

An approach to model reduction of logistic networks based on ranking

Bernd Scholz-Reiter¹, Fabian Wirth², Sergey Dashkovskiy³, Michael Kosmykov³, Thomas Makuschewitz¹, Michael Schönlein²

¹ Planning and Control of Production Systems (PSPS), BIBA – Bremer Institut für Produktion und Logistik GmbH at the University of Bremen, Bremen, Germany

² Institute of Mathematics, University of Würzburg, Würzburg, Germany

³ Centre of Industrial Mathematics, University of Bremen, Bremen, Germany

Abstract Simulations or mathematical analysis of a real-world logistic network require a model. In this context two challenges occur for modelling: First, the model should represent the real-world logistic network in an accurate way. Second, it should foster simulations or analytical analysis to be conducted in a reasonable time. A large size is often a drawback of many models. In the case of logistic networks this drawback can be overcome by reducing the number of locations and transportation links of the graph model. In this paper we present an approach to model reduction of a logistic network based on ranking. The rank of a given location states the importance of the location for the whole network. In order to calculate the importance of a location we introduce an adaptation of the PageRank algorithm for logistic networks. The information about the rank and the structural relations between the locations are used for our approach to model reduction. Depending on the structural relation between locations we suggest three different approaches to obtain a model of lower size.

1 Introduction

Performance and competitiveness of a large-scale logistic network depend on the capability of the network to meet the expectations of the customers [1]. This capability is strongly connected to an effective management of the material flow within the network. The material flow is subject to the complex and often global structure of the network as well as to the dynamics of production and transportation processes [2]. In order to support the management, a better understanding of the dynamics related to the material flow and their consequences for the performance of the logistic network is required. In the literature several methods exist to analyse the material flow.

Three different methods can be utilized for the investigation of the material flow of a logistic network [3]. First, the material flow of the real-world network can be measured. Second, simulations can be carried out in order to analyse changes in the structure and dynamics of the logistic network for different scenarios. Third, mathematical methods can be applied in order to obtain a more precise understanding of the involved processes. Both, simulation and mathematical methods require usually a model of the real-world logistic network.

Model development faces two major challenges. First, the model should exhibit almost the same properties compared to the real-world logistic network and second, it should be tailored to the applied methods in order to enable consolidated findings. The size of a model is often crucial for a successful application of a certain analysis method. A model of lower size facilitates simulations or the application of mathematical methods. Since, logistic networks often consist of a large number of locations and transportation connections between them a representative model of lower size is desired.

Our approach to model reduction is based on a ranking scheme of the locations that takes the material flows within the network and the structure of the connections between the locations into account. For this purpose the PageRank algorithm [4], which has been a core component of Google Internet search engine in its early days, is used. The original ranking algorithm is extended by the results of a material flow analysis. A material flow analysis provides valuable information about the importance of the connections between the locations by analysing the quantities of material flow between the locations [5]. These quantities can be incorporated into the ranking algorithm in order to enhance the ranking. The adapted ranking algorithm provides in terms of its application to a logistic network a reasonable ranking [6], [7]. In order to derive a model of lower size we propose to focus on locations with a low importance for the network. According to their connections to other locations of the network we investigate three different approaches for model reduction. These approaches involve the exclusion and aggregation of individual locations as well as the exclusion of subparts of the network. The paper shows that by applying these approaches for model reduction representative models can be derived.

The outline of the paper is as follows. In section 2 our proposed adaptation of the PageRank algorithm to logistic networks is presented. Three different approaches to model reduction based on the structural properties of the locations are introduced in section 3 and illustrated by examples. Section 4 summarise the findings of this paper and provides an outlook to future research.

2 Adaptation of the PageRank algorithm to logistic networks

Before we introduce a notion of importance rank of a logistic location in a network we describe the network itself as a model.

We model a logistic network as a directed graph in the following way. Let the logistic locations be numbered by $1, \dots, n$ and each of them be a node of this graph. There is an edge from node i to node j if there is a material, information or monetary flow from the i th to j th location. In our approach only the aggregated quantity of a material flow between the locations over a certain period of time is considered. Let a_{ij} be a number quantifying this flow. In particular, if there is no flow from location i to location j then $a_{ij} = 0$. The matrix $\mathbf{A} = (a_{ij})_{i,j=1,\dots,n}$ describes the interconnection structure and the flows of a given network and is called weighted adjacency matrix of the graph. For example the network in Figure 1 has the following weighted adjacency matrix:

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 6 & 0 \\ 7 & 0 & 0 & 8 & 0 & 0 \\ 4 & 5 & 0 & 0 & 7 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 11 & 0 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

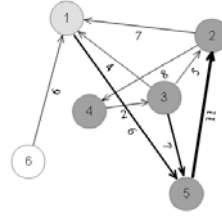


Fig.1. Weighted directed graph of a logistic network. The numbers in the circles are the numbers of nodes and the numbers near the edges are their weights.

The following types of matrices will be used in this paper:

A matrix \mathbf{A} is called *column-normalized* if for all $i=1, \dots, n$

$$\sum_{j=1}^n a_{ij} = \begin{cases} 1, & \text{if there exists } j \text{ such that } a_{ij} \neq 0 \\ 0, & \text{if for all } j \ a_{ij} = 0 \end{cases}$$

A matrix \mathbf{A} is *column-stochastic* if $\sum_{j=1}^n a_{ij} = 1$ for all $i=1, \dots, n$.

It is called *primitive* if there exists a positive integer k such that the matrix \mathbf{A}^k has only positive elements.

We call an adjacency matrix \mathbf{A} *irreducible* if the corresponding graph is strongly connected, i.e., for every nodes i and j of the graph there exists a sequence of directed edges connecting i to j . Note that any primitive matrix is irreducible.

Now we are ready to introduce the notion of importance of logistic locations in a network. In the sequel we call it rank of a node or location. We say that the rank of a certain location depends on the network structure, its position within the network and flows in this network. As the weighted adjacency matrix of a network contains the information about its structure and flows between its nodes we use it to define the importance of the nodes. To define the rank of logistic locations we use the idea of the PageRank [5], which was originally designed for the ranking web pages in the Internet. This idea and the algorithm for its calculation can be adapted in the following way to logistic networks. We say that the rank of a loca-

tion i depends on the flows a_{ij} , from this node to other locations $j=1, \dots, n$ as well as ranks of these locations. The more important locations receive material from a given location the more important it is. In comparison with the original PageRank locations do not share their rank equally between their suppliers but rather proportionally to the flows from their suppliers described by a_{ij} . I.e., the proportions \tilde{a}_{ij}

are $\tilde{a}_{ij} = \frac{a_{ij}}{\sum_{k \in I_j^d} a_{kj}}$. Such proportions were proposed in [9] as a modification of

PageRank. Thus the rank NR_i of the node i should be calculated by

$$NR_i = \sum_{j=1}^n \tilde{a}_{ij} NR_j, \quad i=1, \dots, n. \quad (1)$$

However we introduce a parameter α , $0 < \alpha \leq 1$ and for a given graph with n nodes we define the rank NR_i of a location i by

$$NR_i = \alpha \sum_{j=1}^n \tilde{a}_{ij} NR_j + (1-\alpha) \frac{1}{n}, \quad i=1, \dots, n, \quad (2)$$

i.e., the rank of location i is the sum of the proportion of ranks that locations j contribute and a small positive term, where n is usually a large number. Note that for $\alpha=1$ equation (2) coincides with (1). Thus the value $1-\alpha$, is usually taken close to zero in order to preserve the information about the real structure of the network. It is interpreted as a probability that a location supplies material to locations with which it has no direct link (partnership). This positive term is important for the existence and uniqueness of a solution of (2).

Now we see that the ranks NR_i can be calculated by solving the linear system of algebraic equations (2). However this problem can be nontrivial in the case of large number of nodes n . To solve equation (2) we transform it into an eigenvalue problem that can be solved numerically [7]. The needed transformation steps are:

1. Column-normalization:

$$H = (H_{ij}), H_{ij} = \frac{a_{ij}}{\sum_{k=1}^n a_{kj}} = \tilde{a}_{ij}, \text{ if } \sum_{k=1}^n a_{kj} \neq 0 \text{ and } H_{ij} = 0, \text{ otherwise.} \quad (3)$$

2. Make the matrix stochastic:

$$S = H + \frac{eb^T}{n}, e = (1, \dots, 1)^T \in R^n, b_i = 1, \text{ if } \sum_{j=1}^n H_{ji} = 0 \text{ and } b_i = 0, \text{ otherwise.} \quad (4)$$

3. Making the matrix primitive: $G = \alpha S + (1-\alpha)E$, $E = \frac{ee^T}{n}$ (5)

Let $p = (p_1, \dots, p_n)^T$ be a normalized and nonnegative vector. Multiplying the right side of matrix G by vector p we obtain the right-hand side of equation (2) with $NR_j = p_j$ and the problem (2) is equivalent to the problem of finding the right-

eigenvector $p=Gp$ of G that corresponds to the eigenvalue 1. Since the matrix G is stochastic and irreducible the Perron-Frobenius theorem guarantees that 1 is an eigenvalue of G and all other eigenvalues have absolute value less than 1, i.e., 1 is the spectral radius of G [8]. Since G is primitive and hence irreducible from the Perron-Frobenius theorem in [8] it follows that the corresponding eigenvector is unique up to a scalar factor.

Applying our method to the network given in Figure 1 with $\alpha=0.9$ we obtain the following ranks which are given in Table 1. In parentheses we number the locations by their importance.

The group of the most important locations is highlighted in dark grey. Light grey is used for the group of locations with average importance and white is used for the background of the group of the least important locations. We see that in this example the most important location is location 2. This can be explained by the reason that it delivers large amounts of material to other locations. Location 6 delivers material to only one location with an average rank and therefore has the lowest rank and hence it is the least important location.

Location	Rank
1	0.1040 (5)
2	0.2455 (1)
3	0.2016 (3)
4	0.2056 (2)
5	0.1778 (4)
6	0.0655 (6)

Table 1. Ranks of the original network in Fig.1.

3 Approaches for network model reduction

For the analysis of large logistic networks the reduction of the size of their model is often needed. Our main idea for reduction of a model is to exclude or aggregate the nodes of the lowest ranks, i.e., the less important locations. We propose three different rules how to do this. These rules are chosen in a way to conserve the main structure of the network. In the following we describe these rules and show their implementation on some simple examples. In these examples we will observe the changes of the ranks and especially of the order of the nodes by their ranks.

1. Exclusion of low-ranked locations connected to only one location

Locations with low rank connected to only one location do not describe the major structure of the network. Thus such locations are excluded from the graph as well as their links.

Example 1: Consider the location 6 in Figure 1 that is the least important location and is connected only to the location 1. Applying the given rule it is excluded from the graph. The reduced model is shown in the Figure 2. The ranks of loca-

tions of the reduced graph are given in Table 2. Note that the ranks of the locations in the reduced model have the same order as in the original network.

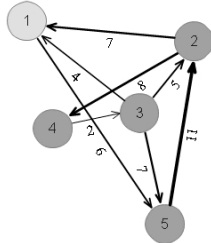


Fig. 2. Reduced network

Location	Rank
1	0.1039 (5)
2	0.2707 (1)
3	0.2202 (2)
4	0.2171 (3)
5	0.1882 (4)

Table 2. Ranks of the reduced network.

2. Aggregation of low-ranked locations

Locations with low ranks appearing in a parallel or sequential connection in the network are aggregated and considered as one location. The weights of incoming and outgoing links of the aggregated location are defined as the sums of the corresponding weights of the former individual links. In this case the information about the original material flow through these locations is kept in the reduced network. In [10] a theoretical analysis of similar rules is presented but it was applied to the original PageRank algorithm without weighted edges of the considered graph.

To illustrate this rule consider the following two examples.

Example 2 (parallel connection): Figure 3a illustrates a logistic network, which consists of only one customer (location 1), who orders products from four OEM's (locations 2-5). The OEM's are connected to one supplier (location 6).

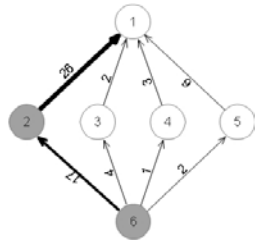


Fig.3a. Original network.

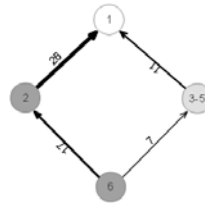


Fig.3b. Reduced network.

Locations 3, 4 and 5 have low ranks, see Table 3a, and are connected in a parallel way to the same locations 1 and 6. Hence, they are aggregated to one location called '3-5' and a reduced model with 4 locations is obtained, see Figure 3b. The new incoming link to '3-5' corresponds to the sum of the former three individual incoming links. The same applies to the outgoing links. Note that the order of the locations by their ranks in the reduced model remains unchanged, cf. the Tables 3a and 3b.

Location	Rank
1	0.0911 (6)
2	0.1456 (2)
3	0.0953 (5)
4	0.0974 (4)
5	0.1037 (3)
6	0.4668 (1)

Table 3a. Ranks of the original network.

Location	Rank
1	0.1375 (4)
2	0.2196 (2)
3-5	0.1723 (3)
6	0.4706 (1)

Table 3b. Ranks of the reduced network.

Example 3 (sequential connection): In the network in Figure 4a the locations 2 and 4 have low ranks and they are connected sequentially. Thus they are aggregated to one location that is denoted by '2,4'. The weight of the incoming link is the weight of the former incoming link of location 4 and the weight of the outgoing link is the weight of the former outgoing link of location 2. Table 4b shows that the ranks in the reduced graph have the same order as in the original graph.

We see that in the last three examples the second rule of model reduction does not change the order of importance between the unchanged locations.

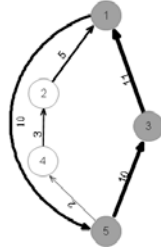


Fig. 4a. Original network.

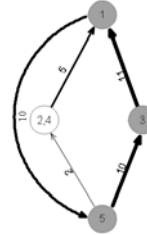


Fig. 4b. Reduced network.

Location	Rank
1	0.2828 (2)
2	0.1051 (5)
3	0.1952 (3)
4	0.1194 (4)
5	0.2975 (1)

Table 4a. Ranks of the original network.

Location	Rank
1	0.3202 (2)
2,4	0.1226 (4)
3	0.2247 (3)
5	0.3326 (1)

Table 4b. Ranks of the reduced network.

3. Exclusion of sub-networks of low-ranked locations connected to only one important location

A sub-network of locations that have low ranks is excluded from the graph if it is connected to one important location only. The material flows between the remaining locations are kept unchanged and therefore the main structure of the original network is preserved.

We illustrate this rule by the following example.

Example 4: The logistic network in Figure 5 contains three locations with high ranks. There are three groups of low rank locations. Each of them has connections to one of the high rank nodes only. By the third rule all three groups are excluded

to obtain the reduced graph, see Figure 6, which consists of 3 locations. The order of the three most important locations is again unchanged, cf. Table 6.

Note that each of the rules defined above can be applied simultaneously to different nodes. Thus the network can be significantly reduced to the desired size. The next example illustrates the application of all three rules to one network.

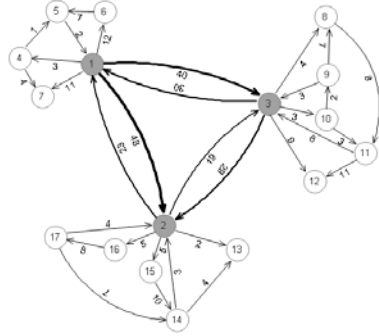


Fig.5. Network with 3 sub-networks.

Location	Rank	Location	Rank
1	0.2946 (01)	10	0.0319 (04)
2	0.2189 (03)	11	0.0165 (11)
3	0.2525 (02)	12	0.0088 (15)
4	0.0123 (14)	13	0.0088 (15)
5	0.0135 (13)	14	0.0162 (12)
6	0.0188 (08)	15	0.0169 (09)
7	0.0088 (15)	16	0.0231 (05)
8	0.0190 (07)	17	0.0168 (10)
9	0.0226 (06)	-	-

Table 5. Ranks of the original network.

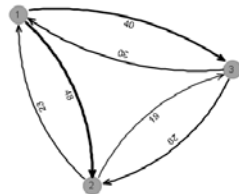


Fig. 6. Reduced network.

Location	Rank
1	0.3890 (1)
2	0.2832 (3)
3	0.3278 (2)

Table 6. Ranks of the reduced network.

Example 5: The graph in Figure 7 consists of 33 locations. Their ranks are shown in Table 7. The most important locations are locations 2, 5 and 14. Locations 1, 3, 6, 8, 27, 28 will be treated as average rank nodes. The rest of the nodes are identified as low rank locations.

Locations 22, 29, 30, 31, 32, 33 are deleted by applying the rule from Section 3.1. The locations 4, 7, 12 are aggregated using the rule of Section 3.2 for sequential connection and locations 9-11 are aggregated using the rule for parallel connections from Section 3.2. Sub-network of nodes 13, 15-21 and sub-network of nodes 23-26 are excluded according to the rule from Section 3.3. The reduced model consists of 11 locations and is shown in Figure 8. The corresponding ranks of the locations are given in Table 8. Locations 2 and 5 are still the most important ones. However the rank of location 14 is of the order of the average rank nodes.

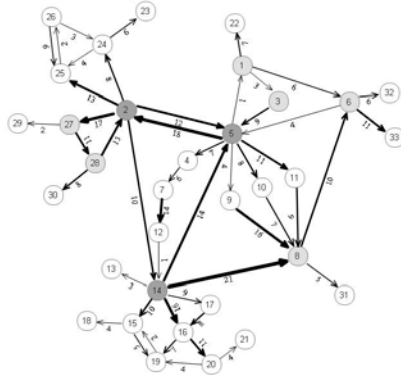


Fig.7. Complex network with 33 locations.

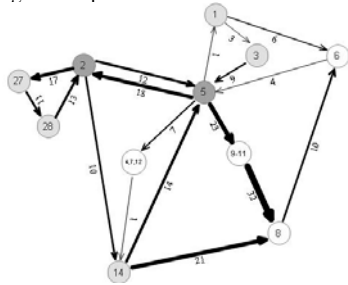


Fig.8. Reduced network with 11 locations.

Location	Rank	Location	Rank
1	0.0567 (04)	18	0.0066 (23)
2	0.1847 (02)	19	0.0087 (22)
3	0.0401 (06)	20	0.0140 (17)
4	0.0230 (10)	21	0.0066 (23)
5	0.1969 (01)	22	0.0066 (23)
6	0.0327 (08)	23	0.0066 (23)
7	0.0193 (12)	24	0.0122 (19)
8	0.0296 (09)	25	0.0189 (13)
9	0.0142 (16)	26	0.0145 (15)
10	0.0099 (21)	27	0.0416 (05)
11	0.0066 (23)	28	0.0346 (07)
12	0.0149 (14)	29	0.0066 (23)
13	0.0066 (23)	30	0.0066 (23)
14	0.1077 (03)	31	0.0109 (20)
15	0.0145 (15)	32	0.0066 (23)
16	0.0218 (11)	33	0.0066 (23)
17	0.0128 (18)	-	-

Table 7. Ranks of the original network.

Location	Rank	Location	Rank
1	0.0762 (06)	14	0.0975 (03)
2	0.2262 (02)	8	0.0320 (09)
3	0.0607 (07)	9-11	0.0301 (10)
4,7,12	0.0212 (11)	27	0.0937 (05)
5	0.2336 (01)	28	0.0943 (04)
6	0.0345 (08)	-	-

Table 8. Ranks of the reduced network.

5 Conclusions and future research

We have proposed three rules for the reduction of the size of the model of a given large-scale logistic network. For this purpose a method for identifying the nodes to be excluded or aggregated was developed that takes the structure of the network and material flows between its locations into account. This method extends the idea of the well known Page Rank algorithm. Examples given in this paper illustrate how this reduction rules can be applied to obtain reasonable models of lower size to describe large logistic networks.

The observed numerical results raise theoretical questions related to the stability of the order of ranks with respect to the application of the proposed reduction

rules. These questions have to be investigated analytically and are a matter of the further research. The development of further reduction rules and analysis of their properties are also planned for the future work.

Acknowledgements. The authors gratefully acknowledge support by Volkswagen Foundation (Project Nr. I/82684 "Stability, Robustness and Approximation of Dynamic Large-Scale Networks — Theory and Applications in Logistics Networks"). S. Dashkovskiy is partially supported by the DFG as a part of Collaborative Research Center 637 "Autonomous Cooperating Logistic Processes - A Paradigm Shift and its Limitations".

References

- [1] Vahrenkamp, R. (2007): Logistik: Management und Strategien. 5. Auflage, Oldenbourg, München.
- [2] Scholz-Reiter, B., Wirth, F., Dashkovskiy, S., Jagalski, T., Makuschewitz, T. (2008): Analyse der Dynamik großskaliger Netzwerke in der Logistik. In: Industrie Management, 3, 37-40.
- [3] Tutsch, D. (2006): Performance Analysis of Network Architectures. Springer, Berlin Heidelberg New York.
- [4] Page, L., Brin, S., Motwani, R., Winograd, T. (1999): The PageRank citation ranking: Bringing order to the web. Technical report, Stanford University.
- [5] Arnold, D., Furmans, K. (2007): Materialfluss in Logistiksystemen. 5. Auflage, Springer, Berlin.
- [6] Scholz-Reiter, B., Wirth, F., Dashkovskiy, S., Makuschewitz, T., Kosmykov, M., Schönlein, M. (2009): Application of the PageRank algorithm for ranking locations of a production network. Proceedings of 42nd CIRP Conference on Manufacturing Systems.
- [7] Langville, A. N., Meyer, C. D. (2006): Google's PageRank and Beyond: The Science of Search Engine Rankings, Princeton University Press, Princeton, NJ.
- [8] Meyer, C. D. (2000): Matrix Analysis and Applied Linear Algebra. SIAM, Philadelphia.
- [9] Baeza-Yates, R., Davis, E. (2004): Web Page Ranking Using Link Attributes. Archive Proceedings of the 13th international World Wide Web conference, Alternate Track Papers & Posters, New York, NY, USA, 328-329.
- [10] Altman, A., Tennenholtz, M. (2005): Ranking systems: the PageRank axioms, Proceedings of the 6th ACM conference on Electronic commerce, Vancouver, BC, Canada, 1-8.