Optimal Bin Width for Autonomous Coverage Estimation Using MDT Reports in the Presence of User Positioning Error

Haneya Naeem Qureshi, *Student Member, IEEE*, and Ali Imran, *Senior Member, IEEE*

***Abstract*— Minimization of drive test (MDT) allows coverage to be estimated at the base station using user equipment measurement reports with the objective of eliminating the need for drive tests. In this letter, we quantify various types of errors in MDT-based autonomous coverage estimation that stem from inaccurate user positioning, for example, as a result of GPS measurement uncertainties and quantization due to dividing the coverage area into bins. By investigating the interplay between quantization and positioning error to estimate coverage, we show that there exists an optimal bin width for coverage estimation and determine it as a function of positioning error and user density. This can enable network operators to configure the bin size for given positioning accuracy that results in the most accurate MDT-based coverage estimation.**

***Index Terms*— Minimization of drive test, positioning error, optimal bin width, autonomous coverage estimation.**

I. Introduction

NETWORK automation or self-organization enables the network to detect changes, such as detection of coverage holes, weak coverage, performance degradation problems and then based on these detected changes, make timely decisions [1]. In conventional cellular networks, cell outage detection mechanisms incur inevitable delay and unreliability that stems from human error and low spatio-temporal granularity of reports gathered via drive tests [2]. This problem is likely to aggravate with the advent of emerging small cells, where the probability of cell outages is expected to increase further.

To overcome the aforementioned challenges, 3GPP has standardized a self-organizing network use case, called minimization of drive test (MDT), which exploits the measurement reports gathered by the user equipment (UE). The UE measurement reports are tagged with their geographical location information, sent to their serving base station (BS) and ultimately used to generate coverage maps [2]–[4]. While far more efficient than drive tests, any MDT based solution for coverage estimation has to overcome following two major errors:

1) Positioning error: The reported geographical coordinates of the UE obtained from any positioning technique, such as assisted global positioning system are susceptible to errors, resulting in the reports being tagged to a wrong location [5]. These locations can also be inaccurate for the purpose of preserving user privacy.
2) Quantization error: Storing all MDT reports from all users is computationally inefficient and leads to unnecessary wastage of valuable memory resources. Therefore, the coverage area is often divided into bins and the average received power from each bin is stored and used to build coverage maps. This results in quantization error due to averaging.

Akbari *et al.* [2], Galindo-Serrano *et al.* [3], Naranjo *et al.* [4], Akbari *et al.* [5], and Akbari *et al.* [6] aim to address the reliability of MDT-based coverage estimation in the presence of positioning errors. However, these studies do not take into account the errors resulting from quantization. Quantization error to estimate cell radius is discussed in [7]. However, the work in [7] does not use MDT-based approach. Sohrabi and Kuehn [8] use regression clustering for construction of RSRP maps from a sparse set of MDT measurements. However, this work [8] assumes perfect user locations and a fixed bin width or grid size. Lin [9] propose a MDT system in which UEs upload the measurement reports periodically. Based on the collected measurement reports, the MDT system learns the knowledge about the communication environment and use it to forecast signal strengths. However, in this work Lin [9] generate the forecast of signal strength given the locations of base stations and UEs are known. Therefore, current studies on MDT-based coverage estimation either assume perfect user locations and no quantization, or consider the effect of positioning and quantization errors independent of each other.

In this letter, we analyze the interplay between the aforementioned errors concurrently in coverage estimation through MDT. While on one hand, decreasing bin size reduces the quantization error, on the other hand, it increases the error in coverage estimation due to incorrect user positioning. This study is the first to show that there exists an optimal bin width for given user positioning error that can minimize the overall error in the MDT based coverage error, i.e., the combined error caused by quantization (dictated by bin size) and user positioning inaccuracy. This calls for an optimization of bin width that would minimize the overall error under positioning error constraints. To the best of our knowledge, this letter is the first to analyze and quantify the interplay between these errors simultaneously and present a framework to determine the optimal bin width that minimizes these errors concurrently.

II. System Model

We consider a system of N base stations uniformly distributed in an area of $A \times A$, where A is the length in meters.

Manuscript received January 24, 2019; accepted February 7, 2019. Date of publication February 12, 2019; date of current version April 9, 2019. This material is based upon work supported by the National Science Foundation under Grant Numbers 1619346, 1559483, 1718956 and 1730650. The associate editor coordinating the review of this letter and approving it for publication was J. Prieto. *(Corresponding author: Haneya Naeem Qureshi.)*

The authors are with The University of Oklahoma, Tulsa, OK 74135 USA (e-mail: haneya@ou.edu).

Digital Object Identifier 10.1109/LCOMM.2019.2899094

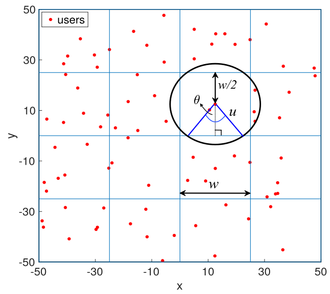


Fig. 1. A portion of the coverage area showing bin width (w) and positioning error radius (u).

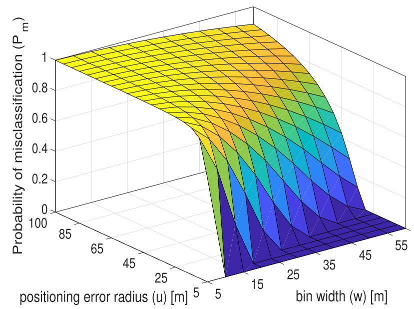


Fig. 2. Probability of misclassification of a user with varying bin width and positioning error radius.

Each base station serves users located in an area of am $\times$ am. Users are distributed according to square point picking process, i.e., two independent sets of points x and y are picked from a uniform distribution and placed at coordinates (x, y). The area served by each base station is further divided into $m \times m$ bins of width w. We assume that the probability density function of the distance and direction of the UEs actual location with respect to its reported position are $\frac{1}{u}$ and $\frac{1}{2\pi}$ respectively. Therefore, given a reported UE position, its actual location is within a circular disc with radius u which is centered at the reported UE position, as illustrated in Fig. 1 for one user. Therefore, the actual position of the ith UE with coordinates (x_i, y_i) can be modeled as $(x_i + u\sqrt{q_i}\cos(2\pi v_i), y_i + u\sqrt{q_i}\sin(2\pi v_i))$, where v_i and q_i are one realization of pseudo random, pseudo independent numbers uniformly distributed in [0, 1].

We consider a small cell environment where propagation conditions are mostly dominated by line of sight. Since MDT measurement reports are based on long term averaged received power [5], fast fading can be considered to be averaged out. Therefore, only the shadowing and path loss effects are taken into consideration in our analysis. The signal propagation model we employ for obtaining the received power measurement is as follows:

$$S\ [\text{dBm}] = T + K - 10\ n\log_{10}\left(\frac{d}{d_o}\right) + X \tag{1}$$

where T is the transmit power in dBm, K is a constant in dB that depends on the antenna characteristics and the average channel attenuation and can be quantified as $-20\log_{10}(4\pi d_o/\lambda)$. n refers to the path loss exponent, d_o is the reference distance and d is the distance between the user and serving BS. The shadowing effect is modeled by the random variable, X which follows a zero mean Gaussian distribution with standard deviation ϕ in dB. A bin or user is considered to be in coverage when its received signal strength is greater than a predefined threshold, γ.

III. Autonomous Coverage Estimation Framework in the Presence of Errors

A. Quantifying User Positioning and Quantization Errors

In this section, we first present insights and methods to quantify the individual effects of user positioning and quantization errors in coverage estimation, followed by quantification of the concurrent effect of these errors in order to determine the optimal bin width.

1) User Positioning Error in the Presence of Bins: Consider the scenario in which the predicted coverage area is divided into $m \times m$ bins. Gathered coverage data from different bins can be represented in a matrix $\boldsymbol{R}$ of dimensions $m \times m$. Thus, the coverage area forms a square matrix, $\boldsymbol{R} \in \mathbb{R}^{(m\times m)}$, where each entry of this matrix, r_i contains the averaged received power in that bin, where $i = 1, ...m^2$. Therefore, $\boldsymbol{R}^{P,Q}$ is a matrix containing measured average received power of users due to positioning uncertainty and $\boldsymbol{R}^{P',Q}$ contains the average received power of users with no positioning uncertainty.

To understand the impact of user positioning error as a function of bin width, consider a user located at the bin center, with coordinates (12.5,12.5) as shown in Fig.1. In the presence of no positioning uncertainty, the user is actually present at this location. However, due to positioning uncertainty, the actual location of the user lies within a circular radius u. Depending on the radius u and bin width, w, the probability of user being actually located in adjacent bins would vary, which would impact coverage estimation. We define this probability of misclassification, P_m as the probability that user's actual position lies in bin j, given that its reported position lies in bin i, where $i \neq j$. Using geometry from Fig.1, three cases of P_m can be distinguished depending on u. By expressing $\theta = 2\cos^{-1}(w/2u)$ and calculating the fraction of area of circle with radius u that lies outside the square with side w, or equivalently, calculating the fraction of user's all possible actual locations that lie outside bin i, P_m when a user is located at the i-th bin center can be derived as follows:

$$\mathrm{P}_m(w,u) = \begin{cases} 0, & 0 < u \le w/2 \\ \dfrac{4u^2\cos^{-1}\left(\frac{w}{2u}\right) - 2u^2\sin\left(2\cos^{-1}\left(\frac{w}{2u}\right)\right)}{\pi u^2}, & w/2 < u < w/\sqrt{2} \\ \dfrac{\pi u^2 - w^2}{\pi u^2}, & u \ge w/\sqrt{2} \end{cases} \tag{2}$$

P_m as a function of u and w is illustrated in Fig. 2. Note that the case when a user is located at the bin center is a lower bound on P_m as P_m will increase as the user moves away from the bin center. Therefore, for any arbitrary user location, the error in coverage estimation due to positioning error in the presence of bins is likely to increase with larger u for the same w or with smaller w for the same u, as the probability of misclassification would increase in these scenarios. It is observed from Fig. 2 that a zero probability of user location being misclassified occurs at the combination of large bin width and small positioning error radius. Note that the RSRP perceived by the users is affected by positioning error since the measured RSRP reports are tagged to wrong locations due to positioning error. This results into error (caused by tagging to wrong location) in the RSRP-location duo reported as part of the MDT reports. This leads to error in the coverage being investigated here. Therefore, the error in coverage estimation due to positioning uncertainty is expected to be the least when bin width is large and positioning error radius is small. In order to capture this effect, we quantify the impact of user positioning error in the presence of bins as follows:

$$E^P = \frac{1}{m^2}\sum_{i=1}^{m^2} |\mathbb{P}[\boldsymbol{r}_i^{P,Q} > \gamma] - \mathbb{P}[\boldsymbol{r}_i^{P',Q} > \gamma]| \tag{3}$$

where the operator $\mathbb{P}$ represents probability, $\boldsymbol{r}^{P,Q}$ and $\boldsymbol{r}^{P',Q}$ are vectorized forms of matrices $\boldsymbol{R}^{P,Q}$ and $\boldsymbol{R}^{P',Q}$ respectively. The i-th element of the vector $\boldsymbol{r}^{P,Q}$, $r_i^{P,Q}$ represents the measured average received power of users in i-th bin in presence of positioning uncertainty and $r_i^{P',Q}$ is the average received power of users in the same bin with no uncertainty.

2) Quantization Error in the Presence of Positioning Uncertainty: In order to quantify the effect of quantization error as a function of positioning error radius, u, it is necessary to analyze the coverage values at the user level for benchmark to investigate the effect of binning. Let $\boldsymbol{r}^{P,Q}$ be the measured received power vector of U users within a cell in the presence of quantization and positioning error and $\boldsymbol{r}^{P,Q'}$ be measured received power vector of those U users without any quantization but the same positioning error. Then the error due to quantization in the presence of a certain positioning error can be quantified as:

$$E^Q = \frac{1}{U}\sum_{i=1}^{U} |\mathbb{P}[\boldsymbol{r}_i^{P,Q} > \gamma] - \mathbb{P}[\boldsymbol{r}_i^{P,Q'} > \gamma]| \tag{4}$$

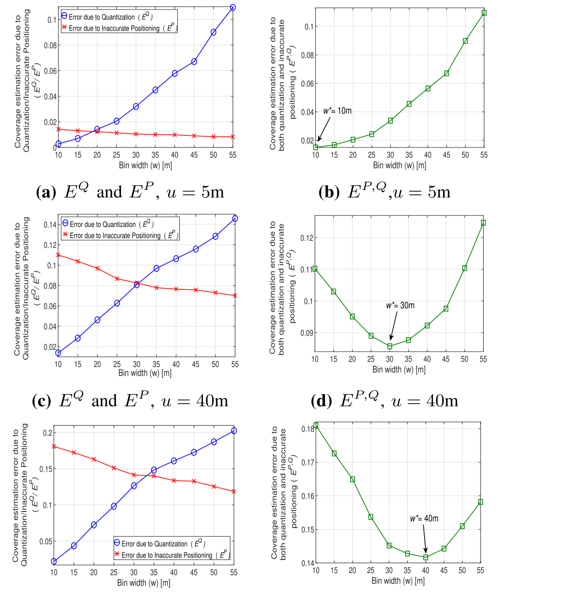
3) Combined Effect of Quantization and Positioning Error: Finally, to quantify the effect of both positioning error and quantization error on coverage estimation, we consider the benchmark to be the received power vector at the user level without any positioning uncertainty (i.e., the user reporting RSRP value from a particular location is actually present at that location), $\boldsymbol{r}^{P',Q'}$. Then the combined effect of both positioning and quantization errors on coverage estimation can be quantified as:

$$E^{P,Q} = \frac{1}{U}\sum_{i=1}^{U} |\mathbb{P}[\boldsymbol{r}_i^{P,Q} > \gamma] - \mathbb{P}[\boldsymbol{r}_i^{P',Q'} > \gamma]| \tag{5}$$

B. Determining Optimal Bin Width

In order to determine the optimal bin width, we want to minimize the total quantization and positioning error. The optimization problem can then be formulated as:

$$\begin{aligned} w^* = {} & \arg\min_{w} \; E^{P,Q} \\ & \text{subject to } w_{min} \leq w \leq w_{max} \\ & \text{GPS error radius} = u \end{aligned} \tag{6}$$



(a) E^Q and E^P, $u = 5$m **(b)** $E^{P,Q}$,$u = 5$m
(c) E^Q and E^P, $u = 40$m **(d)** $E^{P,Q}$, $u = 40$m
(e) E^Q and E^P, $u = 70$m **(f)** $E^{P,Q}$, $u = 70$m

Fig. 3. Errors in coverage estimation for different positioning error radius and $\gamma = -80$dBm.

where the expectation is taken over random variables, x, y, v, q and X. Owing to the small search space, we can solve (6)-(6) via brute force as shown in the next section.

IV. Simulation Results and Analysis

In our simulations, we distribute U users/cell in a system of 9 cells according to square point picking process. The actual position of the ith UE with coordinates (x_i, y_i) is generated as $(x_i + r\sqrt{q_i}\cos(2\pi v_i), y_i + r\sqrt{q_i}\sin(2\pi v_i))$, where v_i and q_i are drawn from a uniform random distribution, $\mathcal{U}[0, 1]$. We consider three different user densities, $U = 3000, 5000$ and 7000. Other simulations parameters are set as follows: $n = 3.5$, $d_o = 1$m, $T = 40$ dBm, $\phi = 4$dB, $N = 9$, $A = 1200$m, $a = 400$m, $w_{min} = 10$m, $w_{max} = 55$m and u is varied from 0m to 70m. Monte-carlo simulations are done over the random variables, x, y, v, q and X.

The error in coverage estimation due to quantization error and incorrect user positioning is shown in Fig. 3 (a), (c), (e) for $u = 5$m, 40m and 70m respectively. On one hand, the coverage estimation error due to quantization increases with increase in bin width owing to greater averaging of user reported measurements as bin width increases. On the contrary, error due to incorrect user positioning decreases with increase in bin with attributing to the fact that for a given positioning error radius, a larger bin width would mean a lesser probability that a particular user is in fact present in adjacent bins as previously illustrated by Fig.2. This trade-off leads to the curves in Fig. 3 (b), (d) and (f). Note that the coverage estimation error due to positioning error is very small in case of $u = 5$m as compared to the quantization error, therefore, the total error in Fig. 3(b) for $u = 5$m is dominated by the quantization error. However, as positioning error increases, it acts as an opposing factor to the increasing quantization error with increasing bin width, yielding an optimal bin width as shown in Fig. 3 (d) and (f).

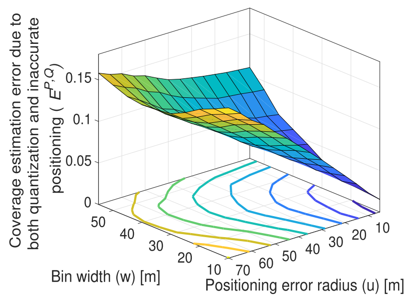


Fig. 4. Total error with varying bin width and positioning error radius.

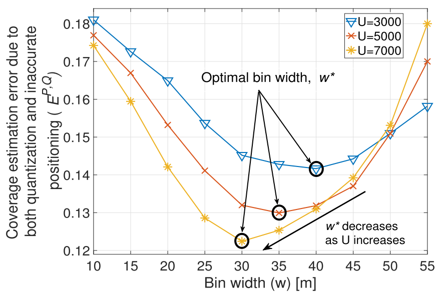


Fig. 5. Total coverage estimation error with varying user densities for $u = 70$m, $\gamma = -80$dBm.

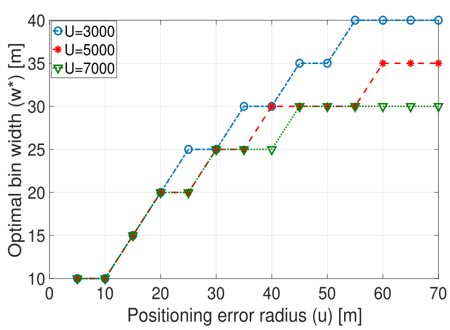


Fig. 6. Optimal bin width with varying positioning error radius.

A 3D graph showing the effect of both errors simultaneously as a function of positioning error radius and bin width is shown is Fig. 4. It is observed that the optimal bin width that minimizes this error increases as positioning error radius increases. The analysis is extended to different user densities. Fig. 5 shows the total error in coverage estimation with varying user densities for $u = 70$m. It can be seen that as the user density increases, the optimal bin width decreases attributing to the decrease in effect of coverage estimation error caused by positioning inaccuracy. This is because the average of users in any bin is likely to differentiate by a small margin when number of users in that bin are more as compared to the scenario with less number of users.

The resulting optimal bin width as a function of positioning error radius is shown in Fig. 6 for different user densities. It is observed that the optimal bin width increases as positioning error radius increases in a non-linear manner. Note that although the overall coverage estimation error increases with increase in positioning error radius, the optimal bin width may be same for certain ranges of positioning error radii. This is because despite increase in $E^{P,Q}$ with increase in u, the bin width which results in minimum $E^{P,Q}$ is same for those positioning error radii. These findings can be used by a network operator to determine the optimal bin width for a given positioning accuracy and user density, that would result in minimum error in MDT-based coverage estimation. This would lead to more accurate coverage estimation, which can then be utilized to design and optimize several aspects of the network, such as minimize total cost of ownership, boost network capacity, detect coverage holes, maximize coverage, minimize power consumption and even optimize handover zones [10].

V. Conclusion

By quantifying the errors in MDT-based coverage estimation that stem from quantization and inaccurate user positioning, we show that there exists an optimal bin width that can be determined to minimize the combined effect of these errors in MDT based coverage estimation. Optimal bin width that minimizes the effect of these errors concurrently is determined as a function of positioning error radius and user density. Thus, for given positioning accuracy and user density, the findings from this study can be directly used by network operators to configure the bin size that results in most accurate MDT based coverage estimation. Depending on the scenario under consideration, framework presented in this study can be extended to varying base station distributions and shadowing standard deviations. Such investigations will be focus of our future work.

VI. Acknowledgment

For more details, please visit: http://www.ai4networks.com.

References

[1] O. G. Aliu, A. Imran, M. A. Imran, and B. Evans, “A survey of self organisation in future cellular networks,” *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 336–361, Feb. 2013.

[2] I. Akbari, O. Onireti, A. Imran, M. A. Imran, and R. Tafazolli, “Impact of inaccurate user and base station positioning on autonomous coverage estimation,” in *Proc. IEEE 20th Int. Workshop Comput. Aided Modelling Design Commun. Links Netw. (CAMAD)*, Sep. 2015, pp. 114–118.

[3] A. Galindo-Serrano, B. Sayrac, S. B. Jemaa, J. Riihijärvi, and P. Mähönen, “Harvesting MDT data: Radio environment maps for coverage analysis in cellular networks,” in *Proc. 8th Int. Conf. Cogn. Radio Oriented Wireless Netw.*, Jul. 2013, pp. 37–42.

[4] J. D. Naranjo, A. Ravanshid, I. Viering, R. Halfmann, and G. Bauch, “Interference map estimation using spatial interpolation of MDT reports in cognitive radio networks,” in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2014, pp. 1496–1501.

[5] I. Akbari, O. Onireti, A. Imran, M. A. Imran, and R. Tafazolli, “How reliable is MDT-based autonomous coverage estimation in the presence of user and BS positioning error?” *IEEE Wireless Commun. Lett.*, vol. 5, no. 2, pp. 196–199, Apr. 2016.

[6] I. Akbari, O. Onireti, M. A. Imran, A. Imran, and R. Tafazolli, “Effect of inaccurate position estimation on self-organising coverage estimation in cellular networks,” in *Proc. 20th Eur. Wireless Conf.*, 2014, pp. 1–5.

[7] P. Bernardin and K. Manoj, “The postprocessing resolution required for accurate RF coverage validation and prediction,” *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1516–1521, Sep. 2000.

[8] F. Sohrabi and E. Kuehn, “Construction of the RSRP map using sparse MDT measurements by regression clustering,” in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[9] P.-C. Lin, “Minimization of Drive Tests using measurement reports from user equipment,” in *Proc. IEEE 3rd Global Conf. Consum. Electron. (GCCE)*, Oct. 2014, pp. 84–85.

[10] A. Taufique, M. Jaber, A. Imran, Z. Dawy, and E. Yacoub, “Planning wireless cellular networks of future: Outlook, challenges and opportunities,” *IEEE Access*, vol. 5, pp. 4821–4845, 2017.