

Project Acronym:	<b>ETS</b>
Contract Number:	<b>14041 - ITEA3 Call 1</b>
Starting Date:	01/01/2016
Ending Date:	31/12/2018

Deliverable Number:	<b>D4.3</b>
Title of Deliverable:	Mobile application based on the real-time model for customer journey optimization
Work-Package of the Deliverable:	<b>WP4: Data Analytics</b>
Deliverable Type:	SW
Distribution:	
Author(s):	IDI EIKON
Contact point:	Gabriel Pérez
Other Contributors:	All the partners

**Abstract:**

The objective of this document is to create mobile applications based on the real-time scalable model developed in D4.1.

<b>Project Co-ordinator</b>	
Company Name:	TMOB
Name of representative:	Duygu Kara
Address:	Yeşilce Yeşilce Mah. Yunus Emre Cad. No: 8 Nil Ticaret Merkezi Zemin Kat, 34418 Kağıthane/İstanbul, Türkiye
Phone Number	+90 541 948 31 06
Fax Number:	
E-mail:	dkara@tmobtech.com
Project WEB site address:	<a href="https://electronicticketing.org/">https://electronicticketing.org/</a>

- \* (a) **PR** = Prototype  
**RE** = Report  
**SP** = Specification  
**OT** = Others
- \* (b) **P** = Public, for wide dissemination (public deliverables shall be of professional standard in a form suitable for print or electronic publication)  
**C** = Confidential, limited to project participants. Irrespective of the status, all the reports and deliverables must be made accessible to the other project participants and responsible EC services.

## Copyright

© 2018 The **ETS** Consortium, consisting of:

- T-Mob, Turkey
- Datong Solution, Korea, Republic of
- Siveco, Romania
- Convex Network, Romania
- Investigación y Desarrollo Informático, Spain

All rights reserved.

This document may not be copied, reproduced, or modified in whole or in part for any purpose without written permission from the **ETS** Consortium. In presence of such written permission, or when the circulation of the document is termed as “public”, an acknowledgement of the authors and of all applicable portions of the copyright notice must be clearly referenced.

This document may change without prior advice.

## Document History

Version <sup>1</sup>	Issue Date	Stage <sup>2</sup>	Content and changes
#1.0	07/05/2018	Draft	Structure and contents defined by WP leader
#1.1	01/06/2018		Integration and last feedback
#1.2	19/03/2019	Final	Final review and submission

For further information related to this Deliverable or to the **ETS** project please visit the project Web site:

**<https://electronticketing.org/>**

or contact the Project Coordinator: **[dkara@tmobtech.com](mailto:dkara@tmobtech.com)**

---

<sup>1</sup> Please use a new number for each new version of the deliverable. Add the date when this version was issued and list the items that have been added or changed. The ‘what’s new’ column will help the reader in identifying the relevant changes. Don’t forget to update the version number and date on the header.

<sup>2</sup> A deliverable can be in either of these stages: “draft” or “final”. For each stage, several versions of a document can be issued. *Draft*: Work is being done on the contents. *Final*: All chapters have been completed.

## Table of Content

1.	ETS Focus .....	4
2.	ETS Real-time Models.....	5
3.	ETS Mobile App.....	7
4.	Journeys Optimisation .....	10
4.1	Data Analytics optimisation .....	11
4.2	Infrastructure optimisation (for changes) .....	13
4.3	Statistics optimisation.....	14
4.4	ETS prediction optimisation .....	15
4.5	Selecting Journeys .....	17
4.6	Customers Satisfaction .....	17
5.	The reality of Data Analytics and Big Data Status.....	20
5.1	Conclusion.....	20

## **1. ETS Focus**

ETS mobile ticketing involves turning the mobile phone into both ticket machine and ticket. This removes the need for passengers to carry cash or wait in line to buy a ticket.

ETS mobile option is intended to allow passengers to travel without tickets, instead using Bluetooth and geolocation technology to track a passenger's movements and automatically charge their travel account when exiting a station.

Integration in ETS (Deliverable 5.1) comes through different modules converging and being unified at one same point: the back end or central system collecting all the information flowing through the different modules.

The flowing of the different modules happens through web services, requiring an invocation at one end and a response from the central system (the other end).

The challenge ETS has at hand when trying to optimise procedures is to minimise the risk of incomplete or even abandoned purchase and to keep the customer coming back.

## 2. ETS Real-time Models

In D4.1 a model for infrastructure occupancy, services and density map, was developed. It could be seen working through a Portal and was based in two simulation types: **object-oriented** and **event-based**

Genetic Algorithms and a Predictive system were used. The architecture chosen for training was ANN (Artificial Neural Network) and for Fraud Detection (automatic detection of fraud) the election was Tensorflow and Keras.

In D4.3 (this document) the optimization functions developed in D4.1 models have been passed to the mobile application, according to users profiles.

For connecting the NoSQL module with the central system, different web services have been prepared:

**ticketUse:** Implemented for the insertion of electronic ticket usage data, it is used from the central platform to send the data to the portal which, after filtering and pertinent checks, inserts it into the Cassandra database.

**insertPosition:** For the insertion of positions during a user tour, once the mobile application grants permission for tracking by GPS to access more advanced features of route optimization and offering information such as nearby points of interest, offers or promotions nearby. It can be used either directly from the mobile application or by grouping these requests in the central system, since it is this the one communicating with the platform.

**getDensityMap:** To obtain the same density data that can be viewed in real time. To be invoked for example from informative panels in the stations that want to draw the occupation in the different lines, the arrival of next trains, etc. to provide useful information for the user.

**getPredictionDensityMap:** To obtain the density data but making a prediction indicating the moment (in the near future) and / or varying some element (like elimination of a train, etc) to use from other modules of the platform some of the learning capabilities of the platform or the simulation of events to study the results in the infrastructure against changes.

**getUserPrediction:** Get response for the prediction of the number of people waiting at a station. The input parameters are the identifier of the station, date and time to make the prediction and a chain as an

indicator of an event or special situation, such as a great show nearby, a special day such as election day or strike, etc.

**insertMapOptimization:** As the system is dynamic, we can load a new map, require information of each station from any of the lines, type of line (bus, train, ferry, ...), connection time between each line, frequency of service (at different day ranges), and other relevant data.

**getOptimization:** For a determined map, origin and destination, to get the optimized route, according time or travel comfort.

As it can be seen the last two services are optimization specific.

### 3. ETS Mobile App

ETS Mobile App is contributed by TMOB. The app is available in Android and includes a mobile ticketing feature that allows users to purchase fares in advance and to display their fares on their smartphone as proof of payment.

Mobile ticketing allows people to buy and use fares in real-time. The app also includes some of most popular trip planning features.

ETS mobile app provides an intuitive mobile ticketing application that allows passengers to see real-time departure information and purchase tickets with just a few taps on their phone. It is designed to incorporate leading edge electronic payment technologies to make transit use more convenient for customers, providing metro operators with a more efficient and secure fare collection platform and responding in this way to the growing use of smartphones to access transportation services.

#### How ETS mobile electronic ticketing works

User should assure their mobile is active when entering a metro station. The mobile phone must be activated before going through a corridor (for the system to detect it and not being considered as a potential fraud threat).

- 1) LOGIN
- 2) **CREATE AN ACCOUNT**
- 3) From TICKETS, you need to ADD FIRSTCARD  
Type can be Anonymous or Personal  
Personal ID can be entered manually or by QR  
ADD NEW CARD (you can add as many cards as you want)
- 4) From CARDS user can manage their smart tickets. From the suboptions inside (clicking on ) user can see Cards, their balance information, their histories, top-up them, edit and save and delete. Also switching cards is in here, since although different cards are allowed the active one is the one selected.
- 5) PAYMENT METHOD is by credit or debit card after selecting an amount. Pre-Payment is the option in ETS. This means ETS card must have a positive balance to be detected as valid.

- 6) When doing the TOP-UP users can change the card to pay by clicking the CHANGE label
- 7) Under the MENU TICKETS it is possible to buy tickets and/or see their history
- 8) In the MENU ROUTES user can reach the transportation map
- 9) In the MENU SETTINGS, user can change Language, Profile, Notifications, Terms & Conditions, get information about app or get in contact for support and/or make comments and suggestions.

### **To use an ETS Smart Card (electronic ticket):**

Passengers need to have active their mobile phone.

The mobile phone can determine user geographical position using GPS (Global Positioning System). The information about user position is used by a number of applications on the mobile phone, such as navigation, the search function and the weather forecast.

But users willing to share their location information (and specially since the entry into force of the new data protection regulation), should activate this function on their mobile and give their express consent to the operator requesting it (can be deactivated at any moment).

In ETS is used for customising ads and/or presenting offers and events along the user route, etc. Using these consumer characteristics is important to time and target marketing. There are numerous examples of how venues define audiences can be effectively targeted by the geo-position function. If not activated these options cannot function.

If user selects GPS, Wi-Fi and mobile networks, the mobile phone can find his/her position quickly and with the highest possible accuracy. Satellite-based GPS requires a clear view of the sky. If user selects Wi-Fi and mobile networks, the mobile phone can find his/hers approximate position using the mobile network and nearby WLANs. If user selects GPS only, the mobile phone can find the exact position but it may take longer as there is no access to supplementary information from the mobile network or nearby WLANs,

## Reminder:

- Tickets are stored on your phone (which makes it possible to use them without an internet connection), so uninstalling the app can permanently erase your tickets!

In ETS main channel is mobile (in itself a multichannel platform (since communication may happen in different ways: text messages, mobile email, app...), but also other channel options are considered. Besides mobiles, according to a forecast from Statista (the Statistics Portal for Market Data), by 2020, the average number of connected devices per person is expected to be 6.58, a whopping spike from 3.47 in 2015. These are not just mobile devices, though and it simply implies that it's not only wise but also imperative, to direct efforts in the direction of creating a **multichannel** experience.

## 4. Journeys Optimisation

The mobile app counts with same features as the web application. Connection takes place through web services at each end (request/response).

Since data analytics deals with forecasting, prediction and optimization, generally speaking, the condition for optimizing routes is the access to many data.

To be able to optimize journeys, it is important to first collate this data into a unified, coherent whole. Looking at the entire journey across all the channels gives an organic view of what customer's journey is actually like, and what could be improved.



ETS customer journey optimization is instrumental for businesses, and its approach aims at ensuring a seamless shopping experience while passengers buy their electronic tickets on their smartphones.

ETS analysis takes into account multiple issues to obtain the best results in terms of efficiency and flexibility.

With ETS data analytics operators can improve their reliability in many areas, as can be reducing downtime and optimising maintenance

The models developed facilitate line modification, removing or inserting trains or even changing the maximum running speeds on those track circuits.

Travel times and line frequencies can be also analyzed, so that finally the best headway to be implemented on the line can be proposed for the user, who is able to know how to improve the main line features like capacity and frequency of trains.

## 4.1 Data Analytics optimisation

In the mobile application, the data model designed is ready to use transport network data and to integrate with the different systems through Web services. Algorithms developed improve the infrastructure management.

To make a prediction on the density and movement of the flow of users in the transport network (beyond current situation), allowing to anticipate possible problems such as bottlenecks in the transport network, historical information combined with online data, such as passenger arrivals and departures at the different stations throughout the day is key. As by now is not yet available in central database, so we have to produce it to be able to create predictions.

We use the historical information produced combined with online data, such as passenger arrivals and departures at the different stations throughout the day, to make a prediction on the density and movement of the flow of users in the transport network, allowing to anticipate possible problems such as bottlenecks in the transport network.

A genetic algorithm is used together with simulation and event-based modelling to improve aspects of the infrastructure, in particular to optimize the distribution of vehicles on different lines or transit frequencies and to study the effect of the incorporation of new vehicles or remove them from a line or similar alterations.

The algorithm developed receives as input the number of days to carry out the simulation or the dates between the process to be performed, the number of users travelling, the number of trips per user and by fixing this number a distribution model taken the average of the trips made by the passengers and the probability that an anomalous data will be generated follows.

Parameters as recharges or varying amount of recharges, new tickets acquisitions, distribution of the hours of use within limits, travel duration based on the route with random modifications, etc. are considered when performing optimisation so that when inserting the desired data in the database, generation takes place through the interfaces created for the integration with the central platform

Central system and mobile application structures follow a relational database design. For data analytics and big data the option was of a NoSQL structure, in concrete Cassandra. To match both structures and to be able to deal with geo-locations and optimization we

designed a single table (see next table structure) where the electronic ticket and the location could be stored:

Field	Type
id	uuid
cardid	bigint
createdate	timestamp
currentbalance	double
currentpasscount	double
customerid	bigint
loyaltypasscount	double
opt_station_list	list< bigint >
opt_station_list_ts	list< timestamp >
stationid_entrance	bigint
stationid_exit	bigint
status	text
travel_time	int
ts_entrance	timestamp
ts_exit	timestamp
validationdate	timestamp
validator_code_entrance	text
validator_code_exit	text

Since NoSQL behaves differently, the optimization and speed of the queries, the information relating to a complete trip is not found in different rows but there is a list of stations of the route and temporary mark in each of the stations.

Data from other modules come to NoSQL data base from the central platform (facial detection, ticket use, etc.)

Regarding the data of the mobile platform referring to the location needed in a NoSQL structure, subject to limitations and privacy, they must be stored as a list of positions per user and trip, in order to optimize the number of rows in the table to take advantage of the characteristics of dynamic growth of the information contained in a row of the NoSQL database. We designed next table containing the minimum requirements needed related to positions for performing analysis and optimization of journeys.

Field	Type
id	uuid
cardid	bigint
customerid	bigint
location_latitude	list< float >
location_longitude	list< float >
location_ts	list< timestamp >
travel_init_ts	timestamp
travel_end_ts	timestamp

In this way in a single row the information relative to a complete trip (between the dates *travel\_init\_ts* and *travel\_end\_ts*) is stored for the same user, such as a list of positions (latitude and longitude) and temporary mark of each of them. To decide if a position belongs to a trip it is checked if there is any row for that user that is not closed (without *travel\_end\_ts*) to update the list of positions or if the temporary mark does not belong to the same trip to close that trip if it was not already and create a new trip.

For the insertion of the data from the central platform, the data of the use of an electronic ticket is received through the *ticketUse* method, which has mandatory insertion parameters such as the ticket identifier or if it is an entry in the validator or a output or a detection waiting for a means of transport and the rest of the optional insertion parameters, such as information related to the balance of the card that will only be sent when a change occurs, for example at the input and/or output station.

The complete list of parameters sent to the web service corresponds to the information stored in the table of the Non-SQL database relating to the location and use of the electronic ticket.

For the insertion of positions there is a web service *insertPosition* with the parameters related to the ticket identifier, user identifier, latitude, longitude and time stamp, with which it will be inserted or updated if the information already exists in the corresponding row of the trip.

The web services developed (invocation) if get not response at the other end, show an ok meaning the service is ok but no data is received. Data availability has been the main handicap along the project.

The exiting of Elektronet in April 2018 did not contributed to make things better, since an interim period was open, conditioning the project continuity, that has not been closed until very recently, waiting for an answer by Turkish authorities.

## **4.2 Infrastructure optimisation (for changes)**

It is a graphical interface to use the event-oriented simulation detailed in the D4.1 deliverable to help in making decisions quickly on variations in the infrastructure, in particular if we want to eliminate a vehicle from a line or if we want to add a new one. We can vary some

data as the average of time frequency between the stations and the capacity in number of passengers of the vehicle.

We must also select the date on which the simulation will be launched as it will use the data of the trips of a specific day that are in the database,

The operation is simple, choose the date to use the data for that day and select a line to add a vehicle and click on *Launch Add* or a line on which to remove a vehicle and click on *Launch Remove*.

The output of the simulation in this case is not visual but measures of behavior and performance, specifically the impact of adding or removing a vehicle from the line, on the number of average passengers (because there are several simulations that are averaged and there is a random component in the data) waiting at each station and if its number has increased or decreased.

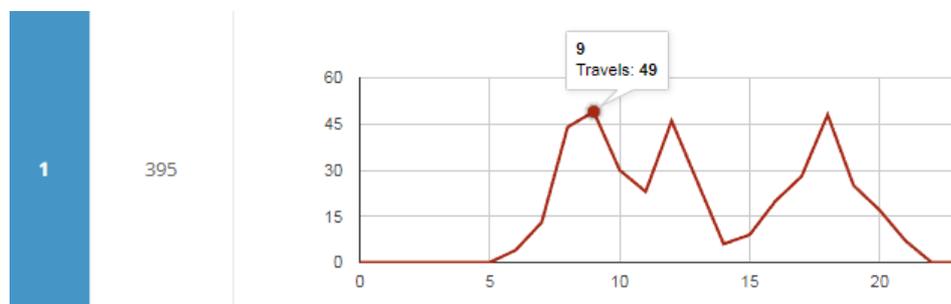
The screenshot displays the 'Infrastructure Change Simulation' interface with the following sections:

- Add vehicle:** Includes a dropdown for 'To line:' (M4), 'Capacity:' (checked 'default' or 'passengers'), and 'Mean travel time:' (checked 'default' or 'minutes'). A green 'Launch Add' button is present.
- Remove vehicle:** Includes a dropdown for 'From line:' (M1a) and a green 'Launch Remove' button.
- Simulate with data:** Includes a 'From day: Day:' input field with the value '18/10/2017'.
- Results:** A scrollable text area showing simulation results for four lines:
  - Line M3: 4.3 passenger (average) MORE waiting in all the stations of the line at the end of the day.
  - Line M2: 18.2 passenger (average) MORE waiting in all the stations of the line at the end of the day.
  - Line M6: 11.5 passenger (average) MORE waiting in all the stations of the line at the end of the day.
  - Line M4: 45.6 passenger (average) LESS waiting in all the stations of the line at the end of the day.

Screen capture for a specific date after pressing Launch Add

### 4.3 Statistics optimisation

As for **Statistics**, actually there is a single report available allowing to obtain the number of trips made in each station, together with a distribution of the number of trips per day in each station. This allows the graphic visualization of these data and the possibility of easily finding saturations and other problems in the infrastructure or simply visualizing the hours when saturation occurs in the different stations, contributing to a better decision making, regarding the need for modification or expansion of some element. This can be done by indicating the specific day (the calendar component only lets us to select between dates where there is data).



Example. Station number 1: 395 Total trips and their distribution during the day, if we place ourselves in one hour the concrete data is shown for example, 49 trips at 9:00

In addition, with respect to the chosen day, other information is shown such as the total number of trips that day in the infrastructure and the minimum and maximum travel time in addition to the average in minutes of the trips.

Partner SIVECO offers other graphical reports on the data, that can be seen in the Reporting module they contribute.

And, lastly, historical data to carry out statistics on the evolution and comparatives and thus controlling trends is ready to go for obtaining indicators as: Trends record; Change in user behaviour; Validation of models; Seasonality ...

#### 4.4 ETS prediction optimisation

Prediction is part of the process when it comes to crafting a well-thought-out path for customers and analytics offer the immense advantage of making informed predictions.

To provide and recommend customers with interest points along the route, not having real data about real geo-located potential points, based in a simulation system, we have designed a Predictive system.

At this point in the work, we have connected with the Central System (now under the control of TMOB) after the exiting of Elektronet (April 2018) and the mobile app (also TMOB). Web Service to connect with the Loyalty Application (Convex) is also ready.

ETS optimization can be used to predict emerging issues and maintenance requirements so operators can address them proactively to keep assets in service and to maximize uptime. Also, advanced analytics showing hidden patterns in the data can be used to quickly and accurately identify root causes.

Inside the station, analytics such as queue management and direction flow can help operators optimize staff utilization and make data-driven marketing decisions. In ETS, station managers (with the solution provided by DTS) can set up video analytics to monitor traffic flow around the station, monitoring the number of people flowing through the ticket barriers and free corridors, for example.

As for congestion, one of the critical points of main stations, it is possible to lower it by using data analytics. Data on peak hour ridership, places with the biggest traffic, and passenger travel behavior collected from station entrances and exits can help operators redeploy trains or make relevant marketing strategies to redirect traffic to less busy hours.

With an ever increasing knowledge, station managers can, for example, adjust their staffing in real-time to ensure customers are served as quickly as possible. Or timely alerts that can enable metro operators to repair or replace assets at the optimal time. Or the creation of comparisons of traffic patterns for different time periods in order to plan their staffing accordingly.

There are plenty of improvements, data analytics can contribute to, although it is true that many require the installation of sensors collecting the appropriate data to be analysed.

For example, to be able to analyze the comfort level of the station, sensors should be in place in them (temperature, humidity, carbon dioxide.....), this would help to make possible the automatization and adjustment of the HVAC system. Also data from sensors can help operators learn more about their customers, contributing to focus their marketing with strategies to optimize revenue and deliver a more personalized customer experience. Adopting cognitive

computing capabilities opens the opportunity to re-route trains, inform passengers of their choices and alternatives, and provide real-time updates through users' preferred channels. No doubt such actions will drive higher traveller satisfaction.

## 4.5 Selecting Journeys

In the casuistry of a case like ETS, once data is collected and put to work, it is likely having to analyze millions of data at one and the same time. With so much data, knowing what to measure is essential. There must be hard facts in numbers that correspond to the journeys. And we suggest, as interesting information to measure (KPIs), the following:

One possible measurement for Customer Journey



We have taken a route and analyzed and mapped customers' journeys in order to differentiate between the important ones and the trivial ones. We have figured out the costliest mistakes and focused on glitches as well as opportunities that we guess may have the highest potential effect on the business.

## 4.6 Customers Satisfaction

In Transport as in every other industry, service is made for its customers. It's created to either solve their problems or fulfil their needs.

This means the service revolves around customers and their experiences, and significant efforts have to be made to provide them with a positive experience.

The journey of providing customers with a positive experience starts from the moment they use one of our services and extends beyond the moment they become our customers. Although it's a never-ending commitment, encouraging them to become a loyal customer by retaining them is definitely considered a milestone of growth.

In ETS it is possible to collect customers' opinions and feedback by means of surveys, allowing us to know the level of satisfaction of customers

Analytics and specially predictions, allow us to create personalized experiences by asking the right questions and by connecting customers' previous interactions.

This is where what we call Customer Satisfaction Survey Questions can help.

In ETS we suggest to add this important feature in order to be able to collect such critical feedback about the service from users. We have created sample surveys for customers to rate experience, product, service, and the like. (See sample next).

#### SAMPLE OF SURVEYS QUESTIONS (for getting customers feedback)

1) Why do you ride metro?

School; Work; Recreation; Errands; Both equally

2) Usage Frequency

Daily , Weekly, Once a month, Every 2-3 months, 2-3 times a month, do not use

3) Based on your most recent experiences , please rate your satisfaction with ETS for each of the following:

Overall quality

Value

Purchase experience

Very dissatisfied, Dissatisfied, Neither satisfied nor dissatisfied, Satisfied, Very satisfied

1	2	3	4	5

What matter to you the most when choosing ETS? (open text)

How can we improve your ETS journey experience? (open text)

Do you have any additional comments or feedback for us? (open text)

In the future, would you be willing to take this survey again?  Yes -  No

In question 3, Satisfaction Scale measures from 1 to 5. 1 is very dissatisfied and 5 meaning the customer is very satisfied

In open-test questions participants are allowed to write out their response within a text box fully expressing their opinions instead of any pre-written response.

While these questions can sometimes be time-consuming to analyze, they encourage the participants to be honest and give them the freedom to address any topic.

Demographics could be also used, although here we enter into personal questions that deal with data protection and require in many countries, of an express consent.

Demographics make it easier for companies to segment customers into buyer personas. By grouping customers together based on key characteristics, this categorization helps to visualize a target audience.

Demographics: Age; Gender; Education; Employment status; Household income; Marital status; Children/dependents; Location (zip code); Ethnic background

## 5. The reality of Data Analytics and Big Data Status

Data Analytics or Business Analytics as many professionals prefer to say, is a forward-thinking way to improve operational efficiency. Decisions can be made faster, and it becomes easier to make sense of **large** volumes of data.

However capturing and analyzing big data is **a complex process** and although a large majority of companies agree that big data will impact their business (78%), only 58% think their company is ready to take advantage of all the potential that big data offers.

### 5.1 Conclusion

Understanding customer journeys is the key to tackling an ETS setup. The way the customers interact and deal with digital channels is not only dynamic, but also highly demanding. To meet their needs right when, and, more importantly, where they need, using data from various channels to create an experience that the user will appreciate, is imperative to survive and thrive amidst current multichannel reality.

In ETS data analytics through simulation scenarios to predict future conditions has been used to demonstrate the usefulness of new trends ongoing in the field, even if actually the real scenario chosen is not well positioned to capitalize on the potential.

Data analytics in ETS follows a technical approach to predict upcoming trends. This process helps find patterns after analyzing previous and current data and is used to devise future courses of action. The tools used to work in this field have been data mining, descriptive modelling, and simulations.