



**HAL**  
open science

# Weighted, Bipartite, or Directed Stream Graphs for the Modeling of Temporal Networks

Matthieu Latapy, Clémence Magnien, Tiphaine Viard

## ► To cite this version:

Matthieu Latapy, Clémence Magnien, Tiphaine Viard. Weighted, Bipartite, or Directed Stream Graphs for the Modeling of Temporal Networks. *Temporal Network Theory*, Springer, pp.49-64, 2019, Computational Social Sciences, 10.1007/978-3-030-23495-9\_3. hal-03261236

**HAL Id: hal-03261236**

**<https://hal.science/hal-03261236v1>**

Submitted on 23 Nov 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Weighted, Bipartite, or Directed Stream Graphs for the Modeling of Temporal Networks

Matthieu Latapy\*, Clémence Magnien, Tiphaine Viard  
Sorbonne Université, CNRS, LIP6, F-75005 Paris, France

## Abstract

We recently introduced a formalism for the modeling of temporal networks, that we call stream graphs. It emphasizes the streaming nature of data and allows rigorous definitions of many important concepts generalizing classical graphs. This includes in particular size, density, clique, neighborhood, degree, clustering coefficient, and transitivity. In this contribution, we show that, like graphs, stream graphs may be extended to cope with bipartite structures, with node and link weights, or with link directions. We review the main bipartite, weighted or directed graph concepts proposed in the literature, we generalize them to the cases of bipartite, weighted, or directed stream graphs, and we show that obtained concepts are consistent with graph and stream graph ones. This provides a formal ground for an accurate modeling of the many temporal networks that have one or several of these features.

## 1 Introduction

Graph theory is one of the main formalisms behind network science. It provides concepts and methods for the study of networks, and it is fueled by questions and challenges raised by them. Its core principle is to model networks as sets of nodes and links between them. Then, a graph  $G$  is defined by a set of nodes  $V$  and a set of links  $E \subseteq V \otimes V$  where each link is an unordered pair of nodes<sup>1</sup>. In many cases, though, this does not capture key features of the modeled network. In particular, links may be weighted or directed, nodes may be of different kinds, etc. One key strength of graph theory is that it easily copes with such situations by defining natural extensions of basic graphs, typically weighted, bipartite, or directed graphs. Classical concepts on graphs are then extended to these more complex cases.

Stream graphs were recently introduced as a formal framework for temporal networks [28], similar to what graph theory is to networks. A stream graph  $S$  is defined by a time set  $T$ , a node set  $V$ , a set of temporal nodes  $W \subseteq T \times V$  and a set of temporal links  $E \subseteq T \times V \otimes V$ . See Figure 1 for an illustration. Each node  $v \in V$  has a set of presence times  $T_v = \{t, (t, v) \in W\}$ . Likewise,  $T_{uv} = \{t, (t, uv) \in E\}$  is the set of presence times of link  $uv$ . Conversely,  $V_t = \{v, (t, v) \in W\}$  and  $E_t = \{uv, (t, uv) \in E\}$  are the set of nodes and links present at time  $t$ , leading to the graph at time  $t$ :  $G_t = (V_t, E_t)$ . The graph induced by  $S$  is  $G(S) = (\{v, T_v \neq \emptyset\}, \{uv, T_{uv} \neq \emptyset\})$ .

---

\*Corresponding author: Matthieu.Latapy@lip6.fr

<sup>1</sup>Given any two sets  $X$  and  $Y$ , we denote by  $X \times Y$  the cartesian product of  $X$  and  $Y$ , *i.e.* the set of all ordered pairs  $(x, y)$  such that  $x \in X$  and  $y \in Y$ . We denote by  $X \otimes Y$  the set of all unordered pairs composed of  $x \in X$  and  $y \in Y$ , with  $x \neq y$ , that we denote by  $xy = yx$ .

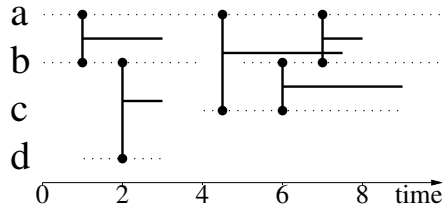


Figure 1: **An example of stream graph:**  $S = (T, V, W, E)$  with  $T = [0, 10] \subseteq \mathbb{R}$ ,  $V = \{a, b, c, d\}$ ,  $W = [0, 10] \times \{a\} \cup ([0, 4] \cup [5, 10]) \times \{b\} \cup [4, 9] \times \{c\} \cup [1, 3] \times \{d\}$ , and  $E = ([1, 3] \cup [7, 8]) \times \{ab\} \cup [4.5, 7.5] \times \{ac\} \cup [6, 9] \times \{bc\} \cup [2, 3] \times \{bd\}$ . In other words,  $T_a = [0, 10]$ ,  $T_b = [0, 4] \cup [5, 10]$ ,  $T_c = [4, 9]$ ,  $T_d = [1, 3]$ ,  $T_{ab} = [1, 3] \cup [7, 8]$ ,  $T_{ac} = [4.5, 7.5]$ ,  $T_{bc} = [6, 9]$ ,  $T_{bd} = [2, 3]$ , and  $T_{ad} = T_{cd} = \emptyset$ .

Stream graphs encode the same information as Time Varying Graphs (TVG) [13], Relational Event Models (REM) [11, 47], Multi-Aspect Graphs (MAG) [52, 53], or other models of temporal networks. Stream graphs emphasize the streaming nature of data, but all stream graph concepts may easily be translated to these other points of views.

A wide range of graph concepts have been extended to stream graphs [28]. The most basic ones are probably the number of nodes  $n = \sum_{v \in V} \frac{|T_v|}{|T|}$  and the number of links  $m = \sum_{uv \in V \otimes V} \frac{|T_{uv}|}{|T|}$ . Then, the neighborhood of node  $v$  is  $N(v) = \{(t, u), (t, uv) \in E\}$  and its degree is  $d(v) = \frac{|N(v)|}{|T|}$ . The average degree of  $S$  is the average degree of all nodes weighted by their presence time:  $d(S) = \sum_{v \in V} \frac{|T_v|}{|W|} d(v)$ .

Going further, the density of  $S$  is  $\delta(S) = \frac{m}{\sum_{uv \in V \otimes V} |T_u \cap T_v|}$ . It is the probability, when one chooses at random a time instant and two nodes present at that time, that these two nodes are linked together at that time. Then, a clique is a subset  $C$  of  $W$  such that for all  $(t, u)$  and  $(t, v)$  in  $C$ ,  $u$  and  $v$  are linked together at time  $t$  in  $S$ , *i.e.*  $(t, uv) \in E$ . Equivalently, a subset of  $W$  is a clique of  $S$  if the substream it induces has density 1.

This leads to the definition of clustering coefficient in stream graphs: like in graphs,  $cc(v)$  is the density of the neighborhood of  $v$ . Equivalently,  $cc(v) = \sum_{uvw \in V \otimes V} \frac{|T_{vu} \cap T_{vw} \cap T_{uv}|}{|T_{vu} \cap T_{vw}|}$ . Likewise, the transitivity of  $S$  is the fraction of all 4-uplets  $(t, u, v, w)$  with  $(t, uv)$  and  $(t, vw)$  in  $E$  such that  $(t, vw)$  is also in  $E$ .

These concepts generalize graph concepts in the following sense. A stream  $S$  is called *graph-equivalent* if it has no dynamics:  $G_t = G(S)$  for all  $t$ . In this case, each stream property of  $S$  is equal to the corresponding graph property of  $G(S)$ . For instance, the density of  $S$  is equal to the one of  $G(S)$ . Graphs may therefore be seen as special cases of stream graphs (the ones with no dynamics).

Like for graphs, the stream graph formalism was designed to be readily extendable to weighted, directed, or bipartite cases. However, these extensions remain to be done, and this is the goal of the present contribution, summarized in Figure 2.

Before entering in the core of this contribution, notice that the set of available graph concepts is huge, much larger than what may be considered here. We therefore focus on a the set of key properties succinctly summarized above. In particular, we do not consider path-related concepts, which would deserve a dedicated work of their own. Rather than being exhaustive, our aim is to illustrate how weighted, bipartite, or directed graph concepts may be generalized to stream graphs in a consistent way, and to provide a ground for further

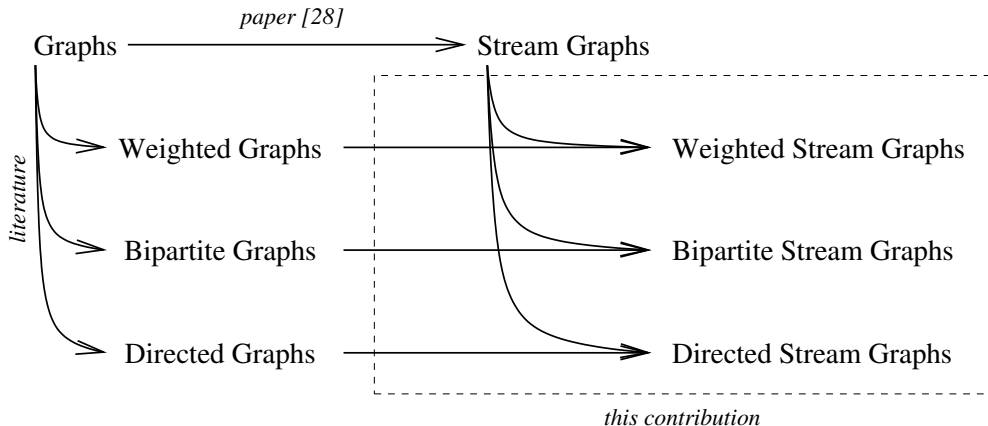


Figure 2: The global positioning of this contribution with respect to state-of-the-art. Left: weighted, bipartite and directed extensions of graph properties are available in the scientific literature. Top: a generalization of graphs to stream graphs was proposed in [28]. Dotted rectangle: in this contribution, we extend weighted, bipartite and directed graph concepts to weighted, bipartite, and directed stream graphs, as well as standard stream graph concepts to the weighted, bipartite and directed cases, in a consistent way.

generalizations.

## 2 Weighted stream graphs

A **weighted graph** is a graph  $G = (V, E)$  equipped with a weight function  $\omega$  generally defined over  $E$ , and sometimes on  $V$  too. Then,  $\omega(v)$  is the weight of node  $v$  and  $\omega(uv)$  the one of link  $uv$ . Link weights may represent tie strength (in a friendship or collaboration network for instance) [36, 45], link capacity (in an road or computer network, for instance) [12, 45], or a level of similarity (in document or image networks, or gene networks for instance) [26, 54]. Node weights may represent reliability, availability, size, etc. As a consequence, weighted graphs are very important and they are used to model a wide variety of networks. In most cases, though, nodes are considered unweighted. We will therefore only consider weighted links in the following, except where specified otherwise.

Even when one considers a weighted graph  $G = (V, E)$  equipped with the weight function  $\omega$ , the properties of  $G$  itself (without weights) are of crucial interest. In addition, one may consider thresholded versions of  $G$ , defined as  $G_\tau = (V_\tau, E_\tau)$  where  $V_\tau = \{v \in V, \omega(v) \geq \tau\}$  and  $E_\tau = \{uv \in E, \omega(uv) \geq \tau\}$ , for various thresholds  $\tau$ . This actually is a widely used way to deal with graph weights, formalized in a systematic way as early as 1969 [16]. However, one often needs to truly take weights into account, without removing any information. In particular, the importance of weak links is missed by thresholding approaches. In addition, determining appropriate thresholds is a challenge in itself [18, 45, 46].

As a consequence, several extensions of classical graph concepts have been introduced to incorporate weight information and deal directly with it. The most basic ones are the maximal, minimal, and average weights, denoted respectively by  $\omega_{\max}$ ,  $\omega_{\min}$  and  $\langle \omega \rangle$ . The minimal weight  $\omega_{\min}$  is often implicitly considered as equal to 0, and weights are sometimes normalized in order to ensure that  $\omega_{\max} = 1$  [37, 54, 21, 1].

In addition to these trivial metrics, one of the most classical property probably is the weighted version of node degree, known as node strength [5, 17, 4]:  $s(v) = \sum_{u \in N(v)} \omega(uv)$ . Notice that the average node strength is equal to the product of the average node degree and the average link weight.

Strength is generally used jointly with classical degree, in particular to investigate correlations between degree and strength: if weights represent a kind of activity (like travels or communications) then correlations give information on how activity is distributed over the structure [5, 42]. One may also combine node degree and strength in order to obtain a measure of node importance. For instance, [39] uses a tuning parameter  $\alpha$  and compute  $d(v) \cdot \left(\frac{s(v)}{d(v)}\right)^\alpha$ , but more advanced approaches exist [12].

Generalizing density, *i.e.* the number of present links divided by the total number of possible links, raises subtle questions. Indeed, it seems natural to replace the number of present links by the sum of all weights  $\sum_{uv \in E} \omega(uv)$  like for strength, but several variants for the total sum of possible weights make sense. For instance, the litterature on rich clubs [2, 40] considers that all present links may have the maximal weight, leading to  $\sum_{uv \in E} \omega_{\max}$ , or that all links may be present and have the maximal weight, leading to  $\sum_{uv \in V \otimes V} \omega_{\max}$ . In the special case where weights represent a level of certainty between 0 and 1 for link presence (1 if it is present for sure, 0 if it is absent for sure), then the weighted density may be defined as  $\frac{\sum_{uv \in E} \omega(uv)}{|V \otimes V|}$  [57].

Various definitions of weighted clustering coefficients have been proposed, and [44, 4, 49] review many of them in details. The most classical one was proposed in [5]:  $cc(v) = \frac{1}{s(v)(d(v)-1)} \sum_{i,j \in N(v), ij \in E} \frac{\omega(vi)+\omega(vj)}{2}$ . A general approach was also proposed in [41]. Given a node  $v$ , it assigns a value to each triplet of distinct nodes  $(i, v, j)$  such that  $iv$  and  $jv$  are in  $E$  and to each such triplet such that  $ij$  is also in  $E$ . Then,  $cc(v)$  is defined as the ratio between the sum of values of triplets in the second category and the one of triplets in the first category. In [41], considered values are the arithmetic mean, geometric mean, maximal value or minimal value of weights of involved links, depending on the application. One may also consider the product of weights, leading to  $cc(v) = \frac{\sum_{i,j \in N(v), ij \in E} \omega(vi) \cdot \omega(vj) \cdot \omega(ij)}{\sum_{i \neq j \in N(v)} \omega(vi) \cdot \omega(vj)}$  as proposed in [54, 26, 1] with normalized weights.

Transitivity is generalized in a very similar way [41] by considering all triplets of distinct nodes  $(i, j, k)$  such that  $ij$  and  $jk$  are in  $E$  and each such triplet such that  $ik$  is also in  $E$ , instead of only the ones centered on a specific node  $v$ . If the associated value is the product of weight, this leads to  $tr = \frac{\sum_{(i,j,k)} \omega(ij) \cdot \omega(jk) \cdot \omega(ik)}{\sum_{(i,j,k)} \omega(ij) \cdot \omega(jk)}$ . If all weights are equal to 1 (*i.e.* the graph is unweighted) this is nothing but the transitivity in  $G$ .

Various other concepts have been generalized to weighted graphs, like for instance assortativity [5], and specific weighted graph concepts, like closeness and betweenness centralities [39], connectability [3], eigenvector centrality [17], or rich club coefficient [2, 40, 56]. We do not consider them here as our focus is on the most basic properties.

We define a **weighted stream graph** as a stream graph  $S = (T, V, W, E)$  equipped with a weight function  $\omega$  defined over  $W$  and  $E$ : if  $(t, v) \in W$  then  $\omega(t, v)$  is the weight of node  $v$  at time  $t$ , and if  $(t, uv) \in E$  then  $\omega(t, uv)$  is the weight of link  $uv$  at time  $t$ . See Figure 3 for an illustration.

If a stream graph represents money transfers, then node weights may represent available credit and link weights may represent transfer amounts; if a stream represents travels, then

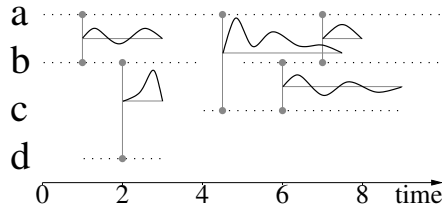


Figure 3: **An example of weighted stream graph.** In this example, nodes are unweighted but links are weighted. Instead of just a straight horizontal line indicating link presence over time, we plot the weight value (assuming that 0 is indicated by the horizontal line).

node weights may represent available fuel, and link weights may represent speed; if a stream represents contacts between mobile device, node weights may represent battery charge and link weights may represent signal strength or link capacity; if a stream represents data transfers between computers then link weights may represent throughput or error rates; like for weighted graphs, countless situations may benefit from a weighted stream graph modeling.

As we will see in Section 3, weighted stream graphs also widely appear within bipartite graph studies. In addition, as explained in [28], section 19, one often resorts to  $\Delta$ -analysis for stream graph studies. Given a stream graph  $S = (T, V, W, E)$  with  $T = [x, y]$  and a parameter  $\Delta$ , its most simple form consists in transforming  $S$  into  $S_\Delta = (T', V, W', E')$  such that  $T' = [x + \frac{\Delta}{2}, y - \frac{\Delta}{2}]$ ,  $T'_v = T' \cap \{t, \exists t' \in [t - \frac{\Delta}{2}, t + \frac{\Delta}{2}], t' \in T_v\}$ , and  $T'_{uv} = T' \cap \{t, \exists t' \in [t - \frac{\Delta}{2}, t + \frac{\Delta}{2}], t' \in T_{uv}\}$ . Then, one may capture the amount of information in  $S$  leading to node and link presences in  $S_\Delta$  with weights:  $\omega(t, v) = |\{t, \exists t' \in [t - \frac{\Delta}{2}, t + \frac{\Delta}{2}], t' \in T_v\}|$  and  $\omega(t, uv) = |\{t, \exists t' \in [t - \frac{\Delta}{2}, t + \frac{\Delta}{2}], t' \in T_{uv}\}|$ .

Like for weighted graphs, in addition to  $S$  itself (without weights), one may consider the thresholded (unweighted) stream graphs  $S_\tau = (T, V, W_\tau, E_\tau)$  where  $W_\tau = \{(t, v) \in W, \omega(t, v) \geq \tau\}$  and  $E_\tau = \{(t, uv) \in E, \omega(t, uv) \geq \tau\}$ , for various thresholds  $\tau$ . One may then study how the properties of  $S_\tau$  evolve with  $\tau$ .

The graph obtained from  $S$  at time  $t$ ,  $G_t = (V_t, E_t)$ , is naturally weighted by the function  $\omega_t(v) = \omega(t, v)$  and  $\omega_t(uv) = \omega(t, uv)$ . Likewise, the induced graph  $G(S)$  is weighted by  $\omega(v) = \frac{1}{|T|} \int_{t \in T_v} \omega(t, v) dt$  and  $\omega(uv) = \frac{1}{|T|} \int_{t \in T_{uv}} \omega(t, uv) dt$ . These definitions correspond to the average weight over time of each node and link. One may define similarly the minimal and maximal node and link weights, and go further by studying  $S$  through the weighted graph  $G(S)$  and the time-evolution of weighted graph properties of  $G_t$ .

These approaches aim to take both the weight and the temporal aspect into account: from a weighted stream graph, the first one provides a family of (unweighted) stream graphs, one for each considered value of the threshold; the second one provides a series of (static) weighted graphs, one for each instant considered. In both cases, the actual combination of weight and time information is poorly captured. We will therefore define concepts that jointly deal with both weights and time. Like with weighted graphs, we simplify the presentation by assuming that only links are weighted (nodes are not).

Since the degree of node  $v$  in a stream graph  $S$  is  $d(v) = \sum_{u \in V} \frac{|T_{uv}|}{|T|}$  and since the strength of node  $v$  in a weighted graph  $G$  is  $s(v) = \sum_{u \in N(v)} \omega(uv)$ , we define the strength of node  $v$  in a weighted stream graph  $S$  as  $s(v) = \sum_{u \in V} \int_{t \in T_{uv}} \frac{\omega(t, uv)}{|T|} dt$ . It is the degree of  $v$  where each neighbor is counted with respect to the weight of its links with  $v$  at the times when it is linked to  $v$ . It is related to the strength  $s_t(v)$  of  $v$  in  $G_t$  as follows:  $s(v) = \frac{1}{|T|} \int_{t \in T_v} s_t(v)$ ; it

is the average strength of  $v$  over time.

With this definition, one may study correlations between degree and strength in stream graphs, as with graphs, and even directly use their combinations, like  $d(v) \cdot \left(\frac{s(v)}{d(v)}\right)^\alpha$  where  $\alpha$  is a parameter.

Unsurprisingly, generalizing density to weighted stream graphs raises the same difficulties as for weighted graphs. Still, proposed definitions for weighted graphs easily apply to weighted stream graphs. Indeed, the sum of all weights becomes  $\sum_{uv \in V \otimes V} \int_{t \in T_{uv}} \omega(t, uv) dt$ , and the maximal weight of possible links may be defined as  $\sum_{uv \in V \otimes V} \int_{t \in T_{uv}} \omega_{\max} dt = \omega_{\max} \cdot |E|$  or  $\sum_{uv \in V \otimes V} \int_{t \in T} \omega_{\max} dt = \omega_{\max} \cdot |V \otimes V|$ . Like with weighted graphs, if all weights are in  $[0, 1]$  then the weighted density may be defined by  $\frac{\sum_{uv \in V \otimes V} \int_{t \in T_{uv}} \omega(t, uv) dt}{|T \otimes V \otimes V|}$ .

Then, one may define the clustering coefficient of  $v$  as the weighted density (according to one of the definitions above or another one) of the neighborhood of  $v$ . One may also consider the time-evolution of one the weighted clustering coefficient in  $G_t$ , according to previously proposed definitions surveyed above. Interestingly, one may also generalize the classical definition [5] as follows:  $cc(v) = \frac{1}{2s(v)(d(v)-1)} \int_{t \in T_v} \sum_{i, j \in N_t(v), (t, ij) \in E} \omega(t, vi) + \omega(t, vj) dt$ , which is the product of link weights of  $v$  with its pairs of neighbors when these neighbors are linked together.

The general approach of [41] also extends: one has to assign a value to each quadruplet  $(t, i, j, k)$  with  $i, j$ , and  $k$  distinct such that  $(t, ij)$ ,  $(t, jk)$  are in  $E$ , and to each such quadruplet such that  $(t, ik)$  also is in  $E$ . As in the weighted graph case, the weighted stream graph clustering coefficient of node  $v$ ,  $cc(v)$  is then the ratio between the sum of values of quadruplets in the second category such that  $j = v$  and the one of quadruplets in the first category such that  $j = v$  too. If the value of a quadruplet is the product of the weights of involved links, we obtain  $cc(v) = \frac{\int_{t \in T_v} \sum_{i, j \in N_t(v), (t, ij) \in E} \omega(t, vi) \cdot \omega(t, vj) \cdot \omega(t, ij) dt}{\int_{t \in T_v} \sum_{i \neq j \in N_t(v)} \omega(t, vi) \cdot \omega(t, vj) dt}$ .

Likewise, we define the weighted stream graph transitivity as the ratio between the sum of values of all quadruplets in the second category above and the one of quadruplets in the first category. If the value of quadruplets is defined as the product of weights of involved links, this leads to  $tr = \frac{\int_t \sum_{(i, j, k)} \omega(t, ij) \cdot \omega(t, jk) \cdot \omega(t, ik) dt}{\int_t \sum_{(i, j, k)} \omega(t, ij) \cdot \omega(t, jk) dt}$ . If all weights are equal to 1 (*i.e.* the stream is unweighted), this is nothing but the stream graph transitivity defined in [28].

If  $S$  is a graph-equivalent stream weighted by a constant function over time, *i.e.*  $\omega(t, v) = \omega(t', v)$  and  $\omega(t, uv) = \omega(t', uv)$  for all  $t$  and  $t'$ , then it is equivalent to the weighted graph  $G(S)$  weighted by  $\omega(v) = \omega(t, v)$  and  $\omega(uv) = \omega(t, uv)$  for any  $t$ ; we call it a weighted graph-equivalent weighted stream. The strength of  $v$  in  $S$  is equal to its strength in  $G(S)$  if  $S$  is a weighted graph-equivalent weighted stream. The same is true for the different notions of density or clustering coefficient: the density of  $S$  is equal to the density of  $G(S)$  and the clustering coefficient of a vertex  $v$  or the transitivity are equal to the clustering coefficient or transitivity in  $G(S)$ .

### 3 Bipartite stream graphs

A **bipartite graph**  $G = (\top \cup \perp, E)$  is defined by a set of top nodes  $\top$ , a set of bottom nodes  $\perp$  with  $\top \cap \perp = \emptyset$ , and a set of links  $E \subseteq \top \otimes \perp$ : there are two different kinds of nodes and links may exist only between nodes of different kinds. Like weighted graphs, but maybe less known, bipartite graphs are pervasive and model many real-world data [24, 27].

Typical examples include relations between client and products [7], between company boards and their members [43, 6], and between items and their key features like movie-actor networks [51, 34] or publication-author networks [35, 36], to cite only a few.

Bipartite graphs are often studied through their top or bottom projections [10]  $G_{\top} = (\top, E_{\top})$  and  $G_{\perp} = (\perp, E_{\perp})$ , defined by  $E_{\top} = \cup_{v \in \perp} N(v) \otimes N(v)$  and  $E_{\perp} = \cup_{v \in \top} N(v) \otimes N(v)$ . In other words, in  $G_{\top}$  two (top) nodes are linked together if they have (at least) a (bottom) neighbor in common in  $G$ , and  $G_{\perp}$  is defined symmetrically. Notice that, if  $v \in \top$  (resp.  $v \in \perp$ ), then  $N(v)$  always is a (not necessarily maximal) clique in  $G_{\perp}$  (resp.  $G_{\top}$ ).

Projections induce important information losses: the existence of a link or a clique in the projection may come from very different causes in the original bipartite graph. To improve this situation, one often considers *weighted* projections: each link  $uv$  in the projection is weighted by the number  $\omega(uv) = |N(u) \cap N(v)|$  of neighbors  $u$  and  $v$  have in common in the original bipartite graph. One may then use weighted graph tools to study the weighted projections [23, 5, 17], but information losses remain important. In addition, projection are often much larger than the original bipartite graphs, which raises serious computational issues [27].

As a consequence, many classical graph concepts have been extended to deal directly with the bipartite case, see [27, 9, 20, 10, 8]. The most basic properties are  $n_{\top} = |\top|$  and  $n_{\perp} = |\perp|$ , the number of top and bottom nodes. The definition of the number of links  $m = |E|$  is the same as in classical graphs. The definitions of node neighborhoods and degrees are also unchanged. The average top and bottom degrees  $d_{\top}$  and  $d_{\perp}$  of  $G$  are the average degrees of top and bottom nodes, respectively.

With these notations, the bipartite density of  $G$  is naturally defined as  $\delta(G) = \frac{m}{n_{\top} \cdot n_{\perp}}$ : it is the probability when one takes two nodes that may be linked together that they indeed are. Then, a bipartite clique in  $G$  is a set  $C_{\top} \cup C_{\perp}$  with  $C_{\top} \subseteq \top$  and  $C_{\perp} \subseteq \perp$  such that  $C_{\top} \times C_{\perp} \subseteq E$ . In other words, all possible links between nodes in a bipartite clique are present in the bipartite graph.

Several bipartite generalizations of the clustering coefficient have been proposed [27, 29, 55, 38, 30]. In particular:

- [27] and [29] rely on the Jaccard coefficient defined over node pairs both either in  $\top$  or  $\perp$ :  $cc(uv) = \frac{|N(u) \cap N(v)|}{|N(u) \cup N(v)|}$  or close variants. One then obtains the clustering coefficient of node  $v$  by averaging its Jaccard coefficient with all neighbors of its neighbors:  $cc(v) = \frac{\sum_{u \in N(N(v)), u \neq v} cc(uv)}{|N(N(v))|}$ .
- for each node  $v$ , [27] and [30] consider all triplets  $(u, v, w)$  of distinct nodes with  $u$  and  $w$  in  $N(v)$  and define the redundancy  $rc(v)$  as the fraction of such triplets such that there exists an other node in  $N(u) \cap N(w)$ .
- similarly, [38] proposes to consider all quintuplets  $(a, b, v, c, d)$  of distinct nodes with  $b, c \in N(v)$ ,  $a \in N(b)$  and  $d \in N(c)$  and defines  $cc^*(v)$  as the fraction of them such that there exists another node in  $N(a) \cap N(d)$ .

Transitivity is usually extended to bipartite graphs [43, 27, 38] by considering the set  $\mathbb{N}$  of quadruplets of nodes  $(a, b, c, d)$  such that  $ab, bc$  and  $cd$  are in  $E$  and the set  $\bowtie$  of such quadruplets with in addition  $ad$  in  $E$ . Then, the transitivity is  $tr(G) = \frac{|\bowtie|}{|\mathbb{N}|}$ . Like above, [38] also propose to consider the fraction of all quintuplets  $(a, b, c, d, e)$  of distinct nodes with  $ab, bc, cd, de$  in  $E$  such that there exists an other node  $f$  with  $af$  and  $ef$  in  $E$ .



We define a **bipartite stream graph**  $S = (T, \top \cup \perp, W, E)$  from a time interval  $T$ , a set of top nodes  $\top$ , a set of bottom nodes  $\perp$  with  $\top \cap \perp = \emptyset$ , and two sets  $W \subseteq T \times (\top \cup \perp)$  and  $E \subseteq T \times \top \times \perp$  such that  $(t, uv) \in E$  implies  $(t, u) \in W$  and  $(t, v) \in W$ . See Figure 4 (left) for an illustration. Each instantaneous graph  $G_t$ , as well as the induced graph  $G(S)$ , are bipartite graphs with the same top and bottom nodes.

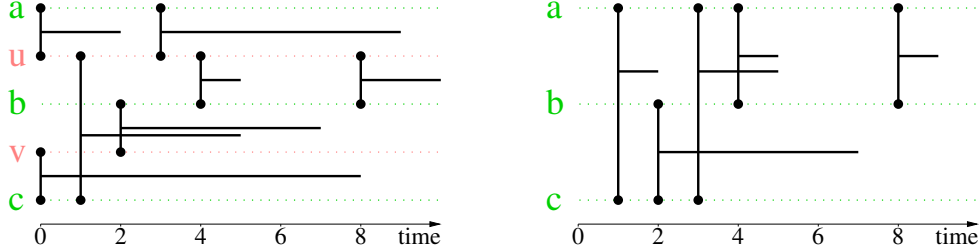


Figure 4: **Left:** a bipartite stream graph  $S = (T, \top \cup \perp, W, E)$  with  $T = [0, 10]$ ,  $\top = \{u, v\}$ ,  $\perp = \{a, b, c\}$ ,  $W = T \times (\top \cup \perp)$ , and  $E = ([0, 2] \cup [3, 9]) \times \{ua\} \cup ([4, 5] \cup [8, 10]) \times \{ub\} \cup [1, 5] \times \{uc\} \cup [2, 7] \times \{v, b\} \cup [0, 8] \times \{vc\}$ . **Right:** its  $\perp$ -projection  $S_\perp$ . For instance,  $a$  and  $c$  are linked together from time 3 to 5 because they both have  $u$  in their neighborhood for this time period in  $S$ .

Bipartite stream graphs naturally model many situations, like for instance presence of people in rooms or other kinds of locations, purchases of products by clients, access to on-line services, bus presence at stations [15], traffic between a set of computers and the rest of the internet [48], or contribution of people to projects, such as software.

The classical definition of projections is easily extended, leading to  $S_\top = (T, \top, W_\top, E_\top)$  and  $S_\perp = (T, \perp, W_\perp, E_\perp)$ , where  $W_\top = W \cap (T \times \top)$ ,  $W_\perp = W \cap (T \times \perp)$ ,  $E_\top = \cup_{(t,v) \in W_\perp} \{(t, uv) \text{ s.t. } (t, uv) \in E \text{ and } (t, v) \in W_\perp\}$  and  $E_\perp = \cup_{(t,v) \in W_\top} \{(t, uv) \text{ s.t. } (t, uv) \in E \text{ and } (t, v) \in W_\top\}$ . In other words, in  $S_\top$  two (top) nodes are linked together at a given time instant if they have (at least) a (bottom) neighbor in common in  $S$  at this time, and  $S_\perp$  is defined symmetrically. See Figure 4 for an illustration. Notice that, if  $v \in \top$  (resp.  $v \in \perp$ ) then  $N(v)$  always is a (not necessarily maximal) clique in  $S_\perp$  (resp.  $S_\top$ ).

One may also generalize weighted projections by considering the number  $\omega(t, uv) = |N_t(u) \cap N_t(v)|$  of neighbors  $u$  and  $v$  have in common at time  $t$  in the original bipartite stream graph. One then obtains weighted stream graphs, and may use the definitions proposed in Section 2 to study them. Still, this induces much information loss, which calls for the generalization of bipartite properties themselves.

The most immediate definitions are the numbers of top and bottom nodes:  $n_\top = \sum_{v \in \top} \frac{|T_v|}{|T|} = \frac{|W \cap (T \times \top)|}{|T|}$  and  $n_\perp = \sum_{v \in \perp} \frac{|T_v|}{|T|} = \frac{|W \cap (T \times \perp)|}{|T|}$ , respectively. We then have  $n = n_\top + n_\perp$ . Like with graphs, the number of links, neighborhood of nodes, and their degree do not call for specific bipartite definitions. We however define the average top and bottom degrees  $d_\top$  and  $d_\perp$  of  $S$  as the average degrees of top and bottom nodes, respectively, weighted by their presence time in the stream:  $d_\top(S) = \sum_{v \in \top} \frac{|T_v|}{|W|} d(v)$  and  $d_\perp(S) = \sum_{v \in \perp} \frac{|T_v|}{|W|} d(v)$ .

We define the bipartite density of  $S$  as  $\delta(S) = \frac{m}{\sum_{u \in \top, v \in \perp} |T_u \cap T_v|}$ : it is the probability when one takes two nodes when they may be linked together that they indeed are. A subset  $C = C_\top \cup C_\perp$  of  $W$ , with  $C_\top \subseteq T \times \top$  and  $C_\perp \subseteq T \times \perp$ , is a clique in  $S$  if all possible links exist between nodes when they are involved in  $C$ : for all  $t$ , if  $(t, u)$  is in  $C_\top$  and  $v$  is in  $C_\perp$  then  $uv$  is in  $E$ .

Like with graphs, defining a bipartite stream graph clustering coefficient is difficult and leaves us with several reasonable choices:

- Extending the Jaccard coefficient to node pairs in a stream graph leads to an instantaneous definition:  $cc_t(uv) = \frac{|N_t(u) \cap N_t(v)|}{|N_t(u) \cup N_t(v)|}$ , which is nothing but  $cc(uv)$  in  $G_t$ . It also leads to a global definition:  $cc(uv) = \frac{|N(u) \cap N(v)|}{|N(u) \cup N(v)|} = \frac{\sum_{w \in \top \cup \perp} |T_{uw} \cap T_{vw}|}{\sum_{w \in \top \cup \perp} |T_{uw} \cup T_{vw}|}$ . We may then define the bipartite clustering coefficient of node  $v$  by averaging its Jaccard coefficient with all neighbors of its neighbors, weighted by their co-presence time:  $\frac{1}{|N(N(v))|} \sum_{u \in \top \cup \perp} \frac{|T_u \cap T_v|}{|T|} cc(uv)$ . Notice that  $cc(uv) = 0$  if the neighborhoods of  $u$  and  $v$  do not intersect. This means that the sum actually is over nodes that are at some time neighbor of a neighbor of  $v$ , which is consistent with the bipartite graph definition.
- Redundancy is easier to generalize to stream graphs: for each node  $v$ , we consider all triplets  $(t, u, v, w)$  composed of a time instant  $t$ , a neighbor  $u$  of  $v$  at time  $t$ ,  $v$  itself, and another neighbor  $w$  of  $v$  at time  $t$ , and we define  $rc(v)$  as the fraction of such quadruplets such that there exists an other node linked to  $u$  and  $w$  at time  $t$ . In other words,  $rc(v) = \frac{|\{(t, u, v, w), u \neq w, u, w \in N_t(v), \exists x \in N_t(u) \cap N_t(w), x \neq v\}|}{|\{(t, u, v, w), u \neq w, u, w \in N_t(v)\}|}$ .
- Similarly, we propose a stream graph generalization of  $cc^*(v)$  as the fraction of sextuplets  $(t, a, b, v, c, d)$  with  $a, b, v, c,$  and  $d$  all different from each other,  $b \in N_t(v)$ ,  $c \in N_t(v)$ ,  $a \in N_t(b)$  and  $d \in N_t(c)$  for which in addition there exists an other node in  $N_t(a) \cap N_t(d)$ .

Finally, if we denote by  $\mathbb{N}$  the set of quintuplets  $(t, a, b, c, d)$  such that  $(t, ab)$ ,  $(t, bc)$  and  $(t, cd)$  are in  $E$  and the set  $\bowtie$  of such quintuplets with in addition  $(t, ad)$  in  $E$ , then the bipartite transitivity for stream graphs is  $tr(G) = \frac{|\bowtie|}{|\mathbb{N}|}$  as before. Like for bipartite graphs, one may also consider the fraction of all sextuplets  $(t, a, b, c, d, e)$  where  $a, b, c, d,$  and  $e$  are distinct nodes with  $(t, ab)$ ,  $(t, bc)$ ,  $(t, cd)$ ,  $(t, de)$  in  $E$  such that there exists an other node  $f$  with  $(t, af)$  and  $(t, ef)$  in  $E$ .

If  $S$  is a graph-equivalent bipartite stream, then its projections  $S_\top$  and  $S_\perp$  are also graph-equivalent streams, and their corresponding graphs are the projections of the graph corresponding to  $S$ :  $G(S_\perp) = G(S)_\perp$  and  $G(S_\top) = G(S)_\top$ . In addition, the bipartite properties of  $S$  are equivalent to the bipartite properties of its corresponding bipartite graph: the density, Jaccard coefficient, redundancy, and  $cc^*$ , as well as transitivity values, are all equal to their graph counterpart in  $G(S)$  if  $S$  is a graph-equivalent bipartite stream.

## 4 Directed stream graphs

A **directed graph**  $G = (V, E)$  is defined by its set  $V$  of nodes and its set  $E \subseteq V \times V$  of links: while links of undirected graphs are unordered pairs of distinct nodes, links in directed graphs are ordered pairs of nodes, not necessarily distinct:  $(u, v) \neq (v, u)$ , and  $(v, v)$  is allowed and called a loop. Then,  $(u, v)$  is a link from  $u$  to  $v$ , and if both  $(u, v)$  and  $(v, u)$  are in  $E$  then the link is said to be symmetric.

Directed graphs naturally model the many situations where link asymmetry is important, like for instance dependencies between companies or species, citations between papers or web pages [32], friendship relations in many on-line social networks [33], or hierarchical relations of various kinds.

Directed graphs are often studied as undirected graphs by ignoring link directions. However, this is not satisfactory in many cases: having a link to a node is very different from having a link from a node and, for instance, having links to many nodes is very different from having links from many nodes. As a consequence, many directed extensions of standard graph properties have been proposed to take direction into account, see for instance [50] section 4.3 and [19, 25].

First, a node  $v$  in a directed graph has two neighborhoods: its out-going neighborhood  $N^+(v) = \{u, (v, u) \in E\}$  and its in-coming neighborhood  $N^-(v) = \{u, (u, v) \in E\}$ . This leads to its out- and in-degrees  $d^+(v) = |N^+(v)|$  and  $d^-(v) = |N^-(v)|$ . These definitions make it possible to study the role of nodes with high in- and out-degrees, as well as correlations between these metrics [33, 22]. Notice that  $\sum_{v \in V} d^+(v) = \sum_{v \in V} d^-(v) = m$  is the total number of links, and so the average in- and out-degrees are equal to  $\frac{m}{n}$ .

The directed density of  $G$  is  $\frac{m}{n^2}$  since in a directed graph (with loops) there are  $n^2$  possible links. Then, a directed clique is nothing but a set of nodes all pairwise linked together with symmetrical links, which, except for loops, is equivalent to an undirected clique. One may also consider the fraction of loops present in the graph  $\frac{|\{(v,v) \in E\}|}{n}$ , as well as the fraction of symmetric links  $\frac{|\{(u,v) \in E \text{ s.t. } (v,u) \in E\}|}{m}$ .

With this definition of density, one may define the in- and out-clustering coefficient of node  $v$  as the density of its in-coming and out-going neighborhood. However, this misses the diversity of ways  $v$  may be linked to its neighbors and these neighbors may be linked together [19, 50]. Table 1 in [19] summarizes all possibilities and corresponding extensions of clustering coefficient. We focus here on two of these cases, that received more attention because they capture the presence of small cycles and the transitivity of relations. Given a node  $v$ , the cyclic clustering coefficient is the fraction of its pairs of distinct neighbors  $u$  and  $w$  with  $(u, v)$  and  $(v, w)$  in  $E$ , such that  $(w, u)$  also is in  $E$ , *i.e.*  $u, v$  and  $w$  form a cycle. The transitive coefficient of  $v$  consists in the fraction of its pairs of distinct neighbors  $u$  and  $w$  with  $(u, v)$  and  $(v, w)$  in  $E$ , such that  $(u, w)$  also is in  $E$ , *i.e.* the relation is transitive.

Finally, these extensions of node clustering coefficient directly translate to extensions of graph transitivity ratio. In the two cases explained above, this leads to the fraction of all triplets of distinct nodes  $(u, v, w)$  with  $(u, v)$  and  $(v, w)$  in  $E$ , such that in addition  $(w, v)$  is in  $E$ , or  $(v, w)$  is in  $E$ , respectively.

We define a **directed stream graph**  $S = (T, V, W, E)$  from a time interval  $T$ , a set of nodes  $V$ , a set of temporal nodes  $W \subseteq T \times V$ , and a set of temporal links  $E \subseteq T \times V \times V$ :  $(t, u, v)$  in  $E$  means that there is a link from  $u$  to  $v$  at time  $t$ , which is different from  $(t, v, u)$ . We therefore make a distinction between  $T_{u,v}$  the set  $\{t, (t, u, v) \in E\}$  and  $T_{v,u}$  the set  $\{t, (t, v, u) \in E\}$ . A loop is a triplet  $(t, v, v)$  in  $E$ . If both  $(t, u, v)$  and  $(t, v, u)$  are in  $E$ , then we say that this temporal link is symmetric. See Figure 5 for an illustration.

Directed stream graphs model the many situations where directed links occur over time and their asymmetry is important, like for instance money transfers, network traffic, phone calls, flights, moves from a place to another, and many others. In all these cases, both time information and link direction are crucial, as well as their interplay. For instance, a large number of computers sending packets at a given computer in a very short period of time is a typical signature of denial of service attack [31]. Instead, a computer sending packets to a large number of other computers in a short period of time is typical of a streaming server.

The directed stream graph  $S = (T, V, W, E)$  may be studied through the standard stream graph  $(T, V, W, \{(t, uv), (t, u, v) \in E \text{ or } (t, v, u) \in E\})$  obtained by considering each directed

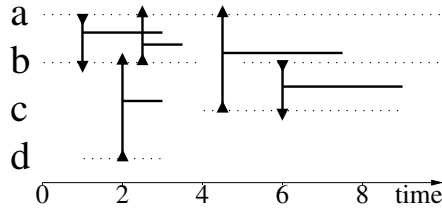


Figure 5: **An example of directed stream graph**  $S = (T, V, W, E)$  with  $T = [0, 10] \subseteq \mathbb{R}$ ,  $V = \{a, b, c, d\}$ ,  $W = [0, 10] \times \{a\} \cup ([0, 4] \cup [5, 10]) \times \{b\} \cup [4, 9] \times \{c\} \cup [1, 3] \times \{d\}$ , and  $E = [1, 3] \times \{(a, b)\} \cup [2.5, 3.5] \times \{(b, a)\} \cup [4.5, 7.5] \times \{(c, a)\} \cup [6, 9] \times \{(b, c)\} \cup [2, 3] \times \{(d, b)\}$ . Notice that  $T_{a,b} = [1, 3] \neq T_{b,a} = [2.5, 3.5]$ , and that links between  $a$  and  $b$  are symmetrical from time 2.5 to time 3.

link as undirected. Likewise,  $S$  may be studied through its induced directed graph  $G(S) = (\{v, \exists(t, v) \in W\}, \{(u, v), \exists(t, u, v) \in E\})$  and/or the sequence of its instantaneous directed graphs  $G_t = (\{v, (t, v) \in W\}, \{(u, v), (t, u, v) \in E\})$ . However, these approaches induce much information loss, and make it impossible to make subtle distinctions like the one described above for network traffic. This calls for generalizations of available concepts to this richer case.

We define the out-going neighborhood of node  $v$  as  $N^+(v) = \{(t, u), (t, v, u) \in E\}$  and its in-coming neighborhood as  $N^-(v) = \{(t, u), (t, u, v) \in E\}$ . Its in- and out-degrees are  $d^+(v) = \frac{|N^+(v)|}{|T|} = \sum_{u \in V} \frac{|T_{v,u}|}{|T|}$  and  $d^-(v) = \frac{|N^-(v)|}{|T|} = \sum_{u \in V} \frac{|T_{u,v}|}{|T|}$ . Like with directed graphs, the total number of links  $m$  is equal to  $\sum_{v \in V} d^+(v)$  as well as to  $\sum_{v \in V} d^-(v)$ .

We extend the standard stream graph density  $\frac{m}{\sum_{uv \in V \otimes V} |T_u \cap T_v|}$  into the directed stream graph density  $\frac{m}{\sum_{(u,v) \in V \times V} |T_u \cap T_v|}$ : it is the fraction of possible links that indeed exist. Then, a clique in a directed stream graph is a subset  $C$  of  $W$  such that for all  $(t, u)$  and  $(t, v)$  in  $C$ , both  $(t, u, v)$  and  $(t, v, u)$  are in  $E$ . These definitions are immediate extensions of standard stream graph concepts.

One may then define the in- and out-clustering coefficient of a node  $v$  as the density of its in- and out-neighborhoods in the directed stream graph. However, like for directed graphs, there are many possible kinds of interactions between neighbors of a node, which lead to various generalizations of clustering coefficient to directed stream graph. We illustrate this on the two definitions detailed above for a given node  $v$ . First, let us consider the set of quadruplets  $(t, u, v, w)$  with  $u, v$  and  $w$  distinct, such that  $(t, u, v)$  and  $(t, v, w)$  are in  $E$ . Then one may measure the cyclic clustering coefficient as the fraction of these quadruplets such that in addition  $(t, w, u)$  is in  $E$ , and the transitive clustering coefficient as the fraction of these quadruplets such that in addition  $(t, w, u)$  is in  $E$ .

Like with directed graphs, this leads to directed stream graph extensions of transitivity. In the two cases above, we define it as the fraction of quadruplets  $(t, u, v, w)$  with  $u, v$  and  $w$  distinct and with  $(t, u, v)$  and  $(t, v, w)$  in  $E$ , such that in addition  $(t, w, v)$  is in  $E$ , or  $(t, v, w)$  is in  $E$ , respectively.

If  $S$  is a directed graph-equivalent directed stream graph, *i.e.*  $G_t = G(S)$  for all  $t$ , then all the properties of  $S$  defined above are equal to their directed graph counterpart in  $G(S)$ .

## 5 Conclusion

Previous works extended many graph concepts to deal with weighted, bipartite or directed graphs. In addition, graphs were generalized recently to stream graphs in order to model temporal networks in a way consistent with graph theory. In this contribution, we show that weighted, bipartite or directed graphs concepts may themselves be generalized to weighted, bipartite or directed stream graphs, in a way consistent with both their graph counterparts and the stream graph formalism. This opens the way to a much richer modeling of temporal networks, and more precise case studies, in a unified framework.

Such case studies may benefit from improved modeling with either weights, different sorts of nodes, or directed links, but may also combine these extensions together. For instance, money transfers between clients and sellers are best modeled by weighted bipartite directed stream graphs. The concepts we discussed then have to be extended even further, like what has already been done for graphs for instance with the directed strength  $s^+(v) = \sum_{u \in N^+(v)} \omega(u, v)$  [39], the directed weighted clustering coefficient [14, 19], or for the study of weighted bipartite graphs [22].

One may also consider other kinds of graph extensions, like multigraphs, labelled graphs, hypergraphs, or multi-layer graphs for instance, which model important features of real-world data and already received much attention. We focused on weighted, bipartite and directed cases because they seem to be the most frequent in applications.

Likewise, we selected only a few key weighted, bipartite or directed properties in order to extend them to stream graphs. Many others remain to generalize, in particular path-related concepts like reachability, closeness, or betweenness, to cite only a few [39].

**Acknowledgements.** This work is funded in part by the European Commission H2020 FETPROACT 2016-2017 program under grant 732942 (ODYCCEUS), by the ANR (French National Agency of Research) under grants ANR-15-CE38-0001 (AlgoDiv), by the Ile-de-France Region and its program FUI21 under grant 16010629 (iTRAC).

## References

- [1] S. E. Ahnert, D. Garlaschelli, T. M. A. Fink, and G. Caldarelli. Ensemble approach to the analysis of weighted networks. *Phys. Rev. E*, 76:016101, 2007.
- [2] Jeff Alstott, Pietro Panzarasa, Mikail Rubinov, Edward T. Bullmore, and Petra E. Vértes. A unifying framework for measuring weighted rich clubs. *Scientific Reports*, 4:7258, 2014.
- [3] Shun-ichi Amano, Ken-ichiro Ogawa, and Yoshihiro Miyake. Node property of weighted networks considering connectability to nodes within two degrees of separation. *Scientific Reports*, page 8464, 2018.
- [4] I. E. Antoniou and E. T. Tsompa. Statistical analysis of weighted networks. *Discrete Dynamics in Nature and Society*, 2008.
- [5] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani. The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences*, 101(11):3747–3752, 2004.

- [6] Stefano Battiston and Michele Catanzaro. Statistical properties of corporate board and director networks. *European Physics Journal B*, 38:345–352, 2004.
- [7] Daniel Bernardes, Mamadou Diaby, Raphaël Fournier, Françoise Fogelman-Soulié, and Emmanuel Viennet. A social formalism and survey for recommender systems. *SIGKDD Explorations*, 16(2):20–37, 2014.
- [8] Phillip Bonacich. Technique for analyzing overlapping memberships. *Sociological Methodology*, 4:176–185, 1972.
- [9] Stephen P. Borgatti and Martin G. Everett. Network analysis of 2-mode data. *Social Networks*, 19(3):243–269, 1997.
- [10] Ronald L. Breiger. The duality of persons and groups. *Social Forces*, 53(2):181–190, 1974.
- [11] Carter T. Butts. A relational event framework for social action. *Sociological Methodology*, 38(1):155–200, 2008.
- [12] Conte A Candeloro L, Savini L. A new weighted degree centrality measure: The application in an animal disease epidemic. *PLoS ONE*, 11, 2016.
- [13] Arnaud Casteigts, Paola Flocchini, Walter Quattrociocchi, and Nicola Santoro. Time-varying graphs and dynamic networks. *IJPEDES*, 27(5):387–408, 2012.
- [14] G.P. Clemente and R. Grassi. Directed clustering in weighted networks: A new perspective. *Chaos, Solitons & Fractals*, 107:26–38, 2018.
- [15] Jeferson Luiz Curzel, Ricardo Lüders, Keiko Veronica Ono Fonseca, and Marcelo de Oliveira Rosa. Temporal performance analysis of bus transportation using link streams. *Mathematical Problems in Engineering*, 2019.
- [16] Patrick Doreian. A note on the detection of cliques in valued graphs. *Sociometry*, 32:237–242, 1969.
- [17] M E J Newman. Analysis of weighted networks. *Physical review E*, 70:056131, 2004.
- [18] Farnaz Zamani Esfahlani and Hiroki Sayama. A percolation-based thresholding method with applications in functional connectivity analysis. In Sean Cornelius, Kate Coronges, Bruno Gonçalves, Roberta Sinatra, and Alessandro Vespignani, editors, *Complex Networks IX*, pages 221–231, 2018.
- [19] Giorgio Fagiolo. Clustering in complex directed networks. *Phys. Rev. E*, 76:026107, 2007.
- [20] Katherine Faust. Centrality in affiliation networks. *Social Networks*, 19:157–191, 1997.
- [21] Peter Grindrod. Range-dependent random graphs and their application to modeling large small-world proteome datasets. *Phys. Rev. E*, 66:066702, 2002.
- [22] Jean-Loup Guillaume, Stevens Le Blond, and Matthieu Latapy. Statistical analysis of a P2P query graph based on degrees and their time-evolution. In *Lecture Notes in Computer Sciences (LNCS), proceedings of the 6-th International Workshop on Distributed Computing (IWDC)*, 2004.

- [23] Jean-Loup Guillaume, Stevens Le Blond, and Matthieu Latapy. Clustering in P2P exchanges and consequences on performances. In *Lecture Notes in Computer Sciences (LNCS), proceedings of the 4-th international workshop on Peer-to-Peer Systems (IPTPS)*, 2005.
- [24] Jean-Loup Guillaume and Matthieu Latapy. Bipartite structure of *all* complex networks. *Information Processing Letters (IPL)*, 90(5):215–221, 2004.
- [25] S.L. Hakimi. On the degrees of the vertices of a directed graph. *Journal of the Franklin Institute*, 279(4):290–308, 1965.
- [26] Gabriela Kalna and Desmond J. Higham. A clustering coefficient for weighted networks, with application to gene expression data. *AI Commun.*, 20:263–271, 2007.
- [27] Matthieu Latapy, Clémence Magnien, and Nathalie Del Vecchio. Basic notions for the analysis of large two-mode networks. *Social Networks*, pages 31–48, 2008.
- [28] Matthieu Latapy, Tiphaine Viard, and Clémence Magnien. Stream graphs and link streams for the modeling of interactions over time. *Social Netw. Analys. Mining*, 8(1):1–61, 2018.
- [29] Pedro G. Lind, Marta C. González, and Hans J. Herrmann. Cycles and clustering in bipartite networks. *Phys. Rev. E*, 72:056127, 2005.
- [30] Christina Lioma, Fabien Tarissan, Jakob Grue Simonsen, Casper Petersen, and Birger Larsen. Exploiting the bipartite structure of entity grids for document coherence and retrieval. In *Proceedings of the 2016 ACM International Conference on the Theory of Information Retrieval, ICTIR '16*, pages 11–20. ACM, 2016.
- [31] Johan Mazel, Pedro Casas, Romain Fontugne, Kensuke Fukuda, and Philippe Owezarski. Hunting attacks in the dark: clustering and correlation analysis for unsupervised anomaly detection. *Int. Journal of Network Management*, 25(5):283305, 2015.
- [32] Robert Meusel, Sebastiano Vigna, Oliver Lehmborg, and Christian Bizer. The graph structure in the web - analyzed on different aggregation levels. *The Journal of Web Science*, 1:33–47, 2015.
- [33] Alan Mislove, Massimiliano Marcon, Krishna P. Gummadi, Peter Druschel, and Bobby Bhattacharjee. Measurement and analysis of online social networks. In *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement*, pages 29–42, 2007.
- [34] Mark E. J. Newman, Stevens H. Strogatz, and Duncan J. Watts. Random graphs with arbitrary degree distributions and their applications. *Physics Reviews E*, 64:026118, 2001.
- [35] M.E.J. Newman. Scientific collaboration networks: I. Network construction and fundamental results. *Phys. Rev. E*, 64:016131, 2001.
- [36] M.E.J. Newman. Scientific collaboration networks: II. Shortest paths, weighted networks, and centrality. *Phys. Rev. E*, 64:016132, 2001.
- [37] Jukka-Pekka Onnela, Jari Saramäki, János Kertész, and Kimmo Kaski. Intensity and coherence of motifs in weighted complex networks. *Phys. Rev. E*, 71:065103, 2005.

- [38] Tore Opsahl. Triadic closure in two-mode networks: Redefining the global and local clustering coefficients. *Social Networks*, 35(2):159–167, 2013.
- [39] Tore Opsahl, Filip Agneessens, and John Skvoretz. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3):245 – 251, 2010.
- [40] Tore Opsahl, Vittoria Colizza, Pietro Panzarasa, and José J. Ramasco. Prominence and control: The weighted rich-club effect. *Phys. Rev. Lett.*, 101:168702, Oct 2008.
- [41] Tore Opsahl and Pietro Panzarasa. Clustering in weighted networks. *Social Networks*, 31(2):155 – 163, 2009.
- [42] Pietro Panzarasa, Tore Opsahl, and Kathleen M. Carley. Patterns and dynamics of users’ behavior and interaction: Network analysis of an online community. *Journal of the American Society for Information Science and Technology*, 60(5):911–932, 2009.
- [43] Garry Robins and Malcolm Alexander. Small worlds among interlocking directors: Network structure and distance in bipartite graphs. *Computational & Mathematical Organization Theory*, 10(1):69–94, 2004.
- [44] J. Saramäki, M. Kivelä, J.P. Onnela, K. Kaski, and J. Kertesz. Generalizations of the clustering coefficient to weighted complex networks. *Physical Review E*, 75(2):027105, 2007.
- [45] M. Ángeles Serrano, Marián Boguñá, and Alessandro Vespignani. Extracting the multi-scale backbone of complex weighted networks. *Proceedings of the National Academy of Sciences*, 106(16):6483–6488, 2009.
- [46] K. Smith, H. Azami, M. A. Parra, J. M. Starr, and J. Escudero. Cluster-span threshold: An unbiased threshold for binarising weighted complete networks in functional connectivity analysis. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2840–3, 2015.
- [47] Christoph Stadtfeld and Per Block. Interactions, actors, and time: Dynamic network actor models for relational events. *Sociological Science*, pages 318–352, 2017.
- [48] Tiphaine Viard, Raphaël Fournier-S’niehotta, Clémence Magnien, and Matthieu Latapy. Discovering patterns of interest in IP traffic using cliques in bipartite link streams. In *Proceedings of Complex Networks IX*, pages 233–241, 2018.
- [49] Yu Wang, Eshwar Ghumare, Rik Vandenberghe, and Patrick Dupont. Comparison of different generalizations of clustering coefficient and local efficiency for weighted undirected graphs. *Neural Comput.*, 29:313–331, 2017.
- [50] Stanley Wasserman and Katherine Faust. *Social network analysis: Methods and applications*. Cambridge university press, 1994.
- [51] D. Watts and S. Strogatz. Collective dynamics of small-world networks. *Nature*, 393:440–442, 1998.
- [52] Klaus Wehmuth, Eric Fleury, and Artur Ziviani. On multiaspect graphs. *Theor. Comput. Sci.*, 651:50–61, 2016.



- [53] Klaus Wehmuth, Artur Ziviani, and Eric Fleury. A unifying model for representing time-varying graphs. In *2015 IEEE International Conference on Data Science and Advanced Analytics, DSAA 2015, Campus des Cordeliers, Paris, France, October 19-21, 2015*, pages 1–10, 2015.
- [54] Bin Zhang and Steve Horvath. A general framework for weighted gene co-expression network analysis. *Statistical applications in genetics and molecular biology*, 4:1544–6115, 2005.
- [55] Peng Zhang, Jinliang Wang, Xiaojia Li, Menghui Li, Zengru Di, and Ying Fan. Clustering coefficient and community structure of bipartite networks. *Physica A: Statistical Mechanics and its Applications*, 387(27):6869–6875, 2008.
- [56] V. Zlatic, G. Bianconi, A. Díaz-Guilera, D. Garlaschelli, F. Rao, and G. Caldarelli. On the rich-club effect in dense and weighted networks. *The European Physical Journal B*, 67(3):271275, 2009.
- [57] Z. Zou. Polynomial-time algorithm for finding densest subgraphs in uncertain graphs. In *In Proceedings of MLG Workshop*, 2013.