

# Analyze Road Traffic Condition by Visualization of Data from Transportation Simulation Systems

Edilson Dias Alves<sup>1</sup>, Xuan Zhaocheng<sup>2</sup>

<sup>1,2</sup>Tianjin University of Technology and Education (TUTE), School Of Information Technology and Engineering

**Abstract:** Nowadays by increasing importance of information in all sectors, illustrating data in a communicative format helps decision makers to understand and analyze effectively large amount of information in a short time. Information visualization, as a way of presenting different data types in a more understandable form, is growing increasingly in various areas. This thesis aims to investigate how information visualization could be used to increase readability and usability of transportation simulation data. Most of the previous studies in this area have focused on visualization of transportation infrastructures such as roads, bridges, etc.; while the main focus of this thesis is visualizing the outputs of transportation simulation systems. In order to study the role of information visualization in the transportation systems, I have investigated the visualization applications in a real implemented agent-based transportation simulator called TAPAS as case. In my case study, the visualization related requirements of users were analyzed, and the visualization tool was designed and developed based on the identified requirements.

**Keywords:** Road traffic simulation, transportation data visualization

## 1. Introduction and Background

### 1.1 Introduction

Nowadays, there is a growing use of simulation in transportation [1]. The idea is to attempt to simulate the real system characteristics in a computer system and observe how applying different scenarios will affect the behavior of the transportation system. This can help in planning process where the decision makers wish to be aware of the consequences of a specific decision on the whole system. The main application of simulation in the transportation context can be cost reduction and increasing quality of service. However, not all the transportation simulation systems are necessarily successful regarding reducing costs and increasing quality. There are some factors which determine the degree of success in simulation models. One important factor is the "Ease of use" of the simulation system by different types of potential users [1]. Typically, a simulation system generates a large amount of data as output. This data must be processed and analyzed in order to be useful for different type of users and decision makers. Hence, the way of analyzing and presenting the output data to users, can be a challenge. As we acquire more information through vision than through all of the other senses [2], visualizing the output data of transportation simulation systems can be an ideal solution.

Information Visualization provides visual representations of large amount of data in order to be more usable and understandable for users. Information visualization could be used in different contexts and for different purposes, such as Network Analysis, Semantic networks and knowledge representation, Cartography, VLSI Circuits design, Genetic maps, etc.

In this thesis, visualization techniques will be applied on output data of transportation simulation systems to increase the understanding of results by different stakeholders. This can facilitate the decision making process for policy makers by visualizing the complex data of simulation results on

screen. In particular, the visualization techniques will be applied on an agent based simulation model for transportation and production called TAPAS [3], Transportation And Production Agent-based Simulator. TAPAS is a micro-level simulation tool which aims to predict the effects of applying different measures including governmental policies like taxes and infrastructure investments and business strategies [4].

TAPAS can act as a decision support system for public authorities and policy makers.

They can monitor what could be the effects of applying a specific set of production and transportation measures and whether it contribute to the government goals such as sustainable economic development or not. Furthermore, simulation can also help enterprises to make both operational and strategic decisions, for example decisions regarding component size and storage locations, etc. [3].

### 1.2 Purpose and Motivation

Information Visualization has an important role in different contexts. In fact, it has been used in different fields and it is an expanding area of knowledge [5].

Obviously, the nature of the data to be visualized is different in various application areas. The choice of selecting a set of visualization techniques and applying them to a data set is highly dependent to the context of data set. As an example, in case of being some inheritance relations between the data elements, we can make use of graphs as a way of representing and visualizing information [6]. In visualization of large and complex graph data, some problems and issues will emerge. Visualizing the whole data in one graph might not be useful for the users and also it causes performance problems and increases the complexity and therefore decreases the readability of the result. Also, it would be more meaningful to display some elements in a different color and shape to increase readability of visualized graph.

Also, some layout algorithms have emerged to solve the problems of visualizing complex and large data. Some of these algorithms have focused on aesthetics criteria in the context of Graph Drawing [7], [8]. But still there is a need to mix these approaches and techniques to reach usability requirements of potential users of this specific context. Thus, it is necessary to investigate how visualization techniques can be applied to transportation systems.

There are other kinds of visualization techniques, such as displaying data through diagrams [9]. Based on the nature of data, some data should be visualized by diagrams.

The focus of this thesis is visualization of data in the context of transportation systems. Indeed, the graph visualization techniques are applied and tailored to visualize data related to transportation systems and link connections between different nodes, in a graph-based shape. Also, we applied other data visualization techniques, such as diagrams. TAPAS, which is a multi-agent based production and transportation simulator is used as a case study [4].

The purpose of this simulator is to show how organizations in supply chain might act after applying different government control policies and whether the outcomes are desirable and near to their goals or not [10]. Currently this simulator does not show the results of simulation graphically, which would make it easier to understand the results.

### 1.3 Quality Attributes

The quality attributes which this thesis has focused on are readability and usability. The readability of a visualization model can be defined as how easy and fast the user can find the information he/she is looking for [11].

Usability is a term which contains several factors. Davis [12] has proposed a model for Information Technology acceptance. He has categorized two main factors which cause users to accept or reject a system. The first factor is the usefulness, which is defined as "The degree to which a person believes that using a particular system would enhance his or her job performance" [12]. This factor can be described by these characteristics [13]:

- 1) Allows working faster.
- 2) Improves job performances.
- 3) Immediate insight.
- 4) Adaptability to task.

The second factor which called the ease of use is defined as "The degree to which a person believes that using a particular system would be free of effort". The characteristics of this factor are listed below:

- 1) Easy to understand.
- 2) Easy to learn.
- 3) Content categories relevance.
- 4) Aesthetic value.

In this thesis I will use the term "usability" which refers to both usefulness and the ease of use of the system.

### 1.4 Goal

This Master thesis project aims to increase usability of transportation simulation system by visualizing the output from different simulation scenarios. The goal of this project is to support the analysis of how actors in a transport chain act given various measures through visualization of information. This goal can be achieved by investigation of the methods for increasing readability and usability of visualized output in this specific context.

### 1.5 Research Question

The central research question is formulated as follow:

- How can visualization techniques be applied to transportation systems?  
Questions that stem from this question are:
  - RQ1: What are the requirements of users who want to view the results from a transportation system simulator in a visual representation?
  - RQ2: What are the important features for increasing readability and usability of a visualization in the context of transportation systems?
  - RQ3: How can visualization techniques be implemented to meet the requirements of the users?

### 1.6 Research delimitation

This research is limited to the visualization of transportation systems and will only consider the techniques which could enhance usability in this context. Therefore, the other parts of supply chains are out of the scope of this thesis.

### 1.7 Results

This Master thesis provides guidelines on how output data of a transportation simulation system could be visualized. Moreover, a tool for visualizing extensive amount of data in the context of transportation systems has been developed. In addition, the final report contains a survey about existing Graph drawing techniques and algorithms. A customized algorithm that suit to the requirements of a complex graph with complex relations and different type of nodes in the context of transportation systems is proposed. This new algorithm covers both aesthetics and readability concerns as well as the usability requirements of users.

## 2. Research Method

This section is aimed to describe the research approach undertaken for this thesis. This section is divided into four sub-sections. In the first sub-section section, the chosen research methodology is described. In the next part, the possible data collection methods are introduced. In the last section, the evaluation methods which are used for the evaluation of this research are described. The figure below (Figure 1) shows the whole research process of this research. This research process is designed based on the findings from [15, 17].

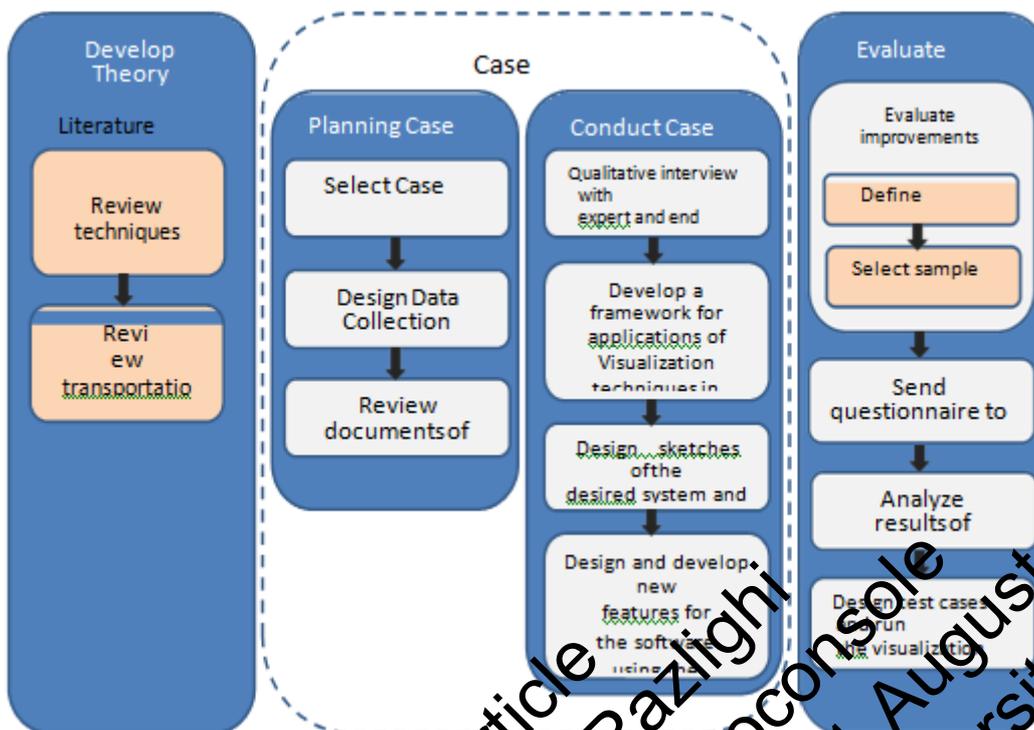


Figure 1: The research process of this thesis (Based on [15, 17])

## 2.1 Literature Review

Because of the fact that there are a tremendous amount of literature in the fields of Information Visualization and Graph drawing algorithms [14], there is a critical need for doing a deep literature review. For this reason, some researchers working on information visualization and graph drawings are identified to use their knowledge published in articles in this field.

## 2.2 Design and creation

In this thesis, a new IT artifact will be developed, so the main research strategy of this thesis is "Design and creation" [15]. As the result of this research we will have a tool for visualizing the simulation results of TAPAS [3], which currently does not have any visualized output. According to March et al [16], there are four types of IT artifacts: Constructs, Models, Methods, and Instantiations. In this research an instantiation, a working computer-based system, was implemented.

Before implementing the visualization options, some sketches from the proposed system were prepared. These sketches were shown to the users to get the approval and their feedback about design alternatives and options. This could be an iterative process that is continued until receiving positive feedback from the users.

Regarding the main research question of this thesis which is "How can visualization techniques be applied to transportation systems?" and two other factors of the above framework, the best research approach for this research could

## 2.3 Case Study

According to Yin R [17], there are three types of case studies: exploratory study, descriptive study and explanatory study. In this research, an exploratory case study will be performed as the first step. The TAPAS simulator [3] will be used as a case to investigate what are the current lacks and possible problems of using visualization techniques in transportation simulation context. In fact, the TAPAS software as existing software for this context will be improved by applying the visualization techniques.

## 2.3 Data Collection methods and techniques

As Yin [17] stated, there are six different resources for evidences which researchers could use in data collection phase: documentation, archival records, interviews, direct observations, participant observation, and physical artifacts. Moreover, Yin [17] has recommended using different resources of evidences related to a fact to increase the quality of research. Therefore, the author will use documents, and qualitative interviews to investigate how visualization can improve transportation simulators. In the table below (Table1), the reasons of selecting these sources of evidences are described. Moreover, it provides a chain of evidences from research questions to data collection and conclusion. The author will use Table 1 to map research questions to the sources such as interview and documentation. Cells of the Table 1 show how different sources should be used to collect data associated with each research question. Furthermore, the literature review which has been done regarding each research question is described.

**Table 1: Mapping research questions to the resource of evidences**

Resources: Research questions	Documentation of TAPAS	Interview	Literature review
RQ1: What are the requirements of users who want to view the results from a transportation system simulator in a visual representation?	Review the technical documents of transportation simulation systems or in general, supply chain simulators and extract potential candidates for visualization	Asking transportation experts and users of system; What are major problems in the process of monitoring the results of simulator which cause difficulties and interrupts in their work? How much time and cost they spend as an outcome to above problems? What do they think about having simulation results visualized? Which information are they interested in?	Review the Information Visualization literature in order to gain knowledge about visualization and also to conduct an effective interview. Review the transportation visualization sources, and extract important challenges of simulation results monitoring.
RQ2: What are the important features for increasing readability and usability of a visualization in the context of transportation systems?	Review the documents of a sample output result to find out if there is any type of visualization or not. What is useful in the existing output format and what problems and difficulties exists?	The results of interviews will be investigated to find which of the graph drawing metrics extracted from literature review, should be used in visualization of transportation simulation results and how these metrics can be applied to this context.	Review the literature of graph drawing techniques to find all possible metrics which could be used in visualization. Evaluate these metrics in terms of usability in the context of transportation systems.
RQ3: How can visualization techniques be implemented to meet the requirements of the users?		By processing the interviews, it will be identified how should be the final presentation of the transportation simulation results and which is important in visualization of this context. With consideration to these results and by having the visualization techniques, it will be clear which techniques should be used in developing a visualization tool.	Review the previous works which have been done for layout algorithms field and find the characteristics and capabilities of each layout algorithm. Find and evaluate required components for developing the Visualization tool.

## 2.4 Evaluation methods

For the evaluation of the product regarding technical aspects, there are some methods like black box testing and white box testing [15]. Black box testing is usually used to ensure that the specific inputs could generate the desired output, while white box testing is usually used when the researcher wants to evaluate the correctness of all possible paths of the code. As this research is about applying visualization to a transportation simulation system and developing a visualization tool, it is possible to test it by both black box and white box testing.

For black box testing, the simulation results of TAPAS can be considered as input for visualization tool. So, it is possible to test the visualization tool regarding the correctness of the created output. It means that we can design test cases to check if a special set of input creates the correct output or not.

Furthermore because the nature of this research is based on user satisfaction and improving the process of monitoring simulation results, we can also evaluate this research from a different point of view. Oates [15] has introduced three type of evaluation of a product from a different point of view:

- Proof of concept: is based on creating a simple prototype to show the feasibility of the researcher's solution in terms of having the specific properties under specific circumstances.
- Proof by demonstration: it deals with evaluation of the product in practice, but not in the real-world. In this approach, the researcher will restrict the context and apply the solution or use the product on a restricted context.
- Real-world evaluation: In this approach, the researcher is examining the product in a real context not in an artificial one.

In this research, the second approach for evaluation of artifact was used. Therefore, we have evaluated the results in a restricted domain with non-real users. We have planned to do real world evaluation in the future, where a real case and real users will be used for evaluation. For this purpose, an existing transportation simulation system, "TAPAS" [3], has improved by enabling users to observe simulation results visually. For evaluating whether these improvements are useful or not, we need another strategy.

Pleasant has categorized common evaluation methods in the field of Information Visualization into four groups [18].

- Controlled experiments comparing design elements.
- Usability evaluation of a tool
- Controlled experiments comparing two or more tools.
- Case studies of tools in realistic settings

So, for the evaluation of improvements of the case, the second approach was used. Indeed, after improving "TAPAS" as a case study, these improvements were evaluated through performing usability evaluation of TAPAS.

This could also be done through Sending a questionnaire to TAPAS users to find out if the improvements could increase readability of TAPAS output or not. This work is considered as a future work for the thesis.

## 3. Literature Review

It is crucial to understand the concept of visualization to develop a tool for data visualization. Thus, first part of this chapter reviews the related concepts of information visualization. The second part previews literature about

principles and theories of graph drawing. Also, previous studies on information visualization in the transportation context will help to use best practices in the design and development phases. Therefore, in the third section, previous works on application of information visualization in transportation systems are presented.

### 3.1 Information visualization

Information Visualization is a relatively new area of study which is acquiring more and more attention in both academic and industrial environments. This concise introduction to the subject explores the use of computer-supported interactive graphical representations to explain data and amplify cognition [19]. In fact, visualization plays an important role in human perception of the environment. "Visualizations have a small but crucial and expanding role in cognitive systems. Visual displays provide the highest bandwidth channel from the computer to the human" [2]. Owing to the importance of visualization in real world problem domains, it has been growing tremendously in recent years. Several books have been published in Information visualization. Also the number of conferences in this field has been steadily increasing [20].

Information visualization is a multi-disciplinary area of science and includes elements of human computer interaction, computer science, graphics, visual design, psychology, and business processes. Nowadays, Information visualization has a large area of applications including: scientific research, digital libraries, data mining, financial data analysis, market studies, manufacturing production control, and crime mapping [21].

### 3.2 Principles and basic theories of graph drawing

Graph Drawing is a way of visualizing datasets. The positioning of nodes (or vertices) and edges can influence usability, readability and aesthetics. The algorithms which are used for positioning are called "Graph Layout Algorithms". There are different layout algorithms with different positioning perspectives. In fact there is no best graph layout algorithm and it depends on the purpose of visualization [22].

This is highly relevant to the approach of visualizing transportation systems in this thesis, because of the fact that the positioning of nodes and links and their attributes are key factors in the context of transportation systems.

In this section, the general metrics and criteria for drawing graphs in an aesthetic way will be introduced. Furthermore, some of the most common graph layout algorithms will be described.

#### 3.2.1 Graph drawing concepts

There are some theoretical guidelines for graph drawing. Battista et al. have categorized most important parameters of graph drawing methods into three graph drawing concepts, to achieve better description of requirements of a good drawing [23]. These principals are as followed:

- Drawing conventions
- Aesthetics
- Constraints

#### Drawing conventions

A drawing convention is a basic rule that the drawing must satisfy to be admissible. [23]. In fact, drawing conventions are the basic rules which must be considered in graph

Plagiarized Article  
Banafsheh Hajinasab Razighi  
Paul Davidsson, Annabella Loconsole  
Master Thesis Project 30p, Spring 2017  
School of Technology Malmö University  
E-mail: Banafsheh.Hajinasab@mah.se

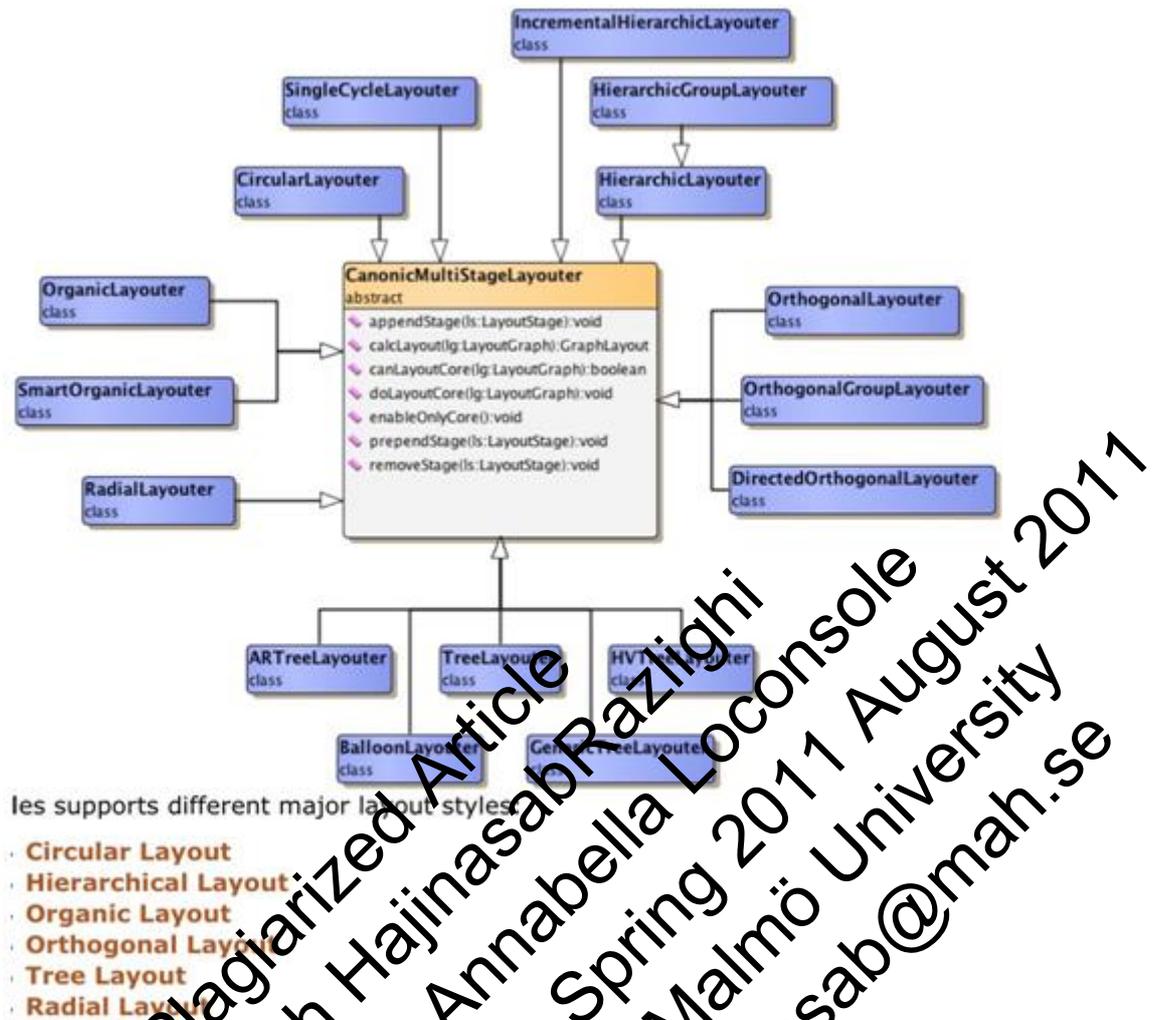


Figure 3: Taken from [7]

### Aesthetics

Graph aesthetics concern parameters which could increase graph drawing readability through focusing on aesthetic issues. “Aesthetics specify graphic properties of the drawing that we would like to apply, as much as possible, to achieve readability.”

Manning [23] has formulated some possible criteria of drawing a “good” graph, for example, these criteria lead to increase uniformity of edge lengths: Minimizing the maximum edge length, maximizing the minimum edge length, minimizing the sum of edge lengths, minimizing the number of distinct edge lengths. In this research, the researcher has tried to create an ideal graph drawing by selecting and combining a set of criteria. In case of having a small set of data, a set of criteria could generate a unique drawing, but if the data set is complex and big, the probability of occurring conflicts in the given criteria increase. In this case, no graph drawing will be generated. Since there are many criteria for drawing a graph [23] and each one could highlight different specifications of a graph, we must identify which graph drawing criteria is the most important for the specific context and application of the visualization. In this case we can select only some of the criteria which have the most importance in a specific context.

Kaufmann et al. have proposed an idea about the level of importance of different graph drawing metrics [8]. They pointed out the fact that selection of graph drawing criteria is

dependent to the context of application. For example, in case of drawing VLSI circuits, some criteria like minimizing the edge length are more important than others.

There are some commonly accepted aesthetics [24, 25, 26]. Applying these general criteria to a graph drawing, will increase the graph readability. Some of these criteria are as below:

- Minimizing the number of edge crossings [27]
- Minimizing bends [28, 29]
- Maximizing symmetry [30, 31]
- Adjacency of nodes which have a connection [24]
- Drawing edges as straight as possible [24]

In case of occurring conflicts between aesthetic criteria in drawing a graph, we need a procedure to select a set of criteria which make a drawing more readable. There are some researched which attempt to grade these criteria base on the effectiveness on human understanding [32, 33]. Purchase [33], has placed a priority order for the importance of the aesthetics. She has done online experiments and measured the ease of understanding drawings. The result of this research indicates that minimizing the number number of edge crossings, is the most important aesthetic criteria, while minimizing bends and maximizing symmetry have less effect on human understandability [33]. Despite the fact that this type of grading could be helpful for building effective drawings, there is still a need for adding the concept of

context or application domain for prioritizing aesthetic criteria. In this thesis, specific criteria for drawing graph of data in the context of transportation systems will be proposed.

### Constraints

Semantic features can be expressed by means of constraints on the drawing.” [34]. Based on this fact, for features that require knowledge about the meaning of the drawing, we can define constraints for graph drawing. As an example, if a set of nodes has more importance than others in a specific application domain, it would be more readable and meaningful to use constraints that place this set of nodes in the center of drawing. Tamassia et al, have classified different type of constraints [34]:

- CENTER: place a set of specified nodes in the center of the graph
- DIMENS: assign the dimension of the symbols representing given nodes
- EXTERN: place given nodes on the outer boundary of the graph
- NEIGH: place a group of nodes as close together as possible
- SHAPE: draw a sub graph with a pre-defined shape
- STREAM: place a sequence of nodes on a straight line

This thesis contributes to this concept by defining the most important constraints in transportation system application domain and uses these constraints to achieve a more readable graph.

### 3.2.2 Graph drawing layout algorithm

There are some general approaches for positioning nodes and edges. The most important and commonly used approaches are listed below [35, 36]:

- Force-base drawing algorithms: In this layout algorithm, the network of nodes and edges are simulated as an electrical system. There are attraction and repulsion forces between nodes and edges. These forces will apply to nodes and the position of nodes will change accordingly. The goal of this algorithm is to reduce the energy associated to graph layout. This algorithm has an iterative nature and it will be repeated until the energy system reaches to a balanced state [30].
- Geographical approach: In this strategy, position of nodes is based on geographical attribute of nodes on a map [37].
- Circular topology: This layout algorithm tries to put nodes on a circle, so the edge crossings will be at the center of the circle [38].

The selection of which approach should be taken is highly dependent on the domain of data to be visualized. For example, in a particular context such as social network, the circular topology is more meaningful than others. In this thesis, the Geographical approach has been used, because of the fact that this approach responds to the requirements of users in the transportation simulation context.

### 3.3 Transportation system visualization

In order to investigate previous works in the area of visualization of transportation systems, different web sites and journals such as IEEE, ScienceDirect, Elsevier, ACM,

and Google Scholar have been reviewed. The keywords used in literature review were “transportation systems visualization”, “transportation chains visualization”, “visualization of transportation simulation” and “visualization of transportation simulation systems”.

There are several research groups working on transportation, such as “Transportation@MIT”<sup>1</sup> and research groups in different universities like Monash<sup>2</sup>

Southampton<sup>3</sup> and Cambridge<sup>4</sup>. Each research group has different approach to transportation, but none of them has focused on applying visualization methods to transportation simulation in order to monitor the effects of input on final results of the simulation.

However, the “Visualization in Transportation Committee”, a sub-committee of the “Transportation Research Board” (TRB<sup>5</sup>), is focused on visualization techniques for transportation issues. Most of the studies on transportation visualization are referred to this committee. After reviewing of related publications, it became clear that the research goals of the relevant work are different from those of this thesis. Most of previous studies have focused on visualization of the physical products (e.g. a new road or bridge) and the impact of these products on transportation system, whereas this thesis emphasizes on visualization of output data from transportation simulation systems. To put it simply, the focus of the mentioned research groups is visualizing and positioning the transportation related products such as bridges and roads to facilitate decision making by stakeholders. In fact, they try to show the picture of improved situation to stakeholders and visualize the outcome of their decision. But this thesis tries to improve readability of information related to transportation simulation results.

Hixson has classified different applications of visualization in transportation [39]. Some of them are listed as below:

- Design
- Identification and evaluation of human factors
- Analysis of outcomes
- Getting approval from stakeholders and relevant agencies
- Security issues

According to Hixson [39], the main uses of visualization are in the design process and it aims to facilitate the work of designers and getting public approval to reduce the rework costs.

Although the approach of using visualization in transportation for design purposes drawing. These conventions are tightly coupled with the context and area of application. One example of a drawing convention could be Orthogonal Drawing in which each graph graph is drawn as a polygonal chain of alternating horizontal and vertical segments (Figure 3).

Kaufmann et al. [8], have described the connectivity of drawing conventions to the traditional drawing style in a specific context. Based on this research, it is important to consider the traditional way of visualizing data of a specific context. For example, it would be more meaningful to draw a graph in the context of electrical circuits, in an orthogonal

shape [8]. Because of the fact that this thesis is about graph visualization of data in a specific context; transportation systems; the above concept is very relevant and useful for this thesis. We can adopt some special conventions for a specific application based on the user requirements. For example, in visualization of transportation systems we can make a rule for representing all the nodes of graph with an icon relevant to their type is also a goal of TAPAS, this approach is relatively different. In the above research, the use of visualization in design process is limited to showing the anticipated final product (e.g. a road, a bridge) with respect to architectural factors as input, while in this the aim of visualization is to show the impact of applying different measures in transportation process, such as taxes in a visual model.

Although visualization has been used in transportation field for many years and several studies have been done for improving and customizing visualization techniques for transportation goals, there is still a lack of research for application of visualization in planning process [40]. Use of visualization in planning process, can cover human and environmental aspects of transportation projects as well as economic impacts. Hughes [40] believes that the research which has been done in this field so far, mainly concerns how to involve stakeholders in design process and investigate what are effective ways of demonstrating design alternatives to different stakeholders. He argues that there is a need to apply visualization techniques in transportation by considering environmental perspectives in planning process [40]. This thesis is focusing on the impact of transportation decisions on environmental aspects and visualizing the results.

There is other research which describes the benefits of applying visualization in the transportation field [41]. In this research, one benefit of visualization is introduced as a complementary way for ensuring achievement of achievements with planned improvements.

Hannon [42] has introduced another concept in visualization named "Virtual Reality". He has made a comparison between simulation and VR. Simulation is defined as a process in

which a model of a real system is created in order to discover the effects of changing system's variables on the whole system and on the future system's behavior. In this research, virtual reality is defined as a technology which output has these characteristics [43]:

- Computer generated
- 3D model.
- The ability to interact with user and the user must be able to navigate the system and feel the effects of variables in a real simulated system.

#### 4. Empirical Case

In this section, we will study the TAPAS case as a real implemented transportation simulation system. First of all, the TAPAS system as the unit of analysis is introduced. Secondly, with regard to RQ1, the visualization requirements of TAPAS users will be identified and analyzed. Next, the potentials of visualization techniques to cover identified user requirements will be discussed. At the next step, the functionalities and features of a TAPAS visualization tool will be elaborated (RQ2). Finally, results of the construction phase will be represented (RQ3). The process for designing the visualization tool is based on the "Systems Engineering" process [48, 49]. We have followed "System Development Life Cycle (SDLC)" as a life cycle of system development; in which we have analysis, Design, Implementation, Testing and Evaluation [50]. Furthermore, we have used the "Spiral Model" for the migration from requirements to design solution [51]. This selection made due to the fact that we have created prototypes of the system to get user feedback during development process. This method has an iterative nature and we repeat each iteration until the final product is made. We have used a simple summarized model of spiral method for this project, which is showed in Figure 7. We have selected this method because this project is a small project and it is not possible and meaningful to follow all the detailed steps of the complete spiral model. As an example, it is not necessary to do risk analysis for a small project. The original model of spiral method which is illustrated in Figure 8, displays the iterations of the "Spiral Model" [51]

Paul Davidsson, Annabellazloconsole@maimbo University  
Master Thesis Project 30p, Spring 2017 August 2017  
E-mail: Banafsheh.Hajinasab@maimbo University

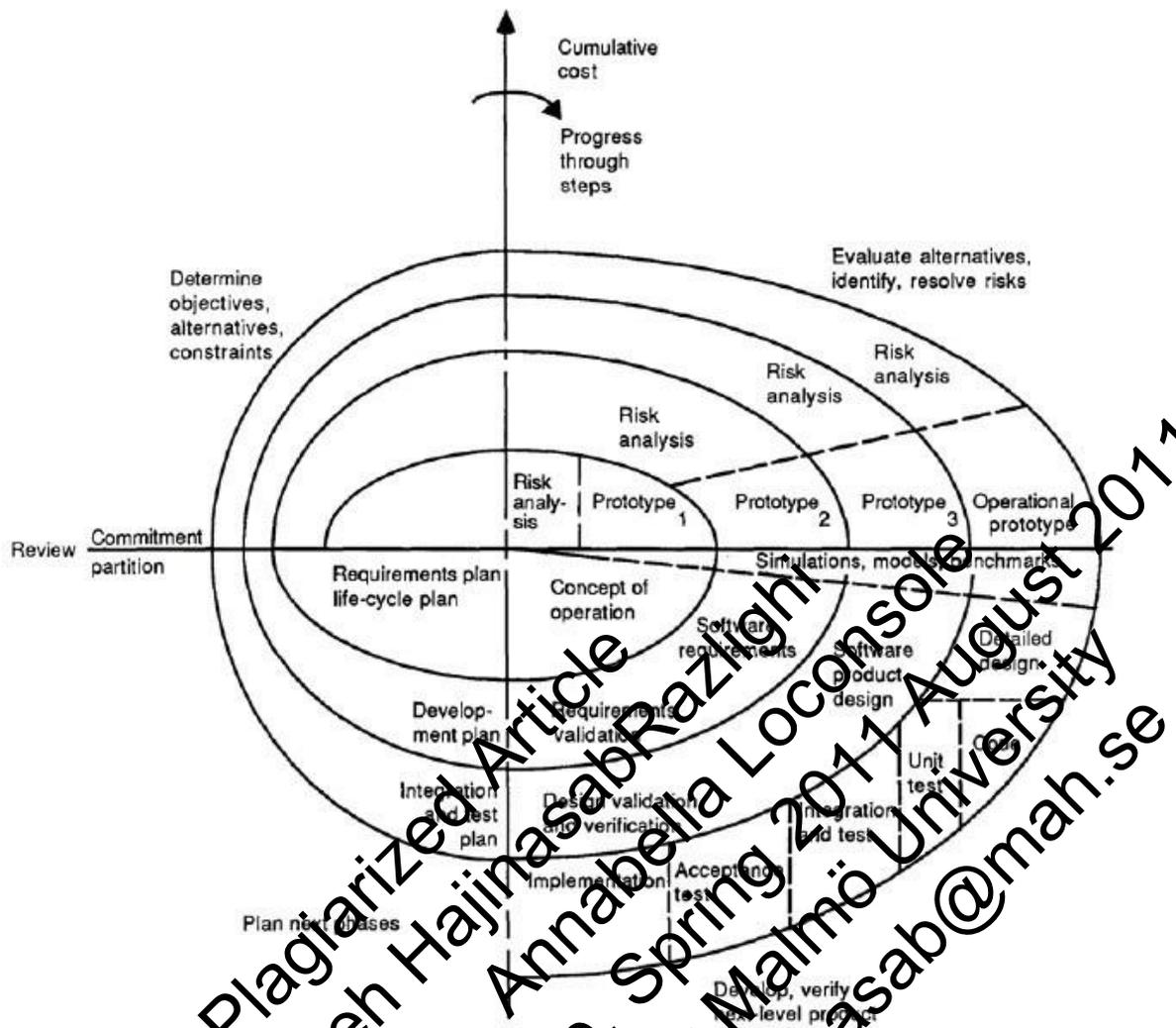


Figure 8: Original spiral model for software development (Taken from [51])

### TAPAS

TAPAS is a Multi-Agent-Based Simulation (MABS) model which aims to analyze the effects of applying production and transportation measures on supply chain activities. There are different types of measures which can be considered as input to TAPAS. TAPAS can be used to predict the consequences of various measures for decision and policy makers, before making any decision. Some of these measures are listed below [52, 4, 10, 54, 55].

- Governmental control policies, such as taxes and fees.
- Infrastructure investments, such as investments in a new road or railway.
- Strategic business related strategies, such as changing the location of a factory.

The output of TAPAS shows how supply chain actors are assumed to react to above measures. So, the users of TAPAS can interpret the results and make more informed decisions and avoid undesired effects. The main outputs of TAPAS simulation are described below:

- Economic effects, such as costs for fuel and taxes
- Logistic effects, such as route and mode choices
- Environmental effects, such as different type of emissions

It is important for users to compare the output results of applying different measures (scenarios).

A TAPAS simulation scenario consists of a set of transportation and production actions, and a set of measures which are applied to the simulation. Thus, each scenario studies the effects of a certain set of measures on the supply chain. TAPAS can have two types of users. The first group is public policy makers, who are interested in predicting the impact of their decisions in approaching their governmental goals, e.g. emission targets. The second group can be private companies and enterprises and other stakeholders of transportation and production, who want to enhance their competitive advantage among other companies. Different type of users may intend to apply different measures. For example, public authorities are more interested in the first group of measures, i.e. governmental control policies, while enterprises might be more eager to predict the effects of their business related decisions, such as changing the location of a factory, on their profit. However, enterprises might be also interested in monitoring the effects of a new transport policy [53, 54]

#### 4.1.1 Definition of TAPAS terms

There are some terms which are used often in the TAPAS documentation, and therefore in this research.

- **Node:** A node can be different kinds of actors in the transportation chain, for example: customer, producer (a factory), or a connection point (terminal) [53]
- **Transportation mode:** A transportation mode represents

the type of transportation (e.g. road, rail, or sea)

- **Link:** A link is a directed connection between two nodes. A link can have exactly one mode of transportation. However, we can have several links between two nodes which differ in their transport mode.

### Requirements analysis

In this section, the efforts which have been done to address the first research question of his thesis; which is: "What are the requirements of users who want to view the results from a transportation system simulator in a visual representation"; are presented. In order to answer RQ1, it is necessary to understand the problems of users regarding readability and usability of TAPAS output, and how a visualization tool could help them. In addition, we need to know the type of target groups who potentially can use TAPAS, because their needs should be considered in the design process of the visualization tool of TAPAS. This can be achieved through a requirement analysis. In order to gather requirements of each target group we conducted interviews with TAPAS users, who are mostly expert in the field of transportation. The first part of this chapter classifies and describes different type of user groups. In the second part, we describe the method of data gathering for each target group. Conclusion of the user requirements will make the last part of this chapter.

### Classification of TAPAS users

We can categorize potential users of TAPAS into two main sections: Public authorities and large companies [54]. They are users of TAPAS simulation system in general. But for this research's purpose, which is developing a visualization tool for TAPAS, we can consider that we have another type of users who are developers of TAPAS. They can be also considered as TAPAS users, because of the fact that they are entering various scenarios to the system and run simulation several times to generate simulation results.

Hence, we can categorize TAPAS users from another perspective:

- Expert users
- End users

Expert users are analysts and researchers who work with the simulation system and manipulate the simulation results. End users are decision makers who can use simulation results to improve their work. Both decision makers in public authorities and large companies are considered as end users in this case. Some part of the end users can be the users who do not work with the visualization tool, but they use the output of visualization tool. This group can be called off-line users, who just see the output of visualization tool as screen dumps or statistical charts. We consider that this type of users have the same requirements as end users.

### Data gathering method

For answering RQ1, it is necessary to analyze the user requirements. Each type of users has their own concerns and expectations from the system. In order to identify user needs, two interviews with expert users of TAPAS have been done. The interviews aim to find the user requirements and discover probable problems which TAPAS users face without visualizing the results. The interview questionnaire and the answers are enclosed to this report as Appendix I.

Furthermore, we have designed a web-based questionnaire for potential end users to understand their special needs and interests of a transportation simulation visualize. The screenshots of this questionnaire are attached to this report as Appendix II.

TAPAS has different output that can be visualized [3,55]. One of the questions of the interview is focused on identifying the interests of users regarding visualization of output. The interviewees suggested some of the output which is important to display in a visual form. Both of interviewees have the same idea about visualization of output, but they expressed them in different languages. The important output to be visualized can be categorized as below:

- 1) Route choices (percentage between different routes)
- 2) Mode choices (percentage between different modes)
- 3) Transportation per link (e.g. number of vehicles per link)
- 4) Share of amount of emissions per route (e.g. CO2)
- 5) Transport cost per route
- 6) Different types of other costs (per route)

As a conclusion of the interviews, the author found that there is an inevitable need for visualization in TAPAS to facilitate the process of analyzing simulation results. Furthermore, it is now clear what requirements of the users are with respect to visualization of output.

### Interview with end users

End users can be defined as decision makers who are working in the transportation field. We designed 7 questions and published them through an online survey. This decision was made because the selected users were in different countries and they were not accessible for a face-to-face interview. Also, we should minimize the number of questions to encourage the respondents to answer the questionnaire. Due to the fact that there were just few end users available, we have sent our questionnaire to three users. Although this questionnaire was easy to answer and it provides visual models to make it more tangible, we received only one response. A possible reason for this is that it was summer vacation and perhaps some of the selected users were on vacation.

The purpose of this questionnaire was to find out which type of simulation data the end users are interested in. We used the knowledge extracted from literature review and the previous interview with expert users to suggest possible data which they may be interested in. Furthermore, we have put some screenshots from our design solutions to get their opinion regarding their preferences to see information of simulation and how the final visualization tool should look like.

From this questionnaire, we extracted their interests for visualization of different kind of the TAPAS output. The user is interested in to see route choices, modal choices, and amount of emissions in a visualized model. Also, this user believes that comparing different scenarios of simulation could be an important feature of the system. The user also feels that visualization of simulation results could contribute to his work.

For the next three questions, we provided mock-ups of our

proposed visualization tool which we had made using Microsoft PowerPoint. The aim of these questions is to gain the user's opinion regarding the importance of showing information related to each route or each link in the transportation network. From these questions we realized that the respondent is interested to have information related to each route.

### Conclusion of user requirements

**Req1. Showing input and output of simulation:** We can conclude from the interviews that both expert and end users are interested in a visualization tool and that they think it could facilitate their work. Also, they need to see which input has resulted in a specific output; just visualization of simulation output will not be meaningful for them. Thus, we conclude that we should visualize the input related to each output as well.

**Req2. Showing different scenarios concurrently:** Both types of interviewees would like to be able to see several different scenarios at the same time. They found it useful to compare various scenarios and their related output. Also, they think this feature can help them in evaluating the output results and facilitate the decision making process.

**Req3. Showing simulation information based on links and routes:** Furthermore, interviewees had some suggestions for which output is more meaningful to be visualized or which not. We can also conclude that they need to see information related to both links and routes. In the next section, in order to understand how visualization techniques could help representing simulation results effectively, we correlated visualization techniques with important information for users in the next section.

### Visualization Techniques versus Requirements

The strategy of answering the second research question (RQ2); which is "What are the important features for increasing readability and usability of a visualization in the context of transportation systems?"; is "correlating the user informational requirements from interviews with visualization techniques from literature review. Thus, after studying the current state of the art of visualization algorithms and by considering the points extracted from interviews, we have decided on which visualization techniques can be

helpful for representing TAPAS outputs. Based on the results, we have proposed a mapping for visualization of different outputs.

This mapping relates the interview results to possible visualization techniques. To put it simply, after analyzing interview transcripts, the different types of TAPAS users and their special requirements regarding visualization of output are determined. Based on the knowledge obtained in the literature review about visualization techniques, the mapping can be completed by relating each user need to a suitable visualization technique. The proposed mapping contributes in the design of the visualization tool. This mapping is presented in the table below (Table 2). This table shows how we can use visualization techniques for addressing the visualization requirements of users.

As stated before, the spiral model is used in this thesis to migrate from requirements to design solution. As a part of this migration, several mock-ups of the tool have been prepared and shown to the users. Based on the user feedbacks, this mapping is revised several times to get the most appropriate visualization solution. Therefore, we have started with an initial simple mapping and we have revised the mapping after getting the user opinions.

In this mapping, each "\*" describes how the user requirements (each column) could be satisfied by the visualization techniques (each row). So, the existence of a "\*" in a cell means that the visualization technique is selected and approved by the user to cover the specific user requirement.

This mapping is a brief illustration of design decisions in response to the user requirements, which is used in the development phase. In fact, this mapping provides a guideline for understanding which visualization technique should be implemented for satisfying user requirements. Furthermore, we have used this mapping for design of the mock-ups.

Furthermore, we have concluded from literature review and user requirements to display all the informational requirements of users on a graph. Also from literature review we found out that a good solution for visualizing a transport network, could be displaying a map as background.

Table 2: Mapping user requirements to best visualization solution

User Groups	Expert and end users informational requirements (Req3)				
	Route Choice	Mode Choice	Transportation per link	Emissions	Transportation Costs
Using colors					
Using pie chart				*	*
Thickness of edges	*	*	*	*	*
Highlighting edges	*				
Using different edge styles		*			
Displaying data on edges				*	*

Furthermore, by analyzing the interview results, we can conclude that there are two ways for presenting information which users are interested in; information on link and information on route. So, we can categorize user requirements with regard to visualization in two parts.

- Link based view
- Route based view

The information is mainly connected to each route, while displaying some information for each link can be also

meaningful.

### Features and Functionalities of TAPAS Visualization tool

Because of the fact that the design decisions could not be discussed with users just by explaining theories, we have designed several mock-ups of the visualization tool. A mock-up could help users to imagine the real system and facilitate the process of designing functions and features of the visualization tool.

Thereby, we designed mock-ups by Microsoft PowerPoint in several steps and discussed with expert users and end users to get their feedback.

The sketch shows the different techniques of visualization, which can be used to show different transportation modes between two nodes. This sketch is illustrated in Figure 12. This sketch is presented to the expert users and they suggested taking the second approach for demonstrating different transportation modes.

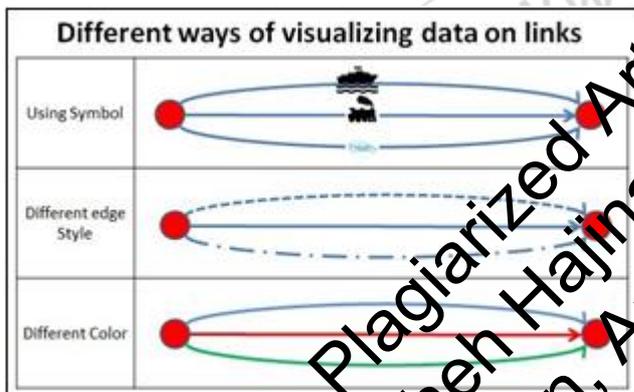


Figure 12: First sketch of proposed Tab View for visualization tool of TAPAS

First of all, for displaying the route choice, the best solution based on literature review and user feedbacks is using graph. To put it simply, the proposed solution displays each customer and producer as a node of transportation graph. For illustrating the percentage of using each route (e.g. the amount of ton/route), the solution could be drawing edges between nodes with different thicknesses. For example, if a specific route has been used more, its edge in the graph will be drawn thicker. As a result of this visualization feature, the user of TAPAS can easily find that which route is being used more than others.

As the next step, based on these main features of visualization system, we designed a more complete sketch. We also used this sketch for preparing the end-users' questionnaire.

The user can select a specific scenario at first when the program runs, then the information related to this specific scenario will be shown. In this way, we can visualize the input of simulation. Furthermore, it will be possible to make a comparison between different scenarios.

The following figures are illustrating the final sketches for showing the cost information on link and route as well as the input of simulation (Figure 13, Figure 14). The other

information; CO2 and amount of transport work; have the same layout but with different data.

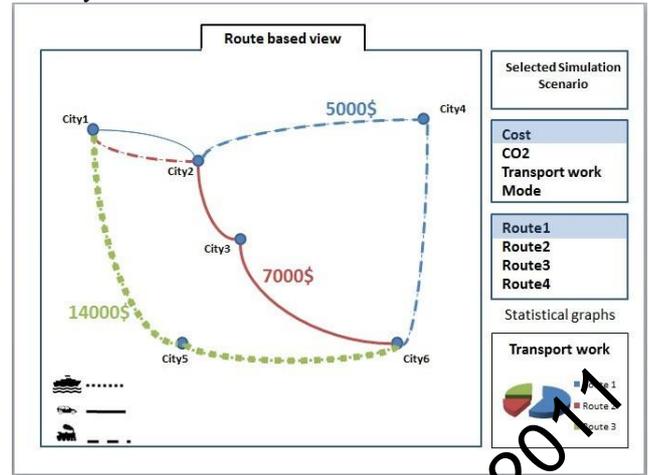


Figure 13: The final sketch which shows the cost information for each link and input of simulation

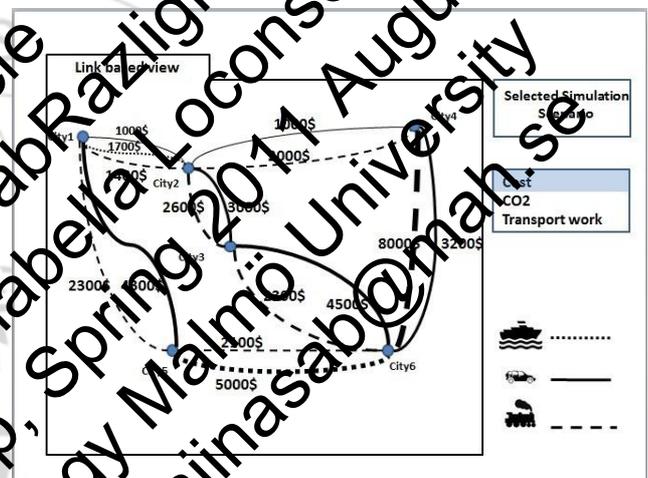


Figure 14: The final sketch which shows the cost information for each route and input of simulation.

Based on user requirements and with consideration to the proposed mapping of user requirements to visualization techniques (Table 2), we proposed to visualize information of each main output (Cost, CO2, Transport work) and allow users to select the desired information from the right side bar in the picture. In Figure 13, the cost information for each link is shown above each link, also we have used different edge styles to show the various transport modes of links. Figure 14, shows the cost information of each route. As we stated in the proposed mapping, we can use edge colors for differentiating different routes. We can also, highlight the selected route. We have a graph in the right bottom corner of the picture, which shows and compares the cost related to all routes.

In general, main functionalities of the system, regarding user requirements and above mapping (Table 2), are designed as below:

**Func1. Show as a graph:** Using color, edge thickness, edge styles to differentiate links; regarding simulation results in response to the Req3.

**Func2. Draw charts:** to illustrate data and make comparison between different scenarios in response to the Req2.

**Func3.Link based information on Graph:** Showing information of cost, emission, amount of transport work considering Req3.

**Func4. Route based information on Graph:** Information of cost, emission, amount of transport work with regard to Req3.

**Func5. Add a map as background:** to increase usability and readability of the graph regarding Req1 and Table 2.

### Visualization tool Construction

The last research question (RQ3) will be approached through designing and developing a complementary tool for TAPAS which aims to visualize the output results in a more understandable way for users. The design of this tool will be based on the results of two previous research questions.

### Development Technologies

As we mentioned before, TAPAS is implemented using the Java language [4, 53], so it would be more integrated to develop this tool in Java language. We have used Swing<sup>8</sup> technology as an Application Programming Interface (API) for designing the user interface. Another option would be to develop the tool on the top of the Google Maps API<sup>9</sup>. The advantage of Google Maps API is that it shows the logistic positions very well and we do not need extra classes and relations with graph drawing libraries anymore. But the problem is we cannot use real positions in a real map for this tool. Because in some cases the positions of two nodes are very far from each other, and if we use real coordinates in a real map, it will decrease readability of the graph. Then we need to customize the map and manage the positions of nodes manually and place them in a way to increase graph usability and readability. The major problem of using Google maps API is that, it contains too much information, which makes it difficult for user to see the simulation results on the map. Furthermore, there are some limitations in using Google Maps API for this tool, for example, we need to bring other classes and libraries for drawing charts and other parts of the tool, which is not possible by using Google Maps API.

Furthermore, because of the nature of this tool, we need to draw graphs and charts. There are some Java components and libraries for the purpose of graph drawing, e.g. UCInet<sup>10</sup> and PAJEK<sup>11</sup>. Each of these mentioned libraries has focused on a specific application area. As an example, UCInet is mainly used for social network analysis. The main application area of PAJEK is also social networks, and it is mainly used for visualizing of very large networks.

For implementing the visualization part of this tool, the JUNG library<sup>12</sup> is regarded as the best solution. JUNG library, Java Universal Network/Graph Framework, is a software library that provides a common and extendible language for the modeling, analysis, and visualization of data that can be represented as a graph or network [56]. Moreover, the JUNG Library has implemented some layout algorithms and uses these layouts to show data in a graph. These layout algorithms are basically mathematical-based algorithms and each of them is focused to solve some specific problems [7]. This selection is motivated by several reasons: First of all, it is well documented and it is easy to download and use; secondly, it is free and it has not any license; thirdly, JUNG has a very good support and it has an "Open discussion forum" and a "Support mailing list", which enables the user to ask questions regarding JUNG and get response very soon.

For the database connection part of the tool, we use PostgreSQL<sup>13</sup>. This decision is made because of the fact that the TAPAS database is made by PostgreSQL. Thus, for designing desirable queries, we have used SQL statements. In order to explore the database, we have used pgAdmin III<sup>14</sup>, which is a PostgreSQL administration and management tool. The selected IDE<sup>15</sup> which is used for developing this tool is NetBeans IDE<sup>16</sup>. This Selection is made because it has a professional Swing GUI Builder<sup>17</sup> which makes it easier to design GUI<sup>18</sup>.

Also, we have selected JFreeChart library<sup>19</sup> for drawing charts and diagrams, because this library contains all type of charts (e.g. Bar charts, Pie charts, Scatter plot charts,...) which can be used in future improvements of this tool and support more types of charts for users. Also, this library is a free open-source java chart library and it is well documented.

### Investigate Current System

To build and develop visualization functions, I investigated the data simulation application software and database. According to my investigation the current database contains just the final results of running different simulation scenarios. In order to visualize information for different routes and links we need detailed intermediate information of cost, emission, and load (transport work) for each link between nodes. Therefore, we decided to design a new table to develop required functions because they could not be implemented by using current database. System. In addition, the new proposed table which contains required information for the visualization tool is illustrated.

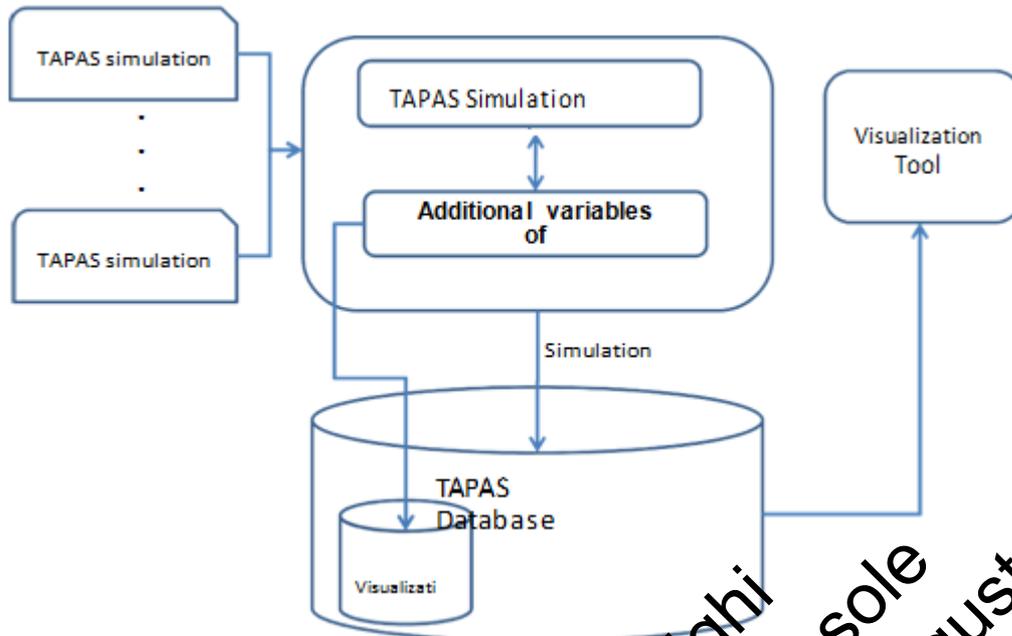


Figure 15: The proposed temporary database schema and its relation with simulation system.

### Developing the Visualization tool

In this section, we will describe the process of developing each functionality of the visualization tool:

#### Func1. Show as a graph:

For developing this functionality, we used D3.js library and we draw the graph based on the information from the database.

We have designed several queries to get desired data and put in the graph. We draw the graph based on the retrieved information. For drawing each aspect of the graph, such as edge thickness, or edge style, a suitable SQL query is designed. Based on the query results we draw the parts of the graph. In some cases we need several nested queries to retrieve desired information. As an example, one of the queries for retrieving the edge styles of each link is described in Appendix III, Part 1.

After designing and running queries, we show the new

information on screen. So, we use the appropriate renderer to update the forms of the visualization tool. For example, in the case of setting the appropriate edge style based on transportation mode, we use the "EdgeStrokeTransformer" renders. An example of a renderer is illustrated in Appendix III, Part 2.

#### Func2. Draw charts

To implement this functionality, we have designed the first form of this tool in a way that the user can open several windows for each scenario. It means that, at the first form of visualization tool, the user can select the scenario name, and then the second form which contains the related scenario will be opened. In this way the user can go back to the first form and select another scenario, while the first selected scenario is still open, and make a comparison between the charts of two or more scenarios. The screenshot of this first form is illustrated in Figure 17.



Figure 17: First form of visualization tool; selecting scenario name

#### Func3.Link based information on Graph:

For illustrating the information on each link, we also, designed several queries and retrieved appropriate data.

Then, we send this data to the Graph Listener to show them in the graph. This Listener will be activated whenever the user pushes the related button of each data to show related

information on each link. One screenshot of the Link Based View is shown below (Figure 18)

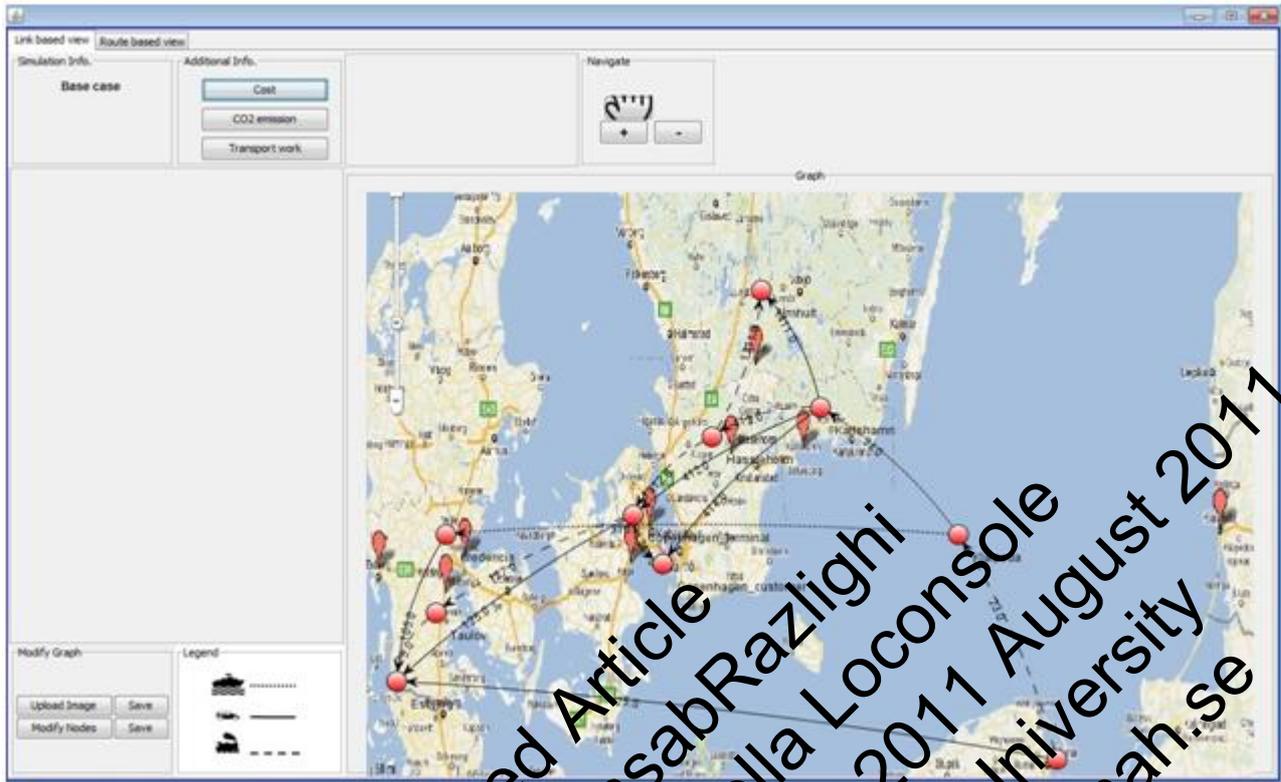


Figure 18: Link Based View form of visualization tool; pushing the cost button.

**Func4. Route based information on Graph:**

For developing this feature we designed more complex queries to retrieve suitable data on each route. Then, we push this data to the Graph Listener to show them in the graph. This Listener will also be activated whenever the user pushes

the related button of each data to show related information on each route. One screenshot of the Link Based View is shown below (Figure 19). Also the charts will be drawn in the Route Based View because the charts are supposed to compare the information for all routes of a scenario to each other.

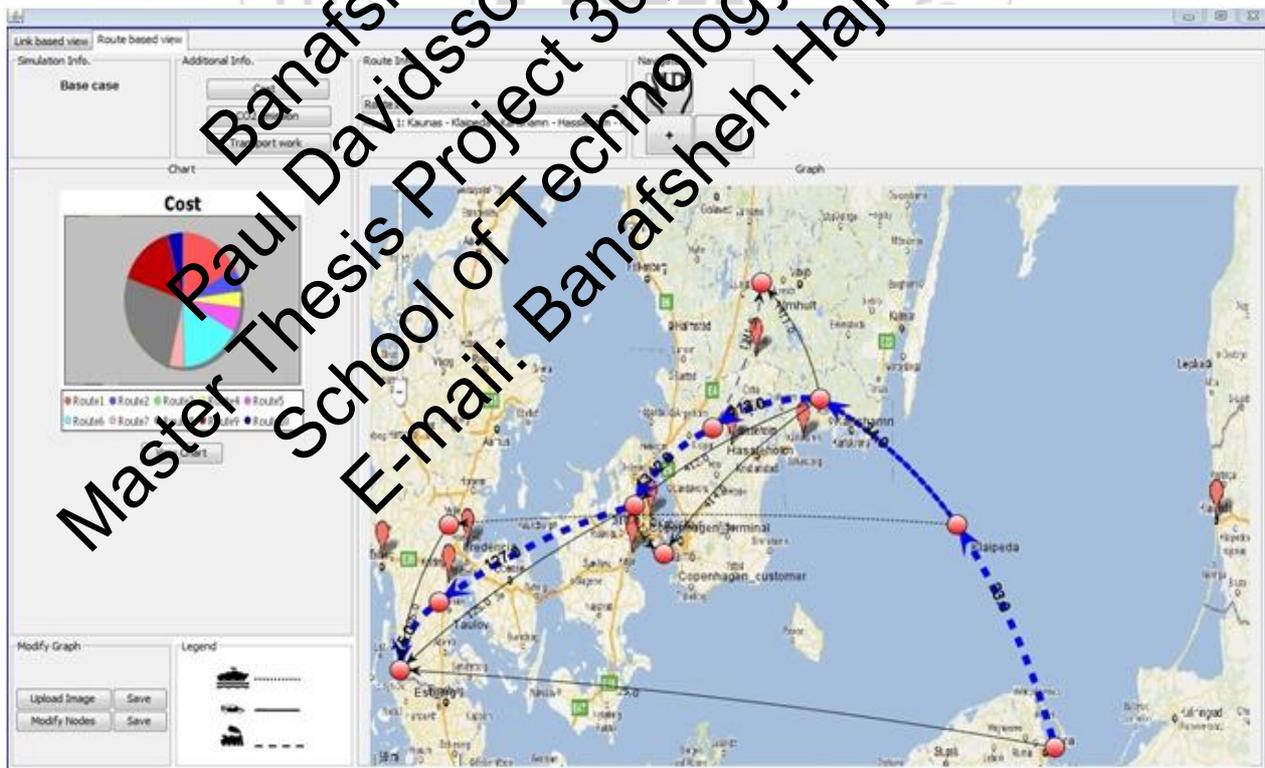
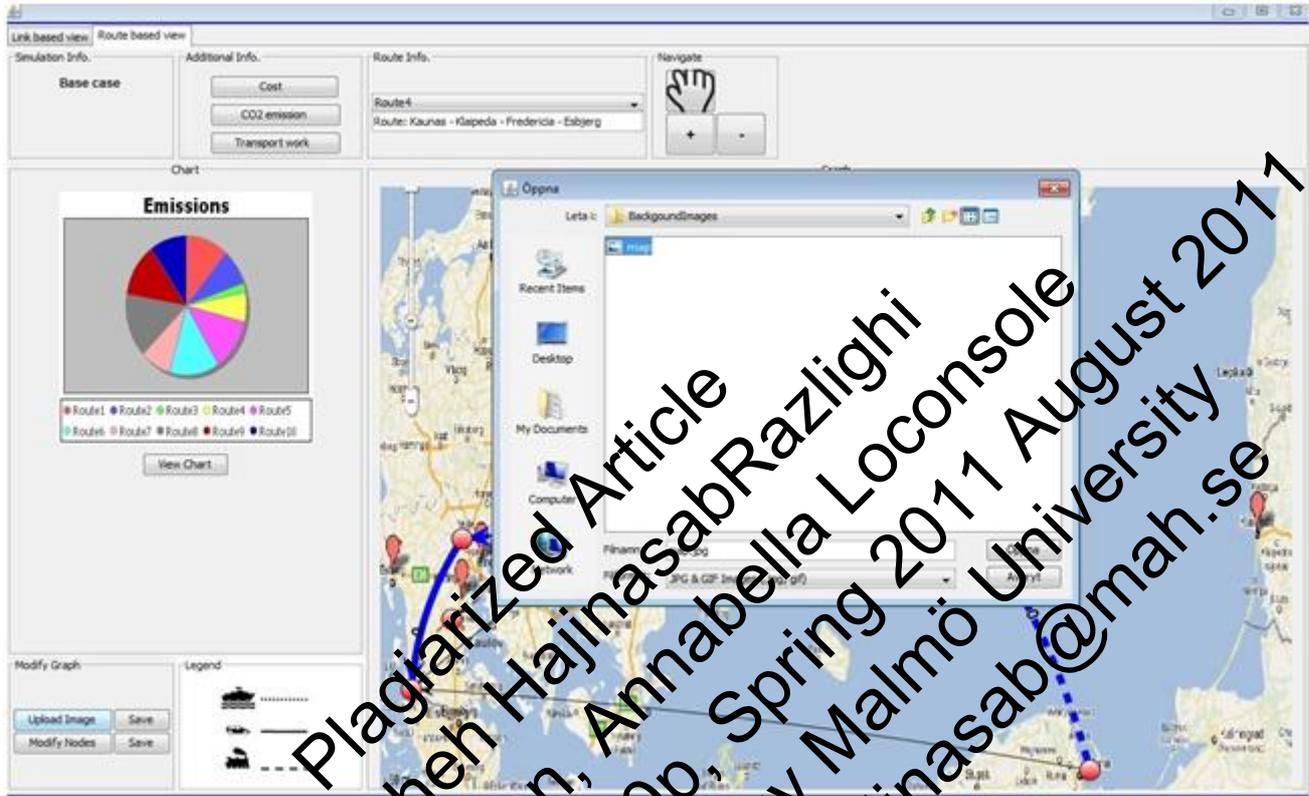


Figure 19: Route Based View form of visualization tool; pushing the cost button, viewing the related chart

**Func5. Add a map as background**

This feature is implemented by adding an ability to the tool which enables user to select an image and upload it. Then for the next run of the visualization tool, the tool will show the right background which is selected and uploaded by user for each scenario. The benefit of this feature is that in case of existence of two nodes which are too far from each other, the user can customize the map and cut the empty gaps between

the nodes and upload it as a new background for the current scenario. This tool has also this capability to move the nodes and change their location on the map. In this case, the user can upload a map and check if the positions of nodes on the map are suitable or not. Then, he/she can move the nodes until reaches a desirable position. Figure 20 shows the file uploading window of the tool.



**4.6 Evaluation**

To evaluate the quality of answers to the research questions we have to investigate validity of findings and reliability of selected methods. In other words, validity means that we have found correct data, and reliability of our research methods could be proved by repeatability of research by other researchers. To increase validity of research we have selected different methods to collect data and observe facts. We have done several interviews with users, experts, and the TAPAS development team. Also, we reviewed the published documents of TAPAS; moreover, the real data of database was reviewed. Therefore, we tried to use triangulation to discover facts in our research. To increase reliability and repeatability of research, we tried to describe all steps and activities in this report which could be used by other researchers. Also, since a considerable part of our research was developing the visualization tool, we have conducted black box and white box testing which help us to evaluate technical aspects of the developed tool.

**5. Conclusion**

As we discussed before, visualization techniques could be used to increase the readability and usability of transportation simulation systems. To investigate how visualization could be applied to transportation context, we studied TAPAS as a

real case and tried to specify user requirements related to visualization, design appropriate functionalities, and develop a visualization tool for TAPAS. Our findings from this case study are concluded for each research question as below:

**RQ1:** What are the requirements of users who want to view the results from a transportation system simulator in a visual representation?

**Req1. Showing input and output of simulation:** Interviews and results of questionnaires show that users of TAPAS are interested in seeing both outputs and inputs of each simulation in the same view.

**Req2. Showing different scenarios concurrently:** In order to compare various scenarios and their related output, users wanted to be able to see results of different simulation scenarios concurrently.

**Req3. Showing simulation information based on links and routes:** users need to see important results of simulation information (cost, emission, and transport work) based on different routes and also based on links between nodes.

**RQ2:** What are the important features for increasing readability and usability of a visualization in the context of transportation systems?

By mapping visualization techniques with informational requirements of users we extracted important features and functionalities for enhancing readability of simulation

outputs as follows:

- Func1. Show as a graph: Displaying destinations, departures, and routes in a graph. Using color, edge thickness, and edge styles to differentiate links; regarding simulation results in response to the Req3.
- Func2. Draw charts: Using charts to illustrate important results of system and make comparison between different scenarios in response to the Req2.
- Func3. Link based information on Graph: Showing information of cost, emission, amount of transport work considering Req3.
- Func4. Route based information on Graph: Information of cost, emission, amount of transport work with regard to Req3.
- Func5. Add a map as background: to increase usability and readability of the graph regarding Req1 and Table 2.

## 6. Acknowledgement

- This thesis would not have been possible without the valuable contribution and guidance of certain people. First and foremost I offer my sincerest gratitude to my supervisor, Professor XUAN ZHAOCHENG, who has supported me throughout my thesis with his patience and knowledge. Secondly, I would like to express my gratitude to all my teachers, for their guidance and precise comments.
- I would like to thank all the teachers of international office for their support. I am also grateful to my parents who have made available their support throughout all my studies.
- Last but not least, I would like to show my deepest gratitude to my classmates and friends for their unconditional love and encouragement in all my pursuits.

**RQ3:** How can visualization techniques be implemented to meet the requirements of the user?

To answer this question, a visualization tool has been developed. The visualization techniques have been used in order to develop the functionalities of the tool.

As an overall conclusion; according to our findings from the TAPAS case; first of all, visualization related requirements of users are identified and analyzed for each group of users. Then the visualization capabilities are applied to each requirement. In the next step, the main features and functions to visualize outputs of system are designed. Next, the current system is studied and evaluated in terms of reusability of database or software functions in the visualization tool. Next step was selecting appropriate technologies for development. Finally, visualization tool is developed and it will be evaluated by system users to ensure that the developed tool could satisfy their requirements.

## Appendix I: Interview with expert users

### 1. Interview questions

As a part of this research, an Internet-based interview is conducted with two of the TAPAS users. The interview has been done as a structured Internet-based interview with a set of pre-defined questions. Interviewees are expert users of

TAPAS and they are also the designers and developers of this simulator. Thus, they are aware of all probable problems and deficiencies of TAPAS. The questions are listed below:

- 1) What is the specification of TAPAS output? Which data is important to show as a visual model?
- 2) How you get output data from TAPAS? Is there currently any visualized output or not?
- 3) What are major problems in the process of monitoring the results of simulation which cause difficulties and interrupts in your work?
- 4) How much time you spend as an outcome of above problems?
- 5) Can visualization help you in understanding the results? If yes, what is your suggestion for design of a visualize model?

## Second interview

- 1) When collecting the output from the database it is easy to do mistakes when formulating the scripts of the desired output, etc. It is quite time-consuming to first collect all data from the database, and then make diagrams or tables of the results. Often you need to make the diagrams many times since you discover some small error, or you want to change something, so that also takes a lot of time. There are large amounts of input and output to handle which makes it complex to get an overview of the results and keep track of the causes and effects. If the results could be visualized in an appropriate way this would facilitate the analysis of the results.
- 2) I don't know, but the whole process as it is today of collecting the results is time consuming.
- 3) Yes, the analysis would be facilitated. If the results could be appropriately visualized it would also facilitate for other people to better understand the results, and TAPAS.
- 4) It's difficult to come up with a suggestion of a design at this moment, but some general points are: Also include the different settings for the studied alternatives/scenarios, and also some of the other input, such as possible transport alternatives/routes.

Having a rather simple model probably makes it easier to understand and analyze the results, but it would be good if it is possible to choose to visualize more details.

- If the results could be presented in a graphical sketch of the transport network (possible transport alternatives, the producer nodes and the consumer nodes), it would be easier to get an overview of the studied scenario and the results.
- To easily view the differences (in terms of route choice, modal split, costs, emissions, etc.) between the alternatives/scenarios is important.

## Appendix III: Sample development codes

A part of designed queries of TAPAS visualization tool  
sql = "select start\_node\_id from visualization\_information where simulation\_nr ='" + selectedSimNr + "' group by start\_node\_id";  
Integer startNodeID = null; Integer endNodeID = null; try  
{

```
pst = conn.prepareStatement(sql); rs = pst.executeQuery();
while (rs.next()) {
startNodeID = rs.getInt("start_node_id"); for (int c = 0; c <
counter; c++) {
if (allNodes[c].getNodeID().equals(startNodeID.toString()))
{ startNode = allNodes[c];
}
}
}
```

```
String sql2 = "select end_node_id from
visualization_information where start_node_id = " +
startNodeID
+ " and simulation_nr = " + selectedSimNr + " group by
end_node_id"; pst2=conn.prepareStatement(sql2);
rs2 = pst2.executeQuery(); while (rs2.next()) {
endNodeID = rs2.getInt("end_node_id"); for (int c = 0; c <
counter; c++) {
if (allNodes[c].getNodeID().equals(endNodeID.toString()))
{ endNode = allNodes[c];
}
}
}
```

```
sql = "select transportation_mode_id,cost from
visualization_information where simulation_nr = " +
selectedSimNr + " and start_node_id=" + startNodeID
+ " and end_node_id=" + endNodeID + " group by
transportation_mode_id,cost";
try {
pst3 = conn.prepareStatement(sql);
pst3.executeQuery();
while (rs3.next()) {
Edge drawEdge = new Edge();
transportModeID = rs3.getInt("transportation_mode_id");
drawEdge.setDrawEdgeCost(rs3.getDouble("cost"));
}
```

```
sql = "select name from transportation_mode where
simulation_nr = " + selectedSimNr + " and id=" +
transportModeID + """;
try {
pst4 = conn.prepareStatement(sql);
pst4.executeQuery();
while (rs4.next()) {
transportModeName = rs4.getString("name");
if (transportModeName.equals("sea")) {
drawEdge.setDrawEdgeColor(drawNode.getY());
drawEdge.setDrawEdgeType("SEA");
}
if (transportModeName.equals("rail")) {
drawEdge.setDrawEdgeType("RAIL");
}
if (transportModeName.equals("road")) {
drawEdge.setDrawEdgeType("ROAD");
}
g.addEdge(drawEdge, startNode, endNode,
EdgeType.DIRECTED);
}
} catch (Exception e) {
JOptionPane.showMessageDialog(null, e);
}
```

```
}
} catch (Exception e) {
JOptionPane.showMessageDialog(null, e);
}
```

```
}
}
} catch (Exception e) {
JOptionPane.showMessageDialog(null, e);
}
```

**A part of designed renderers of visualization tool**

```
BasicVisualizationServer<Node, Edge> vv = new
BasicVisualizationServer<Node,Edge>(layout);
vv.getRenderContext().setEdgeStrokeTransformer(
new Transformer<Edge, Stroke>() { float dash[] = {2.5f};
final Stroke SEA = new BasicStroke(1.0f,
BasicStroke.CAP_ROUND, BasicStroke.JOIN_MITER,
1.0f, dash, 0.0f);
float dash2[] = {10.0f};
final Stroke RAIL = new BasicStroke(1.0f,
BasicStroke.CAP_BUTT, BasicStroke.JOIN_MITER,
10.0f, dash2, 10.0f);
final Stroke ROAD = new BasicStroke(1.0f,
BasicStroke.CAP_BUTT, BasicStroke.JOIN_MITER,
10.0f, null, 0.0f);
public Stroke transform(Edge e) {
if (e.getDrawEdgeType().equals("SEA")) { return SEA;
else if (e.getDrawEdgeType().equals("RAIL")) { return
RAIL;
} else if (e.getDrawEdgeType().equals("ROAD")) { return
ROAD;
}
}
});
```

**References**

- [1] Brunner, P., et al. Toward increased use of simulation in transportation. In proceedings of 1998 Winter Simulation Conference, USA, 1998.
- [2] Ware, Information Visualization: Perception for Design, Morgan Kaufman, USA, 2004.
- [3] Davidsson, P., et al. Multi agent based simulation of transport chains. In Proceedings of the 7th International Joint Conference on Autonomous agents and multi agent systems (AAMAS'08), Estoril, Portugal, 2008.
- [4] Ramstedt, L., Davidsson, P., Holmgren, J., and Persson, J.A., On the Use of Micro-level Simulation for Estimation of the Effects of Governmental Control Policies, 11th World Conference on Transport Research, Berkeley, USA, 2007.
- [5] S.K. Card, J.D. Mackinlay, and B. Shneiderman (eds), Readings in Information Visualization, Morgan Kaufmann Publishers, 1999.
- [6] Herman I, Melancon G, Marshall MS. Graph visualization and navigation in information visualization: a survey. IEEE Transactions on Visualization and Computer Graphics 2000; 6: 24 – 43.
- [7] Di Battista, G., Eades, P., Tamassia, R., and Tollis, I. G., Graph Drawing: Algorithms for the Visualization of Graphs. Upper Saddle River, N. J: Prentice Hall, 1999.
- [8] Kaufmann, M. and Wagner, D., Drawing Graphs: Methods and Models, volume 2025 of Lecture Notes in Computer Science. Springer, 2001.

- [9] Brasseur, Lee E., Visualizing technical information: a cultural critique. Baywood Pub, Amityville, N.Y, 2003.
- [10] M. Bergkvist, P. Davidsson, J. A. Persson, L. Ramstedt, A Hybrid Micro-Simulator for Determining the Effects of Governmental Control
- [11] Policies on Transport Chains, in: Lecture Notes in Computer Science: Vol. 3415. Multi-Agent and Multi-Agent-Based Simulation, Springer, 236–247, 2005.
- [12] Ghoniem, M., et al., On the readability of graphs using node-link and matrix-based representations: a controlled experiment and statistical analysis. Information Visualization, 4 (2).
- [13] Davis, F. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, 13(3): 319-340. 1989.
- [14] Bresciani, S., et al., "Beyond knowledge visualization usability: toward a better understanding of business diagram adoption", Proceedings of 2009 13th International Conference Information Visualisation. Washington, DC: IEEE Computer Society, 2009.
- [15] Di Battista, et al., Algorithms for Drawing Graphs: An Annotated Bibliography. Computational Geometry Theory and Applications. vol. 4. 1994.
- [16] B. J. Oates. Researching Information Systems and Computing. SAGE Publications, 2006.
- [17] March, S.T., et al., Design and Natural Science Research on Information
- [18] Technology, Decision Support Systems (5) 1995, pp. 251-266.
- [19] Yin, R., Case study research; Design and methods. Sage publications, USA, 2003.
- [20] Plaisant, C. The Challenge of Information Visualization Evaluation, Proc. of Conf. on Advanced Visual Interfaces, ACM, New York, 2004.
- [21] Mazza, R., Introduction to Information Visualization Springer, London, 2009.
- [22] C. Chen, Information Visualization: Beyond the Horizon, 2nd edition, Springer, 2004.
- [23] Bederson, B and Shneiderman, B. The Craft of Information Visualization. Readings and Reflections, Morgan Kaufmann, 2003.
- [24] "Graph drawing" Internet. [http://en.wikipedia.org/wiki/Graph\\_drawing](http://en.wikipedia.org/wiki/Graph_drawing) [FEB.20.2011]
- [25] Manning, J.B., Geometric symmetry in graphs. PhD thesis, Purdue University, 1990.
- [26] Sugiyama, K., et al. Method for Visual Understanding of Hierarchical Systems. IEEE Trans. Syst. Man Cybern., SM-11, no. 2, 109-125, 1981.
- [27] Batini, G., et al, What is a Good Diagram? A Pragmatic Approach. In Proc. 4th Internat. Conf. on the Entity Relationship Approach, 1985.
- [28] Purchase, H.C., et al, Validating Graph Drawing Aesthetics, In F. J. Brandenburg, editor, Graph Drawing (Proc. GD '95), vol 1027 of Lecture Notes Comput. Sci., pp. 435-446. Springer- Verlag, 1996.
- [29] D. Ferrari and L. Mezzalana. On drawing a graph with the minimum number of crossings. Technical Report 69-11, Istituto di Elettrotecnica ed Elettronica, Politecnico di Milano, 1969.
- [30] R. Tamassia. On embedding a graph in the grid with the minimum number of bends. SIAM J. Computing, 16(3):421-444, 1987.
- [31] H. Trickey. Drag: A graph drawing system. In Proc. Int. Conf. on Electronic Publishing, pages 171-182. Cambridge University Press, 1988.
- [32] P. Eades. A heuristic for graph drawing., Congressus Numerantium, 42:149-160, 1984.
- [33] R. Lipton, S. North, and J. Sandberg. A method for drawing graphs, In Proc. ACM Syrup. On Computational Geometry, pages 153-160, 1985.
- [34] H.C. Purchase, R.F. Cohen, and M. James. An experimental study of the basis for graph drawing algorithms. A CM Journal of Experimental Algorithmics, 2(4), 1997.
- [35] H. Purchase, Which aesthetic has the greatest effect on human understanding? In G. Di Battista, (Ed.) Graph Drawing 97, 1353 of lecture Notes in Computer Science, Springer Verlag, 284-290, 1997.
- [36] R. Tamassia, et al., Automatic Graph Drawing and Readability of Diagrams. IEEE Trans. Syst. Man and Cybern., SM-18, no. 1, 61-79, 1988.
- [37] Aris, Visualizing & Exploring Networks Using Semantic Subspaces., PhD thesis, Department of Computer Science, College Park, University of Maryland, 2008.
- [38] B. Shneiderman and A. Aris, Network visualization by semantic subspaces. IEEE Transactions on Visualization and Computer Graphics, 12(5), 2006.
- [39] A. Becker, S. G. Lick, G. F. Willis: „Visualizing Network Data“, IEEE Transactions on Visualizations and Graphics, Vol. 4, No. 1, pp. 66-78, 1995.
- [40] B.J. Breitkreuz et al. "Osprey: a network visualization system." Genome Biology 4: R22, 2003.
- [41] CL Hixson III. Visualization for Project Development, . NCHRP Synthesis 361. National Cooperative Highway Research Program. Transportation Research Board, Washington, D.C., 2006.
- [42] Hughes, R.G. Toward an Expanded Research Agenda for Visualization in Transportation: Incorporating State TEA-LU Directives for „Planning“. Institute for Transportation Research and Education, North Carolina State University, Raleigh, NC, 2008.
- [43] Visualization in Transportation: A Guide for Transportation Agencies, AASHTO Task Force on Environmental Design, American Association of State Highway and Transportation Official (AASHTO), December 2001. Internet : <http://design.transportation.org/Documents/VisualizationGuideJuly2003.pdf> ,Accessed [APR.12.2011].
- [44] Hannon, J.J. Emerging Technologies for Construction Delivery. NCHRP Synthesis 372. National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C., 2007
- [45] Sulbaran, T., "Distributed Virtual Reality Scheduling in Future Construction Companies," Proceedings of the 2005 ASCE International Conference on Computing in Civil Engineering, Cancun, Mexico, July 12–15, 2005.
- [46] Hughes, R.G. Freight data visualization: A "Pivotal" point in the development of visualization applications in transportation. Submitted for publication and presentation at "The 2011 Transportation Research Forum", Long Beach, CA, 2010.
- [47] Kutz, M. Handbook of Transportation Engineering,





MALMÖ HÖGSKOLA

School of Technology

Malmö University

Master Thesis Project 30p, Spring 2011

August 2011

## **Master Thesis**

# **Visualization of data from transportation simulation systems**

**By**

Banafsheh HajinasabRazlighi

**Supervisors:**

Paul Davidsson

Annabella Loconsole

**Examiner:**

Jan Persson

## Contact information

*Author:*

Banafsheh HajinasabRazlighi

E-mail: Banafsheh.Hajinasab@mah.se

*Supervisors:*

Paul Davidsson

E-mail: Paul.Davidsson@mah.se

Malmö University, Department of Computer Science.

Annabella Loconsole

E-mail: Annabella.Loconsole@mah.se

Malmö University, Department of Computer Science.

*Examiner:*

Jan Persson

E-mail: Jan.A.Persson@mah.se

Malmö University, Department of Computer Science.

## Abstract

Nowadays by increasing importance of information in all sectors, illustrating data in a communicative format helps decision makers to understand and analyze effectively large amount of information in a short time. Information visualization, as a way of presenting different data types in a more understandable form, is growing increasingly in various areas. This thesis aims to investigate how information visualization could be used to increase readability and usability of transportation simulation data. Most of the previous studies in this area have focused on visualization of transportation infrastructures such as roads, bridges, etc.; while the main focus of this thesis is visualizing the outputs of transportation simulation systems. In order to study the role of information visualization in the transportation systems, we have investigated the visualization applications in a real implemented agent-based transportation simulator called TAPAS as case. In our case study, the visualization related requirements of users were analyzed, and the visualization tool was designed and developed based on the identified requirements.

**Keywords:** Information visualization, transportation simulation visualization.

## Acknowledgement

This thesis would not have been possible without the valuable contribution and guidance of certain people. First and foremost I offer my sincerest gratitude to my supervisor, Professor Paul Davidsson, who has supported me throughout my thesis with his patience and knowledge. Secondly, I would like to express my gratitude to Annabella Loconsole, for her guidance and precise comments. I am indebted to my examiner Jan Persson, whose guidance from the initial to the final level of this thesis enabled me to develop an understanding of the subject.

I would like to thank Dr. Johan Holmgren and Dr. Linda Ramstedt for their input especially in the interviews of this study. They have shared valuable insights in this study. To Blekinge Institute of Technology, thank for giving me the opportunity to work with TAPAS.

I am also grateful to my parents who have made available their support throughout all my studies.

Last but not least, I would like to show my deepest gratitude to my husband Shahram and my daughter Sheida for their unconditional love and encouragement in all my pursuits.

## Table of contents

<b>1</b>	<b>Introduction and background .....</b>	<b>8</b>
1.1	Introduction.....	8
1.2	Purpose and motivation .....	9
1.3	Quality attributes .....	10
1.4	Goal .....	11
1.5	Research question.....	11
1.6	Research delimitation .....	12
1.7	Results .....	12
<b>2</b>	<b>Research Method .....</b>	<b>13</b>
2.1	Literature review .....	14
2.2	Design and creation .....	14
2.3	Qualitative case study.....	14
2.4	Data Collection methods and techniques.....	15
2.5	Evaluation methods .....	17
<b>3</b>	<b>Literature review .....</b>	<b>20</b>
3.1	Information visualization.....	20
3.2	Principles and basic theories of graph drawing.....	21
3.2.1	Graph drawing concepts .....	21
3.2.2	Graph drawing layout algorithms .....	24
3.3	Transportation system visualization .....	25
<b>4</b>	<b>Empirical case.....</b>	<b>32</b>
4.1	TAPAS.....	34
4.1.1	Definition of TAPAS terms.....	36

4.2	Requirements analysis .....	36
4.2.1	Classification of TAPAS users .....	37
4.2.2	Data gathering method .....	38
4.3	Visualization Techniques versus Requirements.....	43
4.4	Features and Functionalities of TAPAS Visualization tool .....	45
4.5	Visualization tool Construction.....	50
4.5.1	Development Technologies .....	50
4.5.2	Investigate Current System.....	52
4.5.3	Developing the Visualization tool.....	55
4.6	Evaluation .....	60
<b>5</b>	<b>Conclusion.....</b>	<b>61</b>
5.1	Future Work.....	62
<b>Appendix I: Interview with expert users.....</b>		<b>64</b>
1.	Interview questions.....	64
2.	Interview transcripts .....	65
<b>Appendix II: Interview with end users .....</b>		<b>70</b>
<b>Appendix III: Sample development codes.....</b>		<b>75</b>
<b>References.....</b>		<b>78</b>

## List of Figures

FIGURE 1: THE RESEARCH PROCESS OF THIS THESIS (BASED ON [15, 17]).	13
FIGURE 2: THE IMPACT OF DIFFERENT CRITERIA ON SELECTION OF RESEARCH METHOD, TAKEN FROM [17]	15
FIGURE 3, TAKEN FROM [7]	22
FIGURE 4: A SAMPLE MAP-BASED VISUALIZATION OF FREIGHT DATA (TAKEN FROM [])	29
FIGURE 5: A SAMPLE SCORE-CARD VISUALIZATION OF FREIGHT DATA BY USING BAR CHART (TAKEN FROM [44])	29
FIGURE 6: A SAMPLE DIAGRAM OF SIMULATION RESULTS WHICH DEPICTS THE DEPENDENCE OF WAITING TIMES ON THE NUMBER OF VEHICLES USED. (TAKEN FROM[46])	30
FIGURE 7: SIMPLE SPIRAL MODEL FOR SOFTWARE DEVELOPMENT (TAKEN FROM <a href="http://www.arctern.com/development_methodologies.aspx">HTTP://WWW.ARCTERN.COM/DEVELOPMENT_METHODOLOGIES.ASPX</a> )	33
FIGURE 8: ORIGINAL SPIRAL MODEL FOR SOFTWARE DEVELOPMENT (TAKEN FROM [51])	34
FIGURE 9: A SAMPLE SCREENSHOT FROM GRACE WHICH SHOWS THE PROCESS OF SIMULATION	39
FIGURE 10: A SAMPLE SCREENSHOT FROM GRACE WHICH SHOWS DIFFERENT BAR CHARTS DURING SIMULATION.	40
FIGURE 11: FIRST SKETCH OF PROPOSED TAB VIEW FOR VISUALIZATION TOOL OF TAPAS	45
FIGURE 12: FIRST SKETCH OF PROPOSED TAB VIEW FOR VISUALIZATION TOOL OF TAPAS	46
FIGURE 13: THE FINAL SKETCH WHICH SHOWS THE COST INFORMATION FOR EACH LINK AND INPUT OF SIMULATION.	48
FIGURE 14: THE FINAL SKETCH WHICH SHOWS THE COST INFORMATION FOR EACH ROUTE AND INPUT OF SIMULATION.	49
FIGURE 15: THE PROPOSED TEMPORARY DATABASE SCHEMA AND ITS RELATION WITH TAPAS SIMULATION SYSTEM.	53
FIGURE 16: THE SPECIFICATION AND FIELDS OF THE NEW PROPOSED TABLE (VISUALIZATION_INFORMATION)	54
FIGURE 17: FIRST FORM OF VISUALIZATION TOOL; SELECTING SCENARIO NAME	56
FIGURE 18: LINK BASED VIEW FORM OF VISUALIZATION TOOL; PUSHING THE COST BUTTON	57
FIGURE 19: ROUTE BASED VIEW FORM OF VISUALIZATION TOOL; PUSHING THE COST BUTTON, VIEWING THE RELATED CHART.	58
FIGURE 20: UPLOADING A MAP AS BACKGROUND IMAGE	59

## List of Tables

TABLE 1: MAPPING RESEARCH QUESTIONS TO THE RESOURCE OF EVIDENCES.	17
TABLE 2: MAPPING USER REQUIREMENTS TO BEST VISUALIZATION SOLUTION	44

# 1 Introduction and background

## 1.1 Introduction

Nowadays, there is a growing use of simulation in transportation [1]. The idea is to attempt to simulate the real system characteristics in a computer system and observe how applying different scenarios will affect the behavior of the transportation system. This can help in planning process where the decision makers wish to be aware of the consequences of a specific decision on the whole system. The main application of simulation in the transportation context can be cost reduction and increasing quality of service. However, not all the transportation simulation systems are necessarily successful regarding reducing costs and increasing quality. There are some factors which determine the degree of success in simulation models. One important factor is the “Ease of use” of the simulation system by different types of potential users [1]. Typically, a simulation system generates a large amount of data as output. This data must be processed and analyzed in order to be useful for different type of users and decision makers. Hence, the way of analyzing and presenting the output data to users, can be a challenge. As we acquire more information through vision than through all of the other senses [2], visualizing the output data of transportation simulation systems can be an ideal solution.

Information Visualization provides visual representations of large amount of data in order to be more usable and understandable for users. Information Visualization could be used in different contexts and for different purposes, such as Network Analysis, Semantic networks and knowledge representation, Cartography, VLSI Circuits design, Genetic maps, etc.

In this thesis, visualization techniques will be applied on output data of transportation simulation systems to increase the understanding of results by different stakeholders. This can facilitate the decision making process for policy makers by visualizing the complex data of simulation results on screen. In particular, the visualization techniques will be applied in an agent-based simulation model for transportation and production called TAPAS [3] (Transportation And Production Agent-based Simulator. TAPAS is a micro-level simulation tool which aims to predict the effects of applying different measures

including governmental policies like taxes, and infrastructure investments and business strategies [4].

TAPAS can act as a decision support system for public authorities and policy makers. They can monitor what could be the effects of applying a specific set of production and transportation measures and whether it contribute to the government goals such as sustainable economic development or not. Furthermore, TAPAS can also help enterprises to make both operational and strategic decisions, for example decisions regarding consignment size and storage locations, etc. [3].

## 1.2 Purpose and motivation

Information Visualization has an important role in different contexts. In fact, it has been used in different fields and it is an expanding area of knowledge [5].

Obviously, the nature of the data to be visualized is different in various application areas. The choice of selecting a set of visualization techniques and applying them to a data set is highly dependent to the context of data set. As an example, in case of being some inheritance relations between the data elements, we can make use of graphs as a way of representing and visualizing information [6]. In visualization of large and complex graph data, some problems and issues will emerge. Visualizing the whole data in one graph might not be useful for the users and also it causes performance problems and increases the complexity and therefore decreases the readability of the result. Also, it would be more meaningful to display some elements in a different color and shape to increase readability of visualized graph.

Also, some layout algorithms have emerged to solve the problems of visualizing complex and large data. Some of these algorithms have focused on aesthetics criteria in the context of Graph Drawing [7], [8]. But still there is a need to mix these approaches and techniques to reach usability requirements of potential users of this specific context. Thus, it is necessary to investigate how visualization techniques can be applied to transportation systems.

There are other kinds of visualization techniques, such as displaying data through diagrams [9]. Based on the nature of data, some data should be visualized by diagrams.

The focus of this thesis is visualization of data in the context of transportation systems. Indeed, the graph visualization techniques are applied and tailored to visualize data related to transportation systems and link connections between different nodes, in a graph-based shape. Also, we applied other data visualization techniques, such as diagrams. TAPAS, which is a multi-agent based production and transportation simulator is used as a case study [4]. The purpose of this simulator is to show how organizations in supply chain might act after applying different governmental control policies and whether the outcomes are desirable and near to their goals or not [10]. Currently this simulator does not show the results of simulation graphically, which would make it easier to understand the results.

### 1.3 Quality attributes

The quality attributes which this thesis has focused on are readability and usability. The readability of a visualization model can be defined as how easy and fast the user can find the information he/she is looking for [11].

Usability is a term which contains several factors. Davis [12] has proposed a model for Information Technology acceptance. He has categorized two main factors which cause users to accept or reject a system. The first factor is the usefulness, which is defined as “The degree to which a person believes that using a particular system would enhance his or her job performance” [12]. This factor can be described by these characteristics [13]:

1. Allows working faster.
2. Improves job performances.
3. Immediate insight.
4. Adaptability to task.

The second factor which called the ease of use is defined as “The degree to which a person believes that using a particular system would be free of effort”. The characteristics of this factor are listed below:

1. Easy to understand.
2. Easy to learn.
3. Content categories relevance.

#### 4. Aesthetic value.

In this thesis we will use the term “usability” which refers to both usefulness and the ease of use of the system.

### 1.4 Goal

This Master thesis project aims to increase usability of transportation simulation system by visualizing the output data from different simulation scenarios. The goal of this project is to support the analysis of how actors in a transport chain act given various measures through visualization of information.

This goal can be achieved by investigation of the methods for increasing readability and usability of visualized output in this specific context.

### 1.5 Research question

The central research question is formulated as follow:

- How can visualization techniques be applied to transportation systems?

Questions that stem from this question are:

- **RQ1:** What are the requirements of users who want to view the results from a transportation system simulator in a visual representation?
- **RQ2:** What are the important features for increasing readability and usability of a visualization in the context of transportation systems?
- **RQ3:** How can visualization techniques be implemented to meet the requirements of the users?

## **1.6 Research delimitation**

This research is limited to the visualization of transportation systems and will only consider the techniques which could enhance usability in this context. Therefore, the other parts of supply chains are out of the scope of this thesis.

## **1.7 Results**

This Master thesis provides guidelines on how output data of a transportation simulation system could be visualized. Moreover, a tool for visualizing extensive amount of data in the context of transportation systems has been developed.

In addition, the final report contains a survey about existing Graph drawing techniques and algorithms. A customized algorithm that suit to the requirements of a complex graph with complex relations and different type of nodes in the context of transportation systems is proposed. This new algorithm covers both aesthetics and readability concerns as well as the usability requirements of users.

## 2 Research Method

This section is aimed to describe the research approach undertaken for this thesis. This section is divided into four sub-sections. In the first sub-section section, the chosen research methodology is described. In the next part, the possible data collection methods are introduced. In the last section, the evaluation methods which are used for the evaluation of this research are described. The figure below (Figure 1) shows the whole research process of this research. This research process is designed based on the findings from [15, 17].

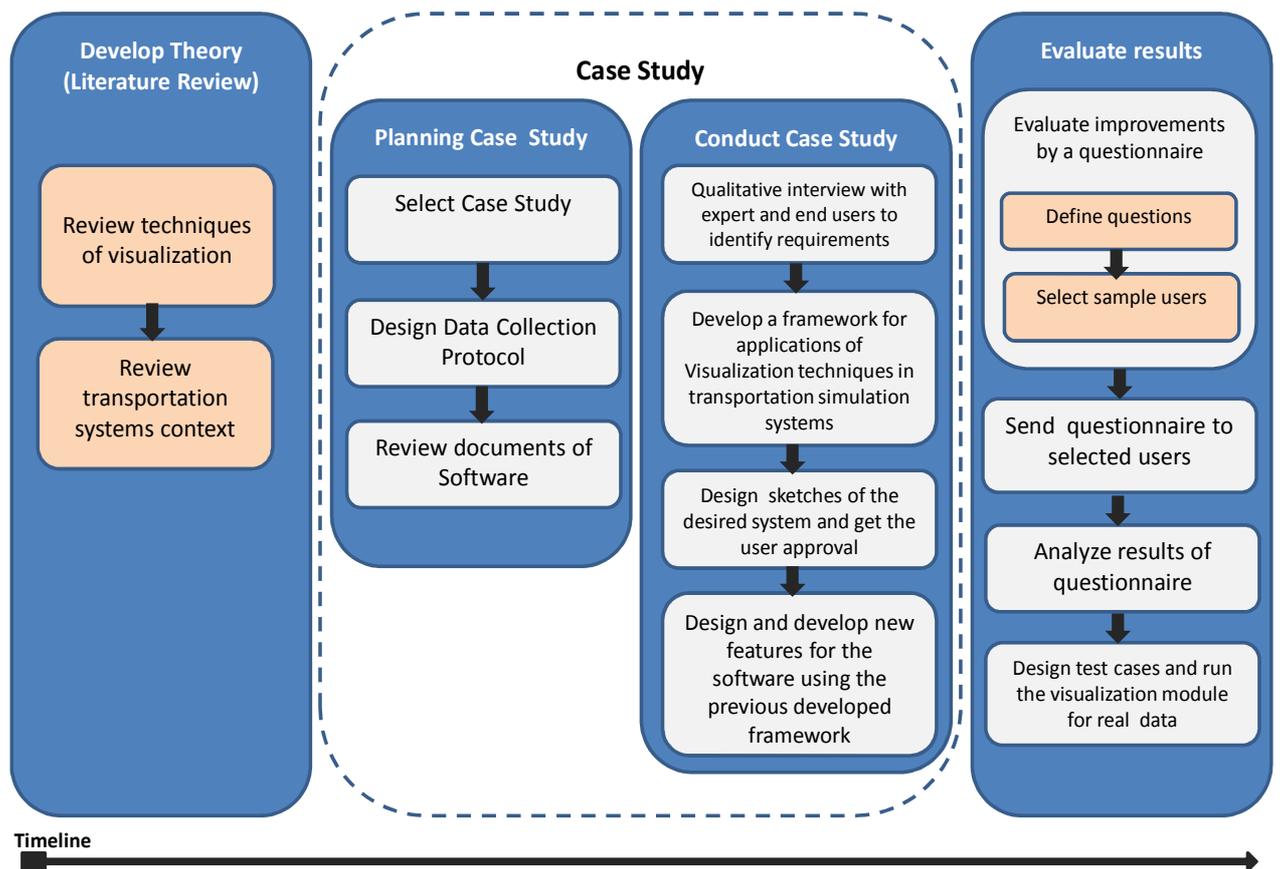


Figure 1: The research process of this thesis (Based on [15, 17]).

## **2.1 Literature review**

Because of the fact that there are a tremendous amount of literature in the fields of Information Visualization and graph drawing algorithms [14], there is a critical need for doing a deep literature review. For this reason, some researchers working on information visualization and graph drawings are identified to use their knowledge published in articles in this field.

## **2.2 Design and creation**

In this thesis, a new IT artifact will be developed, so the main research strategy of this thesis is “Design and creation” [15]. As the result of this research we will have a tool for visualizing the simulation results of TAPAS [3], which currently does not have any visualized output. According to March et al [16], there are four types of IT artifacts: Constructs, Models, Methods, and Instantiations. In this research an instantiation, a working computer-based system, was implemented.

Before implementing the visualization options, some sketches from the proposed system were prepared. These sketches were shown to the users to get the approval and their feedback about design alternatives and options. This could be an iterative process that is continued until receiving positive feedback from the users.

## **2.3 Qualitative case study**

Yin.R, has proposed an idea for selecting the proper research method according to three criteria: (1) the type of research question, (2) whether the researcher has control over behavioral events or not, and (3) focus on contemporary events. His idea is demonstrated in the table below (Figure 2) [17]:

<b>Strategy</b>	<b>Form of Research Question</b>	<b>Requires Control of Behavioral Events?</b>	<b>Focuses on Contemporary Events?</b>
<b>Experiment</b>	how, why?	Yes	Yes
<b>Survey</b>	who, what, where, how many, how much?	No	Yes
<b>Archival analysis</b>	who, what, where, how many, how much?	No	Yes/No
<b>History</b>	how, why?	No	No
<b>Case study</b>	how, why?	No	Yes

**Figure 2: The impact of different criteria on selection of research method, Taken from [17]**

Regarding the main research question of this thesis which is “How can visualization techniques be applied to transportation systems?” and two other factors of the above framework, the best research approach for this research could be Case Study.

According to Yin.R [17], there are three types of case studies: exploratory study, descriptive study, and explanatory study. In this research, an exploratory case study will be performed as the first step. The TAPAS simulator [3] will be used as a case to investigate what are the current lacks and possible problems of using visualization techniques in transportation simulation context. In fact, the TAPAS software as existing software for this context will be improved by applying the visualization techniques.

## **2.4 Data Collection methods and techniques**

As Yin [17] stated, there are six different resources for evidences which researchers could use in data collection phase: documentation, archival records, interviews, direct observations, participant observation, and physical artifacts. Moreover, Yin [17] has recommended using different resources of evidences related to a fact to increase the quality of research. Therefore, the author will use documents, and qualitative interviews to investigate how visualization can improve transportation simulators. In the table below (Table 1), the reasons of selecting these sources of evidences are described.

Moreover, it provides a chain of evidences from research questions to data collection and conclusion. The author will use Table 1 to map research questions to the sources such as interview and documentation. Cells of the Table 1 show how different sources should be used to collect data associated with each research question. Furthermore, the literature review which has been done regarding each research question is described.

<b>Resources</b> <b>Research questions</b>	<b>Documentation of TAPAS</b>	<b>Interview</b>	<b>Literature review</b>
<b>RQ1:</b> <b>What are the requirements of users who want to view the results from a transportation system simulator in a visual representation?</b>	Review the technical documents of transportation simulation systems or in general, supply chain simulators and extract potential candidates for visualization.	Asking transportation experts and users of system; What are major problems in the process of monitoring the results of simulator which cause difficulties and interrupts in their work? How much time and cost they spend as an outcome to above problems? What do they think about having simulation results visualized? Which information are they interested in?	Review the Information Visualization literature in order to gain knowledge about visualization and also to conduct an effective interview. Review the transportation visualization sources, and extract important challenges of simulation results monitoring.
<b>RQ2:</b> <b>What are the important features for increasing read-</b>	Review the documents of a sample output result to find out if there is any type of visualization or not. What is	The results of interviews will be investigated to find which of the graph drawing metrics extracted from literature review, should be used in visualization of transportation	Review the literature of graph drawing techniques to find all possible metrics which could be used in visualization. Evaluate

<p><b>ability and usability of a visualization in the context of transportation systems?</b></p>	<p>useful in the existing output format and what problems and difficulties exists?</p>	<p>simulation results and how these metrics can be applied to this context.</p>	<p>these metrics in terms of usability in the context of transportation systems.</p>
<p><b>RQ3:</b></p> <p><b>How can visualization techniques be implemented to meet the requirements of the users?</b></p>		<p>By processing the interviews, it will be identified how should be the final presentation of the transportation simulation results and what is important in visualization of this context. With consideration to these results and by having the visualization techniques, it will be clear which techniques should be used in developing a visualization tool.</p>	<p>Review the previous works which have been done for layout algorithms field and find the characteristics and capabilities of each layout algorithm. Find and evaluate required components for developing the visualization tool.</p>

**Table 1: Mapping research questions to the resource of evidences**

## 2.5 Evaluation methods

For the evaluation of the product regarding technical aspects, there are some methods like black box testing and white box testing [15]. Black box testing is usually used to ensure that the specific inputs could generate the desired output, while white box testing is usually used when the researcher wants to evaluate the correctness of all possible paths of the code. As this research is about applying visualization to a transportation simulation system and develops a visualization tool, it is possible to test it by both black box and white box testing.

For black box testing, the simulation results of TAPAS can be considered as input for visualization tool. So, it is possible to test the visualization tool regarding the

correctness of the created output. It means that we can design test cases to check if a special set of input creates the correct output or not.

Furthermore because the nature of this research is based on user satisfaction and improving the process of monitoring simulation results, we can also evaluate this research from a different point of view. Oates [15] has introduced three type of evaluation of a product from a different point of view:

- Proof of concept: is based on creating a simple prototype to show the feasibility of the researcher's solution in terms of having the specific properties under specific circumstances.
- Proof by demonstration: it deals with evaluation of the product in practice, but not in the real-world. In this approach, the researcher will restrict the context and apply the solution or use the product on a restricted context.
- Real-world evaluation: In this approach, the researcher is examining the product in a real context not in an artificial one.

In this research, the second approach for evaluation of artifact was used. Therefore, we have evaluated the results in a restricted domain with non-real users. We have planned to do real-world evaluation in the future, where a real case and real users will be used for evaluation. For this purpose, an existing transportation simulation system, "TAPAS" [3], has improved by enabling users to observe simulation results visually. For evaluating whether these improvements are useful or not, we need another strategy.

Plaisant has categorized common evaluation methods in the field of Information Visualization into four groups [18].

- Controlled experiments comparing design elements.
- Usability evaluation of a tool
- Controlled experiments comparing two or more tools.
- Case studies of tools in realistic settings

So, for the evaluation of improvements of the case, the second approach was used. Indeed, after improving "TAPAS" as a case study, these improvements were evaluated through performing usability evaluation of TAPAS. This could also be done through

Master Thesis project: Visualization of data from transportation simulation systems

sending a questionnaire to TAPAS users to find out if the improvements could increase readability of TAPAS output or not. This work is considered as a future work for the thesis.

## 3 Literature review

It is crucial to understand the concept of visualization to develop a tool for data visualization. Thus, first part of this chapter reviews the related concepts of information visualization. The second part previews literature about principles and theories of graph drawing. Also, previous studies on information visualization in the transportation context will help to use best practices in the design and development phases. Therefore, in the third section, previous works on application of information visualization in transportation systems are presented.

### 3.1 Information visualization

Information Visualization is a relatively new area of study which is acquiring more and more attention in both academic and industrial environments. This concise introduction to the subject explores the use of computer-supported interactive graphical representations to explain data and amplify cognition [19]. In fact, visualization plays an important role in human perception of the environment. “Visualizations have a small but crucial and expanding role in cognitive systems. Visual displays provide the highest bandwidth channel from the computer to the human” [2]. Owing to the importance of visualization in real world problem domains, it has been growing tremendously in recent years. Several books have been published in Information visualization. Also the number of conferences in this field has been steadily increasing [20].

Information visualization is a multi-disciplinary area of science and includes elements of human-computer interaction, computer science, graphics, visual design, psychology, and business processes. Nowadays, Information visualization has a large area of applications including: scientific research, digital libraries, data mining, financial data analysis, market studies, manufacturing production control, and crime mapping [21].

In this research, we have focused on visualization of data in the context of transportation. For visualizing a transportation network, we need special techniques of

graph drawing. An important characteristic of this context is positioning the elements on the screen. So, we will investigate the techniques for visualizing a set of nodes and their relations in the next section.

## **3.2 Principles and basic theories of graph drawing**

Graph Drawing is a way of visualizing datasets. The positioning of nodes (or vertices) and edges can influence usability, readability and aesthetics. The algorithms which are used for positioning are called “Graph Layout Algorithms”. There are different layout algorithms with different positioning perspectives. In fact there is no best graph layout algorithm and it depends on the purpose of visualization [22].

This is highly relevant to the approach of visualizing transportation systems in this thesis, because of the fact that the positioning of nodes and links and their attributes are key factors in the context of transportation systems.

In this section, the general metrics and criteria for drawing graphs in an aesthetic way will be introduced. Furthermore, some of the most common graph layout algorithms will be described.

### **3.2.1 Graph drawing concepts**

There are some theoretical guidelines for graph drawing. Battista et al. have categorized most important parameters of graph drawing methods into three graph drawing concepts, to achieve better description of requirements of a good drawing [7].

These principals are as followed:

- Drawing conventions
- Aesthetics
- Constraints

#### **3.2.1.1 Drawing conventions**

“A drawing convention is a basic rule that the drawing must satisfy to be admissible.” [7]. In fact, drawing conventions are the basic rules which must be considered in graph

drawing. These conventions are tightly coupled with the context and area of application. One example of a drawing convention could be Orthogonal Drawing in which each graph is drawn as a polygonal chain of alternating horizontal and vertical segments (Figure 3).

Kaufmann et al. [8], have described the connectivity of drawing conventions to the traditional drawing style in a specific context. Based on this research, it is important to consider the traditional way of visualizing data of a specific context. For example, it would be more meaningful to draw a graph in the context of electrical circuits, in an orthogonal shape [8].

Because of the fact that this thesis is about graph visualization of data in a specific context; transportation systems; the above concept is very relevant and useful for this thesis. We can adopt some special conventions for a specific application based on the user requirements. For example, in visualization of transportation systems we can make a rule for representing all the nodes of graph with an icon relevant to their type.

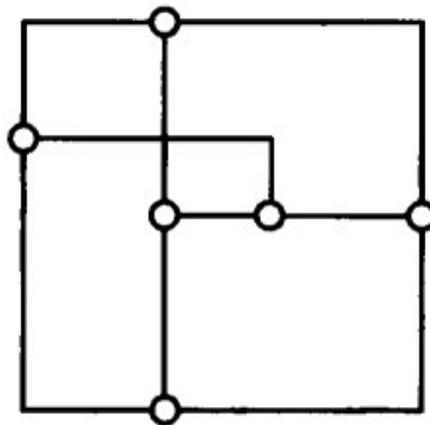


Figure 3, Taken from [7]

### 3.2.1.2 Aesthetics

Graph aesthetics concern parameters which could increase graph drawing readability through focusing on aesthetic issues. “Aesthetics specify graphic properties of the drawing that we would like to apply, as much as possible, to achieve readability.” [7].

Manning [23] has formulated some possible criteria of drawing a “good” graph, for example, these criteria lead to increase uniformity of edge lengths: Minimizing the maximum edge length, maximizing the minimum edge length, minimizing the sum of edge lengths, minimizing the number of distinct edge lengths. In this research, the researcher has tried to create an ideal graph drawing by selecting and combining a set of criteria. In case of having a small set of data, a set of criteria could generate a unique drawing, but if the data set is complex and big, the probability of occurring conflicts in the given criteria increase. In this case, no graph drawing will be generated. Since there are many criteria for drawing a graph [23] and each one could highlight different specifications of a graph, we must identify which graph drawing criteria is the most important for the specific context and application of the visualization. In this case we can select only some of the criteria which have the most importance in a specific context.

Kaufmann et al. have proposed an idea about the level of importance of different graph drawing metrics [8]. They pointed out the fact that selection of graph drawing criteria is dependent to the context of application. For example, in case of drawing VLSI circuits, some criteria like minimizing the edge length are more important than others.

There are some commonly accepted aesthetics [24, 25, 26]. Applying these general criteria to a graph drawing, will increase the graph readability. Some of these criteria are as below:

- Minimizing the number of edge crossings [27]
- Minimizing bends [28, 29]
- Maximizing symmetry [30, 31]
- Adjacency of nodes which have a connection [24]
- Drawing edges as straight as possible [24]

In case of occurring conflicts between aesthetic criteria in drawing a graph, we need a procedure to select a set of criteria which make a drawing more readable. There are some researched which attempt to grade these criteria base on the effectiveness on human understanding [32, 33]. Purchase [33], has placed a priority order for the

importance of the aesthetics. She has done online experiments and measured the ease of understanding drawings. The result of this research indicates that minimizing the number number of edge crossings, is the most important aesthetic criteria, while minimizing bends and maximizing symmetry have less effect on human understandability [33]. Despite the fact that this type of grading could be helpful for building effective drawings, there is still a need for adding the concept of context or application domain for prioritizing aesthetic criteria. In this thesis, specific criteria for drawing graph of data in the context of transportation systems will be proposed.

### 3.2.1.3 Constraints

“Semantic features can be expressed by means of constraints on the drawing.” [34]. Based on this fact, for features that require knowledge about the meaning of the drawing, we can define constraints for graph drawing. As an example, if a set of nodes has more importance than others in a specific application domain, it would be more readable and meaningful to use constraints that place this set of nodes in the center of drawing. Tamassia et al, have classified different type of constraints [34]:

- CENTER: place a set of specified nodes in the center of the graph
- DIMENS: assign the dimension of the symbols representing given nodes
- EXTERN: place given nodes on the outer boundary of the graph
- NEIGH: place a group of nodes as close together as possible
- SHAPE: draw a sub graph with a pre-defined shape
- STREAM: place a sequence of nodes on a straight line

This thesis contributes to this concept by defining the most important constraints in transportation system application domain and uses these constraints to achieve a more readable graph.

## 3.2.2 Graph drawing layout algorithms

There are some general approaches for positioning nodes and edges. The most important and commonly used approaches are listed below [35, 36]:

- Force-base drawing algorithms: In this layout algorithm, the network of nodes and edges are simulated as an electrical system. There are attraction and repulsion forces between nodes and edges. These forces will apply to nodes and the position of nodes will change accordingly. The goal of this algorithm is to reduce the energy associated to graph layout. This algorithm has an iterative nature and it will be repeated until the energy system reaches to a balanced state [30].
- Geographical approach: In this strategy, position of nodes is based on geographical attribute of nodes on a map [37].
- Circular topology: This layout algorithm tries to put nodes on a circle, so the edge crossings will be at the center of the circle [38].

The selection of which approach should be taken is highly dependent on the domain of data to be visualized. For example, in a particular context such as social network, the circular topology is more meaningful than others. In this thesis, the Geographical approach has been used, because of the fact that this approach responds to the requirements of users in the transportation simulation context.

### 3.3 Transportation system visualization

In order to investigate previous works in the area of visualization of transportation systems, different web sites and journals such as IEEE, ScienceDirect, Elsevier , ACM, and Google Scholar have been reviewed. The keywords used in literature review were “transportation systems visualization”, “transportation chains visualization”, “visualization of transportation simulation” and “visualization of transportation simulation systems”.

There are several research groups working on transportation, such as “Transportation@MIT”<sup>1</sup> and research groups in different universities like Monash<sup>2</sup>,

---

<sup>1</sup> <http://transportation.mit.edu>

<sup>2</sup> <http://www.eng.monash.edu.au/civil/research/centres/its/research-activities/public-transport-planning.html>

Southampton<sup>3</sup> and Cambridge<sup>4</sup>. Each research group has different approach to transportation, but none of them has focused on applying visualization methods to transportation simulation in order to monitor the effects of input on final results of the simulation.

However, the “Visualization in Transportation Committee”, a sub-committee of the “Transportation Research Board” (TRB<sup>5</sup>), is focused on visualization techniques for transportation issues. Most of the studies on transportation visualization are referred to this committee. After reviewing the related publications, it became clear that the research goals of the relevant work are different from those of this thesis. Most of previous studies have focused on visualization of the physical products (e.g. a new road or bridge) and the impact of these products on transportation system, whereas this thesis emphasizes on visualization of output data from transportation simulation systems. To put it simply, the focus of the mentioned research group is visualizing and envisioning the transportation related products such as bridges and roads to facilitate decision making by stakeholders. In fact, they try to show the picture of improved situation to stakeholders and visualize the outcome of their decision. But this thesis tries to improve readability of information related to transportation simulation results.

Hixson has classified different applications of visualization in transportation [39]. Some of them are listed as below:

- Design
- Identification and evaluation of human factors
- Analysis of outcomes
- Getting approval from stakeholders and relevant agencies
- Security issues

According to Hixson [39], the main uses of visualization are in the design process and it aims to facilitate the work of designers and getting public approval to reduce the rework costs. Although the approach of using visualization in transportation for design purposes

---

<sup>3</sup> <http://www.trg.soton.ac.uk/index.htm>

<sup>4</sup> <http://www-trg.eng.cam.ac.uk/>

<sup>5</sup> <http://www.trb.org>

is also a goal of TAPAS, this approach is relatively different. In the above research, the use of visualization in design process is limited to showing the anticipated final product (e.g. a road, a bridge) with respect to architectural factors as input, while in this the aim of visualization is to show the impact of applying different measures in transportation process, such as taxes in a visual model.

Although visualization has been used in transportation field for many years and several studies have been done for improving and customizing visualization techniques for transportation goals, there is still a lack of research for application of visualization in planning process [40]. Use of visualization in planning process, can cover human and environmental aspects of transportation projects as well as economic impacts. Hughes [40] believes that the research which have been done in this field so far, mainly concerns how to involve stakeholders in design process and investigate what are effective ways of demonstrating design alternatives to different stakeholders. He argues that there is a need to apply visualization techniques in transportation by considering environmental perspectives in planning process [40]. This thesis is focusing on the impact of transportation decisions on environmental aspects and visualizing the results.

There is other research which describes the benefits of applying visualization in the transportation field [41]. In this research, one benefit of visualization is introduces as a complementary way for ensuring alignment of achievements with planned improvements.

Hannon [42] has introduced another concept in visualization, named “Virtual Reality”. He has made a comparison between simulation and VR. Simulation is defined as a process in which a model of a real system is created in order to discover the effects of changing system’s variables on the whole system and on the future system’s behavior. In this research, virtual reality is defined as a technology which output has these characteristics [43]:

- Computer generated
- 3D model.
- The ability to interact with user and the user must be able to navigate the system and feel the effects of variables in a real simulated system.

This thesis will investigate the visualization of the results of a simulation and it will not cover the visualization of how a system works. Thus it will be a 2D visualization of simulation output and it will show how input data could change the behavior of the system by visual models and diagrams.

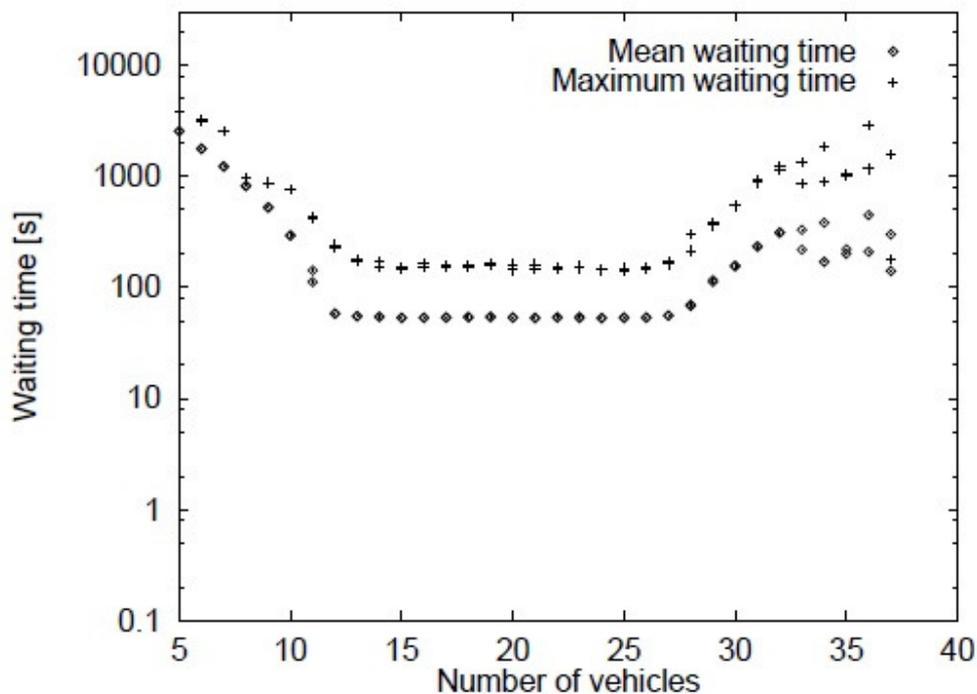
Hughes has reviewed current trends and approaches in visualization of transportation data [44]. Based on this research, two techniques of visualization are mainly used in the context of transportation:

- Map-based visualization: Because of that the nature of data in the context of transportation often is based on geographical information, using map as a background and visualizing transportation data on a map can be considered as a wise solution. Typically, the transportation-related information consists of source, destination and routes between them. Thus, it can be more meaningful to make the use of maps for visualization of freight data. An example of map-based visualization is illustrated in Figure 4.





Hoff, et al have done a research about simulation of Automated Guided Vehicle Systems (AGVS) and visualization of simulation results [46]. They visualized the simulation results in the form of diagrams. One sample diagram is shown in the Figure 6.



**Figure 6:** A sample diagram of simulation results which depicts the Dependence of waiting times on the number of vehicles used. (Taken from[46])

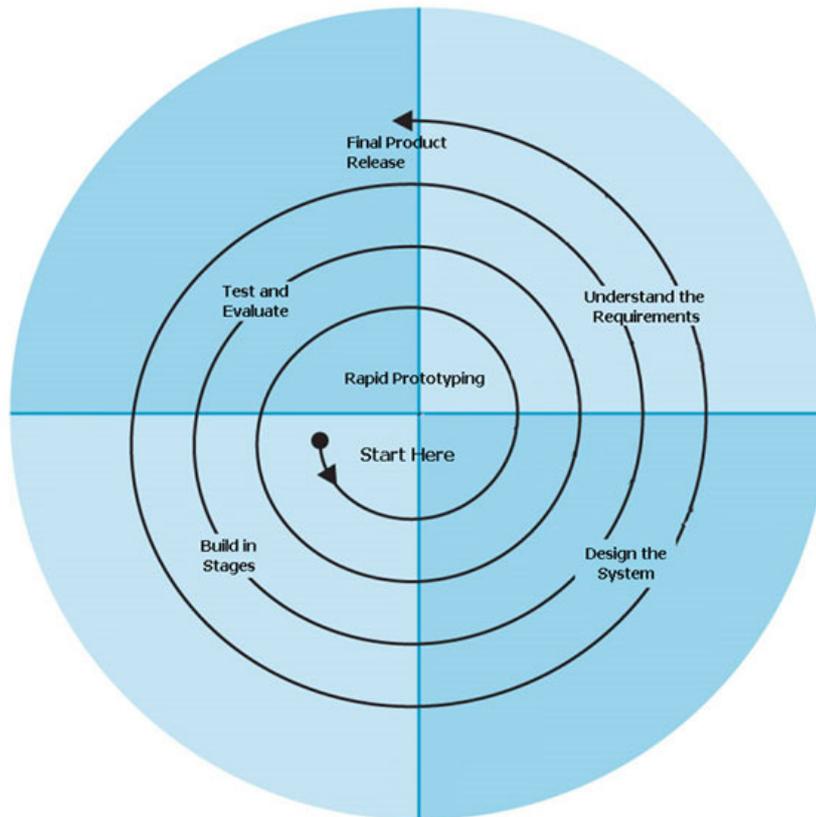
There is research in which the authors have worked on developing a software system for visualizing their multi-agent model of logistics management [47]. Their proposed features for the software, consists of visualizing the simulation process, routes, the appearing of loads and truck movements. This software is the visualization of how the simulation system works. It has the ability to work in two modes, static and dynamic. In the static mode, the user can navigate through the time and in dynamic mode, the movements and actions in simulation are controlled by a timer. The researchers did not provide any picture of this system in the article.

As the scope of this thesis is investigating in visualization techniques for a multi-agent based transportation simulator, where monitoring the effects of applying measures is the goal of simulator, it is very different form the research which have been done in this field so far.

## 4 Empirical case

In this section, we will study the TAPAS case as a real implemented transportation simulation system. First of all, the TAPAS system as the unit of analysis is introduced. Secondly, with regard to RQ1, the visualization requirements of TAPAS users will be identified and analyzed. Next, the potentials of visualization techniques to cover identified user requirements will be discussed. At the next step, the functionalities and features of a TAPAS visualization tool will be elaborated (RQ2). Finally, results of the construction phase will be represented (RQ3).

The process for designing the visualization tool is based on the “Systems Engineering” process [48, 49]. We have followed “System Development Life Cycle (SDLC)” as a life cycle of system development; in which we have Analysis, Design, Implementation, Testing and Evaluation [50]. Furthermore, we have used the “Spiral Model” for the migration from requirements to design solution [51]. This selection made due to the fact that we have created prototypes of the system to get user feedback during development process. This method has an iterative nature and we repeat each iteration until the final product is made. We have used a simple summarized model of spiral method for this project, which is showed in Figure 7. We have selected this method because this project is a small project and it is not possible and meaningful to follow all the detailed steps of the complete spiral model. As an example, it is not necessary to do risk analysis for a small project. The original model of spiral method which is illustrated in Figure 8, displays the iterations of the “Spiral Model” [51].



**Figure 7: Simple spiral model for software development (Taken from [http://www.arctern.com/Development\\_Methodologies.aspx](http://www.arctern.com/Development_Methodologies.aspx))**

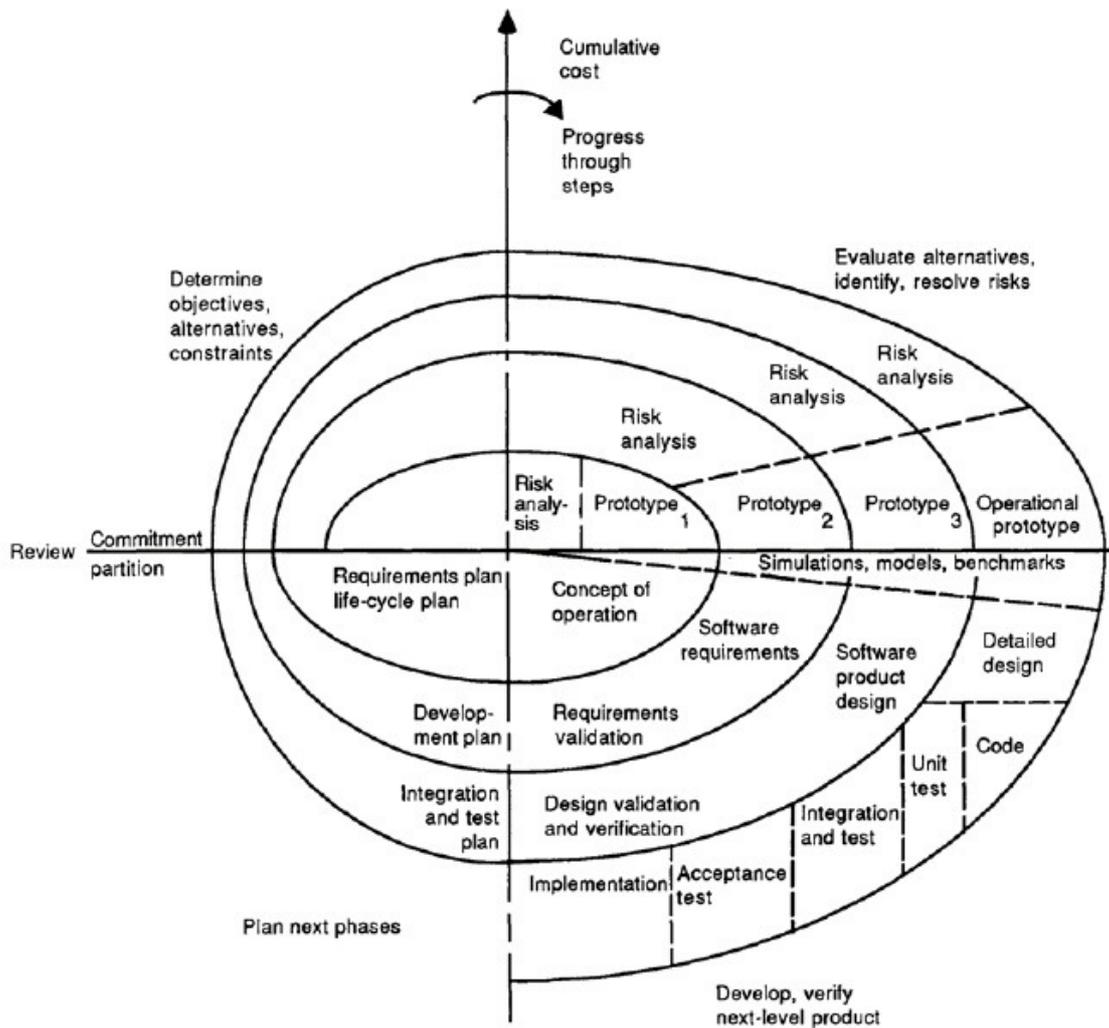


Figure 8: Original spiral model for software development (Taken from [51])

## 4.1 TAPAS

TAPAS is a Multi-Agent-Based Simulation (MABS) model which aims to analyze the effects of applying production and transportation measures on supply chain activities. There are different types of measures which can be considered as input to TAPAS. TAPAS can be used to predict the consequences of various measures for decision and

policy makers, before making any decision. Some of these measures are listed below [52, 4, 10, 54, 55]:

- Governmental control policies, such as taxes and fees,
- Infrastructure investments, such as investments in a new road or railway,
- Strategic business related strategies, such as changing the location of a factory

The output of TAPAS shows how supply chain actors are assumed to react to above measures. So, the users of TAPAS can interpret the results and make more informed decisions and avoid undesired effects. The main outputs of TAPAS simulation are described below:

- Economic effects, such as costs for fuel and taxes
- Logistic effects, such as such as route and mode choices
- Environmental effects, such as different type of emissions

It is important for users to compare the output results of applying different measures (scenarios).

A TAPAS simulation scenario consists of a set of transportation and production actions, and a set of measures which are applied to the simulation. Thus, each scenario studies the effects of a certain set of measures on the supply chain.

TAPAS can have two types of users. The first group is public policy makers, who are interested in predicting the impact of their decisions in approaching their governmental goals, e.g. emission targets. The second group can be private companies and enterprises and other stakeholders of transportation and production, who want to enhance their competitive advantage among other companies. Different type of users may intend to apply different measures. For example, public authorities are more interested in the first group of measures, i.e. governmental control policies, while enterprises might be more eager to predict the effects of their business related decisions, such as changing the location of a factory, on their profit. However, enterprises might be also interested in monitoring the effects of a new transport policy [53, 54].

We have selected TAPAS as our case because

- 1- TAPAS is an implemented transportation simulation system without visualization tool. Thus, it is not easy for users, especially for non-expert users, to interpret the results of simulation, compare output of different simulation scenarios, and make a decision based on their interpretation. A visualization tool can be a complementary tool for TAPAS, and help users to compare results. One of the motivations for selecting TAPAS, is that it is fully implemented and it is currently applied to real cases.
- 2- TAPAS is an outcome of research conducted at Blekinge Institute of Technology<sup>6</sup> and information of the project is accessible for this researcher at Malmö University. Some of the researchers of this project are currently working at Malmö University, so I have access to the resources of TAPAS.
- 3- Users of system are available for interviews and questionnaires.

#### 4.1.1 Definition of TAPAS terms

There are some terms which are used often in the TAPAS documentation, and therefore in this research.

- **Node:** A node can be different kinds of actors in the transportation chain, for example: customer, producer (a factory), or a connection point (terminal) [53]
- **Transportation mode:** A transportation mode represents the type of transportation (e.g. road, rail, or sea)
- **Link:** A link is a directed connection between two nodes. A link can have exactly one mode of transportation. However, we can have several links between two nodes which differ in their transport mode.

## 4.2 Requirements analysis

In this section, the efforts which have been done to address the first research question of this thesis; which is: “What are the requirements of users who want to view the results

---

<sup>6</sup> <http://www.bth.se/com/tapas.nsf/pages/tapas>

from a transportation system simulator in a visual representation”; are presented. In order to answer RQ1, it is necessary to understand the problems of users regarding readability and usability of TAPAS output, and how a visualization tool could help them. In addition, we need to know the type of target groups who potentially can use TAPAS, because their needs should be considered in the design process of the visualization tool of TAPAS. This can be achieved through a requirement analysis. In order to gather requirements of each target group we conducted interviews with TAPAS users, who are mostly expert in the field of transportation. The first part of this chapter classifies and describes different type of user groups. In the second part, we describe the method of data gathering for each target group. Conclusion of the user requirements will make the last part of this chapter.

#### **4.2.1 Classification of TAPAS users**

Based on previous publications of TAPAS, we can categorize potential users of TAPAS into two main sections: Public authorities and large companies [54]. These are users of TAPAS simulation system in general. But for this research’s purpose, which is developing a visualization tool for TAPAS, we can consider that we have another type of users who are developers of TAPAS. They can be also considered as TAPAS users, because of the fact that they are entering various scenarios to the system and run simulation several times to generate simulation results.

Hence, we can categorize TAPAS users from another perspective:

- Expert users
- End users

Expert users are analysts and researchers who work with the simulation system and manipulate the simulation results.

End users are decision makers who can use simulation results to improve their work. Both decision makers in public authorities and large companies are considered as end users in this case. Some part of the end users can be the users who do not work with the visualization tool, but they use the output of visualization tool. This group can be called

off-line users, who just see the output of visualization tool as screen dumps or statistical charts. We consider that this type of users have the same requirements as end users.

#### **4.2.2 Data gathering method**

For answering RQ1, it is necessary to analyze the user requirements. Each type of users has their own concerns and expectations from the system. In order to identify user needs, two interviews with expert users of TAPAS have been done. The interviews aim to find the user requirements and discover probable problems which TAPAS users face without visualizing the results. The interview questionnaire and the answers are enclosed to this report as Appendix I.

Furthermore, we have designed a web-based questionnaire for potential end users to understand their special needs and interests of a transportation simulation visualizer. The screenshots of this questionnaire are attached to this report as Appendix II.

##### **4.2.2.1 Interview with expert users**

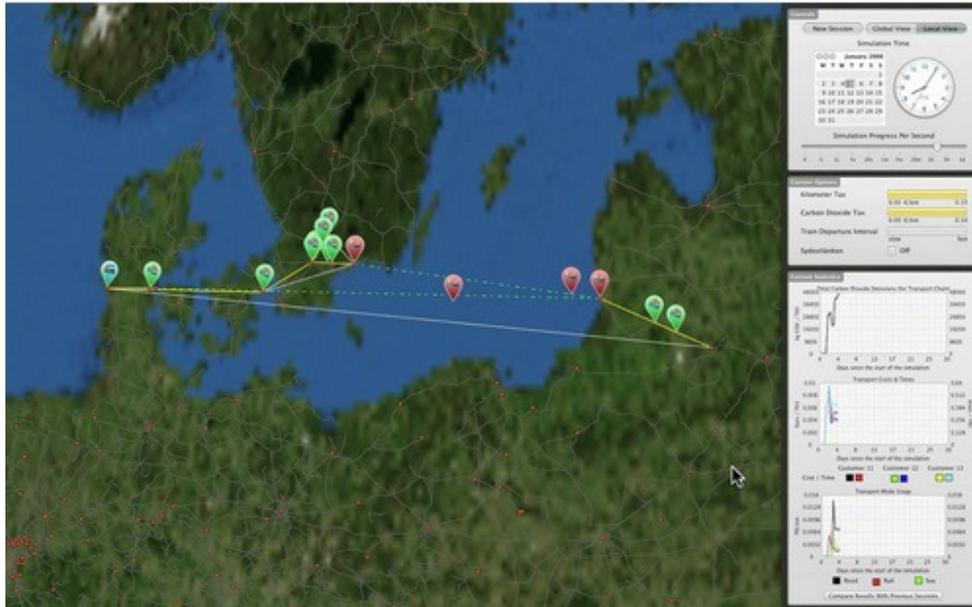
Due to the fact that the expert users were not accessible for a face-to-face interview, interviews conducted via Internet. Therefore, a structured interview with pre-defined questions performed.

From the interviews with expert users, the author found out that currently, there is no Graphical User Interface for presenting the results of simulation. Although they mentioned that there were some attempts to visualize the TAPAS, the purpose of visualizing was relatively different. The aim of that work, which was done at Blekinge Institute of Technology<sup>7</sup>, was to demonstrate the TAPAS simulation process. To put it simply, the visualization tool, called GRACE, is about visualizing the process of simulation and it does not cover the visualization of simulation results. Furthermore, one of the characteristics of GRACE is that it covers only a specific pre-defined scenario and it cannot be used to visualize different simulation scenarios. The interviewees also mentioned that the people who was involved in this project and had knowledge about this

---

<sup>7</sup> <http://www.bth.se/tapas>

visualization tool are not accessible now. Therefore, it is not possible to extend or reuse this tool easily. Some screenshots of GRACE visualization tool are presented in Figure 9 and Figure 10.



**Figure 9: A sample screenshot from GRACE which shows the process of simulation  
(Taken from <http://www.bth.se/tapas>)**



**Figure 10: A sample screenshot from GRACE which shows different bar charts during simulation (Taken from <http://www.bth.se/tapas>)**

Another point that extracted from the interviews is that the input data and output of TAPAS simulation system are stored in SQL database. Thus, to generate a meaningful output result from TAPAS, they have to design some queries. Also, the users have to analyze these data manually to create some graphs, tables and diagrams to show the output in a more understandable way. Since the process of formulating SQL queries and analyzing the collected output data is error-prone, sometimes they need to go back and start from scratch because of small errors. To summarize, one major problem for TAPAS users without a visualization tool is that the whole process of designing visible outputs from TAPAS is quite time-consuming.

Both of the interviewees agreed that visualization can help TAPAS users to have a better and clearer understanding of the results. They had also some suggestions for visualizing. One valuable suggestion is visualizing also the “input data” for different scenarios of simulation. It can help users to realize which input has resulted to this output. It is also a good suggestion to show and compare different scenarios. So that the users can gain a better understanding of how applying different input measures can change the results (e.g. route choice).

TAPAS has different output that can be visualized [3,55]. One of the questions of the interview is focused on identifying the interests of users regarding visualization of output. The interviewees suggested some of the output which is important to display in a visual form. Both of interviewees have the same idea about visualization of output, but they expressed them in different languages. The important output to be visualized can be categorized as below:

1. Route choices (percentage between different routes)
2. Mode choices (percentage between different modes)
3. Transportation per link (e.g. number of vehicles per link)
4. Share of amount of emissions per route (e.g. CO<sub>2</sub>)
5. Transport cost (per route)
6. Different types of other costs (per route)

As a conclusion of the interviews, the author found that there is an inevitable need for visualization in TAPAS to facilitate the process of analyzing simulation results. Furthermore, it is now clear what requirements of the users are with respect to visualization of output.

#### **4.2.2.2 Interview with end users**

End users can be defined as decision makers who are working in the transportation field. We designed 7 questions and published them through an online survey. This decision was made because the selected users were in different countries and they were not accessible for a face-to-face interview. Also, we should minimize the number of questions to encourage the respondents to answer the questionnaire. Due to the fact that there were just few end users available, we have sent our questionnaire to three users. Although this questionnaire was easy to answer and it provides visual models to make it more tangible, we received only one response. A possible reason for this is that it was summer vacation and perhaps some of the selected users were on vacation.

The purpose of this questionnaire was to find out which type of simulation data the end users are interested in. We used the knowledge extracted from literature review and the previous interview with expert users to suggest possible data which they may be

interested in. Furthermore, we have put some screenshots from our design solutions to get their opinion regarding their preferences to see information of simulation and how the final visualization tool should look like.

From this questionnaire, we extracted their interests for visualization of different kind of the TAPAS output. The user is interested in to see route choices, modal choices, and amount of emissions in a visualized model. Also, this user believes that comparing different scenarios of simulation could be an important feature of the system. The user also feels that visualization of simulation results could contribute to his work.

For the next three questions, we provided mock-ups of our proposed visualization tool which we had made using Microsoft PowerPoint. The aim of these questions is to gain the user's opinion regarding the importance of showing information related to each route or each link in the transportation network. From these questions we realized that the respondent is interested to have information related to each route.

#### **4.2.2.3 Conclusion of user requirements**

**Req1. Showing input and output of simulation:** We can conclude from the interviews that both expert and end users are interested in a visualization tool and that they think it could facilitate their work. Also, they need to see which input has resulted in a specific output; just visualization of simulation output will not be meaningful for them. Thus, we conclude that we should visualize the input related to each output as well.

**Req2. Showing different scenarios concurrently:** Both types of interviewees would like to be able to see several different scenarios at the same time. They found it useful to compare various scenarios and their related output. Also, they think this feature can help them in evaluating the output results and facilitate the decision making process.

**Req3. Showing simulation information based on links and routes:** Furthermore, interviewees had some suggestions for which output is more meaningful to be visualized and which not. We can also conclude that they need to see information related to both links and routes.

In the next section, in order to understand how visualization techniques could help representing simulation results effectively, we correlated visualization techniques with important information for users in the next section.

### 4.3 Visualization Techniques versus Requirements

The strategy of answering the second research question (RQ2); which is “What are the important features for increasing readability and usability of a visualization in the context of transportation systems?”; is correlating the user informational requirements from interviews with visualization techniques from literature review. Thus, after studying the current state of the art in visualization algorithms and by considering the points extracted from interview, we have decided on which visualization techniques can be helpful for representing TAPAS outputs. Based on the results, we have proposed a mapping for visualization of different outputs.

This mapping relates the interview results to possible visualization techniques. To put it simply, after analyzing interview transcripts, the different types of TAPAS users and their special requirements regarding visualization of output are determined. Based on the knowledge obtained in the literature review about visualization techniques, the mapping can be completed by relating each user need to a suitable visualization technique. The proposed mapping contributes in the design of the visualization tool. This mapping is presented in the table below (Table 2). This table shows how we can use visualization techniques for addressing the visualization requirements of users.

As stated before, the spiral model is used in this thesis to migrate from requirements to design solution. As a part of this migration, several mock-ups of the tool have been prepared and shown to the users. Based on the user feedbacks, this mapping is revised several times to get the most appropriate visualization solution. Therefore, we have started with an initial simple mapping and we have revised the mapping after getting the user opinions.

In this mapping, each “\*” describes how the user requirements (each column) could be satisfied by the visualization techniques (each row). So, the existence of a “\*” in a cell

means that this visualization technique is selected and approved by the user to cover the specific user requirement.

This mapping is a brief illustration of design decisions in response to the user requirements, which is used in the development phase. In fact, this mapping provides a guideline for understanding which visualization technique should be implemented for satisfying user requirements. Furthermore, we have used this mapping for design of the mock-ups.

Furthermore, we have concluded from literature review and user requirements to display all the informational requirements of users on a graph. Also from literature review we found out that a good solution for visualizing a transport network, could be displaying a map as background.

User Groups Visualization Technique	Expert and end users informational requirements (Req3)				
	Route Choice	Mode Choice	Transportation per link	Emissions	Transportation Costs
Using colors	*				
Using pie chart	*			*	*
Thickness of edges	*	*	*	*	*
Highlighting edges	*				
Using different edge styles		*			
Displaying data on edges			*	*	*

Table 2: Mapping user requirements to best visualization solution

Furthermore, by analyzing the interview results, we can conclude that there are two ways for presenting information which users are interested in; information on link and information on route. So, we can categorize user requirements with regard to visualization in two parts.

- Link based view
- Route based view

The information is mainly connected to each route, while displaying some information for each link can be also meaningful.

#### 4.4 Features and Functionalities of TAPAS Visualization tool

Because of the fact that the design decisions could not be discussed with users just by explaining theories, we have designed several mock-ups of the visualization tool. A mock-up could help users to imagine the real system and facilitate the process of designing functions and features of the visualization tool.

Thereby, we designed mock-ups by Microsoft PowerPoint in several steps and discussed with expert users and end users to get their feedback.

Based on the above mapping, the proposed User Interface of the TAPAS visualization tool can have two Tab Views; Link based view and Route based view. This simple sketch, Figure 11, was produced as the first step. The relevant information of each category can be displayed in their related tab.

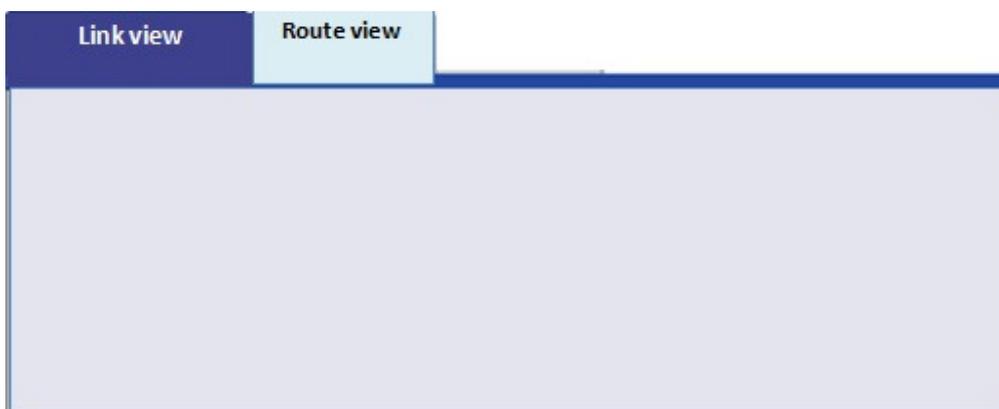


Figure 11: First sketch of proposed Tab View for visualization tool of TAPAS

The second sketch shows the different techniques of visualization, which can be used to show different transportation modes between two nodes. This sketch is illustrated in Figure 12. This sketch is presented to the expert users and they suggested taking the second approach for demonstrating different transportation modes. For the reason that TAPAS users have employed different edge styles as a notation for different transportation modes in the previous articles and documents of TAPAS, they suggested to follow the same notation. In this case, this visualization tool can be more tangible for users.

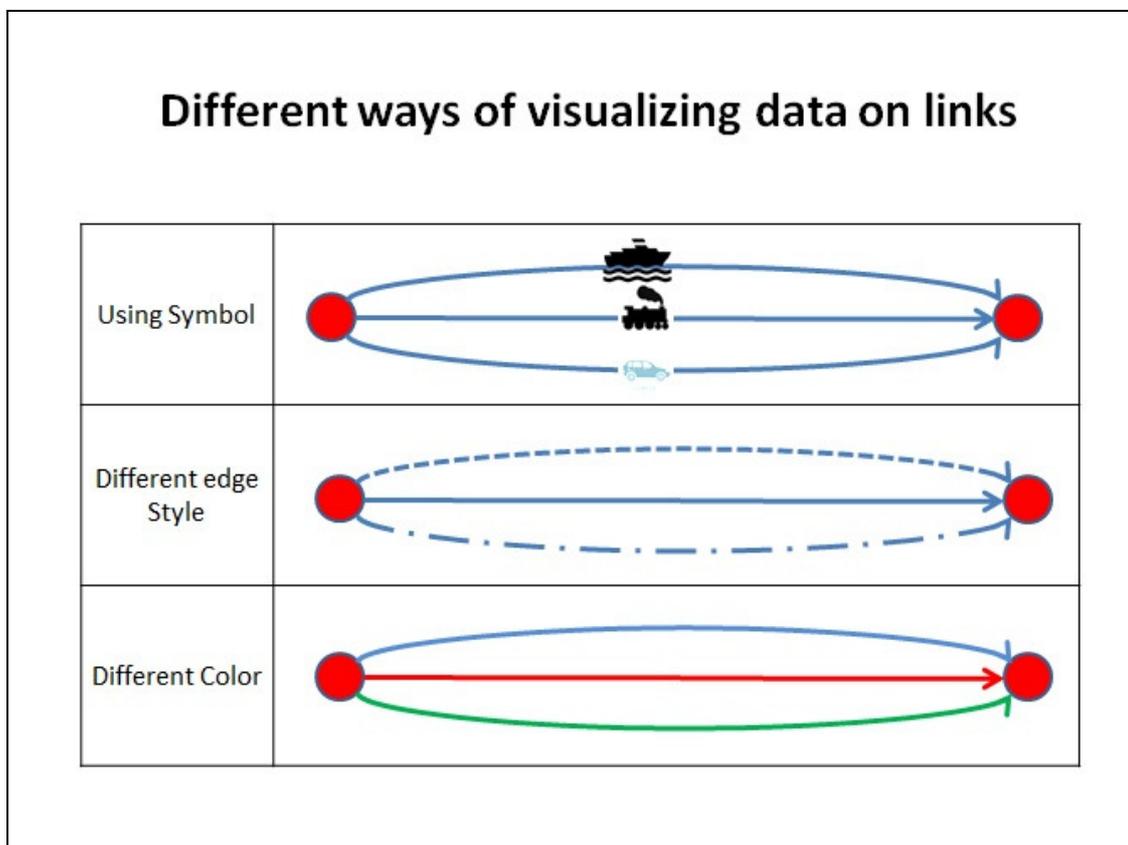


Figure 12: First sketch of proposed Tab View for visualization tool of TAPAS

First of all, for displaying the route choice, the best solution based on literature review and user feedbacks is using graph. To put it simply, the proposed solution displays each customer and producer as a node of transportation graph. For illustrating the percentage of using each route (e.g. the amount of ton/route), the solution could be drawing edges between nodes with different thicknesses. For example, if a specific route has been used more, its edge in the graph will be drawn thicker. As a result of this visualization feature, the user of TAPAS can easily find that which route is being used more than others.

As the next step, based on these main features of visualization system, we designed a more complete sketch. We also used this sketch for preparing the end-users' questionnaire.

The user can select a specific scenario at first when the program runs, then the information related to this specific scenario will be shown. In this way, we can visualize the input of simulation. Furthermore, it will be possible to make a comparison between different scenarios.

The following figures are illustrating the final sketches for showing the cost information on link and route as well as the input of simulation (Figure 13, Figure 14). The other information; CO<sub>2</sub> and amount of transport work; have the same layout but with different data.

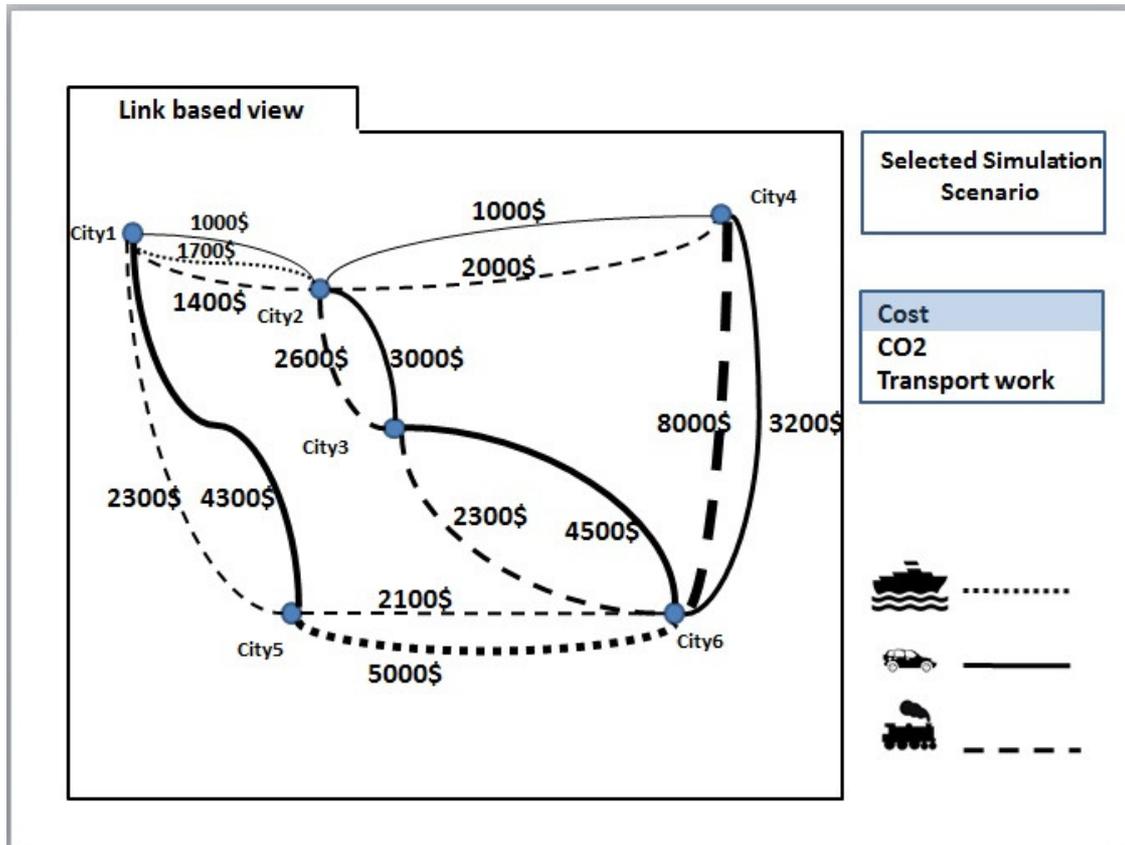


Figure 13: The final sketch which shows the cost information for each link and input of simulation.

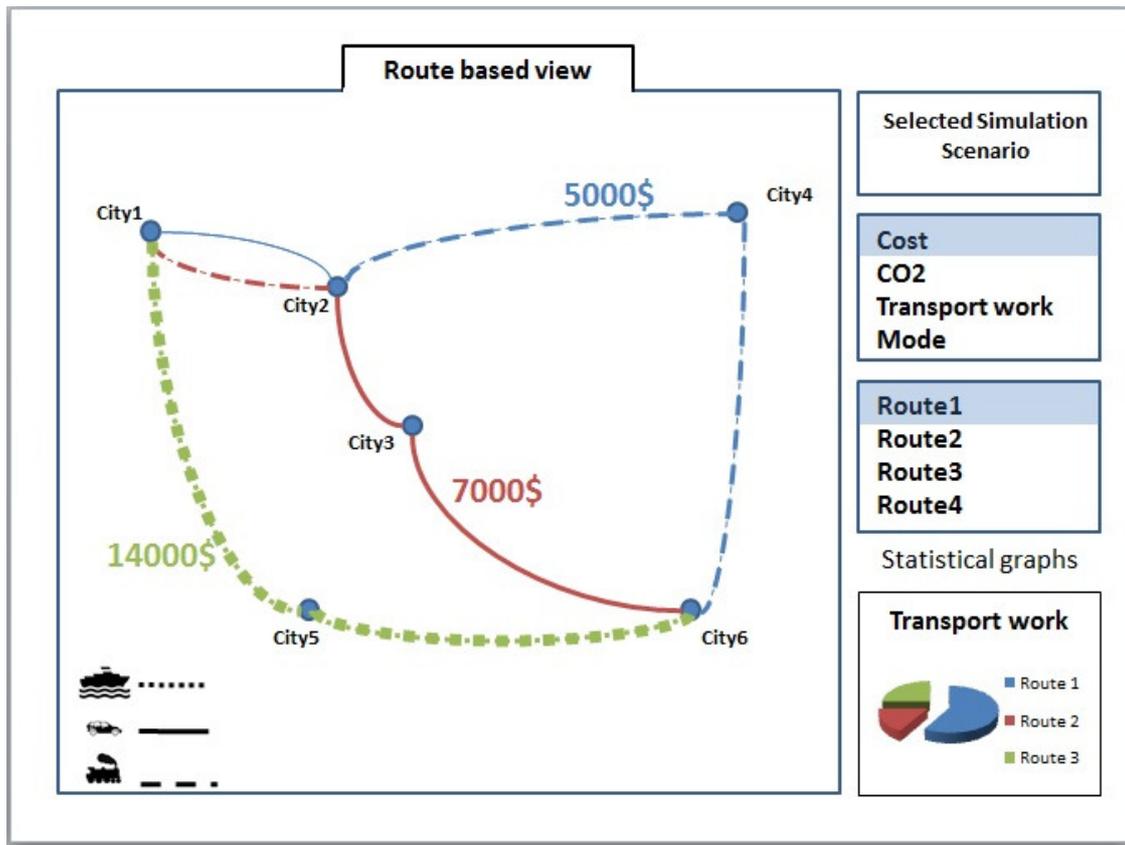


Figure 14: The final sketch which shows the cost information for each route and input of simulation.

Based on user requirements and with consideration to the proposed mapping of user requirements to visualization techniques (Table 2), we proposed to visualize information of each main output (Cost, CO2, Transport work) and allow users to select the desired information from the right side bar in the picture. In Figure 13, the cost information for each link is shown above each link, also we have used different edge styles to show the various transport modes of links. Figure 14, shows the cost information of each route. As we stated in the proposed mapping, we can use edge colors for differentiating different routes. We can also, highlight the selected route. We have a graph in the right bottom corner of the picture, which shows and compares the cost related to all routes.

In general, main functionalities of the system, regarding user requirements and above mapping (Table 2), are designed as below:

**Func1. Show as a graph:** Using color, edge thickness, edge styles to differentiate links; regarding simulation results in response to the Req3.

**Func2. Draw charts:** to illustrate data and make comparison between different scenarios in response to the Req2.

**Func3.Link based information on Graph:** Showing information of cost, emission, amount of transport work considering Req3.

**Func4. Route based information on Graph:** Information of cost, emission, amount of transport work with regard to Req3.

**Func5. Add a map as background:** to increase usability and readability of the graph regarding Req1 and Table 2.

## 4.5 Visualization tool Construction

The last research question (RQ3) will be approached through designing and developing a complementary tool for TAPAS which aims to visualize the output results in a more understandable way for users. The design of this tool will be based on the results of two previous research questions.

### 4.5.1 Development Technologies

As we mentioned before, TAPAS is implemented using the Java language [4, 53], so it would be more integrated to develop this tool in Java language. We have used Swing<sup>8</sup> technology as an Application Programming Interface (API) for designing the user interface.

Another option would be to develop the tool on the top of the Google Maps API<sup>9</sup>. The advantage of Google Maps API is that it shows the logistic positions very well and we do not need extra classes and relations with graph drawing libraries anymore. But, the problem is we cannot use real positions in a real map for this tool. Because in some cases

---

<sup>8</sup> <http://download.oracle.com/javase/tutorial/uiswing/>

<sup>9</sup> <http://code.google.com/apis/maps/documentation/javascript/>

the positions of two nodes are very far from each other, and if we use real coordinates in a real map, it will decrease readability of the graph. Thus, we need to customize the map and manage the positions of nodes manually and place them in a way to increase graph usability and readability. The major problem of using Google maps API is that, it contains too much information, which makes it difficult for user to see the simulation results on the map. Furthermore, there are some limitations in using Google Maps API for this tool, for example, we need to bring other classes and libraries for drawing charts and other parts of the tool, which is not possible by using Google Maps API.

Furthermore, because of the nature of this tool, we need to draw graphs and charts. There are some Java components and libraries for the purpose of graph drawing, e.g. UCINet<sup>10</sup> and PAJEK<sup>11</sup>. Each of these mentioned libraries has focused on a specific application area. As an example, UCINet is mainly used for social network analysis. The main application area of PAJEK is also social networks, and it is mainly used for visualizing of very large networks.

For implementing the visualization part of this tool, the JUNG library<sup>12</sup> is regarded as the best solution. JUNG library, Java Universal Network/Graph Framework, is a software library that provides a common and extendible language for the modeling, analysis, and visualization of data that can be represented as a graph or network [56]. Moreover, the JUNG Library has implemented some layout algorithms and uses these layouts to show data in a graph. These layout algorithms are basically mathematical-based algorithms and each of them is focused to solve some specific problems [7]. This selection is motivated by several reasons: First of all, it is well documented and it is easy to download and use; secondly, it is free and it has not any license; thirdly, JUNG has a very good support and it has an “Open discussion forum” and a “Support mailing list”, which enables the user to ask questions regarding JUNG and get response very soon.

---

<sup>10</sup> <http://www.analytictech.com/ucinet/>

<sup>11</sup> <http://vlado.fmf.uni-lj.si/pub/networks/pajek/>

<sup>12</sup> <http://jung.sourceforge.net/index.html>

For the database connection part of the tool, we use PostgreSQL<sup>13</sup>. This decision is made because of the fact that the TAPAS database is made by PostgreSQL. Thus, for designing desirable queries, we have used SQL statements. In order to explore the database, we have used pgAdmin III<sup>14</sup>, which is a PostgreSQL administration and management tool.

The selected IDE<sup>15</sup> which is used for developing this tool is NetBeans IDE<sup>16</sup>. This selection is made because it has a professional Swing GUI Builder<sup>17</sup> which makes it easier to design GUI<sup>18</sup>.

Also, we have selected JFreeChart library<sup>19</sup> for drawing charts and diagrams, because this library contains all type of charts (e.g. Bar charts, Pie charts, Scatter plot charts,...) which can be used in future improvements of this tool and support more types of charts for users. Also, this library is a free open-source java chart library and it is well documented.

#### 4.5.2 Investigate Current System

To build and develop visualization functions, we investigated the TAPAS software and database. According to our investigation the current database contains just the final results of running different simulation scenarios. In order to visualize information for different routes and links we need detailed intermediate information of cost, emission, and load (transport work) for each link between nodes. Therefore, we decided to design a new table to develop required functions because they could not be implemented by using current database. This new table which is called “visualization\_infromation” is a temporary solution for developing the visualization tool and it can be filled by the

---

<sup>13</sup> <http://www.postgresql.org/>

<sup>14</sup> <http://www.pgadmin.org/>

<sup>15</sup> Integrated development environment

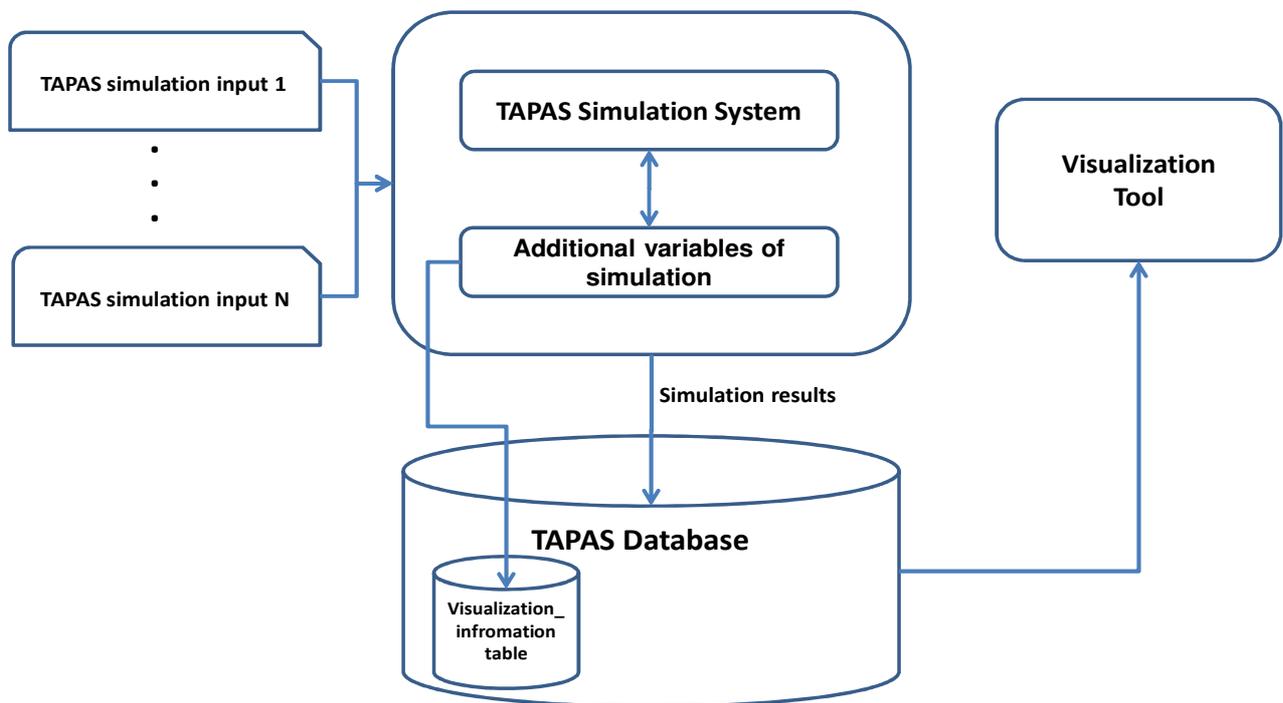
<sup>16</sup> <http://netbeans.org/>

<sup>17</sup> <http://netbeans.org/features/java/swing.html>

<sup>18</sup> Graphical User Interface

<sup>19</sup> <http://www.jfree.org/jfreechart/>

TAPAS system in the future improvements and simulation runs. This table is optimized in order to have minimum effect on the performance of TAPAS. Thus, TAPAS performance will not affect by this new database schema. Figure 15 represents the proposed database schema of TAPAS. This figure shows the existing database and its relation with TAPAS simulation system. In addition, the new proposed table which contains required information for the visualization tool is illustrated.



**Figure 15: The proposed temporary database schema and its relation with TAPAS simulation system.**

This new table provides information related to each link. We have formulated the primary key of this table in a way which enabled us to retrieve the information related to each link from this table. The detail specification of the “visualization\_infromation” table is described in Figure 16.

## Table properties

Property	Value
Name	visualization_information
OID	16803
Owner	tapas
Tablespace	pg_default
ACL	
Of type	
Primary key	simulation_nr, path_id_key, start_node_id, end_node_id, transportation_mode_id
Rows (estimated)	42
Fill factor	
Rows (counted)	43
Inherits tables	No
Inherited tables count	0
Has OIDs?	No
System table?	No
Comment	

```
-- Table: visualization_information
-- DROP TABLE visualization_information;
CREATE TABLE visualization_information
(
    simulation_nr integer NOT NULL,
    path_id_key integer NOT NULL,
    start_node_id integer NOT NULL,
    end_node_id integer NOT NULL,
    cost double precision,
    load double precision,
    co2_emission double precision,
    description character varying,
    transportation_mode_id integer NOT NULL,
    CONSTRAINT visualization_information_pkey PRIMARY KEY (simulation_nr, path_id_key, start_node_id, end_node_id, transportation_mode_id)
)
WITH (
    OIDS=FALSE
);
ALTER TABLE visualization_information OWNER TO tapas;
```

**Figure 16: The specification and fields of the new proposed table (visualization\_information)**

The primary key of this table are pointers to the current tables of TAPAS database. This will enable us to retrieve information of a specific link in a simulation scenario from database. The motivation of adding each field in this table and a brief description of each field is provided below:

- simulation\_nr: This field is representing the selection of simulation scenario. It is a part of foreign key of the visualization table.
- path\_id\_key: This field is representing the route.
- start\_node\_id: It is pointing to the start node of a link in the “link” table.
- end\_node\_id: This field is pointing to the end node of a link in the “link” table.
- cost: It represents the cost information related to each link on a specified route.

- **load:** It represents the amount of transport work related to each link on a specified route.
- **co2\_emission:** It represents the amount of CO2 emission produced on each link on a specified route.
- **description:** This field contains the name of simulation scenario. This name can be a brief description of the scenario specifications.
- **transportation\_mode\_id:** This field represents the transportation mode of each link.

Each field of the table is necessary to find the correct information from database.

Furthermore, for including the geographical positions of a node we have added two fields to one of the existing tables of TAPAS database called “node” table. These two fields are:

- **geographical\_position\_x:** represents latitude of the node.
- **geographical\_position\_y:** represents the longitude of the node.

### 4.5.3 Developing the Visualization tool

In this section, we will describe the process of developing each functionality of the visualization tool:

#### **Func1. Show as a graph:**

For developing this functionality, we used JUNG library and we draw the graph based on the information from the database. We have designed several queries to get desired data and put in the graph. We draw the graph based on the retrieved information. For drawing each aspect of the graph, such as edge thickness, or edge style, a suitable SQL query is designed. Based on the query results we draw all parts of the graph. In some cases we need several nested queries to retrieve desired information. As an example, one of the queries for retrieving the edge styles of each link is described in Appendix III, Part 1.

After designing and running queries, we show the new information on screen. So, we use the appropriate renderer to update the forms of the visualization tool. For example, in the case of setting the appropriate edge style based on transportation mode, we use the “EdgeStrokeTrasformer” renderer. An example of a renderer is illustrated in Appendix III, Part 2.

### Func2. Draw charts:

To implement this functionality, we have designed the first form of this tool in a way that the user can open several windows for each scenario. It means that, at the first form of visualization tool, the user can select the scenario name, and then the second form which contains the related scenario will be opened. In this way the user can go back to the first form and select another scenario, while the first selected scenario is still open, and make a comparison between the charts of two or more scenarios. The screenshot of this first form is illustrated in Figure 17.

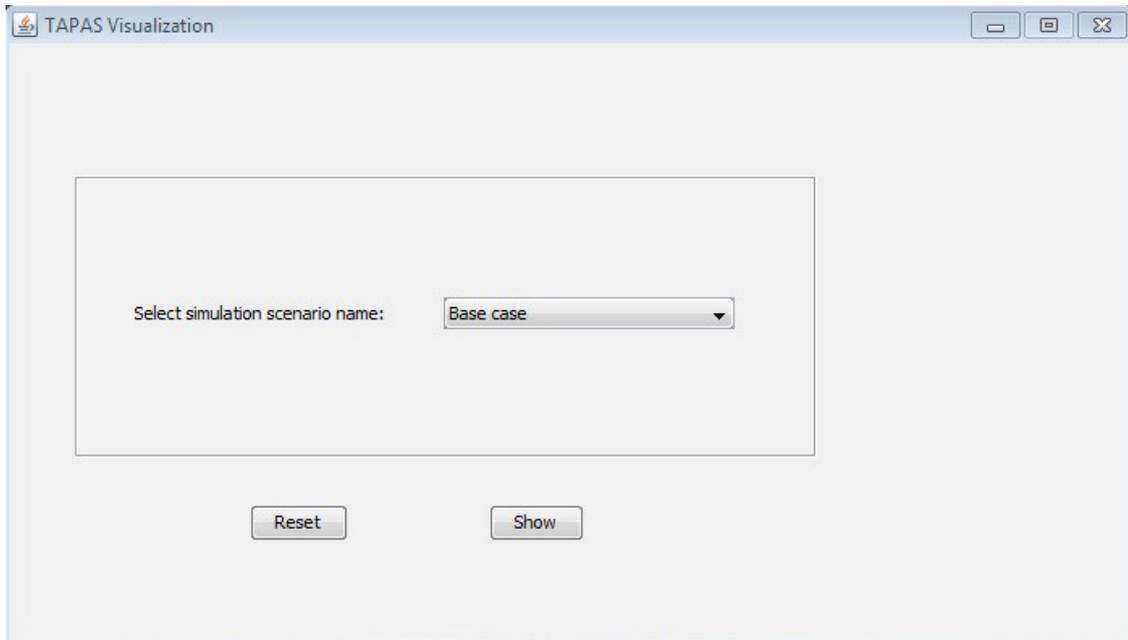


Figure 17: First form of visualization tool; selecting scenario name

### Func3.Link based information on Graph:

For illustrating the information on each link, we also, designed several queries and retrieved appropriate data. Then, we send this data to the Graph Listener to show them in the graph. This Listener will be activated whenever the user pushes the related button of each data to show related information on each link. One screenshot of the Link Based View is shown below (Figure 18).

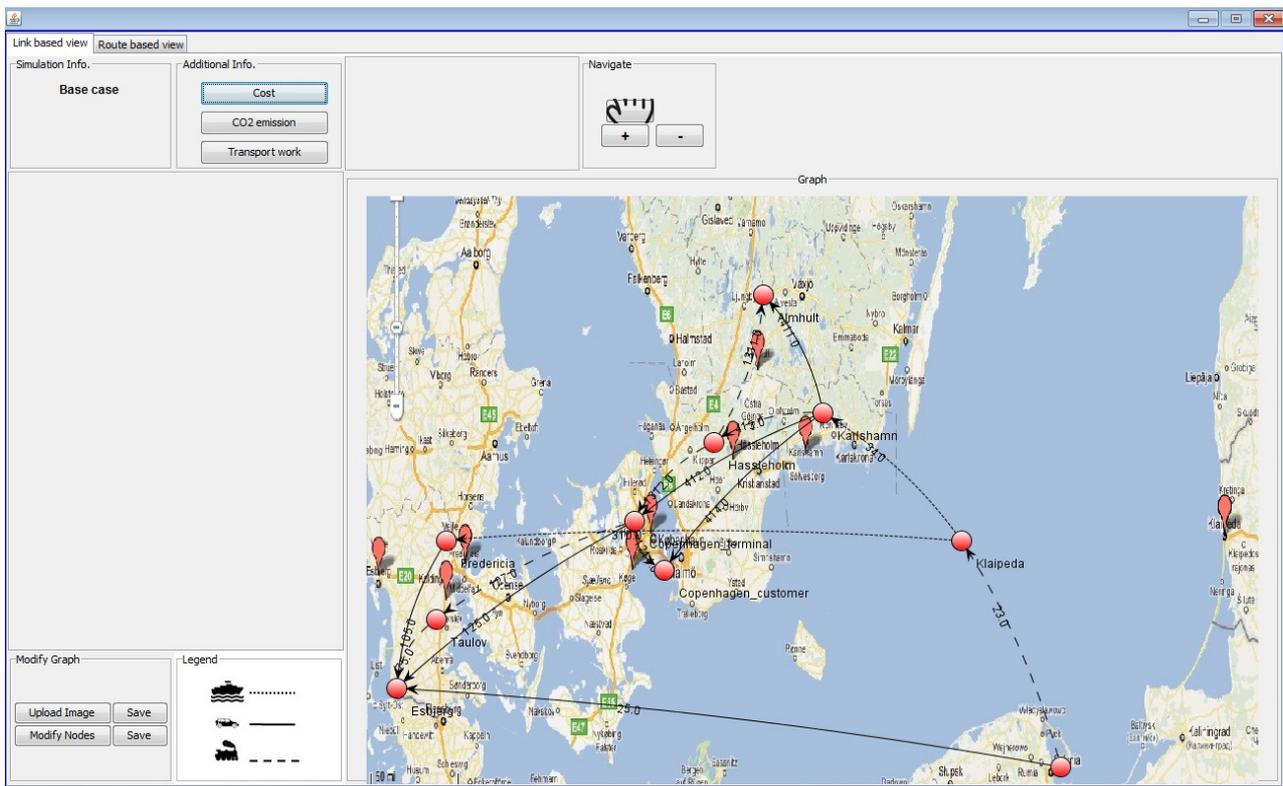


Figure 18: Link Based View form of visualization tool; pushing the cost button

### Func4. Route based information on Graph:

For developing this feature we designed more complex queries to retrieve suitable data on each route. Then, we send this data to the Graph Listener to show them in the graph. This Listener will also be activated whenever the user pushes the related button of each data to show related information on each route. One

screenshot of the Link Based View is shown below (Figure 19). Also the charts will be drawn in the Route Based View, because the charts are supposed to compare the information for all routes of a scenario to each other.

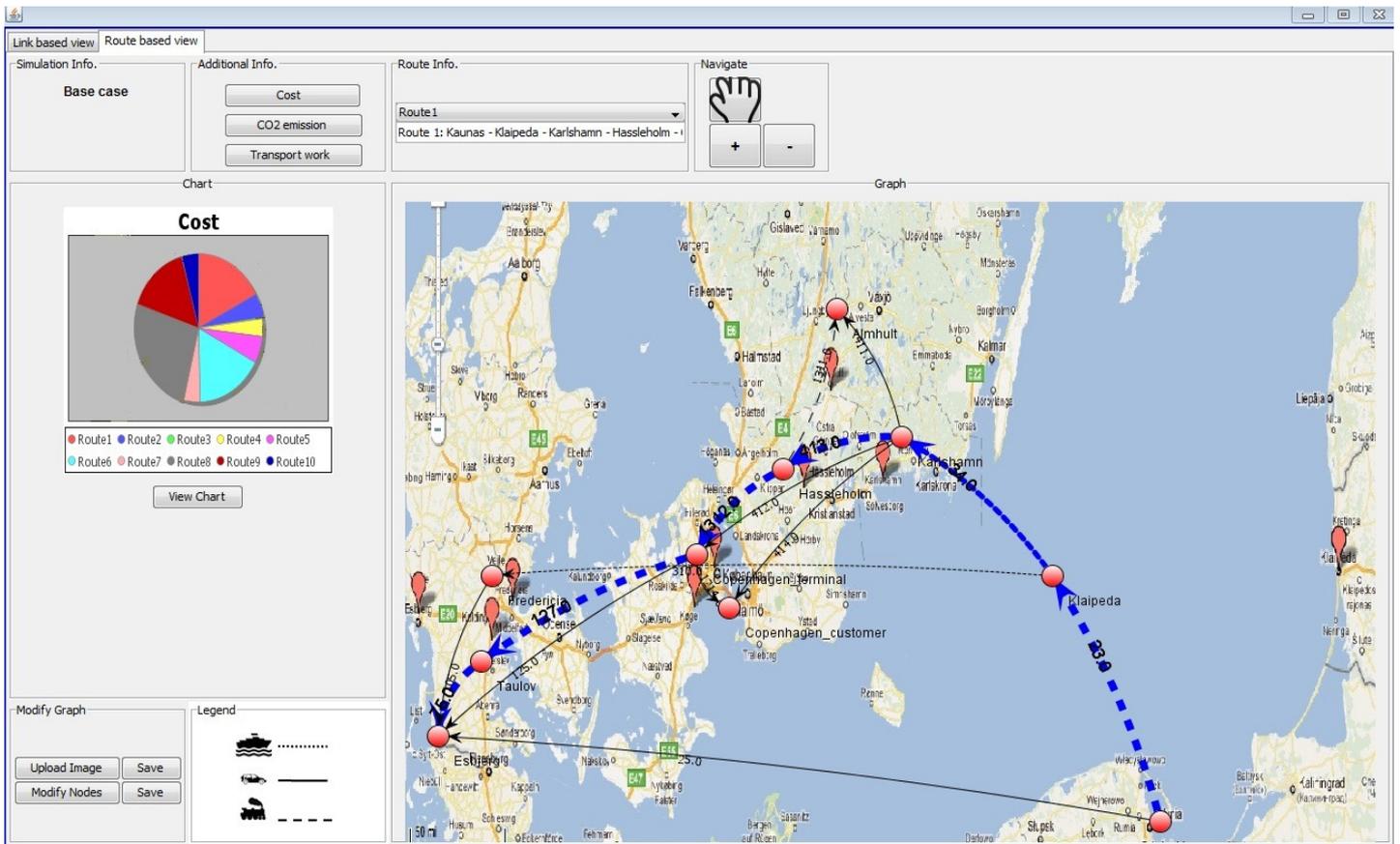


Figure 19: Route Based View form of visualization tool; pushing the cost button, viewing the related chart

#### Func5. Add a map as background:

This feature is implemented by adding an ability to the tool which enables user to select an image and upload it. Then for the next run of the visualization tool, the tool will show the right background which is selected and uploaded by user for each scenario. The benefit of this feature is that in case of existence of two nodes which are too far from each other, the user can customize the map and cut the

empty gaps between the nodes and upload it as a new background for the current scenario. This tool has also this capability to move the nodes and change their location on the map. In this case, the user can upload a map and check if the positions of nodes on the map are suitable or not. Then, he/she can move the nodes until reaches a desirable position. Figure 20 shows the file uploading window of the tool.

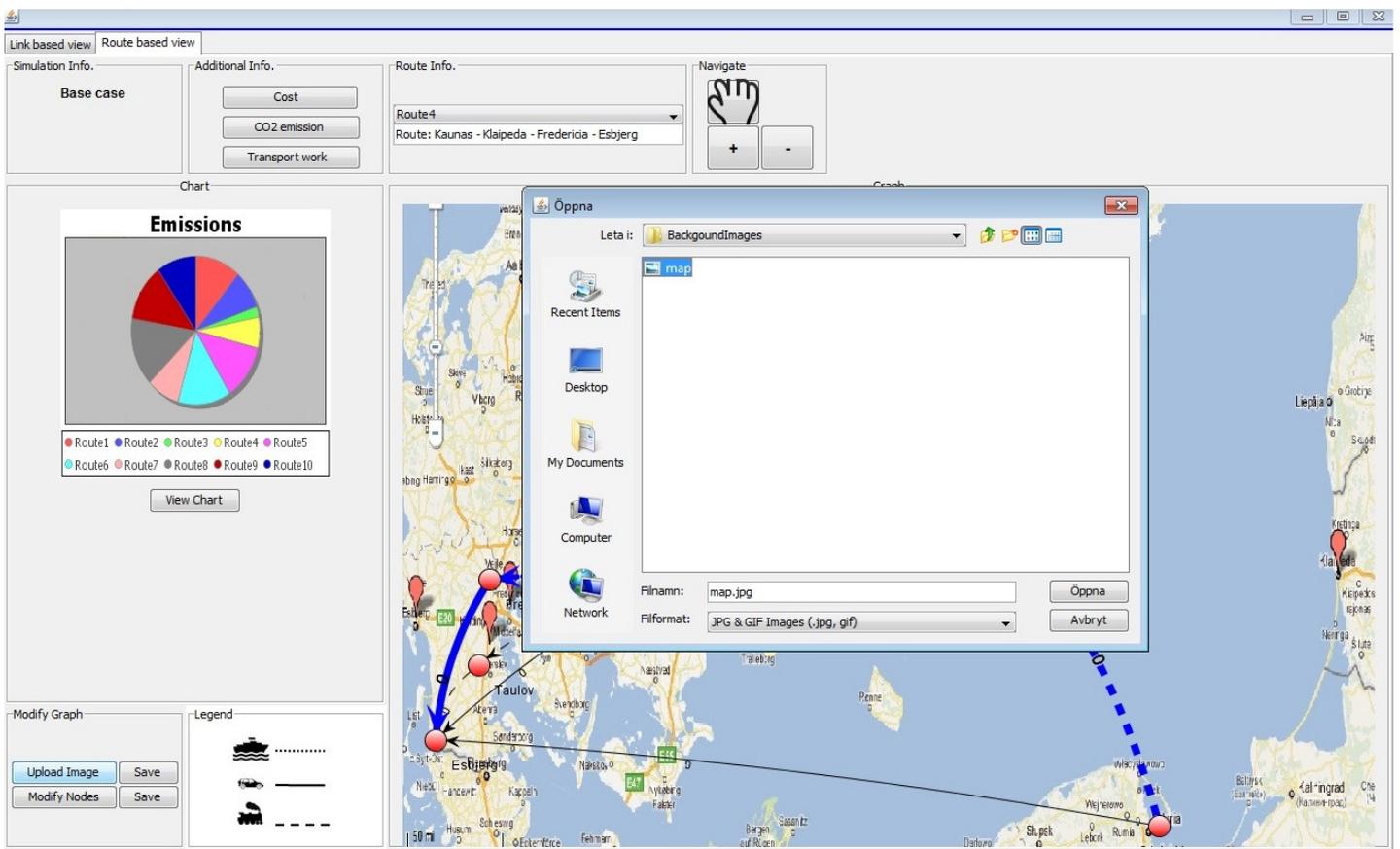


Figure 20: Uploading a map as background image

## 4.6 Evaluation

To evaluate the quality of answers to the research questions we have to investigate validity of findings and reliability of selected methods. In other words, validity means that we have found correct data, and reliability of our research methods could be proved by repeatability of research by other researchers. To increase validity of research we have selected different methods to collect data and observe facts. We have done several interviews with users, experts, and the TAPAS development team. Also, we reviewed the published documents of TAPAS; moreover, the real data of database was reviewed. Therefore, we tried to use triangulation to discover facts in our research. To increase reliability and repeatability of research, we tried to describe all steps and activities in this report which could be used by other researchers. Also, since a considerable part of our research was developing the visualization tool, we have conducted black box and white box testing which help us to evaluate technical aspects of the developed tool.

## 5 Conclusion

As we discussed before, visualization techniques could be used to increase the readability and usability of transportation simulation systems. To investigate how visualization could be applied to transportation context, we studied TAPAS as a real case and tried to specify user requirements related to visualization, design appropriate functionalities, and develop a visualization tool for TAPAS. Our findings from this case study are concluded for each research question as below:

*RQ1: What are the requirements of users who want to view the results from a transportation system simulator in a visual representation?*

**Req1. Showing input and output of simulation:** Interviews and results of questionnaires show that users of TAPAS are interested in seeing both outputs and inputs of each simulation in the same view.

**Req2. Showing different scenarios concurrently:** In order to compare various scenarios and their related output, users wanted to be able to see results of different simulation scenarios concurrently.

**Req3. Showing simulation information based on links and routes:** users need to see important results of simulation information (cost, emission, and transport work) based on different routes and also based on links between nodes.

- **RQ2:** What are the important features for increasing readability and usability of a visualization in the context of transportation systems?

By mapping visualization techniques with informational requirements of users we extracted important features and functionalities for enhancing readability of simulation outputs as follows:

- **Func1. Show as a graph:** Displaying destinations, departures, and routes in a graph. Using color, edge thickness, and edge styles to differentiate links; regarding simulation results in response to the Req3.

- Func2. Draw charts: Using charts to illustrate important results of system and make comparison between different scenarios in response to the Req2.
- Func3.Link based information on Graph: Showing information of cost, emission, amount of transport work considering Req3.
- Func4. Route based information on Graph: Information of cost, emission, amount of transport work with regard to Req3.
- Func5. Add a map as background: to increase usability and readability of the graph regarding Req1 and Table 2.

***RQ3:** How can visualization techniques be implemented to meet the requirements of the users?*

To answer this question, a visualization tool has been developed. The visualization techniques have been used in order to develop the functionalities of the tool.

As an overall conclusion; according to our findings from the TAPAS case; first of all, visualization related requirements of users are identified and analyzed for each group of users. Then the visualization capabilities are applied to each requirement. In the next step, the main features and functions to visualize outputs of system are designed. Next, the current system is studied and evaluated in terms of reusability of database or software functions in the visualization tool. Next step was selecting appropriate technologies for development. Finally, visualization tool is developed and it will be evaluated by system users to ensure that the developed tool could satisfy their requirements.

## **5.1 Future Work**

We will hold a workshop for the system users to present the visualization tool functionalities and requesting them to evaluate the tool and send their feedback to improve the tool and deliver the final version to be used by users. However, this tool could not be used before improving the TAPAS system and we will present the tool using

dummy data. That is, some improvements will be issued to the TAPAS development team to fill the visualization database.

In addition, we should have a real-world evaluation for the visualization tool. We will evaluate the results of this research regarding real-user's feedback.

As a future work, we have planned to send a questionnaire to users of TAPAS to find out if the visualization tool could contribute to readability and usability of results or not.

A possible future improvement of the visualization tool could be visualizing the input of TAPAS. This visualization tool can be improved by investigating in methods and techniques for visualizing input of TAPAS. So that, we can use the visualization tool for designing different scenarios.

## **Appendix I: Interview with expert users**

### **1. Interview questions**

As a part of this research, an Internet-based interview is conducted with two of the TAPAS users. The interview has been done as a structured Internet-based interview with a set of pre-defined questions. Interviewees are expert users of TAPAS and they are also the designers and developers of this simulator. Thus, they are aware of all probable problems and deficiencies of TAPAS. The questions are listed below:

1. What is the specification of TAPAS output? Which data is important to show as a visual model?
2. How you get output data from TAPAS? Is there currently any visualized output or not?
3. What are major problems in the process of monitoring the results of simulation which cause difficulties and interrupts in your work?
4. How much time you spend as an outcome of above problems?
5. Can visualization help you in understanding the results? If yes, what is your suggestion for design of a visualize model?

## 2. Interview transcripts

In this part the interviewees' responses to the interview questionnaire are presented.

### First interview

---

Here the responses of the first interviewee to the above questions are presented:

1. When running a simulation experiment with TAPAS the agents decide and agree upon which transportation and production activities that should be simulated. The simulated transportation and production activities are described as transported distance between the selected producer node and the consumer node, mode choice, transport time, consignment size, frequency, etc. These activities can be further described in terms of different performance measures; costs for transport and production, tax incomes, amount of emissions from transportation, possible delays from the planned delivery, transport work, traffic work, etc.

Some results which are interesting to visualize (which we have been analyzing in simulation experiments) are:

- Share of amount of TEUs per route (and customer).
- Modal split (e.g., in terms of percentages of the amount of vehicle kilometers for the different traffic modes) for different alternatives.
- Share of amount of emissions per route etc.
- The transport cost per route as an indexed value (to see the differences) – and the actual transport costs.

Also, visualizing the settings for the different alternative scenarios would be good. (Please see some TAPAS-publications to get an idea of the kind of results which are produced from TAPAS and which results that are of interest.)

2. We have been using a database where the output is stored to facilitate collecting the desired output. We used scripts to gather the output. (A problem with using the database is that it has to be updated when TAPAS is updated, and that can be time-consuming.)

When we present simulation results from TAPAS, we put together tables and graphs to illustrate the output. (See for instance Section 4 in the attached paper [tapas.pdf](#)).

For the ITS World Congress 2009 there was a project to demonstrate transport chains, where a TAPAS simulation was demonstrated and the results were illustrated. On the welcome page of the TAPAS homepage (<http://www.bth.se/tapas>) you can see some screen shots of the demonstration. We have a demo of this, but it is too large to send over e-mail. A problem with this project is that it was coded just for a specific scenario, so it is not re-usable for other scenarios.

3. When collecting the output from the database it is easy to do mistakes when formulating the scripts of the desired output, etc. It is quite time-consuming to first collect all data from the database, and then make diagrams or tables of the results. Often you need to make the diagrams many times since you discover some small error, or you want to change something, so that also takes a lot of time.

There are large amounts of input and output to handle which makes it complex to get an overview of the results and keep track of the causes and effects. If the results could be visualized in an appropriate way this would facilitate the analysis of the results.

4. I don't know, but the whole process as it is today of collecting the results is time-consuming.
5. Yes, the analysis would be facilitated. If the results could be appropriately visualized it would also facilitate for other people to better understand the results, and TAPAS.

It's difficult to come up with a suggestion of a design at this moment, but some general points are:

- Also include the different settings for the studied alternatives/scenarios, and also some of the other input, such as possible transport alternatives/routes.

- Having a rather simple model probably makes it easier to understand and analyze the results, but it would be good if it is possible to choose to visualize more details.
- If the results could be presented in a graphical sketch of the transport network (possible transport alternatives, the producer nodes and the consumer nodes), it would be easier to get an overview of the studied scenario and the results.
- To easily view the differences (in terms of route choice, modal split, costs, emissions, etc.) between the alternatives/scenarios is important.

## Second interview

---

The second interviewee responds to the questions as below:

1. TAPAS produces different types of output data. Basically it logs information about all physical activities (transports, production/manufacturing, loading, unloading, departures, arrivals of vehicles, etc.) that occurs during a simulation run. It also logs most input.

Not all of this output is relevant to display. It is more important to display different types of performance measures, such as:

- Route choices (e.g., in percentage between different routes, or ton/route)
  - Mode choices (e.g., in percentage between different modes, or ton/mode)
  - Transportation per link (e.g. number of vehicles per link)
  - Transport cost (should be some SQL-script for that)
  - Different types of emissions (CO<sub>2</sub>, etc.)
  - Different types of other costs
2. Currently we only get data by querying a SQL-database where all output (and input) is logged. A project group connected to BTH created a demonstrator (GRACE), which showed the output for a specific scenario that was studied with TAPAS. Unfortunately all people involved in the project left BTH and the knowledge about how to run and work with GRACE disappeared with them. However, a couple of screen dumps can be found on <http://www.bth.se/tapas>, and I have a movie, which I will send to you somehow.
  3. As I see it, the problem is that it takes too much time to generate meaningful output (e.g., in form of graphs) since data has to be processed manually.
  4. It is hard to state exactly how much time is needed to generate meaningful output. Some output needs less time to process than other. But an example might be of help for you. For the papers that Linda sent to you earlier (they contain the same experiments) I spent approximately around one week (maybe a bit less) processing data for 36 simulation runs.

5. I answer that visualization definitely would help understanding the results. However, it is difficult for me to say exactly what an appropriate design for a visualization model would look like. I believe that different types of graphs and charts describing the performance measures mentioned above a natural direction to go.

## Appendix II: Interview with end users

These are the pictures which are provided from the answers of the end user to our online questionnaire:

Response for: Anonymous

Completed: 6/21/2011 4:24 AM EDT

---

**Text Block :**

This questionnaire is being conducted by Banafsheh Hajinasab. I am student of Malmö University, doing my Master thesis in Computer Science under the supervision of Prof. Paul Davidsson. The aim of my thesis is to find the best way of visualizing the outputs of transportation simulation studies. This questionnaire is designed to find out the requirements of end-users of transportation simulation systems. Your contribution by answering this questionnaire is much appreciated.

---

**1 How do you use simulation of transports in your work?**

**Answer**

I am mainly interested in the simulation results.

I work with the simulation software.

Other

---

**2 Which information from simulation, are you interested in?**

**Answer**

Route choices (The percentage of using different routes from one source to a destination which is defined as the amount of goods transported through each route)

Modal choices (The percentage of different transport modes which are used from one source to a destination, such as rail, road and sea)

Amount of emissions

Transport costs.

Other types of costs.

Other: transportwork and trafficwork

3 Are you interested in comparing different scenarios of simulation? (That is, comparing the simulation results of applying different measures, such as, an infrastructure investment, to the transportation system)

Answer
<input checked="" type="checkbox"/> Yes
<input type="checkbox"/> No

4 Do you think visualization of simulation results can help you in your work?

Answer
<input checked="" type="checkbox"/> Yes
<input type="checkbox"/> No

Text Block :

In the questions below, the information "on link" refers to information concerning a link connecting two nodes, which is a part of one or more routes. The information "on route" refers to the information concerning a whole route, from a source to a destination, which may consist of several links. Please tick the box of the most valuable type of visualization below (in 5 to 7). (Tick both if both are equally valuable and tick none if none is valuable). Also feel free to add comments.

5 Transport work ( The information related to the amount of transport work)

**Answer**

Information for each link

**Link based view**

Source City1  
Destination City6

Cost  
CO2  
**Transport work**

.....  
 .....  
 .....

---

Information for each route

**Route based view**

Source City1  
Destination City6

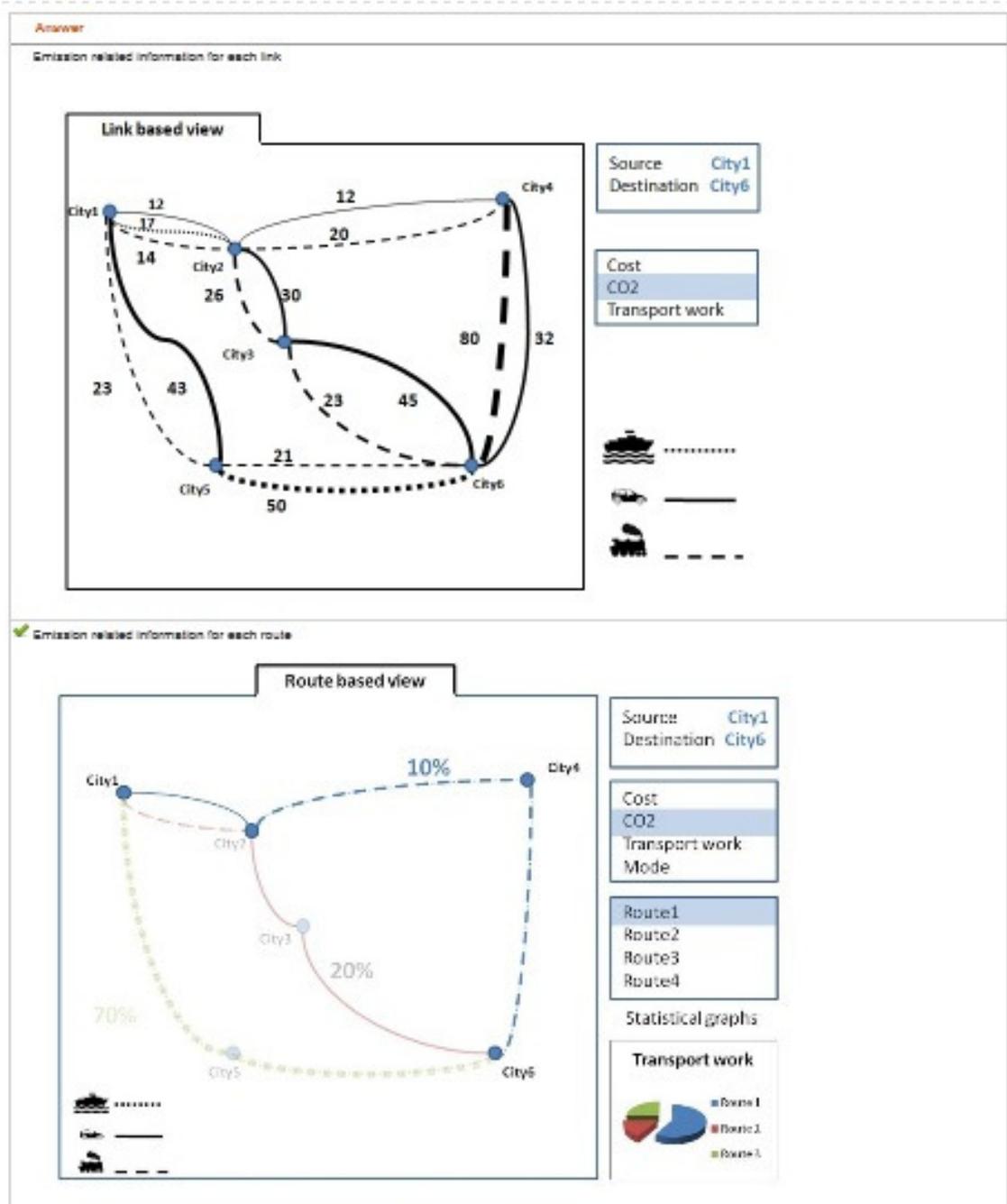
Cost  
CO2  
**Transport work**  
Mode

Route1  
 Route2  
 Route3  
 Route1

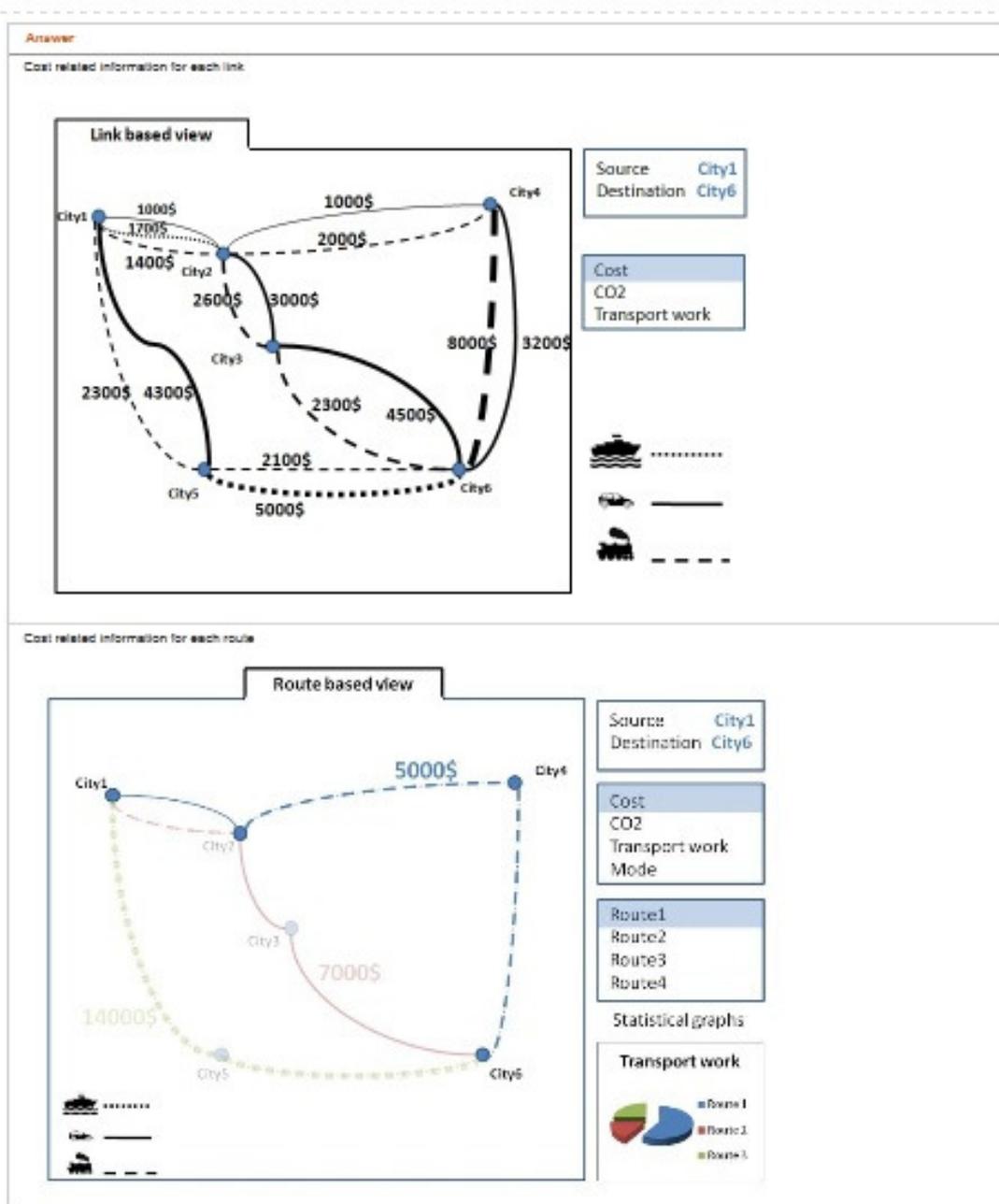
Statistical graphs

**Transport work**

8 Emissions ( The information related to the amount of emissions)



7 Cost ( The information related to the amount of cost )



## Appendix III: Sample development codes

### A part of designed queries of TAPAS visualization tool

```
sql = "select start_node_id from visualization_information where simulation_nr =" + selectedSimNr + ""
      group by start_node_id";
Integer startNodeID = null;
Integer endNodeID = null;
try {
    pst = conn.prepareStatement(sql);
    rs = pst.executeQuery();
    while (rs.next()) {
        startNodeID = rs.getInt("start_node_id");
        for (int c = 0; c < counter; c++) {
            if (allNodes[c].getNodeID().equals(startNodeID.toString())) {
                startNode = allNodes[c];
            }
        }
    }

    String sql2 = "select end_node_id from visualization_information where start_node_id =" + startNodeID
    + "" and simulation_nr =" + selectedSimNr + "" group by end_node_id";
    pst2 = conn.prepareStatement(sql2);
    rs2 = pst2.executeQuery();
    while (rs2.next()) {
        endNodeID = rs2.getInt("end_node_id");
        for (int c = 0; c < counter; c++) {
            if (allNodes[c].getNodeID().equals(endNodeID.toString())) {
                endNode = allNodes[c];
            }
        }
    }

    sql = "select transportation_mode_id,cost from visualization_information where simulation_nr = " +
    selectedSimNr + "" and start_node_id=" + startNodeID + "" and end_node_id=" + endNodeID + "" group
    by transportation_mode_id,cost";
    try {
        pst3 = conn.prepareStatement(sql);
        rs3 = pst3.executeQuery();
        while (rs3.next()) {
            Edge drawEdge = new Edge();
            transportModeID = rs3.getInt("transportation_mode_id");
            drawEdge.setDrawEdgeCost(rs3.getDouble("cost"));
        }
    }
}
```

```
        sql = "select name from transportation_mode where simulation_nr = " + selectedSimNr + " and  
id=" + transportModeID + """;  
        try {  
            pst4 = conn.prepareStatement(sql);  
            rs4 = pst4.executeQuery();  
            while (rs4.next()) {  
                transportModeName = rs4.getString("name");  
                if (transportModeName.equals("sea")) {  
                    drawEdge.setDrawEdgeCO2(drawNodePY);  
                    drawEdge.setDrawEdgeType("SEA");  
                }  
                if (transportModeName.equals("rail")) {  
                    drawEdge.setDrawEdgeType("RAIL");  
                }  
                if (transportModeName.equals("road")) {  
                    drawEdge.setDrawEdgeType("ROAD");  
                }  
                g.addEdge(drawEdge, startNode, endNode, EdgeType.DIRECTED);  
            }  
        } catch (Exception e) {  
            JOptionPane.showMessageDialog(null, e);  
        }  
  
        }  
    } catch (Exception e) {  
        JOptionPane.showMessageDialog(null, e);  
    }  
}  
}  
} catch (Exception e) {  
    JOptionPane.showMessageDialog(null, e);  
}
```

## A part of designed renderers of TAPAS visualization tool

```
BasicVisualizationServer<Node, Edge> vv = new BasicVisualizationServer<Node,Edge>(layout);
vv.getRenderContext().setEdgeStrokeTransformer(
    new Transformer<Edge, Stroke>() {
        float dash[] = {2.5f};
        final Stroke SEA = new BasicStroke(1.0f, BasicStroke.CAP_ROUND,
            BasicStroke.JOIN_MITER, 1.0f, dash, 0.0f);
        float dash2[] = {10.0f};
        final Stroke RAIL = new BasicStroke(1.0f, BasicStroke.CAP_BUTT,
            BasicStroke.JOIN_MITER, 10.0f, dash2, 10.0f);
        final Stroke ROAD = new BasicStroke(1.0f, BasicStroke.CAP_BUTT,
            BasicStroke.JOIN_MITER, 10.0f, null, 0.0f);

        public Stroke transform(Edge e) {
            if (e.getDrawEdgeType().equals("SEA")) {
                return SEA;
            } else if (e.getDrawEdgeType().equals("RAIL")) {
                return RAIL;
            } else {
                return ROAD;
            }
        }
    });
```

## References

---

- [1] Brunner, D., et al. Toward increased use of simulation in transportation. In proceedings of 1998 Winter Simulation Conference, USA, 1998.
- [2] C. Ware, Information Visualization: Perception for Design, Morgan Kaufman, USA, 2004.
- [3] Davidsson, P., et al. Multi agent based simulation of transport chains. In Proceedings of the 7th International Joint Conference on Autonomous agents and multi agent systems (AAMAS'08), Estoril, Portugal, 2008.
- [4] Ramstedt, L., Davidsson, P., Holmgren, J., and Persson, J.A., On the Use of Micro-level Simulation for Estimation of the Effects of Governmental Control Policies, 11th World Conference on Transport Research, Berkeley, USA, 2007.
- [5] S.K. Card, J.D. Mackinlay, and B. Shneiderman (eds), Readings in Information Visualization, Morgan Kaufmann Publishers, 1999.
- [6] Herman I, Melanc, on G, Marshall MS. Graph visualization and navigation in information visualization: a survey. IEEE Transactions on Visualization and Computer Graphics 2000; 6: 24 – 43.
- [7] Di Battista, G., Eades, P., Tamassia, R., and Tollis, I. G., Graph Drawing: Algorithms for the Visualization of Graphs. Upper Saddle River, N. J: Prentice Hall, 1999.
- [8] Kaufmann, M. and Wagner, D., Drawing Graphs: Methods and Models, volume 2025 of Lecture Notes in Computer Science. Springer, 2001.
- [9] Brasseur, Lee E., Visualizing technical information: a cultural critique. Baywood Pub, Amityville, N.Y, 2003.
- [10] M. Bergkvist, P. Davidsson, J. A. Persson, L. Ramstedt, A Hybrid Micro-Simulator for Determining the Effects of Governmental Control Policies on Transport Chains, in: Lecture Notes in Computer Science: Vol. 3415. Multi-Agent and Multi-Agent-Based Simulation, Springer, 236–247, 2005.
- [11] Ghoniem, M., et al., On the readability of graphs using node-link and matrix-based representations: a controlled experiment and statistical analysis. Information Visualization, 4 (2). 2005.

[12] Davis, F. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3): 319-340. 1989.

[13] Bresciani, S., et al., "Beyond knowledge visualization usability: toward a better understanding of business diagram adoption", *Proceedings of 2009 13th International Conference Information Visualisation*. Washington, DC: IEEE Computer Society, 2009.

[14] Di Battista, et al., *Algorithms for Drawing Graphs: An Annotated Bibliography. Computational Geometry Theory and Applications*. vol. 4. 1994.

[15] B. J. Oates. *Researching Information Systems and Computing*. SAGE Publications, 2006.

[16] March, S.T., et al., *Design and Natural Science Research on Information Technology, Decision Support Systems (15) 1995*, pp. 251-266.

[17] Yin, R., *Case study research; Design and methods*, Sage publications, USA, 2003.

[18] Plaisant, C. *The Challenge of Information Visualization Evaluation*, *Proc. of Conf. on Advanced Visual Interfaces*, ACM, New York, 2004.

[19] Mazza, R., *Introduction to Information Visualization*. Springer, London, 2009.

[20] C. Chen, *Information Visualization: Beyond the Horizon*, 2nd edition, Springer, 2004.

[21] Bederson, B and Shneiderman, B. *The Craft of Information Visualization: Readings and Reflections*, Morgan Kaufmann, 2003.

[22] "Graph drawing" Internet: [http://en.wikipedia.org/wiki/Graph\\_drawing](http://en.wikipedia.org/wiki/Graph_drawing) [FEB.20.2011]

[23] Manning, J.B., *Geometric symmetry in graphs*. PhD thesis, Purdue University, 1990.

[24] Sugiyama, K., et al, *Methods for Visual Understanding of Hierarchical Systems*, *IEEE Trans. Syst. Man Cybern.*, SMC-11, no. 2, 109-125, 1981.

[25] Batini, C., et al, *What is a Good Diagram? A Pragmatic Approach*, In *Proc. 4th Internat. Conf. on the Entity Relationship Approach*, 1985.

[26] Purchase, H.C., et al, *Validating Graph Drawing Aesthetics*, In F. J. Brandenburg, editor, *Graph Drawing (Proc. GD '95)*, vol. 1027 of *Lecture Notes Comput. Sci.*, pp. 435-446. Springer-Verlag, 1996.

[27] D. Ferrari and L. Mezzalana. *On drawing a graph with the minimum number of crossings*. Technical Report 69-11, Istituto di Elettrotecnica ed Elettronica, Politecnico di Milano, 1969.

[28] R. Tamassia. *On embedding a graph in the grid with the minimum number of bends*. *SIAM J. Computing*, 16(3):421-444, 1987.

- [29] H. Trickey. Drag: A graph drawing system. In Proc. Int. Conf. on Electronic Publishing, pages 171-182. Cambridge University Press, 1988.
- [30] P. Eades. A heuristic for graph drawing., *Congressus Numerantium*, 42:149-160, 1984.
- [31] R. Lipton, S. North, and J. Sandberg. A method for drawing graphs, In Proc. ACM Syrup. On Computational Geometry, pages 153-160, 1985.
- [32] H.C. Purchase, R.F. Cohen, and M. James. An experimental study of the basis for graph drawing algorithms. *A CM Journal of Experimental Algorithmics*, 2(4), 1997.
- [33] H.Purchase, Which aesthetic has the greatest effect on human understanding? In G. Di Battista, (Ed) *Graph Drawing 97*, 1353 of *lecture Notes in Computer Science*, Springer Verlag, 284-290, 1997.
- [34] R. Tamassia, et al., Automatic Graph Drawing and Readability of Diagrams, *IEEE Trans. Syst. Man and Cybern.*, SMC-18, no. 1, 61-79, 1988.
- [35] Aris, A., *Visualizing & Exploring Networks Using Semantic Substrates.*, PhD thesis, Department of Computer Science. College Park, University of Maryland, 2008.
- [36] B. Shneiderman and A. Aris., Network visualization by semantic substrates. *IEEE Transactions on Visualization and Computer Graphics*, 12(5), 2006.
- [37] R. A. Becker, S. G. Eick, G. J. Wills: 'Visualizing Network Data', *IEEE Transactions on Visualizations and Graphics*, Vol. 1, No. 1, pp. 16-28, 1995.
- [38] B.J. Breitkreut., et al. "Osprey: a network visualization system." *Genome Biology* 4: R22, 2003.
- [39] CL Hixson III. *Visualization for Project Development*, . NCHRP Synthesis 361. National Cooperative Highway Research Program. Transportation Research Board, Washington,D.C., 2006.
- [40] Hughes, R.G. *Toward an Expanded Research Agenda for Visualization in Transportation: Incorporating SAFETEA-LU Directives for 'Planning'*. Institute for Transportation Research and Education, North Carolina State University, Raleigh, NC, 2008.
- [41] *Visualization in Transportation: A Guide for Transportation Agencies*, AASHTO Task Force on Environmental Design, American Association of State Highway and Transportation Official (AASHTO), December 2001. Internet : <http://design.transportation.org/Documents/VisualizationGuideJuly2003.pdf> ,Accessed [APR.12.2011].
- [42] Hannon, J.J. *Emerging Technologies for Construction Delivery*. NCHRP Synthesis 372. National Cooperative Highway Research Program, Transportation Research Board, Washington,D.C., 2007

- [43] Sulbaran, T., “Distributed Virtual Reality Scheduling in Future Construction Companies,” Proceedings of the 2005 ASCE International Conference on Computing in Civil Engineering, Cancun, Mexico, July 12–15, 2005.
- [44] Hughes, R.G. Freight data visualization: A “Pivotal” point in the development of visualization applications in transportation. Submitted for publication and presentation at “The 2011 Transportation Research Forum”, Long Beach, CA, 2010.
- [45] Kutz, M. Handbook of Transportation Engineering, McGraw Hill Publishing, 2004.
- [46] Hoff, A, et al. “Design Simulation and Visualization of Automated Guided Vehicle Systems in a Real Production Environment”, In Proceedings of the 9th European Simulation Symposium, eds. Winfried Hahn and Axel Lehmann, pp. 391-395, 1997.
- [47] P.J. Hoen, et al. Simulation and visualization of a market-based model for logistics management in transportation, In proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems, 2004, pp. 1218–1219.
- [48] IEEE 1220 Standard, Application and Management of the Systems Engineering Process, IEEE Standards Dept., NY, December 1998
- [49] Wymore, W., Model-Based Systems Engineering, CRC Press, Boca Raton, 1993.
- [50] Post, G., et al. Management information systems: Solving business problems with information technology. New York: McGraw-Hill Irwin, 2006.
- [51] B. Boehm; A Spiral Model of Software Development and Enhancement, IEEE Computer, vol. 21, no. 5, May, 1988, pp .61–72.
- [52] Johan Holmgren, On the integration of optimization and agent technology for transportation and production planning, Doctoral Dissertation Series No. 2010:09, Blekinge Institute of Technology, Sweden, 2010.
- [53] Johan Holmgren, et al. TAPAS: A multi-agent-based model for simulation of supply chains, Preprint submitted to Elsevier, 2010.
- [54] L. Ramstedt, Transport policy analysis using multi-agent-based simulation, Doctoral Dissertation Series No. 2008:09, Blekinge Institute of Technology, Sweden, 2008.
- [55] Holmgren, et al. An Agent Based Simulator for Production and Transportation of Products, 11th World Conference on Transport Research, Berkeley, USA, 2007
- [56] “JUNG - Java Universal Network/Graph Framework” Internet : <http://jung.sourceforge.net/> [FEB.15.2011]