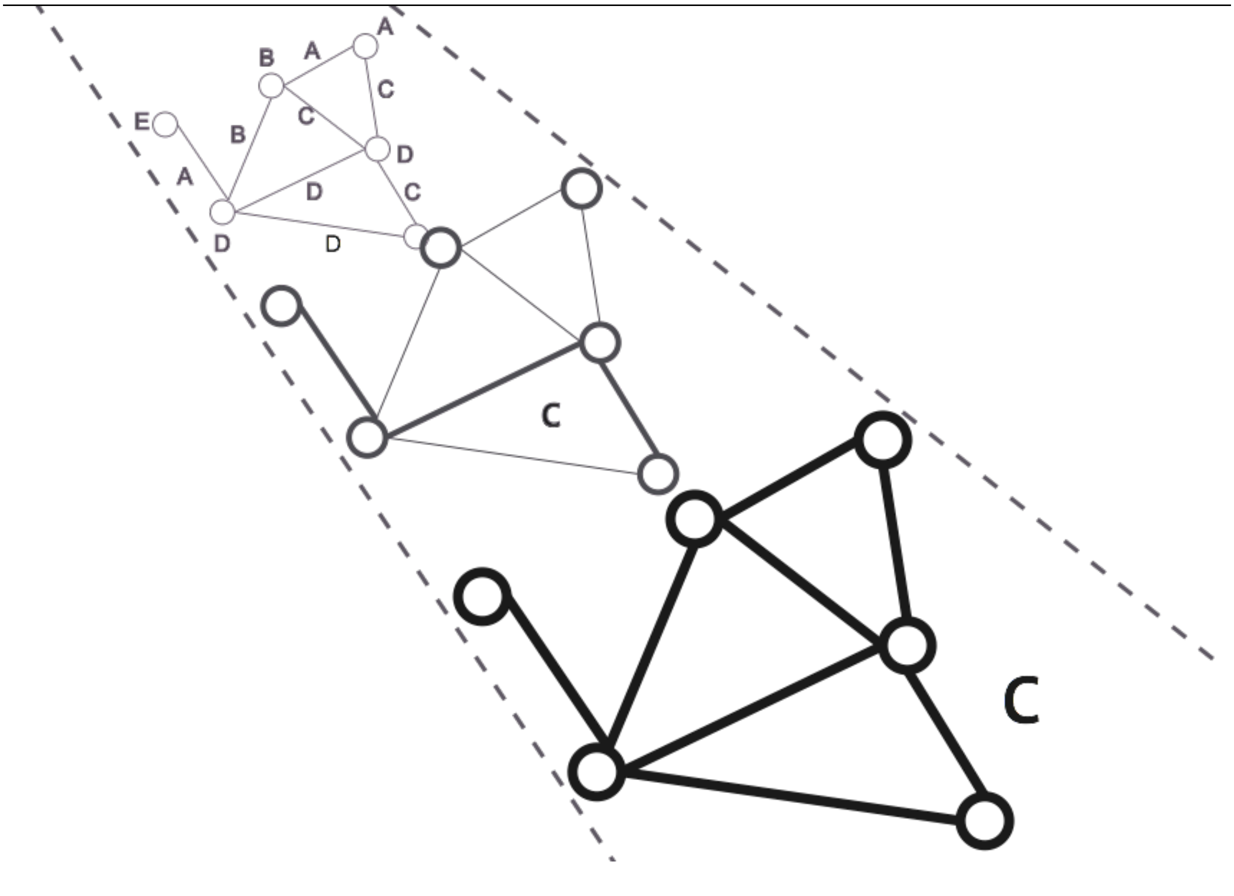


Public Transport Capacity and Quality – Development of an LOS-Based Evaluation Scheme

Hermann Orth, IVT, ETH Zürich
Robert Dorbritz, IVT, ETH Zürich
Ulrich Weidmann, IVT, ETH Zürich

Conference paper STRC 2011

STRC | 11th Swiss Transport Research Conference
Monte Verità, Ascona, May 11 - 13, 2011

Public Transport Capacity and Quality – Development of an LOS-Based Evaluation Scheme

Hermann Orth
ETH Zürich, Institute for Transport Planning and Systems (IVT), Zürich

Phone: +41 44 633 3109
Fax: +41 44 633 1057
email: orth@ivt.baug.ethz.ch

Robert Dorbritz
ETH Zürich, Institute for Transport Planning and Systems (IVT), Zürich

Phone: +41 44 633 6816
Fax: +41 44 633 1057
email: dorbritz@ivt.baug.ethz.ch

Ulrich Weidmann
ETH Zürich, Institute for Transport Planning and Systems (IVT), Zürich

Phone: +41 44 633 3350
Fax: +41 44 633 1057
email: weidmann@ivt.baug.ethz.ch

2011-05-09

Abstract

Public transport providers and government agencies have an interest in viable measurements to evaluate and compare public transport services. With the advent of liberalization in the transportation sector, this interest becomes a requirement as both service providers and service purchasers need agreed-upon measurements for the design of service provision contracts and realization thereof. The objective of this research is to develop standard values and measurements for quality and capacity in roadway-based public transport including tramway and bus services. The focus of the evaluation is primarily on capacity and operational quality, with quality in terms of comfort being considered as far as it impacts capacity. This paper describes the evaluation scheme in development for this purpose.

Conceptually, capacity is viewed as the result of a theoretical maximal threshold value that is reduced by a number of limiting influences arising from the operation in the oftentimes volatile environment that the roadways and passenger behavior are. Based on this, capacity and quality are measured using spatial and temporal indicators. The measurements can be taken to evaluate a transit system at the levels of a single element, e.g. a single bus stop, a set of elements, such as a route or route segment or the whole transit system.

Once calibrated, this evaluation scheme may serve as a tool to evaluate transit service operations.

Keywords

Public transport, public transport quality, capacity and capacity utilization, service measurements

1. Introduction

Performance measurements are required in every kind of operation or service. In public transit, stakeholders have an interest in measuring a variety of indicators, for example operational performance or comfort level provided. Operators and government agencies in particular need standards of operational performance in terms of service reliability, efficiency and capacity. Currently, the main benefit of such measures is the ability to compare a given service to others and thus benchmark and identify best practices, as well as areas with potential for improvement.

In the long term, market liberalization in the transportation sector will additionally require tendering practices using standardized measures for service provision quality and capacity. In the light of this development, service measures will drive business decisions to a much larger extent as service providers become comparable more easily and need to compete for concessions.

This paper is part of an ongoing effort to develop standards for transit service evaluation in Switzerland. A key question in such an endeavour is: How to measure the quality of a service? In this research, a model is developed that evaluates transit services on three levels: individual elements of a transit line, route segments or trips, or a whole transit network. The approach is to study a number of key performance indicators and combine the scores to form a single level of service (LOS) at the single element level. As the measurements are taken at the element level, these can be combined to evaluate all possible lines, line segments or even passengers trips on the second level of the model.

With the perspective of an operator or an agency planning transit services in mind, the focus is on capacity considerations and comfort is included to the extent that it influences capacity and reliability. Conceptually, capacity is the result of a theoretical maximal threshold value that is reduced by a number of limiting influences arising from the operation in the oftentimes volatile environment that the roadways and passenger behavior are. Based on this, capacity and quality are measured in a spatial, i.e. number of passengers per given area, and temporal, i.e. on time performance, dimension. Nonetheless, the model is designed to be modular so that when the need for further indicators, such as purely comfort related ones, arises, these can be added to the evaluation process.

The result is an evaluation model that is simple enough to be used in a standardized environment but at the same time flexible enough so that requirements for additional indicators in certain cases can be included. Operational data gathered by the VBZ, the main transit operator in Zurich, is then used for a case study on parts of the model.

Section two of this paper reviews works previously done in this area and outlines different approaches to transit service evaluation. Also, studies on passenger behavior and characteristics of service quality are briefly reviewed. In section 3, performance measurements are selected and the model development is described in detail. Following this, section 4 presents a partial case study undertaken using operational data. Section 5 concludes this paper, provides a discussion of the work and an outlook on further research.

2. Literature review

2.1 Service capacity and quality measurements

Transit service capacity and quality evaluation has been studied extensively in existing research. In [1], quality is determined using indicators similar to those applied for individual motorized transport, such as speeds and traffic flow. The LOS are determined for individual facilities and differentiate between different types of facilities, e.g. highways or signalized intersections. For transit, irregularities serving transit stops, crowding in vehicles and transit speeds are looked at. Passenger density within the vehicles is measured as a purely comfort-related indicator. The approach taken in [2] is similar in that a wide range of individual measures are evaluated, however quality is studied in much more detail far more passenger-centric. Emphasis is placed on access and comfort and, consequently, many measures such as walking distance to stops, service convenience, information, security and transit stop amenities are included in the evaluation and LOS scales are developed for many of these. These considerations are taken further in [3] where interactions between different components of quality, such as passenger density and transit capacity are briefly introduced. Capacity and comfort are treated as two separate dimensions of transit performance. Transit service capacity is considered in high detail, looking at many factors influencing capacity. Furthermore, the three levels transit stop, route segment and transit system are differentiated. Measures are taken at one of these levels and LOS scales are provided for these individually. A method to combine different measures to form an index is given conceptually. This approach is to build a weighted sum of the individual measures. In [4] the operation of bus lanes is studied, looking at speeds achieved on different types of bus facilities and bus stop capacities.

Similar to these studies is that while a wide range of indicators is looked at, measurements are largely for individual indicators or groups thereof and no systematic approach for developing an overall LOS is provided. Where different system levels are defined, LOS are developed for a number of indicators within the respective levels, but there is no integration between the different systems levels. While this provides the ability to consider many aspects of transit service quality in detail, an overall view is omitted and interactions between indicators are not accounted for in the evaluation.

An approach into the opposite direction is described in [5]. The authors present a model that uses one measure, the travel time quotient between transit and automobile use. This quotient can be calculated for single elements, such as a transit trip between two stops, and multiple elements can be easily combined to measure any trip or whole network. A striking feature of this method is the flexibility and scalability of the evaluation as any trip, segment or route up to the whole system can be evaluated in a consistent manner. Furthermore, by focusing on travel time difference, a measure is chosen that is assumed to be a very influential one on mode choice. A drawback of this method is that it focuses on only one measure, which may be too aggregate and hence omit other decisive indicators.

2.2 Service capacity and quality interactions

In order to be able to consider the interaction between quality and capacity, basic models for transit service capacity are studied. The work carried out in [6] outlines the conceptual process that leads to the realization of quality and capacity realized in service operation. Three phases of service realization are determined: The first phase is project and operational planning, where deterministic influences, such as infrastructure, vehicles and timetables are defined. The second level is the actual service operation and events affecting the service during its operation. Influences in this level can be passenger loads, traffic conditions and subsequent schedule deviations. Many of these influences can be accounted for during the first phase of service realization. The third level are accompanying factors, such comfort and customer service. While the perception of these aspects is highly dependent on the individual passenger, their influence on service capacity and operational quality is quite small. In [7], a model for determining transit capacity is developed that models transit capacity based on a theoretical maximum capacity (vehicle capacity x frequency) and then reduces this value to account for buffer times and acceptable passenger density to yield an operationally viable capacity. The influence of passenger loads on transit performance is studied in detail in [8], where boarding and alighting times of transit vehicles are modelled. The proposed method accounts for passenger density and position and size of doors and therefore may be useful in determining passenger loads critical for service reliability.

3. Quality and capacity modeling

3.1 Influences on quality and capacity

The capacity and quality model is intended to be capacity-focused as this is the main criterion for transit service planning. Quality of service is therefore viewed mainly as operational quality. Quality in terms of comfort is accounted for to the extent that it has an influence on capacity, however the model shall be open for extensions, for example to study quality from the passenger perspective as well.

The goal is to develop a model that covers all major influences in order to conduct an analysis of transit operations. At the same it is needed to be compact enough to be used efficiently. Furthermore, it shall be possible to determine a single level of service score for any network element or sets thereof. This shall provide a standardized evaluation that is consistent across different levels of the transit network under evaluation.

The capacity model developed in [7], derives actual capacity from a theoretical maximum value that is based solely on the maximum vehicle capacity and theoretically minimal possible headways and is then reduced to account for limitations arising from the actual operation. The first level of reduction is inclusion of buffer times to account for some instability from e.g. variation of drivers or different acceleration and braking characteristics of different vehicles, yielding a so called operational capacity. A further reduction is then applied to account for the fact that passengers will rarely accept crowding up to the point of the technically possible level but rather wait for following runs instead or not use transit at all. This is defined in [7] as the comfort-oriented capacity. Further reductions are then made to get the system capacity which accounts for conflicts with other transit lines sharing route segments or stops and finally the mixed traffic capacity which is the capacity than can be reached when operations take place without dedicated right of ways.

During these reduction steps, disruptive influences are taken into consideration. In [6], these influences are summarized and assigned to the service aspects, e.g. time or space, and the network level they affect. The network levels are defined as

- A single element, which can be a transit stop or a line segment between two stops
- Routes or route segments, which can be any combination of consecutive single elements, e.g. a transit line or a passenger transit trip

- The whole transit network or subsets thereof, e.g. tram-networks or express service networks

This results in the collection of influences as shown in Figure 1.

Figure 1: Influences on transit service quality and capacity

Influence	Indicator	Measurement unit	Network level		
			Single element	Route or route segment	network
Time					
Speed	Travel speed	[km/h]	x	x	x
	Acceleration and braking	[m/s²]		x	x
Frequency (service)	Temporal spacing between vehicles	[min]		x	(x)
Vehicle spacing (operationally)	Minimal buffer time between vehicles	[s]	x	x	x
Space					
Available or designated space	Space within vehicle	Seats [1] or standing room [m²]	x	x	x
	Share of dedicated right of ways	[%]		x	x
	Type of road	Qualitative	x	x	
	Type of transit stop	Qualitative	x		
Obstructions					
Passenger density	Density within vehicle	[%] seats or standing passengers [P/m²]	x	x	x
Reliability					
Reliability	On time performance	[%] runs on time		x	x
	Headway adherence	[%] headways as scheduled	x	x	x
Availability					
Availability	Service duration	[h/d]	x	x	x

Within this set, many indicators influence each other or have only minor effect on overall performance. Additionally, some of the indicators, while being important inputs, have an impact that cannot be quantified without very elaborate methods. However, their effect on the operation is measureable, meaning that for the purpose of evaluation, they can be substituted by more easily measurable indicators. Consequently, a set of indicators which is considered to have the strongest influence on operational quality and directly measures performance is selected. These indicators are classified into two groups; temporal and spatial indicators.

3.2 Temporal dimension: reliability and speed

As temporal indicators, the share of on time runs, the share of headways according to schedule as well as actual and scheduled travel speeds are used.

On time runs are considered because this measure can indicate deficiencies in service planning, such as too tight schedules, or hot spots of schedule disruption, e.g. intersections that are constantly above capacity. Furthermore, for passengers traveling by schedule, this is critical. A run is considered on time if it arrives at a stop within a predetermined threshold of the scheduled arrival time. A possible scoring of on time performance is shown in Figure 2.

Figure 2: LOS scheme for transit on time performance

LOS	% runs on time[1] [%]	Description of LOS
A	> 87.5	Almost all runs on time, stable operation
B	75.0-87.4	Few slighlty delayed runs, small impact on operation
C	62.5-74.9	Few moderately delayed runs or many slightly delayed runs, noticeable impact on operation
D	50.0-62.4	Many moderately delayed runs, operation is impacted noticeably
E	37.5-49.9	Many severely delayed runs, operation severely impacted
F	25-37.4	Almost all runs severly delayed, operation not according to schedule

[1] A run is considered not on time if it is ahead of schedule or if it es delayed by more than a predetermined threshold. A common threshold in use at European urban transit operators is 2 minutes

The quota of headways adhering to schedule can also be used to identify critical elements along routes. Furthermore, headway deviations can have critical effects on operational quality as large deviations will quickly lead to bus (or tram) bunching can, in consequence, destabilize operations along a whole transit line. In Figure 3, headway adherence level of service scores are shown.

Figure 3: LOS scheme for headway adherence

LOS	% of headways adhering to schedule[1]	Description of LOS
A	> 87.5	Headways are consistent with schedule headways in almost every
B	75.0-87.4	Few slightly headway deviations
C	62.5-74.9	Few moderately deviating headways or many slight deviations
D	50.0-62.4	Many moderate headway deviations
E	37.5-49.9	Many severe headway deviations
F	25-37.4	Almost all headways severly deviating from schedule headways

[1] A headway is considered not according to schedule headway if it deviates by more than a defined threshold. For now, this threshold is defined as +/- 1 Minute

These two measures also reflect impacts from obstructions, prioritization measures (e.g. signal actuation or dedicated right of ways) and sound planning that accounts for recurring random events.

Finally, speed is included as this is an important influence on line haul performance. It is also one of the main factors affecting the competitiveness of transit against automobile usage and thus a measure for transit service quality. Furthermore, speed directly affects operating costs as it has a direct influence on the number of vehicles required to operate a given service. Due to achievable speeds being highly dependent on road types, number and operation mode of signals and many other factors, absolute measures are not practical. One approach would be to develop reference speeds for a number of road or segment classes. This would, however, result in the need for very elaborate studies of these classes and reference speeds or, with a viable number of classes, likely be too aggregate to describe conditions in a usable way. Therefore, the evaluation of speeds is proposed according to [5] and transit speed is referenced with speed achieved on the same segment by automobiles, as proposed conceptually in Figure 4. This method can directly give insights on the competitiveness of transit services as the transit and automobile speed, and thus travel time, are compared.

Figure 4: LOS Scheme for transit speed

LOS	Transit speed as % of automobile speed	Description of LOS
A	> 87.5	Transit trip can be considered at as fast as automobile
B	75.0-87.4	transit trip barely noticeably longer
C	62.5-74.9	Transit trip slightly longer
D	50.0-62.4	Transit trip longer
E	37.5-49.9	Transit trip at least twice as long
F	25-37.4	Transit trip takes many times as long as automobile trip, uncompetitive

3.3 Spatial dimension: passenger density

As measurement for spatial influences, solely passenger density within vehicles is studied. Passenger density can be used as a measure of quality in terms of comfort, but in this context is selected primarily for its impact on transit operations. As shown in [8], density of standing passengers has a direct effect on boarding and alighting times. Furthermore, it is assumed that while some level of crowding is tolerable, there is a critical level above which stopping times are severely extended leading to major schedule deviations. With respect to level of service considerations, this means that high LOS are assigned for densities at which there is no or only little effect on operational stability. Using the methods developed in [8], critical densities and intermediate steps can be calculated. This results in the initial scoring thresholds for passenger density as shown in Figure 5.

Figure 5: LOS scheme for passenger density within vehicles

LOS	Density [P/m^2]	Description of LOS
A	< 3.0	Density without any effect on stopping duration
B	3.1 – 3.5	Slighly extended stopping durations, any delays can be compensated between stops
C	3.6 – 4.0	Stopping durations extended and slight effect on line performance
D	4.1-5.0	Stopping durations extended considerably, line performance is affected
E	5.1 – 6.0	Stopping durations highly extended, line performance degraded
F	> 6.0	Stopping durations far higher than planned, unreliable operation

3.4 Overall model

The four measurements chosen cover the influential factors on transit performance shown initially because the influence of a number of other indicators is reflected in each measure. Figure 6 shows which indicators have influence on which measurement.

These measurements are taken at the single element level and, on this level, depend on the type of element under investigation. For stops, the applicable measurements are the share of on time runs and the share of schedule-adhering headways at the stops. For line segments, the measures are the average speed between the stops at each end and the standing passenger density within the vehicle between the two stops.

Figure 6: Measure and further influences reflected therein

Measurement	**Further influences reflected in measure**
On time performance	Buffer times, designated right of ways, transit stop type
Headway adherence	Buffer times, designated right of ways, transit stop type
Travel speed relative to automobile speed	Travel speed, Acceleration and braking, designated right of ways, road type
Standing passenger density	Frequency

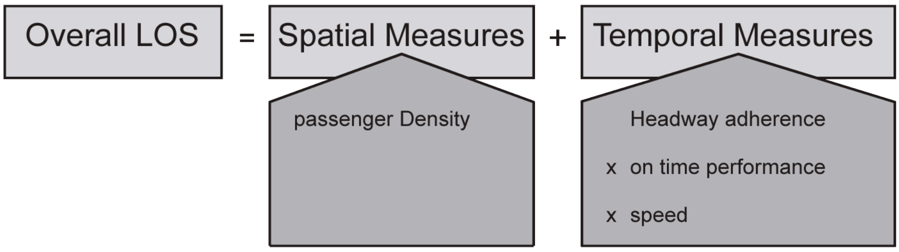
As the evaluation model is meant to yield a single LOS score, a method is required to link the individual measurements together. For this, it needs to be considered what different methods and mathematical operations mean with respect to interactions between indicators.

Summing scores can be understood as the scores summed being independent of each other. The degradation of one measure does has no direct impact on the other measure and it may even be possible for a good and bad score to compensate one another.

On the other hand, multiplication indicates that there is a relationship between the two measures. With one measure being degraded, another multiplicatively linked measure´s maximum score is reduced, meaning that one measures being degraded prevents optimality of the other measure as well. Also, the range of scores is reduced, reducing the impact of any measure when others are degraded.

These considerations mean that within a dimension, scores are multiplied and between dimensions, added, resulting in the scheme as shown in Figure 7.

Figure 7: Overall LOS evaluation scheme



Formally, this is expressed in Eq. 1 where, for the calculation, the LOS grades A to F are referenced with scores of 6 to 1 where 6 would equal A.

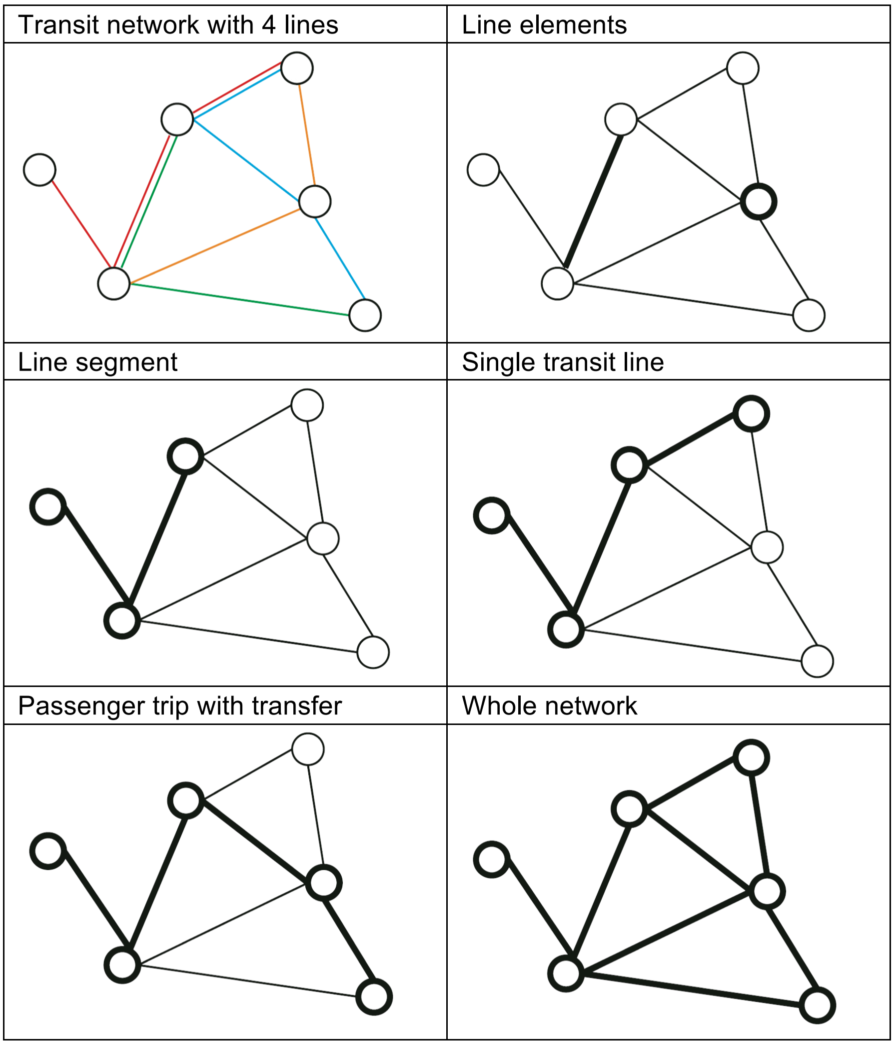
$$LOS = \frac{LOS_{space} + LOS_{time}}{2} \quad \text{(Eq. 1)}$$

Since the focus is on operational performance, measurement of the two dimensions with the greatest impact on performance is sufficient. However, if at a later point further aspects of service, for example ride comfort or amenities are to be evaluated, they can be easily integrated into the model as further "components". The more general calculation model in that case is shown in Eq. 2, with i being the different service quality dimensions.

$$LOS = \frac{1}{n} \sum_{i}^{n} LOS_i \quad \text{(Eq. 2)}$$

The level of service is initially calculated for a single element or for a line segment leading up to a stop. Elements can now be combined variably to compute multi-element LOS scores. These could be parts of a transit line, a whole line or a passenger's trip connecting through many lines. For a more aggregate analysis, parts of a network or even the whole network can be evaluated, by considering numerous or even all elements within the network or network part. This is displayed conceptually in Figure 8.

Figure 8: Network and examples of possible evaluation levels and extents



The procedure to conduct such an analysis depends on the type of the network element: For a single network element, the according measures (speed and passenger density or on time performance and headway adherence) are taken and combined. For analysis of a set of elements, such as routes or trips, the overall element level LOS is determined using all four measures, taken at a line segment and the stop at its end. This method takes into account that when evaluating segments or trips, the performance at a stop in terms of reliability is determined to a large extent by the line element leading up to the stop. With the line and stop element LOS determined, these elements can be combined to form any subset of the whole transit network and the LOS score is then calculated as an average of the scores of the elements contained within the set.

This way, this evaluation model is highly flexible and scalable.

4. Case study

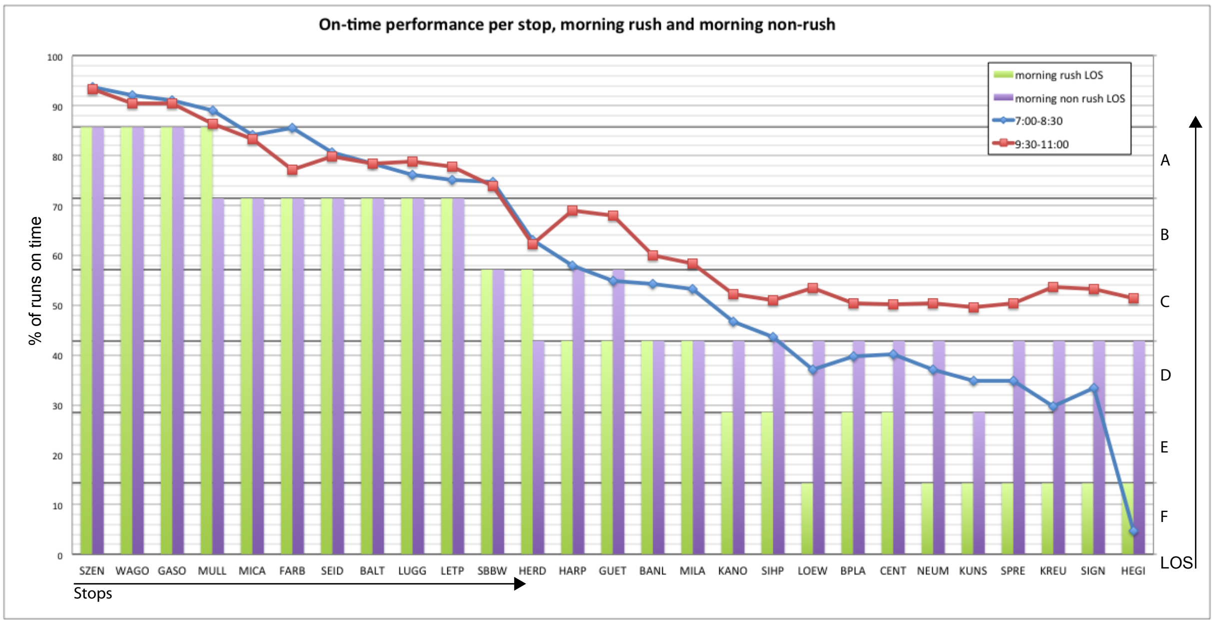
For preliminary analysis, one transit line is studied using data supplied by, the main transit operator in the city of Zurich (VBZ). At this point, an initial and partial analysis is undertaken in order to do a quick study of the levels of service achieved along a transit line. This initial analysis is based to a large extent on a detailed analysis of transit operational performance conducted in [9]. The measures studied are on time performance and speeds.

This analysis studies the performance of line 31 in the VBZ network, a high capacity bus line between two outlying parts of the city passing through the city center. This line is operated with double articulated buses with a capacity of 200 passengers, running at 6 minute headways during rush hour periods.

4.1 On time performance

For this analysis, weekday morning runs are investigated for the direction of the line that tends to carry more passengers. The morning rush period (07.00-08.30) and the following non-rush period (09.30-11.00) are looked at (Figure 9). First of all, and not surprisingly, on time performance degrades over the course of the runs as delays accumulate and cannot be compensated for. For about half of the line, rush and non-rush performance is similar at LOS A-C, indicating that the operational planning accounts for most disruptions during rush hour time quite well. However at the second half of the line, non-rush LOS does not drop below D while during rush hour, the level of service degrades to F.

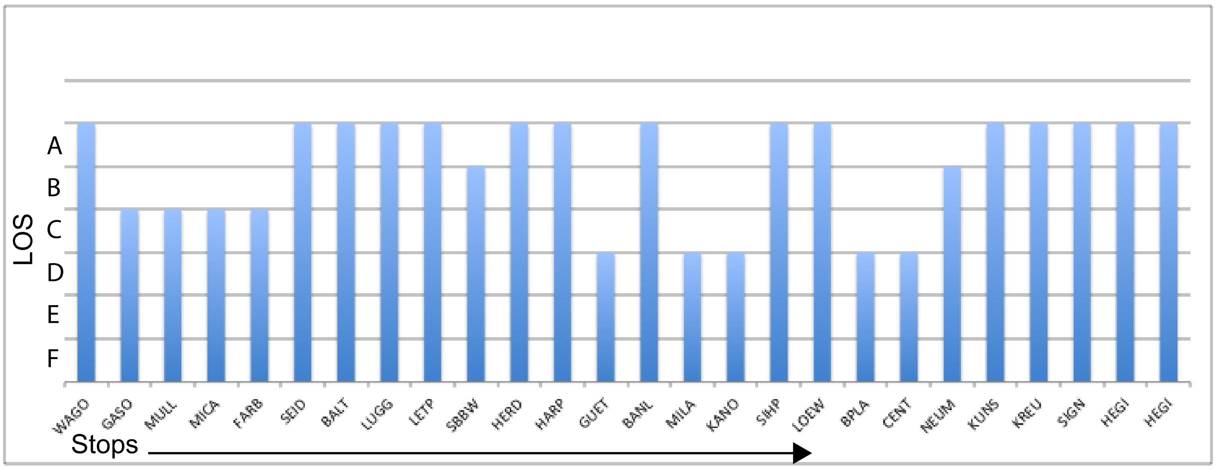
Figure 9: On time performance and levels of service



4.2 Speeds

Speed between stops has been studied and referenced with automobile speeds that are reached in between two transit stops. Figure 10 shows the results of this analysis. It can be seen that overall, the operational performance in terms of travel time is quite high. On about half of all segments, LOS A is achieved and the worst instances are of LOS D. Regarding the operation of the transit line, this indicates that the prioritization of transit in use is effective and ensures that the transit line remains competitive.

Figure 10: Transit speeds and levels of service



5. Further Work

While the model structure is finalized, the LOS thresholds need further studying. At this point, values are chosen largely by experience and have not been thoroughly calibrated with operational data on a large scale yet. This will be done using more elaborate data supplied by transit operators. In the long term, comparability of transit LOS with automobile LOS shall be developed. This applies also to work conducted on bycicle and pedestrian LOS standards currently in development.

6. Discussion

While the model is consistent and measures have been carefully selected, a number of aspects are in need of further evaluation. The approach of measuring speeds on the level of line elements delivers consistency across the different levels of the transit network. However, at more complex and longer routes, especially when considering a passenger trip that includes one or more transfers, it is possible that the same origin-destination trip by automobile would be along a different route and therefore the travel time difference would change. The effect of this needs to be studied further.

In comparison with other evaluation approaches for transit level of service, this approach is somewhat in the middle between the two extremes found in literature. One side is the evaluation of a large number of measures on different levels that cannot be integrated to form a standardized measure, while the other far side is to consider only one measure, however one that can be consistently studied for any level and subset of a network. In this regard, the evaluation method proposed in this paper sacrifices detail on one hand and simplicity on the other hand, however delivers a level of service scoring system that takes into account the major influences on operational capacity and quality and is at the same time highly flexible and scalable.

7. Summary

An approach to evaluating transit service operational performance and building and level of service score for this has been proposed. While the actual scoring thresholds may need further verification, it has been shown that the model is, in general, consistent across all levels of a transit network. It is also flexible and scalable while considering the most important influences on transport capacity and transit quality.

8. References

[1] Forschungsgesellschaft für Strassen- und Verkehrswesen (FGSV) (2005) Handbuch für die Bemessung von Strassenverkehrsanlagen (HBS), 2001 edition, amended 2005

[2] Transportation Research Board (2000) Highway Capacity Manual 2000, Washington, D.C.

[3] Transportation Research Board (2003) Transit Capacity and Quality of Service Manual, Washington, D.C.

[4] St, Jacques, K., H. Levinson (1997) TCRP Report 26: Operational Analysis of Bus Lanes on Arterials, Transportation Research Board, National Academic Press, Washington D.C.

[5] Fu, L., Y. Xin (2007), A New Performance Index for Evaluating Transit Quality of Service, Journal of Public Transportation, 10, No.3, pp. 47-69.

[6] Scherer, S., P. Spacek, U. Weidmann (2010), Multimodale Verkehrsqualitätsstufen für den Strassenverkehr – Vorstudie, project report SVI 2007/005, Institute for Transport Planning and Systems (IVT), ETH Zürich, Zürich

[7] Anderhub, G., R. Dorbritz, U. Weidmann (2008) Leistungsfähigkeitsbestimmung öffentlicher Verkehrssysteme, IVT Schriftenreihe 139, Institute for Transport Planning and Systems (IVT), ETH Zürich, Zürich

[8] Weidmann, U. (1995) Grundlagen zur Berechnung der Fahrgastwechselzeit, Schriftenreihe des IVT 106, Institut für Verkehrsplanung und Transporttechnik (IVT), ETH Zürich, Zürich

[9] Carrasco, N. (2011) Quantifying Public Transport reliability in Zurich, forthcoming, 11th Swiss Transport Research Conference, Ascona, 2011