# Improving Punctuality by Adjusting Timetable Design

**Mayank P. Patel[1] V. R. Gor[2]**
[2]Goverment Engineering College, Modasa

*Abstract*--- Transport authorities try to achieve a high level of quality in public transportation networks by creating performance indicators. The performance indicators can be of various designs, e.g. safety, frequency, arrival punctuality and departure punctuality. A common used performance indicator is departure punctuality. This article shows the impact of various characteristics of timetables on departure punctuality. The elements chosen for research are running time, buffer time and surplus time. This research clearly shows the impact of surplus time on punctuality. Surplus time acts as extra buffer time in timetables and therefore increasing punctuality levels. Furthermore it is shown that punctuality can be described as the sum of these 3 elements. The amount of time rewarded to each of these elements will determine eventually the level of punctuality. Changing these elements has therefore a direct influence on punctuality.

**Keywords:** public transportation; punctuality; simulation; timetable design; performance indicators in public transportation

## I. INTRODUCTION

Mobility in urban environments is becoming an increasing problem. [1] One of the possible solutions to improve mobility in cities is public transportation. To achieve a high service level and lower costs in the public transportation sector, the Dutch government introduced market force in public transportation. This was done in

2001 when the "Wet Personenvervoer 2000" came into force. According to this new law, all public transportation companies had to be privatized. All transportation areas in the Netherlands were divided into so-called concession areas which are governed by transport authorities. These transport authorities issue the transportation concession in its area, to which public transportation companies can enrol.

To safeguard the quality of public transportation in each concession area transport authorities formulate performance indicators. These performance indicators are laid down in the concession agreement between transport authority and public transportation company and vary per concession area. In most cases the

Performance indicators are linked with a reward and penalty system; the public

Transportation Company receives a reward when performing better than agreed and has to pay a fine when performing below agreements.

First, this article describes various performance indicators and their implications to quality in public transportation networks. Secondly, a simulation model is designed to obtain an insight in the influence of different elements used as input for timetables on the level of punctuality of these timetables.

## II. PERFORMANCE INDICATORS

Performance indicators are created by public authorities to measure the quality of public transportation in their concession area. The first question that comes to mind is what the definition of quality is? Is quality defined as "being punctual", or maybe as "being safe". The public authority should attend to the interests of all users of public transportation in the concession area. When travellers are asked what's more important, they would probably rate safety higher than punctuality. [2]

### A. *Safety*

By liberalising the public transportation market public transportation companies had to change its organisational structure in order to survive. Before the liberalisation public transportation companies could mainly focus on the traveller. Earning profit never was the aim of the company, since it was in the hands of the government. After liberalisation this changed. The company suddenly had to make a profit in order to continue to exist. This meant that the performance indicator of public transportation companies was changed into making profit.

As stated by [2] the problem is that the privatized railway, in concentrating on punctuality and reliability whilst trying to achieve the private sector's overriding objective of profit maximization, has embarked on a cost-cutting exercise which eroded expert and tacit knowledge to the extent where safety is seriously compromised. Profit is the main performance indicator for the railway companies yet it appears in none of the government statistics.

The question whether or not liberalization is the best alternative to operate urban public transportation is not answered here. But the government should keep in mind that by changing the rules of the environment companies exist in, it changes the rules within companies as well.

### B. *Frequency*

Frequency is defined as the number of times a modality of a particular line dwells at a stop. Research done by [3] shows that when the frequency is 6 modalities per hour or more travellers do not look at timetables anymore. The estimated waiting time is short enough for travellers to wait at a stop.

This suggests that punctuality no longer is an issue when modalities arrive every 10 minutes. Many urban public transportation lines have a frequency of at least 6 modalities per hour, especially in the centre of cities. When punctuality is not an issue for travellers, why is it that important to transport authorities?

A new idea about the performance indicator therefore is the following:

When the frequency of a line is at least 6 modalities per hour, punctuality as performance indicator is replaced by frequency and inters arrival time. The inter arrival time should be within limits to be determined by the frequency of the line and the transport authority.

Using frequency as performance indicator has another advantage: it's easy to measure. By counting the number of modalities on a course and knowing the time of a course, the frequency can be calculated.

### C. *Arrival punctuality*

Punctuality as defined in this article is departure punctuality. Using departure punctuality as a performance indicator can disrupt running times as defined in the timetable. In many cases drivers speed up

At the end of the trip. This gives drivers more time off at the turning point. This is not in the interest of the traveller, since the modalities arrive and depart too early at the stops at the end of the line.

To solve this problem arrival punctuality can be used as performance indicator. Drivers need to follow to timetable better than when departure punctuality is used as performance indicator. Deviations will become smaller what the reliability of the timetable increases.

### D. *Departure punctuality*

Although other performance indicators might be more useful to improve the quality of public transportation, departure punctuality is chosen in many cases by transport authorities. Punctuality is defined in this article as the percentage of trips that depart on-time at a specific location of a transportation line. The modality that uses this line in this article is a tram, but other methods of transportation can be modelled as well. On-time is defined as departing maximum two minutes after the published time of departure in the timetable. If a trip departs a second too early it isn't on-time, but too early. These definitions are used designing the simulation model and discussing its output.

## III. Effect Of Timetable Design On Punctuality

A review of current literature reveals that no simulation model exists that focuses on the punctuality of urban public transportation networks. To satisfy to a possible need, this article is presented.

Carey and Carville [4] did develop a simulation model to predict the probability distributions of knock-on or secondary delays at stations. These delays occur when no buffer (or too little) is designed in the timetable, so that the delay of one modality affects the departure time of another. However, [4] focuses on train, instead of urban public transportation.

In [5] disturbances on a public transportation line are modelled. The modelling as done by Balbo and Pinson makes it possible to capitalize the knowledge available within a monitoring station and to follow up the evolution of the disturbances in real time. They do not use these disturbances to calculate its influences on punctuality.

When urban public transportation is researched, as is done by [6], punctuality isn't used as performance indicator. Meignan et al. designed a bus-network simulation tool which allows analyzing and evaluating a bus-network at diverse space and time scales, but no relation to punctuality.

This research though uses disturbances and dispersion in running times and dwell times to show the effect of running time design and buffer time design on punctuality. The aim of this article is to show the complex relations between these input characteristics of timetables and punctuality as output of that timetable.

Using Arena simulation software by Rockwell Automation a model is created of a public transportation line. Different timetables then are created based upon various input characteristics. These timetables are used to run the simulation model and the outcomes are compared to show the influences on punctuality.

## IV. Creating The Model

The simulation model is built using Rockwell Automation Arena 12 and is run on a Pentium 4, CPU 3.00 GHz computer. The model simulates a public transportation line using running times to simulate the driving of the vehicle and dwell times to simulate stopping to board and deboard passengers.

The running times and dwell times are based upon real running times and dwelling times of a tram line of a public transport company in the Netherlands. The public transport company delivered empirical data regarding running times, dwell times and stop percentages. This data is transformed into probability distributions. These probability distributions provide dispersal in time to simulate disturbances during operations. These disturbances include traffic lights, other traffic on the shared lanes, weather influences etc.

To change the empirical data into probability distributions that can be used as input for the simulation model a number of steps have to be carried out. These steps are shortly described.

### A. *Step 1: determining dataset*

The dataset for the running time between two stops and the dwell time at a stop can be determined by selecting the data within a particular time period, between two stops. There are 7 different time periods, each with its own data. The dataset is based upon data delivered by a public transport company in the Netherlands and interpreted using Microsoft Excel. Steps 2 and 3 are used to check the dataset and to detect flaws. These flaws then are corrected.

### B. *Step 2: using Arena's Input Analyzer to plot a probability distribution*

After determining the dataset, the data is analyzed using Arena's Input Analyzer.

This is a tool that plots probability distributions to selected datasets. Arena's Input Analyzer performs a Chi-Square test and shows the p-value of a selected probability distribution along with its parameters.

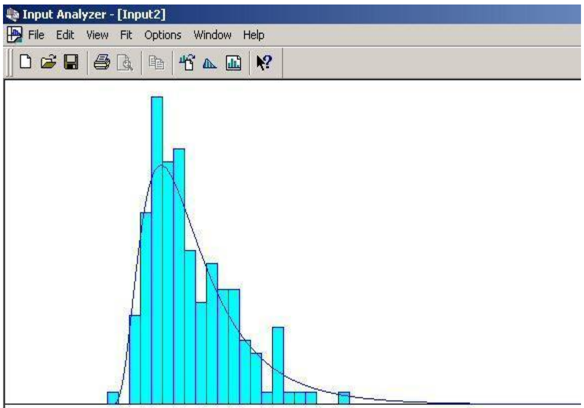


Fig. 1: Plotting probability distributions using Arena's Input Analyzer

Figure 1 shows the probability distribution plotted over the histogram of the dataset. This can be used for a quick determination whether a distribution fits well or not. For a more thorough investigation more analysis is necessary.

The Fit Menu in Arena's Input Analyzer returns the minimum square-error distribution. This square error is the sum of squared discrepancies between histogram frequencies and fitted-distribution frequencies. The best fit is shown in the Input Analyzer, but this could still be a poor fit depending on the p-value as described.

The Chi-Square test is used as a performance indicator for the goodness-of-fit of a probability distribution. First, a probability distribution has to be chosen. The probability distribution with the smallest square error is chosen as the most likely distribution. Second, the corresponding p-value of this distribution is compared to the significance level. As statistical significance level a percentage of 95 is chosen. A Type I error occurs when the researcher rejects a null hypothesis when it is true. The probability of committing a Type I error is called the significance level, and is often denoted by α. This α (level of significance) has a value of 0.05, due to the 95% level of significance. To compare the dataset with

The chosen probability distribution, two hypotheses are defined:

$H_0$: the dataset is consistent with the chosen probability distribution.

$H_1$: the dataset is *not* consistent with the chosen probability distribution.

The p-value is compared to the level of significance, which is 0.05. If the p-value is smaller than 0.05, $H_0$ is rejected, if not $H_1$ is rejected.

This can be shown using the following example:

| Distribution: | Lognormal |
|---|---|
| Expression: | 26.5 + LOGN(8.09, 5.13) |
| Square Error: | 0.004586 |
| *Chi Square Test* | |
| Number of intervals | = 10 |
| Degrees of freedom | = 7 |
| Test Statistic | = 7.41 |
| Corresponding | = 0.401 |
| p-value | |

Table. 1: Performing a Chi-Square test on a chosen probability distribution

In the example shown in Table 1, the corresponding p-value is 0.401, which is bigger than 0.05. This means that $H_1$ is rejected and that the chosen probability distribution is consistent with the dataset. Instead of the dataset the probability distribution is used in the simulation model.

### C. *Step 3: the corresponding p-value is smaller than 0.05*

If the corresponding p-value of the chosen probability distribution is smaller than the level of significance, extreme outliers are identified if present in the dataset. This is done using SPSS Statistics software. To identify en define extreme outliers, the following definitions are used:

1) *75th percentile:*

The value representing the 75th percentile indicates that 75% of the values in a dataset are equal to or below it. The 75th percentile is also called the upper quartile or

$q_n$ (0.75).

2) *Inter Quarter Range or IQR:*

The Inter Quartile Range (IQR) is the distance between the upper quartile and the lower quartile of a dataset.

In formula: IQR = qn (0.75) -

$q_n$ (0.25).

3) *Extreme outlier:*

An outlier is a value that is bigger or smaller than the 75th percentile plus

| | | | |
|---|---|---|---|
| or | minus | 3 times | the IQR. |
| In | formula: | extreme | outlier = |
| $q_n$(0.75) +/- 3*IQR | | | |

Extreme outliers can often be false measurements in a dataset. This can happen when a failure occurs during the process of measuring either technical or human. Other reasons for these values are accidents, infrastructural failures, technical failures of the modality or other disturbances. Extreme outliers can corrupt a dataset and can therefore be removed out of a dataset.

After identifying and eliminating extreme outliers, step 2 is repeated. If possible a probability distribution is plotted. If the corresponding p-value of the chosen probability distribution still smaller is than 0.05 step 4 is carried out. The outliers and the IQR can be represented in a boxplot. The boxplot is an oversight of how the dataset is distributed.

### D. *Step 4: plotting a triangular distribution*

If it still is impossible to plot a dataset, two options remain: using the dataset itself as input or use a probability distribution that has a poor fit. This research uses the triangular distribution to describe processes that do not fit well with another probability distribution. Looking at the

data, a modus can be identified. Furthermore a minimum and a maximum are available. The advantage of using a triangular distribution over the use of the dataset itself is that no data is lost. The dataset has gaps between the values in the set, which are filled using a probability distribution.

## V. SIMULATION PREPARATIONS

At the start of a simulation run the model is empty. Travellers are no entities in the model, since they are modelled using dwell time. This means that no queue is formed at a modality stop. Instead of a queue, the entity will be delayed for a couple of seconds based upon the empirical dwelling time. This simulates the boarding and deboarding of passengers.

At the end of the simulation run the model is empty as well. All entities are disposed out of the model. When the last entity is disposed, the model run will stop. This means that the system is finite.

### A. *Warm-up period*

In the real system there is a startup time, due to the fact that all modalities have to drive from the depot to the A-side. In the model start up time is not modelled. All modalities are available at the A- or B-side at the start of their first trip. They may be delayed however at the start of their second (or more) trip.

### B. *Run length*

According to Banks [7] two types of simulations can de distinguished: Finite-Horizon Simulations and Steady-State Simulations. A Finite-Horizon Simulation is a simulation that starts in a specific state and that is run until a terminating event occurs. The output isn't expected to achieve any steady-state behaviour and the parameters estimated from the output create its value depending on the initial conditions. The purpose of Steady-State Simulations is the study of the long-run

Behaviour of systems. The system researched ends when all trams have driven according to their timetable. This means the system is a Finite-Horizon Simulation. The run length is about a day. Starting at 5.51 hour with the first trip and ending between 0.00 and 1.00 hours, depending on possible delays and the timetable.

### C. *Replications*

A number of independent replications are run with each of them a new observation. Every run starts at 5.51 hours and ends between 0.00 and 1.00 hours. This run is replicated a number of times using different seeds to change the probability distributions.

### D. *Number of replications*

Every replication using different starting points for the probability distributions has a different output. The confidence interval around the average output of the simulation model can be described with the following formula [8]:

$$\mu = [\bar{x} - h, \bar{x} + h]$$

The output of the simulation model is the punctuality of tram line 2. In this formula

$\bar{X}$ can be described as the estimate for the expectancy of the punctuality and *h* as the half-value of the confidence interval. A small value of *h* indicates a high level of confidence of the estimation. If $x_j$ is the result of the $j^{th}$ replication, than the values

Of $\bar{x}$ and $s^2$ for x and $x$ are defined as:

$$\bar{X} = \sum_j \frac{x_j}{n}$$

$$s^2(x) = \sum_i \frac{(x_i - \bar{x})^2}{n-1}$$

Using the Central Limit Theorem which states that the sum of a sufficiently large number of independent random variables (the values of x) will be approximately normally distributed, the confidence interval can be determined using the half-value *h*. This confidence interval is centred

around $\bar{x}$ . To do this, the values of *h* and $s^2(\bar{x})$ need to be calculated.

$$h = t_{n-1,1-\alpha/2} \cdot s(\bar{x})$$

$$s^2(\bar{x}) = \frac{s^2(x)}{n}$$

The value $t_{n-1,1-\alpha/2}$ is found using the

Student-t distribution with (n-1) degrees of freedom. This research uses 5 replications. The average of these 5 replications is shown in the conclusions of this article.

## VI. Designing Test Theories

Using Theory Concerning Running Times And Buffer Times As Developed By [2] (Van Oort/Nes) 4 Test Alternatives Are Created. These Test Alternatives Consist Of 3 Characteristics: Running Time, Buffer Time And Stacking Space Problem. These Characteristics Are Defined As:

### A. *Running Time:*

The Running Time Protocol Determines The Amount Of Running Time Between Stops. When Is Set To 75%, 75 Out Of 100 Modalities Have To Be On-Time. Adjusting This Protocol Can Improve Punctuality, But Travel Times Will Grow As Well.

### B. *Buffer Time:*

Buffer Time Is The Extra Time Added To The Running Time. It Serves As A Recovery Time In Order To Minimize Delays. This Recovery Time Is Added In The Timetable At The Beginning And The End Of A Line.

### C. *Stacking Space Problem:*

The stop at the end of the transportation line has a stacking space problem: the capacity isn't sufficient to allow all lines to dwell using a regular timetable. To overcome this problem, modalities have a maximum dwell time of 2 minutes at the

last stop, while using regular regulations of the public transport company this should be more, e.g. 3 or 4 minutes.

Knowing all characteristics, the test alternatives can be defined:

*Alternative 1*

| Characteristics | |
|---|---|
| Running time | 75% |
| Buffer time | 10% |
| Stacking space problem | Maximum standing still time at the last stop is 2 minutes |

As alternative 1 a 75% running time with a 10% buffer time timetable is designed. This alternative has a restriction on the turnaround side of a course to simulate a capacity problem. Buffer times are set to a maximum of 2 minutes at the turnaround side and the "leftover" buffer time is rewarded at the other side of the course.

*Alternative 2*

| Characteristics | |
|---|---|
| Running time | 75% |
| Buffer time | 10% |
| Stacking space problem | No maximum standing still time at the last stop |

This alternative looks a lot like alternative 1, but shows the impact of maximising the standing still time at the turnaround side of a course to 2 minutes, as is done in alternative 1. This alternative has no restrictions on buffer time and all buffer time is rewarded at both sides of the course

*Alternative 3*

| Characteristics | |
|---|---|
| Running time | 85% |
| Buffer time | 0% |
| Stacking space problem | Maximum standing still time at the last stop is 2 minutes |

As described by [3], buffer time, running time and the probability that a modality can depart on-time are related. They argue that when a lower percentage of running time is rewarded, more buffer time is needed to obtain the same probability of departing on-time. To show the impact of rewarding more running time without buffer time alternative 3 is created. This should lead to a level of punctuality of 85%, according to [3].

*Alternative 4*

| Characteristics | |
|---|---|
| Running time | 75% |
| Buffer time | 0% |
| Stacking space problem | Maximum standing still time at the last stop is 2 minutes |

Buffer times are included in the timetable to improve punctuality and to prevent propagation delays. Due to surplus times in timetables, some extra buffer time is incorporated in the timetable. To research the impact of buffer time in a timetable next to surplus time, the buffer time is excluded from the timetable. When buffer time is set to 0, the impact of buffer time and surplus time on punctuality can be determined.

## VII. RESULTS

Table 2 shows the punctualities of all alternatives after running 5 replications.

| Alternative | Punctuality by the model | Calculated simulation |
|---|---|---|
| Alternative 1 | 87.07% | |
| Alternative 2 | 89.92% | |
| Alternative 3 | 96.17% | |
| Alternative 4 | 79.77% | |

Table. 2: Comparing alternatives regarding punctuality

The outcomes of the simulation model show higher levels of punctuality than theory predicted. This can be explained by surplus time in timetables. The *surplus time* is the time between the next interval and the ready-state of a modality. When the interval time is 10 minutes, every 10 minutes a modality has to leave the starting point at the A-side. If a modality is ready for its next trip at minute 2, it has to wait 8 minutes to start a new trip. These 8 minutes are the surplus time. Surplus time therefore can be influenced by changing the frequency of modalities.

The expected value of alternative 1 is at least 85%, since 75% running time and 10% buffer time is rewarded in the timetable. The remaining percentage (2.07%) can be accredited to the surplus time, which acts as "bonus" buffer time.

The effect of having no maximum standing still time at the last stop in alternative 2 is a slight increase in punctuality compared to a maximum standing still time of 2 minutes. Van Oort and van Nes [3] predict a punctuality of 85% when using a running time of 85% and no buffer time (alternative 3). However, they didn't take into account the surplus time. This alternative clearly shows that a running time of 85% together with the surplus time in the timetable leads to a very high punctuality. Alternative 4 shows that the use of a 75% running time protocol without buffer time a slight increase in punctuality is possible due to surplus times.

## VIII. CONCLUSIONS

This research clearly shows the impact of surplus time on punctuality. Surplus time acts as extra buffer time in timetables and therefore increasing punctuality levels. The outcomes of the simulation model also show that buffer times aren't necessary when increasing running time when surplus time is available in the timetable.

Depending on the level of punctuality required an alternative can be chosen. The chosen alternative and its characteristics can be used as basis for a new timetable.

As is shown in this research punctuality can be described using 3 elements of a timetable: running time, buffer time and surplus time. The amount of time rewarded to each of these elements will determine eventually the level of punctuality. Changing these elements has therefore a direct influence on punctuality.

## REFERENCES

[1] W.W. Veeneman, "Mind the Gap", Proefschrift, DUP Science, Delft, 2002.
[2] B. Cole and C. Cooper, "Making the trains run on time: the tyranny of performance indicators", Production Planning & Control, Volume 16, Issue 2, pages 199-207, 2005.
[3] N. van Oort and R. van Nes, "Betrouwbaar OV begint met goed plannen", Colloquium Vervoersplanologisch Speurwerk, Antwerpen, 2007
[4] M. Carey and S. Carville, "Testing schedule performance and reliability for train stations", Journal of the Operational Research Society, Volume 51, Issue 6, pages 666-682, 2000.
[5] F. Balbo and S. Pinson, "Dynamic modeling of a disturbance in a multi-agent system for traffic regulation", Decision Support Systems, Issue 41, pages 131-146, 2005.
[6] D. Meignan, O. Simonin and A. Koukam, "Simulation and evaluation of urban bus-networks using a multiagent approach", Simulation Modelling Practice and Theory, Issue 15, pages 659-671, 2007.
[7] J. Banks, "Discrete event simulation", Proceedings of the 1999 Winter Simulation Conference, pages 7-13, 1999.
[8] W. David Kelton, Randall P. Sadowski, David T. Sturrock, Simulation with Arena, 3rd edition, New York, 2004.

 