



Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Lecture with Computer Exercises: Modelling and Simulating Social Systems with MATLAB

Project Report

The Swiss Train Network

Mario Vontobel & Elias Rieder

Zurich
Dec 2014

Agreement for free-download

We hereby agree to make our source code for this project freely available for download from the web pages of the SOMS chair. Furthermore, we assure that all source code is written by ourselves and is not violating any copyright restrictions.

Mario Vontobel

Mario Vontobel

Elias Rieder

E Rieder



Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Eigenständigkeitserklärung

Die unterzeichnete Eigenständigkeitserklärung ist Bestandteil jeder während des Studiums verfassten Semester-, Bachelor- und Master-Arbeit oder anderen Abschlussarbeit (auch der jeweils elektronischen Version).

Die Dozentinnen und Dozenten können auch für andere bei ihnen verfasste schriftliche Arbeiten eine Eigenständigkeitserklärung verlangen.

Ich bestätige, die vorliegende Arbeit selbständig und in eigenen Worten verfasst zu haben. Davon ausgenommen sind sprachliche und inhaltliche Korrekturvorschläge durch die Betreuer und Betreuerinnen der Arbeit.

Titel der Arbeit (in Druckschrift):

Swiss train network

Verfasst von (in Druckschrift):

Bei Gruppenarbeiten sind die Namen aller Verfasserinnen und Verfasser erforderlich.

Name(n):

Rieder

Vorname(n):

Elias

Vontobel

Mario

Ich bestätige mit meiner Unterschrift:

- Ich habe keine im Merkblatt „[Zitier-Knigge](#)“ beschriebene Form des Plagiats begangen.
- Ich habe alle Methoden, Daten und Arbeitsabläufe wahrheitsgetreu dokumentiert.
- Ich habe keine Daten manipuliert.
- Ich habe alle Personen erwähnt, welche die Arbeit wesentlich unterstützt haben.

Ich nehme zur Kenntnis, dass die Arbeit mit elektronischen Hilfsmitteln auf Plagiate überprüft werden kann.

Ort, Datum

Zürich, 12.12.2014

Unterschrift(en)

E. Rieder
Mario Vontobel

Bei Gruppenarbeiten sind die Namen aller Verfasserinnen und Verfasser erforderlich. Durch die Unterschriften bürgen sie gemeinsam für den gesamten Inhalt dieser schriftlichen Arbeit.

Contents

1 Abstract	5
2 Individual contributions	5
3 Introduction and Motivations	5
3.1 Idea and Motivation	5
4 Description of the Model	6
4.1 The classical gravity model	6
4.2 The development of the gravity model	6
4.3 Comment on the distance in the gravity model	7
4.4 The Data approach	8
5 Implementation	8
5.1 Network flow-function	8
5.2 Visualization	8
5.3 Investigation-function	9
6 Simulation Results and Discussion	10
6.1 General Network Parameters	10
6.2 Networkflow	10
6.3 Resilience	11
6.4 Capacity	12
7 Summary and Outlook	12
8 References	14
9 Appendix	16
9.1 Seat critical trains	16
9.2 Answers from Nadine Ruch from SBB	17

1 Abstract

This project looks at the Swiss train network. It investigates the flow of people in this transport system which is quiet important for this country. The goal was to find relations between properties like capacity, amount of people travelling, delay and fragility. On one hand this is approached by a general theoretical model. On the other hand by real data. These two methods lead to a model of the Swiss train network. On the basis of this model the properties are analysed. The two approaches merge at the end and allow us to make some statements about the quality of the theoretical model.

2 Individual contributions

We made the whole work in a close teamwork. It was made in continuous exchange and by inspiring each other. Mario set the focus a bit more on coding. Elias on the Concepts and Data. You can see all the steps, code and data on github. Mario has the username mlv19 and Elias is EleSwag.

<https://github.com/EleSwag/Elias-und-Mario>

3 Introduction and Motivations

3.1 Idea and Motivation

Every day we travel from our homes to Zürich. The train in the morning is often very crowded and a lot of delays occur. This circumstance made us think about how the train network works and how you could handle such capacity shortages better. So the general idea was to model the Swiss train network and then look at its properties. This may sound like a very general question. As you can imagine, the whole schedule is quiet complex and consists of a lot of data. At this point we didn't know how much and what kind of data we would have access to. Thats why the we kept the general. We heard in our lecture that the number of inhabitants has an influence on the amount of traffic. The type of model using this approach is called Gravity model.

So we had two approaches to start with: The Gravity model, witch will be explained in detail after, and the effort getting real data.

Because our interest in the problem came from observations in our everyday life, it was very important to us, that our work would have a strong connection to reality. To make this happen, we wrote a mail to SBB, the Swiss rail company, and asked them for real Data. Since we did not know if we would get the data we had two make two plans.

The first possibility would have been getting a rich data set from SBB. Complex models considering the frequency, speed and capacity of the trains and the resulting connections

would be possible. In this case the question of capacity and load, and resilience were very interesting questions.

The second, alternative possibility was getting nothing. But with the gravity model we could be sure to have an angle even without data. This was kind of a backup plan. Of course the questions would have to be asked way more general

So the question that draws through both approaches is:

How do you make a model that allows us to approximate the flow of people in the Swiss train network and enables us to make statements about the traffic flow, the capacity and the resilience of it?

4 Description of the Model

4.1 The classical gravity model

In social science -especially in international economics in trade simulations- it is a common and well established approach to simulate flows with a gravity model. The motivation behind it is the physical gravity force. This is defined by $F_G := G \frac{m_1 m_2}{r^2}$, where G is the gravity constant, m_i is the mass of the i-Th body for $i = 1, 2$ and r is the distance between body 1 and body 2.

This leads to the general ansatz₂:

$$F_{ij} = G \frac{M_i^{\beta_1} M_j^{\beta_2}}{D_{ij}^{\beta_3}}$$

We transferred this idea to our situation by identifying the mass of a city to its population₃ and used different definitions of distance. We dropped the constant at the beginning since we normalized the resulting flow anyway. As exponents we have chosen $\beta_1 = \beta_2 = \beta_3 = 1$ since with no experience this was as suitable as anything else.

4.2 The development of the gravity model

In our first approach we took the 10 biggest cities in regard of the number of inhabitants plus two additional ones (Olten and Arthgoldau) which are well-known important nodes₄ of the real network.

We wondered if they would also become as important in our gravity model as in the real network.

To make the network more realistic we have looked up all real connections between the cities on the online train schedule₅. Our criterion was that two nodes are connected if there exists a direct train connection between them. With this idea we didn't end with

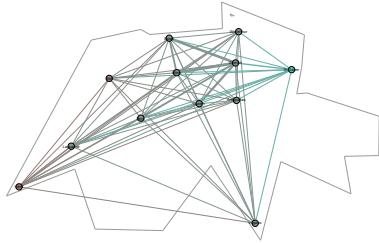


Figure 1: Gravity model 1 (full connected): Model of the 10 biggest cities plus two additional one. Layout by gephi with respecting geographical position.

a fully-connected network which is from our point of view more realistic to simulate the direct flow between cities.

This was the first network we plotted. While looking at the picture it was striking that the network didn't at all cover the hole country. Especially the south-west part was unattended. So we extended the number of cities to 20 and kept the two additional ones from the beginning.

4.3 Comment on the distance in the gravity model

In the general gravity model we used geographical distance as the distance indicator. But we thought it would also be reasonable to use the time it takes to travel from one city to the other as the distance indicator. The idea was that this also takes in account if there is obstacle like a mountain or a lake between the cities. The danger here is, that the network is approximated with the existing network. Anyway we omitted that later because with the number of connection the investment grow nearly immeasurable. Also the time connection was not well-defined any more. In the sense that there were different times for the same connection.

4.4 The Data approach

As mentioned at the beginning the second way looking at our network was real data. Unfortunately the search for data turned out to be quiet difficult. We contacted SBB for data, but in a big company like this things are taking very long. When we finally got in contact with the right persons^{8–10}, there was just to less time left to process all and we could not get all the data we would have liked. Especially data about the load of the network and the capacity seems to be very sensitive. But the people from SBB were very kind and gave us what they could. They sent us a public link¹¹ were you can download the whole timetable in Gtfs-Format (General Transit Feed Specification). Sadly we could not involve that, because of the already mentioned time shortness. We also could sent several questions to one of their networkspecialists¹⁰. Especially the numbers of people boarding and de-boarding for a few big cities made us curious. This was a number we may could relate with our gravity models.

5 Implementation

In general we tried to be as specific as possible with the annotation of the functions in matlab

We created a script called data1 which contains all our relevant data. As already mentioned, all data and code is on GitHub¹.

5.1 Network flow-function

First we implemented a general network function which applies the theory mentioned in the description of the model to an adjacency matrix of a network. This function gave another adjacency matrix back with the simulated flow. This function would also run for every other gravity model based simulation.

For the distance based gravity function we also needed the geographical distance. For that purpose we added to each city its coordinates¹⁵ and computed the distance between each city by just applying Pythagoras. Since Switzerland is a small country we did not take in account that the earth is a sphere and just approximated as if it was a plane. The function gives again an adjacency matrix back, where the ij-th entry is the distance between the point at coordiantes i and tthe point at coordiantes j.

5.2 Visualization

In the lectures it was recommended to use gephi. So some of our pictures are made with it. But a huge disadvantage of gephi is that it is rather complicated to respect geographical position if one uses an adjacency matrix to represent the network. In figure 1 we placed the cities by hand.

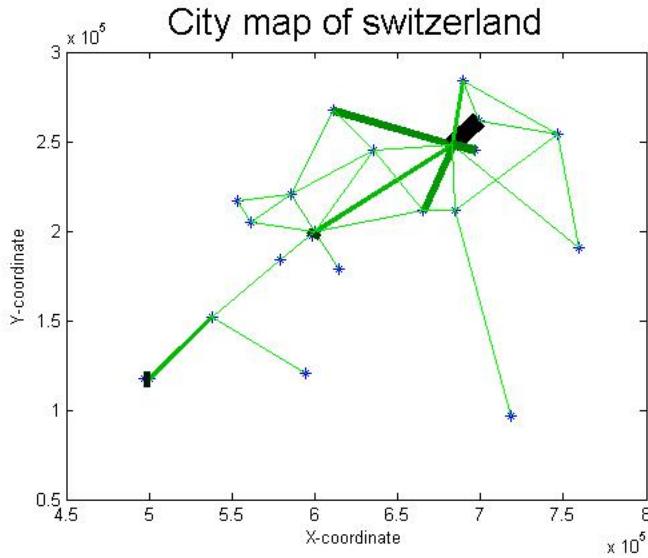


Figure 2: Gravity model 2: Model of the 20 biggest cities plus two additional one. Layout by matlab. The strength of a connection is emphasized by the thickness as well as the depth of the colour from the line. The coordinates are given in the Swiss coordinate system/Swiss grid₁₅

But since this is rather complicated we also wrote function in matlab to illustrate a network given in a adjacency matrix. The result isn't as pretty as in gephi but the geographical representation was more straight forward.

5.3 Investigation-function

To investigate the network we strongly relaid on the paper 'Complexity in human transportation networks: a comparative analysis of worldwide air transportation and global cargo-ship movements'. We implemented the shortest distance function described on page 592/593. At the beginning we had a hard time to find the right function but then Mario remembered that he had in the fall semester 2013 a lecture₇ on algorithm and complexity. Mostly the lecture was about theory from trees (special case of a graph) but there was also a part on general graph theory. In that part was the 'Floyd-Warshall'-algorithm which does exactly what we needed. Which means shortest paths between all nodes. The notes also allowed as to implement a function which checks if a network is still connected by using a 'depth-first-search'-algorithm.

6 Simulation Results and Discussion

Note: When the x-axis is labeled cities with number this is always referring to our city-vector, which is sorted by the size of its population (Zuerich, Genf, Basel, Lausanne, Bern, Winterthur, Luzern, St.Gallen, Lugano, Biel, Thun, Koeniz, Chaux-de-Fonds, Schaffhausen, Freiburg, Chur, Neuenburg, Vernier, Uster, Sion, Olten, Arthgoldau)³

6.1 General Network Parameters

	N	L	σ	ϕ	d_T	$\langle r \rangle$	$\langle w \rangle$	$\langle F \rangle$	$\langle k \rangle$
Full-connected	22	231	0.9545	1	0.8413	$2.6604e + 07$	0.0275	0.5781	21
Connected	22	34	0.1405	2.9134	11.0154	$2.6028e + 07$	0.1037	0.3215	3.1364

Figure 3: N is the Number of Nodes, L is the number of connections, the Networkconnectivity $\sigma = 2\frac{L}{N^2}$, ϕ is the betweenness centrality, d_T is the average shortest pathlength, $\langle r \rangle$ is the average geographical distance $\langle w \rangle$, is the average node weight $\langle F \rangle$, is the average flow per city and $\langle k \rangle$ is the average node degree

The full connected network is the general network from the gravity function, where everything is connected. The connected network is the one developed by us (only connected if there is a direct link).

Beside other obvious observations, it is quite intuitive, that ϕ rises with less links between nodes. One surprisingly observation was that although the mean link rises with a factor 5, the average node flow only sinks by a factor 2.

To investigate more network parameters like node distribution more nodes would be needed. Since in our model each nodeweight or nodedegree appears nearly only once.

The detailed definitions of these parameters are described in reference 6

6.2 Networkflow

To determine how powerful the gravity model is, you have to be very careful about your question. It strongly depends on your point of view. If you want to get an impression of the amount of traffic of the 10 biggest cities in Switzerland, you might get quite good results. But if you slightly change your point of view and ask yourself in which cities of Switzerland a big amount of traffic occurs, your gravity model might not be overall correct.

We tried to get an idea how our model responds to reality. For that we brought our (de-)boarding into play. Because the flows in our model are undirected, the sum of all flows in a node should be equal to the people boarding plus the people deboarding. To check if how good our calculated flow matches with the real flow, we divided one by the other to get a proportion. We did this for both of our models. You can see in the graphics that the

citys with high population are approximated with a similar relative factor. But the for important nodes of the real network both gravity models fail. Have a look at Olten (small number of inhabitants big amount of traffic). Even one of the seat critical trains¹², departs from olten (train were you might have to stand). But precisely for further investigations on capacity, these flows would be important.

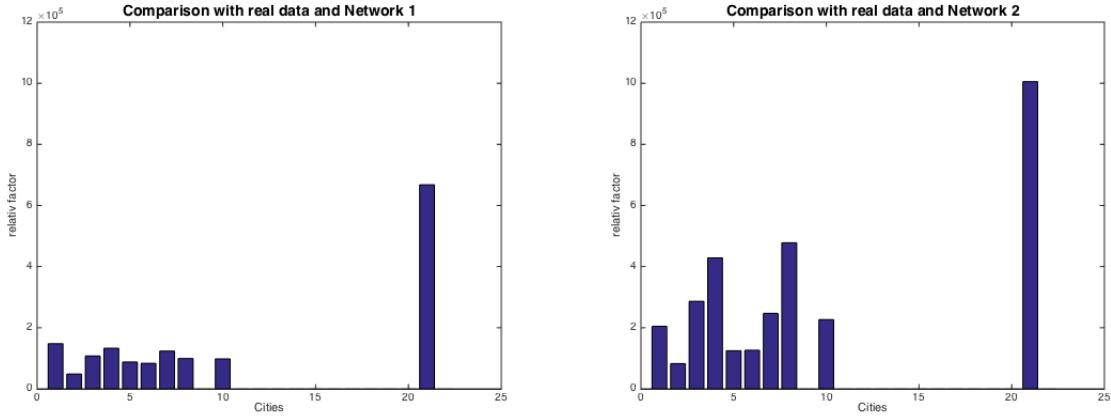


Figure 4: Network is the full connected gravity model based model and network 2 is the one with the connections from the schedule. The relative factor = $\frac{\text{Real data}}{\text{networkflow per City}}$. The networkflow per city is calculated for each city by summing over the columns.

The other, pretty obvious weakness of the gravity model are time variable flows. We learned, that this is a quiet important aspect if you want to find out something about capacity of this network. The fastest growing traffic flows in the network are the peaks in the rush-hour¹⁴.

6.3 Resilience

The resilience of a network can be investigated by looking at a few parameters. One of them is the change of the longest relativ distance when a node is removed. We applied this to two of our network matrices. One was the full connected and the other the connected gravity model. We were interested if the two models would behave similar or not. Network one was besides Zürich very resilience resistant meanwhile network 2 also had nodes where the network was not connected anymore if removed and in general also more sensitive. A remarkable observation was that in the network 2 the diameter phi falls extremly when Lugano is removed.

Besides that SBB assured¹³ us they have replacement train at important nodes.

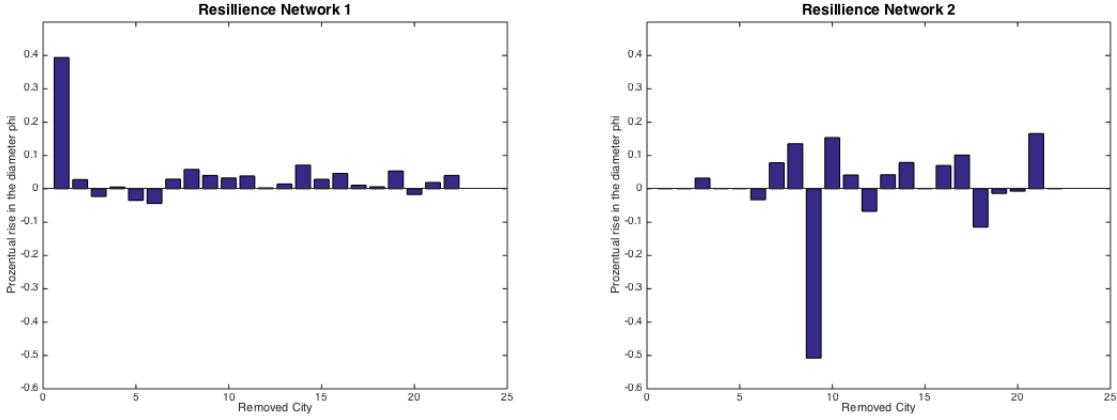


Figure 5: Change of the shortest distance if a node is removed for the full connected gravity model (Network 1) and the connected one (Network 2). For network 2 zero means that the network was not connected anymore

6.4 Capacity

Unfortunatley none of our approaches could answer our questions about capacity. The Gravity model can give Information about the amount of traffic, but it does not tells anything about how big this flows can get. What SBB told us is, how they deal with it: "If the route allows it, the train is maid as long as possible. All Platforms where the train stops have to be long enough. As additional action additional trains are added during the rush our...Alternatively the rolling stock can be optimised. On routes with little space, rolling stock with high capacity is used. If all these actions don't suffice, an extension is required" ¹³. One other thing they mentioned is, that their trying to foster flexible worktime and homeoffices. Their doing this by negotiating with other big employers⁹. This all seems all to be aimed at decresing the rushour peaks, not at handeling them. An other thing they mentioned is, that with higher frequencys and shorter travel times higher the attractivity of line. Which leads back to higer demands and more traffic, a vicious circle.

7 Summary and Outlook

The model of the swiss train network had its strength and weaknesses. Our goal to model the swiss train network was maybe aimed a bit high. We could not answer our questions about resilience and capacity as detailed as desired.

The gravity model gives in general a good approxiamtion for the flow of the cities besides important nodes in the real system like olten which have not a very huge population. Therefor the gravity model is good to approximate the flow of big cities. But this also

means you will not find other high flows in the network. So the question for important nodes can not be modeled with a gravity approach.

One way to receive another flow would be to change the parameters $\beta_1, \beta_2, \beta_3$ in the gravity function. So one could take the geographical distance more in account than the size of population for example.

Our investigations on resilience has shown that the real network is probably not as resistant as the general gravity model. But to real test the network on resilience one would need more nodes since our present network gets disconnected quite early. This would also allow to look at nodedistribution and other parameters. This could may be arrived with the Gtfs-files. Which would also lead to new networks which may have a more realistic flow.

The answers from SBB also let us suppose that they also try to solve problems in the network with reducing the demand of traffic. So they try to regulate the flow more with economic incentives than with improving or changing the network structure.

We found out quite interesting things about the Swiss train network and some interesting trails for further research.

Another aspect who would be worth investigating are the rushour peaks. Maybe it would be possible to approach them with differential equations, but it seem to become very complex there. Even the newokspecialist from SBB could not give us a general answer what the demands for a network are to handle peaks efficient. (ref)

8 References

1. Projectfolder: <http://Github.com/EleSwag/Elias-und-Mario>
(all code and files are accessible here).
2. Gravity Model http://en.wikipedia.org/wiki/Gravity_model
3. List of Cities http://de.wikipedia.org/wiki/Liste_der_Stdte_in_der_Schweiz
4. Important Train Nodes <http://de.wikipedia.org/wiki/Eisenbahnknoten/#Schweiz>
5. SBB online timetable <http://www.sbb.ch/fahrplan.html>
6. O. Woolley-Meza, C. Thiemann, D. Grady, J.J. Lee, H. Seebens, B. Blasius and D. Brockmann: Complexity in human transportation networks: a comparative analysis of worldwide air transportation and global cargo-ship movements?
7. Vorlesung Algorithmen und Komplexität Angelika Steger ETH Zürich
SBB Correspondence
8. Werner Nuber
Director Senior Advisor
Schweizerische Bundesbahnen SBB
9. Matthias Müller
Communication passenger traffic
Schweizerische Bundesbahnen SBB
10. Nadine Ruch
Passenger traffic, corporate development, demand development
Schweizerische Bundesbahnen SBB
SBB Data

11. Link for timetable Data <http://gtfs.geops.ch>
12. Seat critical trains (Appendix)
13. Answers from Nadine Ruch from SBB (Appendix)
14. SBB Konzernlagebericht 2013 page 7 section 5 http://geschaeftsbericht.sbb.ch/fileadmin/user_upload/Downloads/SBB_Konzernlagebericht_GBNB_2013.pdf
15. Swiss Grid http://en.wikipedia.org/wiki/Swiss_coordinate_system

9 Appendix

9.1 Seat critical trains

Sitzplatzkritische Züge			Alternative Verbindungen		
Zug von / nach	Belegung (1.8.14 bis 31.10.14)		Zug von / nach	Belegung (1.8.14 bis 31.10.14)	
	1. Kl.	2. Kl.		1. Kl.	2. Kl.
51 Olten 06.59 – Bern 07.29	81%	87%	957 Olten ab 06.29 – Bern 06.56	57%	70%
1908 Olten 06.58 – Bern 07.25	88%	96%	957 Olten ab 06.29 – Bern 06.56	57%	70%
810 Zürich 07.02 – Bern 07.58	76%	69%	1008 Zürich 06.49 – Bern 07.49	69%	51%
710 Zürich 07.32 – Bern 08.28	89%	84%	1012 Zürich 07.49 – Bern 08.49	67%	63%
812 Zürich 08.02 – Bern 08.58	82%	78%	1012 Zürich 07.49 – Bern 08.49	67%	63%
835 Bern 17.02 – Zürich 17.58	90%	89%	1035 Bern 17.11 – Zürich 18.10	58%	61%
837 Bern 18.02 – Zürich 18.58	73%	89%	737 Bern 18.32 – Zürich 19.28	66%	78%

9.2 Answers from Nadine Ruch from SBB

Questions SBB

Network structure

- How did the structure of the Swiss rail system arise?
We can't answer this question. Basically the network grew historically, the details you have to look up in the literature
- How you determine future connections?
*The BAV (Translator's note: BAV is "Bundesamt für Verkehr" means "Swiss Ministry for Traffic) published the planning process on the Internet. There the planning process is quiet well visible.
<http://www.bav.admin.ch/fabi/index.html?lang=de>*
- Is it the best possible one/Is the structure in change?
Of course we try to optimize the use of the trains and the network continuously. For this also look at the BAV link.
- What are its advantages and disadvantages?
The schedule is the most important criteria for our customers, insofar offering improvements are central. This can be read in the demand which is noticeable rising when attractive Improvements are made (essential is the frequency and the travel time)
- How is the capacity of a line/connection raised? More Frequent trains or longer trains (More people per train)?
It depends. If the route allows it, the train is maid as long as possible. All Platforms where the train stops have to be long enough. As additional action additional trains are added during the rush our. We do this for example between Zürich and Bern, where factual a quarter hour tact is offered. Alternatively the rolling stock can be optimised. On routes with little space, rolling stock with high capacity is used. If all these actions don't suffice, an extension is required. Look at Link from BAV.
- Does a network with high peaks of load require different network structures than a network with constant load?
It depends on the situation. This cant be answered so easy
- What are the critical nodes of the Swiss train system?
Look at the BAV link.

Resilience

- What are the system countermeasures against delays/cancellations?
It depends on the reason of the delay. There is a reserve in the timetable with which the delays can be make up????????? Besides there are standing reserve trains at several Nodes which can be used in case of a "strong" delay. These trains are immediately ready to use and are used if disturbance endangers the stability of the timetable.