical vs. Topological Urban Street Networks
Trivial vs. Non-Trivial Complexities

Street Junctions and Street Segments

From the raw material, *street junctions* and *street segments* emerge spontaneously; see Figure 1a for illustration.

Geometrical Networks: Trivial Complexity

Street junctions and street segments form, respectively, the natural nodes and edges of abstract networks, known as *geometrical networks*; see Figure 1c. Nonetheless, their complexity is trivial: three or four edges for most nodes [3]. Without surprise, the complexity of their dual counterpart the *segment-segment geometrical networks*, is also trivial: four, five or six edges for most nodes; see Figure 1e.

Natural Roads

We (humans) rather reason in terms of *streets* than of street segments. *Named streets* are results of intricate social processes. A *natural road* is an exclusive sequence of successive street segments paired according to some *behavioural based join principle*. Besides the cadastral approach, geometrical ones based on deflection angles are used: a junction centric one (`every-best-fit`) which is almost deterministic; two segment centric ones (`self-best-fit`, `self[-random]-fit`) which have appeared realistic (the latter being the best `fit`) [3]. See Figure 1b for illustration.

Topological Networks: Non-Trivial Complexity

We also plan our trajectories rather in terms of streets than of junctions. A *road-road topological network* reduces natural roads to nodes and links each pair that shares a junction — see Figure 1d. In general, for cadastral and segment centric join principles, road-road topological networks show intelligence behaviours as observed in some technological, biological or social networks [4]; in particular, they generally exhibit small-world and scale-free properties [3, 4].
Dually, *junction-junction topological networks* can be built — see Figure 1f. Nevertheless, their apparent intricacy renders them less appealing to us (human).

Canonical Natural Roads

We will use the adjective *canonical* whenever any involved junction joins at most two natural roads. In Figure 1, the junction j_7 is the only noncanonical junction since it is the only junction that joins together more than two natural roads.

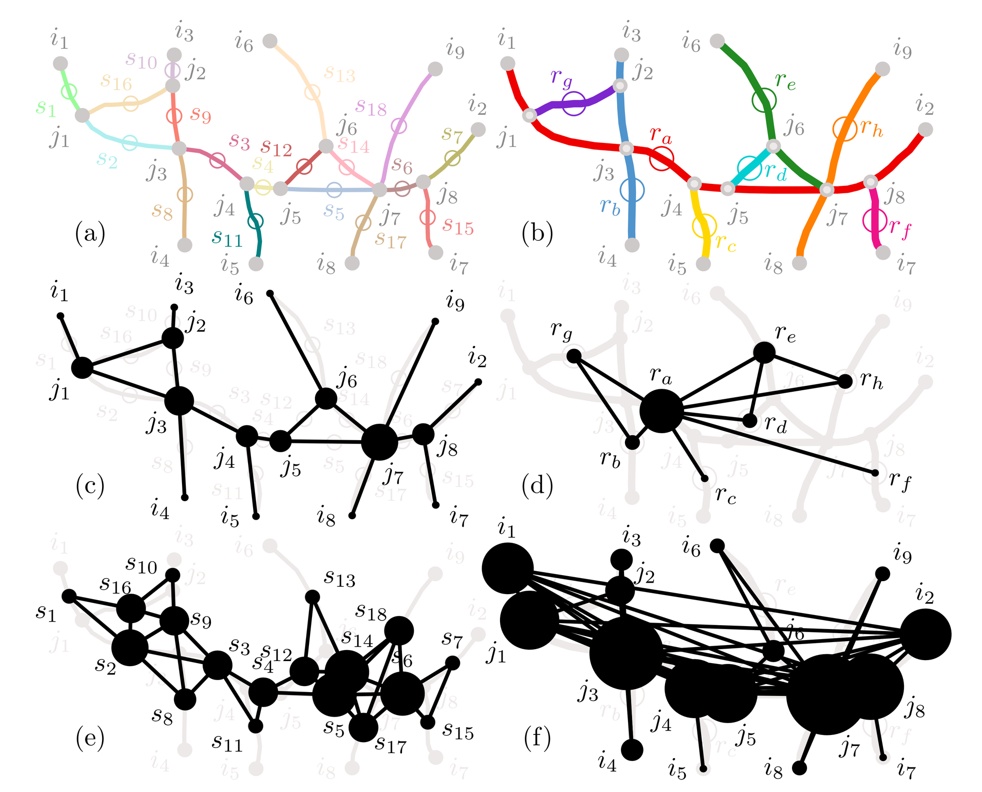

Figure 1: **A *notional urban street network*† and its representations.** *The left column represents the geometrical variants: (a) artificially colored 'raw material' graph displaying street [extended-]junctions (impasses and effective junctions) and segments in grey and pallid colors, respectively; (c) junction-based connectivity graph; (e) segment-based connectivity graph dual to the graph drawn in (c). The right column shows the topological variants for a same generated design: (b) revamped 'raw material' graph exhibiting original junctions and generated natural roads in grey and vivid colors, respectively; (d) natural-road-based connectivity graph; (f) junction-based connectivity graph dual to the graph drawn in (d). Notice that the size of each node is proportional to its valence for the abstract networks (c), (d), (e) and (f).*
† *Inspired by the 'notional road network' introduced in [3].*

Urban Street Galois Lattices
'Structure before measure'

Incidence Relations: Road vs Junction Charts

The construction of any topological network is decomposable in two steps. First step, set the *incidence relation* λ that gathers for each natural road all junctions through which it passes; in practice, λ is better represented by a road/junction cross/dot-chart as exemplified in Table 1. Second step, the adjacency matrix of the road-road (junction-junction) topological network is given by $\lambda \cdot \lambda^{\mathsf{T}}$ ($\lambda^{\mathsf{T}} \cdot \lambda$) where λ is represented as a (0,1)-binary matrix.

λ	j_1	j_2	j_3	j_4	j_5	j_6	j_7	j_8	i_1	i_2	i_3	i_4	i_5	i_6	i_7	i_8	i_9
r_a	●	·	●	●	●	·	●	●	●	●	·	·	·	·	·	·	·
r_b	·	●	●	·	·	·	·	·	·	·	●	●	·	·	·	·	·
r_c	·	·	·	●	·	·	·	·	·	·	·	·	●	·	·	·	·
r_d	·	·	·	·	●	●	·	·	·	·	·	·	·	·	·	·	·
r_e	·	·	·	·	·	●	●	·	·	·	·	·	·	●	·	·	·
r_f	·	·	·	·	·	·	·	●	·	·	·	·	·	·	●	·	·
r_g	●	●	·	·	·	·	·	·	·	·	·	·	·	·	·	·	·
r_h	·	·	·	·	·	·	●	·	·	·	·	·	·	·	·	●	●

Table 1: ***Notional road/junction cross/dot-chart*** *associated to the natural roads generated in Figure 1b.*

Thus the topological networks appear as incomplete, rough representations of incidence relations.

Galois Lattices: Holistic Representation

Relying on the Formal Concept Analysis (CFA) paradigm [5] allows us to plainly represent incidence relations as algebraic structures known as *Galois lattices* [6, 5]. For natural roads, Galois lattices appear to take a very intuitive form composed of two nontrivial layers [7]: the natural roads form the lower one; the junctions the upper one; the *'imply'* ordering relation is "passing through". See Figure 2 for illustration.

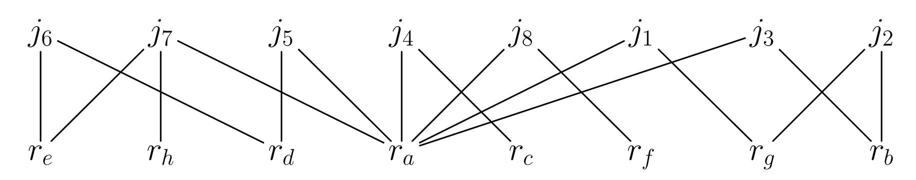

Figure 2: ***Notional (Galois) lattice*** *associated to the natural roads generated in Figure 1b: for clarity, the bottom $\perp$ and top $\top$ elements are not shown.*

Physics Information Theory
'Measuring is the quantification of ordering'

Concerning Galois lattices, while the poset and algebraic perspectives are respectively structural and operational [5], the whole is measurable [2]: imposing natural *consistency constraints* enables us not only to evaluate lattices but also to supercede contemporary information measures.

Briefly, for canonical natural roads (natural roads whose junctions join at most two of them), the structure entropy H holds the functional formula [7]

$$H = \sum_{r} (h \circ w)\,(\mathrm{Va}(r)) + \sum_{j(r,s)} (h \circ w)\,(\mathrm{Va}(r) + \mathrm{Va}(s)) \qquad (1)$$

where the first summation runs over the natural roads r and the second over the junctions $j(r,s)$ joining the pair of natural roads (r,s); the unknown valuation function Va contains the physics of the natural roads, the unknown function w acts as a weight function, and $h\colon x \mapsto -x \ln x$ is the Shannon entropy function.

Entropic Equilibrium
Jaynes' Maximum Entropy Principle

$\ln$-Mean as Moment Constraint

By analogy with [8], let us see natural roads and junctions as balanced social clusters, so USN as balanced mosaics of balanced social clusters. Our initial ignorance yields on the clusters, so the entropy for each cluster c is $\ln(\Omega_c)$ with Ω_c its number of configurations [8]. Applying the *maximum entropy principle* with the $\ln$-mean $\sum_c \Pr(\Omega_c)\,\ln(\Omega_c)$ as moment constraint leads to the Shannon Lagrangian [1, 8]

$$\mathcal{L}\left(\{\Pr(\Omega_c)\}; \nu, \lambda\right) = -\sum_c \Pr(\Omega_c)\,\ln\left(\Pr(\Omega_c)\right) + \\ -(\nu - 1)\left[\sum_c \Pr(\Omega_c) - 1\right] - \lambda\left[\sum_c \Pr(\Omega_c)\,\ln \Omega_c - a_\lambda\right] \qquad (2)$$

where the sums run over the natural roads and junctions c; ν and λ are Lagrangian multipliers. Resolving (2) gives

$$\Pr(\Omega) = \Omega^{-\lambda}\, Z^{-1} \quad \text{with} \quad Z = \sum_c \Omega_c^{-\lambda} \qquad (3)$$

which is a pure power law distribution.

Asymptotic Agent-Based Model

Then, let each cluster to be a hive of agents [1] (drivers, cyclists, pedestrians, suppliers, institutional agents, residents, and so forth) whose very existence relies on the ability for each of its agents to maintain a crucial number of intraconnections crudely equal to a constant number ν, called the *number of vital connections* [1]. For each natural road, being extensive, the number of agents is assumed crudely proportional to its number of junctions; we write

$$\Omega_r = \Omega_r(n_r) \simeq \binom{\frac{1}{2} A\, n_r\, (A\, n_r - 1)}{\upsilon_r} \simeq \frac{A^{2\upsilon_r}}{2^{\upsilon_r} \upsilon_r!}\, n_r^{2\upsilon_r}. \qquad (4a)$$

For each junction, along this spirit, the involved agents are merely the agents of the two joining natural roads combined together; the same crude maneuvers hold

$$\Omega_{j(r,s)} = \Omega_j\left(n_j = n_r + n_s\right) \simeq \frac{A^{2\upsilon_j}}{2^{\upsilon_j} \upsilon_j!}\, n_j^{2\upsilon_j}. \qquad (4b)$$

Thus, the valuation function Va appears to assign to each cluster the number of associated agents while the weight function w asymptotically counts the number of possible vital intraconnection layouts in the involved intranetwork — modulo normalization. Injecting (4) into (3) recovers the pure power-law effectively observed for road-road topological networks [3, 4] and foresees a generalized power law for junction-junction topological networks:

$$\Pr(n_r) \propto n_r^{-\upsilon_r \lambda} \qquad (5a)$$

$$\Pr(n_j) \propto \mathrm{J}(n_j)\; n_j^{-\upsilon_j \lambda} \qquad (5b)$$

where $\mathrm{J}(n_j)$ is the counting function $\sum_{j(r,s)} [n_j = n_r + n_s]$, with Iverson bracket convention; n_r (n_j) is essentially the degree associated to the natural road r (the junction j) in the involved road-road (junction-junction) topological network, while υ_r and υ_j are *a priori* distinct. Preliminary investigations [7] shows evidences for (5b).

Conclusion

A systematic, holistic approach allows us to predict two power law distributions for USN, one already reported in literature [3, 4] and a second under promising investigations [7]. This result opens a perspective on characterizing USN by their pair of scaling exponents.

[1] Dover Y 2004 *Physica A* 334(3-4) 591.
[2] Knuth KH 2011 *AIP Conf. Proc.* 1305(1) 3.
[3] Jiang B, Zhao S and Yin J 2008 *J. Stat. Mech. Theor. Exp.* 2008(7) P07008.
[4] Jiang B, Duan Y, Lu F, Yang T and Zhao J 2014 *Environ. Plan. B* 41(5) 813.
[5] Davey BA and Priestley HA *Introduction to Lattices and Order* 2nd ed.
[6] Ho YS 1982 *Environ. Plan. B* 9(4) 397.
[7] Benoit J and Jabari SE 2018 in preparation .
[8] Kesavan HK 2009 *Jaynes' Maximum Entropy Principle* 2nd ed. 1779–1782.

AGILE 2018 - 21st Conference on Geo-information science - Lund 12-15 June (Sweden)

https://orcid.org/0000-0003-1226-6757