Layered Complex Networks

Maciej Kurant* and Patrick Thiran

Ecole Polytechnique Fédérale de Lausanne (EPFL), CH-1015, Lausanne, Switzerland
(Received 22 October 2005; published 7 April 2006)

Many complex networks are only a part of larger systems, where a number of coexisting topologies interact and depend on each other. We introduce a layered model to facilitate the description and analysis of such systems. As an example of its application, we study the load distribution in three transportation systems, where the lower layer is the physical infrastructure and the upper layer represents the traffic flows. This layered view allows us to capture the fundamental differences between the real load and commonly used load estimators, which explains why these estimators fail to approximate the real load.

In recent years, studies of biological, social, infrastructure, or technological networks have drawn a substantial amount of attention in the physics community. Although these networks are usually considered as distinct objects, they are often a part of larger complex systems, where a number of coexisting topologies interact and depend on each other. For instance, the topologies of the Internet at the IP layer [1], of the World Wide Web (WWW) [2], or of the networks formed by peer to peer (P2P) applications [3], although studied separately, are closely related: Each WWW or P2P link virtually connects two IP nodes. These two IP nodes are usually distant in the underlying IP topology, and the virtual connection is realized as a path found by IP routers. In other words, the graph formed by an application is mapped on the underlying IP network. Moreover, the IP links are in turn mapped on the physical layer [4] that consists of a mesh of optical fibers usually buried in the ground along roads, rails, or power lines. The resulting topologies at the three layers are very different from each other.

Another important class of real-life systems is transportation networks. Graphs derived from the physical infrastructure of such networks were analyzed on the examples of a power grid [5], a railway network [6], road networks [7], or urban mass transportation systems [8]. This approach often gives a valuable insight into the studied topology, but it ignores the real-life traffic pattern. Interestingly, the networks of traffic flows were studied separately, for instance, the flows of people within a city [9] and commuting traffic flows between different cities [10]. These studies, in turn, neglect the underlying physical topology. A comprehensive view of the system often requires one to analyze both layers (physical and traffic) together. Only in some particular cases is one layer sufficient. This is the case, e.g., in airport networks [11], where all traffic flows are one-hop long and the full knowledge of the traffic pattern is introduced into the physical graph by setting the edge weights equal to the amount of traffic they carry. However, in the presence of traffic flows longer than one hop, a weighted physical graph is not sufficient.

Imagine, for instance, that on a path of a long traffic flow f, one physical edge fails, causing f to disappear. This will free some resources on all physical edges originally used by f, which can be evaluated only if we know the full path taken by f. In other words, the dynamical analysis of the transportation network requires the knowledge of the traffic graph and of the actual routes of the traffic flows in the physical graph.

The above examples call for the introduction of additional layers to the description of some complex systems. In this Letter, we propose a general *multilayer model*. We explain it on the example of two layers; all the definitions naturally extend to any number of layers. In the two-layer model, the lower-layer topology is called a physical graph $G^\phi = (V^\phi, E^\phi)$, and the upper-layer topology is called a logical graph $G^\lambda = (V^\lambda, E^\lambda)$. We assume that the sets of nodes at both layers are identical, i.e., $V^\phi \equiv V^\lambda$, but as a general rule, we keep the indices ϕ and λ to make the description unambiguous. Let $N = |V^\phi| = |V^\lambda|$ be the number of nodes. The physical and logical graphs can be directed or undirected, depending on the application. The nodes and edges can have weights assigned to them and denoted by $w(\cdot)$, with $w = 1$ for unweighted graphs. Every logical edge $e^\lambda = (u^\lambda, v^\lambda)$ is mapped on the physical graph as a physical path $M(e^\lambda) \subset G^\phi$ connecting the nodes u^ϕ and v^ϕ, corresponding to u^λ and v^λ. This path may be given explicitly by the data or defined in some general form, such as a shortest path. The set of paths corresponding to all logical edges is called mapping $M(E^\lambda)$ of the logical topology on the physical topology. Now the load l of a node v^ϕ is the sum of the weights of all logical edges whose paths traverse v^ϕ:

$$l(v^\phi) = \sum_{e^\lambda:\ v^\phi \in M(e^\lambda)} w(e^\lambda). \tag{1}$$

In a transportation network, $l(v^\phi)$ is the total amount of traffic that flows through the node v^ϕ [12]; if the logical graph is unweighted, $l(v^\phi)$ counts the number of logical edges that are mapped on v^ϕ. A simple example illustrating the above definitions is given in Fig. 1(a).
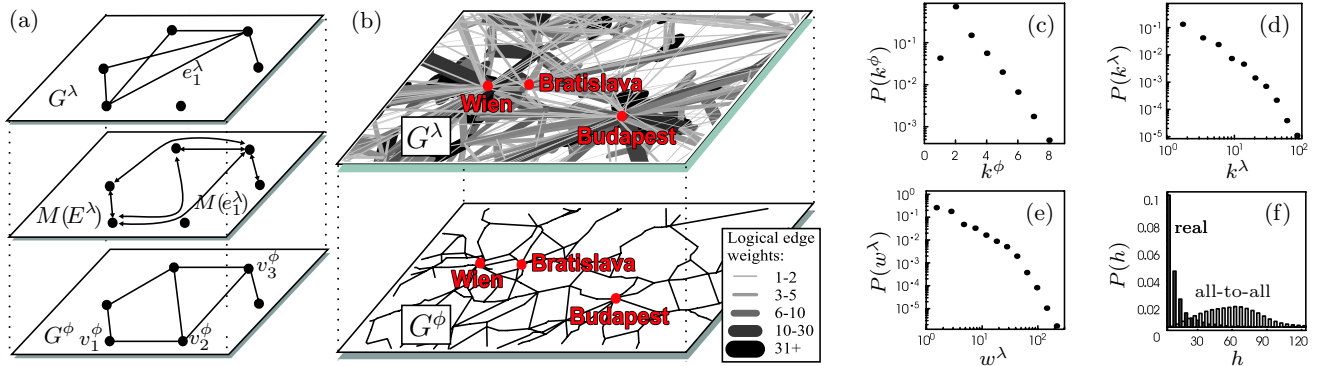
FIG. 1 (color online). Illustration of the two-layer model with the logical graph G^λ and the physical graph G^ϕ. (a) A simple example showing the actual mapping $M(E^\lambda)$ of G^λ on G^ϕ. The logical edge e_1^λ is mapped on G^ϕ as the path $M(e_1^\lambda) = (v_1^\phi, v_2^\phi, v_3^\phi)$. Assuming that G^λ is unweighted, the loads of the three indicated nodes are $l(v_1^\phi) = 3$, $l(v_2^\phi) = 2$, and $l(v_3^\phi) = 4$. (b) Two layers in the EU transportation network. The nodes are the train stations. The edges in the physical graph are the existing rail tracks connecting neighboring stations. Every edge in the logical graph connects the first and the last station of a particular train route; its weight reflects the number of trains following this route in any direction. (c)–(f) Basic characteristics of the EU network (WA and CH yield similar results): (c) node degree distribution in the physical graph; (d) node degree distribution in the logical graph; (e) edge weight distribution in the logical graph; (f) the distribution of the lengths of traffic flows, counted in a number of hops h in the physical graph.

Here we apply this two-layer framework to study transportation networks. The undirected, unweighted physical graph G^ϕ will henceforth capture the physical infrastructure of a transportation network, and the logical graph G^λ will reflect the undirected traffic flows. All data studied in this Letter are extracted from timetables of public transportation systems, with the help of the algorithms described in Ref. [13]. A timetable gives the exact route of each vehicle (bus, train, etc.), which translates directly into a logical edge e^λ (connecting the first and the last station) and its mapping $M(e^\lambda)$. The number of vehicles following the same path in both possible directions defines the flow intensity—the weight $w(e^\lambda)$ of the logical link [14].

We study three examples of transportation networks, with sizes ranging from city to continent. As an example of a city, we take the mass transportation system (buses, trams, and metros) of Warsaw (WA), Poland. At a country level, we study the railway network of Switzerland (CH). Finally, we investigate the railway network formed by major trains and stations in most countries of central Europe (EU); its fragment is shown in Fig. 1(b). The basic characteristics of these networks can be found in Table I and in Figs. 1(c)–1(f). All physical topologies are connected, planar or close to planar, with the diameter d^ϕ in

TABLE I. The studied data sets. N is the number of nodes, $|E^\phi|$ ($|E^\lambda|$) is the number of edges in the physical (logical) graph, d^ϕ is the diameter of the physical graph, and "# Vehicles" is the total number of vehicles.

Data set	N	$|E^\phi|$	d^ϕ	$|E^\lambda|$	# Vehicles
WA - Warsaw	1529	1827	90	324	26 075
CH - Switzerland	1679	1750	142	539	7482
EU - Europe	6276	7273	181	6623	54 073

the order of $O(\sqrt{N})$, and node degree distributions decaying exponentially [Fig. 1(c)]. These features are common to many physical transportation graphs, such as road networks and railway systems. The logical graphs are strikingly different. They are sparse and have multiple components, among which are many isolated nodes (the intermediate stations). Their degree distributions are right-skewed [Fig. 1(d)], meaning that there is a small number of nodes with very high degree. Similar right-skewed distributions are observed for the weights of logical edges [Fig. 1(e)]. Finally, in Fig. 1(f), we compare the length distribution of real traffic flows with the length distribution of all-to-all shortest paths. The former is very much left-skewed, which means that the real flows tend to be local.

Knowing the topologies and the mapping of both layers, we can easily compute the load of a node with Eq. (1). For a comparison, we present below two load estimators based exclusively on the physical graph G^ϕ. Our first metric is node degree k^ϕ. It seems natural that the nodes with high degree carry more traffic than the less connected nodes. Our second metric is betweenness b^ϕ [15]. The betweenness of a vertex v is the fraction of shortest paths between all pairs of vertices in a network that pass through v. If there is more than one shortest path between a given pair of vertices, then all such paths are taken into account with equal weights summing to one. Betweenness aims at capturing the amount of information passing through a vertex. Indeed, many authors take betweenness as a measure of load either directly [16,17] or with slight modifications [18]. The geographical patterns formed by the distributions of the real load, node degree, and betweenness differ substantially [see Figs. 2(a)–2(c)]. To quantify these differences, in Fig. 2(d) we present the scatter plots of the two load estimators versus the real load l. Surprisingly, contrary to the commonly admitted view, in our data set the

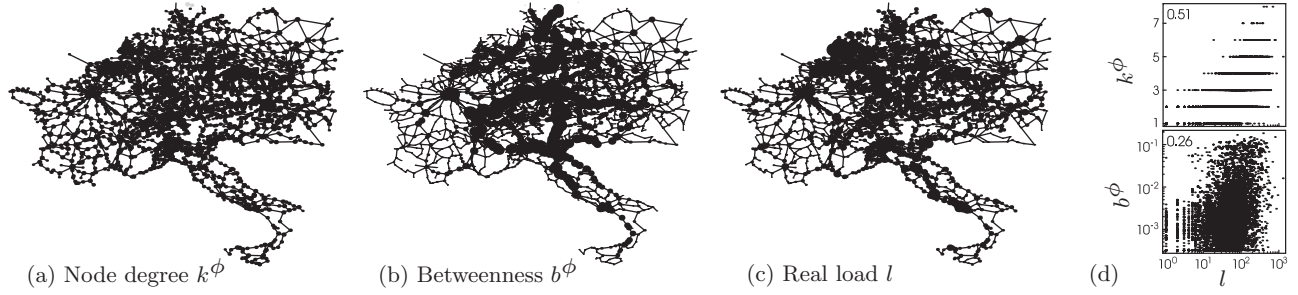(a) Node degree k^ϕ (b) Betweenness b^ϕ (c) Real load l (d)

FIG. 2. EU data set (WA and CH yield similar results). The first three figures present the physical layout of the (a) node degree k^ϕ, (b) betweenness b^ϕ, and (c) real load l. The size of a node is proportional to the measured value. In (d), we present the scatter plots of the (top) node degree k^ϕ and (bottom) betweenness b^ϕ versus the real load l. In the top left corner of every plot, we give the value of the corresponding Pearson's correlation coefficient.

node degree approximates the real load better than the betweenness (its Pearson's coefficient is higher). Nevertheless, both of them are very far from being satisfactory. For a value of the load $l \simeq 10^2$, the corresponding values of betweenness b^ϕ cover more than 2 orders of magnitude. So big disparities may strongly affect the results of the network performance analysis based on the topological load estimators. For instance, the evolution of cascading failures [17] crucially depends on the load distribution.

Why do load estimators fail to mimic the real load pattern? The layered view of the system is very helpful in answering this question. First, observe that the ways we compute node degree, betweenness, and real load can be unified by recasting the first two in the two-layer setting. Indeed, both the node degree and the betweenness can be computed as the node load (1) in systems with specific logical topologies mapped on the physical graph G^ϕ using shortest paths. We denote these logical graphs by $G_{k^\phi}^\lambda$ and $G_{b^\phi}^\lambda$, for the node degree k^ϕ and the betweenness b^ϕ, respectively. They are defined as follows. In the case of the node degree, pick $G_{k^\phi}^\lambda = G^\phi$: The logical graph is identical to the physical graph G^ϕ. Hence, the mapping of $G_{k^\phi}^\lambda$ on G^ϕ reduces trivially to single-hop traffic flows, and (1) boils down to $l(v^\phi) = k^\phi(v^\phi)$. For the betweenness, $G_{b^\phi}^\lambda$ is an unweighted and complete (fully connected) graph. Indeed, the definition of betweenness requires one to find shortest paths between every possible pair of vertices. Note that the mapping defined by betweenness splits the path (and its weight) if there is more than one shortest path, whereas the shortest-path mapping simply returns one of them. However, in large graphs this difference is negligible, especially if the shortest-path algorithm picks one of the possible paths at random.

The same two-layer methodology can, therefore, be used to compute node degree, betweenness, and real load. Moreover, in all three cases we use the same physical graph G^ϕ and a mapping that follows the shortest path. (The real-life flows almost always coincide with shortest paths connecting their end nodes.) Consequently, all the differences between the three metrics are completely captured by the logical graphs $G_{k^\phi}^\lambda$, $G_{b^\phi}^\lambda$, and G^λ. We compare

them in Table II. The graph $G_{k^\phi}^\lambda$ is moderately dense, planar, unweighted, with the degree distribution decaying exponentially. The edge length, counted in the number of hops in the mapping of this edge, is equal to 1 for all edges of $G_{k^\phi}^\lambda$. In contrast, the graph $G_{b^\phi}^\lambda$ is an unweighted, complete graph, with every node of degree equal to $N - 1$. In $G_{b^\phi}^\lambda$, we find both short and long edges; their distribution is bell-shaped, as shown by the "all-to-all" curve in Fig. 1(f). Finally, the real-life logical graph G^λ is sparse, weighted, and has rather local edges. Moreover, the node degree and edge weight distributions of G^λ are both very right-skewed.

There are, thus, a number of fundamental differences between the three logical graphs $G_{k^\phi}^\lambda$, $G_{b^\phi}^\lambda$, and G^λ. They explain why the node degree and betweenness fail to mimic the real load distribution. We expect to observe similar differences in other fields. For instance, in the Internet, the distribution of intensities of traffic flows (which corresponds in this Letter to the edge weights in the logical graph) was shown to be heavy-tailed [19,20]. This is known in the field as "the elephants and mice phenomenon" [19], where a small fraction of flows is responsible for carrying most of the traffic. Moreover, the number of flows originating from a given node (which is equivalent to the node degree in the logical graph) was also shown to follow a power-law distribution [20].

To summarize, in this Letter we have introduced a framework for studying complex systems in which we

TABLE II. The properties of the logical graphs $G_{k^\phi}^\lambda$, $G_{b^\phi}^\lambda$, and G^λ. "Edge length" is the number of hops in the mapping of the edge on the physical graph.

Property	$G_{k^\phi}^\lambda$	$G_{b^\phi}^\lambda$	G^λ
$|E^\lambda|$	$= |E^\phi|$	$= N(N-1)/2$	$< |E^\phi|$
Planar	Yes	No	No
Weights $w(e^\lambda)$	$= 1$	$= 1$	Right-skewed
Degrees $k(v^\lambda)$	Exponential	$= N - 1$	Right-skewed
Edge lengths	$= 1$	Bell-shaped	Exponential

distinguish graphs on two or more layers. We have shown on the example of transportation networks how the layered view can facilitate the description and analysis of such systems. In a traditional one-layer view of a network, we compared three load metrics (degree, betweenness, and real load) and discovered that they yield very different load distributions. Since the ways of constructing the load metrics vary greatly, it seems difficult to speak about the origins of these differences in a systematic way. However, each of the three load metrics can be recast in a two-layer view as a logical, load generating graph mapped onto the physical graph. The fundamental properties of the load metrics are then captured by their logical graphs, providing us with a common ground for a systematic comparison. Our analysis revealed that these logical graphs belong to completely different classes, and, therefore, the three studied load metrics are inherently different.

Understanding the reasons standing behind the load distribution is only one feature made possible by our layered approach. It may also completely change our view on the error and attack tolerance [21] of many systems. A failure of a single physical edge affects all logical edges that are mapped on it. Consequently, a tiny, seemingly unharmful (from one-layer perspective) disruption of the physical graph might destroy a substantial part of the logical graph, rendering the whole system useless in practice. It would be interesting to study how the topological properties at different layers affect this kind of interaction.

---

*Electronic address: Maciej.Kurant@epfl.ch

[1] M. Faloutsos, P. Faloutsos, and C. Faloutsos, in *Proceedings of the ACM SIGCOMM'99 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication, Cambridge, MA, 1999* (Association for Computing Machinery, New York, 1999).

[2] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **401**, 130 (1999).

[3] L. A. Adamic, R. M. Lukose, A. R. Puniyani, and B. A. Huberman, Phys. Rev. E **64**, 046135 (2001).

[4] K. H. Liu, in *IP over WDM* (Wiley, Chichester, England, 2002).

[5] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).

[6] P. Sen *et al.*, Phys. Rev. E **67**, 036106 (2003).

[7] M. Rosvall, A. Trusina, P. Minnhagen, and K. Sneppen, Phys. Rev. Lett. **94**, 028701 (2005); M. T. Gastner and M. E. J. Newman, Eur. Phys. J. B **49**, 247 (2006).

[8] V. Latora and M. Marchiori, Phys. Rev. Lett. **87**, 198701 (2001); J. Sienkiewicz and J. A. Hołyst, Phys. Rev. E **72**, 046127 (2005).

[9] G. Chowell, J. M. Hyman, S. Eubank, and C. Castillo-Chavez, Phys. Rev. E **68**, 066102 (2003).

[10] A. De Montis, M. Barthélemy, A. Chessa, and A. Vespignani, physics/0507106.

[11] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani, Proc. Natl. Acad. Sci. U.S.A. **101**, 3747 (2004); L. Hufnagel, D. Brockmann, and T. Geisel, Proc. Natl. Acad. Sci. U.S.A. **101**, 15 124 (2004).

[12] This definition counts equally the intermediate nodes and the end nodes. Another approach would be to give only half of a traffic unit to the end nodes. The results obtained by both methods are very similar.

[13] M. Kurant and P. Thiran, physics/0510151.

[14] Interpreting the routes of trains and buses as traffic flows gives us the picture at a low level of granularity. We view every vehicle as a traffic unit, regardless of its size or the number of people using it. Moreover, people usually use these vehicles only on a portion of their total journey and not from the first to the last station. However, we believe that they reflect well the general direction and intensity of travels, and we take a vehicle as a basic traffic unit.

[15] L. C. Freeman, Sociometry **40**, 35 (1977).

[16] K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. Lett. **87**, 278701 (2001); G. Szabó, M. Alava, and J. Kertész, Phys. Rev. E **66**, 026101 (2002); B. Bollobás and O. Riordan, Phys. Rev. E **69**, 036114 (2004).

[17] P. Holme and B. J. Kim, Phys. Rev. E **65**, 066109 (2002); L. Zhao, K. Park, and Y.-C. Lai, Phys. Rev. E **70**, 035101(R) (2004); A. E. Motter, Phys. Rev. Lett. **93**, 098701 (2004).

[18] R. Albert, I. Albert, and G. L. Nakarado, Phys. Rev. E **69**, 025103(R) (2004); P. Echenique, J. Gómez-Gardenes, and Y. Moreno, Phys. Rev. E **70**, 056105 (2004).

[19] K. Papagiannaki *et al.*, in *Proceedings of the 2nd Internet Measurement Workshop, Marseille, France, 2002* (Association for Computing Machinery, New York, 2002).

[20] M. Meiss, F. Menczer, and A. Vespignani, in *Proceedings of the Fourteenth International World Wide Web Conference, Chiba, Japan, 2005* (Association for Computing Machinery, New York, 2005).

[21] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **406**, 378 (2000).