Quality of Coverage: A Novel Approach to Coverage for Mobile Crowd Sensing Systems

Sherif B. Azmy*, Nizar Zorba*, Hossam S. Hassanein†
* College of Engineering, Qatar University, Doha, Qatar
† School of Computing, Queen's University, Kingston, Ontario, Canada
Emails:{sherif.azmy, nizarz}@qu.edu.qa, hossam@cs.queensu.ca

***Abstract*—Mobile Crowd Sensing (MCS) is an effective paradigm that utilizes the crowd as an extended instrument for the purpose of collecting data. However, utilizing the crowd comes with the risks stemming from the crowd's heterogeneity. Thus, the MCS administrator must carefully recruit and evaluate MCS participants for the reliable execution of MCS tasks. In this paper, we tackle some of the criteria required for the proper characterization of an Area of Interest (AoI). We propose a coverage metric aimed at MCS systems that takes into consideration the global view of the AoI as a whole, as well as a local picture with regards to the subdivisions with the AoI. The developed coverage metric allows the MCS administrator to identify which regions within the AoI are lacking, in terms of quality, and how they can be compensated by moving participants from neighboring regions. We demonstrate the performance of the presented metric by means of a computer simulation.**

***Index Terms*—mobile crowdsensing; internet of things; sensor networks; coverage quality metric; source quality; data collection.**

I. Introduction

The Internet of Things (IoT) paradigm has led to the development of various sub-paradigms and frameworks that combine the cyber and physical worlds. The integration of sensors in all aspects of life has outlined complex - yet very functional - infrastructures in the premise of smart cities. As a result, the quest for the efficiency of the smart city's operation is becoming more significant as smart city administrators and stakeholders seek increased access to data. One paradigm that caters to these demands is Mobile Crowd Sensing (MCS), which capitalizes on the ubiquity of sensors in *the crowd* of sensor-laden smartphone users. MCS has managed to transform cohorts of smartphone users into an extended instrument [1], with a wide-range of applications within the context of the smart city and its social and physical aspects. This versatility renders MCS as a very useful tool for smart city stakeholders [2].

In MCS, tasks are created by administrators to be assigned by the server to a set of participants who are concerned with its execution. The nature of sensing and task execution can be classified as *participatory* or *opportunistic* sensing [2]. Participatory sensing requires active involvement of the user in order to successfully execute a task. For example, a user can take out the microphone to record a sample of noise pollution [3], or take a photo [4]. On the other hand, opportunistic sensing is executed automatically when the participant's smartphone satisfies the set of conditions for the task, (such as time, place, device capabilities, battery,... etc.) [5]. Various frameworks and models were developed to facilitate the management of MCS and reduce its complexity. One of which is developed in [6] to model the MCS sensing problem over *spatiotemporal cells*, i.e. a cell bounded in space and time.

A large amount of research literature [4], [7] treats MCS as a problem with large scale data, harboring the assumption that data is always available. On the other hand, the authors of [8], [9] argue that data might be scarce, focusing on small data. As the source of MCS data comes from a wide variety of participants, it is prone to significant variations in quality, performance, and capability. To overcome this heterogeneity in data, the authors of [8], [9] have developed small sample quality metrics to quantify and evaluate the reliability of data collected by MCS participants within a spatiotemporal cell. Such a metric allows the elimination of faulty and inconsistent sensors, and provides a *local* picture of the participants' reliability. The quantifiable knowledge of participants' reliability allows the MCS administrators - and the system - to select enough participants for reliable execution of MCS tasks, reducing MCS costs, which relate to the number of participants either in terms of monetary incentives or data consumed.

While having a local picture of a specific spatiotemporal cell is useful for the MCS administrator, the relation to other cells within the region is still absent. Furthermore, some cells have a higher sensing priority than others. This necessitates the versatile recruitment of MCS participants in a manner that caters to the *coverage* needs of the MCS administrators. In this paper, we provide a *global picture* to complement the *local picture*, utilizing the global and local contrast to provide the MCS administrators with as much information as possible about the MCS region. To that end, we develop a coverage metric that extends the metrics developed in [8], [9] to evaluate the quality of the area as a whole during a specific sensing cycle. We have also defined a *relative quality metric* for adjacent cells, that allows the MCS administrator to request participants to move between cells (or to recruit alternative participants) to ensure consistent coverage over the MCS region. Such a metric will enhance the recruitment of participants in an MCS system, increasing its efficiency.

This paper is structured as follows: Section 2 provides an overview of MCS, Quality in MCS, and the Spatiotemporal model; Section 3 addresses the development and description of

978-1-5386-7272-3/18/$31.00 ©2018 IEEE

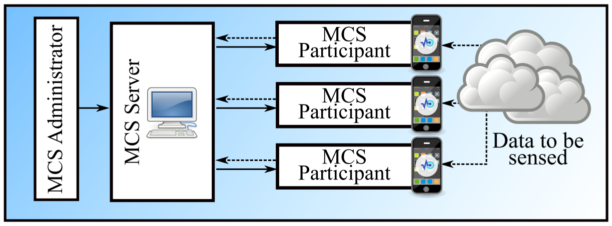

Fig. 1: General MCS Architecture

the proposed coverage quality metric; Section 4 illustrates the usage of the coverage quality metric with relevant examples; and finally section 5 concludes with foreword on coverage in the context of MCS.

II. MCS Overview

A. MCS System Design

Various research efforts have led to the development of many sub-topics under MCS, which have led to the development of a general perception of the architecture of an MCS system [9]. Generally, an MCS system would consist of the following elements:

- **MCS Administrator:** who designs and publishes tasks in an MCS system. Such an entity is closely related to the management of smart cities and aims to cater to the purposes of the smart city for its efficient operation. An MCS administrator may be interested in data collection about a specific phenomenon at a specific location and time.
- **MCS Participants:** are members of the crowd who participate in an MCS system to execute MCS tasks set by the administrator, in exchange for a service or a monetary reward (incentives). They can execute tasks participatorily, being actively involved in the task execution, or opportunistically, by passively being present at the time and place at which the task is automatically executed.
- **MCS System:** which is an automated service that connects the participants to the administrator, recruiting and assigning tasks to them in an automatic manner while benchmarking and evaluating their performance, reliability, and quality.

Fig. 1 illustrates the general architecture of MCS systems in [9].

B. MCS Spatio-Temporal Model

In order to acquire an informative characterization of quality within an Area of Interest (AoI) in an MCS system, the AoI needs to be appropriately divided into geofences or *cells* to which participants are assigned. However, these divisions are not only limited to space, as time - in terms of sensing cycles - is also a dimension upon which such divisions are required. These divisions in space and time can be represented by means of the *spatiotemporal diagram* shown in Fig. 2, similar to that developed and described in [6]. MCS administrators should divide the space in a manner that satisfies the purpose of the MCS system while respecting the limitations of the

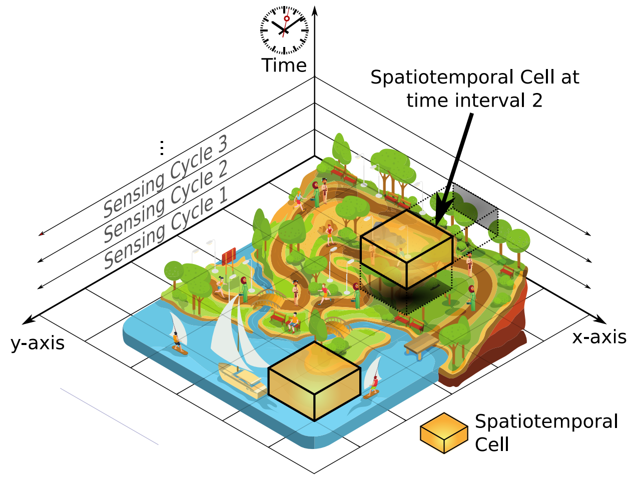

Fig. 2: Spatiotemporal Diagram

Sampling Theorem, which can be extended from the temporal domain into the spatial domain [10], [11]. Tasks are assigned to participants according to their availability, which includes whether they are eligible for performing the task or not, as well as being at the location and the time described in the MCS task description. For example, MCS administrators could be interested in acquiring temperature readings by MCS participants in the m^{th} cell, which the participants will return as a set of readings X_m:

$$X_m = \{x_{m,1}, x_{m,2}, \ldots, x_{m,N_m}\} \tag{1}$$

where the reading of the n^{th} participant is represented as a random variable $x_{m,n}$, and N_m is the number of participants who have successfully executed the task in the m^{th} spatiotemporal cell. Each cell corresponds to a 3-tuple (a, b, c), where a maps to the x-location, b maps to the y-location, and c maps to the c^{th} sensing cycle (temporal location).

The truth of the sensed quantity, μ, is then estimated from the sample obtained in Eq. (1) by computing the mean:

$$\hat{\mu}_m = \text{mean}(X) = \frac{1}{N_m}\sum_{n=1}^{N_m} x_{m,n} \tag{2}$$

while its standard deviation can be estimated by means of the sample error:

$$\hat{\sigma}_m = \sqrt{\frac{\sum_{n=1}^{N_m}(x_{m,n} - \hat{\mu}_m)}{N_m - 1}} \tag{3}$$

III. Coverage Quality Metric for MCS

While MCS can operate under incomplete coverage conditions, it is of significant importance that MCS administrators would be able to identify which spatio-temporal cells are lacking coverage. In this section, we propose a simple scheme that aims to characterize the coverage quality of the AoI as a whole, during a sensing cycle. We also propose a metric to measure quality with respect to adjacent cells. This allows MCS administrators to compensate poor cells from their neighbouring better-off cells. The method developed in this section combines two quantities, the overall coverage quality and the relative quality, to provide the MCS administrator with a detailed map of *which cells have higher quality than others* so that the MCS system would recruit users from cells with satisfied quality criteria to those lacking. Coverage, in that sense, relates to the quality over the whole space, while uniformity implies having a good quality over all the cells.

A. Cell-Specific Quality Metrics

However, before the definition of coverage quality, we briefly describe the definition of quality within a specific cell in this sub-section. In [8], [9] quality metrics were developed and defined in a manner that considers outliers to be threats to quality. The lack of knowledge about the ground truth is what drives the paradigm of MCS to employ MCS participants as proxies in determining the truth. However, the inherent heterogeneity in the crowd of participants is prone to errors and faults, which can be overlooked if the participants are present in an abundance. Nevertheless, this is not always the case it is possible for MCS regions to be underpopulated with participants, or that the number of reliable data points is scarce. In such a case, an outlier resulting from a fault or an abnormality can easily throw off the estimation of the ground-truth. The approach embodied in [8], [9] defines quality as a quantity that relates to the differences between the values acquired within a cell. To that end, robust centrality estimates such as the Median Absolute Deviation-filtered Mean and the Trimmed mean were employed to develop the Mean MAD-Mean Trimmed Mean (MMTM) statistic, θ_{MMTM}. θ_{MMTM} is employed to evaluate small-sample quality in [8], and very small-sample quality in [9], where θ_{MMTM} was defined as:

$$\theta_{\text{MMTM}} = \beta(\bar{x} - \bar{x}_k) + (1-\beta)(\bar{x} - \bar{x}_{\text{MAD}}) \tag{4}$$

with $\beta \in [0,1]$, $\bar{x}$ is the standard mean of the readings within a cell x, $\bar{x}_k$ is the trimmed-mean, and $\bar{x}_{\text{MAD}}$ is the MAD-mean.

Quality can then be described in two manners: it can be described loosely as a measure *how good of a representative the sample is in estimating the ground-truth*, or described precisely as a quantity that reflects the acquired samples' reliability in estimating the ground-truth. Quality, as per [8], [9], was defined for small-sample ranges as:

$$Q = \begin{cases} \log_\gamma\left(\frac{1}{\theta_{\text{MMTM}}}\right) & 11 < N < 30 \\ \frac{1}{2}\log_\gamma\left(\mu_{\text{MMTM}}^{-1}\sigma_{\text{MMTM}}^{-2}\right) & 8 < N < 11 \end{cases} \tag{5}$$

where γ is a scaling factor, μ_{MMTM} and σ_{MMTM} are the mean and variance of the bootstrap distribution of θ_{MMTM}.

B. Overall Coverage Quality

In this subsection, we upgrade the quality metric previously proposed, to tackle the spatial dimension of coverage. For simplicity, the spatial space is assumed to be a $C \times C$ square grid, with each cell having a quality value $Q_{c,c}$, during a specific sensing cycle T. This allows the definition of an overall quality metric, Q_{tot}:

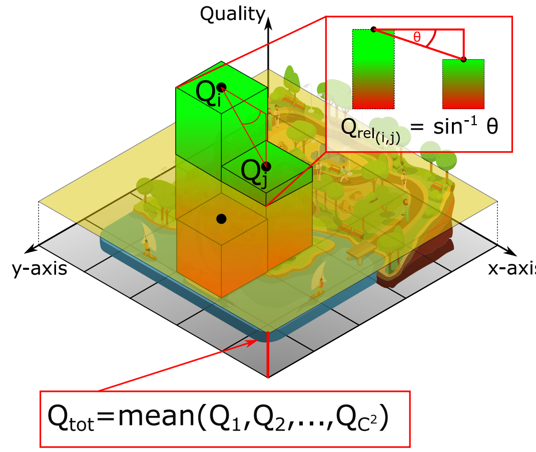

Fig. 3: Coverage Metric Description

$$Q_{\text{tot}} = \text{mean}(\mathbf{Q}_{\text{map}}) = \text{mean}\begin{bmatrix} Q_{1,1} & Q_{1,2} & \dots & Q_{1,C} \\ Q_{2,1} & Q_{2,2} & \dots & Q_{2,C} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{C,1} & Q_{C,2} & \dots & Q_{C,C} \end{bmatrix} \tag{6}$$

where $\mathbf{Q}_{\text{map}}$ is the resulting matrix of all quality evaluations for all cells.

The resulting Q_{tot} can be seen in Fig. 3 where it corresponds to a plane elevated to its value. It is important that an MCS administrator would set a minimum threshold for quality, Q_{min}, which would be useful in two manners: (a) to evaluate whether the overall quality is satisfied, i.e. $Q_{\text{tot}} \geq Q_{\text{min}}$; and (b) to identify specific cells which are below the threshold. Combining these two outcomes together allows the characterization of overall coverage quality, however it does not indicate *how* it can be adjusted to achieve uniformity.

C. Relative Coverage Quality

In order to *how* the quality can be improved over a space, we developed a method that measures the *angle* between each two adjacent quality points and constructs a $C^2 \times C^2$ matrix that maps this relation, denoted $\mathbf{Q}_{\text{rel}}$ whose elements are defined as:

$$Q_{\text{rel},(Q_{i,j},Q_{a,b})} = \sin^{-1}\left[\frac{Q_{i,j} - Q_{a,b}}{\sqrt{[Q_{i,j} - Q_{a,b}]^2 + 1}}\right] \tag{7}$$

where (i,j) corresponds to the coordinates of the of k^{th} cell, and (a,b) correspond to the coordinates of an adjacent cell. Fig. 3 provides an illustration of Eq. 7.

However, to construct the resulting $C^2 \times C^2$ $\mathbf{Q}_{\text{rel}}$ matrix using an automated algorithm, the *2-tuple map*, $\mathbf{Q}_{\text{map}}$ has to be mapped into a *singleton map*, $\hat{\mathbf{Q}}_{\text{map}}$ such that:

$$\underbrace{\begin{bmatrix} Q_{1,1} & Q_{1,2} & \dots & Q_{1,C} \\ Q_{2,1} & Q_{2,2} & \dots & Q_{2,C} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{C,1} & Q_{C,2} & \dots & Q_{C,C} \end{bmatrix}}_{\mathbf{Q}_{\text{map}}} \mapsto \underbrace{\begin{bmatrix} Q_1 & Q_2 & \dots & Q_C \\ Q_{C+1} & Q_{C+2} & \dots & Q_{2C} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{\Delta+1} & Q_{\Delta+2} & \dots & Q_{C^2} \end{bmatrix}}_{\hat{\mathbf{Q}}_{\text{map}}} \tag{8}$$

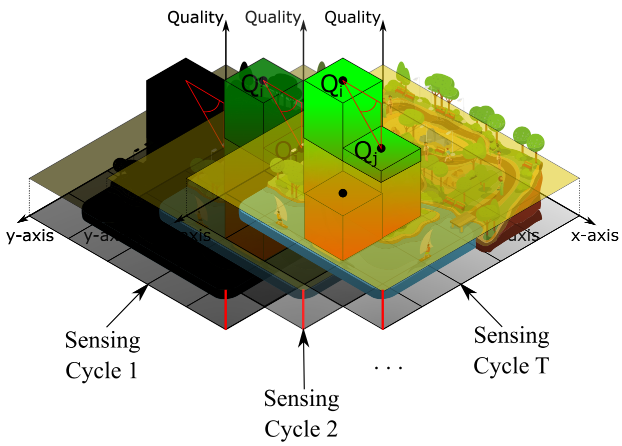

Fig. 4: Descriptive illustration of the change of MCS coverage over time

where $\Delta = C(C-1)$.

Within the singleton map, cells adjacent to the k^{th} cell are corresponding to the north $(k-C)$, west $(k-1)$, east $(k+1)$, and south $(k+C)$, for any cell that is not on the edge nor in the corner. $\mathbf{Q}_{\text{rel}}$ can then be constructed by substituting Eq. 7 for the corresponding elements such that Eq. 7 becomes:

$$Q_{\text{rel}}(Q_k, Q_{k'}) = \sin^{-1}\left[\frac{Q_k - Q_{k'}}{\sqrt{[Q_k - Q_{k'}]^2 + 1}}\right] \tag{9}$$

where k corresponds to the row, and k' corresponds to the column.

The Relative Coverage Quality metric looks at the quality, described by the $\mathbf{Q}_{\text{map}}$ as a surface on which spatial cells correspond to points, and measures the angle between these points. As a result, the synthesized matrix $\mathbf{Q}_{\text{rel}}$ is an anti-symmetric matrix in which transposing elements are a pair of alternate interior angles, thus the sign chance. As a result, $\mathbf{Q}_{\text{rel}}$ is a sparse matrix which can be visualized using color-coded tables, an example is shown in the simulations section, Fig. 5. If the corresponding value for $(k \to k')$ is positive, then the *uniformity* of coverage can be improved by moving participants from cell k to k' participatorily by means of an incentive. On the other hand, a negative value implies that participants could be moved from k' to k. This can also reflect in manipulating the incentives in a manner that would affect the supply and demand of participants in the k^{th} cell, to achieve the MCS system's coverage objectives.

This coverage quality metric, however, changes with time, as every sensing cycle, T^{th}, has a corresponding Q_{tot} value and a corresponding $\mathbf{Q}_{\text{rel}}$ matrix. Fig. 4 provides an illustration of how such coverage change over time.

Algorithm 1 details the steps required to generate the $\mathbf{Q}_{\text{rel}}$ matrix. There exists a number of special cases relating to elements on the edges of the matrix with no preceding or following elements in one or more of the four cardinal directions: North $(i-C)$, East $(i-1)$, West $(i+1)$, and South $(i+C)$. Algorithm 1 embodies a simple decision tree to deal with any of the eight special cases: the four corners $(i = 1, i = C, i = (C-1)C + 1$, and $i = C^2)$; and the four edges: North $(i < C)$, East $(i\%C = 0)$, West $(i\%C = 1)$, and South $(i < C^2$ and $i > (C-1)C + 1)$.

Algorithm 1 Algorithm for Evaluating MCS Coverage Quality

Input: $\mathbf{Q}_{\text{map}}$
Output: Q_{tot} and $\mathbf{Q}_{\text{rel}}$

```
Q_tot ← mean(Q_map(:))
Initialize Q_rel=zeros(length(Q_map))
for i ← 1 : C^2 do
    if i = 1 then
        Q_rel ← Q_rel(i, i+1)
        Q_rel ← Q_rel(i, i+C)
    else if i < C then
        Q_rel ← Q_rel(i, i-1)
        Q_rel ← Q_rel(i, i+1)
        Q_rel ← Q_rel(i, i+C)
    else if i = C then
        Q_rel ← Q_rel(i, i-1)
        Q_rel ← Q_rel(i, i+C)
    else if i%C = 1 then
        if i = (C-1)C+1 then
            Q_rel ← Q_rel(i, i-C)
            Q_rel ← Q_rel(i, i+1)
        else
            Q_rel ← Q_rel(i, i-C)
            Q_rel ← Q_rel(i, i+1)
            Q_rel ← Q_rel(i, i+C)
        end if
    else if i%C = 0 then
        if i = C^2 then
            Q_rel ← Q_rel(i, i-C)
            Q_rel ← Q_rel(i, i-1)
        else
            Q_rel ← Q_rel(i, i-C)
            Q_rel ← Q_rel(i, i-1)
            Q_rel ← Q_rel(i, i+C)
        end if
    else if i < C^2 & i > (C-1)C+1 then
        Q_rel ← Q_rel(i, i-C)
        Q_rel ← Q_rel(i, i-1)
        Q_rel ← Q_rel(i, i+1)
    else
        Q_rel ← Q_rel(i, i-C)
        Q_rel ← Q_rel(i, i-1)
        Q_rel ← Q_rel(i, i+1)
        Q_rel ← Q_rel(i, i+C)
    end if
end for
return Q_tot, Q_rel
```

IV. Computer Simulation

To test the proposed coverage quality metric,we generated a 3×3 map with values obtained from a distribution $N(2.2, 0.5)$

0	64.04	0	15.51	0	0	0	0	0
-64.04	0	-50.14	0	-58.24	0	0	0	0
0	50.14	0	0	0	0.90	0	0	0
-15.51	0	0	0	9.12	0	51.41	0	0
0	58.24	0	-9.12	0	23.42	0	53.85	0
0	0	-0.90	0.00	-23.42	0	0	0	-47.62
0	0	0	-51.41	0	0	0	15.44	0
0	0	0	0	-53.85	0	-15.44	0	-63.79
0	0	0	0	0	47.62	0	63.79	0

Fig. 5: $\mathbf{Q}_{\text{rel}}$ Matrix

to simulate diverse values for the quality metrics described in [8]. The input $\mathbf{Q}_{\text{map}}$ obtained was:

$$\mathbf{Q}_{\text{map}} = \begin{bmatrix} 3.09 & 1.04 & 2.23 \\ 2.81 & 2.65 & 2.22 \\ 1.56 & 1.28 & 3.32 \end{bmatrix} \tag{10}$$

The overall quality metric, Q_{tot} was found to be 2.24, and $\mathbf{Q}_{\text{rel}}$, illustrated in Fig. 5 and Fig. 6, was obtained. We use a color code to identify quality where green denotes that the corresponding cells are close to each other in terms of quality, i.e. the angle is close to $0°$, and red denotes that there is a significant contrast between the quality of the corresponding cells, i.e. far from $0°$. The resultant matrix is anti-symmetric as the relation between each two cells is bidirectional. This result is very important as it shows the status-quo to the administrator and sheds the light on where more participants are needed in order to balance the distribution among the cells.

V. Conclusion

A simple method for the evaluation of coverage within the context of MCS is proposed. It consists of two main quantities to be computed: the overall quality, and the relative quality matrix. Both quantities come from an input map of a cell-specific quality metric. The overall quality provides a global picture of the area of interest within the MCS system, while the relative quality matrix provides insight into the spatial cells of the MCS system, and the contrast of quality between them. This allows the MCS administrator to identify which cells are not yielding reliable readings due to low quality. Furthermore, the cells that satisfied a certain quality threshold can also be identified.

The relationship between adjacent cells was also acquired, which allows the MCS administrator to decide whether to actively request MCS participants to move from one cell to another, or change the incentivization policy in a manner that would introduce new recruits to the area of interest and its cells, or to achieve coverage uniformity. The notion of coverage can further be extended from spatial into spatiotemporal by considering the variation of coverage within time, treating cells as adjacent to each other in both, space and time. The developed metric is of importance for commercial applications as it provides an analysis toolbox of coverage within the context of MCS. This allows the MCS administrator to gain further insights in spite of cost limitations.

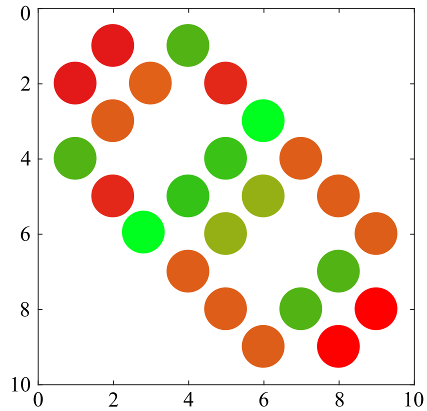

Fig. 6: Spy Plot of $\mathbf{Q}_{\text{rel}}$

Acknowledgment

This work was made possible by NPRP grant NPRP 9-185-2-096 from the Qatar National Research Fund (a member of the Qatar Foundation). The statements made herein are solely the responsibility of the authors.

References

[1] V. Freschi, S. Delpiori, E. Lattanzi, and A. Bogliolo, "Bootstrap based uncertainty propagation for data quality estimation in crowdsensing systems," Springer, pp. 1146–1155, 2017.

[2] B. Guo, Z. Wang, Z. Yu, Y. Wang, N. Y. Yen, R. Huang, and X. Zhou, "Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm," *ACM Comput. Surv.*, vol. 48, no. 1, pp. 7:1–7:31, Aug. 2015.

[3] M. Zappatore, A. Longo, and M. A. Bochicchio, "Using mobile crowd sensing for noise monitoring in smart cities," in *2016 International Multidisciplinary Conference on Computer and Energy Science (SpliTech)*, July 2016, pp. 1–6.

[4] G. Cardone, A. Corradi, L. Foschini, and R. Ianniello, "ParticipAct: A large-scale crowdsensing platform," *IEEE Transactions on Emerging Topics in Computing*, vol. 4, no. 1, pp. 21–32, Jan 2016.

[5] R. F. El-Khatib, N. Zorba, and H. S. Hassanein, "Multi-tasking for cost-efficient mobile crowdsensing under uniformity constraints," in *IEEE ICC, Kansas City, USA*, 2018.

[6] L. Wang, D. Zhang, C. Chen, X. Han, and A. M'hamed, "Sparse mobile crowdsensing: challenges and opportunities," *IEEE Communications Magazine*, pp. 161–167, 2016.

[7] Y. Sun, H. Song, A. J. Jara, and R. Bie, "Internet of things and big data analytics for smart and connected communities," *IEEE Access*, vol. 4, pp. 766–773, 2016.

[8] S. B. Azmy, N. Zorba, and H. S. Hassanein, "Robust quality metric for scarce mobile crowd-sensing scenarios," in *IEEE ICC, Kansas City, USA*, 2018.

[9] S. B. Azmy, R. A. Sneineh, N. Zorba, and H. S. Hassanein, "Small Data in IoT: An MCS Perspective," *Performability in Internet of Things*, 2019, in-press.

[10] C. E. Shannon, "Communication in the presence of noise," *Proceedings of the IRE*, vol. 37, no. 1, pp. 10–21, Jan 1949.

[11] N. Stubbs and S. Park, "Optimal sensor placement for mode shapes via shannon's sampling theorem," *Computer-Aided Civil and Infrastructure Engineering*, vol. 11, no. 6, pp. 411–419, 1996.