Optimal Transport for Mobile Crowd Sensing Participants

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*—Smart cities are becoming more complex and greater volumes of data are required for its efficient operation. Mobile Crowdsensing (MCS) is a paradigm that employs smartphones as instruments to collect data, where the recruitment of participants is based on rewards and incentives. However due to the mobile nature of people, sensing may not be available in a specific area of interest, reducing the quality of the MCS inference of that region. In this paper, we propose a method that utilizes optimal transport so that the MCS administrator could direct participants towards areas with poor quality to improve overall quality. An analysis of optimal transport is presented where the method is evaluated using computer simulations, where it is shown to be efficient for moving participants among spatiotemporal cells.**

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

I. Introduction

The development of the Internet of Things (IoT) paradigm has provided fertile grounds for various sub-paradigms and frameworks that integrate the cyber and the physical aspects of life. This integration has revolutionized the technology behind sensors, completely incorporating it in the structure of tomorrow's smart cities. As a consequence, the pursuit of increased access to data emerged as the administrators and stakeholders of smart cities seek to improve the efficiency of operation. This led to the development of the Mobile Crowd Sensing (MCS) which exploits the presence of sensors in *the crowd*'s smartphones, employing them as an extended instrument [1]. This enabled smart city administrators to *crowdsource* a wide range of cyber-physical data, and provided them with social and physical insights about the smart city [2].

In MCS, administrators have a certain objective for which they create a *task*, which is assigned to a *participant*, who works on its execution, via a centralized *server*. An example of a task can be to acquire samples of noise pollution using microphones [3]. MCS tasks can be classified according to the nature of their execution as *participatory* or *opportunistic*. Participatory sensing actively involves users in the sensing process, such as taking a photo or writing a comment [4]. On the other hand, opportunistic sensing passively conducts the sensing task when the participant's device satisfies the set of conditions defined for the task, such as time, place, device capabilities, battery, etc. [2]. To facilitate the complex implementation of MCS, various frameworks and models were developed. An interesting model, in particular, is developed in [5] which models the MCS sensing problem over *spatiotemporal cells* whose bounds are defined in time and space.

MCS is generally treated as a problem with large scale data under the assumption that participants - and thus data - are always available in abundance. For example, the scale of data employed in [4], [6] is large involving many participants. For a large smart city, this is essentially true at a global scale. However, at the local scale such as a neighborhood, the problem of participant scarcity is a likely possibility. The participant set that MCS data comes from, within a spatiotemporal cell, is diverse and thus prone to variations in quality, performance, and capability. To overcome this heterogeneity, small sample quality metrics were developed in [7] to quantify and evaluate the reliability of participants' readings. The feedback of such metrics allows the elimination of inconsistencies while providing a local picture of the participants' reliability. Furthermore, a coverage metric based on the small sample quality metrics was developed in [8] to allow the MCS administrators to select enough participants to conduct the needed MCS tasks. The coverage metric provides a global picture of the MCS region as a whole, over all cells, while providing a more local picture by assessing how quality varies from one cell to another.

A description of the minimum number of needed readings to achieve a predefined inference quality allows MCS administrators to optimize costs, which are based upon the number of participants. The cost is considered either in terms of monetary incentives, or as the amount of consumed data/energy [9]. Due to the mobility nature of smartphones, the application of MCS on an urban city is challenging as the number of participants is not uniform all over it. Areas which are more active than others will show better MCS inference capability (that we define as coverage), while areas without much participants are going to lack coverage. If the MCS participants' decisions on mobility could be controlled, by means of a reward to achieve a specific objective, then the quality desired by the MCS administrator could be achieved. Thus it is of importance that such mobility recommendations are efficient between cells, incurring the least cost possible. In this paper, we describe how *optimal transport* and its theory [10] can be used to optimize the MCS system performance in terms of quality and cost; and to solve some of the problems that rise during participant recruitment within MCS.

This paper is structured as follows: Section 2 provides

978-1-5386-7646-2/19/$31.00 ©2019 IEEE
Authorized licensed use limited to: Queen's University. Downloaded on June 11,2020 at 17:37:50 UTC from IEEE Xplore. Restrictions apply.

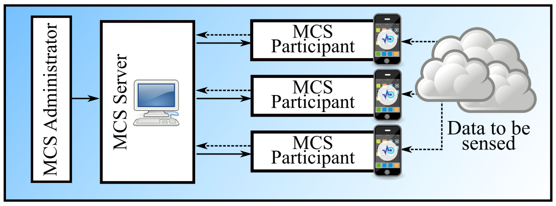

Fig. 1: General Architecture of an MCS System

an overview of MCS architecture, the spatiotemporal model, as well as the aforementioned quality and coverage metrics. MCS Use cases are also discussed in this section. Section 3 introduces the optimal transport theory, the developed optimal transport algorithm, and the potential applications for which optimal transport could be employed within MCS; Section 4 provides a simulated instance of the proposed algorithm. Section 5 concludes with an epilogue on optimal transport in the context of MCS operation.

II. Overview of MCS Framework

As part of IoT, MCS aims to integrate the cyber and physical spaces together. By utilizing smart devices and engaging their users in the sensing process, MCS allows administrators to leverage users and their mobility for the general benefit of the smart city while providing them with a service or a reward in exchange. However, since the implementation of MCS is usually remote, the framework usually involves Administrators, Systems (or server), and participants. Several works such as [2], [11], [12] have provided detailed descriptions of the MCS architecture. Figure 1 shows a simple description of the MCS system architecture. The MCS system generally consists of three main elements, as previously mentioned:

- **MCS Administrator:** is the prime mover of the MCS system who designs and publishes tasks to be picked up and executed by the participants. Within the context of the smart city, the MCS administrator is interested in collecting data about a specific phenomenon at a specific time or location for which they design and publicly post their tasks. Through MCS, the administrator achieves the objectives of the smart city and enhances its operational efficacy.
- **MCS Participants:** they are crowd members who participate in the MCS system by accepting task assignments from the MCS administrator to execute, in exchange for an incentive service or a monetary reward. They are capable of executing tasks participatorily, being actively involved, or opportunistically, passively by satisfying the conditions. They communicate with the MCS administrator via an application on their device that connects them to the MCS server. Furthermore, the mobility of the participants allows the MCS administrator to pick participants whose conditions maximize the potential sensing over the smart city.
- **MCS Server:** is the core of the MCS system, as it is the link that connects the MCS participants and the MCS administrator. The process of participant recruitment, which maps tasks to participants, is automatically performed by the system. The system assigns tasks and recruit participants while evaluating their performance, reliability and quality. It also computes and delivers the reward deserved by the participants.

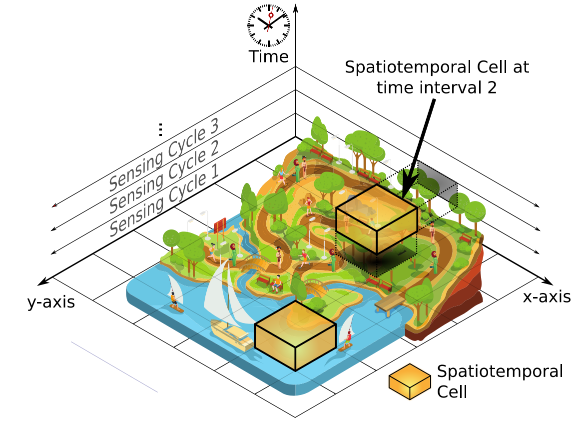

Fig. 2: Spatiotemporal MCS Model

A. MCS Spatiotemporal Model

In MCS, an Area of Interest (AoI) requires to be divided into geofences or cells to which participants are assigned. The granularity of this division reflects in the precision of a phenomenon's characterization within a certain region; i.e. requiring a specific task resolution. However, time is a factor that must be taken into consideration when assigning participants from a dynamic crowd. This led to the development of the *Spatiotemporal Model* that divides an AoI into space and time [5]. Figure 2 illustrates a similar diagram for spatiotemporal cells. MCS administrators are required to divide space and time into cells and sensing cycles that satisfy the purpose of the task at hand, while maintaining consistency with the spatial and temporal versions of Shannon's sampling theorem [13]. In general, within the m^{th} cell, a task will be assigned to N_m participants, who return a set of readings $U_m = \{u_{m,1}, u_{m,2}, \ldots, u_{m,N_m}\}$. In the spatiotemporal model, each reading $u_{m,n}$ is a random variable obtained by the n^{th} participant in the m^{th} spatiotemporal cell. Each cell m corresponds to a 3-tuple $m = (i, j, t)$, which correspond to the location of the m^{th} cell in space and time (a_m, b_m, c_m).

The truth of the sensed quantity in a specific cell, μ_m, is then estimated from X_m by computing the mean:

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

B. Small Sample Quality and Coverage Metrics

The lack of knowledge about the ground truth, μ, is a factor that led to the development of inference models, which employ participants as a proxy to asserting the truth. However, the participant set's heterogeneity poses a challenge to asserting the true value, especially if the number of participants is not enough. To that end, the quality metrics developed in [7]

Authorized licensed use limited to: Queen's University. Downloaded on June 11,2020 at 17:37:50 UTC from IEEE Xplore. Restrictions apply.

provide methods by which a numerical quantity is ascribed to a cell to describe the quality of the set of readings obtained within that cell. The approach employed in [7] defines quality as a quantity, Q, based on the difference between the sample mean and robust centrality estimates: the sample mean after filtering outliers using the median absolute deviation (MAD-mean) and the k-trimmed mean. This difference statistic was called the Mean MAD-Mean Trimmed Mean (MMTM) difference, denoted θ_{MMTM}. It provides the administrator with a degree of control in defining the quality via a sensitivity parameter η which allows the toleration of certain outliers, arising from the phenomenon's natural distribution to be included as a reading, unlike certain abnormalities completely alien to the phenomenon. θ_{MMTM} is employed to evaluate small-sample quality in [7] for $N_m > 11$, while it is combined with the costly statistical bootstrap in [14] for $N_m > 8$, where θ_{MMTM} is expressed as:

$$\theta_{\text{MMTM}} = \eta(\bar{u} - \bar{u}_k) + (1-\eta)(\bar{u} - \bar{u}_{\text{MAD}}) \tag{2}$$

with $\eta \in [0,1]$, $\bar{U}$ is the sample mean of the set of readings U_m within the m^{th} cell, $\bar{U}_k$ is the k-trimmed mean, and $\bar{u}_{\text{MAD}}$ is the MAD-mean.

Quality can then be describes as a measure of how good a sample is in estimating the ground-truth, or precisely as a quantity that reflects the acquired samples' reliability in estimating the ground-truth. Quality, as per [7], is defined as:

$$Q = \begin{cases} \frac{1}{2}\log_\gamma\left(\frac{1}{\mu_{\text{MMTM}}\sigma^2_{\text{MMTM}}}\right) & 8 < N < 11 \\ \log_\gamma\left(\frac{1}{\theta_{\text{MMTM}}}\right) & 11 < N < 30 \end{cases} \tag{3}$$

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

In addition, a coverage quality metric was developed in [8] to extend the notion of quality from a specific cell to that of an AoI during the t^{th} sensing cycle. Quality, for a region of interest, can be described as a matrix at t^{th} sensing cycle, $\mathbf{Q}_{\text{map},t}$ as:

$$\mathbf{Q}_{\text{map},t} = \begin{bmatrix} Q_{1,1,t} & Q_{1,2,t} & \cdots & Q_{1,S,t} \\ Q_{2,1,t} & Q_{2,2,t} & \cdots & Q_{2,S,t} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{S,1,t} & Q_{S,2,t} & \cdots & Q_{S,S,t} \end{bmatrix} \tag{4}$$

where $Q_{i,j,t}$ is the quality for the m^{th} cell $\big($where m is the 3-tuple (i,j,t) as per the spatiotemporal model$\big)$, and S represents the side-length of the square grid into which the AoI is divided.

The corresponding overall coverage metrics for a quality map, $\mathbf{Q}_{\text{map},t}$, is defined as the mean $Q_{\text{tot},t}$:

$$Q_{\text{tot},t} = \text{mean}(\mathbf{Q}_{\text{map},t}) \tag{5}$$

That is represented in Figure 3 as the yellow plane. Another metric, called the relative quality metric, and denoted as $\mathbf{Q}_{\text{rel}}$, is a $S^2 \times S^2$ matrix whose elements are defined as the angle

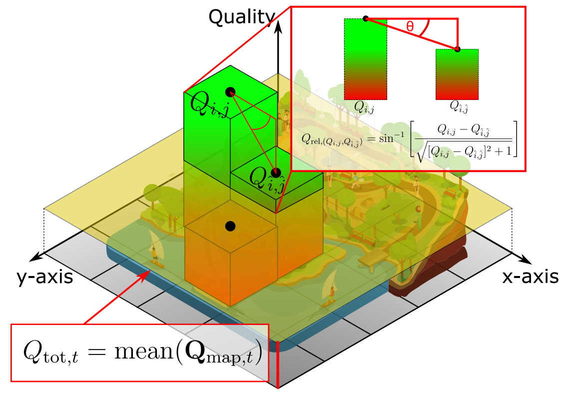

Fig. 3: Coverage Quality Metric: Yellow plane describes overall coverage quality, while angle of depression describes the change of quality between neighbouring cells.

of depression between neighboring quality values:

$$Q_{\text{rel},(Q_{i,j},Q_{\hat{i},\hat{j}})} = \sin^{-1}\left[\frac{Q_{i,j} - Q_{\hat{i},\hat{j}}}{\sqrt{[Q_{i,j} - Q_{\hat{i},\hat{j}}]^2 + 1}}\right] \tag{6}$$

where (i,j) corresponds to the coordinates of the of k^{th} cell, and $(\hat{i},\hat{j})$ correspond to the coordinates of an arbitrary adjacent cell. Figure 3 provides an illustration of Eq. 6.

C. MCS Use Cases

Optimal Transport can be used in any MCS scenario, as MCS is mainly based on smartphones that move and can be "transported". A potential use case are the games based on Augmented Reality (e.g., Pokémon Go). Consider the city council (MCS Administrator) willing to infer about an event in a specific location in the city (e.g., a park) and finds that the number of available sensors is low or null in that location. A possible solution is to trigger people to move to that specific location (Optimal Transport in our case), and AR games are a very good "tool" that can help. By moving the players (and obviously their smartphones) to that location to catch a Pokémon in that park. The availability of several players will enable the needed measurements in the park and quality inference is achieved. Releasing Pokémon creatures with different features (the incentive in MCS) will trigger more or less people (the MCS participants) to move towards a certain location and time (MCS spatiotemporal cell). The Pokémon features/powers/reward would trigger more or less visitors to move, and an optimization of that reward is desired.

A similar use case would be the theme/amusement parks, where the manager (MCS administrator) interested to know the situation in a blind spot can move some crowd (MCS participants) by announcing reduced fees for the rides (the incentive in MCS) at a specific location and time (MCS spatiotemporal cell). The amount of reduction in the fees would trigger more or less visitors to move, and an optimization of that amount is needed.

Authorized licensed use limited to: Queen's University. Downloaded on June 11,2020 at 17:37:50 UTC from IEEE Xplore. Restrictions apply.

III. Optimal Transport for MCS Application

The Optimal Transport problem was found to be particularly useful in resource allocation problems [10], [15]. The discrete version of the transport problem, proposed in 1942 by Leonid Kantorovich, is of relevance to the setting of MCS as the participant recruitment problem is one to which optimal transport can be directly applied to maximize its efficiency. In this section, we provide an overview of optimal transport as well as an algorithm for transporting participants between cells in order to reduce MCS costs while maintaining MCS quality. The feasibility of such transport lies in providing the administrator a degree of control over the participants' mobility, enabling a better characterization of a phenomenon over the area of the smart city. The developed algorithm relies on the previous works that provided metrics for quality and coverage [7], [14]. We also propose alternative usages for it within the MCS context that focus on participants themselves rather than the cells.

A. Optimal Transport

The Optimal Transport theory is the most affordable coupling, or transport plan, to transport between two probability vectors $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ such that the cost of transporting all the elements in $\boldsymbol{\alpha}$ to $\boldsymbol{\beta}$ is minimal. The distributions can be described as:

$$\boldsymbol{\alpha} = \sum_{k=1}^{N_\alpha} \alpha_k \delta_{x_k}, \qquad \boldsymbol{\beta} = \sum_{l=1}^{N_\beta} \beta_l \delta_{y_l} \tag{7}$$

where α_k, β_l correspond to masses located at x_k, y_l, respectively, and δ_x, δ_y correspond to the dirac delta located at position x, y, respectively.

The cost which is to be minimized is a function $c(x_k, y_l)$ that maps the choice of $\alpha \in \boldsymbol{\alpha}$ and $\beta \in \boldsymbol{\beta}$ to a value. It is best described as a matrix $\mathbf{C}$ where $\mathbf{C}_{kl} = c(x_k, y_l) \in \mathbb{R}^{N_\alpha \times N_\beta}$. In most of Optimal Transport problems, the cost is often defined based on the euclidean distance where $c(x_k, y_l) = \|x_k - y_l\|^2$.

The optimal coupling, which is the best transport plan, is to be selected from a set of admissible couplings $\mathbf{U}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ which contains all the possible couplings $\mathbf{P} \in \mathbb{R}_+^{N_\alpha \times N_\beta}$ such that the sum over all the columns of a $\mathbf{P}$ would yield the vector $\boldsymbol{\alpha}$ and over all the rows would yield vector $\boldsymbol{\beta}$. An element, $\mathbf{P}_{kl}$, of a coupling matrices describes the precise amount of mass to be transferred between each mass $k \in \boldsymbol{\alpha}$ to the corresponding mass $l \in \boldsymbol{\beta}$. The set of admissible couplings $\mathbf{U}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is a bounded convex polytope defined by $N_\alpha + N_\beta$ constraints [10]. The optimal transport problem is that of finding an optimal coupling $\mathbf{P}^\star$ such that the average cost is minimal. This can be described using mathematical optimization notation as:

$$\begin{array}{ll} \text{minimize} & \langle \mathbf{C}, \mathbf{P} \rangle \\ \text{subject to} & \mathbf{P} \in \mathbf{U}(\boldsymbol{\alpha}, \boldsymbol{\beta}) \end{array} \tag{8}$$

where $\langle \mathbf{C}, \mathbf{P} \rangle = \sum_{kl} \mathbf{C}_{kl} \mathbf{P}_{kl}$.

Since $\mathbf{U}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is a polytope, linear programming is feasible method for solving optimal transport problems. The authors in [10] cast the optimal transport problem in the standard form of linear programming as:

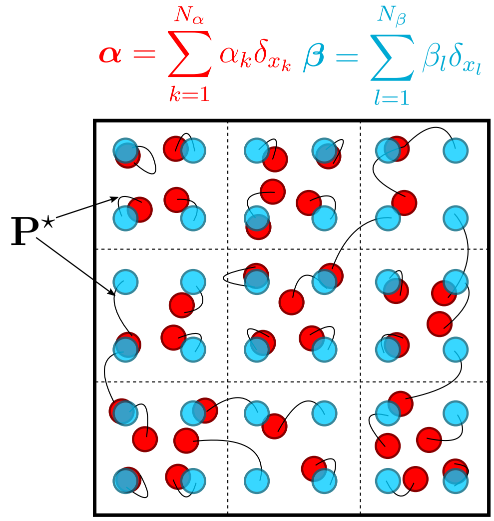

Fig. 4: Rough Illustration of the Optimal Transport between discrete measures

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T \mathbf{p} \\ \text{subject to} & \mathbf{p} \in \mathbb{R}^{n_\alpha n_\beta} \\ & \mathbf{A}\mathbf{p} = \left[\begin{smallmatrix}\boldsymbol{\alpha} \\ \boldsymbol{\beta}\end{smallmatrix}\right] \end{array} \tag{9}$$

where $\mathbf{c}$ is the flattened form of the matrix $\mathbf{C}$, $\mathbf{p}$ is the flattened form of the matrix $\mathbf{P}$, and the matrix $\mathbf{A}$ is defined as:

$$\mathbf{A} = \begin{bmatrix} \mathbb{1}_{N_\alpha}^T \otimes \mathbb{I}_{N_\beta} \\ \mathbb{I}_{N_\alpha} \otimes \mathbb{1}_{N_\beta}^T \end{bmatrix} \in \mathbb{R}^{(N_\alpha + N_\beta) \times N_\alpha N_\beta} \tag{10}$$

where $\mathbb{1}_L$ is the indicator vector of length L, and $\mathbb{I}_L$ is the identity matrix of size $L \times L$.

The Kantorovich formulation of the Optimal Transport problem is more suitable for masses defined discretely, such as $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. Figure 4 provides an illustrative example in which the red circles correspond to the source distribution $\boldsymbol{\alpha}$ while the blue circles correspond to the target distribution $\boldsymbol{\beta}$, and the black lines show a rough example of a non-split coupling. However, for MCS, the splitting of masses needs to be carefully done so that the probability vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ map to whole numbers, as they would represent the transport of participant.

B. Optimal Transport for Coverage Enhancement in MCS

We employ Optimal Transport for the purpose of enhancing quality within MCS systems. In the context of MCS, we describe the m^{th} cell as static mass, δ_m, at a fixed location - the center of a cell - whose weight (or mass) represents the number of participants N_m at a specific time $t = t_i$. However, the cost $c(x_k, y_l)$ will no longer be defined as the ground distance usually considered in Optimal Transport, it would rather be changed to reflect the cost aspect from an MCS perspective. Since $\mathbf{Q}_{rel}$ matrix contains the angles between neighbouring Q_m readings, we utilize it to define the cost $\mathbf{C}$ in the Optimal Transport. This modified cost incorporates quality as defined in Eq. 3, and relates to the number of participants

Authorized licensed use limited to: Queen's University. Downloaded on June 11,2020 at 17:37:50 UTC from IEEE Xplore. Restrictions apply.

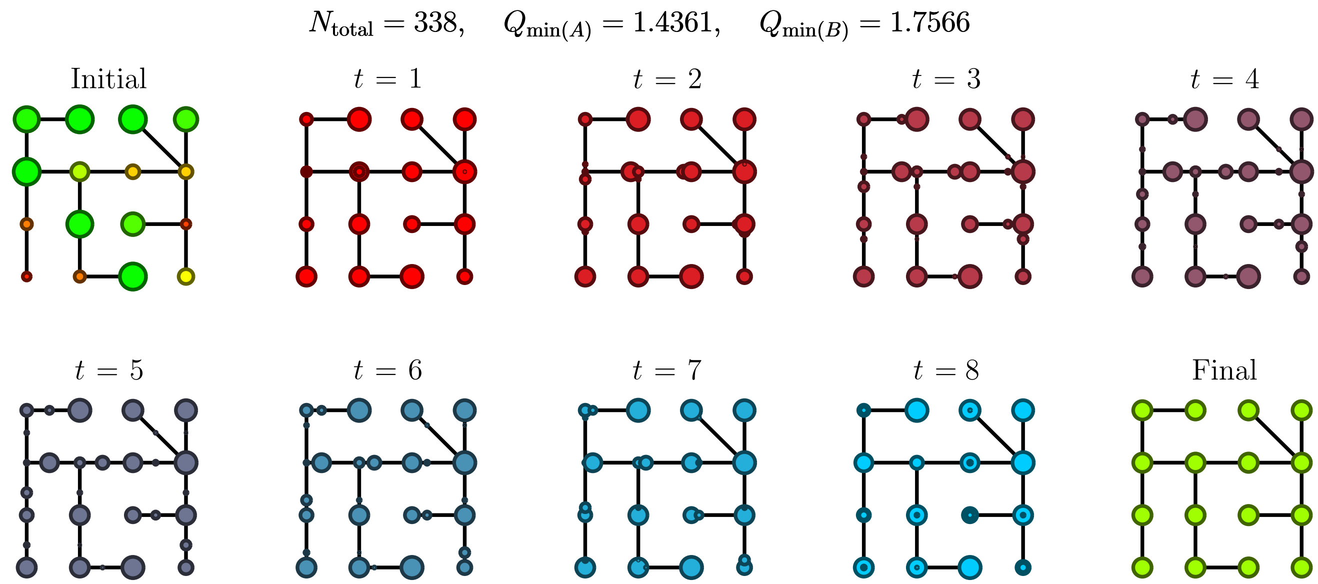

Fig. 5: Simulation of the optimal transport of MCS participants

within a cell and to cost in MCS. Furthermore, the ground distance is proportional to the cost of the incentive that the MCS administrator will offer to induce participants to move from cell k to cell j. The updated cost then becomes:

$$\mathbf{C}_{\text{MCS}} = \{\mathbf{C}_{\text{MCS},kl} : \mathbf{C}_{\text{MCS},kl} = \|x_k - y_l\|^2 \cos^2(\mathbf{Q}_{\text{rel},kl})\} \quad (11)$$

where the cost function $c(x_k, y_l) = \|x_k - y_k\|^2$, and the $\cos^2$ incorporates the cosine into the square, thus modifying the length of ground distance to that of the hypotenuse described in Figure 3.

Nevertheless, the Kantorovich formulation allows the splitting of masses - and thus participants. This requires the optimal transport to be corrected after its computation to ensure that the masses being transported between cells are whole numbers. The probability vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ corresponds to the ratio of participants in each cell to the whole $N_{\text{total}} = \sum_{k=1}^{M} N_k$, similar to the optimal coupling $\mathbf{P}^\star$ that describes ratios of participants to be transported. As such, the algorithm requires to compute actual number of participants being transported and verify that they are whole numbers and that the total number of participants is conserved between the source $\boldsymbol{\alpha}$ and the target $\boldsymbol{\beta}$, being N_{total}. If it fails to do so, a simple solution is to compute the difference between the total of the rounded optimal coupling and N_{total} and take it to the maximum value within the coupling. Algorithm 1 presents the steps through which coverage quality is enhanced. As this algorithm is based on the optimal transport's formulation as a linear program, the simplex method is employed in finding $\mathbf{P}^\star$. The simplex algorithm was found in [16] to have a polynomial smoothed complexity.

Algorithm 1 Algorithm for Optimal Transport among Cells for MCS

Input: $\mathbf{C}, \boldsymbol{\alpha}, \boldsymbol{\beta}, N_{\text{total}}, \mathbf{Q}_{\text{rel}}$
Output: $\mathbf{P}^\star$
$\mathbf{C}_{\text{MCS}} \leftarrow \{\mathbf{C}_{\text{MCS},kl} : \mathbf{C}_{\text{MCS},kl} = \|x_k - y_l\|^2 \cos^2(\mathbf{Q}_{\text{rel},kl})\}$
$\mathbf{c} \leftarrow$ `flat`$(\mathbf{C}_{\text{MCS}})$
$\mathbf{p} \leftarrow$ `flat`$(\mathbf{P})$
Compute $\mathbf{A}$ (Eq. 10)
$\mathbf{P}^\star$ for `f= min` $\mathbf{c}^T\mathbf{p}$`:` $A\mathbf{p} = [\boldsymbol{\alpha};\boldsymbol{\beta}]$`,` $\mathbf{p}$ `>=` 0 using simplex method.
Rounded $\mathbf{P}_{\text{R}} \leftarrow \lceil N_{\text{total}} \mathbf{P}^\star \rfloor$
$N_{R,\text{total}} \leftarrow \sum_{kl} \mathbf{P}_{\text{R},kl}$
$N_{\text{diff}} \leftarrow N_{R,\text{total}} - N_{\text{total}}$
Update $\max \mathbf{P}^\star \leftarrow \max \mathbf{P}^\star - N_{\text{diff}}$
return $\mathbf{P}^\star$

IV. Computer Simulation

To test the proposed algorithm over MATLAB, we generated M random integers between 11 and 30 corresponding to the number of participants present in each cell, N_m, present over the MCS AoI, placed in vectors $\boldsymbol{\alpha}, \boldsymbol{\beta}$. The results are presented in Figure 5, with the transport being illustrated over 8 iterative steps; it is a potential benefit to the MCS administrator to conduct sensing over intermediate cells in sparse systems during the transport of the participants. The sizes of the circles represent the number of participants being transported, while the colors in the beginning and the end represent the quality, where red is low quality and green is high quality.

It can be seen in the "Initial" stage of the optimal transport, the quality is not uniform over all cells, and the number of participants is very low in some cells while others are over-populated. At $t = 1$, the masses split to illustrate the detachments of participants. From $t = 1$ to $t = 8$ the participants movement is depicted over the black lines, which illustrate the presence of a coupling between the k^{th} and the l^{th} cell. At the "Final" stage, the participants join their corresponding cells, yielding a uniform profile of quality over the AoI, while the total number of participants remained the same between "Initial" and "Final". The duration of a cycle

Authorized licensed use limited to: Queen's University. Downloaded on June 11,2020 at 17:37:50 UTC from IEEE Xplore. Restrictions apply.

could vary during MCS.

V. Conclusion

In summary, MCS requires the recruitment of participants who would reliably collect and report information. Over an area of interest, participants might not be available. This requires the MCS administrator to request MCS participants to move from one region to another within the area of interest. Doing so would incur costs for the MCS system, thus the requested transport plans require to be optimal, minimizing the cost while enhancing the coverage. In this paper, a method for transporting MCS participants between cells was proposed. This method considers a current MCS coverage quality of the area of interest and aims to transport participants between cells in a manner that improves the overall coverage quality.

The use of Optimal Transport in MCS is interesting, and it can be of benefit to MCS systems as the cost of the transport could be redefined to include details regarding the participants or their cells, such as their reputation, trust, capability, and others. This paper has examined an initial scenario for which optimal transport was used.

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] 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.

[6] 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.

[7] 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.

[8] S. B. Azmy, N. Zorba, and H. S. Hassanein, "Quality of coverage: A novel approach to coverage for mobile crowd sensing systems," in *2018 Global Information Infrastructure and Networking Symposium (GIIS)*, 2018.

[9] L. Chen, L. Wang, D. Zhang, S. Li, and G. Pan, "Enup: Energy-efficient data uploading for mobile crowd sensing applications," in *UIC-ATC-ScalCom*, 2016.

[10] G. Peyré and M. Cuturi, "Computational Optimal Transport," *ArXiv e-prints*, Mar. 2018.

[11] D. Zhang, L. Wang, H. Xiong, and B. Guo, "4w1h in mobile crowd sensing," *IEEE Communications Magazine*, vol. 52, no. 8, pp. 42–48, Aug 2014.

[12] J. Liu, H. Shen, H. S. Narman, W. Chung, and Z. Lin, "A survey of mobile crowdsensing techniques: A critical component for the internet of things," *ACM Transactions on Cyber-Physical Systems*, vol. 2, no. 3, p. 18, 2018.

[13] 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.

[14] S. B. Azmy, N. Zorba, and H. S. Hassanein, "Bootstrap-based quality metric for scarce sensing systems," in *2018 IEEE Global COmmunications Conference: Ad Hoc and Sensor Networks (Globecom2018 AHSN), Abu Dhabi, UAE*, 2018.

[15] S. Kolouri, S. R. Park, M. Thorpe, D. Slepcev, and G. K. Rohde, "Optimal mass transport: Signal processing and machine-learning applications," *IEEE Signal Processing Magazine*, vol. 34, no. 4, pp. 43–59, July 2017.

[16] D. A. Spielman and S. Teng, "Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time," *CoRR*, vol. cs.DS/0111050, 2001. [Online]. Available: http://arxiv.org/abs/cs.DS/0111050

Authorized licensed use limited to: Queen's University. Downloaded on June 11,2020 at 17:37:50 UTC from IEEE Xplore. Restrictions apply.