Markov Model as Approach to Parking Space Occupancy Prediction

Lea Dujić Rodić[1], Toni Perković[1], Tomislav Županović[1] and Petar Šolić[1]

***Abstract*— One of the most important infrastructures that enable IoT-based Smart Cities is Smart Parking. This paper introduces a machine learning technique based on Hidden Markov Model that applies a Viterbi algorithm for predicting occupancy lot status based solely on collected LoRa RSSI and SNR values.**

***Index Terms*— Smart Parking, HMM, Viterbi, LoRa, RSSI, SNR**

I. Introduction

The Internet of Things (IoT) vision of ubiquitous and pervasive connection of smart things gives rise to a future environment composed out of physical and digital world. In this environment it is possible to receive information about or form the psychical world that was previously not available to us, and moreover, interconnect it to exchange and use this information with the digital world [1]. The IoT applications are being employed in diverse areas of industry, communication, wireless sensor networks, data mining, assisted living, etc. giving rise to the concept of Smart City.

Smart and Sustainable Mobility is one of the central concepts in the vision of the Smart City, where IoT plays an important role [2], [3]. In urban city areas, due to the rise of cars, existing parking systems are inadequate or unable to handle parking loads [4]. Moreover, parking facilities are not accessible in a adequate manner, since it is estimated that drivers spend around 7.8 minutes in finding free parking lots [5]. Studies have shown that in traffic dens environments in urban areas 30- 50% of drivers are in search for free parking [6]. One of the major issue that arise from this is the increase of fuel consumption and air pollution.

The development of dynamic and complex IoT systems has been followed by the exponential growth of big data analytics [7]. Along with it arose Machine Learning (ML) holding a vast potential for data analysis and precise predictions made from the past observations for given new measurements. A commonly and widely used model for sequential or time series data in ML and statistics is the Hidden Markov Model (HMM) [8]. HMM are based on the concept of Markov Chain which can represent any random sequential process that undergoes transitions from one state to another [9]. In the last two decades, HMM has been used in various areas as a data-driven modeling approach in automatic speech recognition, pattern recognition, signal processing, telecommunication, bioinformatics etc. [10]. With that regard, IoT applications that use sensor technology, RFID technology, network communication, data mining and machine learning could prove to be quite efficient in solving the above problem of free parking space [11].

Existing Smart Parking solutions for detecting occupancy include usage of adequate sensing technologies and transmission to a centralized system for further processing (using appropriate radio technology such as LoRa, NB-IoT, Sigfox, BLE5, etc.). Such devices use detection techniques based on sensors such as light, magnetometer, infrared detector, distance sensors or a combination of sensing technologies [12], [13], [14], [15]. However, these solutions are rather power hungry due to the consumption of a large number of sensors, microcontrollers and radio communication peripherals, which impact the lifetime of an otherwise battery-powered device. Recently, some papers started to investigate the possibility of using the simple premise: in electromagnetically harsh environment, if a vehicle occupies a parking lot, the signal strength on the receiver side will be significantly reduced [16], [17]. This way, the device will become a simple beacon device (without any sensor) where occupancy is detected with change in signal strength. In this paper, the data about occupancy and signal strength was collected from five LoRa-based Smart Parking sensor devices via three LoRaWAN gateway devices. Using techniques based on the Hidden Markov Model (HMM) it is possible to estimate parking space occupancy based on signal strength with a probability of 96 %.

II. LoRa-based Smart Parking Sensor Device

In this paper LoRa radio technology was employed for transmitting information about parking lot occupancy. As a a representative of a Low-Power Wide Area Networks (LPWANs), LoRa allows battery-enabled devices such sensors to communicate low throughput data over long distances.

The core of a Smart Parking sensor device is a commercial Libelium LoRaWAN-based Smart Parking sensor device that comprises radar and magnetometer sensor for parking lot occupancy detection. Figure 1 depicts architecture of implemented LoRaWAN parking mechanism. In the implementation of Libelium LoRaWAN parking sensor the device periodically wakes up (every 60 seconds) and activates internal sensor devices (such as radar, magnetometer) for checking the change in parking status. If parking lot status changes (from free goes to occupied or from occupied to free), sensor device initiates message transmission via LoRaWAN protocol over a radio to the gateway. Five Libelium parking sensor devices were placed at faculty parking lots, while three The Things Network gateways were placed within close vicinity of parking sensor devices. Once the message arrives

[1]L. Dujić, T. Perkovic, T. Županović and P. Solic are with Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, University Split, 2100 Split, Croatia {`dujic, toperkov, tzupan01, psolic`}`@fesb.hr`

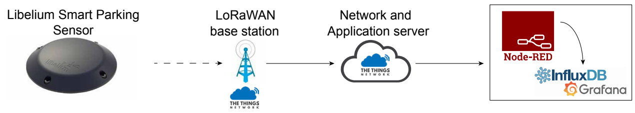

Fig. 1: Network architecture of Libelium Smart Parking sensors.

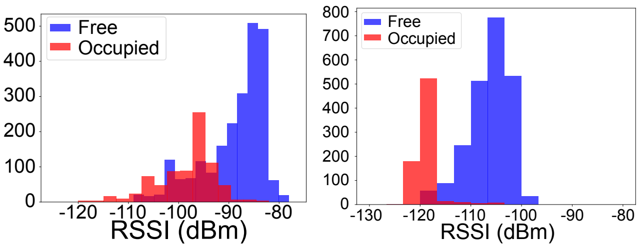

Fig. 2: Histograms of RSSI values for Smart Parking Sensor 2 for free and occupied parking status from (left) Gateway 1 and (right) Gateway 3

to the base station, it is forwarded to the TTN Network and Application server, where the message payload is decoded and prepared for further processing and forwarding using MQTT protocol or HTTP integration. In given implementation, node-red was used for message aggregation, parsing, while InfluxDB was used for storage.

III. Data Analyses and Visualization

Data has been collected from five different sensors and three different gateways in period from 13th of December 2019 until 25th of February 2020. It comprised timestamp, gateway ID, occupancy status (free - 0, occupied - 1), Received Signal Strength Indicator (RSSI) in dBm, Signal to Noise Ratio (SNR) for every gateway. Overall, 130488 raw data was collected from all five sensors. Histograms presented in Figure 2 give an illustration of how Occupancy Status is related with RSSI and SNR[1].

Firstly, it was noticed that parking lots are free considerably more than they are occupied. This is an important property of the parking place indicating its stochastic behaviour. The parking is located on University grounds, and therefore usually free during night time or over the weekend periods. Secondly, from the histograms it was seen that the RSSI values for free and RSSI values for occupied parking status overlap in Gateway 1 (GW1) and Gateway 2 (GW2). The same reasoning applies for SNR values for all sensors in GW1 and GW2. However, result gained for Gateway 3 (GW3) and sensors 2, 3, 4 differ from above mentioned. The aforementioned sensors gave the least overlapping of RSSI and SNR values for a particular occupancy state. It was also noticed that that higher RSSI values indicate a free parking space, while lower indicate occupied one. Difference between the results gained for different gateways could be a consequence of the distance of GW1 (30 m), GW2 (75m) and GW3 (145m) form the parking sensor. GW3 is furthest away and on top of the University building and indoor. This would imply that the closer the gateway is, the channel influences RSSI and SNR stronger than the change of the parking status. Further analyses showed that when parking status does not change, the values of RSSI and SNR change very little or not at all. However, when the parking status does change there is a significant change in RSSI and SNR values.

In light of the above reasoning, conclusions were twofold: 1) RSSI, SNR and occupancy status are considerably correlated and 2) the adequate ML algorithm must be able to comprise the complexity of the data correlations in order to provide appropriate prediction of occupancy status.

IV. Markov Model as a Machine Learning Approach to Prediction

A. Hidden Markov Model

Hidden Markov Models (HMMs) have been known for decades and today are making a large impact with regard to their applications, especially in form of Machine Learning models and applications in reinforcement learning. Markov Chains and process were introduced by the Russian mathematician Markov in 1906 when he obtained a theoretical result for a stochastic process. Markov process can be considered a time-varying random phenomenon for which Markov properties are attained. Its practical importance is the use of the hypothesis that the Markov property holds for a certain random process in order to build a stochastic model for that process [10].

In the broadest sense, a Hidden Markov model (HMM) is a Markov process that can be divided into two parts: an **observable** component and an unobservable or **hidden** component. The observation is a probabilistic function of the state, i.e. the resulting model is a doubly embedded stochastic process, which is not necessarily observable, but can be observed through another set of stochastic processes that produce the sequence of observations. A machine learning algorithm can apply Markov models to decision making processes regarding the prediction of an outcome. In 1986 Rabiner and Juang [18] gave the structure of the first order Hidden Markov Model denoted as $\lambda(A, B, \pi)$, where $A = \{a_{ij}\}$is the matrix of transition probabilities, $B = \{b_j(k)\}$ is the matrix of observation probability distribution in each state and π is the initial state distribution.

Furthermore, one of the central issue of HMM is to fine solution to the decoding problem i.e. to find the optimal sequence of states to a given observation sequence and model used [19]. Most common method to this is by using the **Viterbi algorithm**, introduced by Andrew Viterbi in 1967 as a decoding algorithm for convolution codes over noisy digital communication links. It is the answer to the decoding problem resulting in the Viterbi path, since the algorithm can be interpreted as a search in a graph whose nodes are formed by the states of the HMM in each of the time instant [10].

The Viterbi algorithm finds single best state sequence $q = (q_1, q_2, ..., q_T)$ for the given model and observations. The probability of observing $o_2 o_2 ... o_t$ using the best path that

[1]The data that support the findings of this study are not publicly available due to restrictions that could compromise the privacy of collected data.

ends in state i at the time t given the model λ is:

$$\delta_t(i) = \max_{q_1, q_2, \ldots, q_{t-1}} \mathbb{P}(q_1, \ldots, q_{t-i}, q_t = i, o_2, \ldots, o_t \mid \lambda). \quad (1)$$

$\delta_{t+1}(i)$ can be found using induction as:

$$\delta_{t+1}(i) = b_j(o_{t+1}) \max_{1 \leq i \leq N} [\delta_t(i) a_{ij}] \quad (2)$$

To return the state sequence, the argument that maximizes Equation (2) for every t and every j is stored in a array $\psi_t(j)$. It is important to point out that the Viterbi algorithm can be implemented directly as a computer algorithm and it succeeds in splitting up a global optimization problem so that the optimum can be computed recursively: in each step it maximizes over one variable only, rather than maximizing over all n variables simultaneously.

With regards the first order Markov model, if the past and the present information of the process is known, the statistical behaviour of the future evolution of the process is determined by the present state. Thus, the past and the future are conditionally independent [20]. Therefore, it is reasonable to ask can there be a model which can gather and somewhat keep information from the past. The answer lies within a higher-order Markov models, where the hidden process is a higher order Markov chain and it is dependent on previous states. This gives memory to the model and such a modeling is more appropriate for processes in which memory is evident and important, for example a stock market time series.

B. Model and Results

Hidden Markov model of second order, presented in the following, was designed and used to predict the occupancy status of a parking space, using RSSI and SNR values.

From previously presented and discussed histograms of RSSI and SNR values with regards to occupancy status, it was concluded that when parking status does not change, the values of RSSI and SNR change very little or in most cases not at all. In contrast, when the parking status does change there is a significant change in RSSI and SNR values. Therefore the variables "bring memory" with them that is dependent of the previous state of occupancy. The process it self is of a time series which can be designed and modeled using a second-order HMM. In this model the **Hidden States** are the aim of prediction which is **Occupancy status**. To "bring memory" into our model the **Observable (Visible) States** are defined to be the changes of RSSI (the same reasoning and model applies for SNR) values calculated as the difference between these values form two previous states.

The notation and model illustration is as follows:

- F- free, O - occupied.
- $RSSI_s$- value of Received Signal Strength in occupancy state s in a timestamp.
- $\triangle RSSI = RSSI_s - RSSI_{s+1}$. Difference between two values of RSSI in two consecutive occupancy states.
- FF- state that is free which previous state was free.
- FO- state that is occupied after previously being free.
- OF- state that is free after being occupied
- OO- state that is occupied which previous state was occupied.

States FF, FO, OF and OO bring with them "memory of occupancy" since they remember what was the occupancy status form the "past". These states represent Hidden Layer of states. The Hidden Markov Model model is denoted as $\lambda(A, B, \pi)$, where:

- the probability of transition from one state to another is given with transition matrix A which holds some zero values. This is due to the fact that some transitions are impossible. For instance, you cannot transit from state FF into the state OO,
- initial state distribution (stationary distribution) is denoted as π
- B is the matrix containing the observation probability distribution in each state. In this model, the **observations** are the **changes** of $RSSI$ values in two consecutive occupancy states - $\triangle RSSI$.

Figure 3 visualizes its architecture.

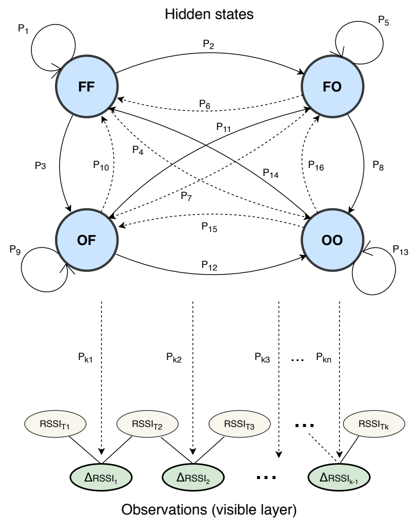

Fig. 3: Illustration of second-order Hidden Markov model for detecting occupancy status based on change of RSSI values

The implementation of the model was done using Python 3 and The Jupyter Notebook. As stated, since it was decided to extract relevant data for each sensor and each of the three different gateways separately, the code had taken into the account all of these possibilities. The used decoding algorithm for finding the optimal sequence of states to a given observation sequence and model is previously defined Viterbi algorithm. All of the data was effectively used a as observation for a chosen step and given as an input to the Viterbi algorithm. The chosen step determines the length of observation sequence. For example, if the chosen step is 4, the whole data set from selected sensor and gateway is divided into subsets of sequences containing four consecutive values of a chosen variable (RSSI or SNR). Every one of this sequences is then given as an observation input to the Viterbi algorithm.

TABLE I: Table of best results obtained for each gateway

Gateway	Variable (sensor number)	Prediction accuracy (best results)
GW1	RSSI (4)	81%
GW2	RSSI (4)	91%
GW3	SNR (2)	96%

The predicted and the true values are stored separately and accuracy score is calculated using *accuracy score* function. This function computes subset accuracy, which is the fraction of samples predicted correctly. The set of labels predicted for a sample must exactly match the corresponding set of labels of true values. Moreover, model's evaluation is done using Mean Absolute Error (MAE). The model was tested for all variables from all sensors and gateways and the best results are given in Table I. As can be seen from the table, the least promising results were gained form the closest gateway GW1. This is due to the previously explained overlapping in the RSSI (or SNR) values with regards to different occupancy status. With regards to GW2, the best results are also consistent with the reasoning of the least overlapping values for different occupancy status.

Finally, the best results were obtained for furthest GW3, giving 96% accuracy for observation values of SNR from sensor 2 and giving 95% accuracy for observation values of SNR from senor 4. Figure 4 illustrates the best result that is obtained using the model.

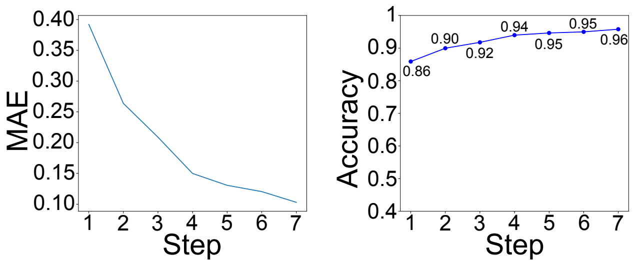

Fig. 4: MAE and prediction accuracy for SNR values from sensors 2 from Gateway 3.

V. Conclusion

This paper presents a novel concept of cost-effective and low-power sensor that achieves free parking occupancy prediction using Hidden Markov model of second order. Namely, LoRa radio technology was employed for transmitting information about parking lot occupancy via Smart Parking sensor devices and three LoRaWAN gateway devices. The analysis of collected data showed correlation between RSSI, SNR and Occupancy status. It was further shown that free parking can be estimated with high accuracy of 96% from signal strength, along with HMM as a Machine Learning technique. Future work will comprise an improvement on the accuracy of presented model using other appropriate ML techniques as well as experiments with a different number of devices. Furthermore, a concept of a novel, low-power, LoRa-based, cost-effective system will be examined.

Acknowledgment

This work was supported by the Croatian Science Foundation under the project "Internet of Things: Research and Applications", UIP-2017-05-4206.

References

[1] M. Zorzi, A. Gluhak, S. Lange, and A. Bassi, "From today's intranet of things to a future internet of things: A wireless-and mobility-related view," *IEEE Wireless Commun.*, vol. 17, pp. 44–51, 12 2010.

[2] T. Perković, S. Damjanović, P. Šolić, L. Patrono, and J. J. Rodrigues, "Meeting Challenges in IoT: Sensing, Energy Efficiency and the Implementation," in *Fourth International Congress on Information and Communication Technology in concurrent with ICT Excellence Awards (ICICT 2019)*, 2019.

[3] S. Nizetic, N. Djilali, A. Papadopoulos, and J. R. J., "Smart technologies for promotion of energy efficiency, utilization of sustainable resources and waste management," *Journal of Cleaner Production*, vol. 231, pp. 565–591, 2019.

[4] A. Singh, A. Kumar, A. Kumar, and V. Dwivedi, "Radio frequency global positioning system for real-time vehicle parking," in *2016 International Conference on Signal Processing and Communication (ICSC)*, 2016, pp. 479–483.

[5] T. Perković, P. Šolić, H. Zargariasl, D. Čoko, and J. J. Rodrigues, "Smart parking sensors: State of the art and performance evaluation," *Journal of Cleaner Production*, p. 121181, 2020. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0959652620312282

[6] V. Paidi, H. Fleyeh, J. Håkansson, and R. Nyberg, "Smart parking sensors, technologies and applications for open parking lots: A review," *IET Intelligent Transport Systems*, vol. 12, 04 2018.

[7] J. Qin, "Process data analytics in the era of big data," *AIChE Journal*, vol. 60, 09 2014.

[8] J. Van Gael, Y. Saatci, Y. W. Teh, and Z. Ghahramani, "Beam sampling for the infinite hidden Markov model," in *Proceedings of the International Conference on Machine Learning*, vol. 25, 2008.

[9] M. K. Mustafa, T. Allen, and K. Appiah, "A comparative review of dynamic neural networks and hidden markov model methods for mobile on-device speech recognition," *Neural Computing and Applications*, vol. 31, no. S-2, pp. 891–899, 2019. [Online]. Available: https://doi.org/10.1007/s00521-017-3028-2

[10] G. L. Kouemou, "History and theoretical basics of hidden markov models," in *Hidden Markov Models*, P. Dymarski, Ed. Rijeka: IntechOpen, 2011, ch. 1. [Online]. Available: https://doi.org/10.5772/15205

[11] Y. Yin and D. Jiang, "Research and application on intelligent parking solution based on internet of things," in *2013 5th International Conference on Intelligent Human-Machine Systems and Cybernetics*, vol. 2, 2013, pp. 101–105.

[12] T. Lin, H. Rivano, and F. Le Mouël, "A survey of smart parking solutions," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 12, pp. 3229–3253, 2017.

[13] T. N. Pham, M. Tsai, D. B. Nguyen, C. Dow, and D. Deng, "A cloud-based smart-parking system based on internet-of-things technologies," *IEEE Access*, vol. 3, pp. 1581–1591, 2015.

[14] H. G. Jung, Y. H. Cho, P. J. Yoon, and J. Kim, "Scanning laser radar-based target position designation for parking aid system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 9, no. 3, pp. 406–424, 2008.

[15] F. Al-Turjman and A. Malekloo, "Smart Parking in IoT-enabled Cities: A Survey," *Sustainable Cities and Society*, vol. 49, p. 101608, 2019.

[16] P. Solic, R. Colella, L. Catarinucci, T. Perkovic, and L. Patrono, "Proof of Presence: Novel Vehicle Detection System," *IEEE Wireless Communications*, 2019. [Online]. Available: http://dx.doi.org/10.1109/MWC.001.1900133

[17] T. Perković, P. Šolić, H. Zargariasl, D. Čoko, and J. Rodrigues, "Smart parking sensors: State of the art and performance evaluation," vol. 262, p. 121181, 2020.

[18] L. Rabiner and B. Juang, "An introduction to hidden markov models," *IEEE ASSP Magazine*, vol. 3, no. 1, pp. 4–16, 1986.

[19] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[20] X. Xi, "Further applications of higher-order markov chains and developments in regime-switching models," 2012.