D1.2 - A framework for developing an objective function for evaluating work package solutions (Cost function)

Grant Agreement N°:	FP7 - SCP0 – GA – 2011 - 265647
Project Acronym:	**ON-TIME**

Project Title: **O**ptimal **N**etworks for **T**rain **I**ntegration **M**anagement across **E**urope

Funding scheme:	Collaborative Project
Project start:	1 November 2011
Project duration:	3 Years
Work package no.:	WP1
Deliverable no.:	ONT-WP01-DEL-002
Status/date of document:	Final, 19/06/2012 (revised 27/06/2013)
Due date of document:	30/04/2012
Actual submission date:	19/06/2012 (revised 27/06/2013)
Lead contractor for this document:	University of Birmingham
Project website:	www.ontime-project.eu

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		
Dissemination Level		
PU	Public	**X**
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Revision control / involved partners

Following table gives an overview on elaboration and processed changes of the document:

Revision	Date	Name / Company short name	Changes
1	14/03/12	UoB	Draft for discussion
2	24/03/12	UoB	Update following feedback from consortium
3	19/06/12	UoB	Final draft. Update after partner comments
4	27/06/13	UoB	Amendments following EC review

Following project partners have been involved in the elaboration of this document:

Partner No.	Company short name	Involved experts
9	UoB	Clive Roberts, Lei Chen, Menglei Lu, Linsha Dai, Chris Bouch, Gemma Nicholson, Felix Schmid.

Executive Summary

The aim of the ON-TIME project is ‘to improve railway customer satisfaction through increased capacity and decreased delays for passengers and freight’. This document details the function that will be used to evaluate the outputs of the ON-TIME project against its high level aim.

The aim of this document is not to definitively define an objective function for the whole project; rather the objective is to identify the parameters that should be measured and assessed to help decide whether one solution (or outcome) is better than another. The level of improvement (or equally deterioration) of the parameters brought about by particular solutions will be considered during the evaluation phase of the project, but should be specified at this stage to allow those in the project to understand how solutions will be assessed. This document, therefore describes an ‘evaluation framework’ that will be used to assess the solutions developed in WP3, WP4, WP5 and WP6.

To address the requirements of the market, infrastructure planning, timetabling and operations (as emphasised in UIC 406), a general high-level measure of ‘Quality of Service’ (QoS) will be adopted in this project.

1 INTRODUCTION

The aim of the ON-TIME project is 'to improve railway customer satisfaction through increased capacity and decreased delays for passengers and freight'. This document details the function that will be used to evaluate the outputs of the ON-TIME project against its high level aim.

In order to assess whether one solution is better than another, a function must be developed to describe what a good outcome looks like. In the area of control engineering, these functions are commonly referred to as objective functions, or specifically as cost functions when the requirement is to minimise a particular outcome (e.g. delay), or as utility functions when the requirement is to maximise a particular outcome (e.g. capacity).

The aim of this document is not to definitively define an objective function for the whole project; rather the objective is to identify the parameters that should be measured and assessed to help decide whether one solution (or outcome) is better than another. The level of improvement (or equally deterioration) of the parameters brought about by particular solutions will be considered during the evaluation phase of the project, but should be specified at this stage to allow those in the project to understand how solutions will be assessed. This document, therefore describes an 'evaluation framework' that will be used to assess the solutions developed in WP3, WP4, WP5 and WP6.

To address the requirements of the market, infrastructure planning, timetabling and operations (as emphasised in UIC 406), a general high-level measure of 'Quality of Service' (QoS) will be adopted in this project. A decomposition of the QoS concept is shown in Figure 1. The top of Figure 1 shows the key performance indicators (KPI) for QoS and the associated quantitative key measures. The bottom of Figure 1 shows the static and dynamic factors that influence the QoS.

The following sections define the individual KPIs and Key Measures and consider how that might be used in each of the work packages.

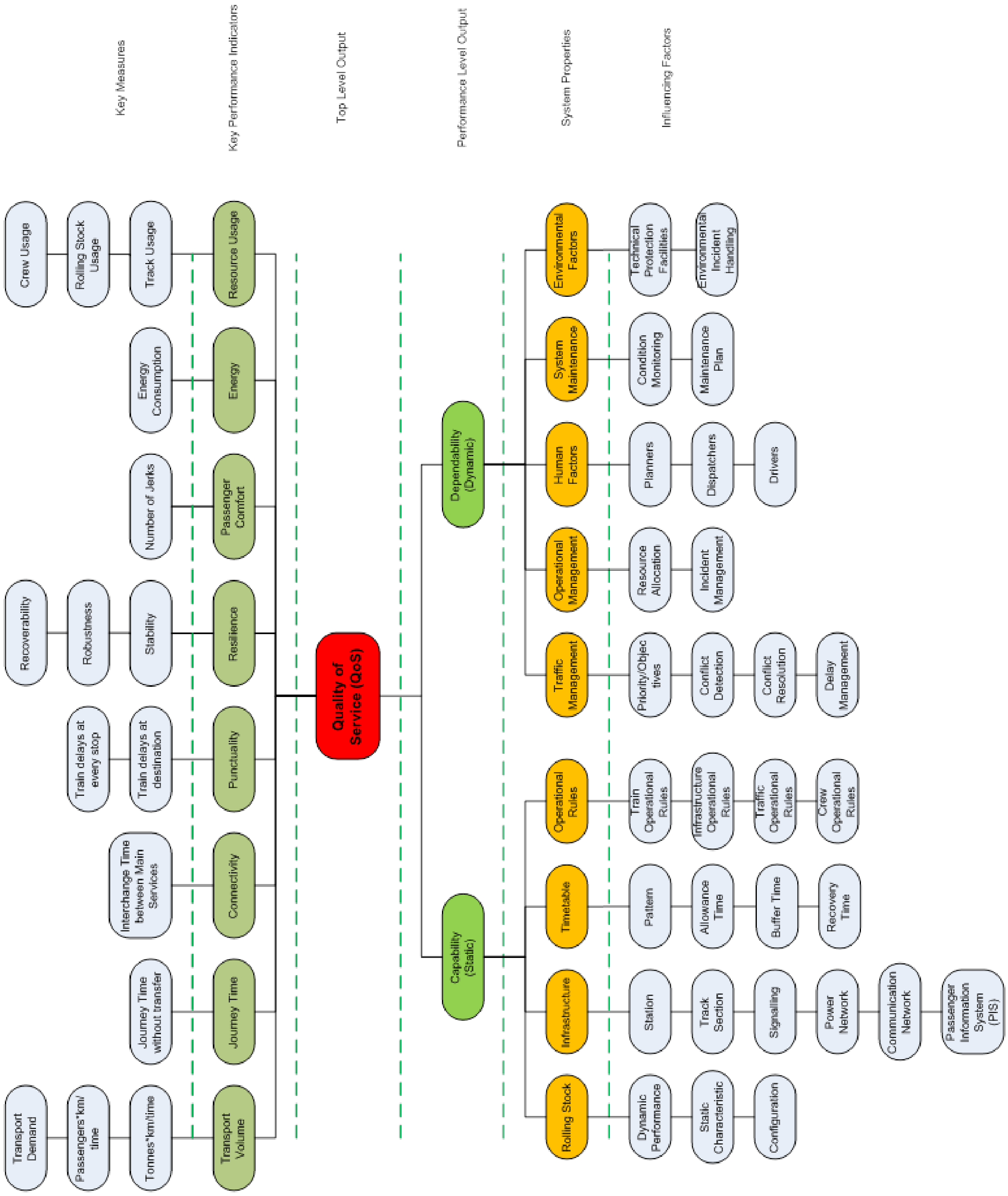

Figure 1 - Quality of Service Diagram

2 REVIEW OF KEY TERMS

2.1 Transport Volume

Transport Volume is defined as the volume of products (passengers or cargo) that can actually be delivered by a transport system over specific infrastructures. In this project, **passenger transport volume**, $\mathbf{V^P}$, is defined as the number of available passenger kilometres per unit time. **Cargo transport volume**, $\mathbf{V^F}$, is defined as the number of available cargo tonne kilometres per unit time. The transport volume should be considered relative to the **passenger** or **cargo transport demand**, $\mathbf{V^{PD}}$ or $\mathbf{V^{FD}}$, defined as the number of passenger kilometres or cargo tonne kilometres demanded over a given infrastructure per unit time.

2.2 Journey Time

In general, journey time is considered as the total practical consumed time for a passenger to complete his trip. In this project the timetabled journey time for passengers or cargo is considered. The *i*th **journey time** between any two stations, $\mathbf{T_i}$, is defined as the total planned time in the timetable between a vehicle's departure from the first station and its arrival at the second station in seconds, with no transfer allowed. The journey time should be considered relative to the *i*th **minimum sectional running time** between the same two stations, $\mathbf{T^M_i}$, computed as the basis of the nominal timetable regulation, using nominal values for all the variables involved in the computation (e.g. Track length, gradient, curvature, traction and braking characteristics, train length, train weight and its distribution, rolling resistances, adhesion, diameter of the wheels, voltage in the power line, etc.). Moreover, the computation is usually done with a "stressed" driving behaviour (i.e. the train starts immediately with no reaction time, accelerates as hard as possible to reach the maximum allowed speed as early as possible, keeps maximum speed as long as possible, and brakes at the latest moment with the service deceleration to come to its halt).

2.3 Connectivity

At a given interchange, where a connection is aimed at from the points of traffic demand and line planning, the *i*th **passenger** or **cargo interchange time** between any two services, $\mathbf{I^P_i}$ or $\mathbf{I^F_i}$, is the timetabled time in seconds between a passenger/cargo arriving on the first service and departing on the second service. The interchange time should be considered relative to the *i*th **minimum necessary passenger or freight interchange time**, between the same two services, $\mathbf{I^{PM}_i}$ or $\mathbf{I^{FM}_i}$.

2.4 Punctuality

Punctuality is the characteristic of being able to complete a required task or fulfil an obligation before or at a previously designated time; it is also an important measure of the performance of train operations.

The **deviation of service i at point j**, $\mathbf{L_{ij}}$, is calculated as the time in seconds between the actual arrival or departure time, t_{ij}, and the arrival or departure time scheduled in the published customer timetable, t^s_{ij}: $L_{ij} = |t_{ij} - t^s_{ij}|$.

2.5 Resilience

Resilience is defined as the ability of a system to withstand stresses, pressures, perturbations, unpredictable changes or variations in its operating environment without loss of functionality. We define three levels of resilience:

Stability (RS1): the ability to recover without active train rescheduling.

Robustness (RS2): the ability to recover with active train rescheduling/ordering.

Recoverability (RS3): the ability to recover with operational management measures such as train cancellation, rolling stock re-allocation etc.

The resilience will be evaluated using three key measures, the **time to recover**, **R**, the **peak delay**, **P**, and the **integral of total delay**, **D**. These three terms are defined in the appendix of this document, together with a description of their measurement and that of related terms using the Graffica HERMES simulator.

2.6 Passenger Comfort

In railway transport, to guarantee a good level of comfort for passengers on their journeys, many factors need to be considered. This deliverable focuses on the influence of the smoothness of the train driving performance on passenger comfort during journeys. **Jerk**, **J(t)**, is the rate of change of acceleration with respect to time,

$$J(t) = \frac{d\boldsymbol{a}}{dt} = \frac{d^2\boldsymbol{v}}{dt^2} = \frac{d^3\boldsymbol{x}}{dt^3},$$

where **a** is acceleration, **v** is velocity, **x** is position and **t** is time. We will consider the number and severity of jerks relative to a threshold value, J_{max}, specified by EC comfort levels.

2.7 Energy

Generally, energy consumption in the railway system includes energy consumed both by running rolling stock and infrastructure such as stations, signalling systems, etc. In this project, only the energy consumed by vehicles is considered. Measure **energy**

consumption by freight or **passenger vehicles**, $\mathbf{E^F}$ or $\mathbf{E^P}$, as the energy consumed in kWh by all vehicles running in a defined simulation area during a given time period.

2.8 Resource Usage

The resources used in the railway system include three main aspects: track usage (RU1), rolling stock usage (RU2) and crew usage (RU3).

In this project the measure of **track usage percentage**, **K**, over a given time window (in minutes) should be based upon the measurement method of capacity consumption outlined in UIC406, section 3.6. The level of **traffic demand** is denoted $\mathbf{K^D}$.

The **rolling stock usage percentage**, **S**, over a given time window is defined as the average percentage of the rolling stock in use over a given time window, relative to the maximum amount of rolling stock available.

The **crew utilisation**, **U**, is a measure of the number of paid man-hours worked by the crew over a given time period.

3 QUALITY OF SERVICE OF RAILWAY OPERATIONS

3.1 Quality of Service Diagram

A structured framework of the objective functions for WP3, WP4, WP5 and WP6 is proposed in this deliverable with a general high level measure QoS, as shown in Figure 1. Quality of Service is an indication of the comprehensive performance of the railway system. It covers Transport volume, Journey time, Connectivity, Punctuality, Resilience, Passenger comfort, Energy and Resource usage. The railway systems are expected to be optimal in terms of all the indicators, however, trade-offs need to be made in practice due to the various constraints in real life railway operations.

On the engineering side, the factors affecting Quality of Service can be broken down into Capability and Dependability. Capability covers all the "static" elements that are relatively hard to change, such as Rolling Stock, Infrastructure, Timetable and Operational Rules. Dependability includes all the dynamic components of the system; Traffic Management, Operational Management, Human Factors, System Maintenance and Environmental Factors. These components can be modified in short term practice.

Deliverable 1.2 aims to provide a framework to form the multi-objective decision making method for WP3, WP4, WP5 and WP6 in terms of the key performance indicators in the QoS measure. For each KPI, an objective function composed of certain key terms (defined in Section 2) and weighting functions, w_m, subject to the necessary constraints, c_j, must be developed. The weightings and constraints must be chosen in such a way, that fairness between operators is guaranteed in accordance with the respective network statement. The general format and a high level description outlining the aim of each objective function are given in Section 3.2 below. Each objective function is loosely defined as a constrained optimisation problem, but could be replaced by a series of unconstrained optimisation problems with appropriate penalty functions.

Each WP needs to consider different performance indicators. The framework is proposed for each WP with different combinations of KPIs. The objective functions should be combined with appropriate weighting functions, W_m, to form an overall decision making method, O_{WPi}, for WP3 to WP6.

3.2 Definition of Objective Function Elements

3.2.1 Transport Volume (TV)

Maximise the available passenger kilometres over a time window

Maximise the available cargo tonne kilometres in a time window

$TV = \max f(V^P, V^F, V^{PD}, V^{FD}, w_m^{TV})$

subject to c_{ji}^{TV};

where the weighting functions are dependent on the operating priorities of a particular line or network.

3.2.2 Journey Time (JT)

Minimise the journey times relative to the minimum sectional running times

$JT = \min f(T_i, T_i^M, w_m^{TV})$

subject to c_j^{JT}.

3.2.3 Connectivity (CN)

Minimise the interchange times between selected services relative to minimum necessary interchange times

$CN = \min f(\mathrm{I}_i^{\mathrm{P}}, \mathrm{I}_i^{\mathrm{F}}, \mathrm{I}_i^{\mathrm{PM}}, \mathrm{I}_i^{\mathrm{FM}}, \mathrm{w}_{\mathrm{m}}^{\mathrm{CN}})$

subject to c_j^{CN}

3.2.4 Punctuality (PT)

Minimise all train delays at selected stations

$PT = \min f(L_{ij}, w_m^{PT})$

subject to c_j^{PT} ,

where the weighting functions could incorporate the passenger numbers of the train.

3.2.5 Resilience (RS)

Minimise the delay propagation in the system

Stability:

$RS1 = \min\ f(R, P, D, w_m^{RS1})$

subject to c_j^{RS1}

Robustness:

$RS2 = \min\ f(R, P, D, w_m^{RS2})$

subject to c_j^{RS2}

Recoverability:

$RS3 = \min\ f(R, P, D, w_m^{RS3})$

subject to c_j^{RS3}

Overall, resilience is a combination of some or all of the terms stability (RS1), robustness (RS2) and recoverability (RS3), depending on the work package.

$RS = f(RS1, RS2, RS3, w_m^{RS})$

3.2.6 Passenger Comfort (PC)

Minimise the number and severity of jerks over a level defined by EC standard comfort levels

$PC = \min f(J, w_m^{PC})$

subject to c_j^{PC}

3.2.7 Energy (EG)

Minimise the sum of energy consumed by trains

$EG = \min f(E^P, E^F, w_m^{EG})$

subject to c_j^{EG}

3.2.8 Resource Usage (RU)

Minimise the track utilisation percentage subject to the minimum traffic demand being met

$RU1 = \min f(K, K^D, w_m^{RU1})$

subject to c_j^{RU1}

Minimise the rolling stock utilisation percentage

$RU2 = \min f(S, w_m^{RU2})$

subject to c_j^{RU2}

Minimise the crew utilisation

$RU3 = \min f(U, w_m^{RU3})$

subject to c_j^{RU3}

Minimise overall resource usage

Resource usage

$RU = f(RU1,\ RU2,\ RU3,\ w_m^{RU})$

4 OBJECTIVE FUNCTION DEFINITION FOR WORK PACKAGES

4.1 Generic definition of objective functions

The KPIs that each work package needs to consider as objectives are shown in Table 1. For each WP, the overall decision making method is formed as follows:

$O_{WPi} = f(TV, JT, CN, PT, RS, PC, EG, RU, W_n)$,

where the weighting function W_n can be 0 if it is not applied in the given WP. The determination of weighting functions needs to consider the main objective trade-off in each WP, and may be changed during the process of the ON-TIME project.

4.2 Objective function for individual WPs

For individual WPs, Table 1 lists the key measures that need to be combined in the objective functions for each WP.

	TV	JT	CN	PT	RS	PC	EG	RU
WP3	✓	✓	✓		✓		✓	✓

WP4		✓	✓	✓	✓		✓	✓
WP5	✓	✓	✓	✓	✓			✓
WP6		✓	✓	✓	✓	✓	✓	✓

Table 1 Key measures applied in WPs

5 CONCLUSIONS

This document identifies the parameters that should be measured and assessed to help decide whether one solution (or outcome) is better than another during the latter stages of the ONTIME project. The level of improvement (or equally deterioration) of the parameters brought about by particular solutions can be assessed using the evaluation framework. This document is therefore for direct use by the individual work packages (WP3, WP4, WP5 and WP6) and relevant sections should be incorporated into the requirements for these work packages.

APPENDIX

Using the Graffica HERMES simulator it is possible to simulate the running of a given timetable with or without artificially induced delay scenarios. The progress of the timetable may be followed and observations of the time at which trains arrive and depart stations and pass signals are made. The following describes a numerical and graphical evaluation method of the delay that develops in the system.

Consider all M journeys that run in a simulation area over a given time period, v_i, and record the time at which train journey i passes its jth observation point, t_{ij}. For all train journeys, record the **delay** of train i at its jth observation point, L_{ij}. This will be calculated as the time in seconds between the actual time and the scheduled time, t^s_{ij}: $L_{ij} = t_{ij} - t^s_{ij}$.

This results in a discrete set of observations of the system. We may define a continuous (step) function for each train journey, representing the most recent delay recorded by that journey. At the time the k^{th} observation point is passed until the $k+1^{th}$ observation point is reached, the **current delay**, $L^c_i(t)$, is the most recent lateness value recorded. Thus for t such that $t_{ik} \le t < t_{i,k+1}$, $L^c_i(t) = L_{ik}$.

Then at any time t, we may define the **sum of delays** (or **delay of the system**) for all M trains in the simulation as

$$L(t) = \sum_{i=1}^{M} L_i^c(t).$$

The overall sum of delays is event driven and will be updated each time a train reaches an observation point.

The output consists of a graph on which L(t) and the lateness of the most severely delayed trains are plotted. The **time to recover**, **R**, defined as the time between the delay of the system increasing above a small threshold e.g. 10 seconds, and it returning below this threshold, is recorded. The **peak delay**, **P**, is the maximum value of the delay of the system,

$$P = \max_t L(t).$$

The **integral of total delay**, **D**, is given by the area under the graph of delay of the system, L(t). In practice this may be calculated by placing all *N* observation times, t_{ij}, in chronological order and redefining them as $[t_1, t_2, \ldots, t_N]$. Then the integral of total delay is:

$$D = \sum_{s=1}^{N-1} (t_{s+1} - t_s)\ .\frac{(L(t_{s+1}) + L(t_s))}{2}$$

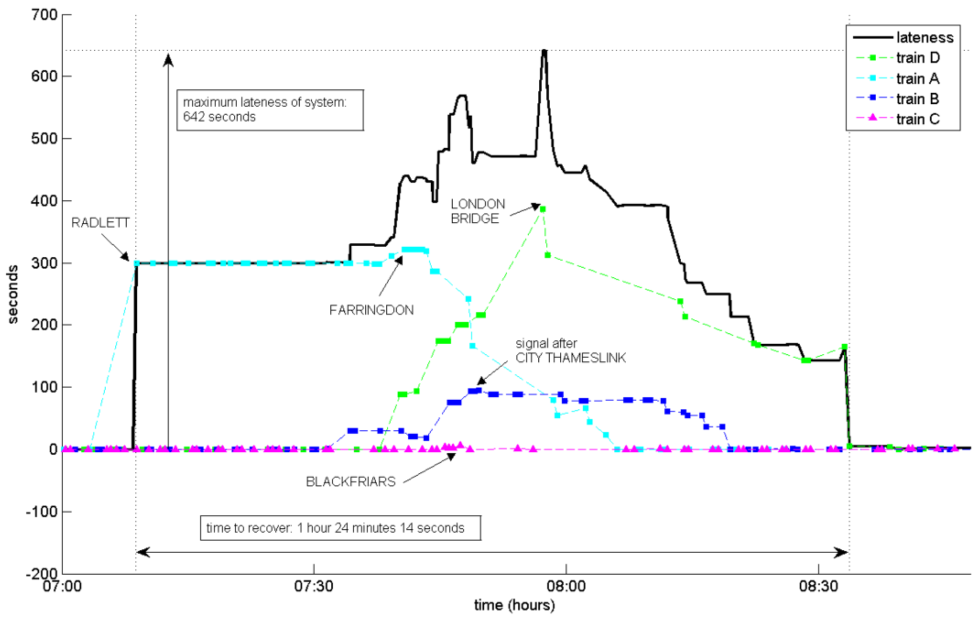

Figure 2 - Example of time to recover (R), peak delay (P) and integral of total delay (D) definitions