On Joint Satellite, Terrestrial, and Delay Tolerant Networks for Railroad Communications

Eugene Tikhonov, Donat Schneps-Schneppe, Manfred Sneps-Sneppe
Ventspils International Radio-astronomy Centre, Ventspils University of Applied Science
Ventspils, Latvia
eugene.tikhonov, donat.shneps-shneppe @abava.net, manfredss@venta.lv

Abstract—**The paper explores data delivery delays in next-generation railway communication networks with various configurations of terrestrial and satellite coverage (including mesh and delay-tolerant forwarding). It proposes a rough analytic estimation (with just two parameters: the total number of trains and the average number of their meetings) and computer simulation (based on realistic train schedules). This approach can serve as a preliminary groundwork for the technical and commercial choice of a preferred 5G railway link type on the stage of implementing: satellite (SatCom) or terrestrial (TerraNet), and assessing the feasibility of the necessary QoS.**

I. INTRODUCTION

Railway communication systems (like other areas) rapidly transform due to the increase in data volume and the development of technologies for the accumulation, storage, processing, and transmission of information. This requires a migration from the existing railway communication system (GSM-R) to the new proposed generations: to the fourth (LTE-R), or immediately to the fifth (5G-R) [1]. The same demand is for customer data transmission: passengers need more coverage of the railway path with a high-speed network. This transition leads to increased competitiveness and new opportunities and services (better forecasts and decisions, fewer errors and costs due to automation, supply chain connectivity, customer availability, IoT and cloud computing, BigData-based services) [2].

Worldwide professional association International Union of Railways (UIC) with the involvement of market participants has developed a project of a possible successor to GSM-R called Future Railway Mobile Communication System (FRMCS) [3]. As one of the components, it considers 5G networks, which are an ecosystem of various technologies. In particular, it is supposed to jointly use both terrestrial networks (with connection to infrastructure stationary base stations, BS) and satellite networks (satellite mobile terminals on trains, ST). Fixed terrestrial Base Stations connections to the outer network can be wired, wireless, or via satellites. This applies to both corporate internal service networks and subscriber networks.

The optimal choice of the type of communication terminals (or their combination) may differ in various geographic locations and is usually assessed by the price/quality ratio. The price depends on the cost of the necessary equipment, overhead work (installation and infrastructure), and maintenance. The corresponding mathematical function defines the expenses dependence on the number of ST and BS.

Quality level is not a simple measurable parameter. The planned applications define network architecture and characteristics. Various quantitative indicators are suitable for the assessment of these characteristics. It is necessary to evaluate them for terrestrial, satellite solutions, or their combination to make a comparison. In other words, to find this quality metric dependence on the number of BS and ST: analytical function or numerically. These indicators (expected characteristics) can also act as an assessment of the feasibility of the QoS (Quality of service) required for the application.

This paper proposes both analytic and simulation methods for such an assessment.Considered simplified cases do not take into account a lot of real parameters: limited volumes of transmitted and stored data, possible communication errors, complex shape of the coverage area and different signal strength in it, deviations from the planned schedules, and so on. But for these simplified cases, very simple theoretical expressions for various types of networks are obtained (including mesh retransmissions and DTN connections), requiring a minimum amount of statistical information. The results are well confirmed by computer simulation, which takes into account more parameters, for example train stops at intermediate points and a given communication range. These assessments may work as quick pre-selection of planned infrastructure deployment, feasibility of QoS, and a basis for more sophisticated models.

The paper is organized as follows. In Section II some related works are considered. Section III compares the effectiveness of increasing network coverage on the railway by increasing the number of base stations (of the terrestrial radio link network) and the number of satellite terminals (on trains). Section IV does the same for Delay Tolerant Networks. Section V presents a summary of all types of networks considered. Section VI compares the number of terrestrial and satellite equipment to ensure equal performance.

II. BACKGROUND AND RELATED WORKS

The idea of a promising high-speed railway (HSR) mobile communication system LTE-R was expressed as an adaptation of LTE capabilities to the peculiarities of railway conditions [4] as a descendant of the existing GSM-R. A comparison of the key differences between the existing GSM-R system with LTE and the proposed railway modifications LTE-R was made [5]. In 2017 a pilot project of such network was launched on the Wonju-Gangneung high-speed train line with a length of 120 km [6].

Further improvements have also been proposed. For example, the Mobile hotspot network enhancement system (with carrier aggregation and better handover, MIMO, physical layer optimizations) [7], or alternative network architectures, for example, NDN-GSM-R (Named Data Network GSM-R) [8] based on the data contents and functions instead of nodes addressing for highly dynamic network topology.

However, at the moment, it may be advisable to go to the gradually introducing fifth-generation networks directly. The proposed technical characteristics and requirements for critical data (control, signaling, monitoring and maintenance) or non-critical information (user applications) were formulated [9]. A large-scale comprehensive analysis of 5G key technologies for smart railways was carried out in [10] (review of a huge number of ideas, proposed architectures, channel models, key technologies, and pending future research).

It is also an obvious idea to use satellite communications on trains in sparsely populated rural areas, where terrestrial ground equipment may not be cost-effective. Such studies have been carried out for a long time (for example, [11] – research, experiments, and open questions on satellite railway communications), but they remain relevant today (for example, a review of commercial prospects and a proposal for own tested equipment system is in [12]).

The railway communication has specifics. That needs to be taken into account in any practical communication system, especially at high frequencies and speeds. Some works are devoted to its study. In [13] observation of various scenarios of radio wave propagation in real conditions was made along the Wuhan-Guangzhou high-speed railway and at stations (for the frequency 930 MHz). Several fundamental scenarios were identified: viaducts (trees can be above or below the tracks on the viaduct), cuttings (in hills and ground irregularities), bridges, tunnels, stations (indoor, outdoor, sort depots), and their combinations. Some particular problems of the Physical and Media Access Control OSI-layers and possible solutions were discussed, for example, in [14] (large ping and Doppler shift affect the Random Access request and initial synchronization) and in [15] (signal losses, delays, Doppler shift for eMBB and IoT – Mobile BroadBand procedures and Narrow Band-Internet of Things). Also, the selection of potential waveforms for the 5G railway was discussed in [16] (with an emphasis on various Filter Bank Multi-Carrier modifications).

Traditional cellular terrestrial networks were completely independent of satellite communications. The 5G-vision changes this approach. According to the European Commission's SaT5G project Horizon 2020 (H2020), 5G goals are an increase in capacity by 1000 times, an increase in data rate by 10-100 times, maximum coverage, service creation in minutes, end-to-end latency less than 1 ms. Since all of them are unlikely to be achievable at the same time, then this generation represents an ecosystem of compatible networks based on various technologies [17]. The key conclusion of the ESPI-ESA conference "Space and SATCOM for 5G" (June 2017) is that SatCom (Satellite Communications) should be integrated into a hybrid communication system and 5G. And it should be normal for the user not to know about which technologies provide him with continuous communication seamlessly – the principle of “connected mobility” [18]. Many advantages make satellite communication usage logical on a significant part of the train path, and somewhere it is the only possible solution. The main ones are coverage of a large area, broadband connection, lack of infrastructure requirements along the way, scalability, and potential integration with terrestrial networks (for example, in tunnels) [11].

A wide review of the features of integrated space, air, and terrestrial 5G networks is carried out in [19] (trends, problems and key technologies, standardization and industry initiatives, research, and open questions). The report of the “Satellite Working Group” of the European technology platform NetWorld-2020 [20] highlights the expected functions of the satellite component of 5G networks: airplanes and high-speed trains, security services, and V2V (Vehicle-to-Vehicle), communication in natural or man-made disasters. The vision of the H2020 SaT5G project [21] is to use SatCom with TerraNet (Terrestrial Network) jointly to speed up 5G deployment and take advantage of new opportunities. A description of the proven practical solutions for railway, satellite, and integrated communications is given in [22] (primarily 3INSAT and ACS, an adaptable communication system considered in the Shift2Rail project - X2Rail-1 as a GSM-R descendant).

An overview [23] of current initiatives and projects, studies, functions of the main stakeholders (ESA, EU NetWorld-2020, 5GPPP, 3GPP, CCSA, Group 4B in ITU-R, IEEE, EU H2020 programs) also proposes a concept of integrating network system based on SDN and NFV (Software Defined Network and Network Function Virtualization), with adaptive balancing threshold of a random choice of available multipath transmission alternatives. Of course, this selection algorithm is relevant only for the area with both terrestrial and satellite connections.

A similar study of the choice of the communication method (TerraNet or SatCom with their specified delays and data rates) is considered in [24]. This study is limited with consideration of receiving variable-bitrate video by passengers on the train. A set of QoS types is specified, each one with the corresponding maximum allowable delay. Several approaches to balance random threshold selection are used: Queuing theory (analytical evaluation), Neural Network regression over the total system, Neural networks modeling divided system. Approaches that use neural networks have some advantages in this simulation. The paper also suggests the use of already deployed networks of both types.

III. 5G SATCOM/TERRANET COMPARISON

A. Parameters and data types for comparison

The following indicators of network efficiency are suggested:

- Physical link characteristics.

- Online time percentage.
- Data delivery delay.

Various network architectures and types of data are possible:

1) Direct connection to BS/ST only – transmission is possible only between the mobile terminal and the receiver in real-time (Real-time Data).

2) Delay/Disruption Tolerant Data transfer – data allows intermediate storage and delivery when possible. Data is only stored on the train-generator.

3) Mesh-connections (mobile Base Stations, MBS) – mobile terminals can act as repeaters for other terminals to communicate with BS or ST [25].

4) Delay Tolerant Network (DTN) routing – data can be stored and delivered by intermediate train-carriers besides self-delivery (an algorithm comparable to Epidemic is assumed) [26].

The impact on the network efficiency indicators of SatCom and fixed TerraNet differs significantly in these different scenarios.

The assessment of that impact is carried out in an approximate analytical form and using computer simulation based on the close-to-real schedules.

B. Physical link characteristics

Basic physical parameters of a wireless network characterize the existing considered connection. These are, first of all: data rate (about Gb/s for 5G), spectral efficiency (depends on coding, modulations, Media Access Control – MAC technologies and so on), coverage and cell size (it is assumed to be about hundreds of meters for terrestrial 5G), carrier frequency and available bandwidth, error correction ratio, transmitted power, limits of mobility (speed limits up to 500 km/h for high-speed trains), and latency (delay in the transmission medium and time for systematic processing, about 250 ms for GEO, about 10s ms for LEO and between these values for MEO satellites). Many of these parameters cannot be maximized jointly at the same time. There is currently no approved standard for railway communication of generations newer than GSM-R. Networks with different combinations of characteristics and various architectures can be considered. The parameters of the network should be balanced for the various specific implementations considered for a specific scenario and based on the intended applications and the associated QoS requirements.

There may be areas on the railway where trains can be not connected to the network as required for some application parameters. For example, when a new faster network is only in the deployment stage, or when such a network with complete coverage is economically unsuitable. In these “offline” zones data delivery delay associated with the network topology occurs. This is studied in more detail further.

C. Considered scenario

An offline section of a railway (a section without high-speed connection of demanded data rate yet) with a length L is considered for the analysis. A stable high-speed communication is assumed at the endpoints. Any railway can be divided into parts: online zones and such offline sections.

For the expanding of the network coverage, some amount of stationary 5G-like Base Stations (BS) can be installed in this offline section. They are considered evenly spaced along the path. These stations are supposed to provide high-rate communication without latency to trains. These BSs allow connection for the trains in a fixed range with radius R.

Another or additional solution is to equip some trains with Satellite Terminals (ST). It was assumed that ST provides guaranteed high-rate connection without latency for those trains on which they are installed.

Mobile terminals installed on trains also communicate with each other directly by implementing mesh-connections [25] and DTN routing [26, 27]. A train can act as an intermediate repeater for close trains if it is connected to the BS (directly or through a chain of the subsequent repeaters). If the train is equipped with a satellite terminal (ST), it can provide access to the outer network – also directly or via a repeaters chain.

The data rate in this simplified scenario is assumed to be the same for all transmissions: terrestrial, satellite, and between trains; and also large enough to transfer all available (and stored) data.

For a theoretical assessment trains were considered as point objects moving at a constant speed V (for all trains at any time), there are no stops. The range of the train-to-train connection was not taken into account, only the fact of connection. Mesh-connection chains were ignored.

For computer simulation, trains were considered as moving points too. But real railway schedules were used with intermediate stops (arrival time, departure time). Trains have a corresponding constant speed between them. Train schedules are cyclical, identical for each day, that corresponds to the real schedules under consideration. That is, trains depart on any day at the same time with the same schedule. Therefore, some trains are already on the railroad at the start of the simulation in the corresponding positions. The train-to-train connection distance was assumed to be R (same as the radius between BS and train). To average the estimates of SatCom, it was assumed that STs are randomly distributed among the trains under consideration and several simulation cycles with different ST-carriers were carried out. Each simulation is carried out with a fixed time step (1 second) during the day (24 h).

A comparison was made for several different real railway schedules in different Russian and European locations, but this paper provides only one typical example: the Surgut-Tyumen railroad (see Table I).

TABLE I. PARAMETERS OF THE SIMULATED SCENARIO

Parameters	Surgut-Tyumen
Number of trains, units	22
Number of meetings, events	174
Average meetings per train, events	7,9
Average speed, km/h	47,57
Length of railway section, km	704
Communication radius R, km	0,30
Cycles with random choice of ST trains	10

D. Online time percentage

To compare SatCom and TerraNet, the indicator "the all-trains average percentage of time online" can be used, depending on the number of installed Base Stations and Satellite Terminals (BSs and STs).

The online time portion in TerraNet corresponds to the time that trains spend inside the coverage areas of base stations (number of stations is N_{BS}). For constant speed (meaning equal travel time L/V) for each train and all of them (see also Fig. 1):

$$online_{BS} = N_{BS}\frac{2R}{L} \tag{1}$$

For complete coverage of the railway section, at least $N_0 = \frac{L}{2R}$ stations are required. This number can be very significant for a small connection distance R (relative to the length of the section L) and reach thousands of units. For example, the Surgut-Tyumen railroad needs not less than 2347 stations (with selected parameters) for only 22 trains.

Online time portion in SatCom corresponds to the percentage of the number of trains equipped with ST (N_{ST}) from the total number of trains N_{trains} (see also Fig. 1):

$$online_{ST} = \frac{N_{ST}}{N_{trains}} \tag{2}$$

Only these trains will be connected to the outer network, the rest will always be offline. For full coverage, all N_{trains} with satellite terminals are required.

Joint operation of BSs and STs has an average online time portion:

$$online(N_{BS}, N_{ST}) = \frac{N_{ST}}{N_{trains}} + N_{BS}\left(1 - \frac{N_{ST}}{N_{trains}}\right)\frac{2R}{L} \tag{3}$$

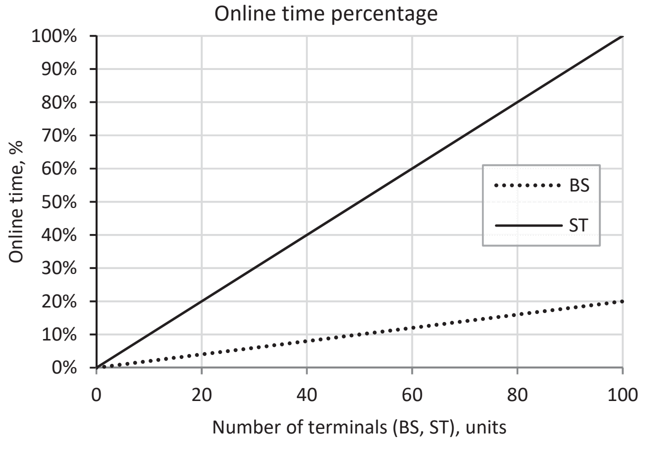


Fig. 1. Time online percentage for TerraNet and Satcom only. Example for 100 trains on a 500 km section with a communication range R = 500 m.

With a small number of trains and a relatively short communication range, the advantage is on the side of SatCom: a lot of base stations (BSs) are needed to provide comparable coverage. However, in this case, communication will be available only for those trains that have a satellite terminal (permanently), others will not have a connection at all on the section. In TerraNet, all trains will be online, but only part of the time. Also, the data that can be transmitted during communication with each BS is limited. The volume is proportional to the connection time (BS range divided by the train speed), and the data rate.

E. Delay/disruption Tolerant Data transfer

Some data types require immediate delivery (online-connection) and cannot be delayed. For example, telephone conversation, real-time broadcasting, some types of monitoring telemetry used for control.

However, a huge amount of information does not require this. This is the transfer of various types of observational telemetry and IoT data, e-mail, messenger messages, and much more. Such data can be stored in the memory of devices while offline to be transferred when there is a connection.

It is possible to estimate the average expected delay for this type of data, assuming that the data (or messages) is generated at a random time (the distribution of which is assumed to be uniform, so the messages (data) represent the Poisson point process) and is equally probable on all considered trains:

$$\langle D \rangle = \frac{1}{N_{trains}} \sum_{n=1}^{N_{trains}} \langle d_n \rangle \tag{4}$$

Where $\langle d_n \rangle$ – average delivery delay for train number n.

If N_{BS} base stations are installed evenly on the section, then they divide it into $(N_{BS}+1)$ correspondingly smaller offline subsections. Neglecting the short communication range of the station itself (with zero delays inside), the average delay for each train will decrease by $\frac{1}{1+N_{BS}}$ times. The average data delivery delay for all trains in TerraNet can be estimated as:

$$\langle D \rangle_{BS} = \frac{1}{N_{trains}} \sum_{n=1}^{N_{trains}} \frac{1}{1+N_{BS}} \langle d_n \rangle = \underbrace{\left(\frac{1}{1+N_{BS}}\right)}_{BS(N_{BS})} \times \langle D \rangle \tag{5}$$

It was assumed that the transmission rate is high enough to transmit all the accumulated data to/from each BS. In practice, the volume is limited by the connection time (antenna range and train speed) and the limited data rate. This can be taken into account by appropriate amendments if a lot of data is generated in future works.

It is possible to estimate the expected delay of Delay Tolerant Data for trains with SatCom also. A train equipped with ST always has zero delays for all data. Not equipped – the "default" average $\langle d_n \rangle$. An arbitrarily chosen train number n with probability $p_n = \frac{N_{ST}}{N_{trains}}$ is equipped with a satellite terminal, and with probability $(1-p_n)$ it is not. Therefore, the average delay with averaging over all trains and all possible random ST assignments:

$$\langle D \rangle_{ST} = \frac{1}{N} \sum_{n=1}^{N} (1-p_n)\langle d_n \rangle = \underbrace{\left(1 - \frac{N_{ST}}{N_{trains}}\right)}_{ST(N_{ST})} \langle D \rangle \tag{6}$$

Joint operation of BSs and STs combines the equations (5) and (6):

$$\langle D\rangle_{BS+ST}(N_{BS},N_{ST}) = ST(N_{ST})\times BS(N_{BS})\times\langle D\rangle \qquad (7)$$

This is very well confirmed by simulations (see Fig. 2):

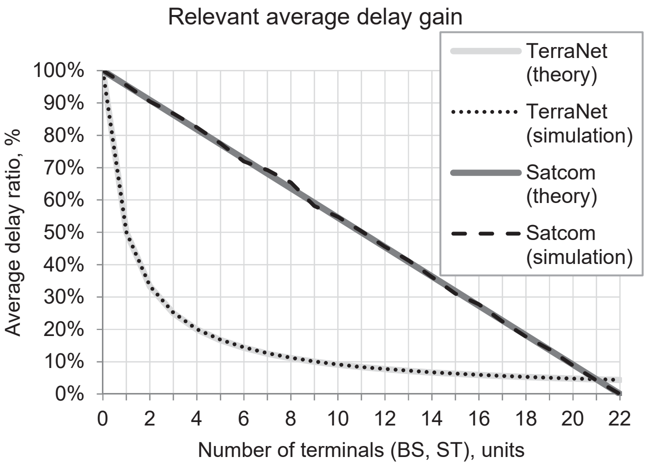


Fig. 2. Relevant delay (percentage to average delay without intermediate BSs) for Tyumen-Surgut schedule (22 trains, 704 km route, communication range 300m). Rank correlation coefficients R^2 = 0,999995 for TerraNet and 0,99986 for SatCom.

F. Mesh-connections (mobile Base Stations)

If the train is offline, but within the connection range of another train that has a connection to the outer network (directly to the BS/ST or through a chain of further repeaters), then it can transmit data through this repeater. Besides, it can become the next link in the chain of the forwarders for even more distant trains.

As shown in [25], for a large connection range of terminals and a significant number of trains, such propagation of communication through mesh-forwarding can be very significant, dynamically “expanding” the coverage range of the base station.

With a short connection range, typical for high-speed connections, for example, 5G, this effect will be negligible and can be ignored in theoretical estimates for BS. However, this capability is important for SatCom and Delay Tolerant Data types. When a train with ST is met, it is possible to send previously generated and stored data through it immediately. Thus, data will be delivered much earlier than at the time of arriving at the endpoint. A train with a satellite terminal acts as a mobile base station – MBS.

The railway schedule dictates the average number of meetings for each train N_{meets} (as the ratio of the total number of meetings to the number of trains). To estimate the average delay, we can assume that encountered trains may be equipped with a satellite terminal with equal probability. Therefore, the number of encountered mobile stations is on average $N_{mBS} = N_{meets}\times\frac{N_{ST}}{N_{trains}}$.

Although the meetings occur in random places on the section, it is possible to estimate roughly their average impact similarly to stationary base stations BS (taking into account the previously considered decrease in the delay from the STs themselves):

$$\langle D\rangle_{BS+ST+mBS}(N_{ST},N_{ST}) = BS(N_{mBS})\times\langle D\rangle_{BS+ST} \qquad (8)$$

$$= \underbrace{\left(\frac{1}{1+N_{meets}\times\frac{N_{ST}}{N_{trains}}}\right)}_{mBS(N_{ST})}\times\langle D\rangle_{BS+ST}$$

This assumption agrees well with the simulation results (shown in Fig. 3):

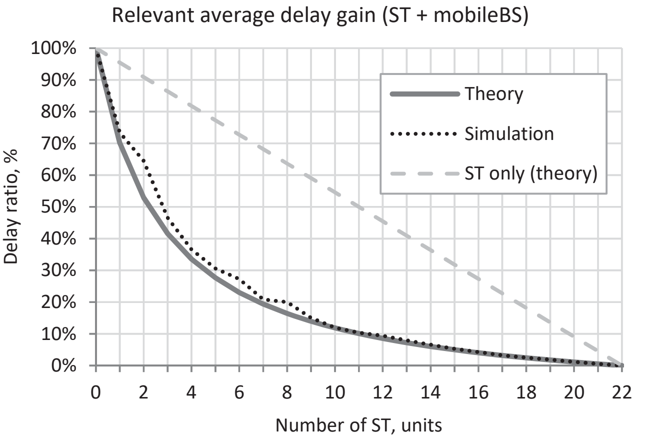


Fig. 3. Relevant delay (percentage to average delay without ST) for Tyumen-Surgut schedule (22 trains, 704 km route, 174 meetings – 7,9 meeting per each train, 10 cycles of randomly assigned ST for each number) without BS. Rank correlation coefficient R^2 = 0,995.

The possibility of using the STs as Mobile Base Stations significantly affects the decrease in the average delay of Delay Tolerant Data, enhancing the efficiency of the network.

It was also assumed that the transmission rate is high enough to transmit all the accumulated data. In practice, the volume is limited by the connection time and the limited data rate. In the case of MBS, the connection time can be estimated as the antennas range divided by the summarized speed of the trains (most of the meetings are with oncoming trains). This means that the amount of data transmitted via one MBS is approximately half that from one fixed BS usually. On the contrary, the transmission volume for rare meetings of passing in the same direction trains will be much larger. It should also be borne in mind that the MBS-train can use part of the satellite connection for its data transmission.

IV. SATCOM AND TERRANET WITH DELAY TOLERANT NETWORK ROUTING COMPARISON

Trains that can communicate with each other, store, transmit and forward delay-tolerant data can implement the network routing that is resistant to disruptions [26] (DTN-protocols). The idea is to transfer the stored data (or a copy of it) to an encountered train, which, presumably, will deliver it earlier than the original sender (independently or taking into account subsequent forwarding). Many DTN protocols try to determine when to send data given resource constraints (memory buffer size and data volumes). Some of them for the railway scenario was discussed in [27].

Computer simulations can determine the maximum reduction in data delivery time – and even implement with unlimited resources or excellent prediction of the future. On average, in the cases considered, DTN routing can reduce 15-20% of the delivery time. The ultimate answer to the optimal protocol has not yet been found, but many decision-making algorithms give results correlated with this reduction.

It is also possible to estimate in principle the theoretical maximum contribution of such a forwarding system.

A. Delivery delay reduction after meeting of trains

Consider an offline railway section that has a connection to the outer network at the endpoints. We can assume that all trains move at the same speed, and their end-to-end time is T. Therefore, all meetings are possible only with oncoming (not passing) trains. The principle of delay calculation is shown in Fig. 4.

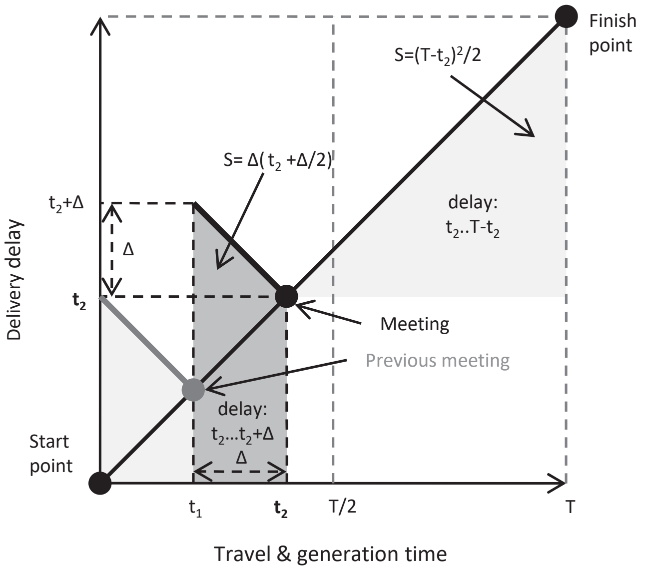


Fig. 4. Visualization of minimum delay calculation for DTN routing with multiple meetings.

If a meeting at the moment t_m occurs earlier than half of the travel time ($T/2$), then the oncoming train will deliver the accumulated data faster. The minimum delivery delay (for the latest data) will be the travel time from the moment of meeting to arrival at the endpoint (for the met train, this is the destination), i.e. the same t_m. The train can give all the data stored after the previous meeting (or from the moment of departure) during the time $\Delta = t_m - t_{m-1}$. Such piece of data will have an average delay (for uniformly generated during travel time T data):

$$\langle d \rangle_m = \frac{1}{T} \times \Delta \left(t_m + \frac{\Delta}{2} \right) \tag{9}$$

The rest of the data after the last (before $T/2$) meeting time t_m, the train delivers itself with an average delay:

$$\langle d \rangle_s = \frac{1}{T} \times \frac{(T - t_m)^2}{2} \tag{10}$$

B. Delivery delay reduction after several meetings

Let on average N train meetings occur on a section of duration T. For an assessment, it can be simplified to assume that the meeting times are uniformly distributed over the travel time (each occurs in time $\Delta = \frac{T}{N+1}$). Then about half of the meetings $M \approx \frac{N}{2}$ will occur on the first half of the way (before $T/2$), and the time of each meeting is $t_m = \Delta \times m$.

Thus, the average delivery delay can be calculated as the sum of M partial delays before the middle of the travel and self-delivery after all M encounters:

$$\langle d \rangle = \langle d \rangle_s + \sum_{i=1}^{M} \langle d \rangle_m = \frac{\Delta^2 (N - M + 1)^2}{2T} \tag{11}$$

$$+ \frac{\Delta^2}{T} \sum_{i=1}^{M} \left(n + \frac{1}{2} \right) \underbrace{\approx}_{M \approx \frac{N}{2}} \langle d \rangle_0$$

$$\times \underbrace{\left(1 - \frac{1}{2} \left(\frac{N}{N+1} \right)^2 \right)}_{DTN_0}$$

Where $\langle d \rangle_0 = \frac{T}{2}$ is the average delivery delay without meetings.

This reasoning is correct for an unlimited data rate (or a relatively small amount of data generated). In practice, it is necessary to take into account the remarks on the volume similar to the MBS, and also that the stored data will include not only own data – but also possible received data from other trains for future forwarding (depending on DTN protocol used).

C. DTN for TerraNet

If the average number of meetings of each train on the section is on average N_{meet}, and the number of base stations N_{BS} divides the section into $N_{BS} + 1$ equal offline subsections, each subsection will have on average $N_1 = \frac{N_{meet}}{N_{BS}+1}$ meetings (of that particular train). Thus, it is possible to estimate the average message delivery delay for BS, taking into account the DTN-forwarding system, using this value in (11):

$$\langle D \rangle_{BS+DTN}(N_{BS}) = BS(N_{BS}) \times \underbrace{\left(1 - \frac{1}{2} \left(\frac{N_{meet}}{N_{meet} + N_{BS} + 1} \right)^2 \right)}_{DTN_{BS}(N_{BS})} \tag{12}$$

$$\times \langle D \rangle$$

Despite the rough assumptions (the most important is the assumption about the uniformity of meetings during the movement – the analysis of any schedule shows that this is rough for a rather chaotic arrangement of meeting times, as well as the assumption of constant speed and the absence of stops), the estimate is usually not far from the results of computer simulation based on a close-to-real schedule (see Fig. 5 example):

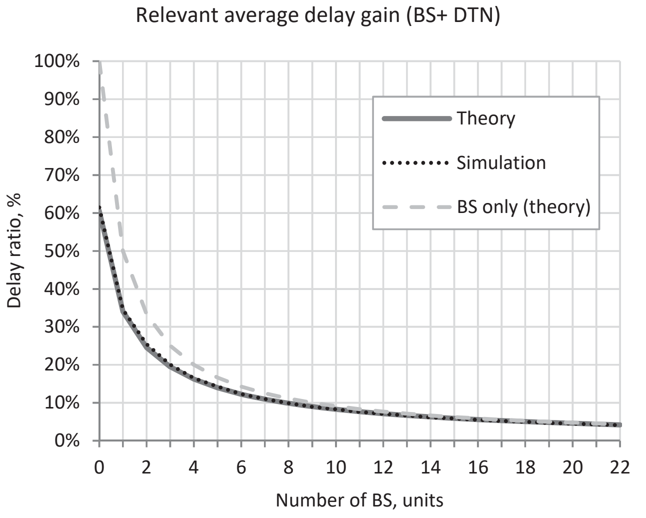


Fig. 5. Delivery delay (relevant to average delay without BS) for Tyumen-Surgut schedule (22 trains, 704 km route, 174 meetings – 7,9 meeting per each train). Rank correlation coefficient $R^2 = 0{,}99988$.

With a small number of base stations, the DTN efficiency is significant, but it decreases quickly with an increase in the number of base stations. This is because there are relatively few meetings, and as a result, they do not occur in a large number of the offline subsections, formed by BSs.

D. DTN for SatCom

When the train has a satellite terminal, it no longer needs a DTN forwarding system to deliver its own messages, the train sends them directly via a satellite link. Therefore, these trains are excluded from the average delay reduced by DTN (both with the presence of several base stations on the section and without them). Accordingly, the percentage of these excluded delays is directly proportional to the percentage of trains equipped with ST. Only trains without ST remain in the average delay. Therefore, there will be a simple direct linear dependence of DTN¬¬BS+ST factor on the portion of trains with ST, starting from the DTN value without any satellite terminal and ending with 100% (no DTN influence), equation (13).

$$DTN_{BS+ST} = DTN(N_{BS}, N_{ST}) = DTN_{BS} + (1 - DTN_{BS}) \times \frac{N_{ST}}{N_{trains}} \qquad (13)$$

$$= DTN_{BS}\left(1 - \frac{N_{ST}}{N_{trains}}\right) + \frac{N_{ST}}{N_{trains}}$$

And thus, the delay of SatCom, taking into account all the components, can be estimated:

$$\langle D\rangle_{ST+mBS+DTN} = ST(N_{ST}) \times mBS(N_{ST}) \times DTN_{BS+ST} \times \langle D\rangle \qquad (14)$$

The estimate meets the simulation results, shown in Fig. 6:

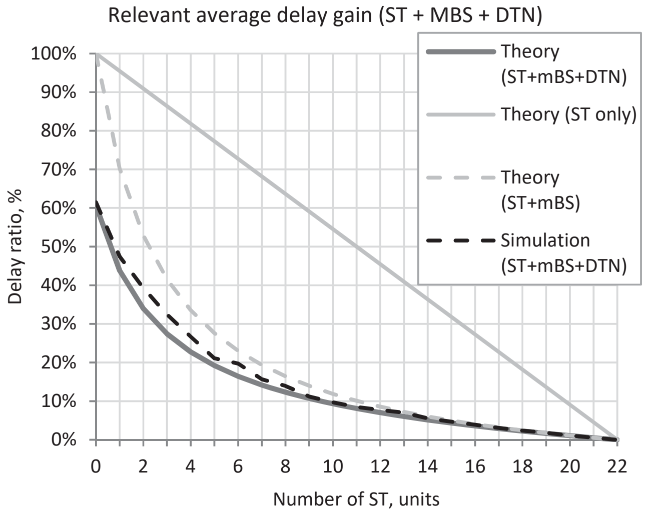


Fig. 6. Delivery delay (relevant to average delay without BS) for Tyumen-Surgut schedule (22 trains, 704 km route, 174 meetings – 7,9 meeting per each train). Rank correlation coefficient $R^2 = 0{,}996$.

The effect of DTN is strongest if the proportion of trains equipped with ST is small and decreases with its growth.

V. JOINT NETWORK DELAY: SATCOM AND TERRANET

The impact of joint SatCom and TerraNet solutions with mesh-forwarding and DTN routing together gives a final estimate of the average delay reduction (the corresponding components are described in equations 5, 6, 8, 13):

$$\langle D\rangle_{BS+ST+mBS+DTN} = BS(N_{BS}) \times ST(N_{ST}) \times mBS(N_{ST}) \qquad (15)$$

$$\times DTN(N_{BS}, N_{ST}) \times \langle D\rangle$$

Visualization example is shown in Fig. 7.

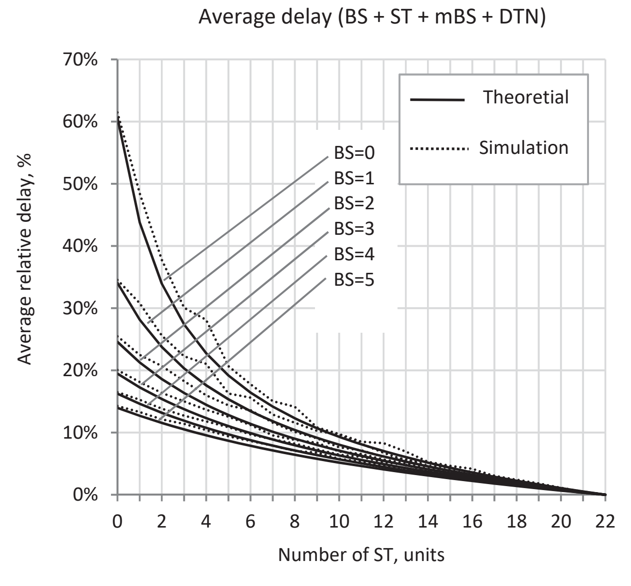


Fig. 7. Delivery delay (relevant to average delay without BS and ST) for Tyumen-Surgut schedule (22 trains, 704 km route, 174 meetings – 7,9 meeting per each train)

Also, this equation compares the characteristics of different network solutions based on BS, ST with each other for different expected parameters and their changes (see Fig. 8).

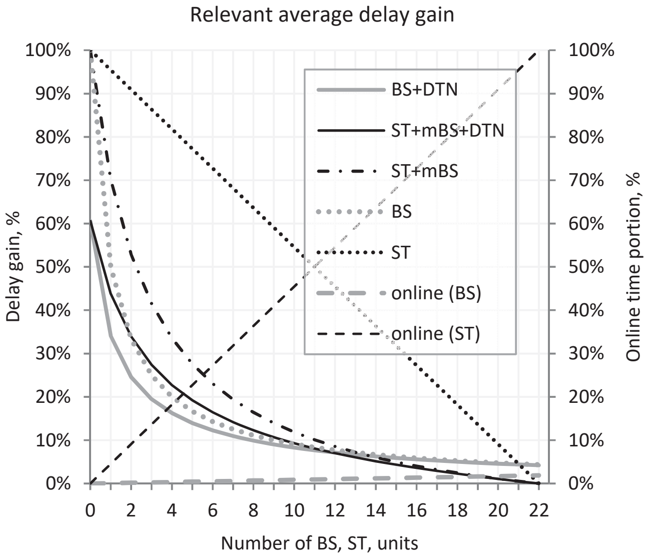


Fig. 8. Analytical comparison of the decrease in the average delivery delay when using various technologies for Tyumen-Surgut schedule (22 trains, 704 km route, 174 meetings - 7.9 meetings per each train). Decreasing grey graphs refer to the left axis, growing black ones to the right.

It is also important to take into account the limits of the amount of stored and transmitted data above (if they are exceeded, the total effect of reducing the delay will be less than the calculated one).

More details can be taken into account in the assessment. For example, mesh-connection chains; or that trains with ST are not only excluded from the beneficiaries of DTN but also increase DTN effectiveness as Mobile Base Stations; or accounting for rare but existing passing meetings.

However, these theoretical estimations already well approximate the results got in a computer simulation based on real schedules and are acceptable under the initial assumptions specified in Section III.B.

Further adaptation and development of the proposed models seem appropriate for specific real cases, their conditions, and features.

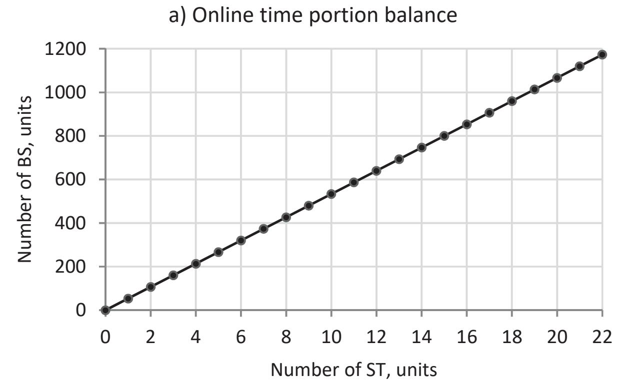
VI. SATCOM AND TERRANET BALANCE

The obtained theoretical estimate allows comparing the number of only terrestrial and only satellite terminals that provide the same characteristics (see Fig. 9). Moreover, this is possible not only for a specific particular schedule, but for possible future changes in the number of trains and/or the intensity of their movement (and, therefore, the number of meetings).

For equal online time, direct delivery delay, mesh delivery delay (using trains with ST as MBSs), this balance can be obtained in an simple analytical form (see Table II). For the equal delay in DTN-routing, the balance of the number of BSs and STs can be found from (15) or in other form from (19). The analytical solution is possible (the equation is quadratic in the number of STs), but lengthy and sophisticated and is not presented in this paper. Also all values can be found from these equations numerically.

TABLE II. EQUATIONS TO COMPARE SATCOM AND TERRANET

Characteristics	Equation
Online time portion	$N_{BS} = \frac{N_{ST}}{N_{trains}} \times \frac{L}{2R} \qquad (16)$
Delay Tolerant Data average delay	
Direct delivery	$N_{BS} = \frac{1}{1 - \frac{N_{ST}}{N_{trains}}} - 1 \qquad (17)$
Mesh delivery	$N_{BS} = \frac{1 + N_{meets} \times \frac{N_{ST}}{N_{trains}}}{1 - \frac{N_{ST}}{N_{trains}}} - 1 \qquad (18)$
DTN (and mesh) delivery	$\begin{aligned} &\left(\frac{1}{1+N_{BS}}\right)\left(1-\frac{1}{2}\left(\frac{N_{meet}}{N_{meet}+N_{BS}+1}\right)^2\right) \\ &= \left(1-\frac{N_{ST}}{N_{trains}}\right)\left(\frac{1}{1+N_{meet} \times \frac{N_{ST}}{N_{trains}}}\right) \\ &\times \left(\left(1-\frac{1}{2}\left(\frac{N_{meet}}{N_{meet}+1}\right)^2\right)\left(1-\frac{N_{ST}}{N_{trains}}\right)+\frac{N_{ST}}{N_{trains}}\right) \end{aligned} \qquad (19)$


a) Online time portion balance

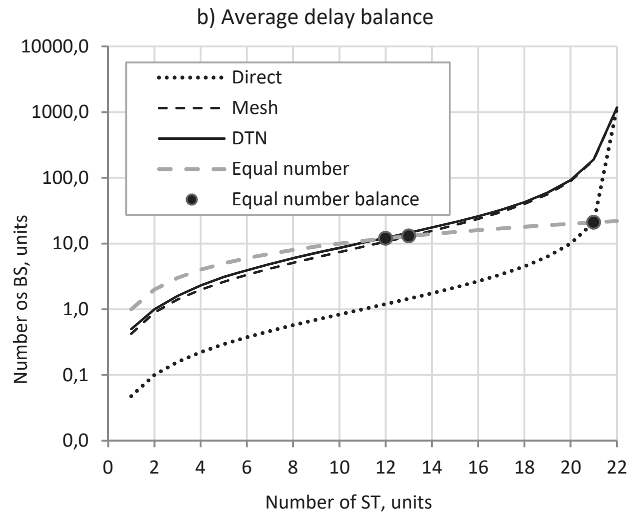

b) Average delay balance

Fig. 9. Balance (equal characteristics) for online time and data (a), and average delivery delay (b) for direct connection only (BS vs. ST), mesh connections (BS vs. ST with mBS mesh option) and DTN-routing (BS with DTN vs. ST with mBS and DTN) for Tyumen-Surgut railroad. Circles show points with equal average delay with equal terminals number.

For example, for the illustrated Surgut-Tyumen railroad (long distance, few trains), the use of mesh technologies (MBS) and additional DTN-routing significantly increase the advantage in delay tolerant data average delay delivery of satellite communications, which is a logical choice in terms of providing more time online with fewer terminals.

VII. CONCLUSION

The online time of the trains in SatCom and TerraNet was compared in conditions with and without repeating mesh-network for Real-time Data with different coverage of terrestrial base stations and satellite terminals. Also, the average data delivery delay with and without DTN routing for Delay Tolerant Data was analyzed. These values were estimated analytically using a minimum number of parameters of railway schedule and computed in more accurate computer simulations. Both methods provide similar results.

The proposed methods allow comparing the SatCom and TerraNet integrated networks, with various combinations of the number of satellite communication terminals and stationary ground base stations (possibly also operating through a satellite connection). The aim is to achieve a given QoS of specified applications for various types of data – both Real-time Data and Delay Tolerant Data. This can be used both in the design of the future 5G network infrastructure and in planning the stages of its deployment.

For critical data (signaling, emergencies) the existing low-speed network of previous generations (that already has complete or large enough coverage) can be used during the migration period.

Similar dependencies were obtained theoretically and checked for compliance by simulation for the Vienna-Salzburg, Moscow-St. Petersburg sections and several synthetic railway schedules. The similarity measure is close to the case chosen for illustration in this paper.

The results obtained are a good starting point for future works. First of all, they can serve as a base to create more detailed models taking into account more features, deliberately ignored in this paper in favor of the simplicity of assessment and conclusions. Among additional parameters there may be resource limitations (data transmission rates, memory buffers volume); imperfection, deviations, and errors (object speeds, schedule accuracy, coverage and communication areas, transmission errors); complex infrastructure and facilities schemes (non-random assignment of satellite terminals, unevenness of terrestrial base stations positioning, various non-linear topologies of railway tracks). For a technical-economic comparison, just general initial considerations are indicated. That requires a deeper study in the future.

There are also a lot of open, unsolved problems related to the topic. For example, the optimal DTN routing algorithm in the railway scenario is unknown yet. The idealized algorithm of this paper can be realized with full and accurate knowledge of all future meetings (Earliest Delivery protocol), or extremely resource-intensive copying of all information to each other during the encounter (Epidemic protocol). Of course, both solutions are impractical and just indicate the goal that a good routing method should strive for.

Also, difficult practical problems may be the problems of seamless handover between SatCom and TerraNet, backward compatibility of new networks with equipment of the old generation, complex interference conditions on the railroad (power poles, trees, buildings, tunnels, cuttings, etc.), satellite communications on high-speed railway and many others.

Difficult practical problems are also the seamless handover between SatCom and TerraNet, backward compatibility of new networks with old generation equipment, complex interference and fading conditions on the railroad (power poles, trees, buildings, tunnels, cuttings, etc.) affecting terrestrial and satellite communications in different ways, and many others.

Many of these questions remain to be explored. Probably, for many of them, the results will be obtained in complex simulations and practical experiments. Therefore, the proposed simple model can act as an initial simplified base for further research.

REFERENCES

[1] M. Sneps-Sneppe, N. Fedorova, G. Sukonnikov, V. Kupriyanovsky, “Digital railway and the transition from the GSM-R network to the LTE-R and 5G-R - whether it takes place?”, International Journal of Open Information Technologies, Vol 5, No 1, 2017.

[2] O. Pokusaev, A. Klimov, V. Kupriyanovsky, P. Morhat, D. Namiot, “Europe's digital railway - from ERTMS to artificial intelligence”, International Journal of Open Information Technologies, Vol 7, No 7, 2019.

[3] UIC official website, Future Railway Mobile Communication System, Web: https://uic.org/rail-system/frmcs/ (Retrieved: 16/12/2020).

[4] G. Tingting, S. Bin, “A high-speed railway mobile communication system based on LTE”, 2010 International Conference on Electronics and Information Engineering, Kyoto, 2010, pp. V1-414-V1-417, doi: 10.1109/ICEIE.2010.5559665.

[5] R. He, B. Ai, G. Wang, K. Guan, Z. Zhong, A. F. Molisch, C. Briso-Rodriguez, C. P. Oestges, “High-speed railway communications: from GSM-R to LTE-R”, IEEE Vehicular Technology Magazine, vol. 11, no. 3, pp. 49-58, Sept. 2016, doi: 10.1109/MVT.2016.2564446.

[6] Samsung Newsroom, World’s First LTE-Railway Service on High-speed Train Goes Live in Korea, Supplied by Samsung and KT, 21.12.2017, Web: https://news.samsung.com/global/worlds-first-lte-railway-service-on-high-speed-train-goes-live-in-korea-supplied-by-samsung-and-kt

[7] J. Kim, H. Chung, S. Choi, I. G. Kim, Y. Han, “Mobile hotspot network enhancement system for high-speed railway communication,” 2017 11th European Conference on Antennas and Propagation (EUCAP), Paris, 2017, pp. 2885-2889, doi: 10.23919/EuCAP.2017.7928212.

[8] Z. Li, Y. Chen, H. Shi, K. Liu, “NDN-GSM-R: a novel high-speed railway communication system via Named Data Networking”, EURASIP Journal on Wireless Communications and Networking, 2016, doi: 10.1186/s13638-016-0554-z.

[9] A. Gonzalez-Plaza, J. Moreno; I. Val, A. Arriola, P. M. Rodriguez, F. Jimenez, C. Briso, “5G communications in high speed and metropolitan railways”, 2017 11th European Conference on Antennas and Propagation (EUCAP), Paris, 2017, pp. 658-660, doi: 10.23919/EuCAP.2017.7928756.

[10] B. Ai, A. F. Molisch, M. Rupp, Z. Zhong, “5G key technologies for smart railways” in Proceedings of the IEEE, vol. 108, no. 6, pp. 856-893, June 2020, doi: 10.1109/JPROC.2020.2988595.

[11] D. Sanz, “Satellite technolgies for broadband internet access onboard high speed train”, SNCF-Direction de l'Innovation et de la Recherche, Paris, France, 2006.

[12] Hughes Network Systems LLC, “Communications on the move”, White Paper, 2014, Web:

https://www.hughes.com/sites/hughes.com/files/2017-04/Comms-on-the-Move-Railways_H51432_HR_07-14-141.pdf

[13] B. Ai, R. He, Z. Zhong, K. Guan, B. Chen, P. Liu, Y. Li, “Radio wave propagation scene partitioning for high-speed rails”, International Journal of Antennas and Propagation, 2012, doi: 10.1155/2012/815232

[14] H. Saarnisaari, C. M. de Lima, “5G new radio in SatCom: an overview of physical and medium access layer issues”, 2020 22nd International Conference on Transparent Optical Networks (ICTON), Bari, Italy, 2020, pp. 1-4, doi: 10.1109/ICTON51198.2020.9203099.

[15] A. Guidotti, A. Vanelli-Coralli, M. Conti, S. Andrenacci, S. Chatzinotas, N. Maturo, B. Evans, A. Awoseyila, A. Ugolini, T. Foggi, L. Gaudio, N. Alagha, S. Cioni, “Architectures and key technical challenges for 5G systems incorporating satellites”, arXiv:1806.02088v1 [cs.NI], 2018.

[16] M. Saideh, M. Berbineau, I. Dayoub, “5G waveforms for railway”, 2017 15th International Conference on ITS Telecommunications (ITST), Warsaw, 2017, pp. 1-5, doi: 10.1109/ITST.2017.7972208.

[17] K. Liolis, A. Geurtz, R. Sperber, D. Schulz, S. Watts, G. Poziopoulou, B. Evans, N. Wang, O. Vidal, B. T. Jou, M. Fitch, S. D. Sendra, P. S. Khodashenas, N. Chuberre, “Use cases and scenarios of 5G integrated satellite-terrestrial networks for enhanced mobile broadband: the SaT5G approach”, International Journal of Satellite Communications and Networking, 2019, vol. 37, pp. 91– 112, doir: 10.1002/sat.1245.

[18] S.Ferretti, H.L. Moeller, JJ. Tortora, M. Vaissiere, “Space and SatCom for 5G – European transport and connected mobility”, Space Capacity Building in the XXI Century, Studies in Space Policy, vol 22, Springer, Cham, 2020, doi: 10.1007/978-3-030-21938-3_32.

[19] S. Zhang, D. Zhu, Y. Wang, “A survey on space-aerial-terrestrial integrated 5G networks”, Computer Networks, vol. 174, 2020, 107212, ISSN 1389-1286, doi: 10.1016/j.comnet.2020.107212.

[20] B. G. Evans, “The role of satellites in 5G”, 2014 7th Advanced Satellite Multimedia Systems Conference and the 13th Signal Processing for Space Communications Workshop (ASMS/SPSC), Livorno, 2014, pp. 197-202, doi: 10.1109/ASMS-SPSC.2014.6934544.

[21] H. Khalili, P. S. Khodashenas, C. Fernandez, D. Guija, K. Liolis, C. Politis, G. Atkinson, J. Cahill, R. King, M. Kavanagh, B. T. Jou, O. Vidal, “Benefits and challenges of software defined satellite-5G communication”, 2019 15th Annual Conference on Wireless On-demand Network Systems and Services (WONS), Wengen, Switzerland, 2019, pp. 1-4, doi: 10.23919/WONS.2019.8795462.

[22] F. Rispoli, “Modern railways: connecting train control systems with mobile and satcom telecom networks”, WIT Transactions on The Built Environment, vol. 199, 2020, pp. 393 – 403, doi: 10.2495/CR200361.

[23] G. Giambene, S. Kota, P. Pillai, “Satellite-5G integration: a network perspective”, IEEE Network, vol. 32, no. 5, pp. 25-31, September/October 2018, doi: 10.1109/MNET.2018.1800037.

[24] I. Bisio, F. Lavagetto, G. Verardo, T. de Cola, “Network slicing optimization for integrated 5G-satellite networks”, 2019 IEEE Global Communications Conference (GLOBECOM), Waikoloa, HI, USA, 2019, pp. 1-6, doi: 10.1109/GLOBECOM38437.2019.9014290.

[25] D. Schneps-Schneppe, E. Tikhonov, “Mesh-network for railways”, International scientific journal “Modern Information Technologies and IT-Education”, vol. 15, num. 2, pp. 516-527, 2019, ISSN 2411-1473, doi: 10.25559/SITITO.15.201902.516-527.

[26] E. Tikhonov, D. Schneps-Schneppe, D. Namiot, “Delay tolerant network potential in a railway network”, 2020 26th Conference of Open Innovations Association (FRUCT), Yaroslavl, Russia, 2020, pp. 438-448, doi: 10.23919/FRUCT48808.2020.9087421.

[27] E. Tikhonov, D. Schneps-Schneppe, D. Namiot, “Delay tolerant network protocols for an expanding network on a railway”, 2020 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies, in press.