EPJ manuscript No.
(will be inserted by the editor)

Structure-preserving model reduction of large-scale logistics networks

Applications for Supply Chains

B. Scholz-Reiter[1], F. Wirth[2], S. Dashkovskiy[3], T. Makuschewitz[1], M. Schönlein[2], and M. Kosmykov[4]

[1] BIBA - Bremer Institut für Produktion und Logistik GmbH, University of Bremen, Hochschulring 20, 28359 Bremen, Germany
[2] Institute for Mathematics, University of Würzburg, Emil-Fischer-Str. 40, 97074 Würzburg, Germany
[3] Department of Civil Engineering, University of Applied Sciences Erfurt, Altonaer Str. 25, 99085 Erfurt, Germany
[4] Center of Industrial Mathematics, University of Bremen, 28334 Bremen, Germany

Received: November 17, 2010 / Revised version: date

Abstract. We investigate the problem of model reduction with a view to large-scale logistics networks, specifically supply chains. Such networks are modeled by means of graphs, which describe the structure of material flow. An aim of the proposed model reduction procedure is to preserve important features within the network. As a new methodology we introduce the LogRank as a measure for the importance of locations, which is based on the structure of the flows within the network. We argue that these properties reflect relative importance of locations. Based on the LogRank we identify subgraphs of the network that can be neglected or aggregated. The effect of this is discussed for a few motifs. Using this approach we present a meta algorithm for structure-preserving model reduction that can be adapted to different mathematical modeling frameworks. The capabilities of the approach are demonstrated with a test case, where a logistics network is modeled as a Jackson network, i.e., a particular type of queueing network.

PACS. 02.01.Ox Graph theory – 89.75.Fb Structures and organization in complex systems – 89.75.Kd Patterns

1 Introduction

In this paper we present an approach to model reduction of dynamic and large-scale logistics networks that aims to preserve major structural characteristics of the original network. Many dynamical systems, whether occurring as models of natural phenomena or designed by man, can be interpreted as an interaction of different components each of which is governed by dynamics admitting inputs from or producing output for other components. The possible interaction is described by a topological structure, which can be interpreted in a natural way as a graph or network. In applications the underlying network is frequently large, especially so in the areas of logistics. In this paper we concentrate on the analysis of supply chains, but we believe that similar tools will be useful in other application areas of logistics. In the following we will freely use the term logistic network, which will, in the context of this paper, refer to supply chain networks. The structure of a logistics network is given by the connections between locations of the network. Within large-scale logistics networks raw material suppliers, manufacturers and the original equipment manufacturer contribute with their factories and facilities to the large number of locations. In addition warehouses and distribution centers of retailers are also called locations in this paper.

The analysis of large-scale networks has been an active research area with respect to different aspects in logistics networks. In the analysis of airline networks for instance, the analysis of hubs has been undertaken using graph theoretic methods. Importance of nodes, i.e. airports, in the graph has been measured by the number of passengers traveling through an airport, see e.g. [1–3]. We note that there is a significant difference between airline networks and the logistics networks under consideration in this paper: in airline networks a significant number of passengers can be expected to travel a certain route both ways. This makes the underlying graph effectively undirected. On the other hand in supply chain networks it is usually possible to determine sources, i.e. first-tier suppliers, and sinks, which represent retailers.

Typical processes taking place in logistics networks are production, storage and shipment of different commodities. Material, information and monetary flows connect the locations of a logistics network and create the structure of the network. The structure of a flow is frequently characterized as *linear, convergent, divergent, or non-linear*. Here linear denotes a simple chain of locations passed one after the other, convergent describes flows orig-

inating from a large number of locations and ending in a few end locations, divergent describes the opposite structure in which a few sources feed a larger number of end locations, while nonlinear in this context simply denotes a more intricate structure which does not fit into the other categories.

In this context it should be noted that the magnitude of a flow between location i and location j could differ from the flow between location i and location k. This observation motivates the assumption that different locations can be of different relevance for the network. However, flows within logistics networks are usually the result of dynamic processes and thus are time-varying. They may change their magnitude over time, for instance, because a factory may dynamically adjust its production rate and output in order to satisfy the given demand, which is changing over time. Hence, every location of the network can be considered as a dynamical system, which changes the magnitude of linked flows. Moreover the dynamics of one location depends on the states at other locations due to interactions. The resulting behavior is typically nonlinear. The number of locations appears to be large in modern logistics networks and their interconnection structure can be rather complicated [4]. This leads to complex dynamics of the whole network.

A mathematical model that enables to predict this behavior can help practitioners to handle and to design such networks or its parts so that an efficient and robust performance will be guaranteed. In addition such a model allows for further analysis of the network, e.g., for stability analysis and to predict the dynamics for a given set of parameters and initial data.

The network approach to the analysis of supply chains has attracted considerable attention in recent years, see e.g. [4–7], in particular focusing on tools from complex systems analysis. In this context it has been noted that it is unreasonable to treat a network as a static entity, but rather small changes in the structure occur on a regular basis. There are different approaches in the literature to model such networks, [8]. Common approaches are queueing networks [9,10], reentrant lines [11,12], approaches based on conservation laws [13,14] and many others [15–17]. Taking the interconnection structure of the network into account this leads to coupled systems of equations describing the behavior of the whole network. Due to the size of such networks the coupled systems contain a large number of equations which are usually nonlinear. This makes it difficult to perform analysis and simulations. One possible way out is to find a reduced size model that approximates the original network and its behavior, i.e., that describes the most important part of its structure and has dynamical properties close to the ones of the original network. This reasoning motivates the research presented in this paper.

Furthermore, even though the network structure in supply chains may be expected to undergo changes with a moderate frequency, we argue that the important features of the network can be expected to be more stable. So the identification and analysis of the important part of the network in combination with tools for a robust design of networks, [18,19] may provide tools for network design despite uncertainty about the future evolution of the network. Also, reduced order models are of interest because they facilitate simulations and also may be more amenable to analytic treatment. We expect that the proposed reduction approach allows for preservation of the essential dynamical properties of the network such as stability, robustness, performance that are relevant in logistics sense. Analysis of these properties should be facilitated by the proposed model reduction. This issue needs further investigation and will be the topic of future research.

A classical approach in the model reduction of linear input-output systems is to describe the approximation error of lower order estimates of stable linear systems in terms of the induced operator norm from the space of square integrable inputs to square integrable outputs. This approach using balanced truncation is well understood and explicit error estimates are available, see for example [20–22]. More recently, Krylov subspace methods have become attractive since they provide methods at reasonable computational cost for very large systems [21, 23]. In case of nonlinear systems the theory is not that well developed. There are specialized methods developed for particular cases or classes of systems, see e.g. [24–26]. In this paper we do not specialize to the classical case of linear systems as for logistics systems in particular other model classes are more appropriate. For this reason we start an investigation of model reduction based on ranking techniques in graphs, which has the benefit of allowing to keep structural features of a given network. These are usually destroyed in approaches using balanced truncation or Krylov subspace methods. We note, however, that recently this problem has also been investigated for linear systems in [27] – although without the use of ranking ideas.

The method we propose in this paper is based on the consideration of flows in a network. The aim of the paper is to put forward a new methodology that might be useful in various circumstances, but which also would have to be adapted to the specific case. This method is in particular designed for the approximation of logistics networks. We represent such networks as graphs, where the logistics locations are vertices and the edges represent the flows between them. Our idea is to find those locations through which the main flows in the network are routed. Then, to reduce the model, we keep this most important part of the network and exclude the less important locations or aggregate them with other vertices. For this purpose we quantify the importance of locations in the following way. We say that the importance of a location is a positive number proportional to the stationary probability that an arbitrarily taken single part of the flow in the network will be found at this location. This stationary probability has to be defined with respect to a suitable Markov chain. This number is called *LogRank* of the location. A similar idea is used by the PageRank algorithm to quantify the importance of web pages in the internet [28]. There are other ranking schemes with similar characteristics, e.g.,

HITS [29] and SALSA [30], see also [31] and the references therein. In contrast to these algorithms we take not only the structure of the network into account but also the magnitude of the flows between the locations.

In the reduction process we aim to preserve the main structure of the network in the sense of preserving the main flows. For this purpose we look for special subgraphs, called motifs [32–34] in the network and apply aggregation of vertices of lower ranks with those of higher rank in these motifs. In general, such an approximation might lead to a change in the importance order of the remaining vertices. We provide a constructive procedure which preserves the ranking order of the remaining vertices. In this paper we introduce a heuristic meta algorithm for the exclusion or aggregation of locations in a logistics network such that the relative order of importance is preserved.

The paper is organized as follows. In the next section we introduce notation and give some background from matrix and graph theory. Section 3 introduces the notion of importance of a logistics location in a network and explains how it can be calculated. Section 4 describes graph transformations that lead to a reduced model and preserve the order of importance. In Section 5 we collect all the steps starting from identification of the locations importance to an algorithm that leads to a reduced model of a logistics network. Up to this point the methods discussed are generic in the sense that no particular modeling setup is considered. To illustrate the approach in a more concrete setting we consider logistics systems modeled by Jackson networks in Section 6. In particular, in this case it is shown that the approach allows to retain characteristics of interest, viz. the stationary distribution, for nodes that are not affected by the order reduction. Discussion and concluding remarks are given in Section 7.

2 Preliminaries on nonnegative matrices and directed graphs

This section provides notions and results about nonnegative matrices and graphs that will be used throughout the paper.

A vector $v \in \mathbb{R}^n$ is called *nonnegative*, if every component is nonnegative, i.e., $v_i \geq 0, i = 1, \ldots, n$ and a matrix $M \in \mathbb{R}^{n\times n}$ is said to be *nonnegative* $M \geq 0$ (resp. *positive* $M > 0$) if its components m_{ij} are nonnegative (resp. positive), i.e., $m_{ij} \geq (>) 0$. The nonnegative orthant of the n-dimensional space $\mathbb{R}^n$ is denoted by $\mathbb{R}^n_+$. The transpose of a vector v respectively a matrix M is denoted by v^T, resp. M^T. In this paper we find it convenient to use the 1-norm of vectors, which for $v \in \mathbb{R}^n$ is given by $\|v\| = \sum_{i=1}^n |v_i|$. The n-dimensional vector with entries all equal to 1 is denoted by $\mathbf{e}_n$. A nonnegative vector v is called a *stochastic*, if its components sum to one, i.e., $\|v\| = 1$. An $n \times n$ matrix $M \geq 0$ is called *column normalized*, if for $j = 1, \ldots, n$

$$\sum_{i=1}^{n} m_{ij} = \begin{cases} 1 & \exists\, i \,:\, m_{ij} > 0, \\ 0 & m_{ij} = 0 \;\forall\, i = 1, \ldots, n. \end{cases}$$

A matrix $M \geq 0$ is called *column stochastic*, or briefly stochastic, if $\mathbf{e}_n^\mathsf{T} M = \mathbf{e}_n^\mathsf{T}$, i.e., columns are stochastic. It is substochastic, if the column sums are bounded by 1. If M^T is colum (sub)stochastic, we call M row (sub)stochastic. The spectrum of a matrix M is denoted by $\sigma(M)$ and the *spectral radius* $\rho(M)$ is

$$\rho(M) = \max\{|\lambda| : \lambda \in \sigma(M)\}.$$

If a matrix M is stochastic, then $\rho(M) = 1$, [35].

A matrix $M \in \mathbb{R}^{n\times n}$ is called *reducible*, if there exists a permutation matrix $P \in \mathbb{R}^{n\times n}$ such that

$$P^\mathsf{T} M P = \begin{bmatrix} A & 0 \\ B & C \end{bmatrix},$$

where the matrices A and C are square. Otherwise, M is said to be *irreducible*. The Perron-Frobenius Theorem, cf. [35, Theorem 2.1.4] provides fundamental insight in the spectrum of positive resp. nonnegative, irreducible matrices.

Theorem 1 *(a) If $M \in \mathbb{R}^{n\times n}$ is positive, then $\rho(M)$ is a simple eigenvalue of M, greater in modulus than any other eigenvalue in $\sigma(M)$.*

(b) If $M \geq 0$ is irreducible, then $\rho(M)$ is a simple eigenvalue, any eigenvalue of M of the same modulus is also simple, M has a positive eigenvector x corresponding to $\rho(M)$, and any nonnegative eigenvector of M is a scalar multiple of x.

The eigenvector to the eigenvalue $\rho(M)$ is called the Perron vector. A matrix M is called *primitive* if there exists a positive integer $k \in \mathbb{N}$ such that $M^k > 0$. Note that any primitive matrix is irreducible. The converse is false, in general. For the following connection between primitive and irreducible matrices we refer to [35, Theorem 2.1.7].

Lemma 1 *For $M \geq 0$ the following are equivalent.*

(a) M is irreducible and $\rho(M)$ is greater in magnitude than any other eigenvalue.

(b) M is primitive.

Another useful connection between irreducible and primitive matrices is the following, cf. [35, Corollary 2.2.28].

Lemma 2 *An irreducible matrix is primitive, if its trace is positive.*

In the following we use the notation from [36]. A directed graph with weights consists of a finite vertex set V and an edge set E, where a directed edge from vertex i to vertex j is an ordered pair $(i, j) \in E \subset V \times V$. The weights can be represented by a $|V| \times |V|$ weighted adjacency matrix A, where $a_{ij} \geq 0$ denotes the weight of the directed edge from vertex i to vertex j. By convention $a_{ij} > 0$, if and only if $(i, j) \in E$. We will denote a directed graph with weights of this form by $G = (V, E, A)$. Further for each vertex $i \in V$ the *indegree* is

$$\mathrm{in}(i) = |\{j \in V \,:\, (j, i) \in E\}|,$$

where $|\cdot|$ denotes the *cardinality* of a set and

$$\mathrm{out}(i) = |\{j \in V \,:\, (i,j) \in E\}|$$

denotes the *outdegree* of the vertex $i \in V$. Additionally, we define for each vertex i the set of *successors* by

$$S(i) = \{j \,:\, (i,j) \in E\}$$

and the set of *predecessors* by

$$P(i) = \{j \,:\, (j,i) \in E\}.$$

A *path* from vertex i to j is a sequence of distinct vertices starting with i and ending with j such that there is an edge in the direction of the path between consecutive vertices. A directed graph is said to be *strongly connected* if for any ordered pair (i,j) of vertices, there is a path which leads from i to j. In terms of the weighted adjacency matrix this is equivalent to A being irreducible, [35]. We call the graph *connected*, if the underlying undirected graph obtained by forgetting the direction of edges is connected.

3 Identification of important locations in large-scale logistics networks

In this section we introduce a scheme to identify the importance of locations of a logistics network. The approach is based on the structure of the network and on a representative flow between the locations, where we consider the material flow. The consideration of the material flow has several advantages in the analysis of logistics networks. It can be easily measured or quantified compared to the information flow. In comparison to the monetary flow within the network it is less confidential and hence more easily accessible. Also to some extent, the material flow allows to deduce information about the monetary and information flow within the network. Note that a real world material flow within a logistics network needs to be normalized to a standard unit. We propose to do this by measuring all segments of the material flow by units of final product. For instance two headlights of a car would be measured as one part, since two headlights are necessary for the final assembly of one car. In the following we mean this normalized material flow whenever we speak of material flow.

As an initial step we number the logistics locations from 1 to n. To model the structure of the network we form a directed graph $G = (V, E)$ where the vertices $V = \{1, \ldots, n\}$ represent the locations. The set of edges is given by

$$E = \{(i,j) \in V \times V : \text{there is material flow from } i \text{ to } j\}.$$

Throughout we are assuming that the graph we are considering is connected, as otherwise there are two independent logistics networks, which should be treated independently. Note that we are not assuming that the network is strongly connected. We do not rule out the possibility of cycles within the graph so that in particular reentrant systems can be treated within our approach.

Associated to the edge set E we define a matrix, denoted by M, that reflects the relative material flow between the locations. To be precise, let f_{ij} denote the observed magnitude of the material flow from location i to j over a pre-defined period. This allows to capture the dynamics of the network. If the pre-defined period is sufficiently long and an ergodicity property holds, an approximation of the stationary behavior of the network is obtained by averaging. As we will place our approach in the framework of Markov chains a stationarity assumption underlies the whole approach.

Then the matrix M is given by

$$m_{ij} := \begin{cases} \dfrac{f_{ij}}{\sum_{i=1}^{n} f_{ij}} & (i,j) \in E, \\ 0 & \text{else.} \end{cases}$$

Note that the normalization is such that the inflow to each location sums to one. Thus the quantity m_{ij} denotes the proportion of the material, which is received by location j from location i, relative to all the material received by location j.

Our interpretation of importance for a location in a logistics network is the following. Consider M as the transition matrix of a homogeneous Markov chain with state space $V = \{1, \ldots, n\}$. This chain evolves according to $x(k+1) = Mx(k)$. Then m_{ij} is the probability to move from vertex j to vertex i. So m_{ij} may be interpreted as the probability that location j places an order with location i. However, in a logistics network driven by a pull policy the flow of orders is stepwise [37]. In particular retailers place their orders at the OEM locations. Based on these orders the OEM location places orders at their suppliers and so on. This process terminates as soon as a source supplier is reached. So the locations take the decision where to order required material. If the chain described above has a stationary distribution then it represents the proportion of orders placed at the locations.

Consequently, the importance of a location for the logistics network is indicated by the stationary distribution of the Markov chain, if it exists. A stationary distribution of a homogenous Markov chain is given by a stochastic eigenvector $r \geq 0$ of the transition matrix M to the eigenvalue one, i.e.,

$$M\, r = r.$$

From the Perron-Frobenius Theorem 1 it follows that, if the transition matrix is irreducible the Markov chain has a unique stationary distribution $r > 0$, [35]. Hence in the irreducible case there is only one solution describing the importance of the locations.

Figure 1 illustrates a small sample network of a logistics network. The matrix M, that contains the relative

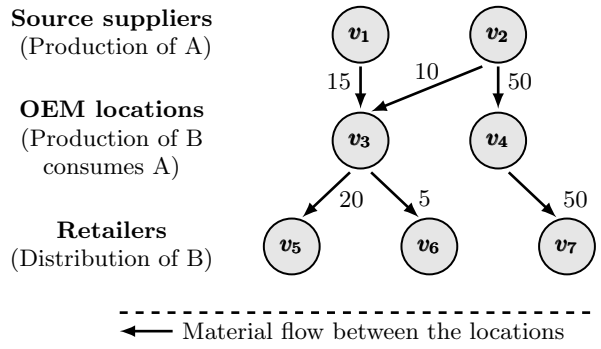

Fig. 1. Logistics network weighted with material flow.

material flows, is for the example scenario given by

$$M = \begin{bmatrix} 0 & 0 & \frac{15}{25} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{10}{25} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Remark 1 From the graph illustrated in Figure 1 it can be concluded that matrix M of the example is not irreducible. There is no unique stationary distribution and indeed as the matrix is not stochastic, it is not even the transition matrix of a Markov chain.

In order to follow the outlined reasoning and to obtain a reasonable result that reflects the importance of the various locations in the logistics network the corresponding graph has to be modified. In the remainder of this section we present modifications of M that ensure irreducibility. These adjustments are based on inherent characteristics of logistics networks.

3.1 Connecting retailers and source suppliers

In this subsection we present necessary modifications of the matrix M for the case that the structure of the captured logistics network does not ensure an irreducible and stochastic matrix. This requirement is mainly driven by the characteristics of source suppliers. Locations modeling source suppliers have indegree zero, i.e., they do not place any orders at other locations and do not receive any material. This leads to zero columns in the corresponding transition matrix, which implies reducibility. In this context ordinary retailers have an outdegree of zero. Hence they represent sinks and do not deliver any material to other locations within the network. We propose to model real world retailers that have outgoing links by two separate locations, one of which has outdegree zero.

In order to obtain a stochastic matrix from M we propose to connect the source suppliers with the retailers. The interpretation of these links is not related to material flow or orders placed from retailers at the source suppliers, as it is doubtful whether such orders will occur at all. Rather these additional links represent an implicit influence of the retailers on the source suppliers. Without the pull exercised by the retailers on the network, material would not flow from the source suppliers. An interpretation of the approach is the implicit assumption that the material flow within the network is also subject to the exchange of information about placed orders between the locations. In particular it can be argued that the source suppliers need information about the orders of the retailers in order to prepare for a timely delivery of material to directly linked locations. This information is motivated by the need of adjusting the processes of raw material sourcing and the requirements to eliminate the Bullwhip-Effect [17].

The question is then with which weights these new additional links should be provided since there is no physical flow attached to them. To describe this consider the set of retailers, denoted $\mathcal{R} \subset V$, and the set of source suppliers, denoted $\mathcal{S} \subset V$. They can be characterized by

$$\mathcal{S} := \{i \in V \,:\, \mathrm{in}(i) = 0\} \text{ and } \mathcal{R} := \{i \in V \,:\, \mathrm{out}(i) = 0\}.$$

Let $i \in \mathcal{R}$ be a retailer and $j \in \mathcal{S}$ be a source supplier, we say that i affects j if and only if the graph $G = (V, E)$ contains a path from source supplier j to retailer i. Given a source supplier $j \in \mathcal{S}$ the set of all retailers affecting j is denoted by $\mathcal{R}_j$. To integrate the new edges to the graph that describes the network we have to amend the edge set. That is, the additional edges are given by

$$E' := \{(i, j) \in V \times V \,:\, j \in \mathcal{S},\, i \in \mathcal{R}_j\}.$$

Note that the sets E and E' do not intersect. This is because E' contains only edges $(i, j) \in V \times V$ which start at vertices $i \in \mathcal{R}_j \subset V$ that have outdegree zero in the graph $G = (V, E)$. Therefore we consider in the following the graph $\mathcal{G} = (V, \mathcal{E})$ that includes the additional structure, i.e., the edge set is given by

$$\mathcal{E} := E \cup E'. \tag{1}$$

We will discuss two possibilities for defining weights for edges in E'. The first defines a matrix $\mathcal{M}_1$ by adding new weights as on the basis of path probabilities. Using the flows $f_{ij}, j = 1, \ldots, n$ we define the proportion of material going from i to j with respect to all the material leaving i by

$$\sigma_{ij} := \frac{f_{ij}}{\sum_{j=1}^{n} f_{ij}}.$$

The weight of the edge from a retailer $i \in \mathcal{R}$ to a source supplier $j \in \mathcal{S}$ is obtained by the addition of the product probabilities of the σ_{ij} over all paths $\tau = (j = i_0, i_1, \ldots, i_k = i)$ from j to i, i.e.

$$m^1_{ij} := \begin{cases} \sum_{\substack{\text{paths from} \\ j \text{ to } i}} \prod_{l=1}^{k} \sigma_{l-1,l} & , (i, j) \in E'. \\ m_{ij} & , (i, j) \notin E'. \end{cases} \tag{2}$$

This may be interpreted as the probability that a source supplier j is the point of origin for the processing of an order of retailer i. This approach assumes that there are no

preferential paths between the particular locations. Equivalently, we wish to consider the probability that an order initiated by a retailer j will arrive at the source supplier i.

By construction the matrix $\mathcal{M}_1$ is column stochastic. It is an easy exercise to see that in the case that Kirchhoff's flow laws apply to the data the matrix $\mathcal{M}_1$ defines a stationary probability, which contains simply the relative flow through the nodes as entries. In this sense, the stationary probability can be read of from the data without any further calculation, see the result for the following example in (6) below. However, in the case of real data it is unreasonable to expect that Kirchhoff's laws hold, for instance because of the delay in receipt of material and the subsequent delivery of products. Thus in a given measurement interval some nodes will effectively stock material, which distorts the overall picture.

The second method we propose for deriving weights can be described measuring the total influence of the retailers through the network. Given a logistics network that is managed with a pull policy, it cannot be guaranteed in general that a certain order of a retailer will originate its' processing at a certain source supplier. Thus, the information about the actual path of a processed order through the network is not available as we only have the information contained in the f_{ij} at our disposition. Note that many different actual paths may lead to the same f_{ij}.

In the transition matrix M we connect the source suppliers with the retailers, if a path between them exists. In this context we assume that a source supplier anticipates an allocation of his production according to the relative demand of the connected retailers. In other words, he will be responsible for the fulfillment of these orders in some sense.

To incorporate this interpretation in the graph $G = (V, E)$ we now specify the relationship between source suppliers and retailers. Given a retailer $i \in \mathcal{R}$ the set of locations that are directly connected to i is denoted by

$$P(i) = \{k \in V \,:\, (k,i) \in E\,, i \in \mathcal{R}\}.$$

So the total inflow of material to retailer i is given by the sum of the material flow of all predecessor of i, i.e.,

$$p_i := \sum_{k \in P(i)} f_{ki}, \quad i \in \mathcal{R}_j.$$

We set $p_i := 0$ if $i \notin \mathcal{R}_j$. Further, for source supplier j the sum of the material flow over all retailers affecting j is given by

$$q_j := \sum_{i \in \mathcal{R}_j} p_i, \quad j \in \mathcal{S}.$$

This quantity is strictly greater than 0, otherwise a source supplier j would be disconnected from the network. A case excluded from our considerations. We use q_j as a normalization factor and define the impact of the retailer i to the source supplier j by

$$\frac{p_i}{q_j}, \quad j \in \mathcal{S},\, i \in \mathcal{R}_j\,.$$

The corresponding weighted adjacency matrix $\mathcal{M}_2$ is defined by the entries

$$m^2_{ij} := \begin{cases} m_{ij} & (i,j) \in E, \\ \dfrac{p_i}{q_j} & (i,j) \in E', \\ 0 & \text{else.} \end{cases} \tag{3}$$

This raises the question which properties the modified adjacency matrix $\mathcal{M}_2$ has.

Proposition 1 *The matrix $\mathcal{M}_2$ is column stochastic.*

Proof. The matrix $\mathcal{M}_2$ is column stochastic if and only if $\mathbf{e}_n^\mathsf{T}\mathcal{M}_2 = \mathbf{e}_n^\mathsf{T}$, i.e., for all $j \in V$

$$\sum_{i=1}^{n} m^2_{ij} = 1\,. \tag{4}$$

Since the matrix M is column normalized it suffices to show that (4) holds for every $j \in \mathcal{S}$, where (4) reads as

$$\sum_{i=1}^{n} m^2_{ij} = \sum_{i=1}^{n} \frac{p_i}{q_j} = \frac{1}{q_j} \sum_{i \in \mathcal{R}_j} p_i = \frac{\sum_{i \in \mathcal{R}_j} \sum_{k \in P(i)} f_{ki}}{\sum_{i \in \mathcal{R}_j} \sum_{k \in P(i)} f_{ki}} = 1.$$

This shows the assertion. □

Further, we investigate the irreducibility of the matrices $\mathcal{M}_1, \mathcal{M}_2$ in terms of the retailers and the source suppliers. To this end we consider the graph, denoted $\mathcal{G}' = (V', \mathcal{E}')$, that consists solely of retailers and source suppliers. That is, the vertex set $V' \subset V$ is given by

$$V' := \mathcal{R} \cup \mathcal{S}.$$

To define the edge set $\mathcal{E}'$ we take on one hand the edge set E', i.e., the edges from the retailers to the source suppliers. On the other hand we incorporate the opposite direction of the material flow. That is, we consider the set $E'' = \{(i,j) \,:\, i \in \mathcal{S}, j \in \mathcal{R},$ there is a path from i to $j\,\}$. The edge set of $\mathcal{G}'$ is then defined by

$$\mathcal{E}' := E' \cup E''.$$

The reversion of the edges in E' corresponds to the fact that the graph $\mathcal{G}'$ is bipartite, i.e., the vertex set V' can be partitioned into the two classes $\mathcal{R}$ and $\mathcal{S}$ such that every edge origins in one class and ends in the other class. A further property that may be assumed for logistics networks is that eventually every material flow ends in some retailer of the network as well as ultimately every location except source suppliers receive material from source suppliers. In terms of the graph $\mathcal{G}$ describing the network this characteristic can be expressed as

$$\begin{gathered} \text{for all } \; i \in V \setminus \mathcal{R} \quad \text{there is a } \; j \in \mathcal{R} \\ \text{such that there is a path from } i \text{ to } j, \text{ and} \\ \text{for all } \; i \in V \setminus \mathcal{S} \quad \text{there is a } \; j \in \mathcal{S} \\ \text{such that there is a path from } i \text{ to } j. \end{gathered} \tag{5}$$

Proposition 2 *Consider the weighted adjacency matrices $\mathcal{M}_i$ $i = 1, 2$ of the graph $\mathcal{G}$ satisfying* (5)*. The following statements are equivalent*

(i) $\mathcal{M}_i$ is irreducible for $i = 1, 2$,
(ii) the graph $\mathcal{G}'$ is strongly connected,
(iii) the graph $\mathcal{G}'$ is connected,
(iv) the graph G is connected.

Proof. As the zero entries of $\mathcal{M}_1$ and $\mathcal{M}_2$ coincide, it is sufficient to prove the result for $\mathcal{M}_1$.

(i)⇒(ii): Assume that $\mathcal{M}_1$ is irreducible. Then the graph $\mathcal{G}$ is strongly connected. This implies that for every pair of distinct vertices (i, j) there is a path from i to j. In particular, for every source supplier s and every retailer r there is a path from s to r. Hence, the fact that $\mathcal{E}'$ contains for every edge (s, r) also its reverse implies that $\mathcal{G}'$ is strongly connected.

The equivalence of (ii) and (iii) is clear, as $\mathcal{G}'$ always has edges in both directions. The implication (iii)⇒(iv) is obvious as $\mathcal{G}'$ cannot be connected if G is not.

To show (iv)⇒(i) we will show equivalently, that (iv) implies that $\mathcal{G}$ is strongly connected. Note that every path from a source supplier to a retailer in G defines a cycle in $\mathcal{G}$ by the additional link back. By assumption (5) this means that whenever there is a path from i to j in $\mathcal{G}$, then by continuing the cycle, there is also a path from j to i in $\mathcal{G}$. Now connectedness of G implies that there for every $i, j \in V$ there is a path from i to j ignoring the directions prescribed by E. By using the cycles available in $\mathcal{G}$ this path leads to a path from i to j in $\mathcal{G}$, so that $\mathcal{G}'$ is strongly connected. □

The modification of the logistics network given in Figure 1 is shown in Figure 2, where the additional edges E' are depicted by dashed lines.

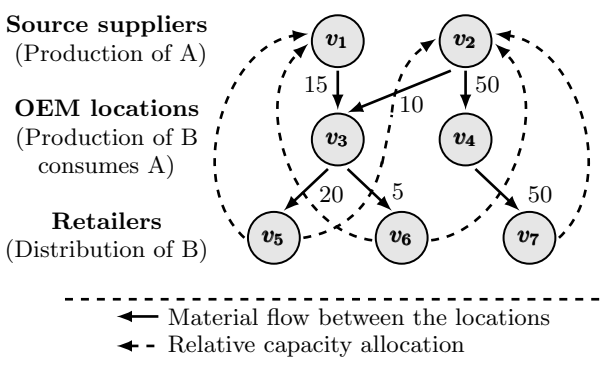

Fig. 2. This figure illustrates the modification proposed in this section for the logistics network example.

The bipartite graph $\mathcal{G}'$ of retailers and source suppliers is shown in Figure 3.

The corresponding modified weighted adjacency matrices are of the form $\mathcal{M}_1, \mathcal{M}_2$ below. Here $\mathcal{M}_1$ corresponds to the proportional division of the flow throughout the network. E.g. the entry $(2, 6)$ is $1/6 \cdot 1/5 = 1/30$, as $1/6$ of the material leaving v_2 goes to v_3 and $1/5$ of the material

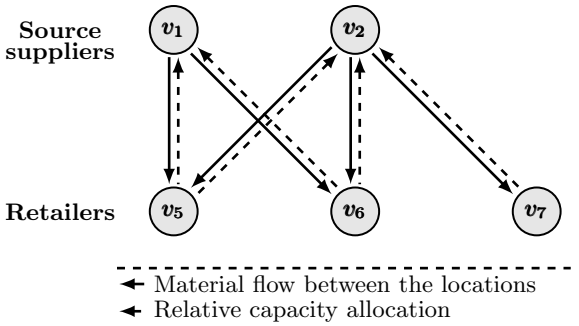

Fig. 3. This figure illustrates the graph $\mathcal{G}'$ for the logistics network example.

leaving v_3 goes to v_6. On the other hand $\mathcal{M}_2$ corresponds to the weighing comparing the outflow. E.g. v_2 delivers goods that can potentially end up in the nodes v_5, v_6, v_7, with a total outflow of 75 and the inflow of v_6 from within that group is 5 so that the entry at position $(2, 6)$ of $\mathcal{M}_2$ is $5/75$.

$$\mathcal{M}_1 = \begin{bmatrix} 0 & 0 & \frac{15}{25} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{10}{25} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{20}{25} & \frac{4}{30} & 0 & 0 & 0 & 0 & 0 \\ \frac{5}{25} & \frac{1}{30} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{5}{6} & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \mathcal{M}_2 = \begin{bmatrix} 0 & 0 & \frac{15}{25} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{10}{25} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{20}{25} & \frac{20}{75} & 0 & 0 & 0 & 0 & 0 \\ \frac{5}{25} & \frac{5}{75} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{50}{75} & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

With Proposition 2 it follows that the matrices $\mathcal{M}_1, \mathcal{M}_2$ are irreducible. The unique stationary distributions for $\mathcal{M}_1$, resp. $\mathcal{M}_2$ are given by

$$r_1 \approx \begin{bmatrix} .0667 & .2667 & .1111 & .2222 & .0889 & .0222 & .2222 \end{bmatrix}^{\mathsf{T}}, \quad (6)$$

$$r_2 \approx \begin{bmatrix} .1111 & .2222 & .1852 & .1481 & .1481 & .0370 & .1481 \end{bmatrix}^{\mathsf{T}}. \quad (7)$$

As predicted the entries of r_1 just give the proportion of the flow going through the respective node compared to the total flow of 225 of the network. The second method places more rank at nodes v_1, v_3, v_5, v_6. It is potentially a matter of some debate which of the two methods is more meaningful. But since the second method is conceptually easier, Kirchhoff's laws cannot be expected to hold in a given measurement of the flows of a network and the information obtained goes beyond the mere calculation of flows through a node, we will concentrate on the second approach in the following. In particular, in the sequel $\mathcal{M} := \mathcal{M}_2$. However, the results of Section 4 apply equally well to $\mathcal{M}_1$ and $\mathcal{M}_2$.

Note that under our standing assumption of connectedness of the logistics network, we have seen in Proposition 2, that for the extended graph $\mathcal{G}$ a unique ranking vector exists. In the next subsection we will perform a further step in the modification of the original matrix by taking account of neglected connections to the outside world of the network.

3.2 Embedding into a larger network

This subsection provides a further modification of the adjacency matrix $\mathcal{M}$, that leads to a primitive matrix. The approach is derived from the recent development in logistics networks, where the distribution of e.g. production facilities and OEM locations in logistics networks has become global. Furthermore, the ability to react swiftly to local economics circumstances (e.g. strikes) forces the logistics network to procure material locally from locations that are not part of the network. The concept to include these exterior effects is to embed the given logistics network into a larger network.

Consider a logistics network $\mathcal{G} = (V, \mathcal{E})$ with n locations, where the relative material flows and the retailers orders are described by the matrix $\mathcal{M} := \mathcal{M}_2$. Let the size of the larger network be $n+m \gg n$, i.e., it consists of $n+m$ vertices. We number the vertices of the overall network in such a way that the first n vertices correspond to the logistics network of interest. As the structure of the large network outside the given logistics network is unknown we assume that every vertex is connected to every other vertex with equal weight. The connection of the original network to the large network is regulated by a parameter $c \in [0,1]$. Here $c = 1$ corresponds to decoupling, and $c = 0$ corresponds to the fact that the original network structure has vanished. First we consider the case where the structure outside the given logistics network is unknown and the interaction of the network of interest with the large network is assumed to be weak. This corresponds to c being close to one.

The interconnection structure of the large network and the embedding is given by the $(n+m) \times (n+m)$ matrix

$$L = \begin{bmatrix} c\mathcal{M} + \frac{1-c}{n+m}\mathbf{e}_n\mathbf{e}_n^\mathsf{T} & \frac{1}{n+m}\mathbf{e}_n\mathbf{e}_m^\mathsf{T} \\ \frac{1-c}{n+m}\mathbf{e}_m\mathbf{e}_n^\mathsf{T} & \frac{1}{n+m}\,\mathbf{e}_m\mathbf{e}_m^\mathsf{T} \end{bmatrix}. \tag{8}$$

The matrix L is column stochastic and primitive. Thus by the Perron-Frobenius Theorem 1 the matrix L has a unique positive normalized right eigenvector corresponding to the eigenvalue 1, which we denote by r. We partition this vector as $r = \begin{bmatrix} r_n^\mathsf{T} & r_m^\mathsf{T} \end{bmatrix}^\mathsf{T}$, $r_n \in \mathbb{R}^n_+, r_m \in \mathbb{R}^m_+$.

Proposition 3 *Given a logistics network described by the matrix $\mathcal{M}$ that is embedded into a larger network described by (8). Let the Perron vector r be partitioned as $\begin{bmatrix} r_n^\mathsf{T} & r_m^\mathsf{T} \end{bmatrix}^\mathsf{T}$. Then r_n is an eigenvector corresponding to the eigenvector 1 of the matrix*

$$c\mathcal{M} + (1-c)\frac{1}{n}\mathbf{e}_n\mathbf{e}_n^\mathsf{T}.$$

Proof. By assumption we have

$$r_n = \left(c\mathcal{M} + \frac{1-c}{n+m}\,\mathbf{e}_n\mathbf{e}_n^\mathsf{T}\right) r_n + \frac{1}{n+m}\,\mathbf{e}_n\mathbf{e}_m^\mathsf{T} r_m\,,$$

and

$$r_m = \frac{1-c}{n+m}\,\mathbf{e}_m\mathbf{e}_n^\mathsf{T} r_n + \frac{1}{n+m}\,\mathbf{e}_m\mathbf{e}_m^\mathsf{T} r_m\,.$$

The second equation implies by multiplication from the left

$$\mathbf{e}_m^\mathsf{T} r_m = (1-c)\frac{m}{n+m}\,\mathbf{e}_n^\mathsf{T} r_n + \frac{m}{n+m}\,\mathbf{e}_m^\mathsf{T} r_m\,.$$

and so

$$\mathbf{e}_m^\mathsf{T} r_m = (1-c)\frac{m}{n}\,\mathbf{e}_n^\mathsf{T} r_n\,.$$

Inserting this expression in the first equation yields

$$\begin{aligned} r_n &= \left(c\mathcal{M} + \frac{1-c}{n+m}\,\mathbf{e}_n\mathbf{e}_n^\mathsf{T}\right) r_n \\ &\quad + \frac{1}{n+m}\,\mathbf{e}_n(1-c)\frac{m}{n}\,\mathbf{e}_n^\mathsf{T} r_n \\ &= \left(c\mathcal{M} + \frac{(1-c)}{n}\,\mathbf{e}_n\mathbf{e}_n^\mathsf{T}\right) r_n. \end{aligned}$$

This is the desired equality. □

However, in general the interconnection with the embedding network is not uniform. That is, some locations sell more products to exterior locations while others buy material from certain exterior locations. So the previous result has to be generalized in the following fashion, cf. [38]. Given a column stochastic matrix $\mathcal{M}$ let $v = \begin{bmatrix} v_n^\mathsf{T} & v_m^\mathsf{T} \end{bmatrix}^\mathsf{T}$ and $w = \begin{bmatrix} w_n^\mathsf{T} & w_m^\mathsf{T} \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{n+m}$ be two stochastic vectors that describe the relationship of the logistics network with the outside world. More precisely, the component w_m reflects the procurement of material from locations that are not contained in the network. The part v_n indicates the sale of material to locations outside the given logistics network. The corresponding weighted matrix for a large network with $n+m$ vertices is defined by

$$L = \begin{bmatrix} c\mathcal{M} + (1-c)w_n\mathbf{e}_n^\mathsf{T} & v_n\mathbf{e}_m^\mathsf{T} \\ (1-c)w_m\mathbf{e}_n^\mathsf{T} & v_m\mathbf{e}_m^\mathsf{T} \end{bmatrix}. \tag{9}$$

As before c describes the strength of the coupling between the original network given by $\mathcal{M}$ and the larger network. We assume that

$$v > 0 \quad \text{and} \quad w_m > 0. \tag{10}$$

Note that this implies the irreducibility of the matrix L. Then we obtain the following generalization.

Theorem 2 *Given a logistics network described by the matrix $\mathcal{M}$ and stochastic vectors v, w such that (10) is satisfied. Let the Perron vector r be partitioned as $[r_n^\mathsf{T} r_m^\mathsf{T}]^\mathsf{T}$. Then r_n is an eigenvector corresponding to the eigenvector 1 of the matrix*

$$\mathcal{M}_c(v,w) := c\mathcal{M} + (1-c)\left(w_n + \frac{\mathbf{e}_m^\mathsf{T} w_m}{1 - \mathbf{e}_m^\mathsf{T} v_m} v_n\right)\mathbf{e}_n^\mathsf{T}. \tag{11}$$

Furthermore, $\mathcal{M}_c(v,w)$ is primitive.

Proof. By assumption we have

$$L \begin{bmatrix} r_n \\ r_m \end{bmatrix} = \begin{bmatrix} r_n \\ r_m \end{bmatrix}$$

and so

$$\begin{aligned} r_n &= \left(c\mathcal{M} + (1-c)w_n \mathbf{e}_n^\mathsf{T}\right) r_n + v_n \mathbf{e}_m^\mathsf{T} r_m, \\ r_m &= (1-c) w_m \mathbf{e}_n^\mathsf{T} r_n + v_m \mathbf{e}_m^\mathsf{T} r_m. \end{aligned}$$

Multiplying the second equation from the right by $\mathbf{e}_m^\mathsf{T}$ we obtain

$$\mathbf{e}_m^\mathsf{T} r_m = (1-c)\mathbf{e}_m^\mathsf{T} w_m \mathbf{e}_n^\mathsf{T} r_n + \mathbf{e}_m^\mathsf{T} v_m \mathbf{e}_m^\mathsf{T} r_m \,,$$

and so

$$\mathbf{e}_m^\mathsf{T} r_m = (1-c)\frac{\mathbf{e}_m^\mathsf{T} w_m}{1 - \mathbf{e}_m^\mathsf{T} v_m} \mathbf{e}_n^\mathsf{T} r_n \,.$$

Note that the previous expression is well defined, as it holds by assumption that $v_n \neq 0$ and so $\mathbf{e}_m^\mathsf{T} v_m < 1$. Inserting this in the first equation we obtain

$$\begin{aligned} r_n &= \left(c\mathcal{M} + (1-c)w_n \mathbf{e}_n^\mathsf{T}\right) r_n \\ &\qquad + v_n (1-c)\frac{\mathbf{e}_m^\mathsf{T} w_m}{1 - \mathbf{e}_m^\mathsf{T} v_m} \mathbf{e}_n^\mathsf{T} r_n \\ &= \left(c\mathcal{M} + (1-c)\left(w_n + \frac{\mathbf{e}_m^\mathsf{T} w_m}{1 - \mathbf{e}_m^\mathsf{T} v_m} v_n \right) \mathbf{e}_n^\mathsf{T} \right) r_n. \end{aligned}$$

This implies the first assertion.

To prove the second assertion, assume that $\mathcal{M}_c(v,w)$ is reducible. Then there exists a permutation matrix $P \in \mathbb{R}^{n \times n}$ such that

$$P^\mathsf{T} \mathcal{M}_c(v,w) P = \begin{bmatrix} Q_{11} & 0 \\ Q_{12} & Q_{22} \end{bmatrix}$$

with quadratic matrices $Q_{ii} \in R^{n_i \times n_i}, i = 1, 2$. This implies that the first n_1 entries of $P^\mathsf{T} v_n$ and $P^\mathsf{T} w_n$ are both zero. For the augmented permutation matrix $\mathrm{diag}(P, I_m)$ it follows that

$$\begin{bmatrix} P^\mathsf{T} & 0 \\ 0 & I_m \end{bmatrix} L \begin{bmatrix} P & 0 \\ 0 & I_m \end{bmatrix} = \begin{bmatrix} Q_{11} & 0 & 0 \\ Q_{12} & Q_{22} & * \\ * & * & * \end{bmatrix}$$

contradicting the irreducibility of L. The assertion follows from Lemma 2 as the second term in (11) has a nonzero diagonal entry. □

The framework derived in this section provides a unique stationary distribution of the discrete Markov chain with transition matrix $\mathcal{M}_c(v,w)$. According to the interpretation of the importance of locations in a logistics network this approach uniquely determines the importance of each location.

Definition 1 *Let $\mathcal{G} = (V, \mathcal{E}, \mathcal{M}_c(v,w))$ describe a logistics network. The normalized eigenvector associated to the eigenvalue one of $\mathcal{M}_c(v,w)$ is called the LogRank of $\mathcal{G}$.*

3.3 Comparison to other ranking schemes

A common feature of ranking schemes such as PageRank, HITS, SALSA, and LogRank is the attempt to extract information from the network structure. Conceptually the steps in the definition of LogRank are comparable to the ranking scheme known as PageRank. Note that the justification of these steps is based on inherent logistics considerations.

The PageRank algorithm is based on the following equation, where the PageRank r_i of vertex i is

$$r_i = \sum_{j \in P(i)} \frac{1}{out(j)} \, r_j. \tag{12}$$

This means that in the PageRank scheme for a vertex the number of the predecessors and their outdegree is of interest and moreover each vertex shares its PageRank equally to all its successors. This is contrary to the LogRank scheme for two reasons. To see this we note that basic part of (11) can be written as

$$r_i = \sum_{j \in S(i)} \frac{f_{ij}}{\sum_{j \in P(i)} f_{ij}} \, r_j. \tag{13}$$

So (i) the LogRank focuses on the indegree and (ii) a vertex shares its LogRank proportionally to the sum of the weights of all ingoing edges. Roughly speaking, a significant difference between the ranking schemes is given by the directions of the edges under consideration. Furthermore, the handling of zero columns in the original matrix formulation is different. Zero columns represent vertices with outdegree zero in the PageRank algorithm, whereas in the LogRank scheme they represent vertices with indegree zero. The PageRank ensures that the transition matrix is column stochastic by the addition of the term $\frac{1}{n}\mathbf{e}_n \, d^\mathsf{T}$. This reflects the idea that from every sink vertex there are edges with the uniform weight $\frac{1}{n}$ to every vertex of the graph. The vector $d \in \mathbb{R}_+^n$ is defined by $d_i = 1$ if $\mathrm{out}(i) = 0$ and $d_i = 0$ else. In contrast to this the LogRank scheme uses network inherent information, as described in Section 3.1. However, the significance of Theorem 2 lies in the fact, that the PageRank procedure for ensuring primitivity of the matrix is equivalent to the assumption that the network under consideration is weakly coupled with a large fully interconnected network. That is, since

$$\mathbf{e}_n^\mathsf{T} \left(w_n + \frac{\mathbf{e}_m^\mathsf{T} w_m}{1 - \mathbf{e}_m^\mathsf{T} v_m} v_n \right) = 1$$

the term $w_n + \frac{\mathbf{e}_m^\mathsf{T} w_m}{1 - \mathbf{e}_m^\mathsf{T} v_m} v_n$ coincides with the teleportation vector of the personalized PageRank algorithm [31]. The embedding interpretation appears to be new even in the context of PageRank.

The standard PageRank scheme applied to the example given by Figure 1 is

$$\left[\alpha(N + \frac{1}{n}\mathbf{e}_n \, d^\mathsf{T}) + (1-\alpha)\frac{1}{n}\mathbf{e}_n \right] r = r,$$

where the matrix N is given by

$$N = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

The result for this case with $\alpha = 0.95$ is

$$r \approx \begin{bmatrix} .0793 & .0793 & .1924 & .1170 & .1707 & .1707 & .1905 \end{bmatrix}^{\mathsf{T}}.$$

It can be seen that the standard PageRank scheme has the drawback, that all source suppliers have the same rank. The reason for this phenomenon is, that the source suppliers represent sources and the term $\frac{1}{n}\mathbf{e}_n\, d^{\mathsf{T}}$ does not use all network inherent information, as described in Subsection 3.1. In particular, each source supplier receives the same share of rank from all other locations.

An application of the method proposed in the Subsections 3.1 and 3.2 to the example from Figure 1, where the interaction with locations outside the logistics network is assumed to be

$$v = \frac{1}{20}\begin{bmatrix} 3 & 1 & 6 & 1 & 2 & 1 & 1 & 2 & 2 & 1 \end{bmatrix}^{\mathsf{T}},$$
$$w = \frac{1}{9}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 3 & 4 \end{bmatrix}^{\mathsf{T}},$$

and $c = 0.95$ leads to the LogRank

$$r \approx \begin{bmatrix} .1278 & .2055 & 0.2066 & 0.1301 & .1558 & .0407 & .1335 \end{bmatrix}^{\mathsf{T}}. \tag{14}$$

Remark 2 (Summary) What all this amounts to is that the LogRank scheme to identify the importance of locations in logistics network has the following characteristics. The embedding of the given network into its external world guarantees the irreducibility of the matrix $\mathcal{M}_c(v, w)$ and thus the uniqueness of the LogRank. Even if the matrix $\mathcal{M}$ is already irreducible, it is reasonable to embed the network into a larger one. The reason for this can be seen by comparing the results of the example. In (6) the network is not embedded into a larger network, whereas in (14) the relation to exterior locations has been taken into account. This information allows a specification of the LogRank such that the LogRank of location four and seven does not coincide.

4 LogRank preserving graph transformations

In this section we make use of the rank and graph structure information for the reduction of network size. In particular, we will apply the key result Theorem 2 provided in the previous Section 3 to calculate the LogRank of each location. Thus we modify the matrix M according to (3) and (11) and obtain the matrix $\mathcal{M}_c(v, w)$. To keep the notation simple we denote this matrix $\mathcal{M}_c(v, w)$ just by A. Let A be column stochastic. Hence the corresponding graph $G = (V, E, A)$ is strongly connected.

In our reduction approach we are going to omit the vertices with low rank. During the process of approximation we aim to preserve the main structure of material flow of the original network. To this end we will introduce aggregation rules for vertices with low rank; for typical subgraphs occurring in the network. Such subgraphs we will call *motifs* [39]. The main object of interest for the structural model reduction of logistics networks is the matrix M that describes an interconnection structure of the network. This matrix will be used to identify the motifs and to describe aggregation rules. We single out the following motifs: parallel connections, sequential connections of vertices and almost disconnected subgraphs. These reduction rules are inspired by the rank properties of motifs in [40,41].

4.1 Aggregation of vertices

By aggregation of vertices we understand the construction of smaller graphs in which vertices may represent nonempty subsets of vertices in the original graph. In the most general case we can assume that we aggregate arbitrary subsets of vertices of the graph $G = (V, E, A)$.

To this end we consider a disjoint partition

$$V = \{1, \ldots, n\} =: J_1 \cup\ J_2 \cup \ldots \cup J_k\,.$$

The vertices of the reduced graph $\tilde{G} = (\tilde{V}, \tilde{E}, \tilde{A})$ are given by

$$\tilde{V} = \{J_1, \ldots, J_k\}\,. \tag{15}$$

The edge (J_j, J_i) is in the edge set $\tilde{E}$ if there are $k \in J_j, l \in J_i$ such that $(k, l) \in E$. In other words, the subsets of vertices J_j keep the connections which were present in the graph G. Finally, we have to define the weights of $\tilde{A}$. Assuming that A is irreducible we use the normalized Perron vector r of A and define

$$\tilde{a}_{ij} := \sum_{\nu \in J_i} \frac{1}{\sum_{\mu \in J_j} r_\mu} \sum_{\mu \in J_j} r_\mu a_{\nu\mu}\,. \tag{16}$$

Theorem 3 *Consider a strongly connected weighted directed graph $G = (V, E, A)$. Let r be the (unique) normalized Perron vector of A. Given a disjoint partition*

$$V = \{1, \ldots, n\} =: J_1 \cup J_2 \cup \ldots \cup J_k\,,$$

and the corresponding construction of the reduced graph $\tilde{G} = (\tilde{V}, \tilde{E}, \tilde{A})$ given in (15), (16)*, then $\tilde{A}$ is irreducible and the unique normalized Perron vector $\tilde{r}$ of $\tilde{A}$ has the property*

$$\tilde{r}_i = \sum_{\nu \in J_i} r_\nu \quad , i = 1, \ldots, k\,. \tag{17}$$

Proof. The proof follows by calculation. By construction it is clear that $\tilde{r} > 0$ and $\sum_{i=1}^{k} \tilde{r}_i = 1$. So compute

$$\begin{aligned}\left(\tilde{A}\tilde{r}\right)_i &= \sum_{j=1}^{k} \tilde{a}_{ij}\tilde{r}_j = \sum_{j=1}^{k} \sum_{\nu\in J_i} \frac{1}{\sum_{\mu\in J_j} r_\mu} \sum_{\mu\in J_j} r_\mu a_{\nu\mu} \sum_{\eta\in J_j} r_\eta \\ &= \sum_{\nu\in J_i} \sum_{j=1}^{k} \frac{1}{\sum_{\mu\in J_j} r_\mu} \sum_{\mu\in J_j} r_\mu a_{\nu\mu} \sum_{\eta\in J_j} r_\eta \\ &= \sum_{\nu\in J_i} \sum_{j=1}^{k} \sum_{\mu\in J_j} r_\mu a_{\nu\mu} = \sum_{\nu\in J_i} \sum_{\mu=1}^{n} r_\mu a_{\nu\mu} = \sum_{\nu\in J_i} r_\nu \\ &= \tilde{r}_i\,,\end{aligned}$$

where in the last equation but one we have used that r is the normalized Perron eigenvector of A. Irreducibility of $\tilde{A}$ is clear by construction. □

Remark 3 (i) The construction in the previous lemma corresponds to: (a) taking the sums of the rows i for $i \in J_j$ to obtain the row associated to i and (b) taking the weighted sums of the columns weighted with the corresponding entry of the normalized leading Perron vector. Assuming that the vertices are ordered so that sets J_j contain j_k consecutive vertices, i.e., $J_1 = \{1, \dots, j_1\}, J_2 = \{j_1 + 1, \dots, j_1 + j_2\}, \dots, J_k = \{1 + \sum_{l=1}^{k-1} j_l, \dots, \sum_{l=1}^{k} j_j\}$ these operations may be given the matrix representation

$$\tilde{A} = \Psi\, A\, \Phi,$$

where

$$\Psi := \begin{bmatrix} 1 \dots 1\, 0 \dots 0 & \dots & & 0 \\ 0 \dots 0\, 1 \dots 1 & 0 & \dots & 0 \\ \vdots \quad \vdots & \ddots & & \\ & \dots\, 0 & 1 & \dots 1 \;\; 0 \\ 0 \dots 0 & \dots & 0 & \dots 0 \;\; I_{\#J} \end{bmatrix},$$

and

$$\Phi := \begin{bmatrix} \tilde{v}_1 & 0 & 0 & 0 & & & & \\ \vdots & \vdots & & & & & & \\ \tilde{v}_{j_1} & 0 & \vdots & & & & & \\ 0 & \tilde{v}_{j_1+1} & & & \dots & & & \\ \vdots & \vdots & 0 & & & & & \\ 0 & \tilde{v}_{j_1+j_2} & & & & & & \\ & & \ddots & & & & & \\ & & & \tilde{v}_{1+\sum_{l=1}^{k-1} j_l} & 0 \\ & & & \vdots & \vdots \\ & & & \tilde{v}_{\sum_{l=1}^{k} j_j} & 0 \\ & & & & I_{\#J} \end{bmatrix}.$$

(ii) General aggregation, as discussed in Theorem 3, may not preserve the main network structure if applied to arbitrary subgraphs. Therefore we develop aggregation rules for typical network motifs.

4.2 Typical motifs

We consider three types of motifs. The first motif consists of vertices that are connected in parallel, see Figure 4 for an example of such a motif. The second motif is given by sequentially connected vertices and is shown in Figure 6. An almost disconnected subgraph is the third type of motif, see Figure 8. In the following subsections we will show how such motifs can be aggregated using (16). The information about the material flow and the structure of a logistics network will be reflected in the graph $G = (V, E, A)$. The result of Theorem 3 will guarantee the rank preserving property.

4.2.1 Parallel connections

Parallel connections are characterized by vertices having the same predecessor and successor sets consisting of a single vertex. That is, the vertices $V_J := \{v_1, \dots, v_k\} \subset V$ are *connected in parallel*, if every vertex has only one ingoing and one outgoing edge and the ingoing edges originate from one vertex $v \in V$ and also the outgoing edges end in solely one vertex $v' \in V$. To be precise, $V_J = \{i \in V : P(i) = v, S(i) = v'\}$. As an example in Figure 4 the vertices $v_1, \dots, v_k$ are connected in parallel.

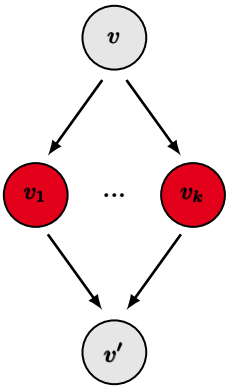

Fig. 4. Parallel connection of vertices $v_1, \dots, v_k$.

To obtain a graph of a smaller size we aggregate the vertices V_J to a single vertex J and do not change the remaining graph. So, we consider the smaller network described by the reduced graph $\tilde{G} = (\tilde{V}, \tilde{E}, \tilde{A})$, where the vertices are given by

$$\tilde{V} = (V \setminus V_J) \cup J, \tag{18}$$

and the edges are given by

$$\begin{aligned}\tilde{E} = E \setminus \{(v, w), (w, v') : w \in V_J\} \\ \cup \{(v, J), (J, v') : J, v, v' \in \tilde{V}\}.\end{aligned} \tag{19}$$

The corresponding cut-out of the reduced graph is shown in Figure 5. In the following we renumber the vertices in V such that the first $l = n - k - 2$ vertices correspond to the

Fig. 5. The vertices $v_1, \dots, v_k$ are aggregated to J.

vertices that are not directly connected with the parallel connection, the vertex v corresponds to $l+1$, the vertex v' corresponds to $l+2 = n-k$ and the vertices $v_1, \dots, v_k$ connected in parallel have the labels $n-k+j, j = 1, \dots, k$. The weights of the reduced weighted adjacency matrix $\tilde{A}$ are obtained according to the formula (16). That is,

$$\tilde{a}_{ij} = a_{ij} \tag{20}$$

for $i, j = 1, \dots, l+2$. The last row $i = l+3$ except for the diagonal element is given by

$$\tilde{a}_{ij} = a_{n-k+l,j} + \dots + a_{n,j}. \tag{21}$$

The last column $j = l+3$ except for the diagonal element is given by

$$\tilde{a}_{ij} = \frac{r_{n-k+1}\, a_{i,n-k+l} + \dots + r_n\, a_{i,n}}{r_{n-l+1} + \dots + r_n} \tag{22}$$

and the diagonal element is

$$\begin{aligned} \tilde{a}_{l+3,l+3} = {} & \frac{r_{n-k+1}\,(a_{n-k+1,n-k+1} + \dots + a_{n,n-k+1})}{r_{n-k+1} + \dots + r_n} \\ & + \dots + \frac{r_n\,(a_{n,n-k+1} + \dots + a_{n,n-k+1})}{r_{n-k+1} + \dots + r_n}. \end{aligned} \tag{23}$$

Note that $\tilde{A}$ matrix is irreducible if A is. By Theorem 3 the LogRank of the reduced network $\tilde{G} = (\tilde{V}, \tilde{E}, \tilde{A})$ has the following properties.

Corollary 1 *Let $G = (V, E, M)$ be a directed graph. Let r be the LogRank calculated using the matrix A obtained by (3) and (11). Assume the vertices $V_J \subset V$ are connected in parallel. Consider the reduced graph $\tilde{G} = (\tilde{V}, \tilde{E}, \tilde{M})$ with the vertex set $\tilde{V}$ given by (18), edge set $\tilde{E}$ given by (19). Let $\tilde{r}$ be the LogRank of $\tilde{A}$ defined by (20)-(23).*

Then the LogRank of $v_J \in \tilde{V}$ is the sum of the LogRanks of the aggregated vertices in V_J, while the LogRank of the unaffected vertices $v_1, \dots, v_{l+2}$ is preserved. That is,

$$\tilde{r} = \left[r_1 \; \dots \; r_{l+2} \; r_{n-k+1} + \dots + r_n \right]^{\mathsf{T}}.$$

4.2.2 Sequential connections

The second motif is given by sequentially connected vertices that we describe as follows. The vertices of the set $V_J = \{v_1, \dots, v_k\}$ are called *sequentially connected*, if there exist vertices $v, v' \in V \setminus V_J$ such that

$$P(v_i) = \begin{cases} v & i = 1, \\ v_{i-1} & i = 2, \dots, k \end{cases}$$

and

$$S(v_i) = \begin{cases} v_{i+1} & i = 1, \dots, k-1, \\ v' & i = k. \end{cases}$$

In Figure 6 a cut-out of a graph is shown, where the vertices $V_J = \{v_1, \dots, v_k\}$ are sequentially connected. Based

Fig. 6. Sequential connection of vertices $v_1, \dots, v_k$.

on this structure a possibility to attain a graph of smaller size is to aggregate the sequentially connected vertices to a single vertex and to leave the structure of the remaining graph as it is. We denote the new vertex by J. A cut-out of the new reduced graph is shown in Figure 7. So we con-

Fig. 7. Vertices $v_1, \dots, v_k$ are aggregated.

sider the reduced graph $\tilde{G} = (\tilde{V}, \tilde{E}, \tilde{A})$, where the vertices are given by

$$\tilde{V} = (V \setminus V_J) \cup J \tag{24}$$

and the edges are given by

$$\begin{aligned} \tilde{E} = {} & E \setminus \{(v, w), (w, v'), (w_1, w_2) : w, w_1, w_2 \in V_J\} \\ & \cup \{(v, u), (u, v') : u, v, v' \in \tilde{V}\}. \end{aligned} \tag{25}$$

In the following we renumber the vertices V such that the first $l = n-k-2$ vertices correspond to the vertices that are not directly connected with the sequential connection, the vertex v is labeled by $l+1$, the vertex v' by $l+2$ and the vertices $v_1, \dots, v_k$ receive the labels $n-k+j, j = 1, \dots, k$.

The corresponding weighted adjacency matrix $\tilde{A}$ can be created from A where the rows and columns corresponding to the vertices $v_1, \dots, v_k$ are replaced by a row and column corresponding to new vertex J. The weights are then given by (16), which coincides with the expressions (20)-(23). Note that $\tilde{A}$ is irreducible. By Theorem 3 the LogRank of the reduced network $\tilde{G} = (\tilde{V}, \tilde{E}, \tilde{A})$ has the following properties.

Corollary 2 *Let $G = (V, E, M)$ be a directed graph. Let r be the LogRank calculated using the matrix A obtained by* (3) *and* (11)*. Assume the vertices $V_J \subset V$ are sequentially connected. Consider the reduced graph $\tilde{G} = (\tilde{V}, \tilde{E}, \tilde{M})$ with the vertex set $\tilde{V}$ given by* (24)*, edge set $\tilde{E}$ given by* (25)*. Let $\tilde{r}$ be the LogRank of $\tilde{A}$ defined by* (20)-(23)*.*

Then the LogRank of $v_J \in \tilde{V}$ is the sum of the LogRanks of the aggregated vertices in V_J, while the LogRank of the unaffected vertices $v_1, \dots, v_{l+2}$ is preserved. That is,

$$\tilde{r} = \begin{bmatrix} r_1 \dots r_{l+2}\ r_{n-k+1} + \dots + r_n \end{bmatrix}^\mathsf{T}.$$

Remark 4 Note that the results for the parallel and sequential case are almost identical, which maybe does not correspond to the intuition coming from other (e.g. electrical networks. The reason for this is the scaling performed in (16) which takes the approach outside of the realm of other network analysis tools.

4.2.3 Almost Disconnected Subgraphs

A further structure in the network, that suggests itself to a reduction is given by subgraphs which are connected to the remainder of the network through just a single vertex. So we consider a set of vertices $V_J = \{v_1, \dots, v_k\}$ and an distinguished vertex $v^* \in V \setminus V_J$ such that any path from $v_i, i = 1, \dots, k$ to the remainder of the vertices in $V \setminus V_J$, and any path from $V \setminus V_J$ to V_J necessarily passes through the vertex v^*. If we assume that the whole graph is strongly connected, this implies in particular, that the subgraph induced by $V_J \cup \{v^*\}$ is by itself strongly connected.

In Figure 8 an example graph is shown, where the vertices $V_J = \{v_1, \dots, v_k\}$ are connected with the rest of the graph only through the vertex v.

To reduce the network size we aggregate the vertices of the subgraph V_J with vertex v^* and do not change the remainder of the graph. We denote the new vertex by J. For the example in Figure 8 the reduced graph is shown in Figure 9. So we consider the reduced graph $\tilde{G} = (\tilde{V}, \tilde{E}, \tilde{A})$, where the vertices are given by

$$\tilde{V} = (V \setminus (V_J \cup v^*)) \cup J \tag{26}$$

and the edges are given by

$$\begin{aligned} \tilde{E} = E \setminus \{&(w_1, w_2), (v^*, w_1), (w_1, v^*) : w_1, w_2 \in V_J\} \\ &\cup \{(J, u) : u \in \tilde{V}, (v^*, u) \in E)\} \\ &\cup \{(u, J) : u \in \tilde{V}, (u, v^*) \in E)\}. \end{aligned} \tag{27}$$

The corresponding weighted adjacency matrix $\tilde{A}$ can be created from A where the rows and columns corresponding to the vertices $v^*, v_1, \dots, v_k$ are replaced by a row and column corresponding to new vertex J. The corresponding weights are given by (16), which coincides with the expressions in front of Corollary 1. By Theorem 3 the LogRank

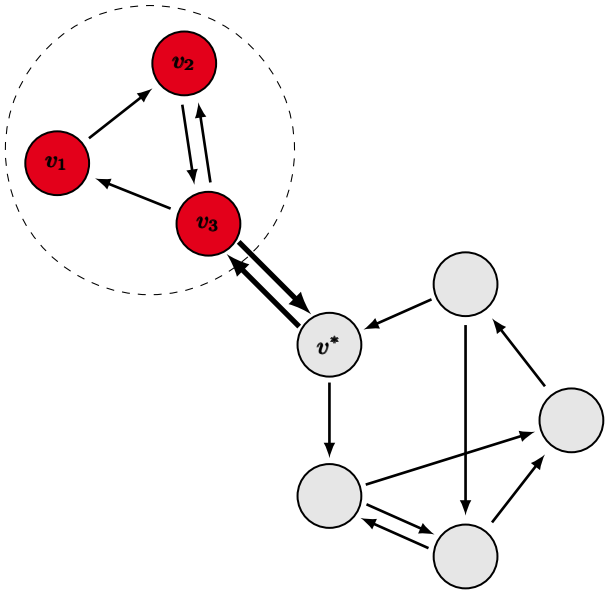

Fig. 8. The subgraph consisting of the vertices $V_J = \{v_1, v_2, v_3\}$ is almost disconnected from the graph.

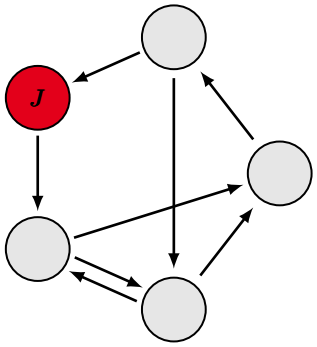

Fig. 9. Subgraph V_J and node v^* are merged to vertex J.

of the reduced network $\tilde{G} = (\tilde{V}, \tilde{E}, \tilde{A})$ has the following properties.

In the following we number the vertices V such that the first $l = n - k - 1$ vertices correspond to the vertices from $V \setminus (V_J \cup v)$, the vertex v^* corresponds to $l + 1$, and the vertices from the subgraph $v_1, \dots, v_k$ have the labels $n - k + 1, \dots, n$.

Corollary 3 *Let $G = (V, E, M)$ be a directed graph. Let r be the LogRank calculated using the matrix A obtained by* (3) *and* (11)*. Assume the vertices $V_J \subset V$ form a subgraph connected with the rest of the graph through the vertex v^*. Consider the reduced graph $\tilde{G} = (\tilde{V}, \tilde{E}, \tilde{M})$ with the vertex set $\tilde{V}$ given by* (26)*, edge set $\tilde{E}$ given by* (27)*. Let $\tilde{r}$ be the LogRank of $\tilde{A}$ defined by* (16)*.*

Then the LogRank of $v_J \in \tilde{V}$ is the sum of the LogRanks of the aggregated vertices in V_J, while the LogRank of the unaffected vertices $v_1, \dots, v_l$ is preserved. That is,

$$\tilde{r} = \begin{bmatrix} r_1 \dots r_l\ r_{n-k} + \dots + r_n \end{bmatrix}^\mathsf{T}.$$

4.3 Discussion of reduction rules

In this section we have introduced a broadly defined aggregation process based on the LogRank. In the most general case this process is hardly structure-preserving. For this reason motifs have been considered as a way of controlling the way of changing the structure. The three typical motifs parallel connections, sequential connections, and almost disconnected subgraphs have been discussed in detail. In logistics contexts further motifs might be of interest and could be discussed based on Subsection 4.1. The reduction rules presented in this section preserve the main structure of the network as well the ranks of the unaffected vertices. In the following section we will utilize the ranking technique from Section 3 and reduction rules for typical motifs from Sections 4.2.1-4.2.3 to introduce a meta algorithm for the reduction of large-scale logistics networks.

5 Meta algorithm for structural model reduction guided by LogRank

Large-scale logistics networks can be modeled as dynamical systems, which can consist of several hundreds of locations and connections. In order to analyze such a network (e.g. in terms of logistic performance or robustness) models are used. A model of smaller size that approximates the characteristics of the original network facilitates the analysis and is therefore desirable. In particular, given the same input a model of smaller size should produce a similar output compared to the original model by not exceeding a predefined acceptable error. The measurement of the error is in general, driven by an underlying cost function, which is foremost in the analysis of approximation error. However, many costs associated to logistics networks actually depend on the input-output behavior, as lead time, level of work in progress, etc. In the context of logistics networks two different aspects of model reduction exist. First, the structure of the network can be simplified, whilst preserving essential structures of the network. Secondly, the accuracy of the modeled dynamics of the network can be reduced in order. Note, that both aspects are interdependent.

In this section we introduce our approach to structure-preserving model reduction. This approach is guided by the LogRank introduced in Section 3. As we do not want to restrict ourselves to a particular model class, we present a meta algorithm for model reduction of logistics networks.

We use the following notation. The input-output map of the network $\mathcal{G} = (V, \mathcal{E}, \mathcal{M}_c(v, w))$ is denoted by f and $\tilde{f}$ denotes the input-output map for the network $\tilde{\mathcal{G}} = (\tilde{V}, \tilde{\mathcal{E}}, \tilde{\mathcal{M}}_c(v, w))$. The comparison of the given network and its approximation is obtained by comparing the error e in the associated cost, input-output behavior or gain predicted by the original model compared to the reduced order model, depending on the situation at hand. Due to the fact that any model reduction is a trade-off between the size of the approximating system and the error, let $\varepsilon > 0$ denote an acceptable upper bound for the reduction error. Further, there is an upper bound Δ for the LogRank of the vertices that are allowed to be reduced. The set of vertices that has a LogRank less than Δ is denoted by R_Δ. That is,

$$R_\Delta = \{v \in V : \text{LogRank}(v) < \Delta\}.$$

The vertices of $R_\Delta = \{v_1, \ldots, v_l\}$ are arranged in ascending order w.r.t. their LogRank, i.e., $v_i \leq v_j$ if and only if $\text{LogRank}(v_i) \leq \text{LogRank}(v_j)$. Consequently, this set contains the candidates of vertices that might be aggregated with other vertices. For the vertex $v_1 \in R_\Delta$ the candidate list $C(v_1, r)$ of reasonable motifs containing v_1 is then set up. These motifs are assessed based on the rank r of the involved locations and ordered in ascending order.

Algorithm 1 summarizes the steps of the structure-preserving approach to model reduction. The meta algorithm is not fixed for a specific class of mathematical models and can be adapted to various modeling approaches in order to capture the context specific characteristics of the given logistics network.

Algorithm 1 Meta algorithm for structure-preserving model reduction guided by LogRank.

```
Compute the LogRank r of the network G = (V, E, M_c(v, w))
and generate R_Δ
repeat
  Delete and consider v_1 ∈ R_Δ;
  Generate candidate list C = C(v_1, r);
  while C ≠ ∅ do
    Delete and consider c_1 from the candidate list C;
    if for c_1 reduction error e ≤ ε then
      Aggregate c_1;
      Clear C;
      Generate new waiting list R_Δ;
    end if
  end while
until R_Δ = ∅
```

The iterative algorithm is initialized with the ranking of the original large-scale logistics network. The obtained ranking is used in order to generate the waiting list R_Δ of locations that might be aggregated. The idea of the algorithm is to keep locations with a high importance and to try to consolidate locations with lower importance. To this end for the first location of the waiting list R_Δ the candidate list $C(v_1, r)$ is created. As long as the candidate list is not empty the first candidate is considered for model reduction. In particular this means that in a first attempt a simplification of the structure is performed. In a second step, if necessary, one tries to reduce the order of the modeled dynamic system. Afterwards the input-output mapping of the original structure and approximated structure is compared. This assessment needs to satisfy the previously defined upper bound for the reduction error ε. If the obtained error should be less than ε, then the model reduction is performed. In addition the candidate list $C(v_1, r)$ is cleared and a new waiting list R_Δ is derived based on the changed ranking of the locations. Otherwise the next

candidate is considered. This process continues until the candidate list is empty. Afterwards the next location of the waiting list R_Δ is considered. The whole process terminates when further model reduction can be performed successfully. The reason can be either that the model reduction error exceeds the upper bound for the reduction error ε or that no locations of lower importance exist that could be aggregated. In Section 6 the meta algorithm is applied to an example of a logistics network.

6 Example: Jackson networks

In this section we model logistics networks as Jackson networks and discuss in particular the effect of deleting network nodes based on the LogRank. To this end we follow the description of Jackson networks given in [10].

An open Jackson network consists of $V = \{1, \dots, n\}$ locations (in the theory of queueing networks these are frequently called *stations*), where orders (*jobs or customers*) are processed. The processing times of orders at each location are identically and independently distributed (i.i.d.), following an exponential distribution with a finite mean that may depend on the queue length. The processing rate, denoted $\mu_i(x_i)$, at location i where x_i products are present is given by a function

$$\mu_i : \mathbb{Z}_+ \to \mathbb{R}_+$$

with $\mu_i(0) = 0$ and $\mu_i(x) > 0$ for all $x > 0$. The evolution of the orders through the network is determined by the routing matrix P, where p_{ij} denotes the probability that an order leaving location i will go to location j. With this convention the row sums of P are bounded by 1. One of the main characteristics of Jackson networks is that all orders at each location belong to the same class and all orders follow the same service time distribution and the same routing mechanism. Hence Jackson networks belong to single class models. So at each location the orders are processed according to their arrivals, that is first-in-first-out (FIFO).

There are three different types of Jackson networks: open, closed and semiopen Jackson networks. They are distinguished by the way in which new orders may arrive from the outside. In the following we focus on open Jackson networks. Here the orders arrive from the outside according to a Poisson process with rate $\alpha > 0$. Each arriving order is independently routed to location i with probability $p_{0i} \geq 0$, with $\sum_{i=1}^n p_{0i} = 1$. This means that at location i the orders are arriving from outside with rate $\alpha\, p_{0i}$. After an order has been served at location i, it either routes through the network according to the routing matrix P, or leaves the network with probability $p_{i0} = 1 - \sum_{j=1}^n p_{ij}$. The matrix P is (row) substochastic and hence has spectral radius $\rho(P) \leq 1$. In the following we assume that $\rho(P) < 1$. This has the effect that the expected time that an order stays in the network is finite.

The *effective arrival rate* λ_i of orders at location i is given by the sum of arrivals from the outside and the internal transitions. This quantity satisfies the *traffic equation*

$$\lambda_i = \alpha\, p_{0i} + \sum_{j=1}^n p_{ji}\, \lambda_j, \qquad i \in V.$$

For $\lambda = \begin{bmatrix}\lambda_1 \dots \lambda_n\end{bmatrix}^\mathsf{T}$ and $a = \begin{bmatrix}\alpha\, p_{01} \dots \alpha\, p_{0n}\end{bmatrix}^\mathsf{T}$ a matrix form expression of the traffic equation is

$$\lambda = a + P^\mathsf{T}\lambda. \tag{28}$$

The number of orders present at location i at time t is denoted by $X_i(t)$ and from the description above it follows that

$$\{(X_i(t))_{i=1}^n : t \in \mathbb{R}_+\}$$

defines a continuous time Markov chain. Let $X = (X_i)_{i=1}^n$, $x = (x_i)_{i=1}^n$ and $\pi(x) = \mathbb{P}[X = x]$ denote the stationary distribution. Then, if for $M_i(k) := \mu_i(1) \cdots \mu_i(k)$ it holds that

$$\sum_{k=1}^\infty \frac{\lambda_i^k}{M_i(k)} < \infty\,,$$

the stationary distribution π of the open Jackson network is of the following product form

$$\pi(x) = \prod_{i=1}^n \left(1 - \sum_{k=1}^\infty \frac{\lambda_i^k}{M_i(k)}\right)^{-1} \frac{\lambda_i^{x_i}}{M_i(x_i)}$$

for all $x \in \mathbb{Z}_+^n$, cf. [10, Theorem 2.1]. Consequently, given the processing rates for the locations all that is needed to write down the stationary distribution is the effective arrival rate λ.

6.1 Effective arrival rates and ranks

In the following we explain that for particular Jackson networks the ranking scheme proposed in Section 3 is in one to one correspondence with the effective arrival rate of the locations.

To this end we define the outflow probability vector $p_o = \begin{bmatrix}p_{10} \dots p_{n0}\end{bmatrix}^\mathsf{T}$, the external inflow probability vector $p_e = \begin{bmatrix}p_{01} \dots p_{0n}\end{bmatrix}^\mathsf{T}$ and the stochastic matrix

$$\overline{P} := \begin{bmatrix} P^\mathsf{T} & p_e \\ p_o^\mathsf{T} & 0 \end{bmatrix}.$$

The interconnection of the Jackson network with the outside world is represented by the vectors $a = \alpha p_e$ and p_o.

In terms of the matrix $\overline{P}$ the outside world is represented by the $n+1$st row and column. The corresponding weighted graph $G(\overline{P}) = (\overline{V}, \overline{E}, \overline{P}^\mathsf{T})$ has the $n+1$ vertices $\overline{V} = V \cup \{n+1\}$ and the edge set $\overline{E}$ is determined by the nonzero entries of $\overline{P}^T$. So, on the subgraph induced by the vertices in V the graph $G(\overline{P})$ coincides with the

graph $G(P)$ representing the Jackson network. The additional vertex $n+1$ represents the outside world and edges from or to that vertex represent inflows to resp. outflows from the Jackson network.

We now show that in the context of logistic networks it is reasonable to assume that $\overline{P}$ is irreducible. To this end note that in logistics networks it is reasonable to assume that every location of the network receives products either from the outside world or from another location within the network. Otherwise there would be parts of the network not engaged in logistic activity which is implausible.

The condition that every location receives products directly from the outside or through other locations of the network can be formulated in terms of $G(\overline{P})$ as

$$\begin{gathered}\text{for all vertices } i \in V \\ \text{there exists a path from } n+1 \text{ to } i.\end{gathered} \tag{29}$$

Proposition 4 *Consider a Jackson network with routing matrix P. If $\rho(P) < 1$ and condition (29) is satisfied, then the matrix $\overline{P}$ is irreducible.*

Proof. Irreducibility of $\overline{P}$ is equivalent to the irreducibility of $\overline{P}^{\mathsf{T}}$ which is equivalent to the graph $G(\overline{P})$ being strongly connected. We show the latter, that is, for every pair of vertices $(i,j) \in V \cup \{n+1\}$ there is a path from i to j. From (29) it follows that there is a path from vertex $n+1$ to every vertex $i \in V$. Further, by definition the matrix $\overline{P}$ is column stochastic and since $\rho(P) < 1$ it follows that $p_{k0} > 0$ for some $k \in V$, so that the last row of $\overline{P}$ is not equal to 0. Assume now that there are vertices i for which there are no paths from i to $n+1$ and if necessary relabel the vertices so that there is no path from $W^- := \{1, \ldots, m\}$ to $n+1$ and for the remaining vertices $W^+ := \{m+1, \ldots, n\}$ such a path exists. Note that $m < n$ as $p_{k0} > 0$ for some $k \in V$. By construction there is no path from the vertices in W^- to those in W^+. Consequently, if we consider the Markov chain with transition properties given by $\overline{P}$ and initial condition supported on W^- then the states of the chain remain in W^- with probability 1. It follows that the chain admits a stationary measure supported on W^- and so the matrix P has an eigenvalue equal to 1. By assumption it follows that $W^- = \emptyset$.

Hence the graph $G(\overline{P})$ admits a path from k to $n+1$ for every vertex $k \in V$. This shows the assertion. □

From now on we assume that $\overline{P}$ is irreducible and return to the relation of Jackson networks to the LogRank. The traffic equation implies that $\alpha = \alpha \|p_e\| = \|(I - P^{\mathsf{T}})\,\lambda\|$, which in turn implies that

$$\alpha = \sum_{i=1}^{n} \lambda_i \left(1 - \sum_{j=1}^{n} p_{ji}\right) = \sum_{i=1}^{n} \lambda_i\, p_{i0} = p_o^{\mathsf{T}}\,\lambda. \tag{30}$$

The combination of the traffic equation (28) and (30) leads to the eigenvalue problem

$$\begin{bmatrix} P^{\mathsf{T}} & p_e \\ p_o^{\mathsf{T}} & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda \\ \alpha \end{bmatrix} = \begin{bmatrix} P^{\mathsf{T}}\lambda + a \\ \sum p_{i0}\,\lambda_i \end{bmatrix} = \begin{bmatrix} \lambda \\ \alpha \end{bmatrix}.$$

As $\overline{P}$ is irreducible, the effective arrival rate λ is the unique solution of

$$(P^{\mathsf{T}} + p_e p_o^{\mathsf{T}})\;\lambda = \lambda, \quad p_o^{\mathsf{T}}\lambda = \alpha. \tag{31}$$

If we compare this condition to the matrix $\mathcal{M}$ defined in (3), then we see a close relation, if we again consider the sets of retailers $\mathcal{R}$ and source suppliers $\mathcal{S}$. Namely, we can assume that $p_{ei} > 0$ if and only if $i \in \mathcal{S}$ and $p_{oi} > 0$ if and only if $i \in \mathcal{R}$. If for every source supplier there exists a path to any retailer, then the matrix $(P^{\mathsf{T}} + p_e p_o^{\mathsf{T}})$ defined in (31) is exactly the matrix $\mathcal{M}$ defined in (3). For this we interpret the Jackson network in terms of unfulfilled orders, so that p_{ij} represents the probability, that location i places an order at location j, if it receives an order. This is in correspondence with the previous definition of m_{ij}. If we now include a further external effect to the Jackson network in the form

$$\overline{P}_c(v,w) := c(P^{\mathsf{T}} + p_e p_o^{\mathsf{T}}) + (1-c)\left(w_n + \frac{\mathbf{e}_m^{\mathsf{T}} w_m}{1 - \mathbf{e}_m^{\mathsf{T}} v_m} v_n\right)\mathbf{e}_n^{\mathsf{T}}$$

this shows that for such Jackson networks the LogRank and the effective arrival rate coincide (up to a constant multiple).

Remark 5 We note that the previous argument is not evident in the general case. As we can see from the discussion in Section 3.1 the weight matrix $\mathcal{M}$ is in general a perturbation of M of higher rank. Whereas in the construction of Jackson networks the matrix $(P^{\mathsf{T}} + p_e p_o^{\mathsf{T}})$ is always a rank-one perturbation of P^{T}.

6.2 Omission of low rank locations

One approach to obtain a network of smaller size is the omission of locations that have an effective arrival rate smaller than a pre-defined number Δ. In this section we derive error estimates for the resulting effective arrival rates. To this end we order the locations of the Jackson network such that the effective arrival rates satisfy

$$\lambda_1 \geq \ldots \geq \lambda_l > \Delta \geq \lambda_{l+1} \geq \ldots \geq \lambda_n.$$

Further, by small abuse of notation let $\lambda_1 = \begin{bmatrix}\lambda_1 \ldots \lambda_l\end{bmatrix}^{\mathsf{T}}$ and $\lambda_2 = \begin{bmatrix}\lambda_{1+1} \ldots \lambda_n\end{bmatrix}^{\mathsf{T}}$. The corresponding eigenvalue problem can then be written as

$$\begin{bmatrix} P_{11}^{\mathsf{T}} & P_{12}^{\mathsf{T}} & p_{e1} \\ P_{21}^{\mathsf{T}} & P_{22}^{\mathsf{T}} & p_{e2} \\ p_{o1}^{\mathsf{T}} & p_{o2}^{\mathsf{T}} & 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \alpha \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \alpha \end{bmatrix}. \tag{32}$$

The reduced eigenvalue problem, where the locations with effective arrival rate smaller than Δ are excluded, can then be written the following

$$\begin{bmatrix} P_{11}^{\mathsf{T}} & D & \frac{p_{e1}}{\|p_{e1}\|} \\ p_{o1}^{\mathsf{T}} & D & 0 \end{bmatrix} \cdot \begin{bmatrix} \widetilde{\lambda} \\ \alpha \end{bmatrix} = \begin{bmatrix} \widetilde{\lambda} \\ \alpha \end{bmatrix}, \tag{33}$$

where $D = \mathrm{diag}(P_{11}\,\mathbf{e}_l + p_{o1})^{-1}$. In particular, we assume that the matrix in (33) has no zero columns and that (29) is applicable.

As a first result we note a minimal error that is necessary if deletions are performed.

Proposition 5 *Consider the effective arrival rate λ_1 defined by (32) and the effective arrival rate $\widetilde{\lambda}$ defined by (33). Then it holds that*

$$\|D\,\widetilde{\lambda} - \lambda_1\| \geq \frac{\|p_{o2}^{\mathsf{T}}\lambda_2\|}{\|p_{o1}^{\mathsf{T}}\|}.$$

Proof. On the one hand, by (32) it holds that

$$\alpha = p_{o1}^{\mathsf{T}}\lambda_1 + p_{o2}^{\mathsf{T}}\lambda_2.$$

On the other hand follows from (33) that

$$\alpha = p_{o1}^{\mathsf{T}}\,D\,\widetilde{\lambda}.$$

The combination of the two equalities implies that

$$\|p_{o2}^{\mathsf{T}}\lambda_2\| = \|p_{o1}^{\mathsf{T}}(D\widetilde{\lambda} - \lambda_1)\| \leq \|p_{o1}^{\mathsf{T}}\|\,\|D\widetilde{\lambda} - \lambda_1\|.$$

This shows the assertion. □

In the next result we obtain lower and upper bounds for the reduction error based on the deletion threshold that is chosen.

Proposition 6 *Consider the effective arrival rate λ_1 defined by (32) and the effective arrival rate $\widetilde{\lambda}$ defined by (33). Then the following holds true*

$$c_\alpha\,(I - P_{11}^{\mathsf{T}})^{-1}\,p_{e1} \leq \lambda_1 - \widetilde{\lambda} \leq \Delta(I - P_{11}^{\mathsf{T}})^{-1}P_{12}^{\mathsf{T}}\mathbf{e}_{n-l},$$

where $c_\alpha = \alpha\left(1 - \frac{1}{\|p_{e1}\|}\right)$.

Proof. From (32) and (33) it follows that

$$\begin{aligned}\lambda_1 &= P_{11}^{\mathsf{T}}\lambda_1 + P_{12}^{\mathsf{T}}\lambda_2 + \alpha\,p_{e1}\\ \widetilde{\lambda} &= P_{11}^{\mathsf{T}}\,D\widetilde{\lambda} + \frac{\alpha}{\|p_{e1}\|}p_{e1}.\end{aligned}$$

By using $c_\alpha := \alpha\left(1 - \frac{1}{\|p_{e1}\|}\right)$ the difference of λ_1 and $\widetilde{\lambda}$ can be represented by

$$\lambda_1 - \widetilde{\lambda} = P_{11}^{\mathsf{T}}\lambda_1 - P_{11}^{\mathsf{T}}\,D\widetilde{\lambda} + c_\alpha p_{e1} + P_{12}^{\mathsf{T}}\lambda_2. \tag{34}$$

First we verify the right inequality. Since $\|p_{e1}\| \leq 1$ it follows that

$$c_\alpha p_{e1} = \alpha\left(1 - \frac{1}{\|p_{e1}\|}\right)p_{e1} \leq 0,$$

where the inequality has to be understood componentwise. Further since the matrix P_{12}^{T} and the vector λ_2 are nonnegative and $\lambda_2 \leq \Delta\,\mathbf{e}_{n-l}$ it holds that

$$\lambda_1 - \widetilde{\lambda} \leq P_{11}^{\mathsf{T}}\left(\lambda_1 - D\widetilde{\lambda}\right) + \Delta\,P_{12}^{\mathsf{T}}\,\mathbf{e}_{n-l}.$$

Another representation of the last inequality is

$$\left(I - P_{11}^{\mathsf{T}}\right)\lambda_1 - \left(I - P_{11}^{\mathsf{T}}\,D\right)\widetilde{\lambda} \leq \Delta\,P_{12}^{\mathsf{T}}\,\mathbf{e}_{n-l}.$$

Since $\rho(P_{11}) < 1$ the inverse of $\left(I - P_{11}^{\mathsf{T}}\right)$ exists and is nonnegative. This implies

$$\begin{aligned}\lambda_1 - \left(I - P_{11}^{\mathsf{T}}\right)^{-1}\left(I - P_{11}^{\mathsf{T}}\,D\right)\widetilde{\lambda}\\ \leq \Delta\,\left(I - P_{11}^{\mathsf{T}}\right)^{-1}P_{12}^{\mathsf{T}}\,\mathbf{e}_{n-l}.\end{aligned}$$

From the fact that $\left(I - P_{11}^{\mathsf{T}}\right)^{-1} = \sum_{k=0}^{\infty}(P_{11}^{\mathsf{T}})^k$ it follows that

$$\left(I - P_{11}^{\mathsf{T}}\right)^{-1}\left(I - P_{11}^{\mathsf{T}}\,D\right) = I + \left(I - P_{11}^{\mathsf{T}}\right)^{-1}P_{11}^{\mathsf{T}}\,(I - D),$$

which implies that

$$\begin{aligned}\lambda_1 - \widetilde{\lambda} \leq \Delta\,\left(I - P_{11}^{\mathsf{T}}\right)^{-1}P_{12}^{\mathsf{T}}\,\mathbf{e}_{n-l}\\ +\,\left(I - P_{11}^{\mathsf{T}}\right)^{-1}P_{11}^{\mathsf{T}}\;(I - D)\,\widetilde{\lambda}.\end{aligned}$$

Furthermore it holds that $P_{11}\,\mathbf{e}_l + p_{o1} < \mathbf{e}_l$ and this implies that the diagonal entries of D are greater than 1. Consequently the matrix $I - D$ is nonpositive. Moreover, the nonnegativity of P_{11}^{T}, $\left(I - P_{11}^{\mathsf{T}}\right)^{-1}$ and $\widetilde{\lambda}$ shows that

$$\left(I - P_{11}^{\mathsf{T}}\right)^{-1}P_{11}^{\mathsf{T}}\,(I - D)\,\widetilde{\lambda} \leq 0.$$

Hence it holds that

$$\lambda_1 - \widetilde{\lambda} \leq \Delta\,\left(I - P_{11}^{\mathsf{T}}\right)^{-1}P_{12}^{\mathsf{T}}\,\mathbf{e}_{n-l}.$$

To prove the left inequality we start again with (34). Since P_{12}^{T} and λ_2 are nonnegative it follows that

$$\lambda_1 - \widetilde{\lambda} \geq P_{11}^{\mathsf{T}}\lambda_1 - P_{11}^{\mathsf{T}}\,D\widetilde{\lambda} + c_\alpha\,p_{e1}.$$

The matrix D is nonnegative and the diagonal entries are greater than one and the off-diagonal entries are zero. Hence as $\widetilde{\lambda}$ is nonnegative it holds that

$$\lambda_1 - \widetilde{\lambda} \geq P_{11}^{\mathsf{T}}\lambda_1 - P_{11}^{\mathsf{T}}\,\widetilde{\lambda} + c_\alpha\,p_{e1}.$$

Further, since $(I - P_{11}^{\mathsf{T}})^{-1}$ exists and is nonnegative it holds that

$$\lambda_1 - \widetilde{\lambda} \geq c_\alpha(I - P_{11}^{\mathsf{T}})^{-1}p_{e1}.$$

This shows the assertion. □

Simple omission of locations provides an easy way of reducing the network without paying too much attention to the error that is involved. However, the effective arrival rate λ determines the stationary distribution π. As the LogRank is closely related to λ, it may be of greater interest to perform the reduction such that the LogRank remains unchanged. We explain how to perform the necessary reduction steps by an application of the meta algorithm from Section 5 in the next section.

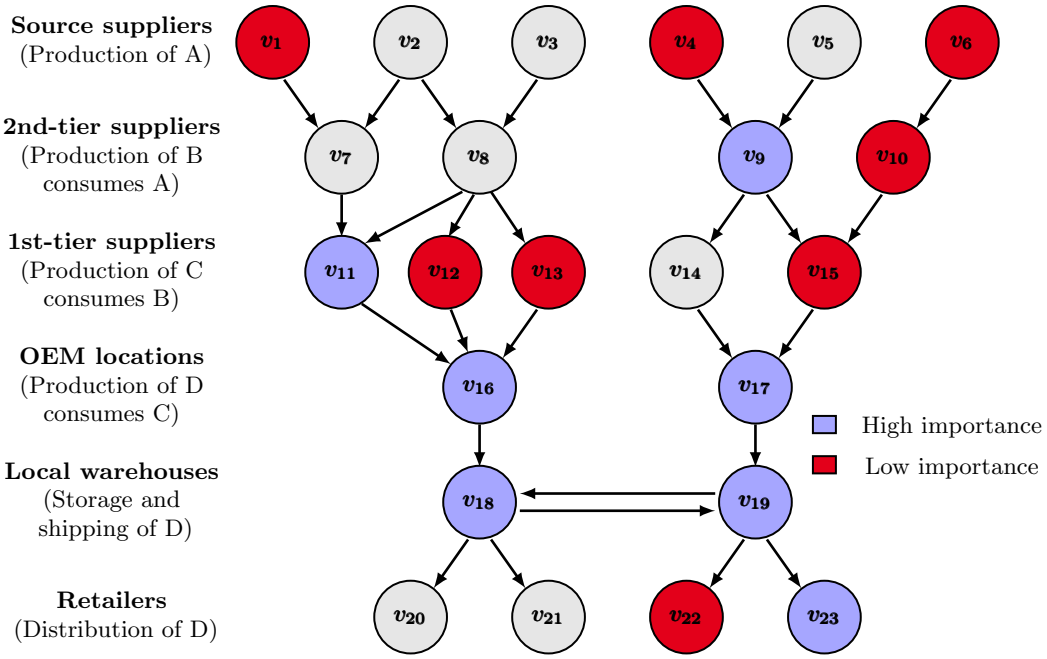

Fig. 10. The topology of the test example.

6.3 Test Case

In this subsection we discuss a test case that illustrates the scheme introduced in Section 5. We consider the following test scenario the topology of which is depicted in Figure 10. We would to point out that the structure of the network is close to symmetric. However, the magnitude of the flows leads to an asymmetric LogRank.

The weighted matrix $\mathcal{M}$ that takes the relation of the source suppliers and the retailers into account is given by

$$\mathcal{M} = \frac{1}{20}\begin{bmatrix} 0 & B \\ A & 0 \end{bmatrix}, \quad A = \begin{bmatrix} 6 & 6 & 6 & 6 & 6 & 6 \\ 5 & 5 & 5 & 5 & 5 & 5 \\ 2 & 2 & 2 & 2 & 2 & 2 \\ 7 & 7 & 7 & 7 & 7 & 7 \end{bmatrix}.$$

where the weighted flow matrix is

$$B = \begin{bmatrix}
10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
10 & 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 14 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 20 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 13 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 7 & 20 & 20 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 20 & 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 11 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 16 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 13 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 7 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 16 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 17 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 20 & 20 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 20 & 20
\end{bmatrix}.$$

The application of Theorem 2 leads to the following

$$v = \frac{1}{230}\begin{bmatrix} 3 \\ 14 \\ 10 \\ 17 \\ 2 \\ 1 \\ 8 \\ 3 \\ 16 \\ 5 \\ 19 \\ 17 \\ 13 \\ 3 \\ 6 \\ 1 \\ 4 \\ 14 \\ 18 \\ 12 \\ 9 \\ 5 \\ 11 \\ 3 \\ 10 \\ 6 \end{bmatrix}, \quad w = \frac{1}{9}\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 3 \\ 4 \\ 2 \end{bmatrix}, \quad \text{LogRank} \approx \begin{bmatrix} .0201 \\ .0355 \\ .0322 \\ .0189 \\ .0451 \\ .0083 \\ .0405 \\ .0445 \\ .0622 \\ .0169 \\ .0620 \\ .0116 \\ .0106 \\ .0500 \\ .0273 \\ .0787 \\ .0772 \\ .1024 \\ .0937 \\ .0489 \\ .0405 \\ .0165 \\ .0563 \end{bmatrix}.$$

The matrix B reflects the relative material flow between the locations and the matrix A represents the modification described in Subsection 3.1. The consideration of the corresponding bipartite graph that consists only of the source suppliers and the retailers shows that it is complete. This corresponds to the fact that the matrix B has

rank one. The embedding of the logistics network into its environment is described by the vectors v and w.

According to the LogRank we divide the set of locations as follows. The important locations are chosen to be those that have a LogRank of at least .055. In Figure 10 the blue vertices correspond to the highly important locations. Further, the locations with a LogRank lower than .03 are considered to be locations of low importance. These are the locations that are candidates for model reduction. In Figure 10 they are depicted by the red.

We have now collected the necessary information to start the meta algorithm. We use the motifs parallel connections, sequential connections, and almost disconnected subgraphs.

We discuss exemplarily the first steps of the Algorithm 1 with respect to the motifs discussed in Subsection 4.2. According to the LogRank the set of vertices that are supposed to be reduced is given by

$$R_\Delta = \{v_6, v_{13}, v_{12}, v_{22}, v_{10}, v_4, v_1, v_{15}\}.$$

In the first step the vertex v_6 is considered. As this vertex is almost disconnected from the network, it is aggregated with vertex v_{10}, denoted v'_{10}. By Theorem 3 the update of the list is

$$R_\Delta = \{v_{13}, v_{12}, v_{22}, v_4, v_1, v'_{10}, v_{15}\}.$$

In the second iteration the vertex v_{13} is part of a parallel connection and hence aggregated with vertex v_{12}. The new vertex is denoted by v'_{12}. The corresponding update is

$$R_\Delta = \{v_{22}, v_4, v_1, v'_{12}, v'_{10}, v_{15}\}.$$

Further steps of reduction based on the LogRanks are:

- aggregation of the almost disconnected node v_{22} with the node v_{19}, denoted by v'_{19};
- aggregation of the almost disconnected node v_4 with the node v_9, denoted by v'_9;
- aggregation of the almost disconnected node v_1 with the node v_7, denoted by v'_7;
- aggregation of the almost disconnected node v'_{10} with the node v_{15}, denoted by v'_{15};
- aggregation of the node v'_{15} and the node v_{14} connected in parallel, denoted by v'_{14}.

The obtained network of smaller size is depicted in Figure 11. Note that although the vertex v'_{12} has still a LogRank smaller than .03 it cannot be reduced further since it is neither part of a sequential or parallel connection nor almost disconnected from the logistics network. However, if further motifs had been under consideration the procedure might continue.

As we have seen in this section our approach to structure-preserving model reduction can be easily adapted and applied to Jackson networks.

7 Conclusions

This paper provides a new heuristic approach for structure-preserving model reduction of large-scale networks with flows that is in particular suitable for logistics systems. For this purpose we consider such networks as graphs with weighted edges and develop the LogRank scheme to evaluate importance of vertices in this graph. Further, we provide a motif based meta algorithm for the reduction in size of the graph such that the main features are preserved. Here we focus on the material flow. We illustrate our methods with the help of examples. In particular, we consider the LogRank and model reduction for Jackson networks.

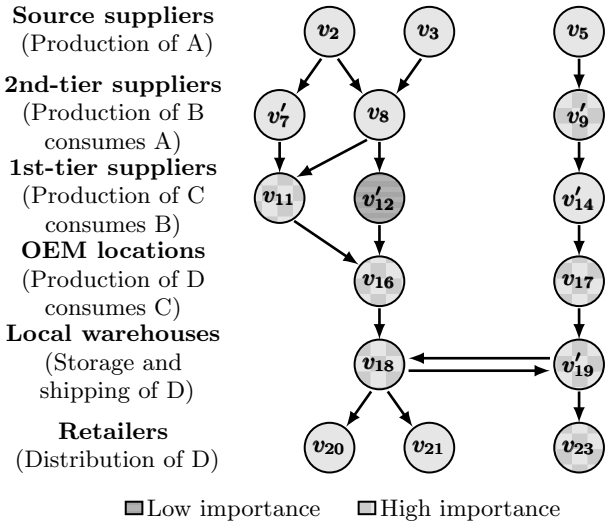

Fig. 11. The reduced network of the test example.

In future work we plan to analyze the relation between dynamical properties of the reduced and the original networks in more detail. In particular, we are going to study conditions under which stability properties are preserved by the LogRank guided approximation. The approach should be further developed to preserve stability properties. Also, estimates for the approximation error need to be investigated.

Acknowledgments

The authors would like to thank the anonymous reviewers whose insightful comments have lead to an improved presentation of the paper.

This work is part of the research project: *Stability, Robustness and Approximation of Dynamic Large-Scale Networks - Theory and Applications in Logistics Networks*, funded by the Volkswagen foundation.

References

1. D. Fleming, Y. Hayuth, Journal of Transport Geography **2**(1), 3 (1994)
2. R. Guimera, L. Amaral, The European Physical Journal B-Condensed Matter and Complex Systems **38**(2), 381 (2004)

3. B. Derudder, L. Devriendt, F. Witlox, Tijdschrift voor Economische en Sociale Geografie **98**(3), 307 (2007)
4. H. Wiendahl, S. Lutz, CIRP Annals-Manufacturing Technology **51**(2), 573 (2002)
5. T. Choi, K. Dooley, M. Rungtusanatham, Journal of Operations Management **19**(3), 351 (2001)
6. A. Surana, S. Kumara, M. Greaves, U. Raghavan, International Journal of Production Research **43**(20), 4235 (2005)
7. S. Pathak, J. Day, A. Nair, W. Sawaya, M. Kristal, Decision Sciences **38**(4), 547 (2007)
8. B. Beamon, International Journal of Production Economics **55**(3), 281 (1998)
9. J. Buzacott, J. Shanthikumar, *Stochastic Models of Manufacturing Systems* (Prentice Hall, Englewood Cliffs, NJ, 1993)
10. H. Chen, D.D. Yao, *Fundamentals of Queueing Networks. Performance, Asymptotics, and Optimization*, Vol. 46 of *Applications of Mathematics* (Springer, New York, NY, 2001)
11. P. Kumar, Queueing Systems **13**(1), 87 (1993)
12. J. Dai, Ann. Appl. Probab. **5**(1), 49 (1995)
13. D. Armbruster, P. Degond, C. Ringhofer, SIAM Journal on Applied Mathematics **66**(3), 896 (2006)
14. S. Göttlich, M. Herty, A. Klar, Communications in Mathematical Sciences **3**(4), 545 (2005)
15. D. Helbing, S. Lämmer, in *Networks of Interacting Machines: Production Organization in Complex Industrial Systems and Biological Cells*, edited by D. Armbruster, K. Kaneko, A. Mikhailov (World Scientific, Singapore, 2005), Vol. 3 of *Lecture Notes in Complex Systems*, pp. 33–66
16. M. Laumanns, E. Lefeber, Physica A **363**, 24 (2006)
17. Y. Ouyang, C. Daganzo, Management science **52**(10), 1544 (2006)
18. B. Scholz-Reiter, F. Wirth, S. Dashkovskiy, M. Kosmykov, T. Makuschewitz, M. Schönlein, *Some Remarks on Stability and Robustness of Production Networks Based on Fluid Models*, in *Dynamics in Logistics. Second International Conference, LDIC 2009*, edited by H.J. Kreowski, B. Scholz-Reiter, K.D. Thoben (Springer, Berlin, Heidelberg, 2009), pp. 27–36
19. B. Scholz-Reiter, F. Wirth, T. Makuschewitz, M. Schönlein, CIRP Annals-Manufacturing Technology **60**, 445 (2011)
20. B. Moore, IEEE Transactions on Automatic Control **26**(1), 17 (1981)
21. A.C. Antoulas, *Approximation of Large-Scale Dynamical Systems* (Advances in Design and Control 6. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2005)
22. P. Benner, V. Mehrmann, D.C. Sorensen, eds., *Dimension Reduction of Large-Scale Systems*, Vol. 45 of *Lecture Notes in Computational Science and Engineering* (Springer, Berlin Heidelberg New York, 2005), proceedings of a Workshop held in Oberwolfach, Germany, October 19-25, 2003
23. Z. Bai, Applied Numerical Mathematics **43**(1-2), 9 (2002)
24. K. Kunisch, S. Volkwein, Numerische Mathematik **90**, 117 (2001)
25. S. Lall, J. Marsden, S. Glavaški, Internat. J. Robust Nonlinear Control **12**(6), 519 (2002)
26. N. Kazantzis, C. Kravaris, L. Syrou, Nonlinear Dynam. **59**(1-2), 183 (2010)
27. W. Schilders, H.A. van der Vorst, J. Rommes, eds., *Model Order Reduction: Theory, Research Aspects and Applications*, Vol. 13 of *Mathematics in Industry* (Springer, Berlin Heidelberg, 2008)
28. L. Page, S. Brin, R. Motwani, T. Winograd, Tech. rep., Technical Report, Stanford University (1998)
29. J.M. Kleinberg, J. ACM **46**(5), 604 (1999), ISSN 0004-5411
30. R. Lempel, S. Moran, Computer Networks **33**(1-6), 387 (2000)
31. A.N. Langville, C.D. Meyer, *Google's PageRank and Beyond: The Science of Search Engine Rankings* (Princton University Press, Princeton, NJ, 2006)
32. D. Donato, L. Laura, S. Leonardi, S. Millozzi, The European Physical Journal B-Condensed Matter and Complex Systems **38**(2), 239 (2004)
33. T.S. Evans, R. Lambiotte, The European Physical Journal B - Condensed Matter and Complex Systems **77**, 265 (2010)
34. S. Fortunato, Physics Reports **486**(3-5), 75 (2010)
35. A. Berman, R.J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Vol. 9 of *Classics in Applied Mathematics* (SIAM, Philadelphia, PA, 1994)
36. J. Bang-Jensen, G. Gutin, *Digraphs: Theory, Algorithms and Applications*, Monographs In Mathematics (Springer, 2009)
37. L. Krajewski, B. King, L. Ritzman, D. Wong, Management Science **33**(1), 39 (1987)
38. S. Serra-Capizzano, SIAM. J. Matrix Anal. Appl. **27**(2), 305 (2005)
39. R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, U. Alon, Science **298**, 824 (2002)
40. A. Altman, M. Tennenholtz, *Ranking Systems: The PageRank Axioms*, in *Proc. 6th ACM Conference on Electronic Commerce* (2005), pp. 1–8
41. A. Altman, M. Tennenholtz, J. Artif. Int. Res. **31**(1), 473 (2008)