A Big Sensor Data Offloading Scheme in Rail Networks

1st Mahdi Saki
School of Electrical and Data Eng.
University of Technology Sydney
Sydney, Australia
Mahdi.Saki@student.uts.edu.au

2nd Mehran Abolhasan
School of Electrical and Data Eng.
University of Technology Sydney
Sydney, Australia
Mehran.Abolhasan@uts.edu.au

3rd Justin Lipman
School of Electrical and Data Eng.
University of Technology Sydney
Sydney, Australia
Justin.Lipman@uts.edu.au

Abstract—In this paper, we propose an offloading scheme to transfer massive stored sensor data from rolling stock to railway data centers. We apply a delayed offloading strategy for non-critical stored data assuming that the critical data has been already separated through an appropriate edge processing task and has been sent via a real-time communication such as cellular networks. We propose train stations as potential and feasible spots for data offloading via available wireless local area networks (WLAN) such as existing WiFi network at stations. Thus, stations will not only be the places of passenger exchange but also data exchange. We develop an analytical model customized for the proposed offloading strategy in rail applications. Then we validate the performance of our model through simulation in various scenarios in Omnet. The simulation results shows an accuracy of %98.67 for the proposed analytical model with reference to the simulation results in Omnetpp. Additionally, by using our proposed scheme, we can theoretically offload up to 5.43 GB per each stopping station.

Index Terms—big sensor data, delayed offloading, IEEE 802.11

I. INTRODUCTION

Future trains will be equipped with many sensors that continuously sense and generate massive IoT (internet of things) data [1]. According to [1], the amount of sensor data produced by only one sensor for sensing the vibration of just one wheel bearing in a train will be as huge as 10 TB during eight operating hours. Thus, for a train with many parts that will be sensed by wide variety of sensors [2], the created data amount will be extremely massive and transmission of such data into data centers will be a challenge.

Based on the risks for passengers and rail equipment, the collected sensor data is classified into two classes including critical data and non-critical data. The critical data can cause serious damages for both people or rolling stocks and should be declared immediately. However, the non-critical data is used for long time analysis and can be evaluated by delay. As trains operate in normal conditions for most of the time, the amount of critical data is tiny and the main part of sensor data is composed of non-critical data. Therefore, if we could classify the sensor data through an appropriate edge processing task (this is the subject of our another work

which is currently accepted and is under final review), we will be able to employ different communication strategies to transfer critical and non-critical data to railway data centers. In this case, it is feasible to send the tiny amount of critical data in a real-time manner (e.g. via cellular networks) while temporarily store the non-critical data and deliver it later via an appropriate offloading strategy [3]. In this way, we will significantly reduce the data traffic over expensive and infrastructure-based communication networks (such as cellular or satellite networks) by offloading the massive part of data through an available cheap WLAN's channel such as WiFi networks at stations (which approximately has no cost).

This is the idea behind our current work which based on that, we propose train stations as potential spots to offload the delay-tolerant non-critical sensor data. In this way, stations as grounded infrastructure, has the feasibility to provide more powerful computation and communication capabilities for our offloading task. Additionally, if we employ the available channels of WLAN in stations, this will cause large cost saving because we will no longer need to install any rail/road side units (RSU). The proposed offloading method will be a train-to-station (T2S) communication between on-board units (OBU) in trains and a data sink system in stations.

Therefore, the main contributions of this paper is:

- we propose a novel scheme for offloading of delay-tolerant part of IoT data in rail networks,
- we develop an analytical model for the proposed offloading scheme that can model the data offloading task for passing stations as well as stopping stations,
- we provide an integrated equation that can estimate the total offloading capacity for a given train during its trip between two terminals including stopping and passing stations
- we embedded the offloading model with a rate control algorithm that enables the data to be offloaded even with the minimum WiFi signal power and therefore, makes the offloading capacity maximum.

The rest of the paper is organized as follows. Firstly, a short literature review is presented in Section II. Then, we explain our proposed offloading scheme in Section III. In Section IV, we develop an analytical model for the proposed offloading

task. The simulation results will be presented in Section V and we conclude the paper in Section VI.
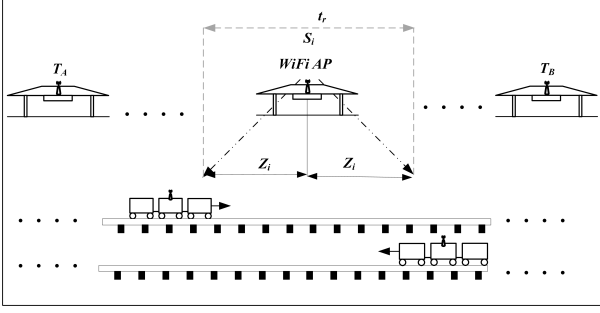

Fig. 1. Overall diagram of the proposed station-based offloading scenario

II. RELATED WORK

Generally, there are two strategies for mobile data offloading via WLAN's including opportunistic and delayed offloading [4]. Opportunistic offloading is applied when a vehicle or a user passes an offloading spot in an opportunistic manner. However, the delayed strategy is for the cases that data transmission can be delayed until it meets a WLAN access point (AP). Based on the scope of our current approach, we review the delayed offloading strategy in the following.

Authors in [5] developed an analytical model for delayed offloading of mobile users via an available WiFi network aiming to maximize the offloading capacity. However, they assumed a fixed data rate for the required WLAN and have not considered the variations of wireless channel states due to user mobility.

Kashihara in [6] has employed a high speed short range communication such as Transfer Jet to collect data from users at bus stops and then offloaded the stored data via fiber optic at terminals. Therefore, the whole collected data must be delayed until bus reaches a terminal and the author has not provided any solution for data offloading at through bus stops. Furthermore, he has only discussed "stopping" bus stops and has not considered "passing" bus stops as extra potential spots for data exchange.

Huang et. al. in [7] have proposed IEEE802.11p WiFi network as an alternative communication method to cellular network for data offloading. This approach fits to small-scale data offloading and does not suit to the offloading of large amount of data due to low throughput of IEEE802.11p standard.

Due to the above issues, we proposed an analytical model that does not only employ stopping stations for data offloading in rail networks, but it will also consider the passing stations to maximize the data offloading capacity. This will also restrict the offloading delays to the short trip time between two consecutive stations rather than long-time journeys among terminals. Additionally, our proposed model will utilize a dynamic data rate scheme vs. fixed data rate for data offloading using a an appropriate rate control algorithm that enables the offloading task to be feasible even with poor WiFi signals. Our

approach will also use IEEE802.11ac-based WiFi networks that: 1) are currently available in train stations and do not need to install any extra infrastructure, 2) can theoretically provide high throughput up to 2.34 Gbps, and 3) is potential to be upgraded by new rapid offloading technologies such as IEEE802.11ay.

III. THE PROPOSED OFFLOADING SCHEME

Figure 1 illustrates the overall diagram of our proposed station-based offloading scenario. In order to obtain an offloading model that can estimate the offloading capacity of each station, we need to estimate the two following parameters:

1) WiFi[1] contact duration, which is the time of presence of a train inside of WiFi communication zone of a given station, and
2) feasible data throughput of offloading session during such WiFi contact duration

For the short range WiFi networks, the contact duration is limited to the duration that trains are sufficiently close to the stations. The main opportunity that can be considered as the contact duration is trains dwelling times for passenger exchange at stations. To increase the efficiency of our offloading task, we add three more time slots for the contact duration. These time slots include when a train is close enough to a station during entering, leaving or passing such station. In fact, our target is to start each offloading session as soon as a train reaches the WLAN communication zone (i.e. upon detecting strong enough beacons from WiFi AP's at the stations).

The dwelling time of a train at each station is an stochastic parameter that varies between a lower and upper bounds [8]. The lower bound, which for each type of train with a given number of doors is the minimum time required for opening and closing its all doors, is a definite quantity. However, the upper bound, which is required for safely exchanging passengers, is variable based on several parameters such as station type, train specifications and hour/day of operation (e.g. peak or off-peak times and weekdays or weekends) [8]. The other time slots are also variable and depend on the speed of train when entering, leaving or passing through a station. Although these time slots are not so long based on the short range of WiFi networks, the amount of offloaded data will be significant thanks to the emerge of rapid offloading protocols such as existing IEEE802.11ac as well as the other new upcoming multi-Gbps standards like IEEE802.11ay [9].

As described, we integrate a rate control scheme into our model that enables the offloading task to be started even with minimum available levels of WiFi signal powers. Generally, there are two types of rate control schemes: PHY layer based schemes that control the data rate based on the parameters of physical layer such as received signal strength (RSS), e.g. Receiver Based Auto Rate (RBAR) algorithm; MAC layer based algorithms that work based on MAC layer parameters such as packet delivery ratio (PDR), e.g. Adaptive Auto Rate

---

[1]Since, in this work, we use IEEE802.11-based networks as the required WLAN, we sometimes use the term "WiFi" instead of "WLAN".

Fallback (AARF). Assuming an RBAR scheme, the maximum data rates of a wireless channel can be theoretically determined based on the mapping tables in the related standards such as what illustrated in IEEE802.11 series [10]. These tables maps the minimum levels of received signal-to-noise ratio (SNR) of a radio signal to a modulation and coding scheme (MSC) index. Each MCS can provide up to a definite data rate based on the carrier frequency, the available channel bandwidth and the number of spatial streams. The level of SNR is estimated based on the level of noise and the received signal strength (RSS). RSS is also estimated based on the transmission power, the distance from AP, environmental (obstacles) and weather (temperature, humidity, etc.) conditions [11] and [12]. Data throughput is theoretically a percentage of maximum data rate called MAC deficiency [13]. The actual amount of data throughput is determined via in-field measurements.

IV. THE ANALYTICAL OFFLOADING MODEL

As explained in Section III, to build the proposed offloading model, we need to find the WiFi contact duration and throughput at each station. Generally, a train will not stop at all through stations between terminals and for some stations, it will only have a short passing. To obtain the maximum efficiency in the offloading process, we consider both types of stations including stopping stations (where a train stops for passenger exchange) and passing stations (where a train just passes with no stop).

For simplicity, in the all following equations, we assume that at each station, a dedicated WiFi network has been allocated only for the task of data offloading and only one train will be permissible to offload data at each session. We also suppose equal speeds and accelerations for a given train during entering and leaving a station. Additionally, we do not affect the location of antennas assuming that there are enough number of antennas for data offloading, e.g. in the first, middle and end parts of trains and station platforms.

A. Offloading Model for Stopping Stations

Figure 2 shows the timing diagrams of the offloading model for both stopping and passing stations. Based on this, the related WiFi contact duration, t_r, for stopping stations is obtained as follows:
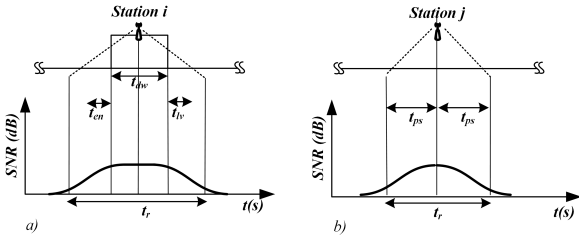

Fig. 2. Timing diagrams of the offloading model for a) stopping stations, b) passing stations

$$t_r^{stp} = t_{en} + t_{dw} + t_{lv} \tag{1}$$

where t_{en}, t_{dw} and t_{lv} are entering, dwelling and leaving times of a train at a given station, respectively.

Assuming that a train enters (leaves) the communication zone of a given station at a given speed and gradually decreases (increases) its speed with an acceleration of a until it stops. If suppose a train enters or leave a station with equal speed and acceleration, then for a maximum communication zone of d_{max} for the WiFi network at that station, entering (leaving) time, $t_{en(lv)}$, is simply obtained:

$$d = \frac{1}{2}at^2 + v_0 t + d_0 \tag{2}$$

where d_0 is the distance between transmitter and receiver during dwelling time for passenger exchange, and d is the distance after t second of leaving the station from the transmitter. For leaving scenario, $v_0 = 0$, as a train starts leaving a station from standstill situation:

$$t_{en} = t_{lv} = \sqrt{\frac{2d_{max}}{a}} \tag{3}$$

For calculation of data throughput, we firstly need to estimate the received signal strength (rss) at distance d from a WiFi access point (AP) using log-normal shadowing path loss model as follows [14]:

$$rss(d) = P_{ref} - 10\gamma log(d/d_{ref}) + X_\sigma \tag{4}$$

where P_{ref} is the received power at reference distance d_{ref} and , γ is the path loss component (PLE), and X_σ is the normally distributed random variable with zero mean and σ standard deviation (SD). γ and σ reflect the environmental conditions and are two and zero for free space, respectively.

P_{ref} can be theoretically obtained by $P_{ref} = P_t - 20log(\frac{4\pi d_{ref}}{\lambda})$ in dBm (supposing free space environment), where P_t is the transmitter power and $\lambda = c/f$ (c is the light speed and f is the radio carrier frequency).

By substituting d from (2) to (4) and considering $d_{ref} = 1m$, rss based on time is obtained as follows:

$$rss(t) = P_{ref} - 10\gamma log(\frac{at^2}{2} + d_0) + X_\sigma \tag{5}$$

According to IEEE802.11ac Standard [10], the maximum bit rate of a WiFi physical link (PHY) is estimated based on the level of signal to noise ratio (SNR). SNR can be calculated as follows:

$$snr^{dBm} = rss^{dBm} - n^{dBm} \tag{6}$$

where n^{dBm}, is the background noise level based on dBm at receiver. From (5) and (6), snr based on time can be obtained:

$$snr(t) = P_{ref} - n^{dBm} - 10\gamma log(\frac{at^2}{2} + d_0) + X_\sigma \tag{7}$$

The maximum bit rate of WiFi PHY link is obtained from MCS mapping tables based unavailable channel bandwidth,

bw, number of spatial streams, N_{ss}, and duration of guard interval (GI) as follows:

$$\forall snr_{min}^i \leq snr(t) < snr_{min}^{i+1} \xrightarrow{\mathcal{F}} bitrate_{max} = r^i \quad (8)$$

where $\mathcal{F}$ is the mapping function that maps every minimum snr to a defined bit rate based on IEEE802.11ac Standard, and $i = \{0, 1, 2, ..., 9\}$ representing the MCS indexes in IEEE802.11ac Standard. r^i is the maximum bit rate that can be reached based on level of snr.

Assuming a MAC efficiency of ρ, the throughput during moving is simply obtained from the maximum bit rate in (8):

$$th_{en(lv)} = \rho r \quad (9)$$

Therefore, the offloading equation for stopping stations is obtained as follows:

$$A_{stopping} = th_{dw}.t_{dw} + 2\sum_{t=0}^{t_{en}} th_{en}.\Delta t \quad (10)$$

In (10), $th_{dw}$is the maximum throughput during dwelling and Δt is the time resolution for calculating snr.

B. Model for Passing Stations

For non-stopping stations, the related equation is simpler than previous section, as the train only passes the station with a constant speed (i.e. $a = 0$) with no stop:

$$t_r^{ps} = 2t_{ps} \quad (11)$$

where $t_{ps} = \frac{d_{max}}{V_{ps}}$, if the train speed during passing is supposed as V_{ps}.

By substituting $d = V_{ps}.t + d_0$ in (4):

$$rss(t) = P_{r_0} - 10\gamma log(V_{ps}t + d_0) + X_\sigma \quad (12)$$

Hence, model for non-stopping stations will be as follows:

$$A_{passing} = 2\sum_{t=0}^{t_{ps}} th_{ps}.\Delta t \quad (13)$$

where $th_{ps}$is obtained through similar steps in (6)-(9).

C. Total Model for Offloading in a Rail Network

Based on equations in (10) and (13), for a train with N_{stp}stopping and N_{pss}passing stations through its trip, the total model will be obtained as follows:

$$A_{total} = \sum_{i=1}^{N_{stp}} (th_{dw}^i.t_{dw}^i + 2\sum_{t=0}^{t_{max}^i} th_t^i.\Delta t^i) + 2\sum_{j=1}^{N_{pss}} \sum_{t=0}^{t_{max}^j} th_t^j.\Delta t^j \quad (14)$$

In (14), i refers to the station number for stopping stations and $1 \leq i \leq N_{stp}$. Similarly, j is to the station number for passing stations and $1 \leq j \leq N_{pss}$.

V. SIMULATION

To validate our developed analytical model, we compare the results of our analytical model with the results obtained from Omnetpp version 5.4.1, which is one of the most powerful open source tool for network simulations. We assume that an IEEE802.11ac-based WLAN is dedicated for data offloading at each train station. IEEE802.11ac Standard can support channel with different bandwidths including 20MHz, 40Mhz, 80MHz and 160MHz. Therefore, we firstly, assume that only 20 MHz channels are available for data offloading, as the worst case scenario. In this case, we obtain the results for different values of transmitter powers and path loss components in both analytical and simulation environments. Then, we estimate the maximum capacity of data offloading for both types of stopping stations and passing stations.

Assuming that the dwelling time of a train is permissible to vary between 20 to 60 seconds, we set $t_{dw} = 20sec$, as the minimum guarantied value and the worst case scenario for data offloading at stopping stations. For acceleration of trains during reaching and leaving a station, we suppose a similar value of $1m/s^2$ with negative and positive signs, respectively. For PLE, we apply different values including 2 (as free space), 2.5, 3 and 3.5 to show the performance of our model for stations in different environments. This is a realistic assumption for a rail network, as every station might be located in places with different environmental conditions. Additionally, we apply different values of transmitter power including $20mw$, $30mw$, $40mw$ and $50mw$ as one of the effecting element in the test results. However, due to the similarities and to avoid repeated figures, we only illustrate the results of some selected scenarios. For the all case studies, the background noise level at receiver is set to $-90dBm$. We also set the MAC efficiency to 44 percent in our analytical model, which is directly obtained from simulation results in Omnetpp. Additionally, to avoid generating results with stochastic elements, we have not considered the shadowing effect at stations as this effect causes random elements at every simulation which is not the scope of this work.

Figure 3 shows SNR versus time for different environments and transmitter powers for both analytical and simulation results. As illustrated, the analytical model can accurately follow the simulation model and can achieve up to 98.67 percent with reference to the simulation results from Omnetpp.

Figure 4 shows the estimated throughput versus time for different values of PLE and transmitter power. The differences between empirical and simulation results are due to employ different data rate control methods. We applied RBAR method in our analytical model to theoretically estimate the maximum data throughput. However, for simulations in Omnetpp, we used AARF algorithm. Additionally, this figure shows the dependency of transmitter power and station environment on data throughput.

To estimate the upper and lower bounds of offloading capacity based on the empirical model, we have illustrated the capacity for different environments for a given transmitter
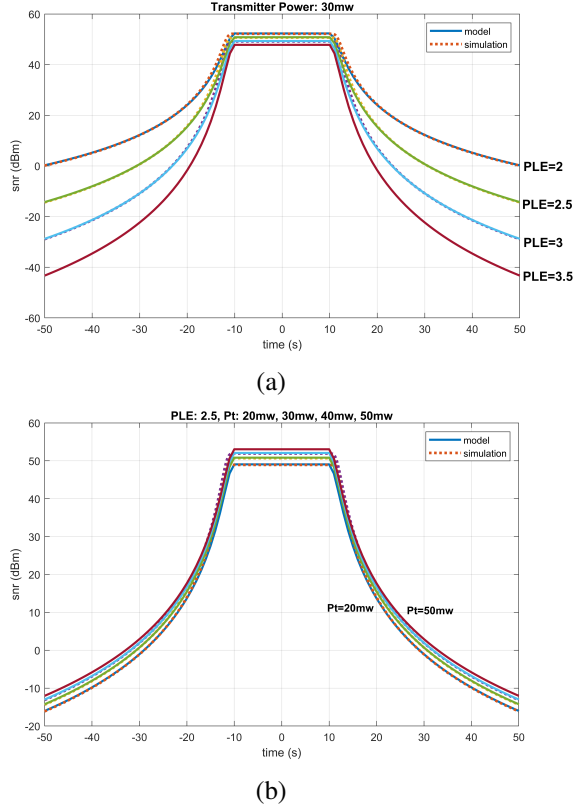
(a)

(b)

Fig. 3. SNR vs. time for different values of a) PLE and b) transmitter power

(a)

(b)

Fig. 4. Throughput vs. time for different values of a) PLE and b) transmitter power

power of $30mw$ in Figure 5. For the lower bound, we assume that only wireless channels with 20MHz bandwidth and one spatial stream is available and $GI = 800ns$. However, for the upper band, we have supposed 160MHz channels, 3 spatial streams and $GI = 400ns$. According to these assumptions, for example for an environment with PLE=2.5, we can theoretically achieve up to 4.42 GB and 1.85 GB offloading capacity for every stopping or passing stations, respectively (Figure 5).

It is imperative to re-emphasize that in the all above results, we have assumed IEEE802.11ac protocol for the physical layer of WLAN. However, by employing new coming standards such as IEEE802.11ay with 100 Gb/s, we can offload much more data using our proposed scheme. With such massive data rate, we can roughly achieve up to 232GB data offloading per each stopping station, in theory. This amount of data offloading is roughly obtained via multiplying the ratio of maximum theoretical data rates of IEEE802.11ay and IEEE802.11ac by the maximum upper bound in Figure 5.

VI. CONCLUSION

In this work, we proposed the existing WiFi networks at stations as intermediate hot spots for delayed offloading of big sensor data from trains to data centers. As the proposed method uses the existing WLAN at stations for data offloading, we will not need to install any extra communication infrastructure as RSU's. We developed an analytical model for the offloading task that can estimate the offloading capacity
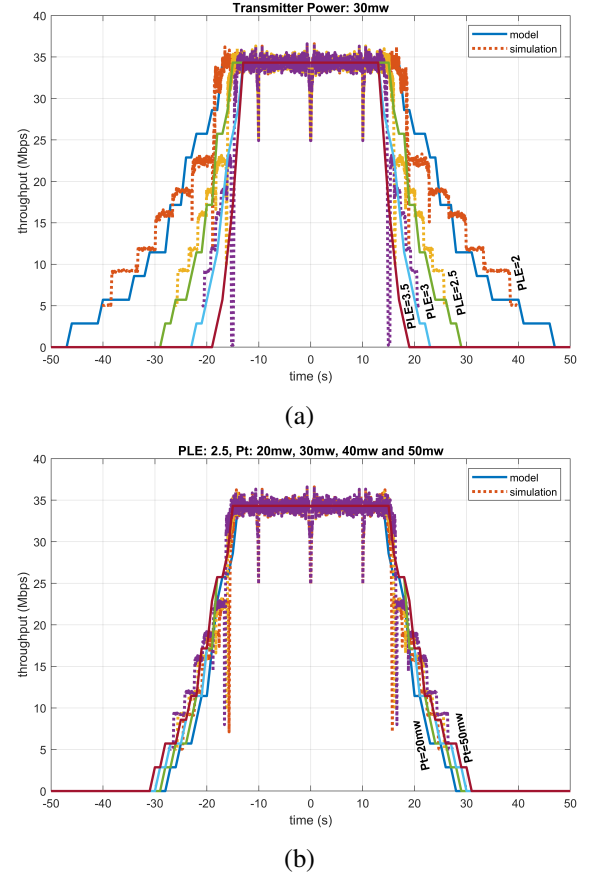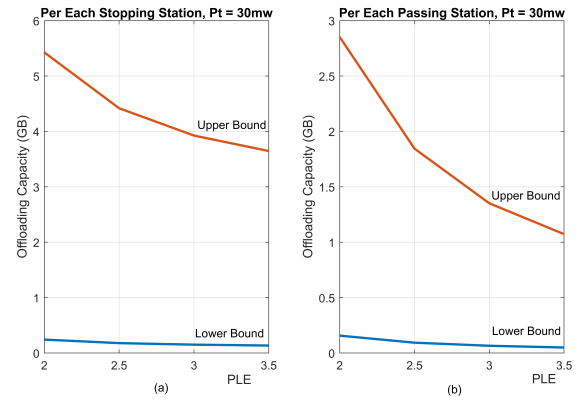

Fig. 5. The theoretical lower and upper bounds of offloaded data estimated by the empirical model for different environments at $P_t = 30mw$: a) for stopping stations, b) for passing stations

for passing as well as stopping stations. Simulation results showed an accuracy of 98.67 percent for our developed model. Additionally, by using the proposed station-based offloading scheme, we can theoretically offload up to 5.43 GB with current offloading standards and several hundreds of GB with the future ultra-fast offloading technologies such as IEEE802.11ay.

In this paper, we did not affect the location of antennas in the offloading model assuming that there are enough antennas along the trains and stations. Therefore, in our future work, we will make the model more realistic by considering the antennas effects on the offloading capacity. We have also supposed a dedicated WLAN at every station that are always available for data offloading for any train. However, WLAN may be unavailable in some situations due to poor signal conditions, lack of free channels because of other communications, etc. Hence, in our future work, we will consider alternative methods such as train-to-train communications in case of WiFi unavailability at stations.

REFERENCES

[1] E. Fumeo, L. Oneto, and D. Anguita, "Condition based maintenance in railway transportation systems based on big data streaming analysis," *Procedia Computer Science*, vol. 53, pp. 437–446, 2015.

[2] V. J. Hodge, S. O'Keefe, M. Weeks, and A. Moulds, "Wireless sensor networks for condition monitoring in the railway industry: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1088–1106, 2015.

[3] H. Zhou, H. Wang, X. Li, and V. C. M. Leung, "A survey on mobile data offloading technologies," *IEEE Access*, vol. 6, pp. 5101–5111, 2018.

[4] D. Suh, H. Ko, and S. Pack, "Efficiency analysis of wifi offloading techniques," *IEEE Trans. Vehicular Technology*, vol. 65, no. 5, pp. 3813–3817, 2016.

[5] H. Deng and I. Hou, "On the capacity-performance trade-off of online policy in delayed mobile offloading," *IEEE Transactions on Wireless Communications*, vol. 16, pp. 526–537, Jan 2017.

[6] S. Kashihara, "Transmission scheduling method for delivery of large-sized data object based on delay constrained request over scheduled transportation vehicles," *EPI International Journal of Engineering (EPI-IJE)*, vol. 1, no. 1, pp. 60–68, 2018.

[7] C.-M. Huang, M.-S. Chiang, D.-T. Dao, H.-M. Pai, S. Xu, and H. Zhou, "Vehicle-to-infrastructure (v2i) offloading from cellular network to 802.11p wi-fi network based on the software-defined network (sdn) architecture," *Vehicular Communications*, vol. 9, pp. 288 – 300, 2017.

[8] D. Li, W. Daamen, and R. M. Goverde, "Estimation of train dwell time at short stops based on track occupation event data: A study at a dutch railway station," *Journal of Advanced Transportation*, vol. 50, no. 5, pp. 877–896, 2016.

[9] S. M. Ashley and A., "Official ieee 802.11 working group project timelines," *Online*, 2016.

[10] "Ieee standard for information technology–telecommunications and information exchange between systems local and metropolitan area networks–specific requirements - part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications," *IEEE Std 802.11-2016 (Revision of IEEE Std 802.11-2012)*, pp. 1–3534, 2016.

[11] S. Fu, Y. Zhang, M. Ceriotti, Y. Jiang, M. Packeiser, and P. J. Marrn, "Modeling packet loss rate of ieee 802.15.4 links in diverse environmental conditions," in *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, April 2018.

[12] J. Luomala and I. Hakala, "Effects of temperature and humidity on radio signal strength in outdoor wireless sensor networks," in *2015 Federated Conference on Computer Science and Information Systems (FedCSIS)*, pp. 1247–1255.

[13] T. Vanhatupa, "Wi-fi capacity analysis for 802.11 ac and 802.11 n: Theory & practice," *Ekahau Inc*, 2013.

[14] S. Wang, T. Lei, L. Zhang, C.-H. Hsu, and F. Yang, "Offloading mobile data traffic for qos-aware service provision in vehicular cyber-physical systems," *Future Generation Computer Systems*, vol. 61, pp. 118–127, 2016.