DATA-AIDED FAST BEAMFORMING SELECTION FOR 5G

João Gante[⋆] *Gabriel Falcão*[†] *Leonel Sousa*[⋆]

[⋆] INESC-ID, IST, Universidade de Lisboa, Portugal
[†] Instituto de Telecomunicações, University of Coimbra, Portugal

ABSTRACT

Millimeter wave frequencies paired up with MIMO antennas employing beamforming are seen as critical enablers of next generation networks. However, selecting the most beneficial beamforming weights in a codebook-enabled downlink transmitter is a lengthy task, as the existing methods rely on some form of channel measurement. In fact, if the used codebook is too large, the traditional methods might fail to select an appropriate entry within the channel coherence time.

In this paper, a new method to assist the beam selection is proposed, based on data obtained from previous connections. Through the continuous update of the set of optimal codebook entries for each position, the search space required for each connection can be greatly reduced if the user position is known. The simulations performed show that retrieving the sets of codebook entries in single user scenarios required less than $51\ ns$. For multi-user scenarios, results exceeding 10 simultaneous users using a 16 entry codebook were achieved, requiring less than $600\ \mu s$. The obtained results show that the proposed method can greatly reduce the beam selection latency and energy requirements, opening the door to powerful millimeter wave networks.

***Index Terms*—** 5G, mmWaves, Codebook, Beamforming

1. INTRODUCTION

The advent of 5G related research reopened the door to the so called millimeter wave (mmWave) frequencies, unlocking a huge chunk of untapped bandwidth [1]. However, some of the existing challenges include achieving non line-of-sight (NLOS) communications, due to reflections on visible obstacles [2], and counteracting these frequencies severe path loss properties [3]. As multiple studies have shown, massive antenna arrays can be employed so as to counteract the aforementioned issues, with the aid of beamforming (BF) [4] [5]. When line-of-sight (LOS) communications are unattainable, the focused beam can be aimed towards obstacles, such that its reflection reaches the desired target only. With careful execution, spatial multiplexing is attainable by simultaneously transmitting multiple streams with reduced interference levels [3]. While the BF and the reflection concepts are straightforward, one question arises: how should the resulting beam be formed? Either the transmitter has some channel state information (CSI), and chooses the most beneficial BF weights according to that information [6], or it doesn't, and needs to perform some beam search algorithm beforehand [7].

Selecting the beamforming weights based on CSI allows for optimal performance [3]. Unfortunately, such methods require a complete or partial estimation of the channel matrix H, usually using pilots. Furthermore, if there is a high path loss between the transmitter and the receiver, it might not be possible to estimate the channel without actually performing some kind of beam search to boost the SNR [8], further hindering these methods' applicability. Thus, the complexity for this type of beam selection method grows with channel estimation, which can quickly become overwhelming as the number of antennas increase.

The beam search methods, while considerably easier to implement, add a significant (time and energy) overhead by probing the codebook entries before establishing a connection [9]. Unfortunately, to harness the spatial diversity benefits of increasingly larger antenna arrays, the codebook size must grow accordingly. In the presence of a small channel coherence time, it might even be impossible to scan all the entries in a large codebook [10]. Thus, the system employing traditional beam search methods must have a small codebook in order to be effective, limiting the BF capabilities.

Since the mmWave propagation is constrained by the obstacles found, a receiver with fixed positioning should measure a constant average received power for each codebook entry, given a constant set of obstacles. For 5G base stations, which are expected to be located in elevated positions in urban scenarios, most of the obstacles are static for a significant amount of time, as they are mostly buildings. Therefore, once found, the set of optimal codebook entries for a given position is expected to remain valid for multiple future connections. Thus, by leveraging such information, we propose to narrow down the beam search procedure, suggesting a set of likely viable solutions if the receiver position is known. With such search reduction, bigger codebooks are enabled, which combined with a proper design (e.g. [11]) can lead to simple, yet powerful solutions. To initialize (and to simulate) the proposed method, the use of ray-tracing techniques [12–14] on accurate 3D maps is also proposed.

The remaining of this paper is organized as follows. In section 2, the system is described and the problem is presented, while Section 3 introduces the proposed algorithms to tackle those problems. Section 4 assesses the used ray-tracing software's accuracy, as well as the proposed algorithm's feasibility. Finally, the conclusions are drawn in section 5.

2. SYSTEM MODEL AND PROBLEM FORMULATION

Consider a system containing a single stationary transmitter, with an antenna array made up of N_T elements, and a set of N_S mobile receivers, totaling N_R receiving antennas. In the frequency domain, we can write the signal at the receiver devices, $\mathbf{y} \in \mathbb{C}^{N_S \times 1}$, as

$$\mathbf{y} = \mathbf{WHFx} + \mathbf{Wz}, \tag{1}$$

where $\mathbf{W} \in \mathbb{C}^{N_S \times N_R}$ and $\mathbf{F} \in \mathbb{C}^{N_T \times N_S}$ are the beamforming matrices for the receivers and transmitter, respectively, $\mathbf{H} \in \mathbb{C}^{N_R \times N_T}$ is the channel matrix, $\mathbf{x} \in \mathbb{C}^{N_S \times 1}$ is the transmitted signal, and $\mathbf{z} \in \mathbb{C}^{N_R \times 1}$ is the noise. Ultimately, to achieve multi-user spatial-

Acknowledgments: This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2013, and FCT Grant No. FRH/BD/103960/2014.

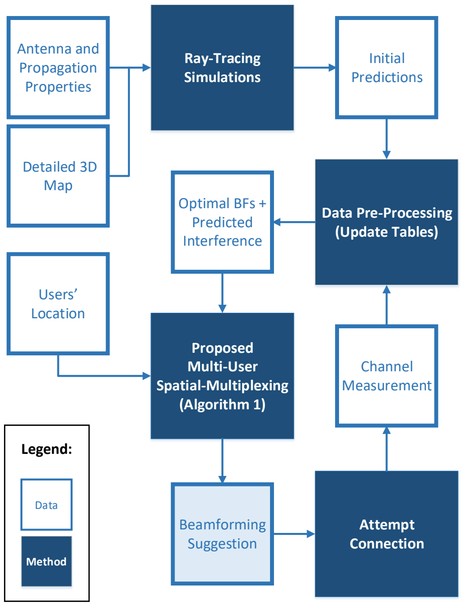


Fig. 1. Relatioship between the used data and methods, resulting in the beamforming suggestions.

multiplexing, the diagonal elements of $\mathbf{WHF}$ should be maximized, while keeping the non-diagonal elements (unwanted interference) at check.

When the information regarding the receivers' antenna arrays is unavailable (or inconvenient to obtain), such as in the establishment of a new connection, the system is unable to jointly optimize both beamforming matrices. In such scenarios, the dimensions of $\mathbf{W}$ are unknown, as the system doesn't know N_R. As result, the best option left for the transmitter is to maximize the desired signals at the receivers' locations and, in that case, $\mathbf{W}$ is a N_S by N_S identity matrix ($\mathbf{I}$). Under those circumstances, equation 1 can be simplified to $\mathbf{y} = \mathbf{HFx} + \mathbf{z}$, where each element of the $\mathbf{HF}$ product corresponds to the power of the j-th signal at the position of the i-th user.

Since $\mathbf{H}$ is sparse in the angular domain [15], it is then heavily influenced by the propagation path. The method introduced in this paper indirectly estimates the channel matrix ($\hat{\mathbf{H}}$) through the measurement of its past instances, which should match the angular sparseness of $\mathbf{H}$ if the obstacle layout remains static. Through this method, $\hat{\mathbf{H}}$ has two major sources of unpredictable distortions: (*i*) incorrect device position estimation and (*ii*) obstacle layout changes.

When the beamforming is defined by finite codebook $\mathbf{C}$, each column of $\mathbf{F}$, $\mathbf{f}(i)$, must take the form of a codebook entry. We can thus formulate our codebook-based constrained optimization problem as follows:

$$\begin{aligned}
\underset{\mathbf{F}}{\text{maximize}} \quad & \text{trace}(\hat{\mathbf{H}}\mathbf{F}) \\
\text{subject to:} \quad & \mathbf{f}(i) \in \mathbf{C}, \\
& (\hat{\mathbf{H}}\mathbf{F})_{i,j} < I_{th},\ i \neq j, \\
& (\hat{\mathbf{H}}\mathbf{F})_{i,j} \geq P_{th},\ i = j,
\end{aligned}$$

where I_{th} is the maximum allowable interference and P_{th} is the minimum acceptable power for the communication.

3. PROPOSED ALGORITHMS

Upon performing the ray-tracing simulations for all propagation patterns in codebook $\mathbf{C}$, the initial estimation of the received power for

Algorithm 1 Multi-User Spatial Multiplexing

I_th = maximum interference threshold;
S = # of top codebook entries to be saved (per location);
L = spatial grid size;
U = # of users attempting to connect;
Load the received power matrix $\mathbf{P}$ (to compute the interferences);
Load the sorted optimal $\mathbf{B}$ matrix, with L by S elements (obtained in the data pre-processing);
Initialize table T with the connected users' location and the respective codebook entry selected;
for $u \leftarrow 0$ **to** $U-1$ **do**
  Receive the location of user u;
  Convert the received location into the correspondent grid position l;
  for $s \leftarrow 0$ **to** $S-1$ **do**
    if $B[l][s] = invalid\_flag$ **do**
      break;
    else if $B[l][s] \in T$ **do**
      skips this codebook entry;
    end if
    $I = max$ (interference on users $\in \{T \cup u\}$) ;
    if $I \leq I\_th$ **do**
      Adds user u and the codebook entry $B[l][s]$ to T;
    end if
  end for
end for

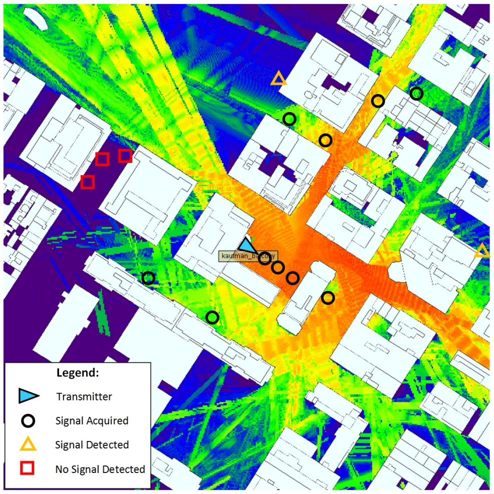
each position in a discrete spatial grid and for all beamforming patterns is obtained, henceforth called matrix $\mathbf{P}$. Since the resulting beam patterns are directive, each location ends up having just a few codebook entries above the minimum power threshold. Therefore, in the data pre-processing stage, a table with the S most suitable codebook entries for each location is generated ($\mathbf{B}$). Since the key task of the beam search method is to obtain the best codebook index for a given connection, the table entries for each location are sorted (in descending order) according to the received power, aiming to minimize the required search. Finally, if a given location has less than S valid codebook entries, a special "invalid flag" is stored at the empty slots. During the beam search process, feedback regarding the tested beam pattern suitability is obtained. Through that feedback, $\mathbf{P}$ is updated using a moving average, allowing the proposed system to cope with changes in the surrounding area.

In order to achieve spatial multiplexing, the transmitter must be confident that the radiation aimed at a target user does not cause too much interference on the remaining connected users. As result, achieving spatial multiplexing is far more difficult than to perform single user beamforming, and the complexity grows exponentially with the number of connected users.

To select the codebook entry for a new user, two separate tasks must be executed: *(i)* assessing which beams might be capable of establishing a connection with the newcomer, knowing the existing interferences that it must be able to sustain, and *(ii)* estimating the additional interference created by the new connection on the previously existing users. The connection can only be established after verifying that all the interference levels are below a target threshold.

The optimization formulation on section 2 covers all the aforementioned concerns. However, since the execution time of the whole beam search is critical and searching through all possible combinations of $\mathbf{F}$ would quickly exceed the timing requirements as the number of users increases, an heuristic was created. While adding sequentially new users, the system will expect that the existing users have locked their beamforming configuration, drastically reducing the required search space while keeping a near-optimal solution.

Considering that the data pre-prossessing step sorted the code-



(a) Propagation in the NYU area. The results shown correspond to the *maximum received power for all possible transmit directions*.

(b) Propagation in the Wall Street area. The results shown correspond to the propagation for the represented transmit direction.

Fig. 2. Ray Tracing simulations for two distinct zones in New York City (Manhattan). As it is observable, the propagation depends entirely on the layout of the area. The power scale starts at -140dBm and ends at -35dBm.

book entries by the resulting received power, the system can simply check the problem restrictions for the sequential entries, until it finds a suitable one[1]. This method, depicted in algorithm 1, will guarantee the optimal solution for the aforementioned problem, provided that the number of saved codebook entries per location is large enough to include a viable entry (if it exists).

Since obtaining a small amount of data from the mobile user (*e.g.* its GPS location) adds a very small delay [9], the resulting latency will be due to the algorithm's execution. With a low execution time, the proposed method is thus able to quickly produce a set of beamforming suggestions that might be tried off before the system resorts to heavier algorithms. As result, adding this method to a

[1]As mentioned before, each element of the $\mathbf{HF}$ product corresponds to the power of the j-th signal at the position of the i-th user, so the interferences can be estimated through $\mathbf{P}$.

Table 1. Simulation Specifications

Parameter Name	Value
Carrier Frequency	28 GHz
Transmit Power	30 dBm
Max. Tx. Gain	24.5 dBi (horn antenna)
HPBW	10.9°
Downtilt	10°
Codebook Size	16 (150° arc with 10° between entries)
Saved BF	4 (per receiver location)
Receiver Grid Size	160801 (400 × 400 m, 1 m between receiver)
# of Executions	10^6

working system should reduce its latency, which should also lead to energy savings.

4. SIMULATIONS AND EXPERIMENTAL RESULTS

4.1. Apparatus and Ray-Tracing Accuracy

Before addressing the results for the proposed algorithm, a major assumption must be verified regarding the reliability of ray-tracing simulations. In order to do so, a recreation of the experimental results obtained in [16] is devised. In [16], directive horn antennas are used to measure the propagation characteristics of the 28 GHz frequency, near the NYU campus. By aiming the transmitter in the direction with the strongest link and by physically rotating the receiver (both in the azimuth plane and in altitude), the authors in [16] were able to assess whether it was possible to establish a communication link with those transmitter-receiver location pairs.

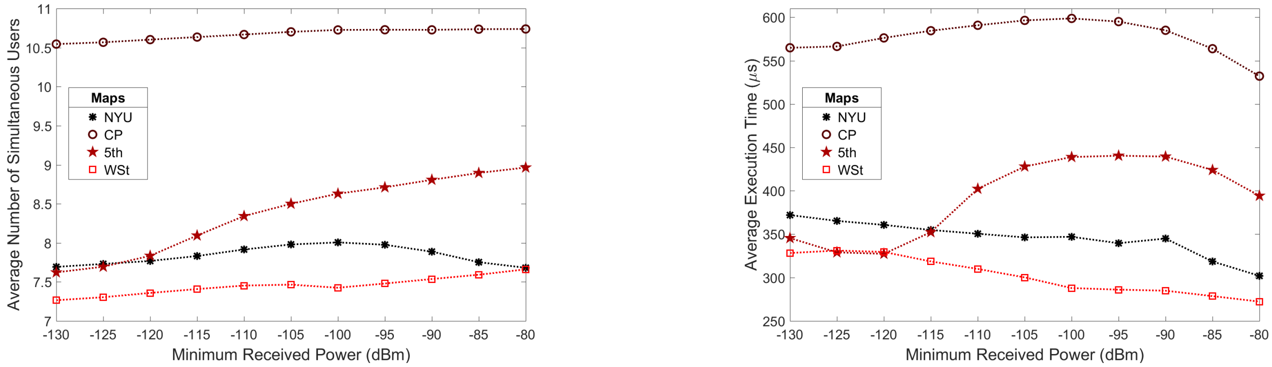
Using the high precision open-source 3D map made available by the New York City Department of Information Technology & Telecommunications [17], as well as the specifications depicted in table 1, the measurements were recreated using the Wireless InSite 3.0.0.1. To compute the received power, the rotating horn antenna receivers were swapped by isotropic receivers which, aside from the antenna gain, should yield similar results. In the actual field measurements, using horn antennas as receivers, successfully received or detected signals correspond to a maximum path loss of 168 or 178 dB, respectively. Considering the aforementioned parameters, this corresponds to a minimum received power of -113.5 or -123.5 dBm at the simulated isotropic receivers.

The recreation's result can be observed in Fig.2a, where the maximum received power for all possible transmit directions is shown. In the results shown, a grid of isotropic receivers was placed on the map, 1 m above the ground, with 1 m of separation between antennas. The colored markings represent the measurement locations and the respective results. As it can be observed inside the markings, the simulated received power matches the experimental measurements. Furthermore, similar results were also obtained in [18] (using a different ray-tracing software), and thus validating the proposed assumption.

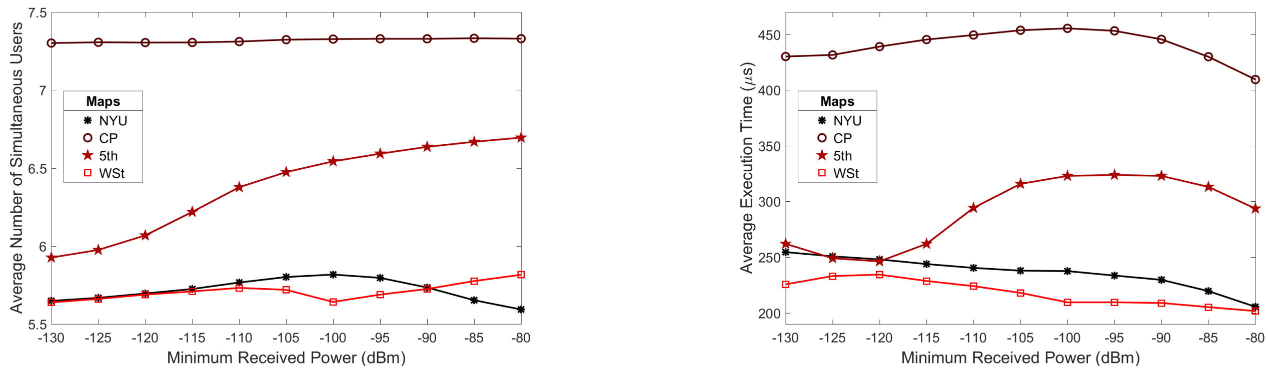
4.2. Simulation Results

Since the dynamic changes in the obstacles are hard to model, this sub-section focuses on the feasibility of the algorithm (its execution time). Unfortunately, the used ray-tracing software doesn't fully support MIMO antennas with beamforming, and thus a physically rotating horn antenna (with the specifications depicted in table 1) is used as transmitter. Even though it is an imperfect representation, the used specifications are similar to the MIMO antenna used in [9].

By predicting the propagation for each entry in the codebook, data pre-processing can then be executed, resulting in a very compact data structure containing the most suitable codebook entries for each location. With the aforementioned data structure and the realistic



(a) Simulation results for a maximum interference of **-5 dB**.



(b) Simulation results for a maximum interference of **-10 dB**.

Fig. 3. Simulation results for 100 users attempting to connect. The maps in the figures' legend correspond to the areas shown in Fig. 2, plus the Central Park (CP) and the 5th Avenue (5th). As the maximum allowed interference increases, so does the number of simultaneous users, which in turn is correlated to the algorithms' execution time (never exceeds $600\mu s$).

simulation specifications presented in table 1, the average execution time for a system with a single randomly positioned user running algorithm 1 was under $51ns$. Considering the obtained results, the proposed single user beamforming method's bottleneck would certainly be time required to attempt the connection given the codebook entry. The results shown do not change with the minimum received power threshold nor with the user locations.

When multi-user spatial multiplexing is preferred to single user beamforming, the algorithm's complexity increases drastically. Besides fetching the optimal codebook entries for each new user, the interference for the multiple users must now be computed, which also requires the full propagation data. As the number of connected users increases, not only the interference computing step gets more demanding but also the likelihood that the new user is able to fit in with its main codebook entries reduces, resulting in increased execution times for each additional user. To mitigate that increasingly difficult interference computing step, algorithm 1 stops checking for the connected users' interferences as soon as it finds one above the threshold.

Considering the parameters in table 1, the average results for 100 randomly positioned users attempting to connect are shown in Fig. 3. To eliminate any bias due to a specific city layout, four different city zones were considered, as partially shown in Fig. 2. As previously discussed, the execution time increases considerably with the number of successful connections, showing an exponential growth. It is also possible to observe that, regarding the number of simultaneous connections, the maximum allowed interference has significantly more impact than the minimum received power threshold. Most zones had similar results, except for the Central Park (the only zone consisting of mostly open space), which consistently allowed more simultaneously connected users.

Even for a low interference threshold, the multi-user spatial multiplexing algorithm had an average of at least 5.5 simultaneous users for the used parameters. In every situation, the average execution time for the 100 users attempting to connect did not exceed $600\mu s$. While the required execution time is far higher than for the single user beamforming, it is still able to comply with the strict 5G requirements of a few milliseconds (given that the communications with the user take negligible time). Therefore, the proposed method has the potential to reduce the required connection time, while keeping a high area spectral efficiency for millimeter waves.

5. CONCLUSIONS

In this paper, a new set of algorithms are proposed to create a set of likely viable beamforming configurations, so as to aid the beamforming selection process. The achieved results show that the single user beamforming suggestions are formulated negligible latency, while the multi-user spatial multiplexing method is able to run in under a millisecond. While these suggestions are prone to errors, the global result should yield a considerable gain, since they greatly reduce the required search space. With the reduced search space, bigger codebooks become possible, allowing for higher area spectral efficiency capabilities and paving the way for next generation networks.

6. REFERENCES

[1] P. Pirinen. A brief overview of 5g research activities. In *5G for Ubiquitous Connectivity (5GU), 2014 1st International Conference on*, pages 17–22, Nov 2014.

[2] T. S. Rappaport et al. Millimeter wave mobile communications for 5g cellular: It will work! *IEEE Access*, 1:335–349, 2013.

[3] T. S. Rappaport et al. *Millimeter Wave Wireless Communications*. Prentice Hall, 2014.

[4] J. C. Liberti and T. S. Rappaport. *Smart Antennas for Wireless Communications: IS-95 and Third Generation CDMA Applications*. Prentice Hall, 1999.

[5] C. D. Durgin. *Space-Time Wireless Channels*. Prentice Hall, 2003.

[6] S. Sun, T. S. Rappaport, R. W. Heath, A. Nix, and S. Rangan. Mimo for millimeter-wave wireless communications: beamforming, spatial multiplexing, or both? *IEEE Communications Magazine*, 52(12):110–121, December 2014.

[7] W. Roh, J. Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar. Millimeter-wave beamforming as an enabling technology for 5g cellular communications: theoretical feasibility and prototype results. *IEEE Communications Magazine*, 52(2), February 2014.

[8] T. E. Bogale, L. B. Le, and X. Wang. Hybrid analog-digital channel estimation and beamforming: Training-throughput tradeoff. *IEEE Transactions on Communications*, 63(12), Dec 2015.

[9] Y. Kim, H. Y. Lee, P. Hwang, R. K. Patro, J. Lee, W. Roh, and K. Cheun. Feasibility of mobile cellular communications at millimeter wave frequency. *IEEE Journal of Selected Topics in Signal Processing*, 10(3), April 2016.

[10] E. Bjornson, E. G. Larsson, and T. L. Marzetta. Massive mimo: ten myths and one critical question. *IEEE Communications Magazine*, 54(2):114–123, February 2016.

[11] Ju-Hong Lee and Jing-Yen Lee. Optimal beamforming-selection spatial precoding using population-based stochastic optimization for massive wireless mimo communication systems. *Journal of the Franklin Institute*, 354(10):4247–4272, 2017.

[12] M. Dong et al. Simulation study on millimeter wave 3d beamforming systems in urban outdoor multi-cell scenarios using 3d ray tracing. In *2015 IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Aug 2015.

[13] S. Baek et al. Comparison analysis of outdoor channel characteristics at 28 ghz and 2 ghz using 3d ray-tracing technique. In *2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall)*, Sept 2014.

[14] C. Wang et al. Ray-tracing based performance evaluation of 5g mmwave massive mimo in hotspots. In *2016 International Symposium on Antennas and Propagation (ISAP)*, Oct 2016.

[15] J. Brady and A. Sayeed. Beamspace mu-mimo for high-density gigabit small cell access at millimeter-wave frequencies. In *2014 IEEE 15th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 80–84, June 2014.

[16] H. Zhao et al. 28 ghz millimeter wave cellular communication measurements for reflection and penetration loss in and around buildings in new york city. In *2013 IEEE International Conference on Communications (ICC)*, pages 5163–5167, June 2013.

[17] Open source new york 3d building model. `http://www1.nyc.gov/site/doitt/initiatives/3d-building.page`. Accessed: 15th of April, 2017.

[18] Z. Zhang, J. Ryu, S. Subramanian, and A. Sampath. Coverage and channel characteristics of millimeter wave band using ray tracing. In *2015 IEEE International Conference on Communications (ICC)*, pages 1380–1385, June 2015.