Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 272 (2025) 61–67

Procedia Computer Science

www.elsevier.com/locate/procedia

The 16th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2025)
October 28-30, 2025, Istanbul, Türkiye

Urban Mobility and the Multiscale Structure of Road Networks

Alberto Amaduzzi[a,1,1], Filippo Dalla[a,1], Lorenzo Di Meco[a,1], Gregorio Berselli[a,1], Davide Micheli[d], Aldo Vannelli[d], Armando Bazzani[a,1,*]

[a]*Dipartimento di Fisica e Astronomia "Augusto Righi", Universita di Bologna, via Irnerio 46, Bologna, Italy*
[b]*INFN, Sezione di Bologna, via Irnerio 46, Bologna, Italy*
[c]*Fondazione Bruno Kessler, Via Santa Croce 77, Trento, Italy*
[d]*TIM S.p.A., Roma, Italy*

Abstract

The development of future smart cities has driven a new research field where different disciplines cope with the complexity of modern cities: Science of Cities. In particular, urban mobility has become one of the most intensely studied topics, thanks to data sets provided by new Information and Communication Technologies (ICT), which make it possible to relate microscopic individual behavior to the statistical laws of urban mobility through the ubiquitous use of mobile phones. In the paper we cope with the problems of pointing out the multilayer structure of an urban road network using a mobile phone GPS data set (the MDT Tim data set) and of building data driven stochastic models for urban mobility using a Maximum Entropy Principle approach. We show as the individual path reconstruction using the GPS data set allows to apply a clustering procedure that divides the urban road network of Bologna city into four connected classes that are related to a hierarchical use of the road network by individuals to perform the urban mobility, and explain the statistical laws on path length and path duration. These results gives possibility of studying the universal properties of congestion formation in an urban road network by using by data driven Markov model that are inferred applying the Maximum Entropy Principle approach of Statistical Mechanics. In this way one reduces the complexity of models and control parameter number and it is possible to study control strategies to limit the congestion spreading.

Keywords: ICT data sets, urban mobility, fuzzy clustering, Markov systems, Maximum Entropy Principle

1. Introduction

Urban mobility has become one of the most fruitful research field in Complex Systems Physics, thanks to the large georeferenced datasets provided by Information and Communication Technologies (ICT)[1], which allow to study the

* Corresponding author. Tel.: +39-0512091198.
E-mail address: armando.bazzani@unibo.it

1877-0509
10.1016/j.procs.2025.10.179

macroscopic statistical laws of urban mobility using the microscopic individual behavior. The challenge is to understand if the methods of Statistical Mechanics apply to Complex Systems or if the peculiar complexity features require a new Physics. The first problem is how to extract the relevant information from the large ICT data sets. In this paper we use the MDT (Minimization of Drive Test) data set provides by TIM (one of the major italian telecommunication company) that give georeferenced, anonymized information on the displacements of $\simeq 5\%$ of mobile devices in the metropolitan area of Bologna (an urban area in the North of Italy with $\simeq 900,000$ inhabitants) , with a time frequency of a few seconds (a typical time frequency is 5 seconds), when they're connected to the telecommunication network. More detailed on the data set can be found in[2]. The data quality is enough to reconstruct the individual paths during the mobility on the road network, but the sample penetration is not sufficient for a direct study of the urban mobility and the congestion problem in the road networks. The complex models for urban traffic (e.g. SUMO (https://eclipse.dev/sumo/)) simulate the vehicle dynamics and require a great amount of information that is not usually available. The main problem of urban mobility is the formation of congestion in the road network and its spreading. A recent paper[4] use realistic simulations to suggest that the congestion formation in a road network is a percolation phenomenon (i.e it has universal features)[5]. However the complexity of the models does not allow to understand how the individuals take advantage from the structure of the urban road network to realize the mobility paths and which are control parameters that could be modified to reduce the congestion effects.

To cope with this problem we propose a road map based on a theoretical frame work that uses the Maximum Entropy Principle (MEP) to infer the transition rates of a class of non-linear random walks (NRW) on network which simulate the congestion transition in a transport network. The application of entropy concept to urban systems has been proposed for long time[6, 7] as key of lecture of the city structure. Here we tries to use the dynamical Entropy concept to infer maximum entropy stochastic models for transport systems. Our idea is to generalize the Parry's theorem[8] that, starting from the connectivity matrix, computes the transition probabilities of a Markov process maximizing the dynamical entropy of the realizations. We propose an extension of this result to the transition rates con continuous processes that considers the application of the Minimum Entropy Production (MINEP) principle[13] to study the Non-equilibrium Stationary Sates (NESS)[9]. The advantage of this approach is that the knowledge of a weighted connectivity matrix for the transport network and of the average traffic loads and the average traffic flows allows building data driven models to study the congestion problem without detailed information on the Origin-Destination (OD) mobility demand in the city. The application of master equation to the social systems has been widely discussed in the literature[11]. The NRW models define queue dynamics models[12] able to simulate the traffic dynamics on road network where the evolution considers the existence of a finite maximum flow rate at intersection points and a finite road capacity. The apriori information on the road network structure is a key issue for the MEP application to infer a Markov model able to correctly describe the evolution out of equilibrium and the phase transitions. We show that the mobility paths reconstructed using the MDT data set, can be used to infer the transition probabilities at crossing points. In particular we highlight the existence of a multilayer structure for the road network that affects the individual behavior and allows to explain the empirical statistical distributions of urban mobility[3]. The study of the mobility paths in the homogeneous road sub-networks[10] confirm the exponential behavior for the path length distributions and support the application of the MEP to urban mobility.

The application of Markov processes to Statistical Mechanics requires has to interpret the relaxation process toward NESS[9] using the MINEP principle[13]. In the paper we show that the presence average unbalanced traffic flows can be included in the NRW models applying the MINEP principle without changing the stationary states. The reductionist features of the NRW models allow to relate the geometrical structure of the road network and the congestion formation through the spectral properties of the weighted connectivity matrix and suggest the control parameters of the transport system that could be used to perform control strategies. The NRW on networks are scalable to simulate different transport systems using a coarse grained description of the city. They define a theoretical framework that could help the interpretation of the data analytic results on the large ICT data sets and offer the possibility to evaluate the their predictive power using the validation procedures of Physics. The paper is organized as follows: in the second section we describe the results of the MDT path analysis of the city of Bologna; in the third section we discuss our maximum entropy approach applied to NRW on network to study the congestion formation in a transport systems: in last section we discuss the perspectives of our approach and the future work.

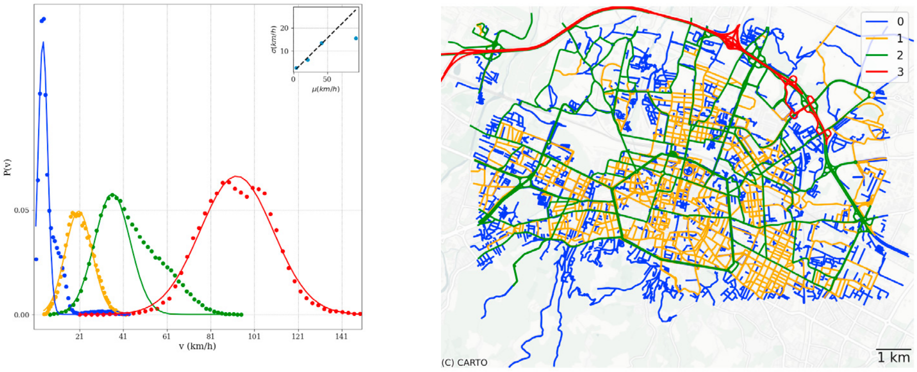

Fig. 1. (Left picture) The empirical distribution of travel speed of the roads belonging to the 4 classes: the continuous lines refers to an interpolation with a Gaussian distribution. The inset show the dependence of the on the mean square value on the average value. (Left picture) Map of the hierarchical multilayer structure of road network in Bologna as inferred using the MDT paths: each road class is denoted by a different color. The spatial scale is shown in the lower right corner of the figure.

2. MDT mobility paths

The MDT TIM data set (see[2] for the detailed information on the data) provide the GPS position of a sample ($\simeq 5\%$) of the mobile phone present in a large area when they are connected to the communication network. Specific algorithms have been developed to reconstruct individual mobility paths on the Bologna metropolitan area during several days(a detailed description of the algorithms can be found in[3]). The mobility paths give information on the individual dynamical during his mobility and one can study how the the structure of the transport network influences the individual behavior. We have analyzed the dynamic features of MDT paths to highlight the multilayer structure of Bologna's road network by applying a fuzzy clustering algorithm described in the paper[3] to define road sub-networks associated with different mobility features. This analysis groups the roads with similar travel speeds and traffic conditions identifying four mobility classes: the slow mobility (class 0), the traffic in the historical center (class 1), the traffic on the main urban roads (class 2), and highways traffic (class 3). In Fig. 1(left) we show the velocity distribution of each class computed from the pass. The Gaussian shape of the travel speed distribution suggests an homogeneous character of the road since the fluctuations could be explained as random effects. The variance of the travel speeds distributions could be explained both by the heterogeneity of individual dynamics but to the effect of vehicle interaction at different traffic load. Indeed in the inset of left plot in Fig. 1 we show the linear dependence of the mean square value on the average velocity for each distribution for the classes 0,1,2 whereas the fourth class (the highways) behaves in a different way. The linear dependence is a typical feature of the exponential probability distribution that could be the consequence of a queue dynamics. The geometrical structure of the sub-networks is shown in Fig. 1(right) where the roads belonging to different classes are denoted by different colors. The highways and the main roads of the Bologna road networks are clearly recognisable in the map The MEP of Statistical Physics implies that the distribution of an observable at equilibrium state is exponential[14]. The big data analysis of individual trajectories proved that in a stationary condition the the individual path distribution is exponential[15] and the power law distribution that has been initially proposed[16] can be explained by the multilayer structure of the transport networks[3]. We checked this feature for the MDT reconstructed paths in the metropolitan area of Bologna. In Fig. 2 we plot the the path length and the path duration distribution[10]. The path length distribution can be interpolated by an exponential law for all the days in the MDT data set, that suggests the existence of a typical length of $\bar{l} = 3.8$ km for the urban mobility in Bologna that can be related to the spatial size of the distance of the historical center from the periphery. The mobility between periphery and historical center seems at the base of the origin-destination mobility demand in Bologna. Conversely the path duration distribution is better fitted by a power law that can be explained by individual behaviors the use the different sub-networks to realize their OD mobility optimizing the travel time[10]. These results are consistent with a Maximum Entropy stationary state for the urban mobility, where the particles (i.e.

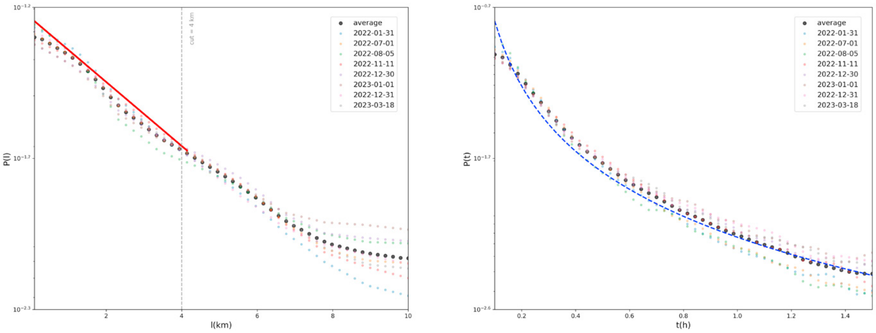

Fig. 2. (Left picture) MDT path lengths daily distribution in a semilog scale: the points refer to the different days in the MDT data set and the continuous line to an exponential interpolation with a characteristic length $\bar{l} = 3.8$ km. (Right picture
MDT path duration daily distribution: the points correspond to the different days in the MDT data set and the continuous line to a power law interpolation $p(t) \propto t^{-\alpha}$ with $\alpha = -1.36$.

the individuals) seem to move independently (i.e. minimizing their interactions) using the travel time as mobility energy[17].

3. Markov models of urban mobility and Maximum Entropy Principle

The Markov processes has been introduced in Statistical Physics as models to study the evolution of systems out of equilibrium[18]. The applicability of Markov processes to social systems has been widely discussed in the literature[11]. From one hand the Markov models cannot take into account the complexity of the social systems, but from the other hand the empirical stationary distributions are consistent with a MEP and the Markov processes can describe the relaxation process toward maximum entropy states. Our point of view is that Markov processes can be useful to study the dynamics of complex systems near maximum entropy stationary states and, in particular, the effect of fluctuations in triggering phase transitions. We associate to a transport system a NRW on network where each node represents an element of the transport system (e.g. a node can be a single road or a public mean station), and the links model the existence of a transport facility among the nodes (e.g. in the case of a road network the links simulate the dynamics at intersection points). We introduce a weight in the links to consider the different relevance of the connections among the nodes. In our reductionist approach, the urban mobility is the realization of the dynamics of many independent particles that perform a random walk on a transport network where the intersections has a maximum flow rate and node has a finite capacity. The dynamics is realized by moving particles among the nodes according to transition probability rates and the dynamical state of the system is given by the particle distribution among the nodes. The connectivity weighted matrix w_{ij} represents the relevance of the connection between the nodes (i, j) in the transport system (e.g. w_{ij} is greater if the nodes represents two main roads); we assume that the matrix w_{ij} is symmetric. To infer the weights w_{ij} one takes advantage from the MDT data analysis that highlights the structure of the road network (see Fig. 1).
The Parry ‘s theorem[8] states that if w_{ij} is a principal matrix (i.e. the network is connected) and v_i is the eigenvector corresponding to the maximum eigenvalue λ then the stochastic matrix

$$\pi_{ij} = \frac{v_i w_{ij}}{\lambda v_j}$$

defines a Markov process constrained by the connectivity matrix w_{ij}, which maximizes the dynamical entropy[19], whose stationary state is $p_i \propto v_i^2$.

We apply this result to time continuous Markov process in reverse way: if w_{ij} is the weight of the link that connects the nodes $j \to i$ and p_i is the stationary probability of an unknown Markov process, we define the transition rates

$$\pi_{ij}^{db} = \tau^{-1} w_{ij} \sqrt{\frac{p_i}{p_j}} \tag{1}$$

for the continuous Markov process that maximizes the information entropy according to the constraints: τ is the evolution time scale. We observe that the definition (1) implies that the spectral properties of the matrix π_{ij} are the same as the symmetric weight matrix w_{ij}: i.e. we not introduced any bias in the definition of the process. It is straightforward to check that p_i is the stationary solution of the process with transition rates (1) and that the process satisfies the detailed balance (DB) condition: i.e. the average traffic flow between any couples of connected nodes is the same in both directions when the system is in an equilibrium state. The DB condition implies that we can apply a MEP to compute the equilibrium state without solving the dynamics[9].
However the restriction to Markov systems that satisfy the DB condition cannot simulate transport systems with net stationary traffic flows among the nodes. This condition happens when we have the loops of the transport network that circulate along a given direction. In this case the stationary states are called NESS and they cannot be explained by a MEP[9]. In this case one can applying the MINEP principle by computing the entropy production of a Markov process as a function of the stationary probability currents J_{ij}^s and to modify the definition (1) applying the MINEP principle[13]

$$\pi_{ij} = \pi_{ij}^{db} + \frac{J_{ij}^s}{2p_j} \tag{2}$$

The stationary probability currents J_{ij}^s can be related to the stationary traffic flows in the network. Our conjecture is that the final system (3) is the Markov model that simulates the traffic dynamics near the equilibrium states maximizing the information entropy with the constrains given by the stationary state and the stationary flows.
The NRW model is defined by the master equation in the one step process approximation

$$\frac{\partial \rho}{\partial t}(\vec{n}, t) = \sum_{(i,j)} \Big(\pi_{ij} \theta_i^c(n_i - 1) \phi_j(n_j + 1) \rho(\vec{n} - \hat{e}_{ij}, t) - \pi_{ji} \theta_j^c(n_j) \phi_i(n_i) \rho(\vec{n}, t) \Big) \qquad \sum_i n_i = N \tag{3}$$

where $\hat{e}_{ij} = \hat{e}_i - \hat{e}_j$ represents the exchange of one particle along the link $j \to i$. The flow function $\phi_i(n) \le \phi_i^{\max}$ simulate the traffic flow outgoing form the node i and the capacity function $\theta_i^c(n)$ introduces the maximum capacity of the node i. These functions could be empirically computed using the dynamics at intersection points of the transport network and the node geometrical dimension. We expect $\phi_i(n) = n$ at low traffic loads since all the particles can move from the node i during a unit time, whereas its value saturates at $\phi_i^{\max}$ when the traffic load increases. The parameter $\phi_i^{\max}$ depend on the transport capacity at the intersections and they turn out the control parameters of the system. The capacity function $\theta_i^c(n) = 1$ at low traffic condition and it reduces rapidly to zero when $n \ge n_i^{\max}$ the maximum capacity of the node. Therefore in the case of low traffic load the master equation (3) simplifies by substituting $\phi_i(n_i) \to n_i$ and $\theta_i^c(n_i) = 1$ and it simulates the dynamics of N independent particles, so that the stationary distribution is a multinomial distribution with average traffic loads $\bar{n}_i = N p_i$. In this case we remark that if we maximize the information entropy for the single particle dynamics, the information entropy is also maximized for the random walk (3).
The estimate of the connectivity weighted matrix is a key issue since it defines the spectral properties of the transition rate matrix (1). In the case of an homogeneous road network w_{ij} could coincide with the adjacency matrix of the network. In the case of a multilayer network one could set the weight w_{ij} proportionally to the product of average travel velocities of the sub-networks containing the i and j roads normalized with respect the average velocity of the network, but this definition should be validated. The equilibrium state p_i is the average traffic load of the node i when the transport network is not a congested state. Since congestion formation is usually limited in the rush hours, we plan to use all the recorded MDT mobility paths to estimate the average traffic load on each road. Even if we have a small sample respect the whole population, the average procedure can provide a good estimate of the average traffic load and the transition rates (1) do not depend on the sample penetration. A similar procedure can be applied to estimate the average traffic flows on a road by using the MDT mobility paths and the definition (2) is computed using the ration between the average traffic flow and the average traffic load on each road.

If one increases that traffic load we have the high traffic load regime in which all the nodes express their maximum outgoing flow, but still we have not a congested states. The NRW dynamics (3) can be simplified by substituting $\phi_i(n_i) \to \phi_i^{\max}$ and $\theta_i^c(n_i) = 1$. This simplified master equation can be used to analytically study the rise of congestion (i.e. $n_i \simeq n_i^{\max}$) in the system. For example it can be proven that if the maximum flow rates $\phi_i^{\max}$ are proportional to the equilibrium state p_i

$$\sum_j \left(\pi_{ij}\phi_j^{\max} - \pi_{ji}\phi_i^{\max}\right) = 0 \tag{4}$$

(we call this the balance condition), all the nodes are dynamically equivalent with respect the congestion formation that depends on the traffic load fluctuations in the nodes. In such a case the congestion spreading is a percolation phenomenon[5] in the transport network. The balance condition (4) is an optimal condition for the transport network since there exist equilibrium states in which all the nodes woks at their maximum flow rates.
Conversely if the balance condition does not hold there are critical nodes in the network that will certainly be congested and the congestion spreading will be triggered by these nodes. This result points out that the maximum flow rates $\phi_i^{\max}$ are the control parameters of the NRW (3): e.g. the can be modulated by changing the traffic lights periods.

4. Discussion and Perspectives

The paper aim is to provide a theoretical framework to define maximum entropy NRW of networks to study the onset of congestion and its spreading in urban transport networks. This result is achieved by generalizing the idea of Parry's theorem in the case of MINEP to consider the existence of NESS for the proposed Markov processes, where the DB condition is violated and we have net stationary traffic flows among the nodes. The reductionist character of the NRW on networks allows an analytical approach to the study of congestion formation highlighting its universal properties and the control parameters of the system. Moreover the simulation results may provide a key of lecture of the results of data analytic for the big ICT data sets to detect the effects of congestion in real cases. The NRW models are scalable since one can describe the urban mobility at different scales and the nodes can represent a single roads or an entire area of the city.
The application to real case assume the possibility to have information of the geometrical structure of the transport network that are synthetized in the flow function that gives the dependence of the outgoing traffic flows on the traffic load of a nodes and the capacity function that limit the incoming flows when the node is congested. The flow functions depend on the dynamics at intersection points that it is the results of the urban traffic rules (e.g. the presence of traffic light). The possibility to infer maximum entropy models requires to measure the stationary states of the transport system using the available data. Our idea is to use the ICT data set that contain GPS data on individual mobility to estimate the average traffic load of the node using a time averaging over a sufficiently long period. The analysis MDT Tim data set for the Bologna metropolitan area (but the analogous data sets are available for other italian cities) allows to reconstruct the mobility paths for a sample of mobile device ($\simeq$ 5% of the whole population), and it shows that the empirical distributions of urban mobility are consistent with the existence of a maximum entropy equilibrium state. Moreover it is possible to highlight the multilayer structure of the urban road network that affects the individual behavior in the realization of the mobility paths and introduces a weighted connectivity matrix for the network. Then we can propose a road map to realize an application of our approach to realistic cases and to perform a validation procedure of the proposed models will be considered in future works. Finally, the simplicity of the models and the reasonable number of control parameters that are the maximum outgoing flow rates of the nodes gives the possibility of performing extensive simulations to implement optimizing control strategies which reduce the congestion impact.

References

[1] Vespignani, A. *Modelling dynamical processes in complex socio-technical systems* (2012) Nature Phys 8, 32–39

[2] Mizzi, C., Fabbri, A., Rambaldi, S. et al. *Unraveling pedestrian mobility on a road network using ICTs data during great tourist events* (2018) EPJ Data Sci. 7, 44

[3] Mizzi, C., Baroncini, A., Fabbri, A. et al. *Individual mobility deep insight using mobile phones data* (2023) EPJ Data Sci. 12, 56

[4] Ambühl, L., Menendez M., González M.C. *Understanding congestion propagation by combining percolation theory with the macroscopic fundamental diagram* (2023) Commun Phys 6, 26

[5] Parongama, S. *Nature of the largest cluster size distribution at the percolation threshold* (2001) J. Phys. A: Math. Gen. 34-41, 8477
[6] Wilson, A.G. *Entropy in Urban and Regional Modelling* (1970) Pion Press, London
[7] Purvis, B.; Mao, Y.; Robinson, D. *Entropy and its Application to Urban Systems* (2019) Entropy, 21, 56
[8] Parry, W. *Intrinsic Markov chains* (1965) Trans. AMS., 112, 55-65
[9] Jaynes, E.T. *The Minimum Entropy Production Principle* (1980) Ann. Rev. Phys. Chem., 31: 579-601
[10] Amaduzzi, A. *Emergent patterns in human mobility: a statistical physics approach* (2025) PHD thesis University of Bologna
[11] G.Haage *Modelling with the Master Equation Solution Methods and Applications in Social and Natural Sciences* (2017) Springer
[12] Grether D., Neumann A., Nagel K. *Simulation of Urban Traffic Control: A Queue Model Approach* (2012) Procedia Computer Science, 10, 808-814
[13] Prigogine, I. *Time, structure, and fluctuations. Science* (1978) 201(4358), 777-785
[14] Jaynes, E. T. *Information Theory and Statistical Mechanics I* (1957) Physical Review, Series II, 106 (4): 620–630
[15] Liang X, Zheng X, Lv W *The scaling of human mobility by taxis is exponential* (2012) Phys A, Stat Mech Appl 391(5):2135–2144
[16] Gonzalez MC, Hidalgo CA, Barabasi AL *Understanding human mobility patterns* (2008) Nature 454:779–782
[17] Gallotti R, Bazzani A, Rambaldi S *Understanding the variability of daily travel-time expenditures using GPS trajectory data* (2015) EPJ Data Sci 4:18
[18] Seifer, U. *Stochastic thermodynamics: principles and perspective* (2008) Eur. Phys. J. B 64(3–4),423
[19] Cover, M.T., Thomas, A.J. *Element of information theory* (2006) Wiley- Interscience publication-USA.