Outage Detection for Millimeter Wave Ultra-Dense HetNets in High Fading Environments

Ahmad Asghar, Hasan Farooq and Ali Imran
Department of Electrical and Computer Engineering
The University of Oklahoma, Tulsa, Oklahoma 74135, USA
Email: {ahmad.asghar, hasan.farooq, ali.imran}@ou.edu

***Abstract*—Millimeter wave spectrum utilization and network densification are two of the fundamental technologies that will enable high user quality of experience required in 5th Generation mobile cellular networks. However, user sparsity in ultra-dense heterogeneous networks and coverage limitations of millimeter wave cells means reliability of such networks will become a key operational challenge. Recent studies have explored the use of machine learning techniques for outage detection in legacy and heterogeneous mobile cellular networks. However, machine learning techniques are highly susceptible to noise in the training data which can affect their outage detection accuracy. To counter these challenges, we present a novel outage detection method based on entropy field decomposition technique first introduced in [1]. The proposed method is able to detect cell outages with at least 96% accuracy even as the level of shadowing in the network is increased which makes it ideal for practical implementation in emerging ultra-dense heterogeneous networks with millimeter wave cells. The proposed solution is compared against k-means clustering for outage detection with results showing that not only does entropy field decomposition return higher true positive results, it also returns fewer false positive results compared to k-means clustering.**

***Index Terms*—Heterogeneous Network; Network Reliability; Millimeter Wave Cells; Entropy Field Decomposition**

I. Introduction

Chief among the requirements for 5th Generation (5G) mobile cellular networks (MCNs) is the enhancement of subscriber Quality of Experience (QoE). QoE enhancement in 5G MCNs is based on a combination of factors including 10x more throughput, less than 1 millisecond latency, and 10x more battery life than 4th Generation MCNs [2]. To meet the capacity enhancement component of these requirements, several solutions have been proposed, with network densification [3] and millimeter wave (mmWave) spectrum utilization [4] among the most popular. It has been demonstrated that a combination of network densification and mmWave spectrum deployment could potentially yield exponential increase in area spectral efficiency [5], thus resolving the capacity crunch 5G MCNs is bound to face.

However, network densification and mmWave spectrum utilization come with their own limitations. Very low user-to-cell ratio can lead to cells being sparsely populated, which makes the performance of ultra-dense heterogeneous networks (UDHNs) exceedingly difficult to estimate [6]. On the other hand, mmWave cells are subject to very high pathlosses due to their operation in 30GHz - 300GHz band. One solution to reduce the impact of high pathloss in mmWave cells is the deployment of highly directional antennas with beam-widths as low as 7° [7]. However, this opens mmWave cell networks to the problem of very large coverage gaps leading to a decrease in network reliability that must be compensated by additional antennas per cell compared to traditional macro cells or by using macro cell overlay [4] adding to network architecture complexity.

The challenges above highlight the difficulties of ensuring reliable and omnipresent coverage in mmWave-UDHNs. In state-of-the-art MCNs, identifying and resolving cellular network outages, which are a consequence of software or hardware failure of network entities, requires highly trained engineers parsing gigabytes of network health logs and network performance indicator data to look for these outages. Due to the continuous growth in cell density and increasing pressure to reduce operational costs, this approach is quickly becoming impracticable.

A. Related Work

Given the significance of the problem in maintaining network service reliability, cell outage detection and prediction has been studied extensively in the last few years. Only a small subset of studies is discussed here that is representative of the larger discussion on state-of-the-art in outage detection research. For a more comprehensive review of outage detection along with outage diagnosis and compensation techniques, the reader is referred to a recent review paper [8].

In [9] the authors employ local outlier factor (LOF) and one-class support vector machines (SVM) techniques to detect coverage anomalies in macro cell environment from the minimization of drive test (MDT) reports data [10]. The authors use the two clustering techniques to separate cells into normal and anomalous based on their received power measurements, and use expert analysis to determine the accuracy of anomaly detection. Using the Receiver Operating Characteristic curves of the two techniques, the authors demonstrate that one-class SVM outperforms LOF considerably. In [11] the authors use handover and radio link failure data to predict network outages in the network. The authors create a diffusion map of changes in user associations due to handover and call failure events and use k-means clustering algorithm to detect the cells with abnormal changes in user associations.

The solutions in [9], [11] use spatial data to detect network outages. However, this limits the observation of outage impact to spatial domain only. In contrast, studies such as [12] propose to use temporal cell level performance metric data such as uplink and downlink throughputs, radio link failures, and handover failures to detect network outages. In [12] the authors use SVM and auto-regressive integrated moving average to identify network anomalies. The authors construct healthy network performance models using the two techniques and predict future cell performance data from those models. If there is a significant deviation between actual and predicted data, the algorithm determines that the cell is in outage.

A key insight from the studies discussed above is that the availability of sufficient data for model training and outage prediction is not a major concern in homogeneous macro cell networks. However, the same does not hold true for UDHNs, especially with very high cell density as it results in very small number of users per cell (< 2 UE/cell) [6]. To address the training data sparsity challenge, the authors of [13] propose to generate cell coverage maps from sparse network coverage data by employing Grey prediction model [14] to interpolate coverage data between randomly distributed locations of users sending MDT reports. The authors use this data to detect anomalies using the predicted user association information and actual user association information.

Despite the advancements in outage detection described above, there exist several key issues that need to be addressed for such solutions to be applicable to mmWave UDHNs:

- Inclusion of Spatio-Temporal Domains: The studies discussed above either use spatial coverage snapshots taken at certain time instants or temporal data of a fixed-location cell for outage detection, thus disregarding the impact of the outage on subscriber QoE in the other domain.
- Sensitivity to shadowing: Most of the existing cell outage detection solutions are highly sensitive to shadowing. This is a key observation in [9] where the impact of shadowing on the accuracy of different outage detection algorithms is investigated. It is shown that as the standard deviation of shadowing increases, machine learning based-outage detection models become less accurate.
- Outage Detection in mmWave-UDHNs: We know that the likelihood of outages increases with cell density as well as complexity of the cell hardware [8] which will be the case in mmWave-UDHNs. However, to the best of our knowledge, an outage detection solution that explicitly targets mmWave-UDHNs while addressing idiosyncrasies of such network does not exist.

B. Proposed Approach and Contributions

In this paper we propose a novel entropy field decomposition (EFD) based solution that can detect cell outages in both space and time, and addresses the other limitations of state-of-the-art solutions discussed above. EFD was first introduced in [1] and has previously been used successfully for brain activity mode detection [15] as well as weather pattern prediction [16], but never in the context of wireless communication. The rationale behind leveraging EFD to solve cell outage detection problem is that it identifies the flow of information in data over both space and time by combining information field theory [17] and entropy spectrum pathways theory [18]. Furthermore, EFD is independent of baseline data model/distribution while also suppressing the effects of noise in activity mode detection process which makes it a natural solution for outage detection in mmWave-UDHN environments marked by heavy shadowing.

The key contributions of this paper can be summarized as follows:

- We present a coverage hole and outage detection solution that is independent of the signal propagation model. This means that the solution can seamlessly be integrated into practical network planning and self-healing solutions without the need for modifying underlying coverage assumptions.
- The proposed solution is designed to identify and minimize the impact of shadowing, fading and noise on outage detection. This allows the solution to detect lapses in coverage even in the event of heavy signal dispersion and shadowing. We demonstrate this by comparing the accuracy of the solution in presence of different shadowing levels.
- The proposed solution incorporates data from both spatial and temporal domains, thus generating coverage estimations that are more accurate in the long term. This spatio-temporal characterization of the outages is a unique feature of this solution that, to the best of the authors' knowledge, has not been achieved so far.
- We analyze the performance of proposed solution for different network topologies including mmWave cell, and mmWave cell-small cell ultra-dense heterogeneous networks. We also compare the results with k-means clustering based coverage hole and outage detection solutions to demonstrate its ability to avoid the issues faced by machine learning-based techniques.

The rest of the paper is organized as follows: Section II describes the system model, Section III presents the proposed EFD solution, Section IV presents the comparative performance analysis. Finally, Section V presents the conclusions of this study.

II. System Model

For the purpose of this study, we declare that a user is in outage when the downlink received power of that user from its associated cell $P^c_{r,u}$ falls below a threshold P^{th}_r i.e., :

$$\text{Outage} := P^c_{r,u} \leq P^{th}_r \tag{1}$$

The log of downlink received power using the exponential pathloss model is expressed as:

$$P^c_{r,u_{dBm}} = f(P^c_t, G_u, G^c_u, b, d^c_u, \beta) + \epsilon^c_u \tag{2}$$

where P^c_t is the transmit power of cell c, G_u is the gain of user equipment, G^c_u is the transmitter antenna gain of cell c,

b is the pathloss constant and depends on the clutter, ϵ_u^c is the shadowing at the location of user u from cell c and usually assumed to be log-normally distributed, d_u^c is the distance of subscriber u from cell c, and β is the pathloss exponent. Assuming each of G_u, G_u^c, b, d_u^c and β remains constant, we can simply re-write (2) as:

$$P^c_{r,u_{dBm}} = f(P^c_t) + \epsilon^c_u \tag{3}$$

Thus, each subset $\hat{\boldsymbol{P}}_{\boldsymbol{t}}$ of the set of cell transmit powers $\boldsymbol{P_t}$ will result in different received powers at the same point. As such, we can define the likelihood of receiving a set of downlink received powers $\boldsymbol{P_r}$ given some set of transmit powers $\hat{\boldsymbol{P}}_{\boldsymbol{t}}$ as:

$$p(\boldsymbol{P_r}) = \int p(\boldsymbol{P_r}|\hat{\boldsymbol{P}}_{\boldsymbol{t}})p(\hat{\boldsymbol{P}}_{\boldsymbol{t}})d\hat{\boldsymbol{P}}_{\boldsymbol{t}} \tag{4}$$

In the event of a cell outage, the loss of transmission from the affected cell will result in a unique set of downlink received powers. Given this set of received powers, we can find the set of transmit powers including the affected cell transmit power using Baye's rule as:

$$p(\hat{\boldsymbol{P}}_{\boldsymbol{t}}|\boldsymbol{P_r}) = \frac{p(\boldsymbol{P_r}, \hat{\boldsymbol{P}}_{\boldsymbol{t}})}{p(\boldsymbol{P_r})} \tag{5}$$

III. Entropy Field Decomposition

For a deterministic system with a fixed signal propagation model and no shadowing, the estimation of (5) is simply a question of going through all the subsets $\hat{\boldsymbol{P}}_{\boldsymbol{t}}$ and calculating the resulting sets $\boldsymbol{P_r}$. However, for a system with random variations in the signal, this estimation becomes more complex. Furthermore, if these random variations affect the system both spatially and temporally, as is the case in real mobile cellular networks, obtaining the conditional probabilities in (5) becomes intractable. However, one method of obtaining an estimate of these probabilities which has been explored in [17] is to use information field theory which represents the probability distributions in terms of an information field. In our case, we must first represent the transmit power or signal data as an information field such that:

$$P_t(x_l, y_l, t_l) \equiv P_t(\zeta_l) = \int \boldsymbol{\vartheta}\delta(\zeta - \zeta_l)d\vartheta \tag{6}$$

where $\zeta_l = x_l, y_l, t_l$ represents the transformation of spatial coordinates $x_i, y_i, i = 1, ..., NM$ and temporal coordinate $t_j, j = 1, ..., O$ as a point on the information field $\boldsymbol{\vartheta}$. The key descriptor of an information field is the Hamiltonian $\mathcal{H}$ which corresponds to the total energy of the field [19] and is defined as:

$$\mathcal{H}(\boldsymbol{P_r}, \boldsymbol{\vartheta}) = -\ln p(\boldsymbol{P_r}, \boldsymbol{\vartheta}) \tag{7}$$

Using the above transformations, (5) can be re-written as:

$$p(\boldsymbol{\vartheta}|\boldsymbol{P_r}) = \frac{e^{\mathcal{H}(\boldsymbol{P_r}, \boldsymbol{\vartheta})}}{\mathcal{Z}(\boldsymbol{P_r})} \tag{8}$$

where $\mathcal{Z}(\boldsymbol{P_r}) = \int e^{\mathcal{H}(\boldsymbol{P_r}, \boldsymbol{\vartheta})}d\boldsymbol{\vartheta}$ is called the partition function. Since the spatio-temporal received powers in a real network are not independent of each other, we consider $P_t(\zeta_l)$ as an interacting field [17] whose Hamiltonian can be derived through Taylor series expansion of (7) as given in [1]:

$$\mathcal{H}(\boldsymbol{P_r}, \boldsymbol{\vartheta}) = \mathcal{H}_0 + \frac{1}{2}\boldsymbol{\vartheta}^\dagger\boldsymbol{\mathcal{D}}^{-1}\boldsymbol{\vartheta} - \boldsymbol{j}^\dagger\boldsymbol{\vartheta} + \sum_{n=1}^{\infty}\frac{1}{n!}\int ... \int V^{(n)}_{\zeta_1...\zeta_n}\vartheta(\zeta_1)...\vartheta(\zeta_n)d\zeta_1...d\zeta_n \tag{9}$$

where $\boldsymbol{\mathcal{D}}$ matrix is the information propagator, the vector $\boldsymbol{j}$ is the information source, the $(.)^\dagger$ notation represents the adjoint of a matrix, and $\mathcal{H}_0$ is the free energy Hamiltonian [19] which can be obtained by integrating the joint probability $p(\boldsymbol{P_r}, \boldsymbol{\vartheta})$ over $\boldsymbol{P_r}$ and $\boldsymbol{\vartheta}$. Since $\mathcal{H}_0$ is a consequence of an interaction-less field, it simply acts as a scaling factor for an interacting field. Also, since we assume that the received power at each point is not independent of other points due to shadowing and fading effects, we can safely ignore $\mathcal{H}_0$ for this study. The terms $V^{(n)}_{\zeta_1...\zeta_n}$ represent the interactions of up to n field components and are integrated over each coordinate. The matrix $\boldsymbol{\mathcal{D}}$ and vector $\boldsymbol{j}$ can be obtained by using the free theory formalism for a Gaussian signal [17] and are given as:

$$\boldsymbol{\mathcal{D}} = \left[{\boldsymbol{\sigma}^{\boldsymbol{2}}_{\hat{\boldsymbol{P}}_{\boldsymbol{t}}}}^{-1} + f(\hat{\boldsymbol{P}}_{\boldsymbol{t}})^\dagger\boldsymbol{\mathcal{N}}^{-1}f(\hat{\boldsymbol{P}}_{\boldsymbol{t}})\right]^{-1} \tag{10a}$$

$$\boldsymbol{j} = f(\hat{\boldsymbol{P}}_{\boldsymbol{t}})^\dagger\boldsymbol{\mathcal{N}}^{-1}\boldsymbol{P_r} \tag{10b}$$

where $\boldsymbol{\sigma}^{\boldsymbol{2}}_{\hat{\boldsymbol{P}}_{\boldsymbol{t}}} = \langle\hat{\boldsymbol{P}}_{\boldsymbol{t}}{\hat{\boldsymbol{P}}_{\boldsymbol{t}}}^\dagger\rangle$ is the covariance of the transmit powers and $\boldsymbol{\mathcal{N}}$ is the covariance of noise in the data.

The interaction terms $V^{(n)}_{\zeta_1...\zeta_n}$ can be obtained using entropy spectrum pathways theory which ranks the optimal pathways within a disordered lattice according to their path entropy [18]. To construct the entropy pathways, we must obtain a coupling matrix $\boldsymbol{\mathcal{Q}}$ of points on the information field lattice. This can be done by generating an adjacency matrix $\boldsymbol{\mathcal{A}_{ij}}$ of spatio-temporal points in the dataset $\boldsymbol{P_r}$ and using the transformation $\zeta_l = x_l, y_l, t_l$ to obtain the components of $\boldsymbol{\mathcal{Q}}$ matrix as follows:

$$\mathcal{Q}(\zeta_i, \zeta_j) = P_r(i)P_r(j)\mathcal{A}_{ij} \tag{11}$$

It is important to highlight here that the $\boldsymbol{\mathcal{Q}}$ matrix can be used to represent any relationship between two or more points in the network regardless of the signal propagation model and the distributions of shadowing and fading. This is a major advantage compared to other techniques such as Bayesian classification that rely on some underlying assumptions regarding data and noise distributions which can lead to very high misclassification error if the actual distribution differs from the assumed one.

The information field can be reconstructed via entropy spectrum pathways that allow the representation of the field in terms of the eigenmodes of $\boldsymbol{\mathcal{Q}}$ using Fourier expansion. In mathematical terms, field components are given as:

$$\vartheta(\zeta_l) = \sum_k^K [a_k\varphi^{(k)}\zeta_l + a_k^*\varphi^{*(k)}\zeta_l] \tag{12}$$

where where $\varphi^{(k)}$ is the k^{th} eigenmode, a_k is the mode amplitude of k^{th} eigenmode and the $*$ operator refers to the conjugate of a number, while K is the number of significant eigenmodes considered for field transformation.

A key insight here is that by only considering the most important eigenmodes and keeping K to a reasonably small value compared to the total number of eigenmodes, we can obtain a decent estimate of the information field, thus reducing the problem complexity significantly. To test the importance of an eigenmode, we can compare the corresponding eigenvalue λ_k with the determinant of the noise covariance matrix $\boldsymbol{\mathcal{N}}$.

Using the above information, we can now obtain the transformed information Hamiltonian $\mathcal{H}(\boldsymbol{P_r}, \boldsymbol{a_k})$ by:

$$\mathcal{H}(\boldsymbol{P_r}, \boldsymbol{a_k}) = \frac{1}{2}\boldsymbol{a_k}^\dagger \boldsymbol{\Lambda} \boldsymbol{a_k} - \boldsymbol{j_k}^\dagger \boldsymbol{a_k} + \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k_1}^{K} \dots \sum_{k_n}^{K} \tilde{V}^{(n)}_{k_1 \dots k_n} a_{k_1} \dots a_{k_n} \tag{13}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix containing the eigenvalues of $\boldsymbol{\mathcal{Q}}$, $\tilde{V}^{(n)}_{k_1 \dots k_n}$ represent the interaction terms of the eigenmodes, and $\boldsymbol{j_k}$ is the amplitude of the k^{th} eigenmode in expansion of the information source $\boldsymbol{j}$:

$$\boldsymbol{j_k} = \int \boldsymbol{j} \varphi^{(k)} d\zeta \tag{14}$$

To calculate the values of mode amplitudes, we solve $\partial \mathcal{H} / \partial \vartheta = 0$ and replace the field with its transformation in terms of its eigenmodes which gives us:

$$\boldsymbol{\Lambda} \boldsymbol{a_k} = \boldsymbol{j_k} - \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{k_1}^{K} \dots \sum_{k_n}^{K} \tilde{V}^{(n)}_{k k_1 \dots k_n} a_{k_1} \dots a_{k_n} \tag{15}$$

If the field interaction terms $V^{(n)}_{\zeta_1 \dots \zeta_n}$ are defined as powers of the coupling matrix $\boldsymbol{\mathcal{Q}}$ such that:

$$V^{(n)}_{\zeta_1 \dots \zeta_n} = \frac{\alpha^{(n)}}{n} \sum_{p=1}^{n} \prod_{\substack{m=1 \\ m \neq n}}^{n} \mathcal{Q}_{\zeta_p} \mathcal{Q}_{\zeta_m} \tag{16}$$

then the mode interaction terms $\tilde{V}^{(n)}_{k_1 \dots k_n}$ are obtained by:

$$\tilde{V}^{(n)}_{k_1 \dots k_n} = \frac{\alpha^{(n)}}{n} \sum_{p=1}^{n} \left(\frac{1}{\lambda_{k_p}} \prod_{m=1}^{n} \lambda_{k_m} \right) \int \left(\prod_{r=1}^{n} \varphi^{(k_r)}(\zeta) \right) d\zeta \tag{17}$$

where coefficients $\alpha^{(n)}$ should be chosen sufficiently small to ensure the convergence of (17).

A. Outage Detection Using Entropy Field Decomposition

The result of applying EFD to the coverage data is an entropy field which identifies how information flows across the spatio-temporal domains. In order to make this resulting field output useful for outage detection in MCNs, algorithm 1 presents our proposed outage detection and localization solution.

The algorithm takes user-cell association $\mathbb{U}_c$ information from the MDT data and the network coverage data when no outage is present in the network $\boldsymbol{P_{r_{norm}}}$, as well as real-time spatio-temporal coverage data $\boldsymbol{P_{r_{RT}}}$. The real-time data is processed using EFD which outputs the data points where high information flows are detected. In simple terms, the EFD algorithm gives the boundary between outage and non-outage effected areas. The points with high energy i.e., the points at the boundary of the outage are passed to a localization module which identifies the cell with which the high energy points are associated. Once the degraded cell is identified, it is passed as an output which can be fed to an outage diagnosis and compensation algorithm.

Algorithm 1 Outage Detection Using EFD

Input: $\boldsymbol{P_t}, \boldsymbol{P_{r_{norm}}}, \boldsymbol{P_{r_{RT}}}, \mathbb{U}_c$
Output: $\vartheta(\zeta_l)$, c in outage

Calculate $\boldsymbol{\mathcal{Q}}$ from (11) using $\boldsymbol{P_{r_{RT}}}$
Obtain eigenvalues and eigenvectors of $\boldsymbol{\mathcal{Q}}$ using (**??**)
Calculate entropy field $\boldsymbol{\vartheta}$ using the eigenvalues and eigenvectors of $\boldsymbol{\mathcal{Q}}$ from (12)
for $l \in 1, \dots, NMO$ **do**
  if $\vartheta_l > 0$ **then**
    $\{u(x_{out}, y_{out})\} = \{u(x_{out}, y_{out})\} + u(x_l, y_l)$
  end if
end for
if $\{u(x_{out}, y_{out})\} = \phi$ **then**
  continue
else
  for $u(x_{out}, y_{out}) \in \{u(x_{out}, y_{out})\}$ **do** $\{c_{out}\} = \{c_{out}\} + \{c = \arg\max_{\forall c \in \mathbb{C}} P^c_{r,u(x_{out},y_{out})_{norm}}\}$
  end for
end if

IV. Simulation and Results

A. Simulation Setup

To test the proposed algorithm, we employed a 3GPP compliant LTE simulator with the pathloss models for small cells specified by the 3GPP, and the pathloss for mmWave directional cells specified in [20]. We only consider a narrow-band mmWave spectrum, therefore frequency-selective fading is not considered. Shadowing is modeled using log-normal distribution with different standard deviations. We use cell transmit power as the input signal for information field generation; however, any other coverage actuation parameter such as transmitter antenna tilts or transmitter antenna gains can be used just as easily if the relationship between the parameter and received signal strength is known. The complete list of simulation parameters are given in Table I.

B. Results

The results presented below compare proposed EFD-based outage detection solution with outage detection using k-means clustering algorithm. Performance of the two solutions is evaluated using two different topologies: 1) mmWave cells only, and 2) UDHN. Outages are simulated by setting the transmit power of the effected cell to 0 dB. In the case of mmWave cells only, one sector of first mmWave base station is in outage, while in the case of heterogeneous network one sector of first mmWave base station and one small cell are in outage. Simulations are carried out for shadowing $\epsilon^c_u \in [0, 10]$ to assess the efficacy of EFD and k-means in mitigating the effects of noise in the data. Note that though the algorithms for

TABLE I: Parameter Settings for Simulation

System Parameters	Value
Transmission Frequency	Small: 2 GHz, mmWave: 38 GHz
Small Cell Transmit Power	Max: 20 dBm, Outage: 0 dBm
mmWave Cell Transmit Power	Max: 41 dBm, Outage: 0 dBm
Antenna Gain	Small: 5dBi, mmWave: 25dBi
Horizontal Beamwidth	Small: 360°, mmWave: 7.8°
Cell Height	Small: 10m, mmWave: 23m
Cell Layout (small cells)	Uniformly distributed
Cell Layout (mmWave cells)	Hexagonal
Transmission Bandwidth	10 MHz
Number of Base Stations	7
Sectors per Base Station	3
Small Cells per Sector	1

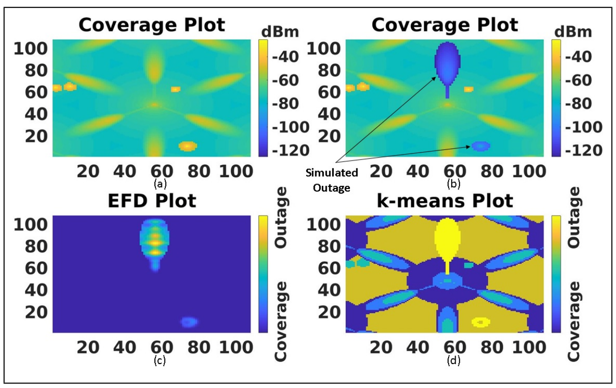

Fig. 1: Small cell and mmWave cell outage simulated in a UDHN with 0dB shadowing.

EFD and k-means clustering are implemented using data from spatio-temporal dimensions, the presented results for coverage and outage detection are averaged over time.

Figs. 1-4 show examples of the simulated outages in a UDHN and mmWave-cell only network under 0dB and 10dB shadowing. For the UDHN cases, one small cell and one mmWave cell is in outage. Subfigures (a) show the network under normal coverage while subfigures (b) show the network in the presence of outages. Subfigures (c) show the results of outage detection using EFD and subfigures (d) show outage detection using k-means clustering.

To truly analyze the impact of EFD and k-means outage detection, we test the two techniques over a range of shadowing values and compare the true positive detection rate (TPR) and false positive detection rate (FPR) of outage effected users. Fig. 5 shows a comparison of the two outage detection schemes for mmWave cell networks and UDHNs under increasing levels of shadowing. We can see that EFD gives highly accurate TPR ($> 96\%$) even under heavy shadowing. This is because of the intrinsic noise countering capabilities of the EFD solution. Conversely, the TPR of k-means clustering is very high for low fading environments ($> 98\%$) but drops off exponentially as the shadowing is increased. This also highlights the limitations of machine learning techniques with respect to the noise in input data, something that EFD effectively counters intrinsically.

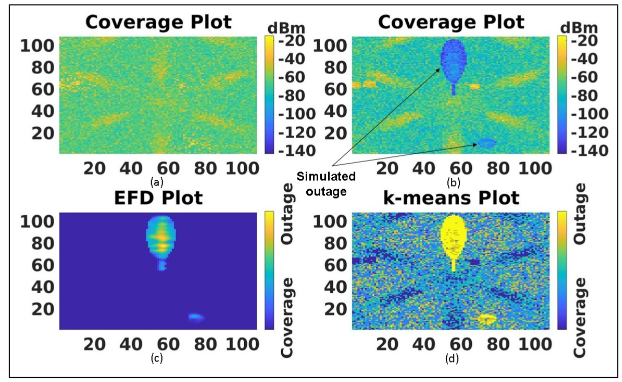

Fig. 2: Small cell and mmWave cell outage simulated in a UDHN with 10dB shadowing.

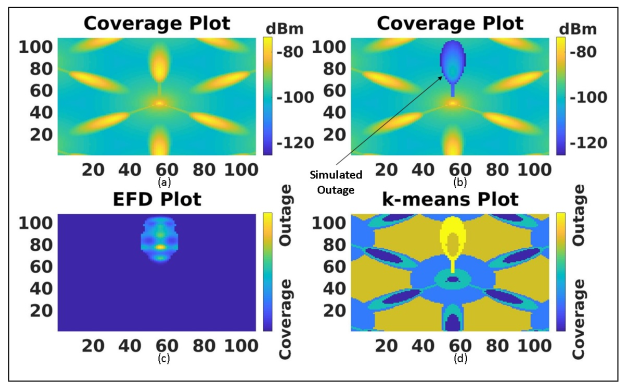

Fig. 3: mmWave cell outage simulated in mmWave-cell only network with 0dB shadowing.

Fig. 6 presents another perspective of evaluating the performance of the two algorithms i.e., by comparing their FPR. This tells us whether the two algorithms actually do perform outage detection or are simply flagging majority of the points as outage affected. From Fig. 6 we can see that the EFD-based outage detection solution has a significantly low FPR for both mmWave cell networks and UDHNs ($< 6\%$). In comparison, k-means clustering has a relatively higher FPR. FPR for EFD is caused due to the detection of maximum entropy boundary which includes the non-outage effected users at the edge of the outage effected areas. On the other hand, the FPR of k-means is primarily due to the effects of shadowing.

V. Conclusion

In this work, we present a novel cell outage detection solution based on entropy field decomposition. The proposed solution is designed to improve network reliability by mitigating the effects of shadowing on outage detection, a feature that is not common in state-of-the-art machine learning based outage detection solutions. The proposed solution is compared with k-means clustering technique for different network topologies and shadowing levels. The results show that EFD-based outage detection performs extremely well even in high shadowing environments compared to k-means clustering making it ideal for outage detection in high density mmWave cell heterogeneous networks.

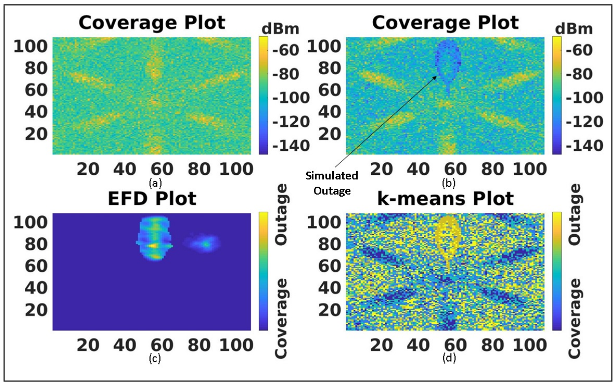

Fig. 4: mmWave cell outage simulated in mmWave-cell only network with 10dB shadowing.

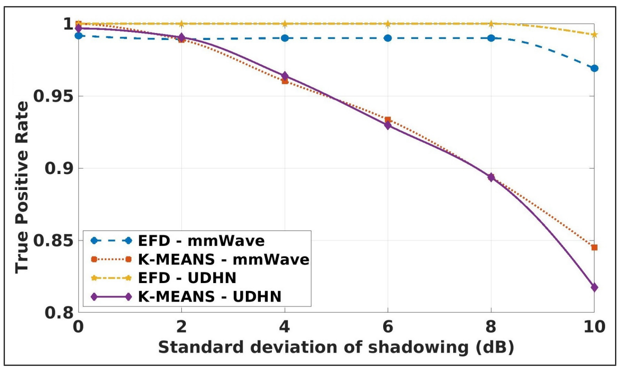

Fig. 5: True positive rate comparison of EFD and k-means clustering outage detection for mmWave cell network and UDHN.

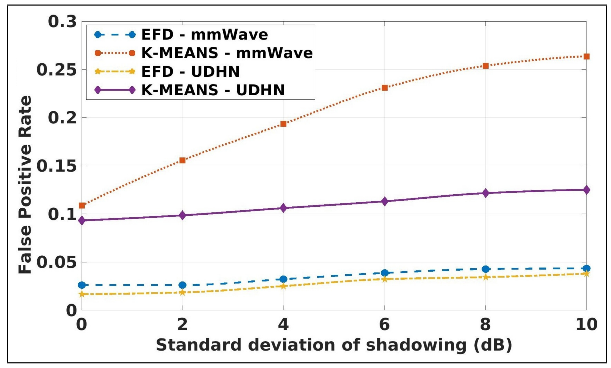

Fig. 6: False positive rate comparison of EFD and k-means clustering outage detection for mmWave cell network and UDHN.

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation under Grant Numbers 1559483, 1619346 and 1730650. For further information about these projects please visit www.bsonlab.com.

REFERENCES

[1] L. R. Frank and V. L. Galinsky, "Detecting spatio-temporal modes in multivariate data by entropy field decomposition," *Journal of Physics A: Mathematical and Theoretical*, vol. 49, no. 39, p. 395001, 2016.

[2] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What Will 5G Be?" *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[3] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Communications*, vol. 18, no. 3, pp. 10–21, June 2011.

[4] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[5] J. G. Andrews, X. Zhang, G. D. Durgin, and A. K. Gupta, "Are we approaching the fundamental limits of wireless network densification?" *IEEE Communications Magazine*, vol. 54, no. 10, pp. 184–190, October 2016.

[6] S. Chernov, M. Pechenizkiy, and T. Ristaniemi, "The influence of dataset size on the performance of cell outage detection approach in LTE-A networks," in *Proc. 10th International Conference on Information, Communications and Signal Processing (ICICS)*, Dec 2015, pp. 1–5.

[7] M. N. Kulkarni, S. Singh, and J. G. Andrews, "Coverage and rate trends in dense urban mmWave cellular networks," in *Proc. IEEE Global Communications Conference (GLOBECOM)*, Dec 2014, pp. 3809–3814.

[8] A. Asghar, H. Farooq, and A. Imran, "Self-healing in emerging cellular networks: Review, challenges and research directions," *IEEE Communications Surveys & Tutorials*, pp. 1–1, 2018.

[9] A. Zoha, A. Saeed, A. Imran, M. A. Imran, and A. Abu-Dayya, "A learning-based approach for autonomous outage detection and coverage optimization," *Transactions on Emerging Telecommunications Technologies*, vol. 27, no. 3, pp. 439–450, 2016.

[10] 3GPP, *TS 37.320 V10.1.0; Evolved Universal Terrestrial Radio Access (E-UTRA); Radio measurement collection for Minimization of Drive Tests (MDT); Overall description; Stage 2*, 3GPP TS, Rev. V8.3.0 Release 8, 2010.

[11] F. Chernogorov, J. Turkka, T. Ristaniemi, and A. Averbuch, "Detection of Sleeping Cells in LTE Networks Using Diffusion Maps," in *Proc. IEEE 73rd Vehicular Technology Conference (VTC Spring)*, May 2011, pp. 1–5.

[12] G. Ciocarlie, U. Lindqvist, K. Nitz, S. Nováczki, and H. Sanneck, "On the feasibility of deploying cell anomaly detection in operational cellular networks," in *Proc. IEEE Network Operations and Management Symposium (NOMS)*, May 2014, pp. 1–6.

[13] O. Onireti, A. Zoha, J. Moysen, A. Imran, L. Giupponi, M. A. Imran, and A. Abu-Dayya, "A Cell Outage Management Framework for Dense Heterogeneous Networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2097–2113, April 2016.

[14] T.-L. Tien, "A new grey prediction model FGM(1, 1)," *Mathematical and Computer Modelling*, vol. 49, no. 7, pp. 1416 – 1426, 2009.

[15] L. R. Frank and V. L. Galinsky, "Dynamic Multiscale Modes of Resting State Brain Activity Detected by Entropy Field Decomposition," *Neural Computation*, vol. 28, no. 9, pp. 1769–1811, 2016, pMID: 27391678.

[16] L. R. Frank, V. L. Galinsky, L. Orf, and J. Wurman, "Dynamic multiscale modes of severe storm structure detected in mobile doppler radar data by entropy field decomposition," *J. Atmos. Sci.*, vol. 75, no. 3, pp. 709–730, Mar. 2018. [Online]. Available: https://doi.org/10.1175/JAS-D-17-0117.1

[17] T. A. Enßlin, M. Frommert, and F. S. Kitaura, "Information field theory for cosmological perturbation reconstruction and nonlinear signal analysis," *Phys. Rev. D*, vol. 80, Nov 2009.

[18] L. R. Frank and V. L. Galinsky, "Information pathways in a disordered lattice," *Phys. Rev. E*, vol. 89, p. 032142, Mar 2014. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.89.032142

[19] W. R. Hamilton, "On a general method in dynamics; by which the study of the motions of all free systems of attracting or repelling points is reduced to the search and differentiation of one central relation, or characteristic function," *Philosophical transactions of the Royal Society of London*, vol. 124, pp. 247–308, 1834.

[20] T. S. Rappaport, G. R. MacCartney, M. K. Samimi, and S. Sun, "Wideband Millimeter-Wave Propagation Measurements and Channel Models for Future Wireless Communication System Design," *IEEE Transactions on Communications*, vol. 63, no. 9, pp. 3029–3056, Sept 2015.