Numerical Algorithms (2020) 83:807–832
https://doi.org/10.1007/s11075-019-00704-y

ORIGINAL PAPER

Edge importance in a network via line graphs and the matrix exponential

Omar De la Cruz Cabrera[1] · Mona Matar[1] · Lothar Reichel[1]

Received: 29 November 2018 / Accepted: 8 April 2019 / Published online: 30 April 2019

Abstract
This paper is concerned with the identification of important edges in a network, in both their roles as transmitters and receivers of information. We propose a method based on computing the matrix exponential of a matrix associated with a line graph of the given network. Both undirected and directed networks are considered. Edges may be given positive weights. Computed examples illustrate the performance of the proposed method.

Keywords Network analysis · Edge importance · Line graph · Matrix exponential

1 Introduction

Many complex systems can be modeled as networks. Informally, a network is a collection of objects or individuals, referred to as nodes or vertices, which are connected to each other in some fashion; the nature of the nodes and the connections may vary widely from application to application. Networks are formalized mathematically as graphs. Connections between nodes are referred to as edges (see Section 2.1 for more details). Like all models, network models leave out many details of reality; however, they are able to capture a substantial part of the complexity of a system in a way that is amenable to mathematical and computational analysis. Networks arise in social science, ecology, telecommunications, transportation, molecular biology, national security, and many other fields (see, e.g., [10, 21] for many examples).

✉ Omar De la Cruz Cabrera
odelacru@kent.edu

Mona Matar
mmatar2@kent.edu

Lothar Reichel
reichel@math.kent.edu

[1] Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA

One often is interested in determining which nodes are "important" in some sense. Aside from any intrinsic characteristics of a node, the importance of a node in a network is due to its connections to other nodes, as well as the importance of the nodes to which it is connected. Various measures have been proposed to quantify the importance of a node in a network. The importance commonly is referred to as the centrality (see, e.g., [3, 7, 12, 14, 19, 20]).

We are interested in measuring the importance of edges in a network. This is a problem that arises in various applications. For example, in [5], the authors are concerned about highway sections with congestion that reduces the overall highway network efficiency. Intuitively, an important edge is a good target for deletion when the goal is to disrupt the network and, therefore, worthy of protection when the goal is to preserve it. On the other hand, unimportant edges may possibly be eliminated (say, to save resources) with a small overall effect. Estrada [9] and Gutman and Estrada [18] describe applications in chemistry.

Our approach to study the importance of edges is to regard them as nodes in a line graph (see Section 3) and apply node centrality measures determined by a matrix function, in particular the matrix exponential. This works in a straightforward way for undirected networks, but becomes more complex for directed ones. We also consider the effect of edge weights.

This paper is organized as follows. Section 2 introduces graphs and associated matrices. Line graphs are discussed in Section 3. Line graphs for both undirected and directed graphs are considered. For directed graphs, we define several line graphs. Section 4 is concerned with graphs that have weighted edges. The identification of the most important edges of an undirected graph with uniform weights by using the exponential function is discussed in Section 5, while Section 6 considers the analogous task for a directed graph with uniform weights. The computation of the most important edges of a directed weighted graph is discussed in Section 7. Computed illustrations are provided in most sections. Section 8 discusses the computations required to apply the described method, and Section 9 contains concluding remarks.

2 Networks as graphs

2.1 Graphs

We represent networks by graphs. A graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists of a set $\mathcal{V}$ of nodes or vertices, and a set $\mathcal{E}$ of edges, which are the links between the nodes. If the edges have a direction, then the graph is said to be directed or oriented; otherwise, it is undirected. In the directed case, each element of $\mathcal{E}$ is an ordered pair $e = (v_1, v_2)$ of elements of $\mathcal{V}$, and we say that e incides on v_2, exsurges from v_1, and connects v_1 and v_2. In the undirected case, each element of $\mathcal{E}$ is an unordered pair $e = \{v_1, v_2\}$ of elements of $\mathcal{V}$, and we say that e incides on both v_1 and v_2, and connects v_1 and v_2 (and also v_2 and v_1). Notice that in either case, it is possible that $v_1 = v_2$; in specific cases, we will require that such "self-loops" do not exist. We assume that there are no multiple edges between any pair of vertices. Both unweighted graphs, in which each edge has weight one, and weighted graphs, in which each edge has a positive weight, are considered. For unweighted undirected graphs, the degree of a node is the number of edges inciding on it; for unweighted directed graphs, we distinguish between the indegree of a node, which is the number of edges inciding on it, and the outdegree of a node, which is the number of edges exsurging from it.

2.2 Matrix representations of graphs

Algebraic Graph Theory uses algebraic methods to study graphs. In particular, the use of Linear Algebra has proved useful for the analysis of networks. Detailed expositions can be found, for example, in [8, 10, 15]. In this section, we will develop only concepts that will be needed below; some notations and definitions are non-standard.

Consider an unweighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with vertices $\mathcal{V} = \{v_1, \ldots, v_n\}$ and edges $\mathcal{E} = \{e_1, \ldots, e_m\}$. The adjacency matrix of $\mathcal{G}$ is an $n \times n$ matrix $A = [a_{ij}]$ with $a_{ij} = 1$ if there exists an edge that connects nodes v_i and v_j, and $a_{ij} = 0$ otherwise. A walk $v_{l_1}, \ldots, v_{l_k}$ is a sequence of vertices such that $v_{l_{q-1}}$ is connected to v_{l_q} for $q = 2, 3, \ldots, k$. The $(ij)^{\text{th}}$ entry of A^k is the number of walks $v_{l_1}, \ldots, v_{l_k}$ with $l_1 = i$ and $l_k = j$. For undirected graphs, the adjacency matrix A is symmetric; for directed graphs, A generally is nonsymmetric. If the adjacency matrix of a directed unweighted graph turns out to be symmetric, then the graph can be identified with an undirected unweighted graph.

Let the graph $\mathcal{G}$ be undirected and unweighted. Then, the incidence matrix of $\mathcal{G}$ is an $n \times m$ matrix $B = [b_{ij}]$ with $b_{ij} = 1$ if e_j incides on v_i, and $b_{ij} = 0$ otherwise. Notice that each column of B has exactly two entries equal to 1 (and the rest zero), unless the corresponding edge is a self-loop, in which case exactly one entry is 1. If there are no self-loops, then $BB^T = A + D$, where $D = [d_{ij}]$ is a diagonal matrix with the diagonal entry d_{ii} equal to the degree of v_i. Here and below, the superscript ^{T} denotes transposition.

Assume now that $\mathcal{G}$ is directed and unweighted. We then define the incidence and exsurgence matrices of $\mathcal{G}$ as the $n \times m$ matrices $B^{\text{i}} = [b^{\text{i}}_{ij}]$ and $B^{\text{e}} = [b^{\text{e}}_{ij}]$, respectively, with $b^{\text{i}}_{ij} = 1$ if e_j incides on v_i, $b^{\text{e}}_{ij} = 1$ if e_j exsurges from v_i, and entries zero otherwise.

Proposition 1 *Let $\mathcal{G}$ be a directed unweighted graph, and let $B^{\text{i}} = [b^{\text{i}}_{ij}]$ and $B^{\text{e}} = [b^{\text{e}}_{ij}]$ denote the associated incidence and exsurgence matrices. Then*

(i) *each column of B^{i} and of B^{e} contains exactly one entry equal to 1, with the remaining entries of the column zero,*

(ii) *$A = [a_{jk}] = B^{\text{e}} B^{\text{i}T}$. Moreover, the entries of B^{e} and B^{i} are such that in each sum*

$$a_{jk} = \sum_{\ell=1}^{m} b^{\text{e}}_{j\ell} b^{\text{i}}_{k\ell}, \quad 1 \leq j, k \leq n, \tag{2.1}$$

there is at most one nonvanishing term. Each nonvanishing element $b^{\text{e}}_{j\ell}$ of B^{e} is paired with precisely one nonvanishing entry $b^{\text{i}}_{k\ell}$ of B^{i}. It follows that each nonvanishing entry a_{jk} can be written as $b^{\text{e}}_{j\ell} b^{\text{i}}_{k\ell}$ for precisely one index $\ell \in$

$\{1, 2, \ldots, m\}$. *Moreover, each entry* $b^{\mathrm{e}}_{j\ell}$ *and each entry* $b^{\mathrm{i}}_{k\ell}$ *determine precisely one entry* a_{jk}.

Proof Statement (i) follows from the definition of the matrices B^{i} and B^{e}. The factorization in (ii) expresses the entries of A in terms of inciding and exsurging edges. Each nonvanishing term of the sum (2.1) represents an edge from node v_j to node v_k. Since the network is assumed to have simple edges only, there can be at most one nonvanishing term in each one of the sums (2.1). The fact that each nonvanishing element $b^{\mathrm{e}}_{j\ell}$ is paired with precisely one nonvanishing entry $b^{\mathrm{i}}_{k\ell}$ follows from the observation that an exsurgent edge has to lead somewhere, and cannot have more than one destination. □

We remark that there is a more commonly used version of the incidence matrix for directed graphs, defined as an $n \times m$ matrix $\hat{B} = [\hat{b}_{ij}]$ with entries $\hat{b}_{i_1 j} = -1$ and $\hat{b}_{i_2 j} = 1$ if e_j connects v_{i_1} and v_{i_2}; see [6, 15]. This definition requires that there be no self-loops. If this is the case, then $\hat{B} = B^{\mathrm{i}} - B^{\mathrm{e}}$.

3 Line graphs

3.1 Line graphs of an undirected graph

Given an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, the line graph of $\mathcal{G}$ is an undirected graph $\mathcal{G}^* = (\mathcal{E}, \mathcal{F})$, in which there is an edge $f \in \mathcal{F}$ that connects the nodes $e, e' \in \mathcal{E}$ if and only if there is a node $v \in \mathcal{V}$ such that both e and e' incide on v in $\mathcal{G}$. Line graphs have particular characteristics. For example, each node v in $\mathcal{G}$ induces a clique (a complete subgraph, that is, a set of nodes that are all connected to each other) in $\mathcal{G}^*$, containing all $e \in \mathcal{E}$ that incide on v. In fact, the collection of cliques produced by nodes in $\mathcal{G}$ with degree at least 2 creates a partition of $\mathcal{F}$ (see [23]). If B is the incidence matrix for $\mathcal{G}$, then it can be easily shown that $E = B^T B - 2I$ is the adjacency matrix for $\mathcal{G}^*$. Throughout this paper, I stands for the identity matrix of suitable order. We refer to E as the line graph adjacency matrix.

3.2 Line graphs of a directed graph

Defining a single line graph for a directed graph is difficult, as there are some non-canonical choices. We therefore introduce four line graphs that capture different relationships between the edges of a directed graph.

Definition 1 Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a directed graph. We define the following associated line graphs:

1. The undirected co-incidence line graph $\mathcal{G}^{\vee} = (\mathcal{E}, \mathcal{F}^{\vee})$, in which distinct edges $e, e' \in \mathcal{E}$ are connected if and only if both e and e' incide on the same node in $\mathcal{G}$.
2. The undirected co-exsurgence line graph $\mathcal{G}^{\wedge} = (\mathcal{E}, \mathcal{F}^{\wedge})$, in which distinct edges $e, e' \in \mathcal{E}$ are connected if and only if both e and e' exsurge from the same node in $\mathcal{G}$.

3. The directed continuation line graph $\mathcal{G}^{\rightarrow} = (\mathcal{E}, \mathcal{F}^{\rightarrow})$, in which the edge e is connected to the edge e' if e incides on v and e' exsurges from v for some node v in $\mathcal{G}$.

For completeness, we also define a fourth line graph: the reverse continuation line graph $\mathcal{G}^{\leftarrow} = (\mathcal{E}, \mathcal{F}^{\leftarrow})$, where $(e, e') \in \mathcal{F}^{\leftarrow}$ if and only if $(e', e) \in \mathcal{F}^{\rightarrow}$. The line graph $\mathcal{G}^{\rightarrow}$ is well known; it is described, e.g., in [17, page 265]. The line graphs $\mathcal{G}^{\vee}$ and $\mathcal{G}^{\wedge}$ are new. It is worth noting that for an undirected graph $\mathcal{G}^{\vee} = \mathcal{G}^{\wedge} = \mathcal{G}^{\rightarrow} = \mathcal{G}^{\leftarrow}$, because there is no distinction between the types of edge connections described above.

The following properties are easily shown:

Proposition 2 *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a directed graph.*

1. *$\mathcal{G}^{\vee}$ and $\mathcal{G}^{\wedge}$ have no self-loops; $e \in \mathcal{E}$ has a self-loop in $\mathcal{G}^{\rightarrow}$ if and only if e is a self-loop in $\mathcal{G}$.*
2. *$\mathcal{G}^{\vee}$ partitions $\mathcal{E}$ into edge- and vertex-disjoint cliques, with one clique for each $v \in \mathcal{V}$ with positive indegree. The same happens with $\mathcal{G}^{\wedge}$, with one clique for each v with positive outdegree.*
3. *Suppose $(e, e') \in \mathcal{F}^{\rightarrow}$. Then $\{e, e''\} \in \mathcal{F}^{\vee}$ implies that $(e'', e') \in \mathcal{F}^{\rightarrow}$, and $\{e', e''\} \in \mathcal{F}^{\wedge}$ implies that $(e, e'') \in \mathcal{F}^{\rightarrow}$.*
4. *$B^{\mathrm{e}} B^{\mathrm{e}T}$ and $B^{\mathrm{i}} B^{\mathrm{i}T}$ are $n \times n$ diagonal matrices containing the outdegrees and indegrees, respectively, of the nodes of $\mathcal{G}$.*
5. *The adjacency matrices for $\mathcal{G}^{\vee}$, $\mathcal{G}^{\wedge}$, $\mathcal{G}^{\rightarrow}$, and $\mathcal{G}^{\leftarrow}$ are given by $E^{\vee} = B^{\mathrm{i}T} B^{\mathrm{i}}$, $E^{\wedge} = B^{\mathrm{e}T} B^{\mathrm{e}}$, $E^{\rightarrow} = B^{\mathrm{i}T} B^{\mathrm{e}}$, and $E^{\leftarrow} = B^{\mathrm{e}T} B^{\mathrm{i}}$, respectively. Note that for undirected graphs $E^{\vee} = E^{\wedge} = E^{\rightarrow} = E^{\leftarrow} = E$.*
6. *Let $B^{+} = [B^{\mathrm{e}} \;\; B^{\mathrm{i}}] \in \mathbb{R}^{n \times 2m}$ and define $E^{+} = B^{+T} B^{+} \in \mathbb{R}^{2m \times 2m}$. Then*

$$E^{+} = \begin{bmatrix} E^{\wedge} & E^{\leftarrow} \\ E^{\rightarrow} & E^{\vee} \end{bmatrix}. \tag{3.1}$$

The matrix $E^{\rightarrow}$ is known as the line graph adjacency matrix associated with the graph $\mathcal{G}$ (see [24]). To distinguish this matrix from the adjacency matrices for other line graphs, we refer to E^{+} as the extended line graph adjacency matrix of $\mathcal{G}$. It is a symmetric matrix, and it can be interpreted as the adjacency matrix of an undirected graph with two copies of the set of edges $\mathcal{E}$.

4 Weights

The algebraic approach is, clearly, crucial for the use of matrix functions like the matrix exponential (see Section 5.1 below) which are important tools to assess the importance of a node or an edge. However, the choice of a particular matrix representation implies some assumptions about the network, as well as limitations in the kinds of networks that can be represented. For example, a node-node adjacency matrix with all entries zero or one cannot accommodate multiple edges between the same pair of nodes, unless we are willing to represent them with weights that count the number of edges.

This section considers directed graphs with weighted edges. The weights are assumed to be positive. The interpretation of the weights depends on the application. In general, edge weights correspond to a capacity or speed of transportation, or the reciprocal of a transfer or communication cost.

Let $\tilde{A}$ denote an edge-weighted adjacency matrix. This matrix is obtained by associating a positive weight with each edge. Thus, the $(ij)^{\text{th}}$ entry of $\tilde{A}$ is the weight of the edge from node v_i to node v_j. We refer to this matrix as edge-scaled. The "unweighted" adjacency matrix A that is associated with $\tilde{A}$ has all edge weights equal to one. Thus, the entries of A belong to $\{0, 1\}$.

Theorem 1 *Let $\tilde{A} = [\tilde{a}_{ij}]$ be the $n \times n$ weighted adjacency matrix of a directed edge-weighted graph $\mathcal{G}$ of n nodes and m edges. Let $z_k > 0$ denote the weight of edge e_k for $1 \leq k \leq m$. Define the diagonal matrix Z with diagonal entries $z_1, z_2, \ldots, z_m$ in some order. Let $A = B^{\mathrm{e}} B^{\mathrm{i}T}$ be the adjacency matrix for the unweighted directed graph associated with $\mathcal{G}$, where $B^{\mathrm{i}} = [b^{\mathrm{i}}_{ij}]$ and $B^{\mathrm{e}} = [b^{\mathrm{e}}_{ij}]$ denote the incidence and exsurgence matrices for the unweighted graph; see Proposition 1. Then $\tilde{A} = B^{\mathrm{e}} Z B^{\mathrm{i}T}$. In particular, each nonvanishing entry of $\tilde{A}$ equals one of the diagonal entries of the matrix Z, and each diagonal entry of Z corresponds to precisely one nonvanishing entry of $\tilde{A}$.*

Proof The result follows from Proposition 1. Each column of B^{e} has precisely one nonvanishing entry 1. Therefore, $B^{\mathrm{e}} Z$ is a weighted exsurgence matrix, with each diagonal entry z_j appearing in exactly one column. The theorem now follows from part (ii) of Proposition 1 with B^{e} replaced by $B^{\mathrm{e}} Z$. □

The matrix $Z = \mathrm{diag}[z_1, z_2, \ldots, z_m]$ of Theorem 1 can be factored according to $Z = Z^{\mathrm{e}} Z^{\mathrm{i}}$, where $Z^{\mathrm{e}} = \mathrm{diag}[z^{\mathrm{e}}_1, z^{\mathrm{e}}_2, \ldots, z^{\mathrm{e}}_m]$ and $Z^{\mathrm{i}} = \mathrm{diag}[z^{\mathrm{i}}_1, z^{\mathrm{i}}_2, \ldots, z^{\mathrm{i}}_m]$ have positive diagonal entries. Then $\tilde{B}^{\mathrm{e}} = B^{\mathrm{e}} Z^{e}$ is a weighted exsurgence matrix, such that each entry z^{e}_j of Z^{e} appears in exactly one column. Similarly, $\tilde{B}^{\mathrm{i}} = B^{\mathrm{i}} Z^{i}$ is a weighted incidence matrix, such that each entry z^{i}_j of Z^{i} appears in precisely one column. The weighted line graph $\mathcal{G}^{\rightarrow}$ is defined by the matrix $\tilde{E}^{\rightarrow} = \tilde{B}^{\mathrm{i}T} \tilde{B}^{\mathrm{e}}$. We will use the matrices $Z^{\mathrm{e}} = Z^{\mathrm{i}} = Z^{1/2}$ in the computed examples reported in this paper, but other choices of Z^{e} and Z^{i} also are possible.

For certain (unweighted) adjacency matrices $A = B^{\mathrm{e}} B^{\mathrm{i}T}$ and weighting matrices Z, the weighted adjacency matrix $\tilde{A} = B^{\mathrm{e}} Z B^{\mathrm{i}T}$ can be expressed by row and/or column scaling of A. We summarize this in the following proposition, whose proof is straightforward.

Proposition 3 *Let $A = [a_{ij}] = B^{\mathrm{e}} B^{\mathrm{i}T}$ be an $n \times n$ unweighted adjacency matrix with m edges, let $\tilde{A} = [\tilde{a}_{ij}] = B^{\mathrm{e}} Z B^{\mathrm{i}T} \in \mathbb{R}^{n \times n}$ be an associated edge-weighted adjacency matrix, and let $W = diag[w_1, w_2, \ldots, w_n]$ be a diagonal matrix with positive diagonal entries. Then (i) $\tilde{A} = AW$ if and only if the weighting matrix Z is* *such that* $\tilde{a}_{ij} = a_{ij} w_j$ *for all* $1 \le i, j \le n$*, (ii)* $\tilde{A} = WA$ *if and only if the weighting matrix* Z *is such that* $\tilde{a}_{ij} = a_{ij} w_i$ *for all* $1 \le i, j \le n$*, (iii)* $\tilde{A} = WAW$ *if and only if the weighting matrix* Z *is such that* $\tilde{a}_{ij} = a_{ij} w_i w_j$ *for all* $1 \le i, j \le n$*.*

The significance of the above result is that we may express the edge-weighting defined by the edge-weighted adjacency matrix $\tilde{A}$ in terms of row or column scaling of the "unweighted" adjacency matrix A.

4.1 Example

Consider the cyclic upper Hessenberg adjacency matrix

$$A = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ & 0 & \ddots & \vdots & 0 & 0 \\ & & \ddots & & 0 & 0 \\ \mathbf{0} & & & 0 & 1 & 0 \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Then, for any weighting matrix Z, the edge-weighted matrix $\tilde{A}$ can be expressed as $\tilde{A} = AW_1$ and $\tilde{A} = W_2 A$ for suitable diagonal weighting matrices W_1 and W_2 with positive diagonal entries.

5 Computing the most important edges in an undirected network by the matrix exponential

5.1 Review of the adjacency matrix exponential

In [12], the authors use matrix functions to rank the nodes in undirected networks. For a symmetric unweighted adjacency matrix $A \in \mathbb{R}^{n \times n}$, the $(ij)^{\text{th}}$ entry of A^p gives the number of walks of length p between nodes v_i and v_j. The $(ij)^{\text{th}}$ entry of the matrix function

$$f(A) = \sum_{p=0}^{\infty} c_p A^p \tag{5.1}$$

gives a weighted average of the number of walks of various lengths between the nodes v_i and v_j. The coefficients c_p are chosen to penalize walks that traverse many edges, because such walks are considered less important than walks that traverse few edges. The coefficients c_p are therefore, generally, chosen to be nonnegative and decreasing as a function of p. A common choice is $c_p = 1/p!$. Then $f(A) = \exp(A)$. The term $c_0 I$ is of no significance.

The entry of $[f(A)]_{ii}$ defines the subgraph centrality of the node v_i, and the entry $[f(A)]_{ij}$, with $i \ne j$, defines the communicability between the nodes v_i and v_j. A relatively large value $[f(A)]_{ii}$ indicates that node v_i is important, and a relatively large value $[f(A)]_{ij}$, $i \ne j$, suggests that communication between the nodes v_i and

v_j is relatively easy (see, e.g., [10, 12]). Another importance measure for nodes is furnished by row sums of the function (5.1), i.e., by the entries $[f(A)\mathbf{1}]_i$, $1 \leq i \leq n$, where $\mathbf{1} = [1, 1, \ldots, 1]^T$. A relatively large value of $[f(A)\mathbf{1}]_i$ suggests that the node v_i is important (see [4] as well as [14, Section 2] and [19]).

5.2 Exponential of the line graph adjacency matrix for undirected graphs

We seek to rank the edges of graphs and first consider undirected graphs. Let $E \in \mathbb{R}^{m\times m}$ be the line graph adjacency matrix for an undirected graph with n nodes and m edges (see Section 3.1 for its definition). We compute the matrix exponential of E and obtain, analogously to (5.1),

$$\exp(E) = \sum_{p=0}^{\infty} \frac{1}{p!} E^p. \tag{5.2}$$

Similarly to the discussion in [12] and above, we can interpret the entries of $\exp(E)$ as indicators of the centrality and communicability of the edges of the graph. For instance, a relatively large diagonal entry $[\exp(E)]_{kk}$ indicates that the edge e_k is important. Similarly, a relatively large off-diagonal entry $[\exp(E)]_{kl}$, $k \neq l$, suggests that information that travels via edge e_k is likely to also travel via edge e_l. One also may define the centrality of the edge e_k as $[\exp(E)\mathbf{1}]_k$. We will use the latter measure and define the edge line graph centrality of an edge e_k between the nodes v_i and v_j as

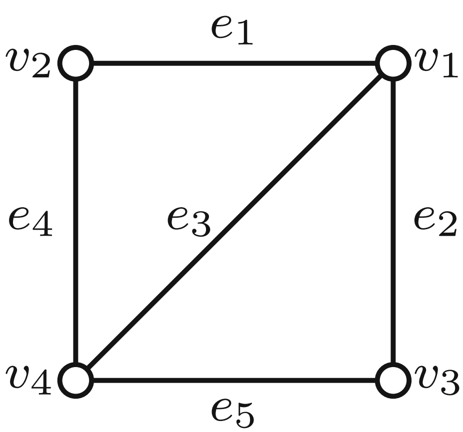
$$^e\mathrm{LC}_k = [\exp(E)\mathbf{1}]_k. \tag{5.3}$$

The following examples illustrate the ranking of edges using this measure.

5.2.1 Example

Consider the graph of Fig. 1 and the associated line graph of Fig. 2. We can see from the graph and the line graph that edge e_3 is the most important edge, because it is directly connected to all the other edges in the graph. The symmetry of the graph and

Fig. 1 Graph of Example 5.2.1



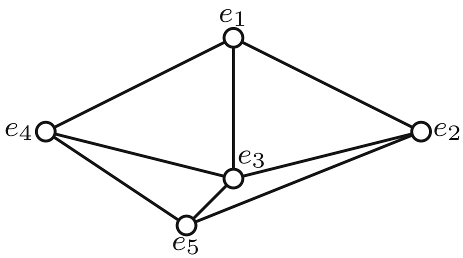


Fig. 2 Line graph of Example 5.2.1

the line graph suggests that the edges e_1, e_2, e_4, and e_5 are equally important. Table 1 confirms this by computing the edge line graph centrality (5.3).

5.2.2 Example

The graph for this example is shown in Fig. 3. Visual inspection suggests that the edge e_2 is the most important edge of the graph. Looking at the graph, one might guess that e_1 is the next most important edge. However, Table 2 shows the edges e_3, e_4, and e_6 to be ranked higher. The line graph for the graph, shown in Fig. 4, sheds light on this ordering. It shows the edges e_2, e_3, e_4, and e_6 to be well connected. This example illustrates that looking at a graph may not always give a good idea of which edges are the most important ones.

5.3 A comparison of downdating methods for undirected graphs

Let the adjacency matrix A define an undirected graph. We would like to remove an edge from this graph so that the total network communicability, defined by

$$\mathrm{TC}(A) = \mathbf{1}^T \exp(A)\mathbf{1}, \tag{5.4}$$

and considered by Benzi and Klymko [4], decreases the least. This problem is discussed by Arrigo and Benzi [2]. Another communicability measure used in the latter paper is the Estrada Index [11, 12] defined by

$$\mathrm{EE}(A) = \mathrm{trace}(\exp(A)) = \sum_{i=1}^{n} [\exp(A)]_{ii}. \tag{5.5}$$

Table 1 Ranking of the edges of Example 5.2.1 by the centrality measure (5.3)

Edge e_k	${}^{e}LC_k$
e_3	29.73
e_2	24.11
e_4	24.11
e_1	24.11
e_5	24.11

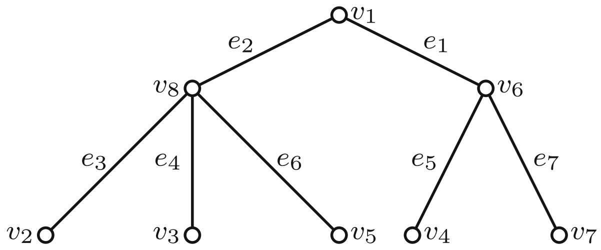


Fig. 3 Graph of Example 5.2.2

There are several ways to measure the importance of an edge. Arrigo and Benzi [2] define the edge total communicability centrality of an existing edge between the nodes v_i and v_j as

$$^{e}\mathrm{TC}(v_i, v_j) = [\exp(A)\mathbf{1}]_i[\exp(A)\mathbf{1}]_j, \tag{5.6}$$

and the edge subgraph centrality of an edge between the nodes v_i and v_j as

$$^{e}\mathrm{SC}(v_i, v_j) = [\exp(A)]_{ii}[\exp(A)]_{jj}. \tag{5.7}$$

In this paper, we propose to compute the exponential of the line graph adjacency matrix and remove edges with the lowest centrality defined by (5.3). For this purpose, we introduce the total line graph centrality measure of the network with line graph adjacency matrix E,

$$\mathrm{LC}(A) = \sum_{k=1}^{m}\{\exp(E)\mathbf{1}\}_k = \mathbf{1}^T \exp(E)\mathbf{1}. \tag{5.8}$$

This approach is shown to be competitive with the approach of removing edges with the lowest edge centralities defined by (5.6) and (5.7), as proposed in [2], and we will show examples where it decreases the total communicability less than the other approaches.

The following examples illustrate the use of the measures described above. We show two examples, for which removing the three least important edges from a graph, using the exponential of the edge line graph to identify these edges, results in a graph with a higher communicability with respect to all the measures (5.3), (5.5), and (5.6).

Table 2 Ranking of the edges of Example 5.2.2 by using the centrality measure (5.3)

Edge e_k	$^{e}LC_k$
e_2	28.07
e_4	23.68
e_3	23.68
e_6	23.68
e_1	17.23
e_7	10.61
e_5	10.61

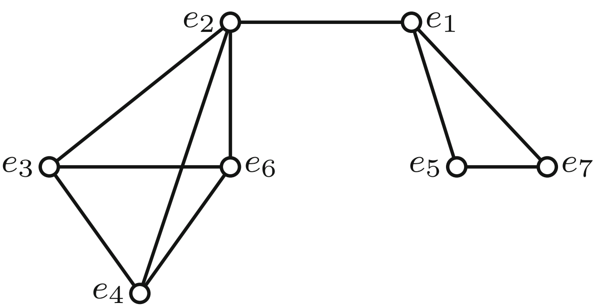


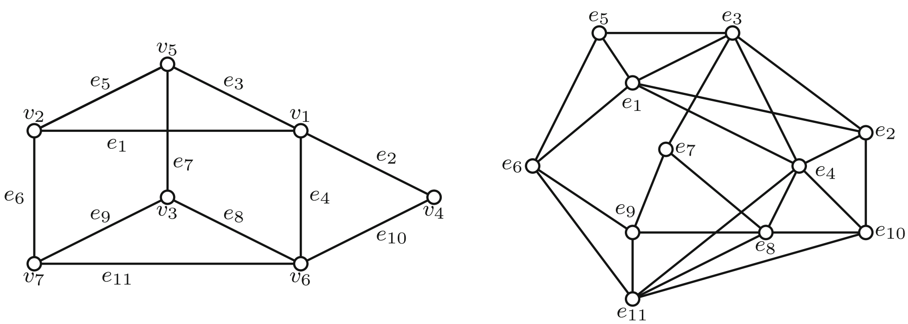
Fig. 4 Line graph of Example 5.2.2

In the first example, we compute the exponential of the line graph, remove the three least important edges, and then calculate the measures after removal. This approach is referred to as nongreedy downdating in [4]. The second example recalculates the measures after the removal of each edge, to make the decision of removal of the next edge based on the updated graph. This approach is called greedy downdating in [4].

5.3.1 Nongreedy downdating example

Consider the connected network shown in Fig. 5. We use different edge centrality measures to decide which three edges should be eliminated from the network so that the total communicability is decreased the least, with the constraint that the graph obtained after edge removal should be connected.

Table 3 shows the edges with the lowest centrality measured by several centrality measures defined above. The first method computes the edge total communicability (5.6), the second one ranks the edges according the edge subgraph centrality measure (5.7), and the third one uses the edge line graph centrality (5.3). We then remove



(a) The graph for Example 5.3.1 (b) The line graph of Example 5.3.1

Fig. 5 **a**, **b** Undirected graph and associated line graph

Table 3 The least important edges in Fig. 5a ordered according to increasing importance by using the edge total communicability, the edge subgraph centrality, and the edge line graph centrality

Edge e_k	${}^{e}\mathrm{TC}(v_i, v_j)$	Edge e_k	${}^{e}\mathrm{SC}(v_i, v_j)$	Edge e_k	${}^{e}\mathrm{LC}_k$
e_2	549.51	e_2	18.82	e_7	85.08
e_{10}	549.51	e_{10}	18.82	e_9	85.08
e_6	581.11	e_6	18.96	e_5	85.08
e_9	581.11	e_9	18.96	e_6	97.91

The edge e_k connects the nodes v_i and v_j

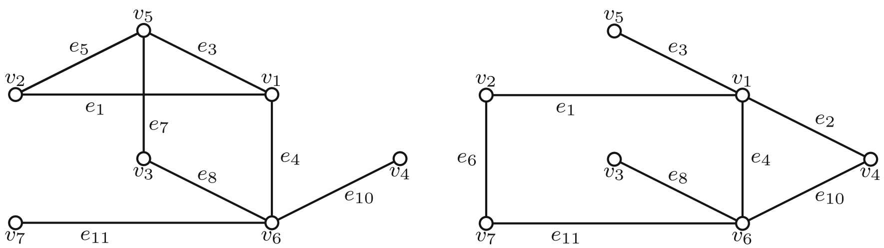
the three edges with the lowest centrality and determine which method decreases the total communicability the least.

When using the edge total communicability centrality measure or the edge subgraph centrality measure, as suggested in [2], the first four columns of Table 3 indicate that the edges e_2, e_6, and e_9 are to be removed. The graph obtained after removal of these edges is shown in Fig. 6a. We note that while some edges have smaller centrality measure than the ones we remove, those edges are not removable because this would disconnect the graph. We then use the edge line graph centrality measure (see the last two columns of Table 3 for the edge ranking) to conclude that the edges e_5, e_7, and e_9 are to be removed. The graph obtained after removal of these edges is shown in Fig. 6b.

We compute the centrality measures $\mathrm{TC}(A)$, $\mathrm{EE}(A)$, and $\mathrm{LC}(A)$ for both graphs of Fig. 6 and report them in Table 4. In each case, we obtain better connectivity of the graph of Fig. 6b than of the graph of Fig. 6a, i.e., when we use the edge line graph centrality measure to determine which edges to remove.

5.3.2 Greedy downdating example

We again consider the network in Fig. 5 and remove the three edges with the lowest centrality using the same measures as in the previous example, but here we update



(a) Graph obtained by removing edges using the technique in [2].

(b) Graph obtained by removing edges using the exponential of the line graph matrix.

Fig. 6 **a**, **b** The nongreedy downdated network of Fig. 5a

Table 4 Comparison of various network connectivity measures for the nongreedily downdated graphs in Fig. 6 obtained by the method in [2] versus the using the exponential of the line graph

	Graph in Fig. 6a	Graph in Fig. 6b
Total network communicability TC(A)	86.22	92.31
Estrada Index EE(A)	20.68	21.16
Total line graph centrality LC(A)	2144.41	3104.77

The reduced graph is required to be connected

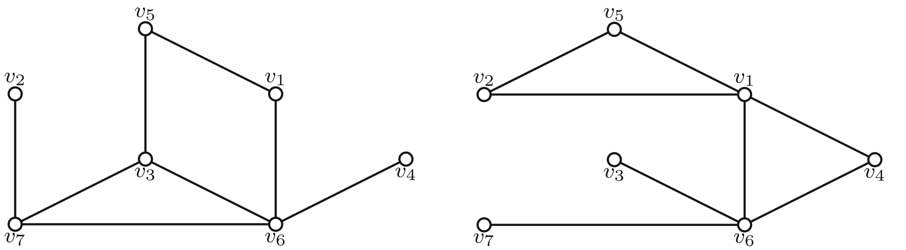
the measures after each edge removal. Like in the previous example, we require the graph obtained after edge removal to be connected. The edge total communicability centrality and the edge subgraph centrality yield the graph in Fig. 7a, while using the exponential of the adjacency matrix for the line graph gives the graph of Fig. 7b.

We compute the network centrality measures TC(A), EE(A), and LC(A) for the graphs of Fig. 7. These measures are reported in Table 5 and are larger than the entries of Table 4 for the nongreedy algorithm. This means that the greedy approach gives graphs with better connectivity than the nongreedy one. Moreover, the graph of Fig. 7b has better connectivity than the graph of Fig. 7a. This also holds for numerous other examples. We conclude that edge removal by the greedy approach based on using the exponential of the line graph adjacency matrix generally is preferable.

5.3.3 Downdating example when the reduced graph is not required to be connected

In Examples 5.3.1 and 5.3.2, we downdated the graphs with the requirement that the reduced graph be connected, as was done in [4]. In this example, we remove three edges with the smallest centrality of the graph of Fig. 5a without requiring the reduced graph to be connected. The edge centrality measures used are the same as in Examples 5.3.1 and 5.3.2.

When measuring edge importance by the edge total communicability centrality and the edge subgraph centrality, we obtain the disconnected reduced graph in Fig. 8



(a) Graph obtained by using the greedy technique in [2].

(b) Graph obtained by using the exponential of the line graph.

Fig. 7 **a**, **b** The greedy downdated network of Fig. 5a

Table 5 Comparison of network connectivity measures for the greedily downdated graphs in Fig. 7 obtained by the method in [2] versus the use of the exponential of the adjacency matrix for the line graph

	Graph in Fig. 7a	Graph in Fig. 7b
Total network communicability TC(A)	89.04	92.52
Estrada Index EE(A)	20.91	22.11
Total line graph centrality LC(A)	2507.77	3119.72

The reduced graph is required to be connected

when we apply both the nongreedy and greedy methods described in Examples 5.3.1 and 5.3.2, respectively. Using the exponential of the adjacency matrix for the line graph gives the graph of Fig. 6b for the nongreedy method, and the graph of Fig. 7b for the greedy method.

Table 6 reports the network centrality measures TC(A), EE(A), and LC(A) for the graphs obtained. The graphs of Figs. 6b and 7b have larger connectivity measures than the graph of Fig. 8, except for the Estrada Index, which is smaller but close. We conclude that edge removal by using the exponential of the line graph adjacency matrix E can be competitive also when we do not require the reduced graph to be connected.

5.4 A comparison of downdating methods for directed graphs

In this section, the adjacency matrix A corresponds to a directed graph. We would like to remove edges of a directed graph so that communicability of the network is decreased the least. A measure used by Arrigo and Benzi [1] is the total network communicability, the same as in (5.4), with A nonsymmetric. They also use a measure that takes into account alternating walks in a directed network. This measure assigns the network a value of its hub strength, referred to as the total hub communicability,

$$\mathrm{T}_h\mathrm{C}(A) = \mathbf{1}^T \cosh(\sqrt{AA^T})\mathbf{1},$$

and a value of its authority strength, referred to as the total authority communicability,

$$\mathrm{T}_a\mathrm{C}(A) = \mathbf{1}^T \cosh(\sqrt{A^TA})\mathbf{1}.$$

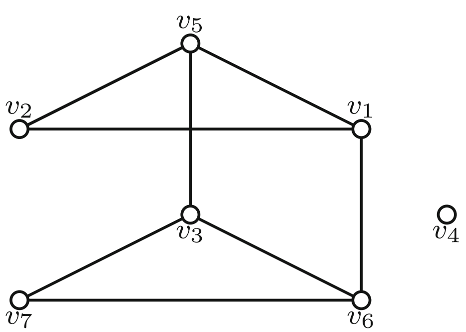


Fig. 8 Downdated graph of Fig. 5a obtained by using the nongreedy or greedy techniques in [2] without requiring the network to stay connected

Table 6 Comparison of network connectivity measures after downdating the graph in Fig. 5a using the methods in [2] versus the use of the exponential of the adjacency matrix for the line graph

	Figure 8	Figure 6b	Figure 7b
Total network communicability TC(A)	91.5	92.31	92.52
Estrada Index EE(A)	22.33	21.16	22.11
Total line graph centrality LC(A)	2501.24	3104.77	3119.72

The reduced graph is not required to be connected

To evaluate the total communicability of the network, the authors of [1] compute the sum $\mathrm{T}_h\mathrm{C}(A) + \mathrm{T}_a\mathrm{C}(A)$. This method is based on the idea of expressing a directed graph by an undirected graph with twice the number of nodes, and applying the matrix exponential to the adjacency matrix associated with the latter graph,

$$\mathcal{A} = \begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}; \tag{5.9}$$

see [3] for details.

This section extends the total line graph centrality measure defined in (5.8) to directed graphs. Taking into account the existing edges e_k, we use the line graph adjacency matrix $E^{\rightarrow}$ to define the total line graph centrality by

$$\mathrm{LC}(A) = \sum_{k=1}^{m} [\exp(E^{\rightarrow})\mathbf{1}]_k = \mathbf{1}^T \exp(E^{\rightarrow})\mathbf{1}. \tag{5.10}$$

Note that when assessing the effect of removing an edge on the communicability of a network, the measure $\mathrm{T}_h\mathrm{C}(A)+\mathrm{T}_a\mathrm{C}(A)$ may yield significantly different values than the measures TC(A) and LC(A), since the latter measures capture the flow deeply in the network, whereas the first measure does not.

We review some of the edge measures used in [1] for ranking edges to identify the least connected edges, which are to be removed. The edge total communicability centrality of an existing edge going from node v_i to v_j is given by

$$^e\mathrm{TC}(v_i, v_j) = [\exp(A)\mathbf{1}]_i [\mathbf{1}^T \exp(A)]_j,$$

and its application to the corresponding matrix (5.9) yields the measure

$$^e\mathcal{TC}(v_i, v_j) = [\exp(\mathcal{A})\mathbf{1}]_i [\mathbf{1}^T \exp(\mathcal{A})]_{n+j}.$$

Based on ideas in [3], the authors of [1] use the generalized hyperbolic sine to define the edge total communicability,

$$^e\mathrm{gTC}(v_i, v_j) = C_h(i)C_a(j),$$

where the total hub communicability of node v_i and the total authority communicability of node v_j are defined by

$$C_h(i) = [U \sinh(\Sigma)V^T\mathbf{1}]_i \quad \text{and} \quad C_a(j) = [V \sinh(\Sigma)U^T\mathbf{1}]_j,$$

respectively. The matrices U, Σ, and V are determined by the singular value decomposition $A = U\Sigma V^T$ (see, e.g., [16] for a discussion of the latter).

Analogously to our definition of the edge centrality (5.3), we define the edge line graph centrality of a directed edge e_k pointing from node v_i to v_j as

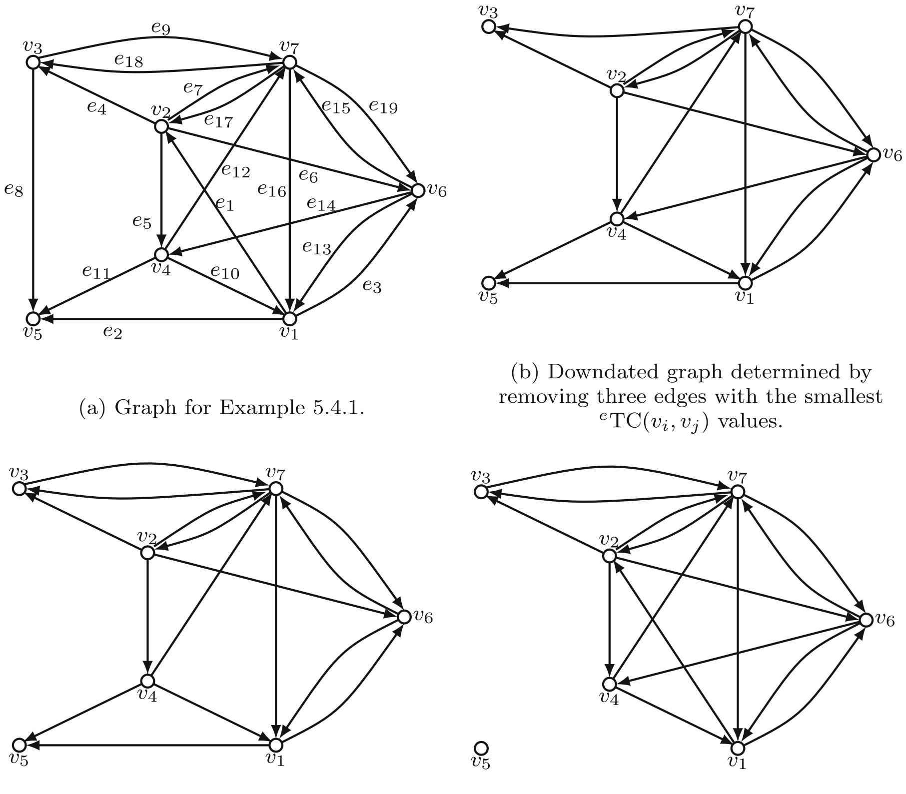
$$^e\mathrm{LC}_k = [\exp(E^{\rightarrow})\mathbf{1}]_k;$$

see also (5.10).

5.4.1 Nongreedy downdating example

Consider the directed graph in Fig. 9a. We would like to remove the three edges with the lowest edge centrality as measured by ^{e}TC, $^e\mathcal{TC}$, ^{e}gTC, or ^{e}LC. The graphs obtained after removing the three edges identified by these measures are shown in Fig. 9b–d. In this example, we allow the resulting graph to be disconnected.

We compute the communicability of the graphs of Fig. 9. Table 7 shows that when removing the edges with the lowest edge line graph centrality, all the measures of network communicability have a larger value than when we remove edges using one



(a) Graph for Example 5.4.1.

(b) Downdated graph determined by removing three edges with the smallest $^e\mathrm{TC}(v_i, v_j)$ values.

(c) Downdated graph determined by removing three edges with the smallest $^e\mathcal{TC}(v_i, v_j)$ and $^e\mathrm{gTC}(v_i, v_j)$ values.

(d) Downdated graph determined by removing three edges with the smallest $^e\mathrm{LC}_k$ values.

Fig. 9 **a–d** The nongreedy downdating of a directed graph

Table 7 Comparison of network communicability measures for the downdated graphs in Fig. 9

	Figure 9b	Figure 9c	Figure 9d
Total network communicability $\mathrm{TC}(A)$	69.62	69.32	94.97
$\mathrm{T}_h\mathrm{C}(A) + \mathrm{T}_a\mathrm{C}(A)$	61	61	62
Total line graph centrality $\mathrm{LC}(A)$	160.13	158.99	261.41

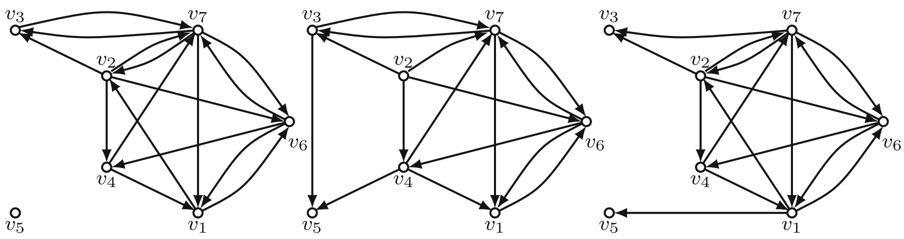
of the other edge centrality measures. While this is not the case for every example we may encounter, it nevertheless suggests that edge line graph centrality is an important measure of edge centrality.

5.4.2 Greedy downdating example

This example differs from the previous one only in that we now update the communicability measures of the graph after each edge removal. Removing edges based on the measure ${}^e\mathrm{LC}(v_i, v_j)$ performed as well as removing edges by using the measure ${}^e\mathrm{TC}(v_i, v_j)$. These two measures outperform the other measures. The downdated graphs obtained by using the various edge centrality measures are displayed in Fig. 10, and the communicability measures for the downdated measures are reported in Table 8. Note that the graph of Fig. 10a is disconnected, while the graphs of Fig. 10b and c are connected.

6 Computing the most important edges in a directed unweighted network using the matrix exponential

We introduced in Section 3 three types of connections between adjacent edges of a directed graph. Each one of these connection types yields a line graph, and the exponentials of the adjacency matrices associated with these line graphs define centrality



(a) Downdated graph determined by removing three edges with the smallest ${}^e\mathrm{TC}(v_i, v_j)$ and ${}^e\mathrm{LC}_k$ values.

(b) Downdated graph determined by removing three edges with the smallest ${}^e\mathcal{TC}(v_i, v_j)$ values.

(c) Downdated graph determined by removing three edges with the smallest ${}^e\mathrm{gTC}(v_i, v_j)$ values.

Fig. 10 **a**–**c** The greedy downdating of the directed graph in Fig. 9a

Table 8 Comparison of network communicability measures for the downdated graphs in Fig. 10

	Figure 10a	Figure 10b	Figure 10c
Total network communicability TC(A)	94.97	60.91	84.78
$\mathrm{T}_h\mathrm{C}(A) + \mathrm{T}_a\mathrm{C}(A)$	62	61	61
Total line graph centrality LC(A)	261.41	127.59	219.59

measures for edges. We are interested in determining which edges are the most and least important ones in a directed network and investigate the application of the matrix exponential to these line graph adjacency matrices.

6.1 The exponential of the extended line graph adjacency matrix E^+

We defined in Section 3.2 the symmetric line graph adjacency matrix E^+ to represent the connections between the edges in a directed graph $\mathcal{G}$ by using the corresponding undirected bipartite graph. A centrality measure for edges based on E^+ is given by the matrix exponential

$$\exp(E^+) = \sum_{p=0}^{\infty} \frac{1}{p!} E^{+p},$$

which has an interesting structure. To exploit this structure, it is helpful to introduce the notion of the total degree of a node in $\mathcal{G}$, which is the sum of the indegree and outdegree of the node.

Proposition 4 *Let the extended line graph adjacency matrix E^+ for a directed unweighted graph $\mathcal{G}$ be given by (3.1). Then $E^+ = B^{+T} B^+$, where $B^+ = [B^{\mathrm{e}} \;\; B^{\mathrm{i}}]$; see Proposition 2. Let m_i denote the total degree of node v_i of $\mathcal{G}$. Then*

$$\exp(E^+) = I + B^{+T} \begin{bmatrix} \frac{\exp(m_1)-1}{m_1} & 0 & 0 & \cdots \\ 0 & \frac{\exp(m_2)-1}{m_2} & 0 & \cdots \\ 0 & 0 & \frac{\exp(m_3)-1}{m_3} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} B^+. \tag{6.1}$$

Thus, the structure of the matrix $\exp(E^+)$ *is related to the structure of $B^{+T} B^+$, but the entries of the former matrix depend on the total degree of the nodes. Since the function $x \to \frac{\exp(x)-1}{x}$, $x > 0$, is increasing, the importance of a node generally increases with its total degree.*

Proof We have

$$(E^+)^2 = (B^{+T} B^+)(B^{+T} B^+) = B^{+T} (B^+ B^{+T}) B^+ = B^{+T} (B^{\mathrm{e}} B^{\mathrm{e}T} + B^{\mathrm{i}} B^{\mathrm{i}T}) B^+,$$

where we note that the matrix $M = B^{\mathrm{e}} B^{\mathrm{e}T} + B^{\mathrm{i}} B^{\mathrm{i}T}$ is diagonal. Its nontrivial entries are the total degrees of the nodes of $\mathcal{G}$. It is easy to show that for any integer $p \geq 1$, we have $(E^+)^p = B^{+T} M^{p-1} B^+$, from which it follows that

$$\begin{aligned}\exp(E^+) &= I + (B^{+T} B^+) + \frac{(B^{+T} M B^+)}{2!} + \frac{(B^{+T} M^2 B^+)}{3!} + \ldots \\ &= I + B^{+T} M^{-1} (\exp(M) - I) B^+.\end{aligned}$$

This shows (6.1). □

Let us take the simple example in Fig. 11 to test how meaningful the use of $\exp(E^+)$ is for measuring the importance of the edges of this graph. Proposition 2 describes the construction of the adjacency matrix E^+ associated with the graph. Its matrix exponential, with non-vanishing entries rounded to three significant decimal digits, is given by

$$\exp(E^+) = \begin{bmatrix} 2.72 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4.19 & 0 & 3.19 & 0 & 0 \\ 0 & 0 & 4.19 & 0 & 3.19 & 0 \\ 0 & 3.19 & 0 & 4.19 & 0 & 0 \\ 0 & 0 & 3.19 & 0 & 4.19 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2.72 \end{bmatrix}.$$

The top right quarter of the above matrix represents the propagation of signals from edges traveling through the network. The fact that the entry $[\exp(E^+)]_{16}$ vanishes indicates that the edge e_1 cannot connect to the edge e_3 by any number of steps. But we easily see in Fig. 11 that these edges are connected by two steps: e_1 to e_2 followed by e_2 to e_3. This illustrates that the matrix $\exp(E^+)$ is poorly suited to indicate the importance of edges. We conclude that a matrix other than E^+ is needed to make the exponential function meaningful for edges.

6.2 The exponential of the line graph adjacency matrix $E^{\rightarrow}$

Thulasiraman and Swamy [24] discuss the line graph adjacency matrix $E^{\rightarrow} = B^{\mathrm{i}T} B^{\mathrm{e}}$. It has an entry 1 in position (i, j) if and only if the edge e_j passes information to the edge e_i through a node, i.e., if and only if the head of edge e_j coincides with the tail of edge e_i. The entries of the matrix $(E^{\rightarrow})^2$ tell us whether information is passed from an edge to another one through two nodes. In other words, $[(E^{\rightarrow})^2]_{ij} = 1$ if there exists an edge pointing from the target node of e_i to the source node of e_j. Similarly, $[(E^{\rightarrow})^p]_{ij}$ counts the number of ways information is transferred from the edge e_i to the edge e_j through p nodes. The matrix exponential $\exp(E^{\rightarrow})$ is a weighted sum of positive powers of $E^{\rightarrow}$, with transfers of information via many nodes having a smaller weight than transfers via few nodes (cf. (5.2)). Note that the matrix $E^{\rightarrow}$ generally is nonsymmetric. Each row of $\exp(E^{\rightarrow})$ represents an edge in its emitter role, and each column expresses its role in receiving information. Similar to the

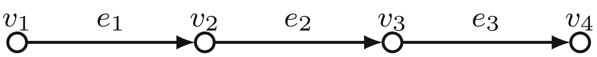
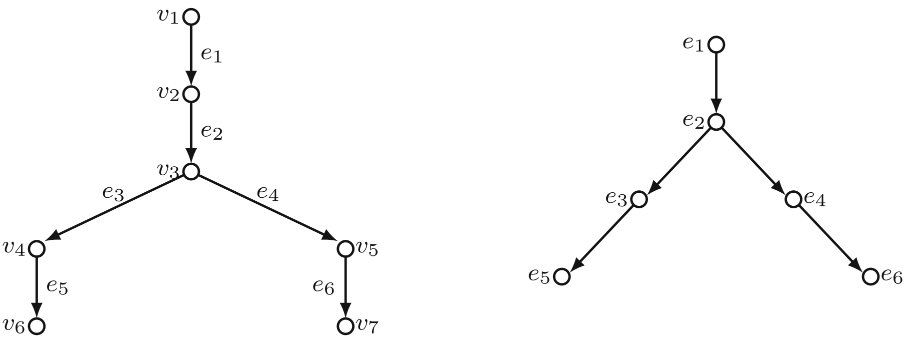


Fig. 11 Simple directed network



(a) Graph of a directed tree. (b) Line graph of a directed tree.

Fig. 12 **a**, **b** Graph and line graph for Example 6.2.1

discussion in Section 5.3, we can determine the ability of an edge to transmit information through the network by ordering the elements of the row sums of the matrix $\exp(E^{\rightarrow})$, i.e., of $\exp(E^{\rightarrow})\mathbf{1}$. The largest entry of this vector corresponds to the most important transmitter. Similarly, the vector $\mathbf{1}^T \exp(E^{\rightarrow})$ provides an ordering of the edges in their role as information receivers, where the largest entry corresponds to the most important receiver. We will use these measures in the following examples.

We note that powers of $E^{\vee}$ and $E^{\wedge}$ defined in Proposition 2 do not add any information about propagation through the network because of the nature of connections between edges that they provide. Moreover, $E^{\leftarrow} = (E^{\rightarrow})^T$. Therefore, calculating $\exp(E^{\leftarrow})\mathbf{1}$ is equivalent to evaluating $\mathbf{1}^T \exp(E^{\rightarrow})$.

6.2.1 Example

We consider the ranking of edges of a graph $\mathcal{G}$ that is the small tree shown in Fig. 12a. The corresponding line graph is displayed in Fig. 12b. Table 9 displays the centrality measures $\exp(E^{\rightarrow}\mathbf{1})$ and $\mathbf{1}^T \exp(E^{\rightarrow})$. The edge e_2 contributes the most to broadcasting information through the network, because it is the only edge that points to a node from which two edges emerge. The edge e_1 is the next most important edge, because it is one step further away from the split at the node v_3 than the edge e_2. The

Table 9 Ranking of the edges of Fig. 12 using the exponential of the matrix $E^{\rightarrow}$

Edge e_k	$[\exp(E^{\rightarrow})\mathbf{1}]_k$	Edge e_k	$[\mathbf{1}^T \exp(E^{\rightarrow})]_k$
e_2	4	e_5	2.67
e_1	3.33	e_6	2.67
e_3	2	e_3	2.5
e_4	2	e_4	2.5
e_5	1	e_2	2
e_6	1	e_1	1

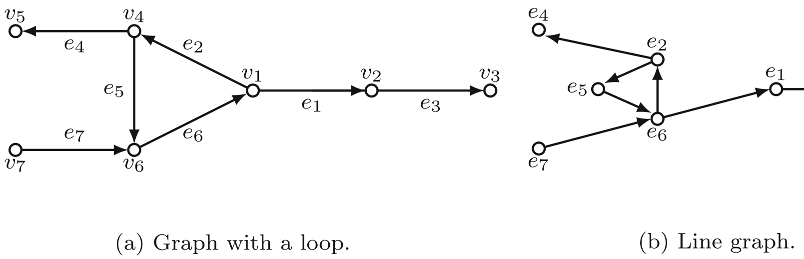


Fig. 13 **a**, **b** Graph and line graph for Example 6.2.2

edges e_5 and e_6 are dead ends and therefore are ranked as the least important transmitters. On the other hand, the latter edges have the highest capability of receiving information and therefore are ranked as the most important receivers. We conclude that the ranking of transmitters and receivers determined by the measures $\exp(E^{\rightarrow})\mathbf{1}$ and $\mathbf{1}^T \exp(E^{\rightarrow})$ is in agreement with intuition based on the graphs in Fig. 12.

6.2.2 Example

This example illustrates the effect of a loop in a directed graph shown in Fig. 13. Table 10 shows the edge e_6 to be both the highest ranked broadcaster and the highest ranked receiver. The directed loop between the nodes v_1, v_4, and v_6 makes edges between these nodes be strong broadcasters.

6.2.3 Flight example I

The last example of this section uses a directed unweighted network determined by domestic flights in the USA during year 2016, as reported by the Bureau of Transportation Statistics of the US Department of Transportation [22]. The airports are nodes and the flight segments are edges. This yields a nonsymmetric adjacency matrix $A \in \mathbb{R}^{705\times705}$. Since most flights have return flights, the matrix A is close to symmetric.

We determine the matrix $E^{\rightarrow}$ using the adjacency matrix for the flights network, and rank the departing flights by computing $\exp(E^{\rightarrow})\mathbf{1}$. The six largest entries of

Table 10 Ranking of the edges of Fig. 13 using the exponential of the matrix $E^{\rightarrow}$

Edge e_k	$[\exp(E^{\rightarrow})\mathbf{1}]_k$	Edge e_k	$[\mathbf{1}^T \exp(E^{\rightarrow})]_k$
e_6	4.78	e_6	3.76
e_2	3.97	e_1	3.23
e_5	3.56	e_2	3.23
e_7	3.56	e_3	2.89
e_1	2	e_4	2.89
e_3	1	e_5	2.89
e_4	1	e_7	1

this vector determine the most important flights. They are displayed in Table 11. We rank the arriving flights by computing the largest entries of $\mathbf{1}^T \exp(E^{\rightarrow})$. However, the computation of $\exp(E^{\rightarrow})\mathbf{1}$ and $\mathbf{1}^T \exp(E^{\rightarrow})$ may result in numerical overflow on many computers. To avoid this difficulty, we compute the spectral radius μ of $E^{\rightarrow}$ and evaluate $\exp(E^{\rightarrow} - \mu I)\mathbf{1}$ and $\mathbf{1}^T \exp(E^{\rightarrow} - \mu I)$ instead. This eliminates overflow and does not affect the relative size of the entries of the computed vectors. Therefore, the ordering is not affected by this modification. This approach of avoiding overflow has previously been applied in computations reported in [14].

We conclude that a flight from the George Bush Intercontinental Airport in Houston to the Hartsfield-Jackson Airport in Atlanta transmits the best through the network of all domestic flights in the USA. The flight from the Cleveland-Hopkins International Airport to Atlanta comes second, followed by the flight from the Tampa International Airport to Atlanta. Similarly, the flight from the Raleigh-Durham International Airport in North Carolina to Atlanta absorbs the most flights through the network of national flights. We remark that although the Los Angeles International Airport and the Chicago O'Hare International Airport are among the three top ranked airports according to the Federal Aviation Administration [13], these airports are not among the ones shown in Table 11. This depends on that our model disregards the number of flights (if larger than one) and the number of passengers on each segment.

7 Computing the most important edges in a directed weighted network using the matrix exponential

Similarly as in the previous section, we can rank the edges of a directed weighted network in their role as transmitters of information by calculating the entries of the vector $\exp(\tilde{E}^{\rightarrow})\mathbf{1}$, where the matrix $\tilde{E}^{\rightarrow} = \tilde{B}^{\mathrm{i}T} \tilde{B}^{\mathrm{e}}$ is determined by taking weights into account as described in Section 4. The relative size of the entries of the vector $\mathbf{1}^T \exp(\tilde{E}^{\rightarrow})$ provides an ordering of the edges in their role as information receivers and the relative size of the elements of $\exp(\tilde{E}^{\rightarrow})\mathbf{1}$ furnishes an ordering of the edges as information transmitters.

Table 11 Ranking of flight segments in the domestic flights network using the exponential of the matrix $E^{\rightarrow}$

Most departing flights		Most landing flights	
From airport	To airport	From airport	To airport
IAH	ATL	ATL	RDU
CLE	ATL	ATL	DSM
TPA	ATL	ATL	BHM
GRR	ATL	ATL	GSP
ALB	ATL	ATL	FSD
PIT	ATL	ATL	BTV

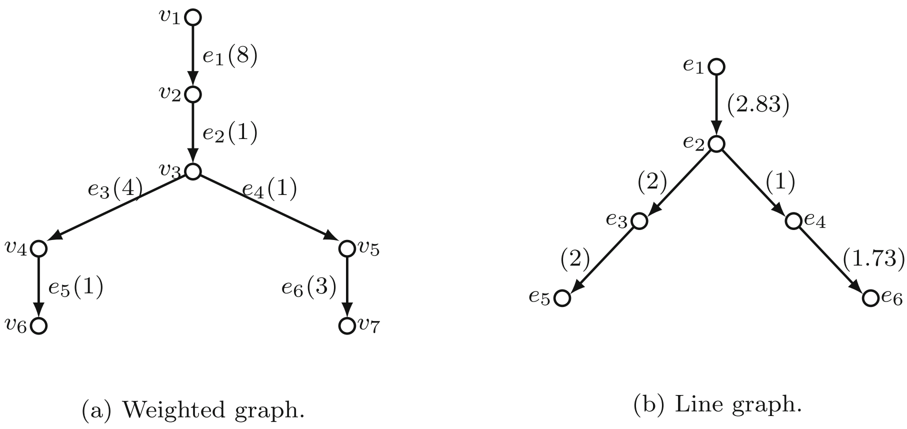


Fig. 14 **a**, **b** A directed weighted tree with line graph

7.1 Example

The graph of this example is displayed in Fig. 14a, with the edge weight shown for each edge. Let the diagonal entries of the diagonal matrix Z contain positive edge weights. Figure 14b shows the associated line graph for $Z^{\mathrm{i}} = Z^{\mathrm{e}} = Z^{1/2}$ (see Section 4 for the definition of the matrices Z^{i} and Z^{e}).

Table 12 ranks the edges according to the importance as transmitters and receivers. Although the edge e_5 has weight 1, and e_6 has weight 3, and these edges are positioned in a similar way, the output of our algorithm suggests that the edge e_5 is more absorbing of information than edge e_6.

7.2 Flight example II

We take the same example studied in Section 6.2.3, but this time we include a weight with each edge. The weight is set equal to the total number of enplanements on all the flights for that segment, as reported by the Bureau of Transportation Statistics [22]. The weights define the diagonal matrix Z and determine an edge-weighted adjacency

Table 12 Ranking of the edges of Fig. 14 using the exponential of $\exp(E^{\rightarrow})$

Edge e_k	$\{\exp(E^{\rightarrow})\mathbf{1}\}_k$	Edge e_k	$\{\mathbf{1}^T\exp(E^{\rightarrow})\}_k$
e_1	10.77	e_5	6.89
e_2	6.87	e_3	5.83
e_3	3.00	e_6	4.41
e_4	2.73	e_2	3.83
e_5	1.00	e_4	3.41
e_6	1.00	e_1	1.00

matrix $\tilde{A}$ as in Theorem 1. Due to the weights, a flight segment with thousands of passengers will affect the flow more than a flight segment with only a few passengers.

To avoid overflow, we evaluate $\exp(E^{\rightarrow} - \mu I)$, where μ is the spectral radius of $E^{\rightarrow}$, similarly as in Section 6.2.3. Table 13 displays the six top ranked edges of the network. All of the segments of the table start and end at one of the top airports as described by the Federal Aviation Administration [13]. In particular, the segment from the Hartsfield-Jackson Airport in Atlanta to the O'Hare Airport in Chicago is the one that dissipates the highest number of passengers through the network of domestic flights in the USA, and the same segment in the opposite direction receives the most passengers through the network. These two airports are among the top three busiest airports in the USA according to the Federal Aviation Administration of the US Department of Transportation [13]. This example attests to the validity of our model.

8 Computational aspects

We comment in this section on the computations required to evaluate

$$\exp(E^{\rightarrow})\mathbf{1} \quad \text{or} \quad \mathbf{1}^T \exp(E^{\rightarrow}). \tag{8.1}$$

Generally, the matrix $E^{\rightarrow}$ is nonsymmetric. For small networks, this matrix is small and can be explicitly formed. The exponential $\exp(E^{\rightarrow})$ then easily can be evaluated, such as by the MATLAB function $\texttt{e}xpm$, and the desired quantities (8.1) can be determined. If the matrix $E^{\rightarrow}$ is large enough so that overflow may occur when evaluating its exponential, the spectral factorization of $E^{\rightarrow}$ may be computed. This yields the spectral radius μ of $E^{\rightarrow}$. Moreover, the spectral factorization can be used to evaluate $\exp(E^{\rightarrow} - \mu I)\mathbf{1}$ and $\mathbf{1}^T \exp(E^{\rightarrow} - \mu I)$. These matrices may only be computable with reduced accuracy when the eigenvector matrix of $E^{\rightarrow}$ is severely ill-conditioned. This has not been an issue in our computations.

When the matrix $E^{\rightarrow}$ is large, it may be attractive to evaluate approximations of the quantities (8.1) with the aid of the nonsymmetric Lanczos process or the Arnoldi process. Their application does not require the matrix $E^{\rightarrow}$ to be formed; only matrix-vector products with $E^{\rightarrow}$, and possibly with its transpose, have to be computed (see

Table 13 Ranking of the segments in the domestic flights network, taking the passengers enplanement as the segments weights, and using the exponential of the line graph adjacency matrix

Most dissipating flights		Most absorbing flights	
From airport	To airport	From airport	To airport
ATL	ORD	ORD	ATL
DTW	ORD	ORD	DTW
OGG	LAX	LAX	LAS
PHL	DEN	DEN	PHL
LAS	LAX	LAX	SEA
SEA	LAX	DEN	LAX

[7, 14]) for details. The eigenvalues of the reduced matrix computed with the nonsymmetric Lanczos or Arnoldi processes yield sufficiently accurate approximations of the spectral radius to avoid overflow in the computation of $\exp(E^{\rightarrow} - \mu I)\mathbf{1}$ and $\mathbf{1}^T \exp(E^{\rightarrow} - \mu I)$.

9 Conclusion

This paper discusses the determination of the most important edges of an undirected or directed graph by using an associated line graph. For directed graphs, several line graphs are described and their usefulness for ranking edges is discussed. We also consider the task of removing unimportant edges. Computed examples illustrate the feasibility of the methods described.

Funding information This research is financially supported in part by NSF grants DMS-1720259 and DMS-1729509.

References

1. Arrigo, F., Benzi, M.: Edge modification criteria for enhancing the communicability of digraphs. SIAM J. Matrix Anal. Appl. **37**, 443–468 (2016)
2. Arrigo, F., Benzi, M.: Updating and downdating techniques for optimizing network communicability. SIAM J. Sci. Comput. **38**, B25–B49 (2016)
3. Benzi, M., Estrada, E., Klymko, C.: Ranking hubs and authorities using matrix functions. Linear Algebra Appl. **438**, 2447–2474 (2013)
4. Benzi, M., Klymko, C.: Total communicability as a centrality measure. J. Complex Networks **1**, 124–149 (2013)
5. Chen, C., Jia, Z., Varaiya, P.: Causes and cures of highway congestion. IEEE Control. Syst. Mag. **21**, 26–32 (2001)
6. Chen, W.-K.: Graph Theory and Its Engineering Applications. World Scientific, Singapore (1997)
7. De la Cruz Cabrera, O., Matar, M., Reichel, L.: Analysis of directed networks via the matrix exponential. J. Comput. Appl Math. **355**, 182–192 (2019)
8. Diestel, R.: Graph Theory. Springer, Berlin (2000)
9. Estrada, E.: Edge adjacency relationships and a novel topological index related to molecular volume. J. Chem. Inf. Comput. Sci. **35**, 31–33 (1995)
10. Estrada, E.: The Structure of Complex Networks: Theory and Applications. Oxford University Press, Oxford (2012)
11. Estrada, E., Hatano, N.: Statistical-mechanical approach to subgraph centrality in complex networks. Chem. Phys. Lett. **439**, 247–251 (2007)
12. Estrada, E., Higham, D.J.: Network properties revealed through matrix functions. SIAM Rev. **52**, 696–714 (2010)
13. Federal Aviation Administration (FAA). Passenger boarding (enplanement) and all-cargo data for US airports, 2016
14. Fenu, C., Martin, D., Reichel, L., Rodriguez, G.: Network analysis via partial spectral factorization and Gauss quadrature. SIAM J. Sci. Comput. **35**, A2046–A2068 (2013)
15. Godsil, C., Royle, G.F.: Algebraic Graph Theory. Springer, New York (2013)
16. Golub, G.H., Van Loan, C.F.: Matrix Computations. John Hopkins University Press, Baltimore (2013)
17. Gross, J.L., Yellen, J.: Graph Theory and Its Applications, 2nd edn. Taylor & Francis, Boca Raton (2006)
18. Gutman, I., Estrada, E.: Topological indices based on the line graph of the molecular graph. J. Chem. Inf. Comput. Sci. **36**, 541–543 (1996)

19. Katz, L.: A new status index derived from sociometric analysis. Psychometrika **18**, 39–43 (1953)
20. Kleinberg, J.M.: Authoritative sources in a hyperlinked environment. J. ACM **46**, 604–632 (1999)
21. Newman, M.E.J.: Networks: An Introduction. Oxford University Press, Oxford (2010)
22. Bureau of Transportation Statistics. Research and Innovative Technology Administration/Transtats
23. Orlin, J.: Contentment in graph theory: covering graphs with cliques. Indagationes Mathematicae (Proceedings) **80**, 406–424 (1977)
24. Thulasiraman, K., Swamy, M.N.S.: Graphs: Theory and Algorithms. Wiley, New York (1992)