INTELLIGENT FORECASTING METHODOLOGIES FOR HARDCORE RAILWAYS DELAYS WITH POWERFUL DEEP LEARNING SCHEMES

LOKAIAH PULLAGURA[1], DR. R.P. SINGH[2]
Research Scholar[1], Vice Chancellor[2],
CSE Department[1,2],
Sri Satya Sai University of Science and Technology, Madhya Pradesh, India[1,2]
lokaiah.p@gmail.com[1], prof.rpsingh@gmail.com[2]

Abstract:-Now-a-days Train-Delay (TD) forecasting frameworks don't exploit cutting edge instruments and systems for taking care of also, removing valuable and noteworthy data from the huge measure of endogenous (i.e., created by the railroad framework itself) and exogenous (i.e., identified with railroad task yet produced by outside wonders) information accessible. Moreover, they are not planned so as to manage the inherent time shifting nature of the issue (e.g., normal changes in the ostensible timetable, and so forth.). The reason for this paper is to manufacture a dynamic information driven TD forecasting framework that misuses the latest instruments and strategies in the field of time shifting huge information investigation. Specifically, we delineate TD forecasting issue into a period shifting multivariate relapse issue that permits abusing both verifiable information about the train developments and exogenous information about the climate given by the national climate administrations. The execution of these techniques have been tuned through the best in class threshold out procedure, a ground-breaking strategy which depends on the differential security hypothesis. At long last, the execution of two productive usage of shallow what's more, profound outrageous realizing machines that completely abuse the later in-memory vast scale information preparing innovations have been contrasted and the present best in class TD forecast frameworks. Results on certifiable information originating from the Italian railroad organize demonstrate that the proposition of this paper can amazingly enhance the cutting edge frameworks.

Keywords: Apache Spark, enormous information, Deep Extreme Learning Machine (DELM), postpone forecast, dynamic shifting frameworks, in-memory processing, savvy transportation frameworks, display determination, railroad, Shallow Extreme Learning Machine SELM), threshold-out.

I. INTRODUCTION

Present research slants in railroad transportation frameworks have appeared expanding enthusiasm toward new advances ready to gather, store, process, and picture a lot of information, just as toward new approachs originating from machine learning, man-made brainpower, and computational insight to dissect that information so to remove significant data. Precedents are: condition-based support of railroad resources [1][2][3], programmed visual review frameworks [4], [5], organize limit estimation [6], streamlining for vitality productive rail route activities [7], promoting investigation for rail cargo transportation [8], use of ontologies and connected information in rail lines [9], [10], enormous information for rail assessment frameworks [11], complex occasion handling over train information streams [12], blame conclusion of vehicle on-board hardware for fast railroads [13]– [15] and for customary ones [16], inquire about on capacity and recovery of a lot of information for rapid trains [17], improvement of an online geospatial danger display for rail route systems [18], train marshaling enhancement through hereditary calculations [19], and explore on new innovations for the railroad ticketing frameworks [20]. Specifically, this paper centers around building a dynamic train defer forecast framework going for giving helpful data to traffic the board and dispatching forms through the use of cutting edge devices and methods ready to coordinate heterogeneous information sources and to manage dynamic shifting frameworks.

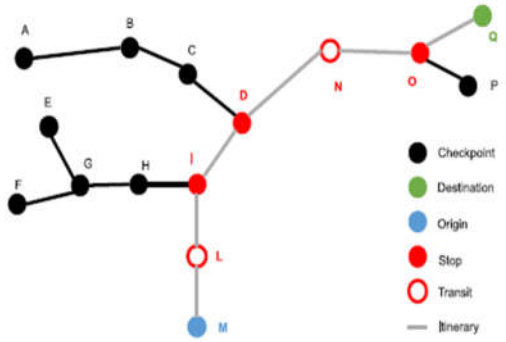

Fig.1. Railway network depicted as a graph, including a train itinerary from checkpoints M to Q.

Postponements can be because of different causes: disturbances in the tasks stream, mishaps, failing or harmed gear, development work, fix work, and serious climate conditions like snow and ice, floods, and avalanches, to give some examples. Despite the fact that trains should regard a settled calendar called ostensible timetable train delays happen every day and can influence contrarily railroad activities, causing administration disturbances and misfortunes in the most pessimistic scenarios. Rail traffic the board frameworks [21] have been created to help the administration of the intrinsic unpredictability of rail administrations and systems by giving a coordinated and all encompassing perspective of operational execution, empowering large amounts of rail activities effectiveness.

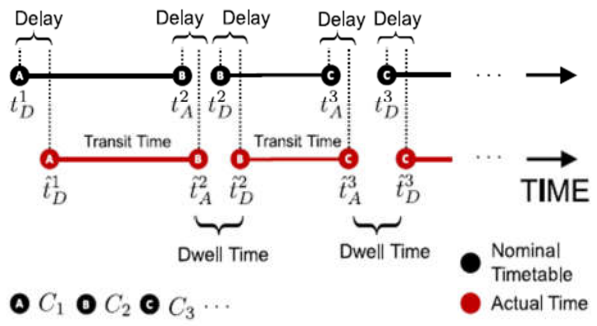

Fig.2. From the train itinerary to mathematical notation

By giving a precise DTDPS to TMSs, it is conceivable to incredibly enhance traffic the board and dispatching regarding the accompanying. 1) Passenger data frameworks, expanding the view of the dependability of railroad traveler administrations and, if there should arise an occurrence of administration disturbances, giving legitimate options in contrast to travelers searching for the best train associations [22], [23]. 2) Freight following frameworks, evaluating products' a great opportunity to entry effectively so as to enhance clients' decision-making forms. 3) NT arranging, giving the likelihood of refreshing the train trip planning to adapt to repetitive TDs [24]. 4) Delay the board (rescheduling), permitting traffic directors to reroute prepares so as to use the railroad arrange bitterly [25], [26]. Because of its key job, a TMS stores the data about each train development, i.e., each train landing and takeoff timestamp at "checkpoints" observed by flagging frameworks (e.g., a station, a switch, and so on.). Datasets created by TM records have been utilized as basic information hotspots for each work tending to the issue of building a DTDPS.

For example, Milinkovi'c et al. [27] built up a fluffy Petri net model to gauge TD-put together both with respect to master learning and on recorded information. Berger et al. [28] displayed a stochastic model for TD spread and conjectures dependent on coordinated non-cyclic diagrams. Pongnumkul et al. [29] dealt with information driven models for TD expectations, regarding the issue as a period arrangement estimate one. Their framework depended on autoregressive coordinated moving normal and closest neighbor models, despite the fact that their work reports the utilization of their models over a restricted arrangement of information from a couple of trains. At long last, Goverde [30], Hansen et al. [31], and Kecman et al. [32], [33] built up an escalated research with regards to TD expectation and proliferation by utilizing process mining systems dependent on imaginative coordinated occasion diagrams, on chronicled TM information, and on master learning about railroad framework. Be that as it may, these models depend on guileless information driven methods, and they just consider TMs information to make their forecasts. Different components influencing railroad activities (e.g., drivers' conduct, travelers volumes, strikes and occasions, and so on.) are in a roundabout way considered (e.g., explicit models for quite a long time), or even not considered, and now and again they can't be effectively incorporated in the models.

In addition, the characteristic unique nature of the marvels, due for the most part to the adjustments in the conduct of the travelers and to the adjustments in the NT, is never considered. Rather, utilizing best in class apparatuses and procedures, it is conceivable to play out a more profound investigation over information originating from various sources however identified with a similar marvel, seeking after the possibility that the more data is accessible for the making of the model, the better the execution of the model will be. Hence, this paper researches the issue of building a DTDPS by misusing cutting edge devices and methods that permit information driven models to bargain productively and successfully with a lot of information originating from railroad arrange dynamic frameworks, incorporating additionally exogenous information with specific reference to climate data. Conversely to the best in class DTDPS models, which depend on earlier learning of the railroad arrange, the proposed model is totally information driven and does not require any earlier data about the railroad organize.

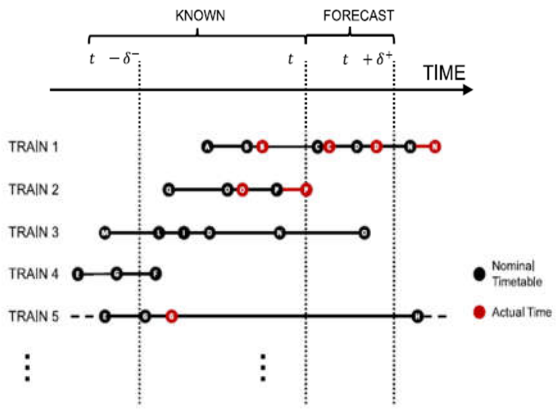

Fig.3. Data available to the TD prediction models for the network of Fig. 1.

Truth be told, creators arrangement considers the issue of anticipating TDs as a dynamic time arrangement conjecture issue, where each TM speaks to an occasion in time. TMs information recognize a dataset of TD profiles for each train, from which it is conceivable to manufacture a lot of information driven models that, cooperating, play out a relapse examination on the past TD profiles and thusly foresee future ones by considering the time fluctuating nature of the issue. Also, this arrangement is general enough to make conceivable to incorporate information about WCs identified with the agendas of the thought about trains, for instance of the mix of exogenous factors into the determining models.

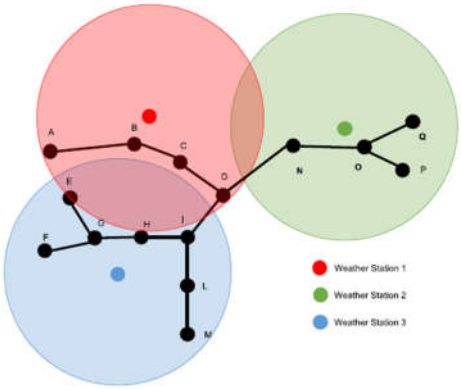

Fig.4. Weather Information.

In the relapse system, and more all in all in the regulated learning structure, extraordinary learning machines (ELMs) speak to a cutting edge instrument [34] regardless of whether different methodologies exist in the field of neural systems [35]– [41]. ELM [42]– [44] were acquainted with conquer issues presented by back-spread preparing calculation [45]– [48]: conceivably moderate intermingling rates, basic tuning of improvement parameters, and nearness of nearby minima that call for multistart and retraining systems. The first ELM are likewise called "shallow" ELM (SELM) in light of the fact that they have been created for the single-shrouded layer feedforward neural systems [49]– [51], and they have been summed up so as to adapt to situations where ELM are not neuron alike. SELM were later enhanced to adapt to issues unmanageable by shallow designs [52]– [56] by proposing different profound ELM based upon a profound engineering [57]– [59], so to make conceivable to separate highlights by a multilayer include portrayal structure. This paper considers, out of the blue, both SELM and DELM for anticipating TDs, and proposes an adjustment of their normal learning systems to misuse Big Data parallel structures so as to meet the high-requesting computational prerequisites of dynamic vast scale railroad systems. Specifically, the proposed usage exploit and completely misuse the ongoing Apache Spark [60], [61] in-memory expansive scale information handling innovation upon a condition of-theart huge information engineering [62] (Apache Spark on Apache YARN [63], [64]) running on the Google Cloud framework [65]. Note that writers have picked SELM since they have just been analyzed, in a prior primer work of [66], with other best in class approaches. Results have demonstrated that SELM exhibitions are practically identical to other best in class approachs in any case, as creators will appear in this paper, SELM are progressively suited for adapting to the enormous information nature of the issue. At long last, by utilizing DELM, creators will demonstrate that it is conceivable to additionally enhance the ELM exhibitions. Demonstrate choice is another key issue in gaining from information, going for tuning and evaluating the execution of a learning method [67], [68]. Resampling strategies like wait, cross approval and bootstrap [68] are regularly utilized by experts since they function admirably by and large, in spite of the fact that they may prompt serious false revelation issues [69], particularly when the quantity of hyperparameters is high. Differential security permitted achieving an achievement result by interfacing the field of protection safeguarding information investigation and the speculation capacity of a versatile learning method [70]– [72]. From one hand, it demonstrated that a learning calculation which indicates DP properties likewise sums up [70], [73]. From the other hand, if a calculation isn't DP, it permits to express the conditions under which a hold out set can be reused without danger of false revelation through a DP method called thresholdout [71], [74], [75]. Consequently, without precedent for writing as per the learning of the writers, we will exploit the thresholdout methodology so as to improve the hyperparameters of SELM and DELM in a versatile route by demonstrating the points of interest regarding utilizing conventional resampling techniques. The depicted methodology and the expectation framework execution have been approved dependent on the genuine verifiable information given by Rete Ferroviaria Italiana, the Italian foundation chief that controls all the traffic of the Italian railroad arrange [76], and on chronicled information about WCs and figures, which is openly accessible from the Italian climate administrations [77], [78]. For this reason, a total arrangement of novel and explicit key execution markers concurred with RFI and dependent on the prerequisites of their frameworks, however broad enough to be reached out to other railroad systems, has been structured and utilized. A while of TM records and WC information from the whole Italian railroad arrange have been misused, demonstrating that the new proposed strategy outflanks the present system utilized by RFI, which is to a great extent dependent on the best in class approach of [33], to foresee TDs as far as generally exactness.

II. LITERATURE SURVEY

In the year of 2013 the authors "K. Noori and K. Jenab" proposed a paper titled "Fuzzy reliability-based traction control model for intelligent transportation systems [15]" in that they described such as a fluffy Bayesian footing control framework was created for rail vehicles with speed sensors in savvy transportation frameworks. The framework included three fundamental segments to detect, process, and order the footing conditions. The data got from the speed sensors is utilized to maintain a strategic distance from any blunder that may cause benefit interference and pointless upkeep.

There are, nonetheless, events when these signs may not be detected, transmitted, or got decisively because of surprising conditions, for example, clamor. In this manner, in this investigation, the γ-level fluffy

Bayesian model was proposed for sensor-based footing control frameworks. So as to apply the fluffy Bayesian idea, the wheel increasing speed was thought to be a fluffy irregular variable for enrollment work with fluffy earlier circulation.

Utilizing the fluffy signs, the canny model figures the danger of grouping for the framework that outcomes in deciding the misclassification choice at least expense. The model's motor includes a scientific issue which can be fathomed in any programming dialect in installed or inserted PCs. The reasonable model was connected to a contextual investigation with promising outcomes, which can be utilized for target frameworks or reenactment.

In the year of 2015 the authors "Z. Bin and X. Wensheng" proposed a paper titled "An improved algorithm for high speed train's maintenance data mining based on MapReduce [16]" in that they described such as every day support of the fast Electric Multiple Units (EMU) trains creates a lot of information which can be used for EMU's blame analysis. The current parallel incessant example development calculation for information mining has a few shortcomings in this application. In this paper, an enhanced calculation is proposed by utilizing the nearby incessant example tree (FP-Tree) rather than the worldwide FP-Tree. This calculation embraces parallel handling in each datum preparing steps. The creation guidelines of the neighborhood FP-Tree are upgraded, and the looking methodology of the successive examples are additionally moved forward. This calculation demonstrates quick, exceedingly productive and precise in the trials during the time spent EMU's blame finding.

In the year of 2015 the authors "B. Wang, F. Li, X. Hei, W. Ma, and L. Yu" proposed a paper titled "Research on storage and retrieval method of mass data for high-speed train [17]" in that they described such as to guarantee the safeguards of rapid train, a ton of sensors are introduced in the train to ongoing screen its task status. These sensors gather measures of information which might be a few GBs for one time task of one train. Such huge sum information should be put away and made do with high proficiency. In this paper, a half and half stockpiling design by consolidate database and document framework is proposed, in which the detachment on the line and segment is performed separately while putting away the crude observed information. Further, an enhanced pressure calculation is utilized to pack twofold buoy information in the observed information. The reenactment result demonstrates that the proposed strategy is compelling and plausible.

In the year of 2015 the authors "Y. Qingyang and Y. Xiaoyun" proposed a paper titled "Scheduling Optimization Model and Algorithm Design for Two-Way Marshalling Train [19]" in that they described such as Train task alteration is a sort of continuous multi-target expansive scale blend improvement issue with multi-confinement. An epic resistant hereditary calculation dependent on fake insusceptible calculation and hereditary calculation is advanced. At that point planning improvement show for two-way marshaling train is proposed, which is fathomed by the proposed safe hereditary calculation. At long last, a case dependent on the two way marshaling train is developed. The reproduction results demonstrate the practicality of the model and the calculation.

In the year of 2014 the authors "Y.-T. Zhu, F.-Z. Wang, X.-H. Shan, and X.-Y. Lv" proposed a paper titled "K-medoids clustering based on MapReduce and optimal search of medoids [20]" in that they described such as: at the point when there are commotions and anomalies in the information, the customary k-medoids calculation has great strength, in any case, that calculation is appropriate for medium and little informational collection for its unpredictable computation. MapReduce is a programming model for handling mass information and appropriate for parallel registering of enormous information. Accordingly, this paper proposed an enhanced calculation dependent on MapReduce and ideal hunt of medoids to bunch huge information. Initially, as indicated by the fundamental properties of triangular geometry, this paper decreased computation of separations among information components to help seek medoids rapidly and diminish the estimation unpredictability of k-medoids.Also, as per the working guideline of MapReduce, Map work is in charge of ascertaining the separations between every datum component and medoids, and doles out information components to their groups; Reduce capacity will check for the outcomes from Map work, seek new medoids by the ideal inquiry methodology of medoids once more, and return new outcomes to Map work in the following MapReduce process. The trial results demonstrated that our calculation in this paper has high proficiency and great adequacy.

In the year of 2017 the authors "M. Dotoli, N. Epicoco, M. Falagario, C. Seatzu, and B. Turchiano" proposed a paper titled "A decision support system for optimizing operations at intermodal railroad terminals [23]" in that they described such as: we present a choice help apparatus to advance two of the most basic exercises in multi-purpose railroad holder terminals, in an iterative and coordinated structure gave to the terminal benefit enhancement. To begin with, the model permits improving the cargo trains structure, amplifying the organization benefit, while regarding physical and financial limitations, and setting in the train head/tail holders indicting to consequent goals. Subsequently, in light of the subsequent train arrangement, the choice emotionally supportive network permits advancing the compartments allotment in the terminal stockpiling yard, so as to boost the filling level while regarding physical requirements. The model is effectively tried on a genuine

contextual analysis, the inland railroad terminal of a main Italian multi-purpose coordinations organization.

In the year of 2014 the authors "T. Dollevoet, F. Corman, A. D'Ariano, and D. Huisman" proposed a paper titled "An iterative optimization framework for delay management and train scheduling [25]" in that they described such as: Postpone the executives figures out which associations ought to be kept up in the event of a deferred feeder train. Ongoing postpone the board models fuse the constrained limit of the railroad framework. These models acquaint progress imperatives with ensure that wellbeing controls are fulfilled. Sadly, these progress requirements can't catch the full subtleties of the railroad framework, particularly inside the stations. We along these lines propose an advancement approach that iteratively explains a plainly visible postpone the executives show from one viewpoint, and a minuscule train booking model then again. The naturally visible model figures out which associations with keep up and proposes an aura timetable. This mien timetable is then approved infinitesimally for a bottleneck station of the system, proposing an attainable calendar of railroad tasks. We assess our iterative enhancement system utilizing true examples around Utrecht in the Netherlands.

In the year of 2014 the authors "X. Li, B. Shou, and D. Ralescu" proposed a paper titled "Train rescheduling with stochastic recovery time: A new track-backup approach [26]" in that they described such as: Train rescheduling is an essential choice process in railroad the executives. It plans to limit the negative impacts emerging from the aggravations by means of ongoing traffic the executives. Two fundamental difficulties are the means by which to figure the dynamic and complex rescheduling issue as an enhancement model, and how to get a decent arrangement inside a brief timeframe limit. Concentrating on the stochastic limit recuperation times of blocked tracks, we propose another track-reinforcement rescheduling (TBR) approach which ideally relegates each influenced train a reinforcement track, in view of the estimation of recuperation time, the first timetable, and track evolving cost. At that point, we define a blended whole number programming (MIP) model to acquire a contention free timetable which limits the defer cost and the normal track evolving cost. A ravenous calculation is intended to reorder prepares and reschedule the landing and flight times, and afterward we utilize a MIP calculation to tackle the ideal track reinforcement procedure. In light of the Beijing-Shanghai rapid railroad line, we lead broad trial thinks about which demonstrate that the TBR approach can diminish the rescheduling cost by a normal of 10.17% contrasted and conventional methodologies. Increasingly essential, the eager based calculation is appeared to have the capacity to acquire great arrangements (with a normal blunder of just 2.85%) inside 1.5 s, which suggests the high capability of our methodology in an ongoing rush hour gridlock the board framework where quick reaction is basic.

In the year of 2013 the authors "S. Milinkovi´c, M. Markovi´c, S. Veskovi´c, M. Ivi´c, and N. Pavlovi´" proposed a paper titled "A fuzzy Petri net model to estimate train delays [27]" in that they described such as: Indeed, even with the most exact timetable, prepares frequently work with deferrals. The running and hanging tight occasions for trains can increment startlingly, making essential defers that reason thump on postponements and deferrals for different trains. The precise estimation of train delays is vital for making timetables, dispatching trains, and arranging frameworks. In this work, we proposed a fluffy Petri net (FPN) show for evaluating train delays. The FPN display with attributes of progressive system, shading, time, and fluffy thinking was utilized to reenact traffic procedures and train developments in a railroad framework. The trains were shaded tokens, the track segments were named places, and discrete occasions of train development were named advances. The train essential postponements were recreated by a fluffy Petri net module in the model. The fluffy rationale framework was consolidated in the FPN module in two different ways. To begin with, when there were no verifiable information on train delays, master learning was utilized to characterize fluffy sets and principles, changing the skill into a model to figure train delays. Second, a model dependent on the Adaptive Network Fuzzy Inference System (ANFIS) was utilized for frameworks where the authentic information on train delays were accessible (from recognition frameworks or from the train dispatcher's logs). The postpone information were utilized to prepare the neuro-fluffy ANFIS demonstrate. After the aftereffects of the fluffy rationale framework were checked, the ANFIS display was recreated by a fluffy Petri net. The reenactment was approved by energizing the train development and plotting the time-separate chart of the trains. Aftereffects of the reproduction were sent out to a database for extra information mining and near investigation. The FPN demonstrate was tried on a piece of the Belgrade railroad hub.

In the year of 2011 the authors "A. Berger, A. Gebhardt, M. Müller-Hannemann, and M. Ostrowski" proposed a paper titled "Stochastic delay prediction in large train networks [28]" in that they described such as: in day by day activity, railroad traffic dependably goes amiss from the arranged calendar to a specific degree. Essential introductory postponements of trains may cause an entire course of auxiliary deferrals of different trains over the whole system. In this paper, we propose a stochastic model for defer proliferation and estimates of entry and takeoff occasions which is appropriate to all sort of open transport (not exclusively to railroad traffic). Our model is genuinely reasonable, it incorporates general holding up approaches (to what extent do trains sit tight for postponed feeder trains), it utilizes driving

time profiles (discrete appropriations) on movement bends which rely upon the flight time, and it consolidates the make up for lost time capability of support times on driving segments and train stops. The model is suited for an online situation where a huge stream of refresh messages on the present status of trains arrives which must be engendered through the entire system. Proficient stochastic proliferation of postponements has vital applications in online timetable data, in defer the executives and train attitude, and in steadiness examination of timetables. The proposed methodology has been actualized and assessed on the German timetable of 2011 with holding up approaches of Deutsche Bahn AG. A total stochastic defer proliferation for the entire German train organize and an entire day can be performed in under 14 seconds on a PC. We tried our spread calculation with counterfeit discrete travel time circulations which can be parameterized by the measure of their vacillations. Our estimates are contrasted and genuine information. For reasons unknown, stochastic spread of postponements is sufficiently effective to be pertinent by and by, yet the gauge quality requires further alterations of our counterfeit travel time circulations to gauges from genuine information.

III. CONCLUSION

This framework manages the issue of building a DTDPS dependent on cutting edge apparatuses and procedures ready to quickly get a handle on the learning covered up in verifiable information about TM and exogenous climate information. Specifically, the proposed arrangement enhances the cutting edge techniques really misused from the IM like RFI. Results on certifiable TM information given by RFI and climate information recovered from the national climate administrations demonstrate that cutting-edge examination methodologies can perform up to twice superior to current best in class strategies. Specifically, misusing just recorded information about TM gives vigorous models with superior regarding the real TD expectation arrangement of RFI. The execution of these models can be additionally enhanced by considering likewise WI.

We have likewise appeared at productively and adequately tune the hyper parameters engaged with the learning calculations because of an ongoing procedure called threshold out, which depends on the DP hypothesis. At last, by misusing the Apache Spark in memory innovation, we have possessed the capacity to assemble a framework with elite likewise as far as the required preparing time for building every one of the models required for managing a largescale railroad organize. Future works will consider different exogenous data accessible from outer sources, for example, data about traveler streams by utilizing touristic databases, about railroad resources conditions, or some other wellspring of information which may influence rail route dispatching activities.

References

[1] E. Fumeo, L. Oneto, and D. Anguita, "Condition based maintenance in railway transportation systems based on big data streaming analysis," in Proc. INNS Big Data Conf., San Francisco, CA, USA, 2015, pp. 437–446.

[2] H. Li, B. Qian, D. Parikh, and A. Hampapur, "Alarm prediction in large-scale sensor networks—A case study in railroad," in Proc. IEEE Int. Conf. Big Data, Silicon Valley, CA, USA, 2013, pp. 7–14.

[3] H. Li et al., "Improving rail network velocity: A machine learning approach to predictive maintenance," Transp. Res. C Emerg. Technol., vol. 45, pp. 17–26, Aug. 2014.

[4] H. Feng et al., "Automatic fastener classification and defect detection in vision-based railway inspection systems," IEEE Trans. Instrum. Meas., vol. 63, no. 4, pp. 877–888, Apr. 2014.

[5] C. Aytekin, Y. Rezaeitabar, S. Dogru, and I. Ulusoy, "Railway fastener inspection by real-time machine vision," IEEE Trans. Syst., Man, Cybern., Syst., vol. 45, no. 7, pp. 1101–1107, Jul. 2015.

[6] S. A. Branishtov, Y. A. Vershinin, D. A. Tumchenok, and A. M. Shirvanyan, "Graph methods for estimation of railway capacity," in Proc. IEEE 17th Int. Conf. Intell. Transp. Syst., Qingdao, China, 2014, pp. 525–530.

[7] Y. Bai, T. K. Ho, B. Mao, Y. Ding, and S. Chen, "Energy-efficient locomotive operation for Chinese mainline railways by fuzzy predictive control," IEEE Trans. Intell. Transp. Syst., vol. 15, no. 3, pp. 938–948, Jun. 2014.

[8] X. Zhang and D. Gong, "Application of big data technology in marketing decisions for railway freight," in Proc. ICLEM Syst. Plan. Supply Chain Manag. Safety, Shanghai, China, 2014, pp. 1136–1141.

[9] C. Morris, J. Easton, and C. Roberts, "Applications of linked data in the rail domain," in Proc. IEEE Int. Conf. Big Data, Washington, DC, USA, 2014, pp. 35–41.

[10] J. Tutcher, "Ontology-driven data integration for railway asset monitoring applications," in Proc. IEEE Int. Conf. Big Data, Washington, DC, USA, 2014, pp. 85–95.

[11] Q. Li, Z. Zhong, Z. Liang, and Y. Liang, "Rail inspection meets big data: Methods and trends," in Proc. Int. Conf. Netw. Based Inf. Syst., Taipei, Taiwan, 2015, pp. 302–308.

[12] M. Ma, P. Wang, C.-H. Chu, and L. Liu, "Efficient multipattern event processing over high-speed train data streams," IEEE Internet Things J., vol. 2, no. 4, pp. 295–309, Aug. 2015.

[13] F. Wang, T.-H. Xu, Y. Zhao, and Y.-R. Huang, "Prior LDA and SVM based fault diagnosis of vehicle on-board equipment for high speed railway," in Proc. IEEE Int. Conf. Intell. Transp. Syst., 2015, pp. 818–823.

[14] Y. Zhao, T.-H. Xu, and W. Hai-Feng, "Text mining based fault diagnosis of vehicle on-board equipment for high speed railway," in Proc. IEEE Int. Conf. Intell. Transp. Syst., Qingdao, China, 2014, pp. 900–905.

[15] K. Noori and K. Jenab, "Fuzzy reliability-based traction control model for intelligent transportation systems," IEEE Trans. Syst., Man, Cybern., Syst., vol. 43, no. 1, pp. 229–234, Jan. 2013.

[16] Z. Bin and X. Wensheng, "An improved algorithm for high speed train's maintenance data mining based on MapReduce," in Proc. Int. Conf. Cloud Comput. Big Data, Shanghai, China, 2015, pp. 59–66.

[17] B. Wang, F. Li, X. Hei, W. Ma, and L. Yu, "Research on storage and retrieval method of mass data for high-speed train," in Proc. 11th Int. Conf. Comput. Intell. Security, Shenzhen, China, 2015, pp. 474–477.

[18] J. Sadler et al., "GeoSRM—Online geospatial safety risk model for the GB rail network," IET Intell. Transp. Syst., vol. 10, no. 1, pp. 17–24, Jan. 2016.

[19] Y. Qingyang and Y. Xiaoyun, "Scheduling optimization model and algorithm design for two-way marshalling train," in Proc. Int. Conf. Intell. Transp. Big Data. Smart City, 2015, pp. 705–708.

[20] Y.-T. Zhu, F.-Z. Wang, X.-H. Shan, and X.-Y. Lv, "K-medoids clustering based on MapReduce and optimal search of medoids," in Proc. Int. Conf. Comput. Sci. Educ., Vancouver, BC, Canada, 2014, pp. 573–577.

[21] E. Davey, "Rail traffic management systems (TMS)," in Proc. IET Prof. Develop. Course Railway Signalling Control Syst., 2012, pp. 126–143.

[22] M. Müller-Hannemann and M. Schnee, "Efficient timetable information in the presence of delays," in Proc. Robust Online Large-Scale Optim., 2009, pp. 249–272.

[23] M. Dotoli, N. Epicoco, M. Falagario, C. Seatzu, and B. Turchiano, "A decision support system for optimizing operations at intermodal railroad terminals," IEEE Trans. Syst., Man, Cybern., Syst., vol. 47, no. 3, pp. 487–501, Mar. 2017.

[24] J.-F. Cordeau, P. Toth, and D. Vigo, "A survey of optimization models for train routing and scheduling," Transp. Sci., vol. 32, no. 4, pp. 380–404, 1998.

[25] T. Dollevoet, F. Corman, A. D'Ariano, and D. Huisman, "An iterative optimization framework for delay management and train scheduling," Flexible Serv. Manuf. J., vol. 26, no. 4, pp. 490–515, 2014.

[26] X. Li, B. Shou, and D. Ralescu, "Train rescheduling with stochastic recovery time: A new track-backup approach," IEEE Trans. Syst., Man, Cybern., Syst., vol. 44, no. 9, pp. 1216–1233, Sep. 2014.

[27] S. Milinkovi´c, M. Markovi´c, S. Veskovi´c, M. Ivi´c, and N. Pavlovi´c, "A fuzzy Petri net model to estimate train delays," Simulat. Model. Pract. Theory, vol. 33, pp. 144–157, Apr. 2013.

[28] A. Berger, A. Gebhardt, M. Müller-Hannemann, and M. Ostrowski, "Stochastic delay prediction in large train networks," in Proc. 11th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems, OASIcs—OpenAccess Series in Informatics, vol. 20. Dagstuhl, Germany: Schloss Dagstuhl Leibniz-Zentrum für Informatik, 2011, pp. 100–111.

[29] S. Pongnumkul, T. Pechprasarn, N. Kunaseth, and K. Chaipah, "Improving arrival time prediction of Thailand's passenger trains using historical travel times," in Proc. Int. Joint Conf. Comput. Sci. Softw. Eng., Chon Buri, Thailand, 2014, pp. 307–312.

[30] R. M. P. Goverde, "A delay propagation algorithm for large-scale railway traffic networks," Transp. Res. C Emerg. Technol., vol. 18, no. 3, pp. 269–287, 2010.