Experiences with Establishing a Simulation Scenario of the City of Duisburg with Real Traffic Volume

Xiaoyi Ma, Xiaowei Hu, Thomas Weber and Dieter Schramm

Department of Mechatronics, University of Duisburg-Essen, 47047 Duisburg, Germany; xiaoyi.ma@uni-due.de (X.M.); thomas.markus.weber@uni-due.de (T.W.); dieter.schramm@uni-due.de (D.S.)
* Correspondence: xiaowei.hu@uni-due.de

1. Introduction

Traffic simulation tools can be divided into macroscopic, microscopic, mesoscopic scale simulations. Microscopic traffic simulation models each vehicle and its dynamics individually [1]. Moreover, microscopic traffic simulation has been developed into a cost-effective and powerful tool for traffic management and planning as well as the automotive industry. However, the accuracy of the simulation scenario is closely related to the usefulness of the simulation results. For realistic simulation scenarios, a road network, traffic demand, real traffic lights, and infrastructures are needed as input data [2]. The quality of all the input data is crucial for the accuracy of the simulation.

SUMO (Simulation of Urban MOBility) is a microscopic, open-source and space-continuous, time-discrete road traffic simulation package [1]. With the advantages of easy programming, scalability, input data extensiveness, and verification convenience, it has been widely used by many researchers in building realistic simulation scenarios. Krajzewicz et.al. [3] introduced the network import of the city of Gothenburg and the evaluation of the traffic flow with the scenarios in and around Berlin. Bieker [2] imported three regions’ scenarios of city Bologna from Vissim, another traffic simulation software. The scenario area was combined with three small scenarios, while the traffic demand contained one hour of a day. The traffic demand of this scenario was based on induction loop data provided by the local municipality. Codeca [4] rebuilt the city of Luxembourg in SUMO with buses and bus stations in a 24-h scenario. This scenario covers an area of almost 156 km² with 931 km of roads. The important traffic demand was generated from population and geographical location information and was not calibrated or verified with real traffic data. Asano [5] reproduced Kobe city with an area of about 130 km². The traffic...
demand was also generated by geographical location information, gathered from mobile phone position data. These simulation results were compared with the measurements from the traffic censuses. However, the simulation results did not reflect the real traffic situation in every respect.

In the establishment of realistic traffic scenarios, there are always some factors that affect the overall accuracy. The previous studies of the authors about the traffic scenarios in the city of Wuhan [6] also illustrated that the traffic demand generated from the geographical population distribution only roughly represents the status of the traffic and that high precision simulation scenarios must be built from precise traffic input data. In this study, thanks to the support from Wirtschaftsbetriebe Duisburg and the city of Duisburg, the traffic light positions, dynamic traffic light plans, OD (Origin/Destination) matrix and induction loop data are all available for this work. On this basis, this article is aiming to present the experiences of building a high-precision simulation scenario of the whole city of Duisburg with OD matrix data. The simulated traffic volume was also compared with the real traffic volume detected by the induction loops.

2. Road Network

The city road network can be obtained by several methods: using Wireless Sensor Networks (WSN) [7], acoustic signal [8], etc. Like in our previous study about the city Duisburg inner ring area [9], the road network used in this work was extracted from OSM (Open Street Map) [10]. A rectangular area encompassing the area we need was extracted from OSM. We used JOSM [11] to edit the exported network from OSM and determined the boundaries of the road network. During the process of importing the road network into SUMO, the public transportations lines such as railway, subway, and bus-only lanes were ignored. This study focuses on the individual vehicles on the road, which are the main reason for the occurrence of traffic jams. What is more, there were no pedestrians or cyclists simulated.

The calibration of the road network is a time-consuming task. However, it is important for the precision of the simulation system.

To ensure that each street is correct, we compared the network from OSM with Google Street View [12]. Common defects of the road network are:

- Divided intersections. Large intersections are always divided into smaller ones, especially those with a cover (e.g., viaducts and green belts) in the middle. The extra intersections are often accompanied by the extra unlinked traffic lights, which increase the possibility of unrealistic traffic jams.
- Occurrence of extra roads. Sometimes, the non-motorized roads or the dedicated roads are included in the normal motorways, however in street view, some isolation piles or stop signs can be found at the road entrance.
- Missing roads. There are no two-directional ways in SUMO. However, they are common in residential areas. Many people have a single-lane road leading to their garage, responsible for both incoming and outgoing directions. In SUMO, we had to add the reverse lanes for these kinds of roads. What is more, there were also other road-missing conditions. An example can be seen in Figure 1. In this case, a part of a ring road in the parking lots is missing. It causes a dead-end in Figure 1a. After comparing with the satellite map in Figure 1b and the street view map in Figure 1c, the missing sections were manually filled in. The modified road network of the parking lots is in Figure 1d.
- Reverse directions. In some cases, some lanes on a road may be in the wrong direction, which affects their connection with other roads. This kind of mistake is less likely than the other three but is also a common error in road network imported from OSM.
These mistakes of the road network can only be compared with real scene and adjusted manually.

Figure 1. Example of a missing road in the road network. (a). Road network of the parking lot imported from Open Street Map (OSM). (b). Satellite map from Google map of the parking lot. (c). Street view from Google map, there is indeed a road missing in the imported map. (d). The missing road is manually added into the road network.

3. Traffic Demand

An OD matrix is an effective data form in the field of traffic planning with a wide range of applications. The OD matrix specifies the traffic demand between the origin and destination points in a network. The OD matrix data usually consists of two parts: one describes the distribution and numbering of points/areas in the road network, the other presents the volume of traffic from one point/area to another. The OD matrix used in this work contains 750 points/areas and covers an area over 800 km². Of all the areas, 596 are in Duisburg, and 153 are around Duisburg. The outer city areas contain the cities of Mülheim an der Ruhr, Oberhausen, Dinslaken, Hünxe, Wesel, Rheinberg, Kamp-Lintfort, Neukirchen-Vlyn, Moers, Krefeld, as well as the northern part of city Düsseldorf and Ratingen. The time length of the OD matrix is 24 h a day. The passenger cars and trucks are divided into two O-format files.

3.1. Polygons and TAZ

The original OD matrix in this study was given in shapefile. Hence, we used polyconvert in SUMO to transform the boundary lines in the shapefile into the polygons in SUMO. From polygons, traffic assignment zones (TAZ) are described and listed by their id-number. In principle, the OD matrix describes the traffic volume from one point to another. However, there are many roads in each point/area. Thus, an additional tool written in python called edgesInDistricts.py [3] in SUMO was used for allocating each point/area of the OD matrix into the roads in the point/area according to its length. Therefore, the traffic demands between the points/areas in the OD matrix were converted into traffic volume between roads in the road network. Figure 2 shows the distribution of TAZ in SUMO. The 596 inner-city TAZ are shown in five different colors, while the 154 outer city TAZ are in olive. The road networks in the city of Duisburg are shown in black.
SUMO. The 596 inner-city TAZ are shown in five different colors, while the 154 outer city TAZ are in olive. The road networks in the city of Duisburg are shown in black.

Figure 2. Overall range of the simulated area.

3.2. Trip Generation and Distribution

A trip is used for describing the departure and arrival road/edge of each vehicle. We used od2trips in SUMO to generate trips from the OD matrix according to the distribution of the TAZ. To split an OD matrix over a day into 24 h, an additional option called “timeline” was used. A common daily timeline retrieved from cities in the western part of Germany was used [13]. For passenger cars and trucks, timelines named TGw2_PKW and TGw_LKW were used separately for describing the traffic distribution for a workday. The reading sequence of files in SUMO is line-based. It is necessary to record all the trips and sort them by the departure time. At this point, each vehicle generated according to the traffic demand has its own trip containing the departure and arrival road/edge. In the next section of this paper, specific routes are determined.

3.3. Route Generation and Dynamic Equilibrium

Dijkstra’s shortest path algorithm [14] is commonly used to find the shortest path to all nodes from a single source node in a network. In principle, the Dijkstra algorithm measures each possible path from the departure point to all the points around it at each step and finds the shortest path to the destination of the whole network in every step. In SUMO, we use “duarouter” [3] to compute routes for each vehicle. However, if all vehicles follow the shortest path generated by duarouter, vehicles with similar departure and arrival roads/edges will experience severe congestion on their shortest road, while no vehicles are on the second shortest road. In real-world, drivers certainly like to drive a shorter distance, but the current traffic situation is also an important consideration in route selection. An example can be found in car navigation; some drivers prefer the shortest distance route, while some others prefer the shortest time route. What is more, if a driver knows a special path from work to home, which avoids traffic jam on the main road during the peak hours, he would like to choose that despite a possible detour.

4. Simulation and Verification

Apart from the road network and traffic demands, traffic light plans also have a great influence on traffic simulation results [15]. A traffic light control algorithm is a complex and detailed research topic, which contains multiple kinds of control strategies. In Duisburg, not only static traffic light control systems are used, in some intersections, dynamic systems are also installed. In those intersections, the pedestrians can shorten their waiting time at the red light by touching a button on the traffic light pole. In this case, the linked traffic light group of the entire intersection will change their strategy. Nonetheless, the dynamic traffic light control system mainly affects the relationship between pedestrians and vehicles.
In this study, since there are no pedestrians in the simulation, we prefer to use static traffic light control systems for higher simulation efficiency.

The extent of the road network (out of city Duisburg) is an issue we encounter which has been launched a debate on. The range of our induction loop data are limited in the city of Duisburg. Meanwhile, the range of the OD matrix is far more extended. Three ways are carried out for the most realistic approach. First, both road network and traffic demand are set just in city Duisburg (Case A). Then, we perform a simulation with the road network in the city of Duisburg, but traffic demand with a larger extent (Case B and C). Last, both road network and traffic demand are set into a larger range (Case D).

4.1. Duisburg Interior Simulation

A simulation interior city Duisburg is performed. In this 233 km² area, there are over 23,000 roads/edges and more than 9000 nodes. Both, highways and urban ways, are included. The road network of the city Duisburg can be seen in Figure 3. For traffic demand, only interior city activities from the OD matrix are used in this case. Only interior city traffic happens in this situation, incoming and outgoing traffic of the city is set to zero. In this case, we receive a smooth simulation result. Even in peak hours, the traffic in the road network is not heavy. Although it gives us a hypothetical result, that is, assuming that no vehicle enters or leaves the city, urban traffic will be very smooth, the simulated scenario, in this case, is far from the traffic reality.

Figure 3. Map (Source: HERE map), road network, and traffic demand partition of city Duisburg.

4.2. Duisburg Simulation with Intercity Traffic

In this section, intercity traffic is added to the traffic demand. As the road network remains the same as in the last section, some trips’ departure and arrival roads/edges are out of the network. Under this circumstance, these departure or arrival roads/edges are moved to the closest edge on the network. For example, if a trip starts from Duisburg center to Düsseldorf center, the road network does not contain the city of Düsseldorf, but the trip remains as a traffic demand. Here we change the destination from Düsseldorf center to the road/edge nearest Düsseldorf center in city Duisburg. To achieve this purpose, we redefined the TAZ out of city Duisburg and distributed the outer city TAZ onto different roads/edges of interior city areas.

The simulation results show that traffic jams are more likely to occur on the streets/edges at the city limits. The reason could be, that the outer city traffic puts extra pressure on the
boundary of the city. Another reason could be that, if a trip has its departure and arrival road/edge both out of the city, originally there is a chance to finish this trip by staying out of the city. However, if we link all the outer city TAZ with its nearest road/edge (Case B), all trips with both departure and arrival road/edge outer the city must come onto the road network. It generates unrealistic traffic jams for the simulation, especially on the boundary of the city, where lots of TAZ are linked. To solve this problem, in Case C, we remove all the trips with both departure and arrival road/edge out of the city. Traffic interior the city, traffic of incoming and outgoing city Duisburg are retained in the simulation. The simulation result this time no longer has the problem of a crowded city boundary for Case C. However, in this case, the trips that would originally pass the city are also ignored in the simulation. This circumstance assumes that all the trips that have both departure and arrival road/edge out of the city do not use the road network interior of the city. Detailed simulation results will be compared with other cases and real traffic data in the following sections.

4.3. Duisburg Plus Surrounding Area

In the end, a road network of the same size as the given OD matrix is performed, and all the traffic demand from the OD matrix is implemented in the simulation. About traffic light data, only the ones in the city Duisburg are from real traffic light plans, the outside ones are the default traffic light plans provided by SUMO. Fortunately, the relatively sparse out-of-city road network requires less accuracy in traffic lights.

Table 1 shows the comparison of the traffic simulation of city Duisburg with different simulation scope. Case A is the inner city simulation that includes inner-city traffic and inner-city road network. Case B contains inner-city traffic, traffic from inside city to outside city or vice versa, and traffic from outside city to outside city. Meanwhile, Case C does not contain the traffic both departure and arrival road/edge out of the city. Both Case B and Case C have the same road network as Case A, i.e., the inner-city road network. Case D has all the traffic demand and all road networks (inside and outside of the city) in simulation.

| Simulation Scope | Case A | Case B | Case C | Case D |
|------------------|--------|--------|--------|--------|
| Traffic          |        |        |        |        |
| Inner-city       | √      | √      | √      | √      |
| In to out/out to in | x    | √      | √      | √      |
| Out to out       | x      | √      | x      | √      |
| Map              |        |        |        |        |
| Inner-city       | √      | √      | √      | √      |
| Outer city       | x      | x      | x      | √      |

4.4. Verification Points

To verify the precision of the simulation of each case, real traffic data are used. The verification data are from detectors equipped on the roads in Duisburg, supported and operated by the Wirtschaftsbetriebe Duisburg and the city of Duisburg. The detectors contain induction loops and cameras. The induction loops are installed beneath the road surface, mainly composed of induction coils and communication devices. When a metallic object (such as a vehicle) drives over the induction loops, the metallic parts of the vehicle inverse the inductance of the loop and the voltage of the circuit is decreased. Therefore, the induction loops can directly measure the time each vehicle passes. With double loops separated by a fixed distance, the detector can measure the speed of the vehicle [16,17].

Figure 4a shows the position of detectors of an intersection on the city planning map. The cameras are shown with the dotted line circle and named with K plus number. The induction loops are shown with rectangles and named with D plus number. The detectors in the city Duisburg are always located near the intersections. To compare the real data with simulated traffic, we put detectors in the simulation. In SUMO, the detectors are
added in additional files. The cameras are shown each as a yellow rectangle with a white border and middle cutting line, the induction loops are shown in pink, as can be seen in Figure 4b.

In the selection of verification points, eight roads in different areas of the city Duisburg are chosen. Intersection No. 116, 903, 834, 656, and 751 represent the roads in the north, south, west, east, and middle part of city Duisburg, respectively. Intersection No. 721, 723, 726, and 742 represent the different parts of the inner ring of city Duisburg. The locations of the selected eight road sections are shown in Figure 5. Each intersection has 2 to 12 detectors, in total, 46 detectors are used in verification.

**Figure 4.** Position of detectors of an intersection. (a) on map, (b) in SUMO.

**Figure 5.** Selected road section for verification.
The detectors keep recording traffic volume during 24 h, for each detector, the data of a whole week are provided for this work. Figure 6 shows data from an induction loops detector near the intersection of Schwanenstraße and Poststraße of the city Duisburg. Figure 6 shows that the morning peak (around 8 a.m.) and the afternoon peak (around 5 p.m.) are the busiest time of the day for this road section.

![Figure 6. Traffic volume of an induction loop.](image)

In the four simulation cases in Table 1, it is hard for the traffic demand of Case D to realize in the network because of the large-scale simulation scope. Because some routes have no access to their destinations, the traffic demand assigned to this route cannot be reflected in the simulation. Therefore, the simulation results of Case D are not representative. Other three cases are simulated separately. Figure 7 shows the simulation result of Case A, B, and C in black, blue, and red lines, and separately the real data in green box plots in four verification points. From the comparison of the real data and simulated traffic volume, the traffic volume of Case B at the beginning time is much more than other cases and the real data for intersection No. 751 and No. 656. However, the large traffic volume in Case B causes traffic jam and makes a decrease of traffic volume in the following simulation. As a result, the accuracy of the Case B simulation drops quickly as simulation time progresses. During the simulation process, individual traffic jams first appear on the fringe of the city, and then spread to the entire city. In contrast to Case B’s excessive traffic, Case A has the problem of insufficient traffic. The black lines in Figure 7 always remains at a lower vehicle quantity. Case C shows the closest result to the real data in the comparison of the three cases.

To facilitate the calculation, the detectors in the simulations record vehicle quantity every 600 s. If the simulated traffic volume is more than the minimum real traffic volume of the same time period and less than the maximum real traffic volume of the same time period of the recorded week, the time period of this case is counted by “in real range”. Figure 8 shows the calculation of the proportion of the simulation results “in real range” as simulation time increases. This figure can also represent the simulation accuracy of Case A, B, and C. Case C has always the best performance comparing to Case A and B. For the first eight simulation hours, the best reproduction rate of Case C is 72.22%.

Eventually, the simulations of city Duisburg faced the problem of large-scale traffic congestion. Therefore, the simulation after eight hours is not representative. The large-scale traffic congestion in simulation can result in several reasons. First, the traffic light plans can be a major. In this work, traffic light plan data are not available, and all the traffic lights in the simulation are using default plans in SUMO. In reality, Duisburg has both static and dynamic control traffic lights, the reasonably planned traffic lights make urban roads still run well under heavy traffic. In this work with limited traffic light data, only hours with low traffic flow are simulated together. Moreover, the road network has been modified
manually according to the street view, but there may be still mistakes in the road network, causes insufficient access roads between two OD matrix areas.

![Figure 7. Comparison of simulation results and real data.](image1)

Figure 7. Comparison of simulation results and real data.

![Figure 8. The proportion of simulated data in the real data range.](image2)

Figure 8. The proportion of simulated data in the real data range.

5. Conclusions and Outlook

This article presents a high-precision microscopic traffic simulation scenario of the city Duisburg. Road network from OSM and traffic demand from OD matrix are combined in four different ways for simulation. To find out the most precise scenario, data from detectors including induction loops and cameras are used for verification. The simulation results show that Case C with part of the traffic demand (contains trips either departure or arrival point is inside the city) and interior city road network are closest to the data collected in the real world. The reproduction rate of the simulation differs from the different simulated range in four cases and the simulation time. The percentage of simulated traffic volume in the real traffic volume range can be up to 72.22%. The traffic volume after 6 o’clock has increased significantly. Meanwhile, the accuracy of the reproduced traffic volume has decreased. The method of generating traffic flow from the OD matrix in SUMO may affect the accuracy under a large amount of traffic and a complex road network.

Even though we tried our best to achieve the goal of high precision for traffic simulation scenarios, there are still some factors affecting our precision. First, it is hard to balance
the scope of simulation. Second, traffic light data are not available for the entire city, and the lack of other traffic users such as pedestrians, railways, subways, and buses make the roads more convenient for private vehicles. Next, the selection of verification points also has effects on the precision of results. The data of detectors is from one particular week but weather conditions, time of the year, holiday or not. Instead, all have an impact on the traffic volume. Traffic is a special research object that cannot be reproduced 100 percent, many factors impact the final scenario, but with the high precision scenario of the city of Duisburg, lots of research can be carried out. The traffic simulation of the city of Duisburg can provide a reference for the simulations using an OD matrix and traffic plan guidance for the urban transportation departments.

Our next step will try to decrease traffic demand and then simulate 24 h traffic of the city of Duisburg. The traffic simulation based on the scenario of city Duisburg will also be the focus on the new technologies of vehicles. Electric vehicles and autonomous vehicles will be the major research objects. What kind of impact these vehicles with new technologies will have on the traffic flow is part of our further research work.

Author Contributions: X.M.: conceptualization, data curation, formal analysis, methodology, software, writing and editing. X.H.: conceptualization and revision. T.W. & D.S.: project supervision co-design of the content and revision. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: We like to thank Wirtschaftsbetriebe Duisburg and Stadt Duisburg for their support with the traffic data of city Duisburg. A part of this work is based on the researches within this cooperation. Map data copyrighted by OpenStreetMap contributors and available from https://www.openstreetmap.org.

Acknowledgments: We acknowledge support by the Open Access Publication Fund of the University of Duisburg-Essen.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lopez, P.A.; Wiessner, E.; Behrisch, M.; Bieker-Walz, L.; Erdmann, J.; Flotterod, Y.-P.; Hilbrich, R.; Lucken, L.; Rummel, J.; Wagner, P. Microscopic Traffic Simulation using SUMO. In Proceedings of the 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018; pp. 2575–2582.

2. Bieker, L.; Krajzewicz, D.; Morra, A.; Michelacci, C.; Cartolano, F. Traffic simulation for all: A real world traffic scenario from the city of Bologna. In Modeling Mobility with Open Data; Behrisch, M., Weber, M., Eds.; Lecture Notes in Mobility; Springer: Cham, Switzerland, 2015.

3. Krajzewicz, D.; Erdmann, J.; Behrisch, M. Recent development and applications of SUMO-Simulation of Urban MObility. Int. J. Adv. Syst. Meas. 2012, 5, 3–4.

4. Codeca, L.; Frank, R.; Engel, T. LuST: A 24-hour Scenario of Luxembourg City for SUMO Traffic simulations. In Proceedings of the SUMO User Conference 2015—Intermodal Simulation for Intermodal Transport, Berlin, Germany, 7–8 May 2015.

5. Asano, Y.; Ito, N.; Inaoka, H.; Imai, T.; Uchitane, T. Traffic simulation of Kobe-city. In Proceedings of the International Conference on Social Modeling and Simulation, Plus Econophysics Colloquium 2014; Springer: Cham, Switzerland, 2015; pp. 255–264.

6. Ma, X.; Hu, X.; Schramm, D. Effects of School Activities on Traffic Flow on Example of Traffic Scenario from the City Wuhan. In Neue Dimensionen der Mobilität; Springer Gabler: Wiesbaden, Germany, 2020; pp. 153–163.

7. Tubaishat, M.; Zhuang, P.; Qi, Q. Wireless sensor networks in intelligent transportation systems. Wirel. Commun. Mob. Comput. 2009, 9, 287–302. [CrossRef]

8. Lefebvre, N.; Chen, X.; Beausoy, P. Traffic flow estimation using acoustic signal. Eng. Appl. Artif. Intell. 2017, 64, 164–171. [CrossRef]

9. Ma, X.; Hu, X.; Weber, T.; Schramm, D. Traffic Simulation of Future Intelligent Vehicles in Duisburg City Inner Ring. Appl. Sci. 2021, 11, 29. [CrossRef]

10. OpenStreetMap Contributors. Planet Dump [Data File from 01.10.2017]. 2015. Available online: https://planet.openstreetmap.org (accessed on 19 August 2020).
11. JOSM. Available online: https://josm.openstreetmap.de/ (accessed on 8 July 2018).
12. Google Street View. Available online: https://www.google.com/streetview/ (accessed on 19 August 2020).
13. Schmidt, G.; Thomas, B. Hochrechnungsfaktoren für manuelle und automatische Kurzzeitzählungen im Innerorts bereich. In Forschung Straßenbau Straßenverkehrstechnik; Bundesministerium für Verkehr: Bonn, Germany, 1996.
14. Dijkstra, E.W. A note on two problems in connexion with graphs. Numer. Math. 1959, 1, 269–271. [CrossRef]
15. Seyed Sajad, M.; Michael, S.; Enda, H. Traffic light control using deep policy-gradient and value-function-based reinforcement learning. IET Intell. Transp. Syst. 2017, 11, 417–423.
16. Westerman, M.; Litjens, R.; Linnart, J.P.Z. Integration of Probe Vehicle and Induction Loop Data: Estimation of Travel Times and Automatic Incident Detection; PATH Research Report UCB-ITS-PRR-96-13, Institute of Transportation Studies; University of California: Berkeley, CA, USA, 1996.
17. Treiber, M.; Arne, K. Traffic flow dynamics: Data, models and simulation. In Physics Today; Springer: New York, NY, USA, 2014; pp. 7–23.