Smart public transit insights

Alexandru SIROMASCENKO,
Ingenios.ro, București, România
alex@ingenios.ro

Cosmin POPA,
Ingenios.ro, Galati, România
cosmin@ingenios.ro

Corina DIMA,
„Dunărea de Jos" University of Galați, România
Corina.Bocaneala@ugal.ro

Abstract

Local public administrations should pay special attention to the implementation of the smart mobility concept because public transport is a key factor for the sustainable development of cities. This paper explores the analysis of public transportation performance using intelligent algorithms that process data obtained from GPS devices and various other sensors operating on buses. The results obtained from this analysis form the basis of an intelligent management of the public transport system. The proposed analysis model aids in evaluating the transport system's performance and improving the feasibility and suitability of the transport plan. The analysis of historical monitoring data facilitates timely, fact-based decisions. Our work explains why advanced IT solutions for transport management based on analytics should be present in all smart city projects, given that the efficient mobility of citizens increases the level of socio-economic development.

Keywords: smart cities, intelligent public transportation, mobility.

1. Introduction

The mobility of citizens is a key factor influencing the development of cities. Public transportation ensures access to workplaces, hospitals, schools, culture and public services. Urban mobility comes with a series of challenges that consist of increased traffic congestion, pollution leading to climate change and energy crises.

For this reason, public transportation comes with a number of advantages: public transport can be used by all citizens, regardless of income, easy access to workplaces and other points of interest in big cities, increased road safety due to control and professional drivers, lower costs for the community, reduced traffic congestion, reduced costs for fuel consumption, reduced pollution [1], [2], [3].

The idea of developing an intelligent public transport system has appeared since the 1980s. A history of this concept's evolution is presented in [4]. Intelligent transport systems (ITS) are software applications that help with traffic management and supply information of interest to both administrative staff and passengers.

An ITS can provide historical data as well as real-time information for passengers and also for those who plan, design, build, and operate the transportation system [5]. One of the approaches of these transportation applications is studying the environmental impact of vehicle traffic [6]. Modern management systems take into account the fact that current transport fleets also include electric vehicles. Some of the ITS also address other aspects such as smart parking which requires careful planning [7] A classification of ITS can be consulted in [8].

The use of sensors in many important domains that require urgent responses such as transport or medical assistance was possible due to the technological development that made these devices accessible from a technical point of view and in terms of price [9], [10]. An intelligent transport system (ITS) for organizing transport that reduces traffic congestion, time to destination, fuel consumption and pollution was proposed in [11]. This approach includes the traffic surveillance, vehicle surveillance, passenger surveillance and driver surveillance. ITS technology involving wireless communication can lead to security and privacy challenges, authentication, integrity, location privacy, identity privacy, anonymity issues [12], [13].

Our paper approaches the design of intelligent transport systems from two perspectives: timetables and passenger volumes. The timetable is essential for the quality of public transportation because it fulfills several functions: it helps travelers plan their trips, ensures an evenly distributed passenger load, reduces transportation costs, helps the management to evaluate the performance of the public transportation system.

The proposed approach helps to verify if trips are taking place as planned. The succession interval of vehicles in the stations is an important indicator of the good performance of public transportation. Delays are not desirable, but it is also not appropriate for buses to arrive before the planned time. When gaps are noticed between the timetable and the actual trip, the cause must be checked (moments of the day with heavy traffic, technical problems, the driver's driving style) and, from this point, decisions can be made on whether the problems can be easily fixed operationally, or if the timetable must be revised. In order to make it easier to understand, the compliance with the timetable and the information about the trips can be represented graphically.

The work is structured in three sections. The first section describes the framework for analyzing the performance of intelligent transportation systems. The second section presents a proposed approach for transportation system monitoring. The last part is a conclusion section.

2. Analysis of the performance of the public transport system

The analysis of a public transport system’s performance is done from several perspectives:

- compliance with the trip schedule: the planned trips and vehicle allocation assume certain travel times and technical conditions which might not be achievable in practice, so certain planned trips might not be fulfilled;
- compliance with vehicle succession intervals on the route: similarly to the above perspective, but at a more detailed level, from a station perspective; this metric is important because it directly affects passenger satisfaction for several reasons: waiting time, travel time predictability, vehicle crowding;
- analysis of the number of passengers: the data received from the counting sensors installed on vehicles is aggregated at several levels – per vehicle, station, line and

the performance at each level is evaluated separately and combined in order to assess the adequacy of the transport planning to the demand;

- analysis of aggregated data using specialized charts and reports: the data must be presented from several interconnected perspectives, with each one being presented in a specialized, graphically suggestive way, in order to ensure both overviews and detailed analyses

2.1. The schedule

The traffic schedules are drawn up for each line and include the following stages:

- aggregating demand data based on demographics, land usage, surveys etc.;
- creating spatio-temporal origin-destination matrices;
- determining the stations and lines;
- creating the trip schedule by taking into account the temporal variability of demand;
- calculation of the necessary fleet on each line to cover the schedule from a cost efficient perspective, by taking into account depot locations, movement times, refueling, maximum usage time restrictions etc.;
- allocating drivers to the schedule while taking into account working conditions.

The allocation of vehicles is a critical step in establishing the schedule and is prone to generating problems if the input data does not reflect the reality on the ground: vehicle conditions and especially movement times throughout the day. Although certain problems can be resolved in real-time through vehicle reallocation, on the long run these create cost inefficiencies: a trip which is unfulfillable using the vehicles already on the road due to time desynchronizations should be rescheduled by taking into account real movement times instead of being assigned an available vehicle from the depot on the fly. This is where the feedback loop using real monitoring data allows the plan to be fixed.

2.2. Succession of vehicles on the route

The trips are not identical in most cases, either in length (as the duration per half-trip), or in density (the number of trips per hour is different across the day). Likewise, their succession and return to the terminus points is carried out in a strictly fixed order (but which changes depending on the periods of the day). The distance between two trips represents the circulation interval and can be represented down to station level.

This metric is important because it directly influences the system’s ability to manage passenger inputs at every point in the network and every moment of the day. An inadequate vehicle availability interval, at each station, can cause both overcrowding, if the interval is too long, and wasted resources, if the interval is too short, so an appropriate frequency must be planned and followed during the day.

2.3. Analysis of the number of passengers

All the streets of a city constitute its street network. The street network has a certain structure and is used by all traffic participants. The totality of passengers transported in one hour, in one direction or in both constitutes passenger traffic. Passenger traffic is measured in passengers per hour and per direction. The analysis of the potential of transportation of a neighborhood leads to the following relationship:

$U_j = L_j \frac{M}{L}$, where: (1)

- L_j is the number of inhabitants of the analyzed neighborhood;
- M is the total number of jobs in the city;
- L is the total number of inhabitants (of the city).

Using various measuring methods, such as bus sensors, surveys or street video cameras, the traffic flows for a certain urban area can be analyzed using appropriate mathematical models.

In addition to data collection, the scientific interpretation of recorded data is of particular importance. Comparing them with the values recorded in previous years, extrapolating them, establishing conclusions based on previous experiences constitute the elements of data processing and generalization. The basic principles for obtaining a technical-economic efficiency and a maximum precision of the measurements are the following:

- increasing the number of measurements, which has the consequence of increasing the probability of having accurate information about the studied phenomenon;
- the integration of information from various sources, which allows both data enrichment and cross-validation.

Using passenger counting sensors mounted on each vehicle door is a straightforward approach to creating a feedback loop for public transit performance evaluation and improvement. Such sensors, however, have certain limitations, especially in crowded scenarios and due to heterogenous passenger behavior. The data accuracy can be improved by using video surveillance with computer vision. This latter type of data can also be used for improving planning and scheduling, as it can also provide information for areas and time intervals not yet served by public transit.

2.4. Charts and reports

The data from the monitoring of public transit can be used in two main directions

- as a ready-aggregated input for the replanning and rescheduling activities;
- as a data repository with several aggregation levels available for aiding in problem identification and decision making.

For the former usage category, the data must be presented in such a way as to:

- ensure an adequate visualization of data, specific to the type of information being represented: movement times, deviations from plan, passenger numbers etc.;
- allow navigation from one level of aggregation to another;
- allow the visualization of correlations between different variables: individual trips, movement times, distance covered, vehicles, drivers, lines etc.

The reports and charts that the transport company management could use are:

- percentage of trips fulfilled;

- average movement times to/from stations and depots;
- deviations from real average and planned indicators;
- passenger counts per line and station;
- mesoscopic and microscopic events down to individual data packets received from the vehicle, in order to gain insights on the causes of deviations and anomalies: driver faults, missing sensor data, synchronization errors etc.

3. Proposed solutions

Mobility and passenger counter data in a public transport system can be obtained by installing devices on each public transport vehicle that can measure various indicators, such as GPS, counters for each access door, mobility data such as travel speed, deceleration force, odometer, fuel consumption, weight on each axle, data specific to the transport management: route, driver, bus, station at which it stopped, but also calculated data such as computed average speed, and distance traveled using GPS coordinates. Each bus transmits data continuously at a predetermined interval or discontinuously depending on a certain event that triggered the sending.

Multiple problems may arise, caused by the errors inherent to the hardware devices, the incorrect entry of certain details by the driver or the impossibility of data transmission due to signal obstruction. From a hardware point of view, there are a multitude of subsystems that work together, coordinated by a central system.

The following hardware systems can be placed within a vehicle for public transport:

- sensor that reads the data provided by the CAN bus, containing specific mobility information: acceleration, speed, braking etc.;
- sensors that read the number of passengers boarding or disembarking at each vehicle door;
- GPS sensor;
- mobile communication device.

The centralizer system installed on each vehicle collects a snapshot of these data and sends them according to preset scenarios: at a certain frequency or triggered by a certain event, eg: entering the station perimeter or opening the doors triggers the sending of a data package with the current indicator values.

In practice, problems may arise in the case of this hardware architecture, which will be detailed further on. The sensors that read the data provided by the CAN bus may not read the mobility data due to some hardware problems. The sensors that read the number of passengers can fail (only certain doors will transmit data). In the case of motion sensors, they can also be affected by the irregular movement of passengers or by crowding, which affects the measurement accuracy. The GPS sensor may have acquisition problems if the vehicle moves in certain areas where the GPS signal is not available or is available with very low accuracy.

This transmission can cause problems such as the incorrect detection of station stops, the incorrect reporting of non-compliance with the route, a speed/distance that is different from reality, etc. If the signal for the transmission of this data may be lost, certain routes may appear as missed.

Except for these potential problems due to the abnormal operation of the hardware components, there can be situations that arise because the driver does not stop at certain stations, does not correctly enter certain necessary data such as his ID or the selected route, drives faster than the planned schedule, etc.

The passenger counter data can be used either to determine the number of passengers in the respective vehicle, thus allowing for the generation of a heatmap map of the passenger density in a certain area or station.

We propose a model for data aggregation and analysis based on both hierarchical and network type relations, which aims to:

- take into account the different types of raw data:
    a) driver input data: driver id, selected line,
    b) vehicle driving indicators: odometer value,
    c) spatial awareness: GPS coordinates,
    d) transit indicators: door status, passenger counters;
- use operational context for data aggregation, matching and uncertainty reduction:
    a) GPS coordinates and structures of transit lines, stations and depots,
    b) vehicles and drivers,
    c) transit plan;
- provide resilience and minimize the impact of sensor errors and operational deficiencies:
    a) missing or unsynchronized data packages,
    b) missing or invalid GPS data,
    c) faulty driver operation;
- ensure a good performance in the process of data aggregation and analysis by pre-calculating various indicators and avoiding excessive fragmentation of aggregated data;
- retain the links between aggregated and raw data;
- allow for flexibility in data visualization, roll-up, drill-down, cross-inspection.

Our model is based on aggregating the raw data obtained from the vehicles into several types of time continuous events, allowing for several benefits:

- reducing the volume of records when analyzing the data;
- pre-calculating certain indicators for each event type;
- establishing various relations between events based on temporal intersection.

An event is generated by a succession of data packets which have something in common: the value of an indicator, a missing indicator, a certain matching between indicators and context etc. Each event is generated per vehicle and can be analyzed in relation to other events for the same vehicle, or, in more advanced scenarios, the events of other vehicles. Each event has a time interval and may or may not be interrupted by other events, depending on their types.

3.1. Event types

The events in our model are grouped in several categories:

1. Driver events:
    - driver: describes the time interval in which a vehicle is driven by a specific driver, based on the current logged in driver mentioned in the data packets;
    - missing driver: describes the time interval when data received from the vehicle does not contain the driver identifier, due to non-synchronization / non-functioning of certain equipment or due to the end / non-start of a driver's work session.
2. Events about the movement of vehicles:
    - on the move: a time interval in which there is a mileage difference between any two consecutive data packets received from the vehicle. Outside stations, this event indicates that there is no time interval greater than the standard data transmission interval without mileage difference. The time difference may be smaller inside station perimeters due to more frequent data transmission. Additionally, this type of event can be parameterized in order to allow for a minimum duration required for triggering (more than two consecutive packages);
    - stationary: a continuous time interval containing at least two data packets in which there is no mileage difference between its start and end moments.
3. Events about the route:
    - additional attributes per event: selected route – route identifier, transmitted in each raw data packet as a result of its selection by the driver;
    - missing route: a time interval in which the received data does not contain the route identifier attribute, due to the omission of route selection by the driver or equipment malfunction. For the traceability of activities, it is necessary to have one or more routes for each type of activity, even for activities that do not describe the following of a transport line: withdrawal to the depot, shunting, special runs, etc.;
    - GPS error: raw data does not contain valid GPS positions due to non-initialization / malfunction of the device. For this reason, it is not possible to check compliance with the route, as it is not possible to confirm the presence of the vehicle in its vicinity;
    - maneuver: a route is selected that does not have a predefined route, so the proximity to it cannot be checked, and the GPS positions are valid;
    - outside the route: a route with a predefined route is selected, the GPS positions are valid, but they are not located at a distance less than *x* meters (parameter) from at least one of the segments that make up the route or from one of the polygons representing the stations related to the route;
    - on route: a predefined route is selected; GPS positions are valid and are less than x meters from a route segment or station polygon.

4. Events about stations / trip details:
    - additional attributes per event:
        a) selected route: by default, the selected route is the one carried out, since the presence in stations is checked only for stations of the selected route
        b) direction: forward/return – for two-way routes; forward only – for circular routes
        c) station number: sequence number of the current station within the route
        d) station reference: link to a database record that represents the station in the context of the current route (a station can belong to several routes)
        e) passengers boarded / disembarked
        f) planned station time interval: the time interval in which the vehicle was supposed to be present at the station, according to the traffic plan
    - station: the time interval in which the vehicle is inside the polygon that describes a station on the selected route;
    - doors open: generated by a set of consecutive reports indicating doors as open; due to the risk of desynchronization of equipment at the vehicle level, the range of the open door event may be partially or completely outside the range of the station;
    - closed doors: generated by a set of consecutive reports indicating that the doors are closed;
    - passengers movement: generated by a set of consecutive packages, where each package indicates a different value for the passengers counter than the previous package;
    - between stations: the time interval in which the vehicle is between two consecutive stations. Within a trip on a route with n stations there should be *n* station events and *n*-1 inter-station events;
    - invalid station: indicates the presence of the vehicle, within a trip, in an incorrect station due to and invalid order of travel compared to the previous station. This type of events can occur either due to a real deviation from the route or it can be an anomaly generated by GPS errors, in the case of routes with a certain topology (stations within the same route in proximity, on adjacent streets or in different directions of the same street).
5. Realized trip events:
    - additional attributes per event:
        a) number of stations planned,
        b) number of stations reached,
        c) planned trip identifier – may be missing if no matching planned trip was found,
        d) planned trip interval;
    - Trip: the time interval in which the vehicle made a trip on the selected route.
6. Unrealized events

Unlike the events listed above, unrealized events are generated based on the lack of data from the vehicle or the mismatch with the traffic plan of the vehicle. These events are generated after the creation of the corresponding realized events, by checking the plan. Unlike realized events, in the time frame of unrealized events, data packets may indicate vehicle activities unrelated to them: off-route, GPS error, missing data, etc. Their purpose is to mark non-fulfillment of the plan through the same structure as the achieved data, in order to have a unified analysis and viewing approach within the graphical interface and reports. Unrealized events have the same additional attributes as their realized counterparts. They can be of the following types:

- missed station: the presence of the vehicle was not detected at all within the station in question, during a trip, in a time interval that ensures the correct order of the stations, by reference to the other validly reached stations. The missed station can be part of a completed trip, together with completed stations, or a missed trip, which contains only missed stations;
- missed trip: no station was reached within a planned trip, so the realized trip event could not be created.

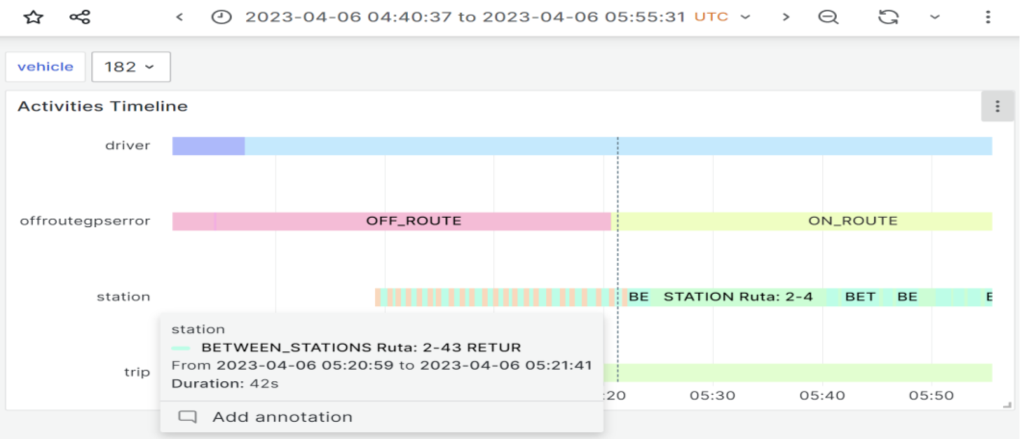
3.2. Proposed graphical user interfaces

In order to leverage the benefits of data aggregation through event detection and statistics generation, we propose the following Graphical User Interfaces (GUIs)

3.2.1. Timeline of events

The event timeline GUI includes a graphical representation of individual event durations, as well as their overlapping, for a particular vehicle. The events are grouped into multiple timelines, based on their types, with events within a timeline being non-overlappable. The GUI allows zooming in and out in order to cover the desired time interval for inspection. A tooltip pops up when hovering over an event, allowing the user to see various information related to the event, such as its time interval and duration.

The timeline from factor of this GUI allows the user to get a visual representation of events' durations and overlapping, but this can also limit the amount of information that can be viewed in a glance, as events can be too long or too short, depending on the zoom level. The next proposed GUI overcomes these shortcomings.



2023-04-06 04:40:37 to 2023-04-06 05:55:31 UTC

vehicle 182

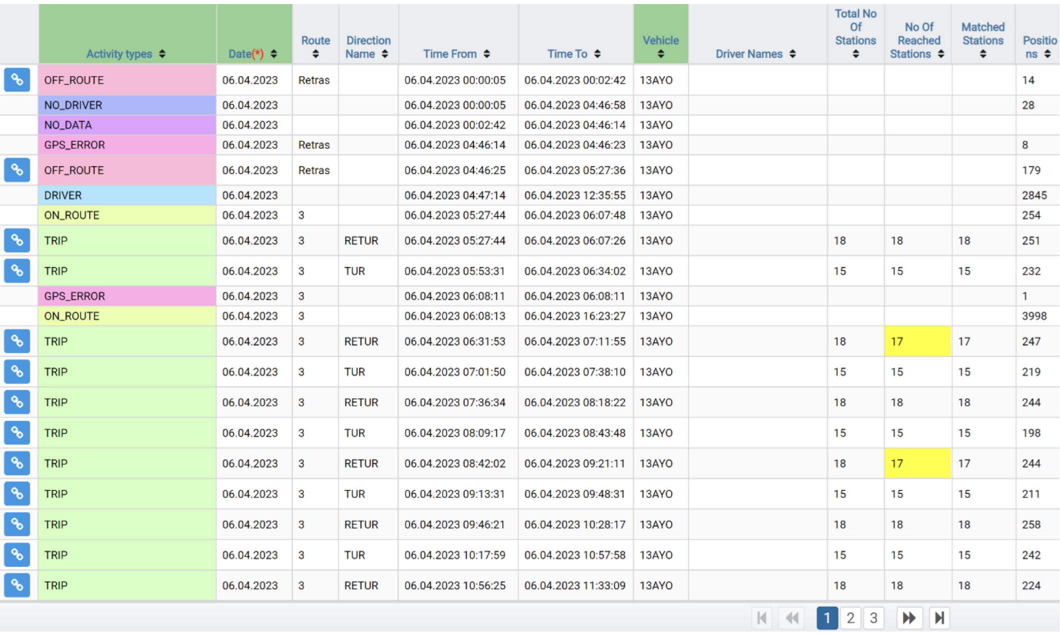
station
BETWEEN_STATIONS Ruta: 2-43 RETUR
From 2023-04-06 05:20:59 to 2023-04-06 05:21:41
Duration: 42s

Fig. 1. Visual representation of events' durations and overlapping

3.2.2. General table of events

As a general feature, all the proposed Table GUIs allow pagination, and also filtering and ordering for each column. This particular Table GUI allows listing individual detected events, being a mesoscopic data GUI. In its basic use case it allows the user to inspect the timeline of events for a certain vehicle, allowing to hide and show types of events by preference. In more advanced scenarios, the column filtering and ordering combinations make possible insights such as: top trips by duration for certain routes, total number of missed trips for a route or vehicle, top off route events by duration etc.

The table form factor of this GUI allows an equal representation of events, regardless of their duration, but it doesn't feature a graphical representation of their time spans or overlapping.



	Activity types	Date(*)	Route	Direction Name	Time From	Time To	Vehicle	Driver Names	Total No Of Stations	No Of Reached Stations	Matched Stations	Positions
🔗	OFF_ROUTE	06.04.2023	Retras		06.04.2023 00:00:05	06.04.2023 00:02:42	13AYO					14
	NO_DRIVER	06.04.2023			06.04.2023 00:00:05	06.04.2023 04:46:58	13AYO					28
	NO_DATA	06.04.2023			06.04.2023 00:02:42	06.04.2023 04:46:14	13AYO					
	GPS_ERROR	06.04.2023	Retras		06.04.2023 04:46:14	06.04.2023 04:46:23	13AYO					8
🔗	OFF_ROUTE	06.04.2023	Retras		06.04.2023 04:46:25	06.04.2023 05:27:36	13AYO					179
	DRIVER	06.04.2023			06.04.2023 04:47:14	06.04.2023 12:35:55	13AYO					2845
	ON_ROUTE	06.04.2023	3		06.04.2023 05:27:44	06.04.2023 06:07:48	13AYO					254
🔗	TRIP	06.04.2023	3	RETUR	06.04.2023 05:27:44	06.04.2023 06:07:26	13AYO		18	18	18	251
🔗	TRIP	06.04.2023	3	TUR	06.04.2023 05:53:31	06.04.2023 06:34:02	13AYO		15	15	15	232
	GPS_ERROR	06.04.2023	3		06.04.2023 06:08:11	06.04.2023 06:08:11	13AYO					1
	ON_ROUTE	06.04.2023	3		06.04.2023 06:08:13	06.04.2023 16:23:27	13AYO					3998
🔗	TRIP	06.04.2023	3	RETUR	06.04.2023 06:31:53	06.04.2023 07:11:55	13AYO		18	17	17	247
🔗	TRIP	06.04.2023	3	TUR	06.04.2023 07:01:50	06.04.2023 07:38:10	13AYO		15	15	15	219
🔗	TRIP	06.04.2023	3	RETUR	06.04.2023 07:36:34	06.04.2023 08:18:22	13AYO		18	18	18	244
🔗	TRIP	06.04.2023	3	TUR	06.04.2023 08:09:17	06.04.2023 08:43:48	13AYO		15	15	15	198
🔗	TRIP	06.04.2023	3	RETUR	06.04.2023 08:42:02	06.04.2023 09:21:11	13AYO		18	17	17	244
🔗	TRIP	06.04.2023	3	TUR	06.04.2023 09:13:31	06.04.2023 09:48:31	13AYO		15	15	15	211
🔗	TRIP	06.04.2023	3	RETUR	06.04.2023 09:46:21	06.04.2023 10:28:17	13AYO		18	18	18	258
🔗	TRIP	06.04.2023	3	TUR	06.04.2023 10:17:59	06.04.2023 10:57:58	13AYO		15	15	15	242
🔗	TRIP	06.04.2023	3	RETUR	06.04.2023 10:56:25	06.04.2023 11:33:09	13AYO		18	18	18	224

Fig. 2. Representation of events

3.2.3. Event details

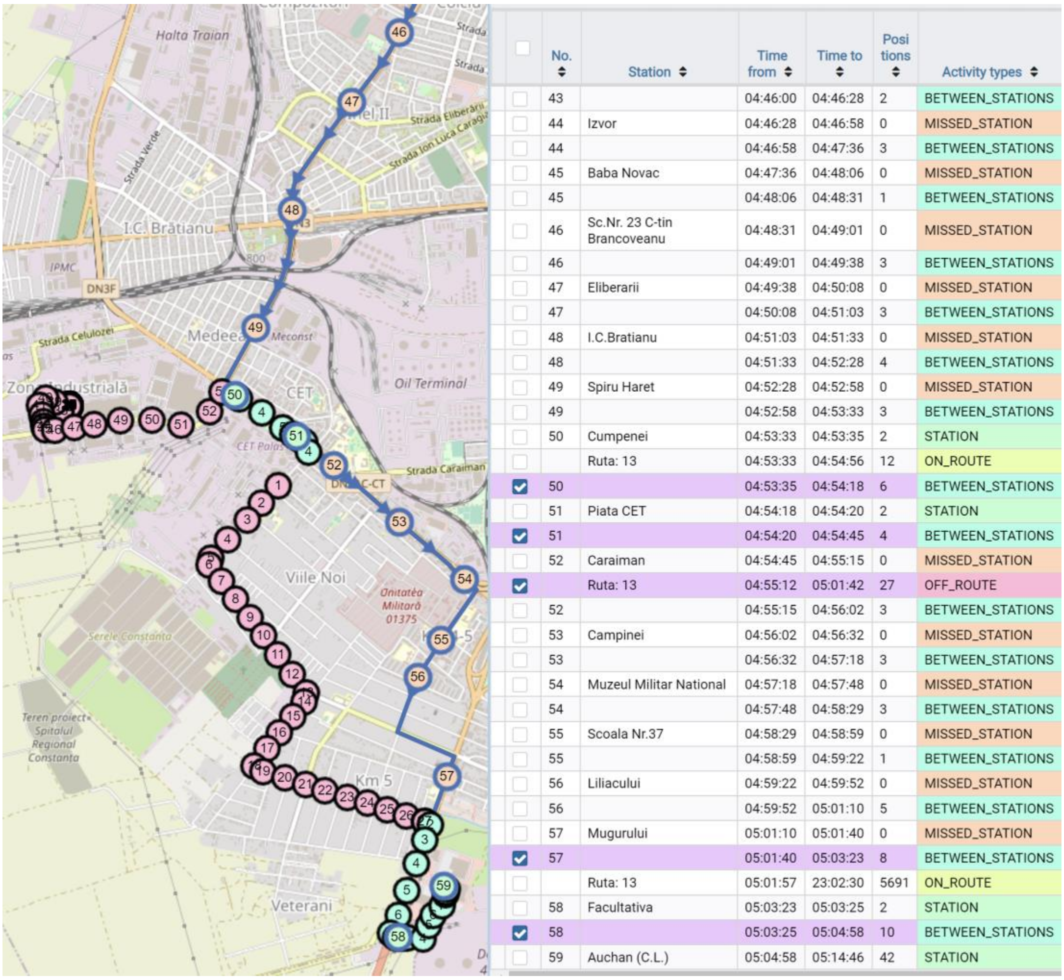
This GUI is accessible by selecting an event from the General Table of Events. It allows the visualization of several layers of information through multiple visualization types, from the perspective of the selected event:

- a list of overlapping events;
- a map layer for the overlapping events;
- an additional layer for displaying the individual data packets that were the basis for each event in the list.

This flexible interface allows the inspection of both granular and aggregated data, as shown in the screenshot below. In this featured scenario, the user was inspecting an individual trip which had an abnormal number of missed stations. On a closer inspection, it was revealed that the detected trip event was not a true trip, but it was detected automatically by the algorithm because:

- the driver left the depot with an actual trip line selected on the device;
- the vehicle entered the selected line's route at some point and continued on that route, passing by a number of stations in the correct order (50, 51);
- the vehicle exited the route after a number of stations and continued off-route until it reached the route again (stations 58, 59).

This situation is an operational fault on the driver's part, because the current task must be confirmed manually on the vehicle's external display. Even if the vehicle is not performing a trip, it must display the current task correctly (maneuver from the depot to the start of the line). From a passenger's point of view, this situation could have created confusion, especially when waiting in a station and seeing a vehicle passing by without stopping, but with the trip line shown on the display matching its current location.



☐	No.	Station	Time from	Time to	Positions	Activity types
☐	43		04:46:00	04:46:28	2	BETWEEN_STATIONS
☐	44	Izvor	04:46:28	04:46:58	0	MISSED_STATION
☐	44		04:46:58	04:47:36	3	BETWEEN_STATIONS
☐	45	Baba Novac	04:47:36	04:48:06	0	MISSED_STATION
☐	45		04:48:06	04:48:31	1	BETWEEN_STATIONS
☐	46	Sc.Nr. 23 C-tin Brancoveanu	04:48:31	04:49:01	0	MISSED_STATION
☐	46		04:49:01	04:49:38	3	BETWEEN_STATIONS
☐	47	Eliberarii	04:49:38	04:50:08	0	MISSED_STATION
☐	47		04:50:08	04:51:03	3	BETWEEN_STATIONS
☐	48	I.C.Bratianu	04:51:03	04:51:33	0	MISSED_STATION
☐	48		04:51:33	04:52:28	4	BETWEEN_STATIONS
☐	49	Spiru Haret	04:52:28	04:52:58	0	MISSED_STATION
☐	49		04:52:58	04:53:33	3	BETWEEN_STATIONS
☐	50	Cumpenei	04:53:33	04:53:35	2	STATION
☐		Ruta: 13	04:53:33	04:54:56	12	ON_ROUTE
☑	50		04:53:35	04:54:18	6	BETWEEN_STATIONS
☐	51	Piata CET	04:54:18	04:54:20	2	STATION
☑	51		04:54:20	04:54:45	4	BETWEEN_STATIONS
☐	52	Caraiman	04:54:45	04:55:15	0	MISSED_STATION
☑		Ruta: 13	04:55:12	05:01:42	27	OFF_ROUTE
☐	52		04:55:15	04:56:02	3	BETWEEN_STATIONS
☐	53	Campinei	04:56:02	04:56:32	0	MISSED_STATION
☐	53		04:56:32	04:57:18	3	BETWEEN_STATIONS
☐	54	Muzeul Militar National	04:57:18	04:57:48	0	MISSED_STATION
☐	54		04:57:48	04:58:29	3	BETWEEN_STATIONS
☐	55	Scoala Nr.37	04:58:29	04:58:59	0	MISSED_STATION
☐	55		04:58:59	04:59:22	1	BETWEEN_STATIONS
☐	56	Liliacului	04:59:22	04:59:52	0	MISSED_STATION
☐	56		04:59:52	05:01:10	5	BETWEEN_STATIONS
☐	57	Mugurului	05:01:10	05:01:40	0	MISSED_STATION
☑	57		05:01:40	05:03:23	8	BETWEEN_STATIONS
☐		Ruta: 13	05:01:57	23:02:30	5691	ON_ROUTE
☐	58	Facultativa	05:03:23	05:03:25	2	STATION
☑	58		05:03:25	05:04:58	10	BETWEEN_STATIONS
☐	59	Auchan (C.L.)	05:04:58	05:14:46	42	STATION

Fig. 3. Detailed view of an event

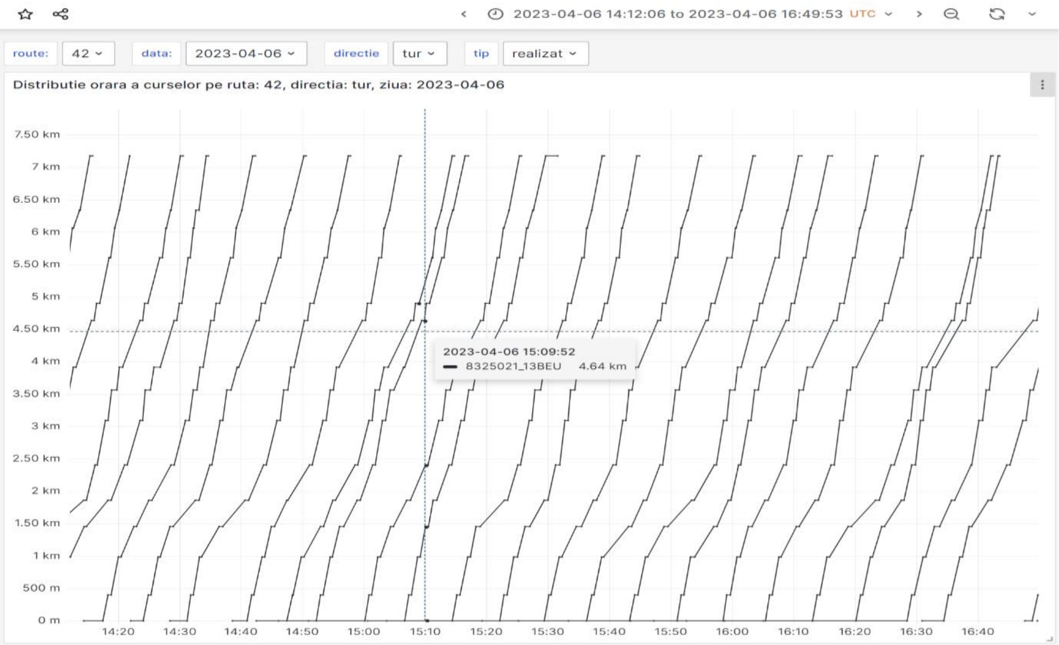
3.2.4. Trip timeline

In Figure 4, several runs by different vehicles on the same line and direction are shown simultaneously. The horizontal axis illustrates the temporal dimension (the start and end times of the constituent events of the journey – stops at stations and movements between stations). The vertical axis illustrates the spatial dimension (distance traveled within the trip).

- a steeper slope of the chart lines means a faster travel speed. Station stops are marked by horizontal lines (there is no variation on the spatial axis, only on the time axis). Illustrating the spatial dimension based on the travel distances facilitates:
    a) visualization of distance differences between stations;
    b) correct display of speeds (slopes of lines);
- at the level of each station, it is possible to view the sequence of vehicles:
    a) the horizontal line indicates the presence of a vehicle in the station;
    b) the absence of the horizontal line indicates the absence of a vehicle in the station (on the selected route).

By simultaneously illustrating several runs on the same line and direction, it is possible to visualize:

- the differences between the travel speeds for the route sections;
- the differences between the times spent in the stations;
- the approaching or distancing of some vehicles on the temporal scale: in Figure 4, the emphasis is on the approaching around km 5 of the route for the two trips that started around 14:50. In certain cases, it is even possible for a vehicle to overtake another vehicle that started its trip before it.



2023-04-06 14:12:06 to 2023-04-06 16:49:53 UTC

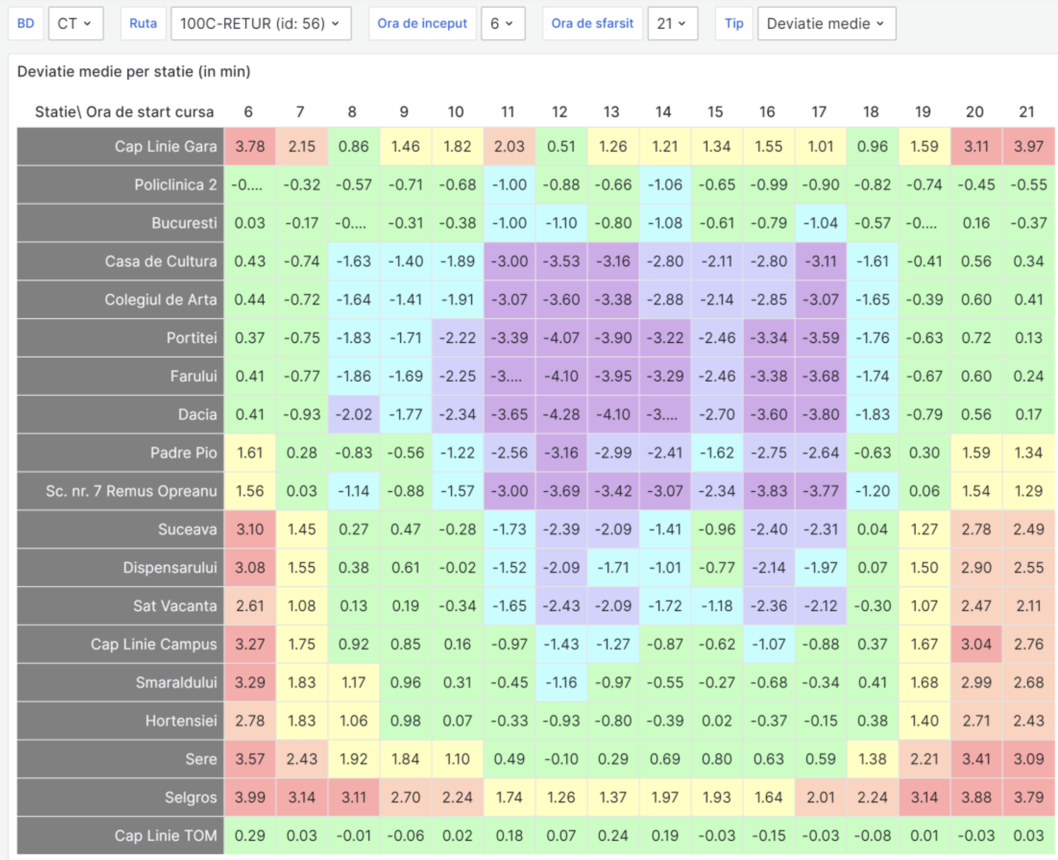
route: 42 data: 2023-04-06 directie tur tip realizat

Distributie orara a curselor pe ruta: 42, directia: tur, ziua: 2023-04-06

Fig.4. The trips timeline

3.2.5. Line statistics heatmap

This GUI allows for a condensed visualization of statistics for a selected line. The data is grouped on the columns by time intervals (hourly in the displayed case) and on rows by the constituent stations of the route. The user can select several statistics from the rightmost dropdown (time between stations as absolute real value, planned value, real value as percentage of the planned value etc.). In Figure 5, the accumulated delays relative to the planned times are shown in minutes. The color palette illustrates advances relative to the plan by warm colors, values within an acceptable margin of error in green (+/- 1 minute) and delays from the plan by cold colors. In the scenario depicted below, one can notice rushed starts of the trips for almost all hourly intervals, and also rushed reachings of the stations in the second half of the route for the intervals 6-7 and 19-21. Surprisingly, the final stations of the route are reached at the planned time, which might indicate anomalous driver behavior. The mid-section of the chart shows a generalized delay corresponding to the first half of the route, across several time intervals centered on the afternoon. This can be explained by the lack of correlation between the trip plan and the real traffic conditions in the affected time intervals and city areas. This type of visualization allows the transit operator to assess the level of plan suitability and the underlying data can be used in the planning calibration.



BD CT

Ruta 100C-RETUR (id: 56)

Ora de inceput 6

Ora de sfarsit 21

Tip Deviatie medie

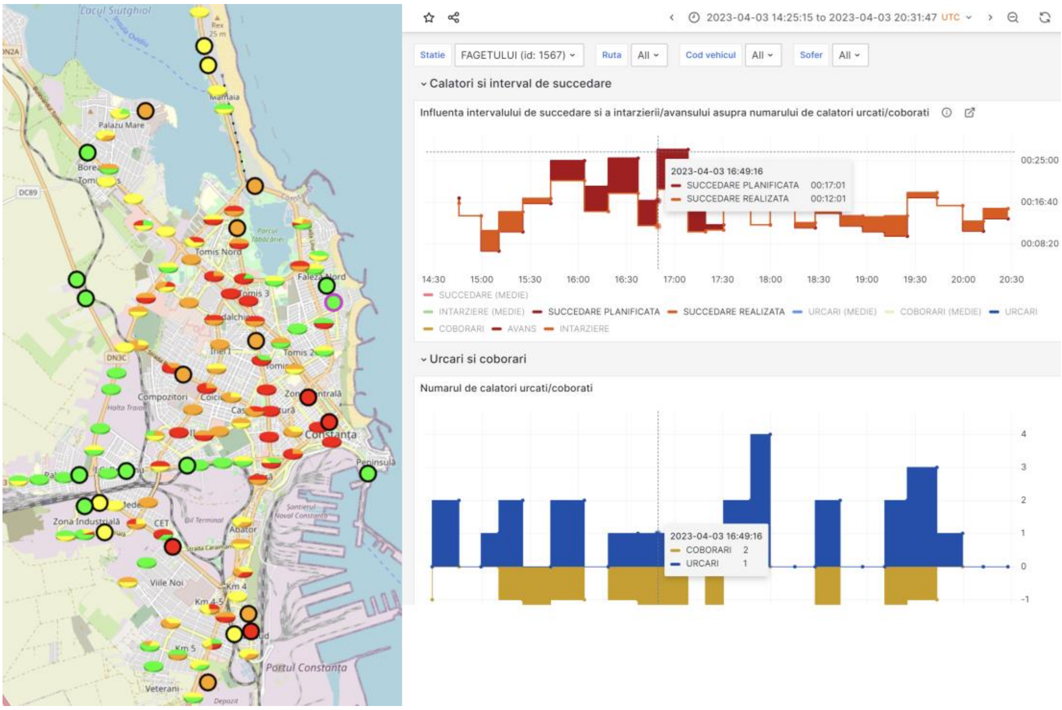
Deviatie medie per statie (in min)

Statie\ Ora de start cursa	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Cap Linie Gara	3.78	2.15	0.86	1.46	1.82	2.03	0.51	1.26	1.21	1.34	1.55	1.01	0.96	1.59	3.11	3.97
Policlinica 2	-0....	-0.32	-0.57	-0.71	-0.68	-1.00	-0.88	-0.66	-1.06	-0.65	-0.99	-0.90	-0.82	-0.74	-0.45	-0.55
Bucuresti	0.03	-0.17	-0....	-0.31	-0.38	-1.00	-1.10	-0.80	-1.08	-0.61	-0.79	-1.04	-0.57	-0....	0.16	-0.37
Casa de Cultura	0.43	-0.74	-1.63	-1.40	-1.89	-3.00	-3.53	-3.16	-2.80	-2.11	-2.80	-3.11	-1.61	-0.41	0.56	0.34
Colegiul de Arta	0.44	-0.72	-1.64	-1.41	-1.91	-3.07	-3.60	-3.38	-2.88	-2.14	-2.85	-3.07	-1.65	-0.39	0.60	0.41
Portitei	0.37	-0.75	-1.83	-1.71	-2.22	-3.39	-4.07	-3.90	-3.22	-2.46	-3.34	-3.59	-1.76	-0.63	0.72	0.13
Farului	0.41	-0.77	-1.86	-1.69	-2.25	-3....	-4.10	-3.95	-3.29	-2.46	-3.38	-3.68	-1.74	-0.67	0.60	0.24
Dacia	0.41	-0.93	-2.02	-1.77	-2.34	-3.65	-4.28	-4.10	-3....	-2.70	-3.60	-3.80	-1.83	-0.79	0.56	0.17
Padre Pio	1.61	0.28	-0.83	-0.56	-1.22	-2.56	-3.16	-2.99	-2.41	-1.62	-2.75	-2.64	-0.63	0.30	1.59	1.34
Sc. nr. 7 Remus Opreanu	1.56	0.03	-1.14	-0.88	-1.57	-3.00	-3.69	-3.42	-3.07	-2.34	-3.83	-3.77	-1.20	0.06	1.54	1.29
Suceava	3.10	1.45	0.27	0.47	-0.28	-1.73	-2.39	-2.09	-1.41	-0.96	-2.40	-2.31	0.04	1.27	2.78	2.49
Dispensarului	3.08	1.55	0.38	0.61	-0.02	-1.52	-2.09	-1.71	-1.01	-0.77	-2.14	-1.97	0.07	1.50	2.90	2.55
Sat Vacanta	2.61	1.08	0.13	0.19	-0.34	-1.65	-2.43	-2.09	-1.72	-1.18	-2.36	-2.12	-0.30	1.07	2.47	2.11
Cap Linie Campus	3.27	1.75	0.92	0.85	0.16	-0.97	-1.43	-1.27	-0.87	-0.62	-1.07	-0.88	0.37	1.67	3.04	2.76
Smaraldului	3.29	1.83	1.17	0.96	0.31	-0.45	-1.16	-0.97	-0.55	-0.27	-0.68	-0.34	0.41	1.68	2.99	2.68
Hortensiei	2.78	1.83	1.06	0.98	0.07	-0.33	-0.93	-0.80	-0.39	0.02	-0.37	-0.15	0.38	1.40	2.71	2.43
Sere	3.57	2.43	1.92	1.84	1.10	0.49	-0.10	0.29	0.69	0.80	0.63	0.59	1.38	2.21	3.41	3.09
Selgros	3.99	3.14	3.11	2.70	2.24	1.74	1.26	1.37	1.97	1.93	1.64	2.01	2.24	3.14	3.88	3.79
Cap Linie TOM	0.29	0.03	-0.01	-0.06	0.02	0.18	0.07	0.24	0.19	-0.03	-0.15	-0.03	-0.08	0.01	-0.03	0.03

Fig. 5. Visualization of statistics for a selected bus line

3.2.6. Passenger volumes analysis

This GUI offers both an overview and a detailed analysis of passenger counts. On the left side, the map highlights the relative passenger volumes for all the stations on the network: the stations are colored according to the counter value quartiles, with the green stations having within 25% of the maximum passenger volumes. The map uses clustering in order to properly represent stations that are too close to each other. On the right side, there are two synchronized timelines, which display data for a single station selected from the map. The top timeline illustrates the deviations from the planned stop times, and the bottom one shows the passenger counter values for each individual stop within the station. This synchronized view allows the user to assess the influence that the deviation from the plan has over the passenger volumes. The passenger volume statistics can be used to reevaluate the transport planning on a citywide scale.



2023-04-03 14:25:15 to 2023-04-03 20:31:47 UTC

Statie: FAGETULUI (id: 1567)
Ruta: All
Cod vehicul: All
Sofer: All

Calatori si interval de succedare

Influenta intervalului de succedare si a intarzierii/avansului asupra numarului de calatori urcati/coborati

2023-04-03 16:49:16
SUCCEDARE PLANIFICATA 00:17:01
SUCCEDARE REALIZATA 00:12:01

Urcari si coborari

Numarul de calatori urcati/coborati

2023-04-03 16:49:16
COBORARI 2
URCARI 1

Fig. 6. Passenger volumes analysis

4. Conclusions

The improvement of public transport is a continuous concern of the public administrations of modern cities. Public transport systems can lead to traffic relief, reduced costs and environmental protection. This has enabled the development of intelligent transport systems (ITS) for public transit. Despite all the progress made, most public transport systems are missing the opportunities of using analytics on the operational monitoring data in a feedback loop for the improvement of transportation plan. Our proposed event-based approach offers a solid, standardized foundation for the development of data analytics that can aid in improving the performance of transportation systems based on real-world feedback. The proposed model helps public transit operators in monitoring the system's performance, identifying and analyzing anomalies at multiple levels of detail, by leveraging the benefits of data aggregation and suggestive graphical visualizations. The event-based model also serves as a basis for the continuous generation of statistics which can facilitate the improvement of the public transit planning.

Acknowledgements

This research was supported by the project AUDITORIUM, POC 221, code 143690.

References

[1] K. Shaaban, M. Elaminb and M. Alsoubb, “Intelligent Transportation Systems in a Developing Country: Benefits and Challenges of Implementation,” pp. pp. 1373-1380..

[2] J. So, K. M. T. Kim and S. Son, “Approach to Prioritizing Intelligent Transport Systems (ITS) Services in Developing Countries,” *The Mongolia Case, Transport and Communications Bulletin for Asia and the Pacific,* pp. 32-42, 2018.

[3] S. Taie and A. Elazb, “Challenge of Intelligent Transport System,” *International Of Modem Engineering Research UMBR,* vol. 6, no. ISSN: 2249-6645., 2016.

[4] A. Auer, S. Feese, S. Lockwood and A. Vann Easton, “History of Intelligent Transportation Systems,” 2023. [Online]. Available: www.its.dot.gov/index.htm.

[5] Q. Ali, N. Ahmad, A. Malik, G. Ali and W. Rehman, “Issues, Challenges, and Research Opportunities in Intelligent Transport System for Security and Privacy, Applied Sciences,,” 2018.

[6] M. Sami and K. Sara, “Intelligent Transportation Systems for Sustainable Urban Environments,” *International Journal of Advanced, Natural Sciences and Engineering Researches,* Vols. vol. 7, no. 9, pp. pp. 166-177.

[7] K. Ahmad, H. Khujamatov, A. Lazarev, N. Usmanova, M. Alduailij and M. Alduailij, “Internet of Things-Aided Intelligent Transport Systems in Smart Cities,” *Challenges, Opportunities, and Future, Wireless Communications and Mobile Computing,* vol. vol. 2023, 2023.

[8] G. Pauer, „ Development Potentials And Strategic Objectives Of Intelligent Transport Systems Improving Road, Transport and Telecommunication,” Vols. vol. 18, no. 1, 2017, pp. pp. 15-24.

[9] A. Naz and I. Hoque, “Integration of Intelligent Transportation Systems (ITS) with Conventional Traffic Management in Developing Countries,” [Online]. Available: https://engrxiv.org/preprint/view/3154.

[10] D. Kinane, F. Schnitzler, S. Mannor and e. al, Intelligent Synthesis an Real-time Response using Massive Streaming of Heterogeneous Data (INSIGHT) and its anticipated effect on Intelligent Transport System (ITS) in Dublin City, Ireland,, Helsinki.: Proceedings of the 10th ITS European Congress, 2014.

[11] A. Putra, H. Warnars, F. Gaol, B. Soewito and E. Abdurachman, “A Proposed surveillance model in an Intelligent Transportation System (ITS,” Jakarta, Indonesia:, Indonesian Association for Pattern Recognition International Conference (INAPR), TEEE, 2018, pp. pp. 156-160.

[12] D. Hahn, Munir, A. and V. Behzadan, “Security and Privacy Issues in Intelligent Transportation Systems,” *Classification and Challenges, IEEE Intelligent Transportation Systems Magazine,* vol. 13, pp. 181-196, 2021.

[13] N. Audeechya, N. Shah and D. Jain, “Design, Technologies & Challenges in Intelligent Transportation System,” *International Journal of Engineering Research & Technology (LERT),NCETECE 14 Conference Proceedings,* no. ISSN: 2278-0181..