Review paper

Hyperloop vehicles spacing control challenges and possible solutions

M. Pawlik[1], M. Kycko[2], K. Zakrzewski[3]

Abstract: Amount of works and activities tending towards defining new transport mode on the basis of the hyperloop system concept is growing significantly. They assume use of individual vehicles, offering space for several dozen passengers, running with speeds near speed of sound in a closed space with significantly lowered air pressure, utilizing magnetic levitation. Simultaneously it is fairly from economic point of view assumed, that first implementations should link locations between which traffic demand is expected to be very high. Assumed short spacing between hyper-vehicles, which are frequently declared to be ad-hoc adjusted to transport demand, to the knowledge of the authors gained in railway transport, seems to be in conflict with high speed safety related spacing in view of the line infrastructure capacity operational rules defined in the UIC (International Union of Railways) documents. That is the challenge, that formed the basis for authors' investigations described in the paper. Several thesis regarding future new mode of transport based on hyperloop concept form an outcome of those investigations presented in conclusions.

Keywords: hyperloop, vehicles spacing, infrastructure capacity, hyper-trains, virtual coupling.

[1] Assistant professor, PhD., Eng., Warsaw University of Technology, Faculty of Civil Engineering, al. Armii Ludowej 16, 00-637 Warsaw, Poland, e-mail: m.pawlik@il.pw.edu.pl, ORCID: https://orcid.org/0000-0003-3357-7706
[2] M.Sc., Eng., Railway Research Institute, Chłopickiego 50, 04-275 Warsaw, Poland, e-mail: mkycko@ikolej.pl, ORCID: https://orcid.org/0000-0002-5630-2812
[3] M.Sc., Eng., Railway Research Institute, Chłopickiego 50, 04-275 Warsaw, Poland, e-mail: kzakrzewski@ikolej.pl, ORCID: https://orcid.org/0000-0002-1418-7350

1. Introduction

Transport is subject to continuous development, but it is still difficult to ensure appropriate level of capacity for some connections having high traffic demand. Comparisons show, that highest capacity is offered by maritime transport e.g. container ships for connections between east Asia and Europe and by railway transport for connections between inland locations. Goods and passengers transport respectively between shipping yards and cities offered by railway start to be inadequate in view of globalization of industry, commerce, work and culture. The impact of the modernization of railway lines on their capacity was described in [1]. Limiting activities to rebuilding railway lines does not always allow achieving expected capacity. Therefore, more and more frequently experts are proposing new concepts to meet capacity requirements [4, 5, 6, 10, 13, 15]. Proposals are starting from modifications of existing transport modes and go to new proposed modes of transport, especially the hyperloop.

In order to place following discussion on hyperloop capacity in a hyperloop system development time frame let's point, that the concept to run vehicles using totally separated infrastructure (e.g. tunnels) were present already in the nineteen century. However presently low pressure based systems known as hyperloop are being elaborated in USA partly as a way to gain some technical experience supporting space technologies and in EU by several start-up groups which have managed to establish at the end of 2020 joint technical committee JTC20 (Joint Technical Committee no 20) under CEN and CENELEC European Standardisation Organisations. The aim of the JTC20 is to define hyperloop principles and point standards which already exist and are expected to be applicable for hyperloop before starting detail research works dedicated to hyperloop technical challenges. Polish Standardisation Committee PKN is involved in JTC20 works. It has to be pointed, that according to authors knowledge, up to now there were no control command or signalling works dedicated to hyperloop systems.

The hyperloop concept assumes moving hypervehicles (pods / capsules) in a tube with significantly lowered internal pressure reducing air resistance and with magnetic levitation eliminating wheel/rail friction resistance, enabling high speeds reaching nearly speed of sound – over 1200 km/h. Some experts, especially industrial experts, have doubts and point expected high costs and risks which have to be investigated in details. This is unfortunately not reflected in publications, as it requires credible cost drivers and values, which are not available. One of the issues that gives rise to doubts is ensuring safe distance between the hypervehicles, which will slip inside the tube with high speeds and high frequencies. Disbelievers point, that proposed approach would significantly shorten travel time, but

at the same time the level of risk associated with maintaining a safe distance between hypervehicles would grow unacceptably. Let's analyse whether or not they are right.

In rail transport, in order to increase the capacity, infrastructure managers utilize subdivision of the track sections between stations into block sections, which allow keeping distance between moving trains. Block sections are the basis of line block signalling systems. In normal operational conditions only one train is allowed to run on a section. Each section is supervised whether it is occupied or free. It could be used to form train route only when it is free. Set of successive block sections forms the train route. The route is reflected as an authority for the train driver on trackside colour light signals by signalling systems [2]. Subdivision of the line section into e.g. four block sections theoretically enables running four trains between two stations on such track, however block sections cannot be too short. Due to the fact that the block sections must be of an appropriate length to enable train braking to standstill it is frequently not possible to increase the capacity by adding block sections. In such cases it is theoretically possible to add extra track. There are some e.g. four track lines in Europe, but only is some cases adding additional tracks is possible along whole connection. If it is only on part of the route enlarging capacity, despite the need, is not possible. Hence, to increase capacity, alternative ways for managing trains spacing were proposed and partly endorsed.

2. Space between trains in railway transport

Increasing train speed results in longer braking distances, which can pose a challenge to drivers in terms of signal visibility. Increasing the capacity causes the block sections to be shorter, which also increases the complexity of the line block signalling systems. Long braking distances in combination with short block sections result in the use of multiple block sections to brake a train. On the other hand, maintaining longer block sections as speeds increase results in a long distance between trains, which of course limits line capacity.
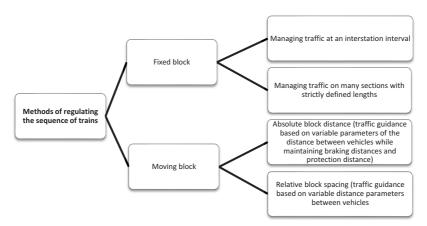
Fig. 1. Division of the applied methods of train sequence control.

2.1. Capacity of the railway lines and stations

Capacity is defined as the number of trains that can travel a certain section of a line in a given time. The bandwidth is also influenced by quality factors such as, for example, infrastructure features, signalling system, traffic distribution. One of the documents that defines the method of calculating the consumption of production capacity is the International Union of Railways UIC leaflet [3], in which the capacity depends on:

– the number of trains per time interval,
– the average speed of the trains,
– the stability of timetable,
– the heterogeneity of traffic.

The relationship between these parameters is shown at Figure 2.
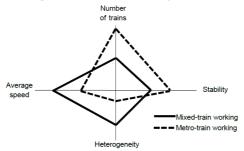

Fig. 2. Railway capacity balance [3].

Increasing traffic volume and speed reduces the quality and safety of traffic. Stability of the timetable is based on margins added to the running time of trains and margins added to the time required for setting following routes by signalling system. Stability of the timetable decreases when time margins are smaller and as a result even trains' minor delays are being propagated to other trains causing further usually bigger delays leading to operational disturbances. On the other hand, homogeneity of traffic means that each vehicle would travel each block section at the same speed, which is unlikely, since all vehicles would have to have the same dynamic running characteristics, and the sections should have the same length, cant, etc.

When differences between train dynamic running characteristics and differences between sections are growing, differences in running time of trains of different types are also growing. When differences between sections are growing trains can follow each other in a homogenous way if margins added to the routes are big enough. When heterogeneity of traffic increases the capacity consumption for the same number of trains increases proportionately, as shown at Figure 3.
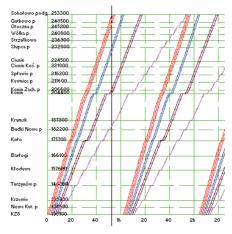

Fig. 3. Graphic timetable example.

Where:
horizontal axis represents time while vertical axis represents distance along investigated railway line
red, blue, brown and violet lines represent respectively intercity trains, long distance trains, interregional trains
and local trains stopping shortly on many stations and therefore significantly lowering railway line capacity

During capacity planning and infrastructure planning, it turns out that the use of infrastructure is unprofitable. From a timetable point of view, capacity considerations bring together a given infrastructure and existing demands for train paths. From an operational point of view, on the other hand, the capacity situation is always in flux and depends on current infrastructure availability, delays, diversions and the number of additional trains and their characteristics [3].

2.2. Fixed block signalling

The method of regulating the sequence of trains based on the principle of a fixed block distance is commonly used by railway infrastructure managers. Fixed block signalling divides the track into small blocks which determines how far apart trains will be kept for safety and how frequently stations will be served. The challenge for signal engineers is to size these blocks for optimum headway and safety at the same time; adjust the block for safety and headway is affected, adjust the block for headway and safety is affected [4]. The train is occupying block section by block section. Train starts to occupy block section when its first axle enters the track behind the trackside colour light signal transmitting information about the current state of the block section. The signal automatically starts to display stop aspect informing that track ahead is occupied. This information is transmitted to the preceding trackside signalling devices (backwards) informing the train running behind the first train about the number of free block sections, and thus providing information about the permitted speed. Signal aspects on the trackside colour lights signals are adapted to the linear block rate. [5]

Fig. 4. Blocks are defined as the distance between two signals. Trains are prevented
from colliding by ensuring that only one train is allowed into a block at once [6].

Figure 5 shows occupation time for individual block section and factors which are influencing that time. Each block section has to be individually assigned for a train before train is authorized to use it and has to be individually released after train leaves such section.
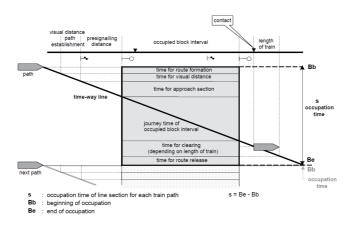
Fig. 5. Use of capacity – fixed block [3].

In order to achieve a higher capacity, shorter block sections are introduced. This is technically possible but cannot be accepted in terms of safety as long as the signalling is based on trackside light signals. In order to increase the level of safety, the basic signalling has been supplemented with digital data transmission, which is used to control the train. Digital data transmission is based on the communication of data transmission from the trackside systems to the computer systems installed in the driver's cab, which provides full information to drivers about driving restrictions [7], [8], [9].

2.3. Moving block signalling

The method of controlling the sequence of trains based on the moving block principle is based on a virtual variable block distance, the boundaries of which are not defined in a fixed manner (fixed reference points with a specified distance of the route) but dynamically change depending on the current traffic situation. With the use of this method, traffic is carried out depending on the position and running parameters of the preceding train (the second train receives via radio "movement authority" containing data specifying the maximum speed and distance to the end of the permission. In practice, two variants of the moving block are possible [10]:

− absolute moving block section, when the second running train receives permission to move to the place closer than it results from the last report on the position of the preceding train. This is a consequence of the assumption that the preceding train may stop immediately after sending the report on the position of the end of the train (in the so-called "zero time"). In this case, the

permission to drive must take into account the full stopping distance of the other vehicle and additionally the protection distance;

−   relative moving block section, when the second-going train receives permission to run to a place closer than the position that the end of the preceding train will reach after sending the last end-of-train report when this vehicle brakes as violently as it is physically possible (non-zero time stop). In this case, the distance between the two vehicles is shorter than the estimated braking distance of the second vehicle. In the implementation of the sequence based on this principle, the reaction time is important in this case (delay in transmitting information about the end position of the preceding train).

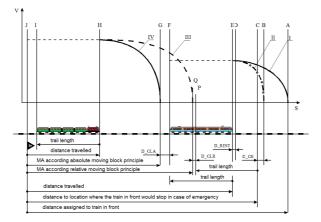The diagram of the moving block section is shown at Figure 6.

Fig. 6. Use of capacity – moving block [11].

Where:
"A" to "J" represent track locations moving with vehicles, while
"curve I" represents dynamic movement authority for the train in front,
"curve II" - shortest possible braking of the train in front in case of emmergency,
D_CE – distance representing underestimation of distance travelled from reference point in case of emergency stop,
D_RINT – distance representing underestimation of the location of the train in front during emergency brake initiation,
"curve III" – ad-hoc calculated braking curve limit according to relative moving block principle,
D_CLR - distance representing safety margin applicable to relative moving block principle,
"curve IV" – ad-hoc calculated braking curve limit according to absolute moving block principle,
D_CLA - distance representing safety margin applicable to absolute moving block principle.

Location of the train from safety point of view for the train in front means safe data about location of the end of the train and for the train which follows means safe data about location of the front of the train. Special safety relevant devices are therefore required to ensure on-board train consistency checking instead of track-side block section track occupancy checking. The moving block concept can be defined as an automated control system allowing each train to receive a movement authority

info from the control centre. The control centre has to have a continuous dialogue with all the trains to know their speed and position continuously. For this system, European Railway Traffic Management System (ERTMS) application level 3 is regarded as an example, as well as Communication Based Train Control CBTC system [12].
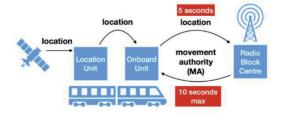

Fig. 7. Overview of ERTMS moving block railway signalling [13].

The use of a moving block principle for managing trains spacing allows to reduce the distance between trains by about 20%. Unfortunately, on some lines it is insufficient to meet the demands related to the capacity of the railway lines.

2.4. Communication based train control

The continuous development of signalling systems and the increase in speed on railway lines have resulted in the introduction of communication-based train control. This allows driving restriction information to be passed to the driver's cab via the display. Full set of limits is called Movement Authority (MA).

Before journey a set of train data has to be entered by the driver (weight, length, brake type, …). This data are used together with MA authorities received from track-side to calculate dynamic movement characteristic limits. Such limits are displayed on the driver's desk and used for verification of the train movement and the way the driver drives the train.

Communication based train control systems mostly use MA authorities based on fixed block sections which are being used by primary signalling. However, cab signalling opened possibility of frequent update of the MA authorities and therefore practical implementation of the moving block trains spacing principle. The impact of modern systems such as ERTMS / ETCS on capacity was presented in [14].

2.5. Reconfigurations of trains' compositions in motion

It is possible to introduce smaller distances between trains, especially when they are mechanically coupled by couplers and pneumatically by joining braking pipes. However, with such a solution, they are no longer separate trains. We do have such trains which are being split into two on an intermediate station and vice versa trains which are joint on the intermediate station to continue running as a single train. Joining and splitting at intermediate stations take a lot of time and hence lower capacity. The solution to this problem has been defined as reconfigurations of trains compositions in motion. Such proposal was developed based on existing characteristics of the automatic coupling systems and presented at the ESREL European Safety and Reliability Conference [15]. Analyses taking into account different cases (trains running together serving different independent locations as well as trains running together serving and not serving intermediate stations) have shown, that without new coupling technologies minimum and maximum coupling speeds of the existing technical solutions make trains joining in motion not acceptable in operation. As a result, as all trains from the timetable point of view have to run in both directions (from station A to station B and back from B to A latest the day after) theoretical possibility, proven by calculations, to split trains in motion is not enough to introduce such approach in operation.

3. Hyperloop capacity considerations

Hypervehicles are seen as individual vehicles able to serve single dozens of passengers. Usually 30 passengers are assumed per one hypervehicle. Usually assumed number of passengers per train is much higher – usually 400 passengers per train. So, there is a difference of an order.

There is no doubt, that hyperloop journey have to be as safe for the passengers as the train journey. Managing spacing between hypervehicles have to be as safe as the fixed block in case of railways. Applying fixed block principle will result with long block sections due to high speeds of running – up to the speed of sound, about 1200 km/h. Assumed hypervehicles spacing is frequently below two minutes. Such spacing, going down to 90 seconds and even below, is available in case of relatively low speeds up to 80 km/h and full heterogeneity of vehicles and low differences between characteristics of the sections. In other words, it is the case of metro within a city. It is hardly possible in case of railways. There are of course places, where time distances between trains are in single minutes. As an example one could point station Bruxelles Central on a four track line having only six tracks for stopping over one thousand trains per day at platforms. Traffic is really dense in rush hours,

but as a consequence the speeds are low at least near station. Passengers on platforms frequently can see on one platform end train leaving the station and on the other train entering the station at the same track. For hypervehicles therefore if fixed sections are to be used they have to be adjusted to time intervals as shown at figure 8.
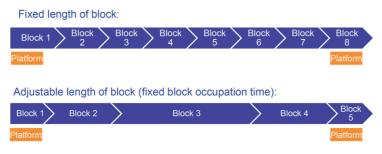

Fig. 8. Adjusting fived block to time intervals [5].

Most of the trains spacing limitations applicable to fixed block concept are arising from limitations of the trackside colour light based signalling, which have to be without any doubt perfectly recognized by the train drivers. That is not the case for hyperloop systems. Assumed speeds up to 1200 km/h impose driverless operation. Therefore, moving block approach to hypervehicles spacing would be relatively easy for implementers. However also in case of moving block required distances grow with speeds as a result of enlarged braking distances. Moreover, using relative moving block instead of an absolute moving block leads to difficult questions regarding accidents involving many hypervehicles. One could compare it to "carambol" of many cars on a motorway involving cars which are not able to stop before the place of an accident as they are running too quickly. It seems therefore that relative moving block concept may not be accepted as a solution for hyperloop.

Coming back to ad-hoc adjustment of the timetable. Some experts assume, that hypervehicles could be available at station and could be put on a run immediately when passengers are there. Adjusting timetable on demand is a nice concept, but has to respect safety related margins for hypervehicles spacing. As a result, authors propose considering creation of the hypertrains out of hypervehicles. It could be easy to add hypervehicles to hypertrains on demand. From the capacity point of view that requires new type of coupling system without mechanical or pneumatic connection. Such virtual coupling is technically possible, although presently not available. It would be based on data transmission and processing directly impacting propulsion and braking systems. The hypertrains could be regularly scheduled or running on demand e.g. when assumed minimum of hypervehicles are ready for run coupled together.

The virtual coupling would open once again the possibility to consider once more changing trains compositions in motion. For hypertrains based on virtual couplings limitations which were pointed in [15] would not apply. As a result, having hyperinfrastructure between two locations it would be possible for instance to form a hypertrain to final destination and stop e.g. last hypervehicle or two last hypervehicles on intermediate station or chosen intermediate stations without affecting run of the rest of the hypertrain. That would be possible however only if it is foreseen from the very first standardization documents as virtual coupling would have to ensure appropriately safe and reliable cooperation with propulsion and braking systems of the hypervehicles.

4. Conclusions

Above considerations show, that operational assumptions regarding speeds and vehicle spacing usually pointed by experts describing hyperloop concept are not fully reasonable especially as up to now there were no control command or signalling works dedicated to hyperloop systems. Offering safe hyper connections with reasonable traffic capacity requires taking into account possibility to form hypertrains out of hypervehicles. That has to be considered on an early stage of standardisation works due to several reasons:

- hyperloop stations would need not only platforms or other places where hypervehicles are waiting for passengers but also tracks or other infrastructure where hypertrains are formed with hypervehicles quickly and automatically just before start;
- virtual coupling systems for hypervehicles have to be fully developed in conjunction with propulsion and braking systems of the hypervehicles;
- traffic supervision systems as well as vehicles automatic operating systems have to take into account at least splitting in motion for serving intermediate stations.

That would require safety related considerations starting from risks identification and analyse and ending with safety related application conditions. Appropriate approach is already in use in case of electronic programmable solutions utilized by railways e. g. for signalling systems providing electronic movement authorities for trains. Therefore, procedural approach to standardisation should be based on RAMS standards – EN 50126-1 & -2:2018, EN 50128:2020 and EN 50129:2019. Such approach is being proposed by authors within common CEN and CENELEC works dedicated to hyperloop solutions.

References

[1] Rychlewski J., "Wpływ modernizacji linii kolejowych na ich przepustowość", TTS Technika Transportu Szynowego 7-8/2009, pp 23-27
[2] Dąbrowa-Bajon M.: Control command and signalling basics (Podstawy Sterowania Ruchem Kolejowym). OWPW, Warsaw 2005, ISBN: 978-83-7814-320-8
[3] UIC leaflet 406 - Capacity, 1st edition, Paris, September 2004
[4] Naeem M Ali, "Moving Block vs Fixed Block Signalling – Which is better?", September 2015 https://www.linkedin.com/pulse/moving-block-vs-fixed-signalling-which-better-naeem-ali/
[5] Toruń A., Lewiński A., "Metoda lokalizacji pojazdów w procesie sterowania ruchem kolejowym", Czasopismo Logistyka 4/2010
[6] Harriss Lydia, "Moving Block Signalling", Houses of Parliament, Parliamentary Office of Science &Technology, Number 20, 26 April 2016
[7] System Requirements Specification for the Eu-ropean Train Control System, filing no. A200/SRS.04-A5499A5-03.01-960809, Utrecht 1996
[8] Winter P., Compendium on ERTMS, Eurail-press, Hamburg 2009, ISBN: 978-3777103969
[9] Pawlik M., European Rail Traffic Management System, functions and technical solutions over-view – from an idea to implementation and ex-ploitation (Europejski System Zarządzania Ruchem Kolejowym, przegląd funkcji i rozwiązań technicznych - od idei do wdrożeń i eksploatacji), KOW, Warsaw 2015, ISBN 978-83-943085-1-3
[10] Toruń A., Lewiński A. Informacja przestrzenna w procesie sterowania ruchem kolejowym, VII Konferencja Naukowo - Techniczna LOGITRANS – 2010, pp 2365-2373
[11] European Railway Agency, ETCS System Requirements Specification, subset 026-1, issue 3.4.0, 2014
[12] S. Midya and R. Thottappillil, 'An overview of electromagnetic compatibility challenges in European Rail Traffic Management System', Transportation Research Part C: Emerging Technologies, vol. 16, no. 5. pp. 515–534, 2008
[13] Basile D., Maurice H., Ferrari A., "Modelling and Analysing ERTMS L3 Moving Block Railway Signalling with Simulink and Uppaal SMC", Formal Methods for Industrial Critical Systems, 24th International Conference, FMICS 2019, Amsterdam, The Netherlands, August 30–31, 2019
[14] Białoń A., Gradowski P., Gryglas M., "Wpływ system sterowania ruchem kolejowym ERTMS/ETCS na zwiększenie przepustowości linii kolejowych", TTS Technika Transportu Szynowego, 7-8/2010, pp 17-23
[15] Pawlik M., Safety margins for the reconfigurations of trains' compositions in motion, Risk, Reliability and Safety: Innovating Theory and Practice - Proceedings of the 26th European Safety and Reliability Conference, ESREL 2016, ISBN: 978-113802997-2, 2017 Taylor & Francis Group, London

Sterowanie następstwem pojazdów hyperloop, wyzwania i możliwe rozwiązania

Słowa kluczowe: hyperloop, następstwo pojazdów, przepustowość infrastruktury, hyper-pociągi, wirtualny sprzęg.

Streszczenie:

Intensyfikowane są różnego rodzaju działania zmierzające do zdefiniowania nowego rodzaju transportu na bazie koncepcji systemu hyperloop. Zakłada się wykorzystywanie pojedynczych pojazdów, korzystających z lewitacji magnetycznej, oferujących miejsca dla kilkudziesięciu pasażerów, poruszających się z prędkością bliską prędkości dźwięku w zamkniętej przestrzeni z istotnie obniżonym ciśnieniem powietrza. Jednocześnie przyjmuje się, że pierwsze wdrożenia obsługiwać będą połączenia o bardzo wysokim zapotrzebowaniu na transport. Przyjmowane małe, na przykład dwuminutowe, ad hoc zmieniane, odstępy pomiędzy hyper-pojazdami, wedle wiedzy autorów pochodzącej z transportu szynowego, wydają się sprzeczne z zasadami zarządzania następstwem pojazdów poruszających się z dużymi prędkościami opisanymi w dokumentach Międzynarodowego Związku Kolei UIC definiujących eksploatacyjne zasady zarządzania przepustowością infrastruktury liniowej. Związane z tym wyzwania legły u podstaw rozważań autorów, które zaowocowały kilkoma tezami dotyczącymi przyszłego systemu hyperloop.

Przedstawione w artykule rozważania wychodzą od różnych zasad zarządzania następstwem pociągów w transporcie kolejowym jako podstawy analizy potrzeb i możliwości zarządzania następstwem pojazdów w systemie hyperloop.

Uwzględniono zarządzanie następstwem według zasady stałego odstępu blokowego oraz według zasady ruchomego odstępu blokowego rozróżniając względy i bezwzględny ruchomy odstęp blokowy. Przedstawiono podstawy definiowania przepustowości dla liniowej infrastruktury transportowej oraz koncepcję zwiększania przepustowości poprzez łączenie i dzielenie składów w biegu. We wnioskach wskazano, że w odniesieniu do systemów typu hyperloop należy uwzględnić potrzebę tworzenia hyper-pociągów z hyper-pojazdów oraz powiązania pomiędzy trzema systemami technicznymi: wirtualnymi sprzęgami, napędami pojazdów i systemami hamowania pojazdów. Na końcu artykułu wskazano ścieżkę dalszych prac przywołując normy definiujące formalne podejście do związanych z bezpieczeństwem systemów elektronicznych i programowalnych stosowanych w transporcie kolejowym. Podejście takie autorzy proponują w ramach prac wspólnego komitetu Technicznego CEN i CENELEC dedykowanego dla systemów hyperloop.