

Finding the Maximal Day-Time Dependent Component of a Subway System

Marian Sorin Nistor¹(✉), Doina Bein², Horia Nicolai Teodorescu^{3,4},
and Stefan Wolfgang Pickl¹

¹ Department of Computer Science, Universität der Bundeswehr München,
85577 Neubiberg, Germany

{sorin.nistor, stefan.pickl}@unibw.de

² Department of Computer Science, California State University,
Fullerton, CA 92831, USA
dbein@fullerton.edu

³ Institute of Computer Science of the Romanian Academy, Iași, Romania

⁴ ETTI Department, “Gheorghe Asachi” Technical University of Iași,
700050 Iași, Romania
hteodor@etti.tuiasi.ro

Abstract. We analyze the properties of a transportation network with time-dependent flows with the aim of finding the maximal component along the duration of a period of the variation of the flow. We apply this analysis method to the subway network of Munich. The analysis is detailed and results are included. Beyond the theoretical interest in this type of network with periodically varying properties, the paper presents a well-structured analysis method that can be widely applied to numerous other networks of applicative interest.

Keywords: Subway network · Network analysis · Maximal component · Longest shortest path

1 Introduction

The subway systems in general are critical urban infrastructures, exposed to frequent disruptions. Various factors contribute towards serious schedule delays, including having the entire system shut down. A temporary shutdown of a means of public transport is relatively rare, but has a serious impact over the activity of a city. Moreover, if this unfortunate event was not due to natural hazards or temporary failures of the system, but purposely targeted attacks, it poses then security issues.

In this paper, we analyze the scenario where a threat propagates through the network using the subway system. For this, we tackle two research questions: 1. What is the longest ride one can have in a threat propagation, such as a flee attempt using the subway system within a limited travel time and at a specific time of the day? 2. What is the maximal number of stations one can reach in a flee attempt using the subway system within a limited travel time and at a specific time of the day?

We conducted the analysis on a subway system modeled as a mathematical network where the stations are nodes and the connection between pair of stations are edges [1].

The train frequency at each station, the corresponding subway line(s) of each station, and the travel time between adjacent stations are encoded as weights on edges. Hence, the research questions can be redefined as:

1. What is the longest day-dependent and time-dependent (i.e., day- and time dependent) shortest path of a subway network. (Recall that the shortest path is the path between two nodes in a network having the smallest edge-weight sum.)
2. What is the maximal day-time dependent sub-network (component) of a subway network? (Recall that the maximal component is a connected sub-network that starting from a root node of the network covers a range of nodes that satisfies a given maximal edge-weight sum.)

The first question is a particular application of the *weighted shortest path problem* [2] that can be solved efficiently using Dijkstra's algorithm [3]. The second question is a particular application of the *constrained maximum-weight connected graph problem* [4]. The latter is the focus of this paper. Here, the first problem is treated as a sub-problem of the second. Namely, a rooted day-time dependent component is a collection of the corresponding shortest paths starting from a root node to each node of a maximal edge-weight range of nodes. We consider the maximal day-time dependent component (MDTDC) to be the main problem, and the longest day-time dependent shortest path (LDTDSP) to be the sub-problem.

Various optimized solutions to this problem are available in [5–7]. The application range is very diversified and similar analyzes are applied to biological systems [8–10], wildlife corridors design [11, 12], forest planning [13], communication network design [4], computer vision [14], and few others.

Our main contribution is a well-defined analysis method that we apply to subway networks. The subway networks are special because of the variation of the edge weight value. Thus, we consider the weights to have two sets of values, depending on which case they fall into. One case is when two adjacent nodes are part of the same subway line; in that case the weight value on the edge represents the travel time between nodes. In the second case, when two adjacent nodes are not part of the same subway line, then the weight value on the edge represents the sum of waiting times at the starting node and the travel time between nodes.

A detailed description of the conducted analysis is given in Sect. 2. The Munich subway network proves to be a good candidate [15, 16] for our analysis due to its size, ridership [17], and train frequency variation over train time schedules [18]. More details on the data collection is given in Sect. 3. The results are presented and discussed in Sect. 4. The paper concludes with considerations on the application on this network and remarks for other subway networks.

2 The MDTDC Problem with the LDTDSP Sub-problem

2.1 Formulation

Let $G = (V, E)$ be a graph and $v_i \in V$ a random starting node. Knowing all edges e_{ij} of that node, the time of the day (or the hour of the schedule) t_D , the corresponding

frequencies $f(t_D, e_{ij})$, and their travel times along the edges $d(v_i, v_j)$, we define the component G_c of G of all nodes that can be reached in time t starting at $t = 0$ from v_i .

All the vertices v_h in G_c satisfy the condition that there is a shortest path from v_i to v_h such that it can be traveled along in time less than t :

$$P(v_i, v_h) = \min_{P(v_i, v_h)} \left(\sum_{e_{ij} \in P(v_i, v_h)} \alpha \frac{60}{f(t_D, e_{ij})} + d(v_i, v_j) \right) \leq t \tag{1}$$

where $P(v_i, v_h)$ denotes the shortest paths from v_i to v_h , $\alpha = 0$ if v_i and v_j belong to the same subway line, else $\alpha = 1$.

From (1), for a given t_D and t , MDTDC is defined as $\sum P(v_i, v_h)$, and LDTDSP is defined as $P_{max}(v_i, v_h)$. In Fig. 1 (left) is an example of the MDTDC on a simple network. In this example, the MDTDC of the network starts from the red (center) node, and all the orange nodes represent the visiting range for a given maximal edge-weight of three. In Fig. 1 (right) is an example of the LDTDSP on the same simple network. Here, the LDTDSP of the network starts from the red (center) node, and all the orange nodes represent the visiting path for a given maximal edge-weight of three.

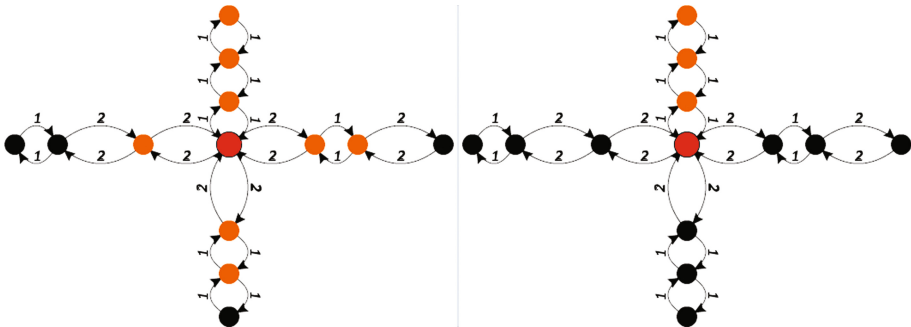


Fig. 1. A simplified version of the application for a given maximal edge-weight of three without considering subway lines or train frequencies. MDTDC is exemplified on a simple network to the left. LDTDSP is exemplified on a simple network to the right.

2.2 Algorithm

Input: An input file containing the adjacency list of the subway network for one schedule with multiple weights, and an array of the travel time limits. The input file has the following information on each column: col_1 = station from, col_2 = station to; col_3 to col_26 = train frequency on that edge by hour from 00 h to 23 h (24 columns); col_27 = subway line; col_28 = travel time on edge in min. The array has the following elements: 5 min, 10 min, 15 min, 20 min.

Output: The results for MDTDC and LDTDSP for each travel time limit.

```

Begin
read master_list
read travel_limit
for each hour H do
  extract adjacency_list_H from master_list
  convert adjacency_list_H to graph_H
  for each node(graph_H) N do
    calculate visited_nodes with
      breadth_first_search(graph_H, from_node=N)
    order visited_nodes
    for each node(visited_nodes) V do
      calculate crnt_sp_V with
        shortest_path(graph_H, from_node=N, to_node=V)
      calculate crnt_travel_time for crnt_sp_V
      save crnt_travel_time to travel_time_nodes_V
      save crnt_sp_V to list_sp_V
    {end for}
    save list_sp_V & travel_time_nodes_V to LDTDSP_N
    save travel_time_nodes_V to travel_time_N
  {end for}
  calculate MDTDC_H with sum(LDTDSP_N)
  calculate LDTDSP_H with max(LDTDSP_N)
{end for}
for each travel_limit T do
  for each node(graph_H) N do
    save MDTDC_T from MDTDC_H with
      travel_time_N < travel_limit_T
    save LDTDSP_T from LDTDSP_H with
      travel_time_N < travel_limit_T
  {end for}
{end for}

```

3 Data Collection

The analysis described in the previous section was applied to the Munich subway network. The network ridership in average is above typical, with more than 398 million passengers per year [17] for a network with only 100 stations [17]. The train frequency of this network varies along the subway lines, and sometimes is different along the same line but in the opposite direction. Namely, the number of trains of a subway line entering a station in a day is not the same from both directions, because sometimes trains are redirected on other lines sharing the same destination station. This makes the Munich subway network a very good candidate for our analysis.

For this paper, the train frequency data for year 2017 was collected for every station of the network. This data is publicly available and it can be accessed from the website of the Munich's association of public transport authorities (MVV) [18].

The subway network has four main schedules, corresponding to *Monday-Thursday*, *Friday*, *Saturday*, and *Sunday/Holiday*. The latter one is the most relaxed schedule in terms of train frequency, while *Friday* schedule is the busiest. MVV provides few side schedules with a more relaxed train frequency during the children's vacation, or more intense during the carnival days, Christmas Eve and New Year's Eve [18]. In this paper, only the regular schedules are considered. This network has eight subway lines (U1–U8), but only six (U1–U6) have a daily schedule. The other two (U7–U8) reuse stations of other lines, with a limited train frequency, because these lines were created only to avoid overcrowding the central stations by supplementing the number of trains, e.g. during peak hours. Therefore, the train frequency of U7–U8 is assimilated by the lines having a regular schedule at the stations reused by these lines [19]. E.g. the line U7 only reuses stations of U1, U2, and U5, therefore in our analysis we do not treat this particularity, and the train frequency of U7 is attributed to the other three lines whose stations have been used.

For each edge, the train frequency starting or ending in one of the two corresponding stations is saved for each hour of the day. Along with the frequency on each edge, the travel time and the subway line are also stored. The train number handled by a station in an hour varies between 0–30 trains. The travel time between every two adjacent stations varies between 1–4 min. The lines are numbered between 1–6 for the stations corresponding to one line and 9 for shared stations by multiple subway lines. Figure 2 presents an example of cleaned input data for the analysis.

From station	To station	00:00	01:00	...	15:00	16:00	...	22:00	23:00	Line	Time
...
Maillingerstraße	Stiglmaierplatz	8	2	...	16	18	...	12	12	1	1
Moosfeld	Messestadt West	4	2	...	10	12	...	6	6	2	2
Scheidplatz	Bonner Platz	5	1	...	8	12	...	6	6	3	1
Prinzregentenplatz	Max-Weber-Platz	3	1	...	10	12	...	6	6	4	2
Neuperlach Zentrum	Therese-Giehse-Allee	5	1	...	12	12	...	6	6	5	2
Westpark	Partnachplatz	4	1	...	9	12	...	6	6	6	1
Münchner Freiheit	Giselastraße	10	2	...	23	24	...	12	12	9	1
...

Fig. 2. Example of cleaned input data for the analysis of the Munich subway network.

4 Results and Discussion

In this section, we present and discuss the results obtained after the data set was cleaned and we applied the analysis method from Sect. 2. MDTDC and LDTDSP are computed for the Munich subway network during the week days and weekend days. For the week days, two related schedules are analyzed: *Monday-Thursday* and *Friday*. For the weekend days, the other two schedules are analyzed: *Saturday* and *Sunday/Holiday*.

For our analysis, MDTDC and LDTDSP are calculated for a maximal given travel time corresponding to 5 min, 10 min, 15 min, and 20 min, in order to have a more diversified range of travel windows. For MDTDC, the results below 5 min travel time limit are flat, while for a travel time limit above 20 min, MDTDC covers more than half of the stations.

In Tables 1 and 2, the results show a linear growth of MDTDC from one travel time limit to another. One can observe that as the travel time limit increases, MDTDC varies from one hour to another. The results in Table 1 are more diversified than in Table 2 because the week days have a higher train frequency variation on the subway lines. For example, taking the 20 min travel limit, the difference between 4 am and 7 am on a *Monday-Thursday* schedule is more than double, while on a *Sunday/Holiday* schedule the difference is very small.

Table 1. MDTDC of the Munich subway network during week days, *Monday-Thursday* (M-T) and *Friday* (Fr) schedules, for a maximal given travel time of 5/10/15/20 min

Hour	5 min M-T	5 min Fr	10 min MT	10 min Fr	15 min MT	15 min Fr	20 min MT	20 min Fr
00:00	12	12	22	22	27	27	34	34
01:00	12	12	21	21	25	25	27	27
02:00	0	12	0	21	0	25	0	27
03:00	0	0	0	0	0	0	0	0
04:00	12	12	18	21	21	25	23	27
05:00	12	12	22	22	27	27	34	34
06:00	12	12	23	23	31	31	41	41
07:00	12	12	25	25	38	38	47	47
08:00	12	12	25	25	38	38	47	47
09:00	12	12	24	24	35	35	44	44
10:00	12	12	24	24	33	33	42	42
11:00	12	12	24	24	33	33	42	42
12:00	12	12	24	24	33	34	42	44
13:00	12	12	24	25	33	38	42	46
14:00	12	12	24	25	33	38	42	46
15:00	12	12	24	25	34	38	46	46
16:00	12	12	25	25	38	38	47	46
17:00	12	12	25	25	38	38	47	46
18:00	12	12	25	25	38	38	47	46
19:00	12	12	25	25	35	35	45	45
20:00	12	12	22	22	32	32	39	39
21:00	12	12	22	22	30	30	39	39
22:00	12	12	22	22	30	30	39	39
23:00	12	12	22	22	30	30	39	39

Notice that MDTDC formed in a 5 min travel time window is quite large for all train schedules except 2 am to 4 am. That is, if one finds the starting station that generated the size of this component, in only 5 min one can reach at any time of the week a range of up to 12 stations. This represents 12% of the size of the network. Furthermore, in up to 20 min time, one can reach up to 47 stations during the peak hours of the week days, which is almost half the size of the network.

This variation can be better explored in the two charts of Fig. 3. The left side chart corresponds to the MDTDC results from Table 1, while the right side chart corresponds to the MDTDC results from Table 2.

Table 2. MDTDC of the Munich subway network during weekend and holidays, *Saturday* (Sa) and *Sunday/Holiday* (S&H) schedules, for a given maximal travel time of 5/10/15/20 min.

Hour	5 min Sa	5 min S&H	10 min Sa	10 min S&H	15 min Sa	15 min S&H	20 min Sa	20 min S&H
00:00	12	12	22	22	27	27	34	34
01:00	12	12	21	21	25	25	27	27
02:00	12	0	21	0	25	0	27	0
03:00	0	0	0	0	0	0	0	0
04:00	12	12	21	21	25	25	27	27
05:00	12	12	21	21	25	25	29	29
06:00	12	12	21	21	25	25	29	29
07:00	12	12	21	21	26	25	34	29
08:00	12	12	22	21	30	25	39	30
09:00	12	12	22	22	30	30	39	39
10:00	12	12	23	22	30	30	39	39
11:00	12	12	24	22	31	30	39	39
12:00	12	12	24	22	31	30	40	39
13:00	12	12	24	22	31	30	40	39
14:00	12	12	24	22	31	30	40	39
15:00	12	12	24	22	31	30	40	39
16:00	12	12	24	22	31	30	40	39
17:00	12	12	24	22	31	30	40	39
18:00	12	12	24	22	31	30	40	39
19:00	12	12	23	22	30	30	39	39
20:00	12	12	22	22	30	30	39	39
21:00	12	12	22	22	30	30	39	39
22:00	12	12	22	22	30	30	39	39
23:00	12	12	22	22	30	30	39	39

Notice that, since the network has about 100 stations, 47 stations in the maximal component means about 47% of the network. Thus, the maximal component is between 12% for 5 min and 47% for 20 min., see Fig. 3.

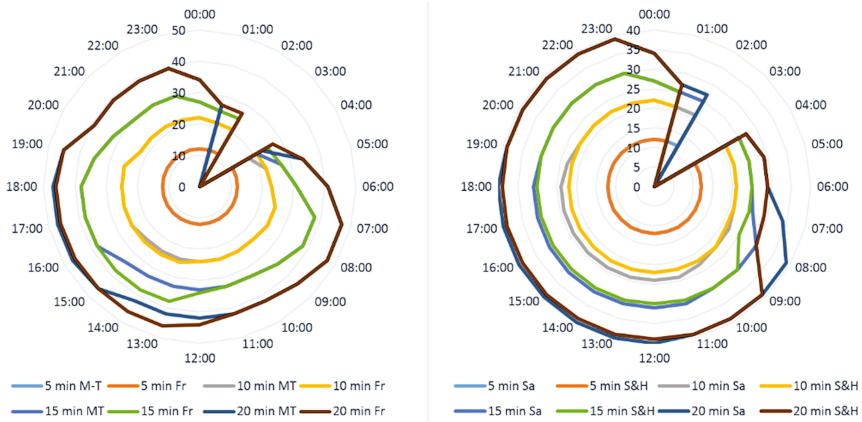


Fig. 3. Variation of MDTDC of the Munich subway network. The left side chart corresponds to week days (Table 1), and the right side chart corresponds to weekend and holidays (Table 2). The numbers along the radial axis represent the number of stations in the maximal component.

LDTDSP of the Munich subway network for each schedule is presented in Table 3. The results are the same for all the schedules. LDTDSP changes only from one time travel limit to another, but not from one schedule to another.

Table 3. LDTDSP of the Munich subway network of each schedule for a given maximal travel time of 5/10/15/20 min.

Schedule	5 min	10 min	15 min	20 min
Monday to Thursday	5	8	11	13
Friday	5	8	11	13
Saturday	5	8	11	13
Sunday/holiday	5	8	11	13

As mentioned before, there is a connection between LDTDSP and MDTDC. Thus, Table 3 can be read together with Tables 1 and 2. For example, if on *Friday* 8 o'clock MDTDC for a 20 min travel time is 47, then 13 (~28%) is given only by LDTDSP. The latter can be seen visual represented in Fig. 4.

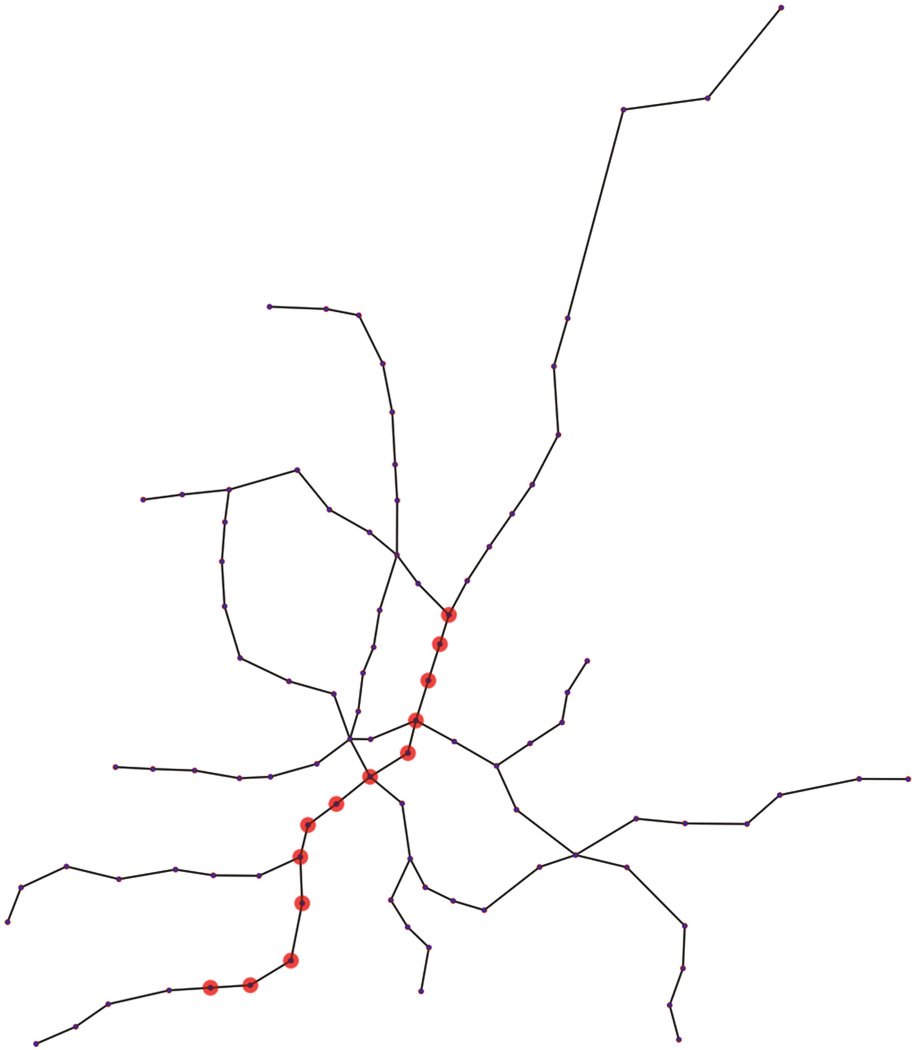


Fig. 4. The Munich subway network with the stations part of LDTDSP highlighted in red during the *Friday* schedule at 8 am for a 20 min travel time.

5 Conclusions

We analyzed the scenario where a human threat propagates through the network and we applied it on a subway system. We analyzed the properties of the system with propagation constrains in the attempt to find the maximal day-time dependent component and the longest day-time dependent shortest path for various schedules. The maximal component at the rush hours of the day, during the weekdays is very large, for a travel duration of 30 min; in fact, the maximal component is roughly half the total

network. Even for 15 min, the maximal component is almost 40%. This raises serious concerns and consequences for the security of people in the network. The application to the Munich subway system shows that the selected data samples and analysis criteria were sufficient for this application. The names of the starting stations creating the maximal components and the longest paths were removed on purpose, the results should be used only in scientific terms.

This paper is yet another example suggesting that the activity of a subway system is fragile and justifies its presence in the critical urban infrastructures category. Here, we established a well-defined analysis method that can be applied to other type of transportation systems or to other systems with similar properties.

Acknowledgments. MSN was supported by the German Federal Ministry of Education and Research (BMBF), project RE(H)STRAIN (Grant No. 13N13786). HNT was supported by a DAAD travel and exchange grant. Doina Bein acknowledges the support by Air Force Office of Scientific Research under award number FA9550-16-1-0257.

References

1. Teodorescu, H.-N.L.: Revisiting models of vulnerabilities of the networks. *Stud. Inform. Control* **25**(4), 469–478 (2016)
2. West, D.B.: *Introduction to Graph Theory*. Prentice Hall, Upper Saddle River (1996)
3. Dijkstra, E.W.: A note on two problems in connexion with graphs. *Numer. Math.* **1**(1), 269–271 (1959)
4. Lee, H.F., Dooly, D.R.: Algorithms for the constrained maximum-weight connected graph problem. *Nav. Res. Logist. (NRL)* **43**(7), 985–1008 (1996)
5. El-Kebir, M., Klau, G.W.: Solving the maximum-weight connected subgraph problem to optimality. arXiv preprint [arXiv:1409.5308](https://arxiv.org/abs/1409.5308) (2014)
6. Álvarez-Miranda, E., Ljubić, I., Mutzel, P.: The rooted maximum node-weight connected subgraph problem. In: Gomes, C., Sellmann, M. (eds.) *International Conference on AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems*, pp. 300–315. Springer, Heidelberg (2013)
7. Li, Z., Zhang, S., Zhang, X.-S., Chen, L.: Exploring the constrained maximum edge-weight connected graph problem. *Acta Math. Applicatae Sin.* **25**(4), 697–708 (2009). English Series
8. Dittrich, M.T., Klau, G.W., Rosenwald, A., Dandekar, T., Müller, T.: Identifying functional modules in protein-protein interaction networks: an integrated exact approach. *Bioinformatics* **24**(13), i223–i231 (2008)
9. Guo, Z., Li, Y., Gong, X., Yao, C., Ma, W., Wang, D., Li, Y., Zhu, J., Zhang, M., Yang, D., et al.: Edge-based scoring and searching method for identifying condition-responsive protein-protein interaction sub-network. *Bioinformatics* **23**(16), 2121–2128 (2007)
10. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T.: Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**(11), 2498–2504 (2003)
11. Conrad, J.M., Gomes, C.P., van Hove, W.-J., Sabharwal, A., Suter, J.F.: Wildlife corridors as a connected subgraph problem. *J. Environ. Econ. Manag.* **63**(1), 1–18 (2012)
12. Dilkina, B., Gomes, C.P.: Solving connected subgraph problems in wildlife conservation. In: Lodi, A., Milano, M., Toth, P. (eds.) *International Conference on Integration of Artificial Intelligence (AI) and Operations Research (OR) Techniques in Constraint Programming*, pp. 102–116. Springer, Heidelberg (2010)

13. Carvajal, R., Constantino, M., Goycoolea, M., Vielma, J.P., Weintraub, A.: Imposing connectivity constraints in forest planning models. *Oper. Res.* **61**(4), 824–836 (2013)
14. Chen, C.-Y., Grauman, K.: Efficient activity detection with max-subgraph search. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1274–1281 (2012)
15. Nistor, M.S., Bein, D., Bein, W., Dehmer, M., Pickl, S.: Time-based estimation of vulnerable points in the munich subway network. In: Doerner, K., Ljubic, I., Pflug, G., Tragler, G. (eds.) *Operations Research Proceedings 2015: Selected Papers of the International Conference of the German, Austrian and Swiss Operations Research Societies (GOR, ÖGOR, SVOR/ASRO)*, Vienna, Austria, 1–4 September 2015, pp. 355–360. Springer, Cham (2017)
16. Nistor, M.S., Pickl, S.W., Raap, M., Zsifkovits, M.: Network efficiency and vulnerability analysis using the flow-weighted efficiency measure. *Int. Trans. Oper. Res.* (2017). doi:[10.1111/itor.12384](https://doi.org/10.1111/itor.12384)
17. Münchner Verkehrsgesellschaft mbH (MVG): MVG in figures (2016). <https://www.mvg.de/dam/mvg/ueber/unternehmensprofil/mvg-in-figures-s>. Accessed 12 Mar 2017
18. MVV - Münchner Verkehrs- und Tarifverbund: MVV GmbH (2017). <http://www.mvv-muenchen.de/>. Accessed 12 Mar 2017
19. MVV - Münchner Verkehrs- und Tarifverbund: Alle Informationen zu den Bahnhöfen im MVV (2015). <http://www.mvv-muenchen.de/de/netz-bahnhoeefe/bahnhofsinformation/index.html>. Accessed 12 Mar 2017