Enhanced Data-Driven LoRa LP-WAN Channel Model in Birmingham

AlaaAllah ElSabaa
Department of Engineering
Birmingham City University
Birmingham, UK
alaaallah.elsabaa@mail.bcu.ac.uk

Florimond Guéniat
Department of Engineering
Birmingham City University
Birmingham, UK
florimond.gueniat@bcu.ac.uk

Wenyan Wu
Department of Engineering
Birmingham City University
Birmingham, UK
wenyan.wu@bcu.ac.uk

Michael Ward
Department of Engineering
Birmingham City University
Birmingham, UK
michael.ward@bcu.ac.uk

***Abstract*—Innovative solutions providing better coverage and minimized power consumption by end nodes such as Low Power Wide Area Networks (LP-WAN) have facilitated the advances towards improved IoT connectivity. Long Range Wide Area Network (LoRaWAN) technology stands out as one leading platform of LP-WANs receiving vast attention from both industry and academia. Performance evaluation of LoRaWAN is promising, in particular in the field of outdoor localization and object tracking. Limitations of node ranging and tracking without the need of energy-draining solutions like GPS, however, has not been tackled thoroughly. In this work, we explore the performance of the LoRa LP-WAN technology using real-life measurements in Birmingham, UK, using commercially available equipment. We present a channel attenuation model that can be utilized to estimate the path loss in 868 MHz ISM band in urban-similar areas. The proposed channel model is then compared to previously well-identified empirical path loss models and enhanced by detecting and eliminating outlier data from the obtained real measurements for an optimal fitted model. We, further, propose a novel RSSI distribution-based and k-means clustering to enhance the power-to-distance prediction accuracy that improves absolute errors by 4% and 18%.**

***Index Terms*—Internet of Things, LoRa, LoRaWAN, Channel Models, Outliers , Distribution Clustering.**

I. Introduction

With the exponential growth of Internet of things (IoT) devices connectivity on the road of digitalizing everyday life, existing communication technologies are faced with new challenges to fulfill the life quality promised. Those challenges include assured connectivity with low-power demand to increase the devices' lifetime, long range, low data rate and cost-effectiveness. The recent advances in Low-Powered Wireless Area Networks (LPWAN) technologies such as Sigfox [1] and Long Range (LoRa) [2] offering optimised device cost, higher coverage when compared to LTE and low-power utilization, has facilitated the maturation of Smart Cities.

A LoRaWAN network consists mainly of a gateway (GW), end nodes and a server. End nodes are embedded with LoRa modules to communicated wirelessly to the GWs. Gateways, then, collect the received packets and forward them to server through an IP network, as shown in fig 1. The LoRaWAN communication protocol embed its own LoRa modulation [3] at the physical layer. Based on CSS (Chirp Spread Spectrum) [4] mechanism, LoRa modulation spreads the narrow band signal over a larger frequency band. It results in a robust signal transmission, a long lifetime of battery powered end devices, and low cost installation. Data rates lie between 0.3 kbit/s and 50 kbit/s [5]. The actual range depends on the data rate and frequency used as well as on the propagation conditions found on the installation site [6].

In [7], authors presented a theoretical evaluation of the capabilities and the limitations of LoRaWAN. Capacity and scalability studies were presented in [8], [9]. These work highlighted that LoRaWAN systems should be configured thoroughly to reach a proper tradeoff between scalability and efficiency [10]. An essential phase when designing a Smart City network is the knowledge of the propagation characteristics. The knowledge of the expected losses a signal will undergo plays an important role in an optimal network design [11]. With the correct consideration, better coverage, lower costs and acceptable quality can be achieved. For this purpose, empirical path loss models are used to offer accurate predictions for the possible received signal strengths [11]. While the presence of established empirical path loss model such as; Okumura Hata model, ITU Advanced, Cost 231-Hata ,etc, provide a basis for the channel propagation. Yet, depending on the type of environment the Smart City is located in, these channels models may lack accuracy, as their independent parameters such as reflection, diffraction and scattering factors need to be tuned or corrected.Therefore, various LoRa-based smart networks have recently experimented and formalized their own path loss model such as, Oulu, Finland [11] and Dormund, Germany [12]. Due to the nature of the city, geographical location and the surrounding terrain, both researches proposed their own propagation model for further usage of LoRaWAN on their grounds and for similar environment.

Therefore, to examine and evaluate the requirements of deploying an urbanized Smart City communication network, Birmingham City University (BCU) has set up an LPWAN within Birmingham, UK, using LoRa as a network technology.

In this work, we test various published path-loss models and inspect their effectiveness against real signal measurements from our LoRa test deployment within Birmingham. While our

978-1-6654-8453-4/22/$31.00 ©2022 Crown

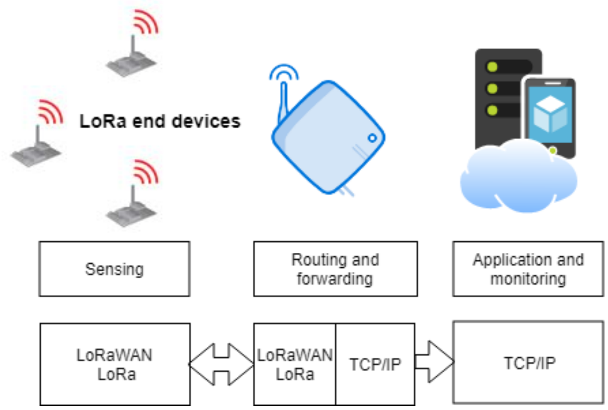


Fig. 1: Star topology of LoRaWAN

results present a new model for our area, we further expand our work to include clustering as a

The model also, embeds outliers detection and elimination before verifying and finalizing our initial path loss model. We extend our work to include clustering algorithms to investigate their influence on power to distance estimation.

The main contributions of this paper are;

- Critical analysis of three established models to compare and evaluate their performance accuracy in Birmingham.
- Presenting an initial model based on real measurements to represent a well defined model for an optimal fit.
- Proposing a RSSI K-mean and DBC clustering for enhancing accuracy and comparing it to all models mentioned.

The remainder of this paper is organized as follows. Related and recent work is discussed in Section II. Section III describes the deployment setup of our experiment. Sections IV and V provide performance measurement results and the proposed enhanced channel model. At last, Section VI concludes the paper.

II. LP-WAN: LoRa Related Work

A. Background

Previous experimental work using LoRaWAN such as [13], tested the technology capabilities in the most rural area of Oulu, Finland. The work done utilized commercially available equipment, where the end node was either stationed on top of a car (for on-land measurements) or attached to the radio mast of a boat. Real measurements were reported to the gateway installed. Under different circumstances considering high density urban areas, authors in [14] evaluated the performance of LoRaWAN in Melbourne's Central Business District. Their initial measurements showed a guaranteed and successful communication only within a radius of approximately 200 meters from the gateway. The communication totally fails around 600 meters revealing the challenging task of network deployment in high density urban areas. For better understanding of LoRAWAN limitations, the authors then enhanced the gateway coverage using a higher building installation in a medium density urban area. This experiment resulted in an improved LoRa propagation reaching up to 5.8 km.

Fundamental LoRaWAN empirical work testing the signal coverage and packet loss ratio are found in [13] and [15]. In [16], authors extended the work to test various modulation parameters on a single link of communication. LoRA modulation, and hence its data rate, depends mainly on the following factors.

Spreading Factor (SF): ranging between 7 to 12, this affects the slope of the variation of frequency and so the energy per chirp symbol. The higher the SF the longer the time taken to transmit a symbol, which in turn decrease the data rate yet increases the coverage range.

Bandwidth (BW): To modulate the LoRa data signal, the chirp has to range through f_{min} to f_{max}, also known as the chirp rate. BW can be set as 125 kHz, 250 kHz or 500 kHz.

Coding Rate (CR): It is the forward error correction rate (FEC). LoRa implements CR to increase the signal resilience against interference at the cost of introducing redundancy bits.

B. Path Loss and Channel Models

A fundamental principle when understanding the channel properties of RF propagation is the usage of the Received Signal Strength Indicator (RSSI) for ranging. RSSI reflects the difference between the transmitted and received power of a signal that is used to estimate the distance propagated. Another vital parameter is the Path Loss (PL), PL is defined the power taken by the signal to overcome the communication medium and can be calculated using the link budget equation:

$$P_{tx}(dBm) = RSSI(dBm) + G_{system}(dB) - PL(dB). \quad (1)$$

The link budget equation shown in eq. (1) relates all gains and losses from the transmitter, through all transmission mediums, to the receiver. $RSSI$ is the expected power at the receiver, P_{tx} the transmitted power, G_{system} is the system gains such as those associated with directional antennas for both the transmitter and/or the receiver. PL are losses due to the propagation channel, either calculated via a wide range of channel models or from empirical data.

1) Free Space Path loss (PL_{FS}): PL_{FS} (Friis equation [17]) model provides the baseline and most ideal path loss as it assumes a line-of-sight and obstacle-free connection. It is also the elemental base of all upcoming models. The equation is given in the logarithmic domain as follows:

$$PL_{FS} = 20\log_{10}(f) + 20\ \log_{10}(d) + 32.45. \quad (2)$$

Here and henceforth, the distance (d) is given in km, the frequency (f) in MHz and the path loss (PL_x) is in (dBm).

2) Okumura-Hata & Cost 231-Hata Path loss Model (PL_{OH}, PL_H): Calculating the path loss value is usually referred to as a prediction process given that the exact prediction is only possible for simpler cases, such as the previously mentioned PL_{FS} or for a flat-earth model. Practically, the path loss is modelled using experimental data incorporating statistical methods that are based on measured and averaged losses along typical environment of radio links. An anchor signal propagation model, namely Okumura-hata model, was originally the results of real data collected in Japan, for signals propagating at a frequency range of 150 to 2000 MHz [18]. The Okumura-hata model was also used in [19] as the authors simulated LoRa packets to predict RSSI in an urban area, where it outperformed other models. It extends the path loss model PL_{FS} by taking into consideration the antenna gain factor, propagation gains due to node and base-station antennas' heights, and a correction gain.

$$PL_{OH} = 39.55 + 26.16\log_{10}(f) - 13.82\log_{10}(h_{tx}) - a(h_{rx}) + (44.9 - 6.55\log_{10}(h_{tx}))\log_{10}(d), \quad (3)$$

and

$$PL_{H} = 46.3 + 33.9\log_{10}(f) - 13.28\log_{10}(h_{tx}) - a(h_{rx}) + (44.9 - 6.55\log_{10}(h_{tx}))\log_{10}(d) + C, \quad (4)$$

where the term $a(h_r)$ in eq.(3) and (4) is calculated as $a(h_r) = 3.2\log(11.75h_r)^2 - 4.97$ shows the correction factor for large cities and for $f \geq 400MHz$, [20]. h_{tx} and h_{rx} stand for the transmitter and the receiver antenna heights in meters, respectively. Cost 231-Hata model extends Okumura–Hata model for medium to small cities to cover the band [1500–2000] MHz, where C is 0 for medium cities and 3 for metropolitan areas [21].

3) Experimental Path loss Models: LoRa has been recently experimented in Oulu, Finland [11] and Dormund, Germany [12]. Due to the nature of the city, geographical location and the surrounding terrin, both research proposed their own suitable propagation model for further usage of LoRaWAN on their grounds and for similar environment. Authors in [19] used their measurements to compare various empirical models in search of the optimal model to represent thier data.

This work considers these models for performance analysis against real experimental data within Birmingham.

III. Experimental Setup

A. Stage I: Experimental Setting

1) LoRa Base-station and Node: In this work we utilize a LoRa off-shelf node "The Things Uno" (shown in fig. 3a). The Things Uno node, thereafter referred to as $Leo1$, is based on the Arduino Leonardo with added LoRaWAN module Microchip (RN2483 - class A protocol stack) [22]. We further enhance $Leo1$ with a GPS module for primary measurements. The Adafruit 746 GPS module uses MTK3339 GPS System on Chip (SoC) that can can track up to 22 satellites on 66 channels and has a sensitivity of -165 dBm.

$Leo1$ is used to measure the RSSI and SNR using LoRaWAN modulation operating at 868 MHz (within the ISM band in Europe) with a receiver sensitivity up to -148 dBm. The measurement system consists of a singular LoRa Multi-Connect Conduit IP67 Base Station (shown in Fig. 3b). While the TTN node is considered mobile and easy to carry around the vicinity of the campus, the LoRa MultiConnect Conduit is based indoors on campus. Table I lists the characteristics of both, the $Leo1$ node and the gateway used in our setup.

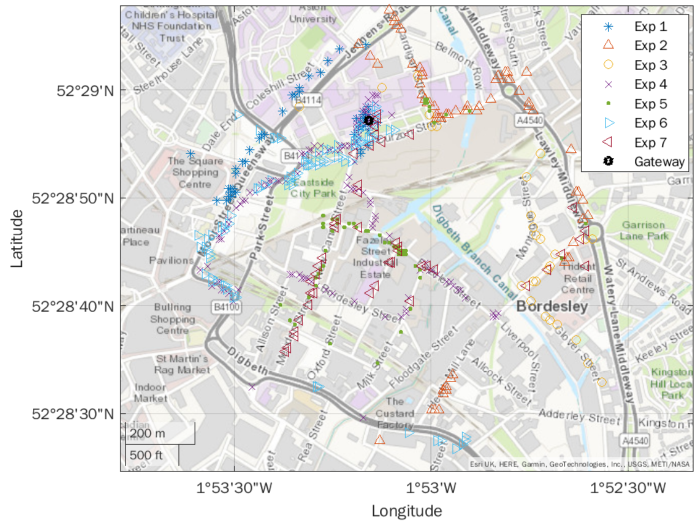
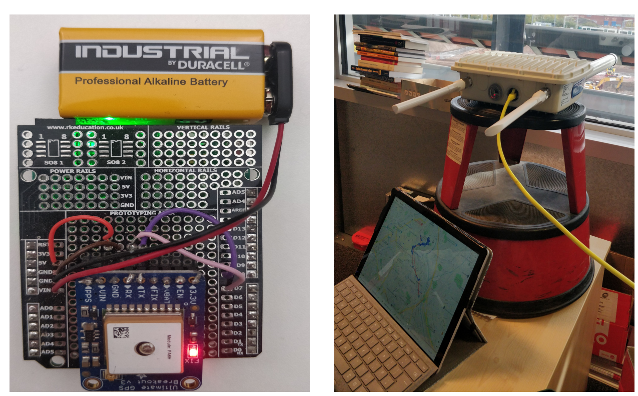


Fig. 2: Map showing the gateway and locations of the data collected from all experiments

TABLE I: Characteristics of MultiTech Gateway and The Things Uno node

Characteristic	Gateway	Node
Module	MTCDIP-LEU1-267A-868	The Things Uno
LoRa Chip		Microchip RN2483 LoRa
Operating Frequency	867.9	867.9
Modulation	LoRa	LoRa
SF	12	12
Coding Rate	4/5	4/5
Rx Input Sensitivity (dB)		-148
Tx Power	14dBm	
Rx current consumption		14.2 mA

2) Measurement Setup: The measurements took place in the city center of Birmingham, UK. Population of Birmingham is over 1.2 million and the maximum building height limit at 242 metres. During all measurements the position of the base station was fixed. Powered by a 9V battery, $Leo1$ was operating various conditions (travelling by car, bus along major roads, or by foot) to assure a better and thorough coverage. The $Leo1$ was periodically sending packets to the gateway, that includes a packet counter and GPS coordinates.



(a) (b)

Fig. 3: Field setup. a): $Leo1$ LoRa node. b): LoRa gateway.

B. Stage II: Antenna Pattern

As the $Leo1$ node has a built-in antenna, an initial and fundamental step would be exploring the antenna pattern efficiently. This action would help understand the effect of changing the direction and polarity of the node wherever it is moving or when installing it for a period of time. We set up the node outside the Millennium Point building, in a fixed position with a compass aligned to a certain direction for a period of 4 minutes. During this period of time, $Leo1$ transmits the SNR and RSSI values with empty packets back to the TTN server. RSSI and SNR values are then extracted from the packets received and fed into a MatLab program that averages the values of the RSSI and subsequently illustrate the antenna pattern. Fig. 4 shows the illustration of the directivity of the embedded antenna. It is worth mentioning that the average RSSI values when changing the direction of $Leo1$ lies within less than 0.017% of the mean RSSI value revealing that, the embedded antenna in $Leo1$ can be well approximated as an isotropic antenna, and hence, the moving and/or fixing position of the node does not affect the output power and would not be considered as a factor affecting the RSSI.

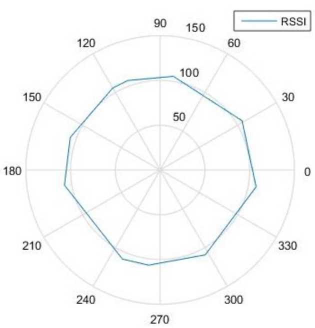


Fig. 4: Directivity plot.

IV. Data Acquisition and Regression Channel Model

A. Regression Model

We consider the parameters h_r, h_t, P_{tx} and f to be constant in our evaluation. The initial fit model of the sample measurements is determined using eq. 1 as the baseline with the average PL:

$$PL_{avg} = B + 10n\log_{10}(d/1000), \quad (5)$$

with the distance d in m , B defining the PL intercept and n being the PL exponent.

To identify the model, a regression based on the minimization of the mean square errors is done. The cost function by $J(\zeta) := \sum_i ||RSSI_{model}(d_i,\zeta) - RSSI_i||^2$, where i corresponds to the ith measurement, and $\zeta \in \mathbb{R}^2$ are the parameters of the path loss model: $\zeta := (B, n)$. The optimum parameters $\zeta^\star$ of the model are then obtained by minimization of J, such that $\zeta^\star = \underset{\zeta\in\mathbb{R}^2}{\arg\min}\, J(\zeta)$, following the derivative-free Nelder-Mead method [23].

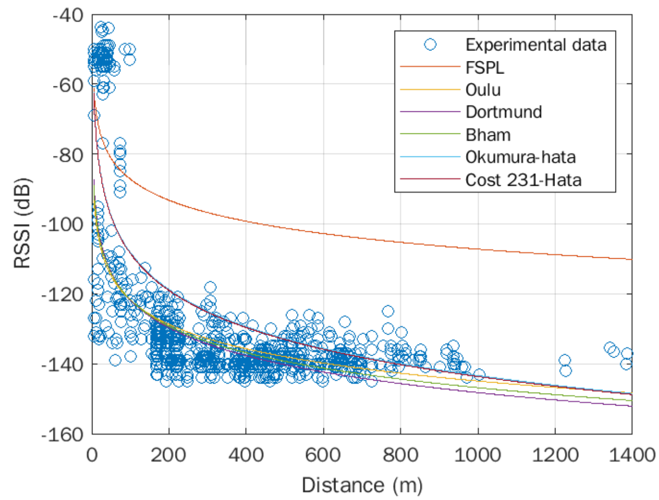


Fig. 5: Comparing RSSI of different models with measurements in Bham for 868 MHz LoRa

TABLE II: Propagation parameters used for different environment in literature and the proposed initial model

Metric	Free Space	Oulu [11]	Dortmund [12]	B'ham
Path loss exponent (n)	2.00	2.32	2.65	2.52
Path loss intercept (B)	91.22	128.95	132.25	130.85

B. Data Processing

After obtaining an initial regression model, we perform some data analysis to enhance the model further, that is done utilizing outlier data detection and elimination. In [24], authors presented a data outlier detection method that involves the median absolute deviation (MAD) as an alternative and more robust technique to measure the dispersion of data points. The median (M), like the mean (σ), is a measure of central tendency but offers the advantage of being very insensitive to the presence of outliers [25]. After calculating the MAD of the RSSI values, we then define the criteria upon which a specific data point should be rejected or tagged as an outlier. Following the suitable decision criteria proposed in [26], we implement $\frac{RSSI_i - M}{MAD} > |\pm 3|$. After obtaining and sorting the outliers, an iterative process of removing them is implemented that resulted in an optimal MSE value of 24.907, compared to an initial value of 29.3520.

C. Comparative Performance

The LoRa packets are compared with the following path loss models: FS, OH, Cost-231, Oulu and Dortmund. Fig.5 shows the RSSI of the packets received within Birmingham, it also shows the operation of the various models considered and the regression model for Birmingham. For accuracy comparison between the propagation models implemented in this study, we present the following statistical parameters $\Delta_i = RSSI_{estimate} - RSSI_{measured}$, $|\Delta y| = \frac{1}{N}\sum_{i=1}^{N}|\Delta_{yi}|$ and $\sigma_e = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\Delta_i - |\Delta y|)^2}$. Where $|\Delta y|$ is the mean absolute error and the σ_e is the standard deviation of the estimated error. N indicates the total number of data samples collected and Δ_i represent the difference between the estimated and measured samples. Table III summarizes the results obtained. The Oulu model with $|\Delta y|$ = 9.966 and has the closest accuracy to our model, as shown in fig. 5 the model estimate the RSSI values for a close distance up to 600m with a good accuracy then deviates at longer distances. While, both OH and cost-231 models provides the most imprecise estimation of the RSSI received in B'ham with $|\Delta y|$ = 13.52 and 13.34 showing a notable variation in both close and far distances.

Table II shows a comparison between the pathloss exponents (n) and intercepts (B) for the different channel models considered. While the Oulu and Dortmund models show a close initial PL intercept B, the proposed regression model for Birmingham shows higher B. The PL intercept is a parameter relating to the shadowing effect that is due to the power losses caused by barriers the signal propagates through. That might be due to the fact that BCU is surrounded by few tall and large buildings while Oulu is located at a water shore. It is also worth mentioning that Dortmund had the advantage of a better probability of LOS transmission, with a transmitting antenna height placed at 30m, while in Oulu the antenna was placed at 24m. In our experiment, the antenna was placed an an approximate height of 15m.

On the contrary to the PL exponent n, where the B'hm model shows the least exponent value of all. A lower exponent translates to a lower rate of RSSI decrease with the increasing of the distance spanned, resulting in a less inclined slope. The small value of n reflects the open space surrounding BCU, where the GW is installed, which indicates the appropriateness of the location for better coverage.

Our experiment shows a loss of LoRa packets around the 1.1 to 1.3 km range, this loss of can be an indication of high density of signal barriers causing the attenuation of the received signal power at the exact angle the measurements were taken from. A constructive approach would be to retry gathering more measurements at the same distance but with variable deviations from the gateway. The higher power received at longer distances in Birmingham (around 1.4 km) when compared to other models had an influence on the proposed channel model, showing a higher probability of LOS communication at longer distances accompanied by relatively higher RSSI value when compared to short distances, differentiate the Birmingham environment from other models.

TABLE III: Statistical Performance Metrics

Error Parameters	Okumura Hata	Cost 231-Hata	Dortmund	Oulu	B'ham	K-mean B'ham	DBC B'ham
$\lvert\Delta y\rvert$	13.52	13.34	10.05	9.96	9.9	8.09	9.5
σ_e	17.43	17.48	22.88	22.57	22.63	15.17	17.5

D. Experimental Repetition

The repetition principle is an important step in scientific and experimental research, as the observations and readings are affected by natural variables that can be stationary or random, which require a fair amount of samples to reveal their changing regularity. Stabilizing the mean and the standard deviation are the foundation of repetition, resulting in a population fairly represented by the sample [27].

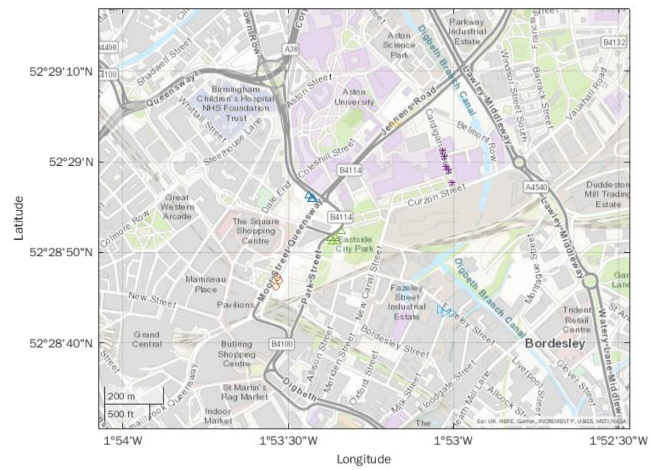


Fig. 6: Close-by locations achieved from different experiments

To achieve the aforementioned, we maintained two methods, the first approach was collecting a series of measurements for each location during the same experiment, then averaging the values for a reliable estimate. The second approach was collecting the measurements from similar locations but at different times. To achieve that, the data collected from 7 successful experiments throughout a time span of 15 months has been separately analysed as shown in fig. 2, where each experiment's data points retrieved is presented in a different color. Even though each experiment was concerned with its own path, there was enough points overlapping. The analysis included an algorithm that was used to go through all data-location pair, calculate the haversine distance, which is the distance from the GPS co-ordinates to the gateway location. Then, cross-reference the resulting distance with the GPS co-ordinates to find close by locations, since the distance on its own can represent any point on the circumference of a circle centered from the gateway. The algorithm identified the close-by locations using a threshold of approximately 15m. This threshold was chosen after considering the accuracy of the GPS chip used in the experimentation and the GIS system that is 3m and 11m, respectively.

Fig. 6 shows the first output of the algorithm, as every cluster of points is the output of overlapping close-by locations from at least 2 different experiments. Additionally, to maintain an even analysis, the algorithm was assembled to output clusters in different directions and distances of the spanned area. The results of this analysis can be found in table IV, where the average distance of each cluster from the gateway, the average and variance RSSI from various experiment are computed to be compared to the actual measurements found using the Bham model when applied to the same average distance. The results show that the variance of measurements from different

TABLE IV: Analysis from experimental repetition

	Area 1	Area 2	Area 3	Area 4	Area 5	Area 6
Mean Distance (m)	309.5	517.8	176.84	290.53	450.62	646.75
Mean RSSI	-109.5	-108.5	-101.8	-108.37	-105.66	-108.67
Variance RSSI	0.25	1.66	31.77	0.23	0.333	0.333
Bham Model RSSI estimate	-104	-109.6	-97.89	-103.32	-108.13	-112
K-mean Clustering Model	-110.96	-108.99	-98.38	-109.55	-108.99	-108.99
Distribution-based Clustering Model	-105.98	-108.23	-103.5	-105.71	-107.62	-109.2

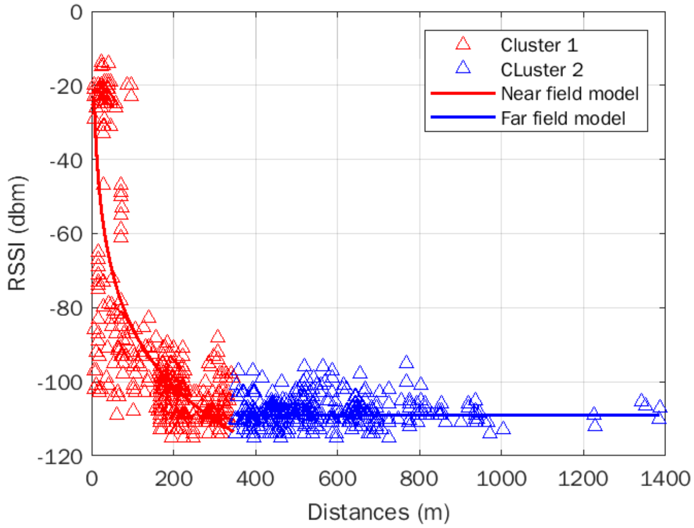


Fig. 7: K-mean clustered data with regression models for both clusters

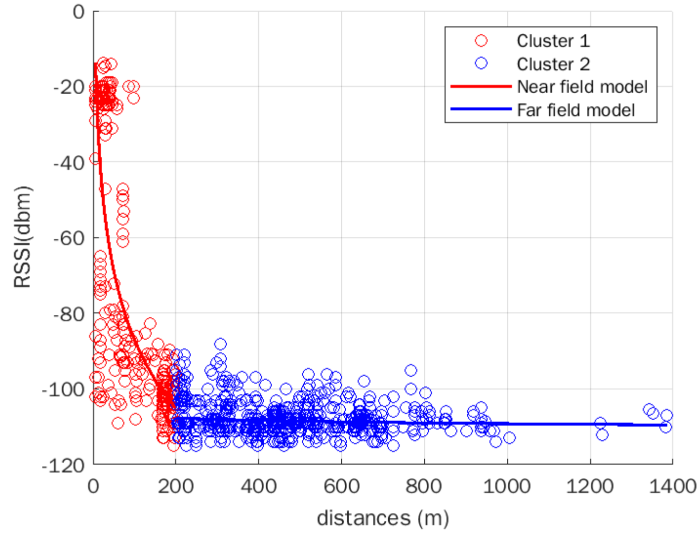


Fig. 8: Distribution-based clustered data with regression models for both clusters

experiments around the same location are quite similar which conveys the validity of the assumption that the RSSI for every location is stationary. Moreover, when comparing the average RSSI to the RSSI yielded when applying the B'ham model, a trend of overestimating the true value of RSSI in close-by distances to the gateway, i.e. distances less than or equal to 300m, can be found. In the same way, the model underestimate the power received in longer distances.

V. Distribution-based Clustering

After conducting the work done in the previous section, we noticed that the model represented deflected from the average RSSI retrieved during repeating experiments, even when considering the variance of those measurements. The model values would tend to overestimate the predicted received power is close-by distances and vice versa when it comes to longer distances. This observation raised the question of whether regression models produce the optimal representation of the collected data. Consequently, we explored the prospect of grouping data that share some features or similarities forming clusters.

Initially, we implemented the k-mean algorithm to cluster our data using the suitable measures of both the squared euclidean distance and sum of absolute differences as they were the most suitable measures to our data. We started by normalizing the data to limit redundant data and ensures that good quality clusters are generated which can improve the efficiency of the clustering algorithm [28]. Then by defining our 3 variables as distance, RSSI and SNR, we uzed the k-means algorithm to cluster the data into 2 clusters as shown in fig. 8. Ending with two clusters was not an assumption, it was the result of an iterative process to find the optimal number of clusters to represent the collected data. We evaluated the outputs for clustering in 3, 4 and 5 clusters and with different centroid calculation methods and this was found to be the most convenient. While k-means chooses the centroid of the cluster based on the mean or median distance from the rest of the cluster points, the DBC considers the distribution of the data presented. In our DBC algorithm, we divided the RSSI data into two clusters, based on their distribution. The first cluster showed a binomial distribution up till the distance of 200 m away from the transmitter, while distances exceeding 200 m and reaching 1400 m showed a normal distribution.

Fig. 7 and fig. 8 show the results of DBC and K-means clustering, the new clustered models are referred to as near and far field models. On one hand, the kmean algorithm near cluster covered up to 346 m, the far field model shows a constant model for the higher distances. While the differences between power in longer distances might seem slight, a constant model would not capture the variations present. On the other hand, the DBC show quite distinct variation between near and far field models.

To evaluate the clustering algorithms used in modelling and for fair comparison, We reused the formula for the mean absolute error $|\Delta y|$ that showed that the DBC model yielded a 4% better absolute error than the initial model presented and the k-means clustering revealed a 18%. Moreover, we computed the estimated RSSI for both models and compared them to the results obtained from the repetition analysis in sec. IV-D, that is shown in table IV. Although, the overall absolute error of the DBC was as little as 4%, it still showed a noticeably more accurate power estimation to the results from the repetition analysis i.e for Area1, the DBC estimate is -105.9 dB when the true measurement was -104 dB. While the initial model estimate is further away also as the K-mean clustering. This pattern can also be seen in Areas 2 and 3. A worth mentioning observation is that the K-mean clustering outperformed DBC in shorter distances to the gateway such as Area 4. Based on our findings, it is shown that proposing a clustering algorithm enhances the prediction accuracy when compared to a regression model. Both, K-means or DBC surpasses regular regression models, but the decision of which is more adequate is perhaps more related to the distances covered.

VI. Conclusion

There are various challenges and factors to be considered when implementing smart cities using LP-WAN. LoRaWAN has the potential to build successful LPWANs but only when its strengths and weakness limitations are considered. In this work, we presented a real-world experimental study and proposed a channel model evaluating the performance of LoRa to understand the its suitability for Smart cities in Birmingham city and similar urban areas. The experiments show the success of LoRa signals reception up to 1.4 km range, with blind spots between 1.1 to 1.3 km range. Extending the work to compare various empirical channel models showed insufficient adequacy, subsequently an initial channel model is presented and evaluated, showing an increase of accuracy of 4% and 18%. Further to that, studying the RSSI values received and evaluating their distribution had a positive influence on enhancing the estimated measurements. Utilizing our findings on a bigger scale of RSSI fingerprinting or localization shows a promising prospect.

Acknowledgment

Authors want to acknowledge the funding support from European Union's Horizon 2020 Research and Innovation programme under the Marie Sklodowska-Curie IoT4WIN Grant Agreement No. 765921.

References

[1] "Sigfox.com," Mar 2021. [Online]. Available: https://www.sigfox.com/en

[2] L. Alliance, "A technical overview of lora and lorawan," *White Paper, November*, vol. 20, 2015.

[3] S. Corporation, "An1200. 22 lora modulation basics," 2015.

[4] L. Vangelista, "Frequency shift chirp modulation: The lora modulation," *IEEE Signal Processing Letters*, vol. 24, no. 12, pp. 1818–1821, 2017.

[5] Y. Bagariang, M. I. Nashiruddin, and N. M. Adriansyah, "Lora-based iot network planning for advanced metering infrastructure in urban, suburban and rural scenario," in *2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*. IEEE, 2019, pp. 188–193.

[6] R. Sanchez-Iborra, J. Sanchez-Gomez, J. Ballesta-Viñas, M.-D. Cano, and A. Skarmeta, "Performance evaluation of lora considering scenario conditions," *Sensors*, vol. 18, no. 3, p. 772, 2018.

[7] F. Adelantado, X. Vilajosana, P. Tuset-Peiro, B. Martinez, J. Melia-Segui, and T. Watteyne, "Understanding the limits of lorawan," *IEEE Communications magazine*, vol. 55, no. 9, pp. 34–40, 2017.

[8] K. Mikhaylov, J. Petaejaejaervi, and T. Haenninen, "Analysis of capacity and scalability of the lora low power wide area network technology," in *European Wireless 2016; 22th European Wireless Conference*. VDE, 2016, pp. 1–6.

[9] M. C. Bor, U. Roedig, T. Voigt, and J. M. Alonso, "Do lora low-power wide-area networks scale?" in *Proceedings of the 19th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, 2016, pp. 59–67.

[10] R. El Chall, S. Lahoud, and M. El Helou, "Lorawan network: Radio propagation models and performance evaluation in various environments in lebanon," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2366–2378, 2019.

[11] H. Linka, M. Rademacher, O. G. Aliu, and K. Jonas, "Path loss models for low-power wide-area networks: Experimental results using lora," 2018.

[12] P. Jörke, S. Böcker, F. Liedmann, and C. Wietfeld, "Urban channel models for smart city iot-networks based on empirical measurements of lora-links at 433 and 868 mhz," in *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE, 2017, pp. 1–6.

[13] J. Petajajarvi, K. Mikhaylov, A. Roivainen, T. Hanninen, and M. Pettissalo, "On the coverage of lpwans: range evaluation and channel attenuation model for lora technology," in *2015 14th International Conference on ITS Telecommunications (ITST)*. IEEE, 2015, pp. 55–59.

[14] P. J. Radcliffe, K. G. Chavez, P. Beckett, J. Spangaro, and C. Jakob, "Usability of lorawan technology in a central business district," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*. IEEE, 2017, pp. 1–5.

[15] A. J. Wixted, P. Kinnaird, H. Larijani, A. Tait, A. Ahmadinia, and N. Strachan, "Evaluation of lora and lorawan for wireless sensor networks," in *2016 IEEE SENSORS*. IEEE, 2016, pp. 1–3.

[16] M. Bor and U. Roedig, "Lora transmission parameter selection," in *2017 13th International Conference on Distributed Computing in Sensor Systems (DCOSS)*. IEEE, 2017, pp. 27–34.

[17] H. T. Friis, "A note on a simple transmission formula," *Proceedings of the IRE*, vol. 34, no. 5, pp. 254–256, 1946.

[18] M. Hata, "Empirical formula for propagation loss in land mobile radio services," *IEEE transactions on Vehicular Technology*, vol. 29, no. 3, pp. 317–325, 1980.

[19] E. Harinda, S. Hosseinzadeh, H. Larijani, and R. M. Gibson, "Comparative performance analysis of empirical propagation models for lorawan 868mhz in an urban scenario," in *2019 IEEE 5th World Forum on Internet of Things (WF-IoT)*. IEEE, 2019, pp. 154–159.

[20] J. Milanovic, S. Rimac-Drlje, and K. Bejuk, "Comparison of propagation models accuracy for wimax on 3.5 ghz," in *2007 14th IEEE International Conference on Electronics, Circuits and Systems*. IEEE, 2007, pp. 111–114.

[21] C. Action, "231,"digital mobile radio towards future generation systems, final report," european communities," EUR 18957, Tech. Rep, Tech. Rep., 1999.

[22] *Low-Power Long Range LoRa® Technology Transceiver Module*, Microchip Technology, 2015-2020. [Online]. Available: https://ww1.microchip.com/downloads/en/DeviceDoc/RN2483-Data-Sheet-DS50002346E.pdf

[23] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the nelder–mead simplex method in low dimensions," *SIAM Journal on optimization*, vol. 9, no. 1, pp. 112–147, 1998.

[24] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median," *Journal of Experimental Social Psychology*, vol. 49, no. 4, pp. 764–766, 2013.

[25] D. L. Donoho and P. J. Huber, "The notion of breakdown point," *A festschrift for Erich L. Lehmann*, vol. 157184, 1983.

[26] J. Miller, "Reaction time analysis with outlier exclusion: Bias varies with sample size," *The quarterly journal of experimental psychology*, vol. 43, no. 4, pp. 907–912, 1991.

[27] L. Hu, X. Bao, and Q. Wang, "The repetition principle in scientific research." *Zhong xi yi jie he xue bao= Journal of Chinese Integrative Medicine*, vol. 9, no. 9, pp. 937–940, 2011.

[28] D. Virmani, S. Taneja, and G. Malhotra, "Normalization based k means clustering algorithm," *arXiv preprint arXiv:1503.00900*, 2015.