SFN Gain Prediction by Neural Networks for Enhancing Layer 2 Coverage in LDM Systems

Yosvany Hervis Santana, David Plets, *Member*, *IEEE*, Toon de Pessemier, Rodney Martinez Alonso, *Student Member*, *IEEE*, Glauco Guillen Nieto, Luc Martens, *Member*, *IEEE* and Wout Joseph, *Senior Member*, *IEEE*

***Abstract*—LTE-eMBMS systems efficiently deliver multicast/broadcast services using Layered Division Multiplexing (LDM) technology. In a two-layer LDM system, Layer 1, with higher power allocation delivers mobile services, and Layer 2 in a Single Frequency Network scheme provides local content. The challenge is to reduce the gap in the layers' coverage areas caused by the use of different constellations, and SFN gain for Layer 2. Hence, the precision in the coverage area estimation is crucial for the successful planning and deployment, particularly regarding the SFN gain contribution in Layer 2. For this purpose, a real digital TV broadcasting SFN system was used as a model to design a method based on Machine Learning algorithms, aiming to enhance the coverage area precision for the Layer 2 in eMBMS. The method is able to estimate SFN gain value with a Mean Absolute Error (MAE) of 0.72 dB and certainty in positive or negative contribution in 93% of the cases.**

***Index Terms*—eMBMS, LTE, LDM, Layers' Coverage Gap, Machine Learning, SFN.**

I. Introduction

DUE to the rapid increase in the number and type of mobile devices, the emerging mobile applications, and the demand for higher video service quality, mobile broadband systems are currently under great pressure to increase the service capacity to meet the fast-growing data traffic requirement, particularly on video services.

Multimedia Broadcast Multicast Service (MBMS) is the point-to-multipoint transmission specification defined by the 3rd Generation Partnership Project (3GPP) to provide efficient delivery of multicast and broadcast services in cellular networks[1]. An enhanced version (eMBMS) was defined for the Long-Term Evolution (LTE) network, which can provide higher throughputs, enhanced broadcasting services, and flexible carrier configuration [1]. In the LTE network, the eMBMS are delivered in sub-frames of the available Physical Resource Blocks in TDM mode [2]. However, the inclusion of LDM in 5G-MBMS [3], enables the provision of unicast, multicast or broadcasting services, multiplexed across the multi-layer architecture. In [4], the authors proposed the application of Layered Division Multiplexing (LDM) technology to the LTE-eMBMS. LDM consists of a multi-layer signal structure for providing different services simultaneously over the same frequency channel, with multiple signal layers and different transmission powers.

LDM technology can efficiently deliver multiple broadcasting services with different bandwidth requirements [5]. Compared to traditional Timing Division Multiplexing (TDM) and Frequency Division Multiplexing (FDM) systems over an LTE-eMBMS network, the LDM can achieve higher capacity when delivering multiple services in the same channel. For instance, the authors demonstrated in terms of Spectral Efficiency that for the same Signal to Noise Ratio (SNR) value of 15 dB, LDM outperforms F/TDM by 1.6 bps/Hz under the assumption of ideal channel coding [6]. In [6], the authors compare LDM and FDM/TDM in terms of the mobile service's capacity-coverage tradeoff. Theoretically, it was demonstrated that, under the assumption of ideal channel coding, LDM outperforms FDM/TDM in both capacity and coverage.

In a Single Frequency Network (SFN) system, the same signal is transmitted synchronously from all transmitters at the same channel [7][8]. A SFN allows achieving a higher spectrum utilization efficiency (only one instead of several channels), less total transmission power, and a more homogeneous distribution of electric-field (E) strength [9][10]. This technology optimizes the spectrum resources because it provides the required coverage through multiple transmitters operating at the same frequency and carrying the same content. It introduces a certain contribution to the reception signal level and can improve service availability and a more homogeneous E-strength distribution throughout its coverage area. These characteristics of a SFN enable a potential network gain [9].

The advantages of this broadcast network structure have enabled its introduction in recent broadband technologies for broadcast and multicast applications. For LTE and LTE-Advanced services, SFN plays a key role. This was introduced in Release 9 of the 3GPP (3rd Generation Partnership Project)

The work of Y. Hervis Santana was supported in part by the Special Research Fund (BOF), and in part by LACETEL.

Y. Hervis Santana and R. Martinez Alonso are with the LACETEL, Havana 19200, Cuba, and also with the INTEC, Ghent University-IMEC, 9052 Gent, Belgium (e-mails: yosvany.hervissantana@ugent.be; rodney.martinezalonso@ugent.be).

D. Plets, T. de Pessemier, L. Martens and W. Joseph are with INTEC, Ghent University-IMEC, 9052 Gent, Belgium (e-mail: david.plets@ugent.be; toon.depessemier@ugent.be; luc1.martens@ugent.be; wout.joseph@ugent.be).

G. Guillen Nieto is with the LACETEL, Havana 19200, Cuba (e-mail: glauco@lacetel.cu).

as Multimedia Broadcast Single Frequency Network (MBSFN), where the same content is transmitted to a group of users in a cell using a subset of available resources [11].

Using SFN, a reduction in the transmitted power can be made, keeping the same coverage [12][13]. However, the SFN gain is not homogenous and might be even negative in some areas. In [14], the authors proposed a methodology to calculate the SFN gain (SFNG). The same article also proved that SFN gain is closely related to some of the measured parameters, i.e., received Electrical field (E), differences of *E* between the two strongest transmitters (E_{diff}), enabling its estimation. Also, in [15], the authors demonstrated that the transmission delay difference as a considerable impact on the *SFNG*. Using these parameters to train/validate a Machine learning (ML) model, the SFN gain can be estimated, and the results can be further applied in different scenarios.

We analyze here Layer 2 of the LDM system. Layer 1 or the Upper Layer (with higher power allocation) is used to deliver mobile services (LTE). The Layer 2 or Lower Layer is designed to provide high data rate services (UHDTV or multiple HDTV services) to fixed reception terminals, where the operational SNR is usually high [5]. A relevant challenge in such approach is to reduce the gap in the layers' coverage areas identified in [16]. Hence, the precision in the coverage area estimation is crucial for planning and deploying an eMBMS system, particularly regarding the SFN gain contribution in Layer 2.

Our research's novelty is a method based on an Artificial Intelligence algorithm for accurately estimating the SFN gain for its further application in balancing the Layer 1 to Layer 2 coverage ratio in LTE-eMBMS networks. The three significant contributions of this paper are as follow: (i) the feasibility of using ML to estimate the *SFNG* based on real data, (ii) a method based on ML to identify the *SFNG* with an error of 0.72 dB, and (iii) enhancing the coverage area precision for the SFN Layer 2 in eMBMS.

The paper is organized as follows. Section II presents the scenarios and methods used in the research. Sections III describes the most relevant performance results. Finally, Section IV highlights the main conclusions of this work.

II. Scenarios and Methods

A. Measurement Area

For estimating the SFN gain, the field measurements in [14] are considered. The transmitting network is located in Ghent (Belgium). The SFN consists of three transmitters operating at a frequency of 602 MHz [14]. Fig. 1 shows a map of Ghent with the location of the three Transmitters (Tx) (black dot markers in Fig. 1) and the 50 km route (black line) where the measurement campaign was performed. The route starts from the center of Ghent to the surrounding municipalities. A total of 389 locations were measured, and the resulting dataset is used to train/validate the ML models. The work is focused on estimating the SFN gain value with six different ML models, with reference models and determine the best model to estimate the SFN gain, compared with the actual results from the measurement campaign.

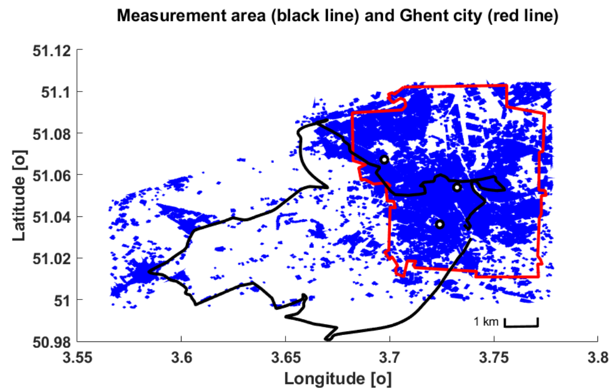

Fig. 1. Map of Ghent (route inside the red line) with the three Transmitters (black dot markers) and the measurement route (black line).

B. Data recording and processing

The field tests were based on mobile Digital Video Broadcasting (DVB) measurements along a 50 km route. The exact same measurement route was repeated four times, each time with a different transmitter configuration (Transmitters 1, 2, 3 and SFN scenario). In order to properly compare the recorded measurements along each route, a synchronization procedure is required, in both time and space. Compared to [14], an improvement in the spatial synchronization is applied in this work, based on [17]. Using the algorithm from [17], the four trajectories were first map matched, meaning that all GPS location sequences were mapped to the most likely trajectory on road segments, accounting for the instantaneous velocities, speed limits, one-way streets, road type and location. It was visually confirmed that this procedure resulted in the four same routes. The remainder of the synchronization process consists of time synchronization of the samples along the route and remains unchanged compared to [14]. This time synchronization consists of a division of one of the four routes into smaller segments of a fixed segment length, determining the sample at the end of each segment and determining the matching border samples of the other three trajectories, resulting in a set of corresponding road segments for all four routes. Similarly to [14], due to statistical relevance, data were discarded where the SFN or MFN trajectory did not contain more than 5 samples. Also, segments are discarded for which the resulting actual segment length of the SFN and MFN trajectories differ more than 20%. Using a similar reasoning as in [14], a segment length of 100 m was chosen. Once the segment lengths are chosen, a database is created which for each segment contains all the information required for training and validating the neural network. A total of 21 variables in 389 points, related to SFN and MFN parameters as well as geographical information were obtained.

Before starting with the training and validating process, we must analyze the available data to avoid outliers associated with possible measurement errors. All the values out of the range *median* ± 9.6 dB (highest possible contribution of an

SFN with three transmitters) were removed. Fig. 2 shows a histogram of the measured *SFNG*. Among the 367 samples, 316 (86%) samples are between -5 dB and 5 dB (inside the blue rectangle). For having a better understanding of the *SFNG*, some other metrics need to be highlighted. For instance, the average of the samples is 3.16 dB, with a standard deviation of 3.96 dB, and the median value is 0.22 dB.

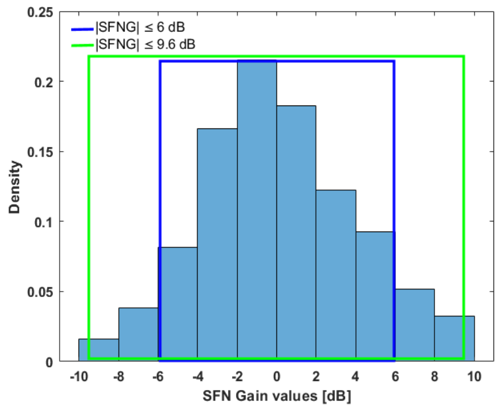

Fig. 2. Histogram of the *SFNG* measured values. The image highlights the *SFNG* values below to 6 dB (blue rectangle) and 9.6 dB (green rectangle).

Among all the measured parameters, we want to highlight E_{diff} because of its influence in the *SFNG* estimation. In [14], the authors defined this parameter as the difference between the dominant transmitter and the second strongest transmitter (1).

$$E_{diff} = E_{median}^{Dominant\ Tx} - E_{median}^{SecondStrongestTx} \quad [\text{dB}] \qquad (1)$$

Another parameter to highlight is the *SFNG*. It is close related to some measured parameters. In [14] and [10] the SFNG is define as:

$$SFNG = MER_{SFN} - MER_{MFN} \qquad [dB] \qquad (2)$$

where in the particular case of [14], the MER_{SFN} is the median MER in a segment when all the three transmitters are active and MER_{MFN} is the maximum of the median MER values with only one transmitter active.

C. Proposed Machine Learning Algorithms

Artificial intelligence and Machine Learning (ML) algorithms are inexpensive and powerful tools, widely used to learn data patterns by exploiting vast relevant information from a previously collected dataset [18]. These algorithms have recently been applied to plan and optimize multiple kinds of telecommunication networks and services [18][19][20], proving their advantage over theoretical and/or empirical propagation models.

For designing the model, we use the WEKA tool [21]. It is a collection of machine learning algorithms and data preprocessing tools [21]. We use the algorithms with the best performance for our dataset in the training/validation process. Those algorithms are Gaussian Process (GPs), Linear Regression (LR), Instance-Based Learning (IBK- commonly known as a k-nearest neighbour algorithm), SMOreg, KStar, and Multilayer Perceptron (MLP).

The GPs are a nonparametric classification method based on a Bayesian methodology [21]. It is a powerful algorithm for both regression and classification. LR is a classical statistical method that computes the coefficients or "weights" of linear expression, and the predicted ("class") value is the sum of each attribute value multiplied by its weight [22]. k-Nearest Neighbor (kNN) is one of the most popular data classification algorithms [23] because it is easy to implement. There are only two parameters to implement, the value of k (number of neighbors), and the distance function (Euclidean or Manhattan) [23][24]. SMOreg implements the support vector machine for regression. The parameters can be learned using various algorithms by setting the RegOptimizer [25]. KStar is an instance-based classifier based upon the class of those training instances similar to it, as determined by some similarity function. It differs from other instance-based learners in that it uses an entropy-based distance function [26]. MLP can model non-linear problems and map sets of input data into a set of appropriate outputs using a supervised algorithm called Levenberg-Marquardt [27]. The activation function and the number of hidden layers can be changed directly on the Weka tool.

D. Methods and Reference Models

To know how accurate the proposed ML models are, we compare the obtained results with different benchmarking models.

- "Ideal Reference Model" (IRM), where we added all possible inputs (20) to estimate one output (SFNG). This model will help deciding whether the SFN gain could be estimated using an ML algorithm trained with the available dataset. After the comparison with the field measurement, the inputs highly correlated to the SFNG will be removed and the final model will be compared with the following two models.
- "Basic Reference Model" (BRM), by predicting the SFN gain as training dataset average. Each error in the validation set will be the average in the training dataset minus the actual measured SFN gain.
- "Linear E_{diff} Model" (LEM). It is based on [14] and estimates the SFNG based on a linear relationship with the recorded E_{diff} values. In [14], a correlation of -0.48 was obtained for the entire dataset. In this paper, the linear relationship will be determined based on the training set to ensure a fair comparison with the ML model.

E. Dataset Inputs and Output

The whole dataset consists in 20 Inputs (distances from the Tx, Electrical Field (E) values for every steady point with the Tx working on MFN mode and SFN mode, delta distances between the two strongest transmitters, and so on) related to each measured point and one Output (*SFNG*) (See Fig. 3). Initially, the whole dataset is assessed to verify if the SFN gain can be estimated using ML. Finally, we find which inputs have more influence on the SFN gain estimation. We use the k cross-validation algorithm [28], where the best performance was obtained with k = 20. For k = 20, it has 19/20 (349 samples for training) of the data for learning. In contrast, for k = 5, it has only 4/5 (294 samples for training) of the data.

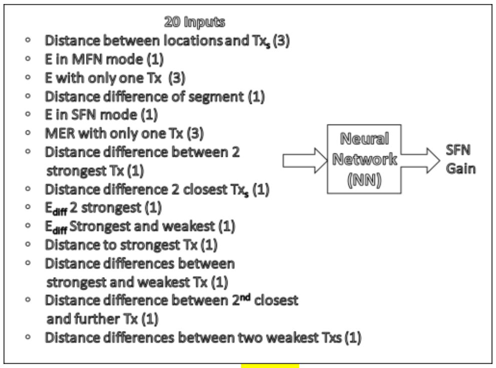

Fig. 3. Representation of the ML with twenty inputs and one output. E = Electrical field, MFN = Multi Frequency Network, SFN = Single Frequency Network, MER = Modulation Error Rate, E_{diff} = Electrical field difference.

F. Model Evaluation

The Mean Absolute Error (MAE) is used to control the training process and have a result in the validation process. This value is calculated by employing (3).

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Ti - Oi| \tag{3}$$

Where n is the total number of samples within the samples population, Ti and Oi are the i^{th} values of the target *SFNG* inside the training/validation dataset and output (estimated SFN gain value by the ML), respectively [29]. One way to avoid overfitting [30] the system is to split the whole dataset in different configurations to have a number of samples for training (i.e., 70%, 75%, or 80%) and the rest for validation (respectively 30%, 25% or 20%). However, the *MAE* value does not show the points where the error was higher than the *MAE*, being a problem in the SFN design. This uncertainty can be solved by calculating the standard deviation of the error over the whole validation set.

The next step is to investigate the six algorithms (Section II-C) to estimate the SFN gain and compare their results. The machine learning is based on updating the weights and bias from randomly chosen values every time the system is started up. During the training process, these values are updated to reduce the *MAE* and finally having an output closer to the target (measured) value. Hence, in every iteration, we will have a different result of the *MAE*. For validation, the algorithm does not choose the best data combination only. We divided the 367 samples in ten different datasets with the proposed combinations for calculating the Progressive Average (PAVE) of the *MAE* value. It should be pointed out that each output is considered after the whole process is completed (load the proper dataset among the ten available, training/validating, and calculate the *MAE* value). At the end of the process, the *PAVE* should lead to an approximately constant value (low standard deviation), or more simulations are needed to obtain the *MAE* value with its respective standard deviation.

For employing ML as a solution to estimate the SFN gain, some elements have to be considered. The "Ideal Reference Model" (IRM) performance needs to be close to the measurement campaign's available data. Hence, the expected MAE between the target (measured data) and the predicted data (estimated values by the ML) we propose it to be below 0.5 dB. Second, the *MAE* of the proposed model in this research has to be lower than the *MAE* of the "*BRM*" and "*LEM*". Otherwise, the better solution to estimate the SFN gain could be one of these options. Then, to consider our model as a better solution for SFN gain estimation, we propose that the *MAE* should be equal or less than the *MAE* of the *IRM* plus its standard deviation ($MAE_{\text{proposed model}} \leq MAE_{IRM} + \text{Standard Deviation}_{IRM}$).

G. LDM performance in LTE-eMBMS

A critical problem in LTE-eMBMS networks is the coverage gap between Layer 1 (delivering mobile services, LTE) and Layer 2 (delivering TV, using SFN) in the LDM system. The main problems are, i) comparing with ATSC 3.0 the cell of the LTE-eMBMS network is significantly reduced and, as a consequence of that, ii) the co-channel interference is a more severe issue comparing with Digital TV system. Hence, the difference in SNR plays an important role in the coverage parity [31]. Therefore, to estimate the Layer 2 coverage, an improved *SFNG* prediction model has been proposed in this paper to bridge the gap.

In [31], the authors made a simulation-based analysis with different configurations. They established different SNR thresholds with different injection levels for Layer 2. In [14], the authors used for the measurements a DVB-T network where the constellation used for the tests was 16- Quadrature Amplitude Modulation (QAM), by using the rate-1/2 TC and MPE-FEC (Multi-Protocol Encapsulation-Forward Error Correction). This configuration has a similarity with one used in [31]. Then, starting with the hypothesis that in the Layer 2 we have the same configurations used in [31], it might be valid to use the *SFNG* estimation model to analyze the coverage area in Layer 2 of LTE-eMBMS. For the propagation path loss, we use the model described in [32], adjusted for the city of Ghent, Belgium.

III. Results and Discussion

A. Neural Network accuracy

In Section II-B (Fig. 2), the resulting dataset from the measurement campaign was analyzed. It contains 21 parameters, 20 inputs, and one output. The dataset was divided into ten different datasets for every setup (70/30, 75/25, 80/20) and evaluated for every algorithm using k-cross validation with k = 20. TABLE I shows the *MAE* for every algorithm with 80% training/20% validation setup. It is equal to 311 (295 samples using k-cross validation with k = 20) samples for training and 78 samples for validating every one of the ten different datasets.

TABLE I. MAE OF THE TRAINING/VALIDATION PROCESS AND THE STANDARD DEVIATION FOR THE IDEAL REFERENCE MODEL. 80% TRAINING/20% VALIDATION. STD = STANDARD DEVIATION

Algorithm	Training 80%		Validation 20%	
	MAE [dB]	StD [dB]	MAE [dB]	StD [dB]
GPs	0.64	0.79	0.64	0.78
LR	1.12	1.50	1.10	1.38
MLP	0.58	0.73	0.49	0.56
SMOreg	0.79	0.85	0.82	1.06
IBK	0.86	0.87	0.80	0.83
KStar	0.90	1.12	0.94	0.99

The algorithm with the best performance for this dataset is MLP with a *MAE* of 0.49 dB (lower than 0.5 dB) and a standard deviation of 0.56 dB. The MLP uses a supervised learning technique called feedforward, widely used in datasets where training with known data an/several output(s) can be estimated. The LR has the worst performance where the MAE and the standard deviations are 1.10 dB and 1.38 dB, respectively. Every input has its contribution (wide dispersion of the values), and finding an approximation for all the inputs is difficult. The algorithms based on k-nearest neighbor (IBK and KStar) have a modest performance compared with MLP. Those algorithms find the neighbor based on the distance locations, and the rest of the parameters have a significant dispersion making the estimation difficult.

B. Proposed model

Section III-A demonstrated that the *SFNG* could be estimated using ML with a *MAE* lower than 0.5 dB (estimated value in TABLE I). Hence, due to the high correlation between the SFNG and the MER inputs, and the need of especial equipment to measure the MER the next step was to remove the MER inputs. The *MAE* for the resulting 17 inputs and one output is 0.86 dB and 0.62 dB of standard deviation. The next step is to decide the inputs with a higher contribution to the SFNG. Using the algorithm "*ClassifierAttributeEval*" in the Weka tool [33], we obtain the inputs that contribute positively to every algorithm under test. After applying the algorithm, the input with the highest contribution (ranked by Weka with 0.634) to the *SFNG* is E_{diff} in agreement to the results found in [15]. Despite the strong relationship between the *SFNG* and the E_{diff} parameter, other parameters contribute to the SFNG (i.e., E_{diff} between strongest and weakest transmitters (ranked by Weka with 0.515), E in MFN mode (ranked by Weka with 0.495) and the distance difference between closest and farthest transmitters). In contrary, the input with less contribution to the *SFNG* the received electrical field in SFN mode (ranked by Weka with 0.187).

To improve the results, some parameters were modified in the IBK algorithm to minimize the *MAE*. The first modified parameter is k (number of neighbors), with the best performance for k = 4. The second parameter is the distance weighting. To improve the results, some parameters were modified in the IBK algorithm to minimize the MAE. The first modified parameter is k (number of neighbors), with the best performance for k = 4. The second parameter is distance weighting, which is set to 1/distance. The last parameter is the algorithm to search the neighbor where the best performance was achieved with the Cover Tree algorithm using a Euclidean function. Therefore, the IBK algorithm performs better than the rest of the algorithms in the proposed model. The main difference between IBK and the other evaluated algorithms is the ability to work with smaller datasets (in our case, the dataset is considered as a relatively small dataset). The IBK algorithm works better with a small dataset because it learns from the dataset only at the time of making the real time prediction. This makes the IBK converge faster to the final result. It means a shorter calculation time and best algorithm's performance to achieve a similar or better accuracy. The opposite occurs with the rest of the algorithms, which require a large dataset.

Choosing the best data combination by the algorithm might lead to overfitting [30]. This phenomenon appears if the *MAE* in the training process decreases, while the *MAE* in the validation process increases. A balance between the MAE values, both in training and validation processes, should be found to avoid this situation. TABLE II shows the results of the six algorithms with the three chosen setups after ten times running. The best performance was obtained with the IBK algorithm that is proposed, where the MAE = 0.72 dB < 1.3 (MAE_{IRM} + $Standard\ Deviation_{IRM}$) dB in the 80% training/20% validating.

TABLE II. MAE OF THE VALIDATION DATASET BASED ON THE PROGRESSIVE AVERAGE AFTER TEN SIMULATIONS. THE STANDARD DEVIATION SHOWS HOW THE *MAE* HAS CHANGED THROUGH THE TEN SIMULATIONS. *DISTANCE WEIGHTING SET TO 1/DISTANCE

Algorithm	training/ validation (%)					
	70/30		75/25		80/20	
	MAE [dB]	StD [dB]	MAE [dB]	StD [dB]	MAE [dB]	StD [dB]
GPs	2.03	1.78	1.93	1.84	1.83	1.90
LR	2.12	1.56	1.95	1.49	1.93	1.78
MLP	1.78	1.35	1.82	1.35	1.60	1.48
SMOreg	1.91	1.57	1.90	1.51	1.81	1.69
IBK	0.98	0.61	0.89	0.70	0.86/0.72*	0.62/0.58*
KStar	1.23	0.99	1.14	0.94	1.12	0.77
BRM	-	-	-	-	3.17	2.17
LEM	-	-	-	-	2.98	4.00

C. Proposed model versus Reference Models

Section II-D proposed two basic reference models (*BRM* and *LEM*). The *BRM* assumed an area where the *SFNG* was the average of the training set, and each error in the validation dataset was the average minus the measured *SFNG*. Compared with the average of the measured data, the resulting *MAE* is 3.17 dB, with a standard deviation of 2.37 dB. The second one was based on a linear equation ($Y = Mx + N$) with the form $Y = -0.275x + 2.0353$. Where Y is the *SFNG*, and x refers to E_{diff} values. We previously mentioned that the dataset was divided in ten different datasets. This means that we have 10 different values of M and N. To have the final values we apply the *PAVE* technique. The resulting *MAE* for this model is 2.98 dB, and the standard deviation is 4.00 dB. TABLE II summarizes the *MAE* values and standard deviations for the two models compared with ML models. The IBK model has a better performance than the two reference models because some parameters were modified, the number of neighbors, and the function to calculate the distance among neighbors. Hence, the IBK model can estimate values of *SFNG* with a 0.72 dB accuracy. Fig. 4 shows a scatter plot with measured and estimated *SFNG* values. In 73 of the 78 samples (93%) the estimate was correct as to whether or not there is a positive contribution to the *SFNG*.

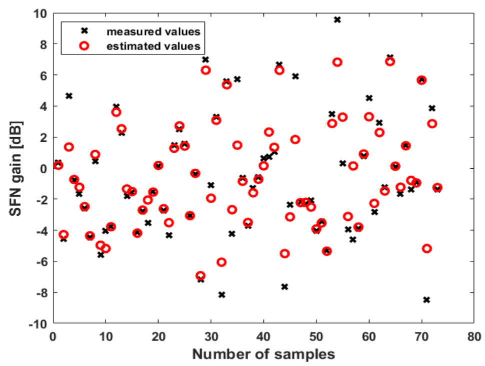

Fig. 4. Scatter plot of the measured SFNG and the estimated SFNG value. The represented values correspond to the validation dataset with a distribution of 80% training/20% validation.

D. SFNG and Covered area in L2 of LDM network

1) SFNG estimation

Section II assessed the performance of the six algorithms under test. The IBK achieved the best performance (lowest *MAE* and Standard Deviation) among the six, for our dataset. The dataset was built with the measured data along the route represented in Fig. 1. Fig. 5 shows the predicted *SFNG* with the newly developed model, applied to the city center of Ghent and the measured values. The area highlighted with the red line in Fig. 1 (Ghent city) was selected to generate a color-based map of the *SFNG*. A total of 102 points were chosen to cover the entire area. Using the model described in [32] the values of E were calculated. Finally, the *E* in SFN mode and the *SFNG* was estimated using the IBK model.

The estimated values are similar to the measured ones in the area near the map's pixels corresponding to a measured location (*i.e.*, Zone 1, Fig. 5). The obtained values for the 102 locations using the IBK algorithm have a similar *SFNG* value as those obtained in training/validating process. Zone 1 is located in the middle of the three transmitters, where all of them contribute to the received signal level and to the *SFNG* (lower E_{diff} value). However, in this area the average of the estimated *SFNG* values is 0.29 dB. On the contrary, in Zone 2, Fig. 5, we can notice an area with negative values of the *SFNG*. The area is close to the transmitter M (downtown of Ghent city). In this zone the E_{diff} has a lower value ($E_{S\_Tx} - E_{M\_Tx} \sim 7$ dB), causing that there is not a significant influence on the *SFNG*. Furthermore, the influence of two of the three transmitters to the received signal level compared with the strongest one, it is not significant to contribute the *SFNG*. In Zone 3, the *SFNG* is around 6 dB. The zone is out of the city center in a less populated area, and the three transmitters contribute to the received signal level. To conclude, in 74 (72.5%) of the 102 points, the *SFNG* is near zero dB or positive. In this way, the improvement made when using SFN over MFN is demonstrated.

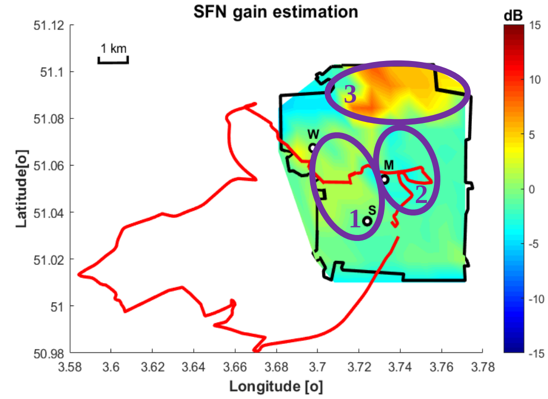

Fig. 5. Estimated SFNG in Ghent city (the area inside the black border and measured values (along the 50 km of route highlighted with the red line). Zones 1 and 3 represent areas where the SFNG is positive, and in Zone 2, the SFNG is negative. S = Strongest transmitter, M = Middle transmitter, and W = Weakest transmitter in terms of transmitted power.

2) Covered area in LDM network

When planning a wireless network, the major question is the service area and the population covered (i.e., 95% of locations covered, 99% of the time at the cell edge). The main problem using an LDM over an LTE eMBMS network is the coverage gap between Layer 1 (LTE) and Layer 2 (LTE+SFN).

Section III-C, demonstrated that the IBK model could estimate the *SFNG*. Based on the hypothesis assumed and described in Section II-G, we analyzed and compared the coverage area working in MFN and SFN. The goal is to manage the coverage area of Layer 2 in LDM. By varying the injection level, the coverage area can be modified, and the gap between layers could be reduced. TABLE III shows how the coverage area varies either for MFN or SFN. Moreover, TABLE III shows that by adding the *SFNG*, the coverage area can be increased.

TABLE III: LAYER 2 COVERAGE AREA MANAGEMENT USING THE INJECTION LEVEL AND ADDING THE SFNG.

Injection Level	**4 dB**	**5 dB**	**6 dB**	**7 dB**
Covered Area				
MFN	5.44 km	5.67 km	5.75 km	5.8 km
SFN	7.65 km	7.86 km	8.03 km	8.17 km

We show in Fig. 6 the Layer 2 coverage area for the three transmitters working in MFN and SFN mode to have a better understanding. Using 23 dBm as transmitted power, 7 dB of injection level for the Layer 2, and the rest of the configuration described in Section II-G, the maximum coverage distance is 5.8 km for every transmitter in MFN mode. On the other hand, based on the maximum estimated value of the *SFNG* (6 dB) in in the entire area, the maximum coverage distance is 8.17 km. Hence, the coverage distance difference between both technologies is 2.37 km. This demonstrates that using SFN in Layer 2 of the LDM, the coverage area can be improved by 41%.

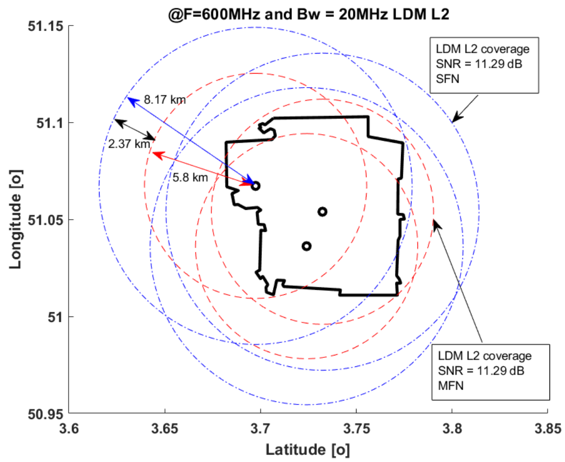

Fig. 6. The coverage area for each transmitter in MFN and in SFN mode. Red circles correspond to each transmitter's coverage area working in MFN mode and the Blue circles correspond to SFN mode.

3) *Multimedia Broadcast SNR*

Until this point, our research focused on working with only three transmitters in an SFN configuration. However, it would be valuable to evaluate our model over Low-Power Low Tower-based MBSFN, describing the delivery of MBMS (Multimedia Broadcast Multicast Services) over a radio network which is synchronized in order to minimize interference. The technique allows a group of cells to transmit the same multicast content utilizing the same radio sub-carriers, which essentially means that a group of cells all appear as one large cell since the content transmitted from each individual cell is identical (and synchronized). This makes the MBSFN transmission appear to a user equipment as a transmission from a single large cell, increasing the Signal-to-Interference Ratio due to the absence of inter-cell interference. Our proposed model can here be used to estimate the MBSFN gain and make a trade-off between the coverage area, the required number of Base Stations, and the power consumption.

IV. Conclusion

This paper studied the coverage gap between Layer 1 and Layer 2 in an LDM system of the LTE eMBMS network. Six ML algorithms' performance was compared with the available dataset from a field measurement. The IBK model was the algorithm with the best performance for our dataset with a *MAE* = 0.72 dB and a standard deviation of 0.58 dB. In 73 of the 78 samples (93%) the estimation was correct as to whether or not there is a contribution to the *SFNG*. The contribution of using our AI-driven prediction method to manage the Layer 2 coverage area on an SFN network configuration in an LTE eMBMS was demonstrated.

Feature work will consist in update the existing dataset with field measurements and use the proposed model to make a network planning.

References

[1] M. Uitto, A. Heikkinen, S. J. Rantala, and J. Mäkelä, “Evaluation of Evolved Multimedia Broadcast Multicast Service for More Efficient Mobile Video Streaming,” in *2019 IEEE International Symposium on Multimedia (ISM)*, 2019, pp. 103–1034, doi: 10.1109/ISM46123.2019.00024.

[2] L. Zhang, Y. Wu, W. Li, K. Salehian, A. Florea, and G. K. Walker, “Improving LTE eMBMS system spectrum efficiency and service quality using channel bonding, non-orthogonal multiplexing and SFN,” in *2016 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2016, pp. 1–8, doi: 10.1109/BMSB.2016.7521908.

[3] M. Simon, E. Kofi, L. Libin, and M. Aitken, “ATSC 3.0 Broadcast 5G Unicast Heterogeneous Network Converged Services Starting Release 16,” *IEEE Transaction on Broadcasting*, vol. 66, no. 2, pp. 449–458, 2020, doi: 10.1109/TBC.2020.2985575.

[4] L. Zhang, Y. Wu, G. K. Walker, W. Li, K. Salehian, and A. Florea, “Improving LTE-eMBMS With Extended OFDM Parameters and Layered-Division-Multiplexing,” *IEEE Transaction on Broadcasting*, vol. 63, no. 1, pp. 32–47, 2017, doi: 10.1109/TBC.2016.2630269.

[5] L. Zhang *et al.*, “Layered-Division-Multiplexing: Theory and Practice,” *IEEE Transaction on Broadcasting*, vol. 62, no. 1, pp. 216–232, 2016.

[6] D. Gómez-Barquero and O. Simeone, “LDM Versus FDM/TDM for Unequal Error Protection in Terrestrial Broadcasting Systems: An Information-Theoretic View,” *IEEE Transaction on Broadcasting*, vol. 61, no. 4, 2015.

[7] D. Plets, W. Joseph, L. Verloock, L. Martens, P. Angueira, and J. A. Arenas, “SFN gain in broadcast networks,” in *2011 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2011, pp. 1–6.

[8] P. J. Orbea, J. M. Matías, M. Hernández, Y. Valdés, and S. Landeros, “Effect of Receiving Antennas Radiation Pattern on an SFN Network Coverage,” in *2019 IEEE Symposium on Computers and Communications (ISCC)*, 2019, pp. 1–7, doi: 10.1109/ISCC47284.2019.8969760.

[9] I. T. U. ITU-R, “Handbook on Digital Terrestrial Television Broadcasti ng Networks and Systems Implementation.” 2016.

[10] B. Ruckveratham and S. Promwong, “Evaluation of SFN gain for DVB-T2,” in *2017 International Conference on Digital Arts, Media and Technology (ICDAMT)*, 2017, pp. 85–88, doi: 10.1109/ICDAMT.2017.7904940.

[11] I. T. U. ITU-R, “Report ITU-R BT.2386-3. Digital terrestrial broadcasting: Design and implementation of single frequency networks,” 2020.

[12] J. Montalban *et al.*, “LOCAL CONTENT DELIVERY IN SFNS USING LAYERED DIVISION MULTIPLEXING (LDM),” 2016.

[13] S. Park *et al.*, “Field Comparison Tests of LDM and TDM in ATSC 3.0,” *IEEE Transaction on Broadcasting*, vol. 64, no. 3, pp. 637–647, 2018, doi:

10.1109/TBC.2017.2755375.

[14] D. Plets, W. Joseph, P. Angueira, J. A. Arenas, L. Verloock, and L. Martens, “On the Methodology for Calculating SFN Gain in Digital Broadcast Systems,” *IEEE Transaction on Broadcasting*, vol. 56, no. 3, pp. 331–339, 2010.

[15] D. Plets, W. Joseph, L. Verloock, L. Martens, P. Angueira, and J. A. Arenas, “Influence of handover mechanisms and transmission delay difference on SFN gain,” in *IEEE international Symposium on Broadband Multimedia Systems and Broadcasting*, 2012, pp. 1–5, doi: 10.1109/BMSB.2012.6264258.

[16] J. Montalban *et al.*, “Error propagation in the cancellation stage for a multi-layer signal reception,” in *2014 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting*, 2014, pp. 1–5.

[17] J. Trogh, D. Botteldooren, B. De Coensel, L. Martens, W. Joseph, and D. Plets, “Map Matching and Lane Detection Based on Markovian Behavior, GIS, and IMU Data,” *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–15, 2020, doi: 10.1109/TITS.2020.3031080.

[18] L. Dai, H. Zhang, and Y. Zhuang, “Propagation-model-free Coverage Evaluation via Machine Learning for Future 5G Networks,” in *2018 IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2018, pp. 1–5, doi: 10.1109/PIMRC.2018.8580992.

[19] M. Ribero, R. W. Heath, H. Vikalo, D. Chizhik, and R. A. Valenzuela, “Deep Learning Propagation Models over Irregular Terrain,” in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 4519–4523, doi: 10.1109/ICASSP.2019.8682491.

[20] C. E. G. Moreta, M. R. C. Acosta, and I. Koo, “Prediction of Digital Terrestrial Television Coverage Using Machine Learning Regression,” *IEEE Transaction on Broadcasting*, vol. 65, no. 4, pp. 702–712, 2019, doi: 10.1109/TBC.2019.2901409.

[21] R. R. Bouckaert *et al.*, *WEKA Manual for Version 3.8.5*. Hamilton, 2016.

[22] S. Kavitha, S. Varuna, and R. Ramya, “A comparative analysis on linear regression and support vector regression,” in *2016 Online International Conference on Green Engineering and Technologies (IC-GET)*, 2016, pp. 1–5, doi: 10.1109/GET.2016.7916627.

[23] M. A. Mohammed *et al.*, “Solving vehicle routing problem by using improved K-nearest neighbor algorithm for best solution,” *Journal of Computational Science*, vol. 21, pp. 232–240, 2017, doi: https://doi.org/10.1016/j.jocs.2017.04.012.

[24] S. Zhang, L. Xuenlong, M. Zong, Z. Xiaofeng, and D. Cheng, “Learning k for kNN Classification,” *ACM Transactions on Intelligent Systems and Technology*, vol. 8, no. s, 2017, doi: 10.1145/2990508.

[25] S. M. A. M. Gadal and R. A. Mokhtar, “Anomaly detection approach using hybrid algorithm of data mining technique,” in *2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE)*, 2017, pp. 1–6, doi: 10.1109/ICCCCEE.2017.7867661.

[26] A. de Haro-García, G. Cerruela-García, and N. García-Pedrajas, “Instance selection based on boosting for instance-based learners,” *Pattern Recognition*, vol. 96, p. 106959, 2019, doi: https://doi.org/10.1016/j.patcog.2019.07.004.

[27] J. Bilski, B. Kowalczyk, A. Marchlewska, and J. M. Zurada, “Local Levenberg-Marquardt algorithm for learning feedforwad neural networks,” *Journal of Artificial Intelligence and Soft Computing Research*, vol. 10, 2020.

[28] K. Wang, J. Yang, G. Shi, and Q. Wang, “An Expanded Training Set Based Validation Method to Avoid Overfitting for Neural Network Classifier,” in *2008 Fourth International Conference on Natural Computation*, 2008, vol. 3, pp. 83–87, doi: 10.1109/ICNC.2008.571.

[29] M. Lindholm, F. Lindskog, and F. Wahl, “Estimation of conditional mean squared error of prediction for claims reserving,” *Annals of Actuarial Science*, vol. 14, no. 1, pp. 93–128, 2020, doi: DOI: 10.1017/S174849951900006X.

[30] H. Li, J. Li, X. Guan, B. Liang, Y. Lai, and X. Luo, “Research on Overfitting of Deep Learning,” in *2019 15th International Conference on Computational Intelligence and Security (CIS)*, 2019, pp. 78–81, doi: 10.1109/CIS.2019.00025.

[31] L. Zhang *et al.*, “Coverage analysis on improved LTE eMBMS with layered-division-multiplexing and longer cyclic prefix,” in *2017 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*, 2017, pp. 1–7.

[32] D. Plets, W. Joseph, E. Tanghe, L. Verloock, and L. Martens, “Analysis of propagation of actual DVB-H signal in a suburban environment,” in *2007 IEEE Antennas and Propagation Society International Symposium*, 2007, pp. 1997–2000, doi: 10.1109/APS.2007.4395915.

[33] S. Gnanambal, M. Thangaraj, V. Meenatchi, and V. Gayathri, “Classification Algorithms with Attribute Selection: an evaluation study using WEKA,” *Int. J. Advanced Networking and Applications*, vol. 09, no. 06, pp. 3640–3644, 2018.

Yosvany Hervis Santana was born in 1988 in Matanzas, Cuba. In 2012, he obtained a B.Sc. degree in Telecommunications and Electronics Engineering and an M.Sc. degree in Digital Systems (2017) from the Higher Polytechnic Institute CUJAE, Havana, Cuba. Since 2012, he is a fellow researcher at LACETEL, R&D ACM Transactions on Intelligent Systems and Technology Telecommunications Institute. He also has collaborated on DTV engineering projects with KitKing and KONKA (China). Currently, he is a Ph.D. student at the WAVES group (Department of Information Technology – INTEC, Ghent University).

David Plets was born in 1983 in Belgium. He received a Master degree and Ph. D. degree in 2006 and 2011 respectively. Currently, he is a member of the IMEC – WAVES group (Department of Information Technology – INTEC, Ghent University). His current research interests include low-exposure wireless indoor network planning, cognitive networks, WiFi QoS optimization, and localization algorithms.

Toon De Pessemier received an MSc degree in computer science engineering at Ghent University in 2006. In 2013, he obtained a PhD in his research "improved online services by personalized recommendations and optimal quality of experience parameters". In 2015, he obtained a post-doc scholarship for his research, "Design and validation of innovative techniques to assist people in the decision-making process by analyzing data of user behavior and contextual features". From 2018, he is a part-time professor and teaches Recommender Systems.

Rodney Martinez Alonso was born in 1987 in Havana, Cuba. In 2010, he obtained a B.Sc. degree in Telecommunications and Electronics Engineering and an M.Sc. degree in Digital Systems (2015) from the Higher Polytechnic Institute CUJAE, Havana, Cuba. Since 2010, he is a fellow researcher at LACETEL, R&D Telecommunications Institute. He also has collaborated on DTV engineering projects with HAIER and KONKA (China). In 2016 he started as a Ph.D. student at the WAVES group (Department of Information Technology – INTEC, Ghent University). The work there leads him in 2020 to receive his Ph. D. degree.

Glauco Guillen Nieto was born in 1961 in Havana, Cuba. In 1985, he obtained a B.Sc. degree in Radiocommunications and Broadcasting Engineering and a Ph.D. degree (1989) from the former Electro-Technical Institute of Communication, currently Ukrainian Academy of Communication. He is currently a senior researcher at LACETEL and an elected member of the Cuban Academy of Science (2012-2024). From 2012 to 2018, he was the advisor of the local network operators. He is a member of the Communication Ministry and Communication Enterprise Group Scientific Councils. He also has collaborated in different projects with Fineline Technology (Canada), DTVNEL (DTV National Engineering Lab - Beijing), HAIER, and KONKA (China).

Luc Martens (M'92) was born in Gent, Belgium on May 14, 1963. He received the M. Sc. degree in electrical engineering from Ghent University (Belgium) in July 1986. From September 1986 to December 1990, he was a research assistant at the Department of Information Technology (INTEC) of the same university. During this period, his scientific work was focused on the physical aspects of hyperthermic cancer therapy. His research work dealt with electromagnetic and thermal modeling and with the development of measurement systems for that application. This work led to a Ph. D. degree in December 1990. Since January 1991, he is a member of the permanent staff of the Interuniversity Micro Electronics Centre (IMEC), Ghent, and is responsible for the research on experimental characterization of the physical layer of telecommunication systems at INTEC. His group also studies topics related to the health effects of wireless communication devices. Since April 1993 he is Professor in electrical applications of electromagnetism at Ghent University.

Wout Joseph was born in Ostend, Belgium on October 21, 1977. He received the M. Sc. degree in electrical engineering from Ghent University (Belgium), in July 2000. From September 2000 to March 2005, he was a research assistant at the Department of Information Technology (INTEC) of the same university. During this period, his scientific work was focused on electromagnetic exposure assessment. His research work dealt with measuring and modeling of electromagnetic fields around base stations for mobile communications related to the health effects of the exposure to electromagnetic radiation. This work led to a Ph. D. degree in March 2005. Since April 2005, he is a postdoctoral researcher for iMinds-UGent/INTEC. From October 2007 to October 2013, he was a Post-Doctoral Fellow of the FWO-V (Research Foundation – Flanders). Since October 2009, he is a professor in the domain of Experimental Characterization of wireless communication systems. His professional interests are electromagnetic field exposure assessment, in-body electromagnetic field modeling, electromagnetic medical applications, propagation for wireless communication systems, antennas, and calibration. Furthermore, he specializes in wireless performance analysis and Quality of Experience.