Sense-making from Distributed and Mobile Sensing Data: A Middleware Perspective

Santanu Sarma, Nalini Venkatasubramanian, and Nikil Dutt

Department of Computer Science, University of California, Irvine, USA

[santanus, nalini, dutt]@ics.uci.edu

Abstract

This paper presents a scalable and collaborative mobile crowdsensing framework for efficient collective understanding of users, contexts, and their environments. Collaborative mobile crowdsensing enables information to be gathered and shared by users who are directly involved (participatory sensing) or integrated seamlessly as needed (opportunistic sensing) through user mobile platforms. To address the scalability needs of the mobile ecosystem, we additionally employ compressive sensing techniques for approximate gathering and processing of sensor data - this requires new mechanisms for sensor data collection, tunable approximate processing, and mobile networking architecture, to create a compressive collaborative mobile crowdsensing platform called SenseDroid. The proposed framework is build using a multi-tired hierarchical architecture to sense spatial variations of a parameter of interest, perceive spatio-temporal fields, and enable energy efficient local mobile sensing with a small number of measurements. This approximate, yet tunable approach combines different sensing approaches opportunistically while trading scalability (and coverage) for data accuracy (and energy efficiency). In this paper we propose and discuss the framework and the challenges associated with compressive and collaborative mobile sensing for multi-tired hierarchical mobile network architecture for emerging mobile collaborative applications.

General Terms Mobile Sensing, Distributed Middleware, Internet-of-Things

Keywords Compressive Sensing, Participatory Sensing, Collaborative Sensing, Sensor Network, Mobile Phone Sensing.

1. Introduction

The proliferation of device platforms and mobile applications (12 billion devices and over 3 million apps by 2017) has changed how humans interact. These new interactions enable the mobile device/user to be an active participant in the collection, sharing and dissemination of information- end platforms/users capture and process local context and communicate this information to other platforms/services using heterogeneous connectivities. We envision that the next generation mobile ecosystem will be far more sophisticated, complex and diverse than what is currently used. With each passing year, we see increasingly powerful smartphones being manufactured, which have a plethora of powerful embedded sensors like microphone, camera, digital compass, GPS, accelerometer, temperature sensors and many more [8]. Such platforms also support seamless external sensor connectivity (e.g. on body for health and wellness monitoring) that further equips such devices with rich and unique sensing capabilities. Moreover, the ability to easily program today's smartphones, enables us to exploit these sensors, in a wide variety of application such as personal safety, emergency and calamity response, situation awareness, remote activity monitoring, transportation and environment monitoring [7, 8]. Furthermore, the ability to share the sensed content with other users and applications has enabled crowds/humans (and their devices) to become information providers in a *crowdsensing* ecosystem. Effective use of the sensed data relies on effective "sensemaking" that transforms the gathered data to meaningful information for improved situational awareness, decision making and control. This paper focuses on enabling such "sensemaking" from mobile users and devices.

Broadly, the emerging field of mobile phone sensing or crowdsensing can take multiple forms [7]. In *participatory sensing*, the user is directly involved in the sensing activity; this burden is alleviated in the *opportunistic sensing* paradigm by delegating and automating the sensing task to the mobile phone sensing system. In this paper, we argue for a collaborative sensing approach where the users collaborate or cooperate to have better and reliable sensing information and obtain missing sensing information when specific sensors are not available in their own devices. Collaboration can be useful in generating more accurate and reliable information of spatial fields distributed across geographical areas and region, e.g., multiple temperature sensor readings in a space would be more reliable than a single reading. We discuss some specific applications of collaborative sensing and sensemaking in following use case scenarios.

Disaster and emergency response: Mobile intelligent networks can play a key role in emergency response, surveillance and security, and battlefield operations. Consider a fire scenario where information from in-situ and mobile sensors can help in incident perimeter assessment as well as rapid localization of regions with high impact. Coordination among fire fighters is another important aspect in fire rescue operations. A collaborative mobile crowdsensing framework can be used to coordinate among the firefighters for their own safety and as well as quick evacuation. Collaborative sensing can provide situation awareness of different users in a facility during the rescue operation. Based on the situations, rescue operations can be coordinated more effectively to reduce response time and save precious lives.

Personal health monitoring and wellness: Mobile phone sensing has the potential to continuously collect/sense data for health and wellness analysis. UbiFit Garden [3] is a mobile phone sensing system jointly developed by Intel and University of Washington, which uses small inexpensive on-body sensors and mobile phones along with machine learning techniques for activity modeling to infer people's activities throughout everyday life. In [12], stress level of mobile user was measured using mobile phones, while [11] explored the use of smartphones in predicting the mode of the users. This can be extended to a family or a group of related people to jointly infer their moods, and exercise routines, exposures to pollutants etc. to find combined stress quotient. The same can also be used to achieve a family health indicator.

DAC'14, June 01 - 05 2014, San Francisco, CA, USA.

http://dx.doi.org/10.1145/2593069.2596688

Smart spaces and their effective utilization: Smart buildings and smart spaces can use a collaborative sensing framework to monitor dynamic environmental conditions and requirements (e.g air conditioning and lighting preferences) and allows individuals to tailor lighting levels to their personal preferences and tasks to save energy footprints [18]. It can be used to understand the pattern of a facility usage (e.g. a library or a museum) and understand group behavior to improve the facility and its service.

The rapid growth in mobile sensors (in addition to sensors in our surrounding environment) and sensing data poses a serious challenge to the existing and traditional sensemaking paradigm. Sensemaking from large numbers of heterogenous sensors, data, and mobile platforms is an extremely challenging task - there is a need for new architectures and softwares that can support sensemaking both effectively and at scale. Another interesting aspect has to do with the accuracy of sensing/sensemaking. Applications require information at different levels of accuracy and resolution - these tolerances can be leveraged to tradeoff accuracy for scale. In this paper, we propose a hierarchical and extensible framework to support collaborative sensing at scale. In particular, we discuss two strategies - the use of hierarchical sensing and the application of compressive sensing techniques to address scalability/accuracy concerns in mobile crowdsensing. We implement the framework as a distributed middleware for mobile platforms, called SenseDroid, and explore its use in several emerging mobile collaborative applications. The key benefits of the proposed collaborative sensing framework are as follows:

- ability to opportunistically set different sparsity levels to exploit regional fluctuations
- ability to analyze a region with more emphasis based on criticality or knowledge of events. Multi-resolution compressive thresholds i.e. number of sensing samples collected from a region based on the size and importance.
- ability to use different basis and sensing matrix by exploiting prior available data of different regions
- ability to use heterogeneous sensors with different characteristics and quality (as in different mobile phone)
- enable more energy-efficient sensing in the framework and context-aware information exchange

2. Related Works

Mobile phone sensing has recently attracted extensive research attention from both academia and industry due to its attractive applications. A comprehensive review of these applications can be found in a recent survey paper [8]. Several issues relating to sensing and coverage have been well studied for mobile sensor networks [31]. However, only few works have addressed collaborative and compressive sensing specifically with mobile phones. In [15], the authors presented analytical results on the rate of information reporting by uncontrolled mobile sensors needed to cover a given geographical area, and demonstrate the feasibility of using existing software and standard protocols for information reporting and retrieval to support a large system of uncontrolled mobile sensors using a testbed. In [25], the authors proposed a protocol, Aquiba, that exploits opportunistic collaboration of pedestrians to achieve energy efficiency and reduce data redundancy. Its performance was studied via simulations. In [28], the authors introduced mechanisms for automated mapping of urban areas that provide a virtual sensor abstraction to applications. They also proposed spatial and temporal coverage metrics for measuring the quality of acquired data.

Luo et al. [13] were the first to examine the notion of compressive sensing over large scale wireless sensor networks (WSN) to reduce the number of transmission. Their data gathering compressive scheme reduced the number of transmission from $\mathcal{O}(N^2)$ transmission to $\mathcal{O}(NM)$ where the number of measurement $M \ll N$, the cluster size. However, they assume that the data field is smooth with uniform sensor characteristics, negligible sensor noise and heterogeneity, and global constant sparsity without leveraging the local or regional fluctuations of the signal field. Due to these assumptions as well as unique differences between traditional WSN from mobile phone sensing (e.g. static vs high mobility, limited computation and power resource vs considerable computational and rechargeable energy resource, and mostly broadcast mode operation vs bidirectional multi-network operation, limited number of sensors per node to varied types of sensor in a node), naive and plain implementation of their technique can introduced redundant data communications (e.g. from the leaf nodes) and reduction in overall network throughput [14]. Moreover, the assumption of uniform compression threshold across the network regardless of the data field characteristics and the inability to exploit regional fluctuations along with the sensor characteristics in more realistic situations can results in poor compression efficiency and thereby impacting the energy efficiency. We therefore propose a hierarchical distributed architecture where the local field sparsity and sensing characteristics can be effectively and jointly exploited at different levels for efficiency and scalability. Unlike WSN nodes that usually lack computation and power resources , the mobile phones being considerably resourceful can adopt compressive sensing at each nodes for example in energy efficient context processing as dicussed in the subsequent sections.

3. Hierarchical Architecture and Middleware for Sense-Making

The key idea of our approach to achieve scalable sense-making is to exploit hierarchical architecture combined with tunable/configurable compressive sensing both in spatial and temporal dimensions at different levels. The conceptual architecture of the our framework supporting multi-tiered collaborative and compressive sensing is illustrated in Fig. 1 where the network is hierarchically organized and spatially distributed through multiple local clouds (LCs) which in turn is formed from spatial distribution of nano clouds (NCs). The NCs consists of mobile nodes connected to a central head or a broker. The the head broker in the LCs in turn communicate with other LCs and the public cloud in the next hierarchy. The NCs are formed in the region of interest and the workload of the sink nodes (i.e. broker) is distributed among multiple sink nodes in the LCs such that all the mobile nodes need not flow the information to a single node to overcome network range and scalability bottlenecks. This hierarchy allows the nodes to collaborate through the broker (performing the operation of the sink node/collector) and concatenate the results of the NCs for the local region.

The hierarchical approach is based on the observation that the number of random observations from any region should corresponds to the local spatio-temporal sparsity as well as the NC size in stead of the global sparsity. Intuitively, this should work better than the global scheme as the local correlation among the nodes can be exploited in the local area (i.e. LCs) than global area. Besides, local sparsity is easy to compute and often prior available data about the local regions can be exploited to improve the sensing efficiency and data transmission requirements thereby saving energy.

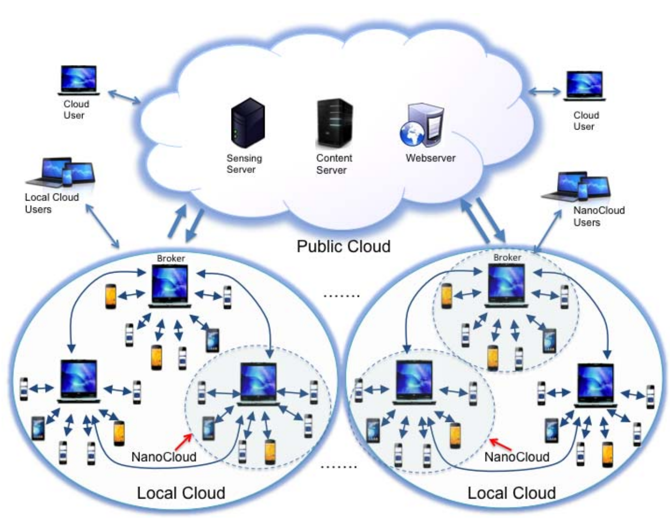

Figure 1. Multi-tiered hierarchical structured mobile cloud architecture for scalable collaboration and compressive sensing. The multi-tiered architecture consists of hierarchy of local cloud which in turn consists of NanoClouds.

The SenseDroid Middleware: We next present **SenseDroid**, a collaborative sensing middleware platform that enables mobile phone based physical sensing and sensemaking. The following are the key features in SenseDroid:

- **Mobile Phone Sensing**: SenseDroid enables and provides data capture from different sensors on (or attached to) mobile phones by providing configurable sensing probes. The user can configure the sensing probes and sampling techniques through a sensing API.
- **Context Determination, Analysis & Processing**: SenseDroid enables the use of the sensed information to determine high level features such as user activities, physiological parameters, events, and their correlations. The shared sensing and context are used to determine group context, behavior, and preferences.
- **Communication and Collaboration**: SenseDroid provides libraries and APIs for communication, service discovery, and collaboration among mobile phones for different network topologies (e.g. client-server and peer-to-peer).
- **Data Logging and Retrieval**: SenseDroid provides data management routines and interface to a light weight database such as SQLite for data logging and efficient sensor data processing and storing.
- **Query and Filtering**: SenseDroid supports on-demand query and filtering functionality from different participating users. Filtering helps deliver only the relevant information to collaborating users.

We have developed and implemented an initial prototype application for Android smartphones to test much of the above functionalities. In general, the architecture and design features can be ported to other platforms. SenseDroid is distributed across the different levels of the hierarchical architecture described above.

Fig. 2 shows the details of the broker functionalities and the mobile node middleware and application components necessary in an NC to supporting collaboration and compressive sensing over proposed hierarchical network. Unlike a traditional WSN, the mobile NC supports bidirectional data flow between the nodes and the broker using multiple networks like WiFi, GSM, bluetooth etc. that enables dissemination of collective information and collaboration among the mobile nodes through the broker. However, in case of the compressing sensing approach, the broker performs stochastic (random) spatial sampling in various nodes. If N mobile sensors nodes are uniformly and randomly distributed in a NC such that the compressive sample of given sparsity requires M random measurements from these N nodes, the broker initiates these measurements by commanding and telemetering the selected nodes with the sensor. The broker can also use measurement from infrastructure sensors in absence of either enough sensor in the mobile modes or to off-load the burden of sensing cost from the mobile nodes.

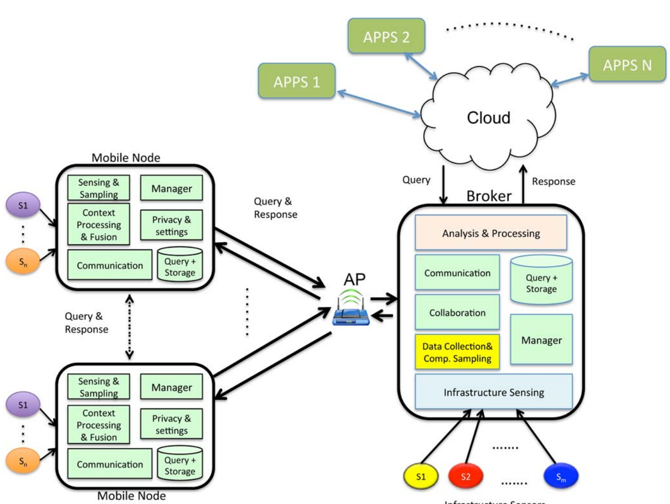

Figure 2. Basic components of the NanoCloud: mobile nodes with thin client and the broker orchestrating between other neighbors.

The mobile phones /nodes in the NCs provide unique sensing abilities and capabilities for both physical sensors as well as computationally enabled virtual sensors [23] as shown in Fig. 3. The SenseDroid framework provides individual probes for available physical sensors along with their configurable measurement parameters such as sampling rate, duration etc. and fuse these physical sensors measurements to construct more meaningful sensors (e.g. orientation, compass and inclinometer sensors in Fig. 3). In similar ways, indirect sensing by computational means can be used to derive computationally enabled virtual sensors such as situation specific contexts pertaining to user location, activity, environment, health, emotions, and social scenarios. Thus, in addition to basic physical sensor probes [1], SenseDroid provides several virtual sensing probes corresponding to different types of contextual informations. Moreover, unlike some of the recent works in mobile context processing [11, 12, 17], SenseDroid employs compressive sensing in the temporal dimension to exploit the temporal correlation in the sensor measurements to achieve energy efficient contexts determinations. As an example, we use compressive sampling instead of continuous uniform measurement of the GPS and WiFi to derive the 'IsIndoor' flag with similar accuracy while saving energy consumptions. This 'IsIndoor' flag spatial field can be used, for instance, during an earthquake to assess the potential dangers to human life. Thus, unlike WSN and the work in [13], due to the considerable computational and programable resources in mobile phone, the use of configurable compressive sensing at each node enables the unique ability to jointly perform spatio-temporal compressive sensing of both physical and virtual sensors in the proposed framework. As an example, Fig. 4 shows the reconstruction accuracy of a accelerometer signal of 256 samples from just 30 random samples in determining the 'IsDriving' context of the mobile node. When the same is applied using the spatial compressive sensing over a region, can provide indications to the traffic situations. We explain the details of spatio-temporal compressive sensing approach and its fundamentals in the next section.

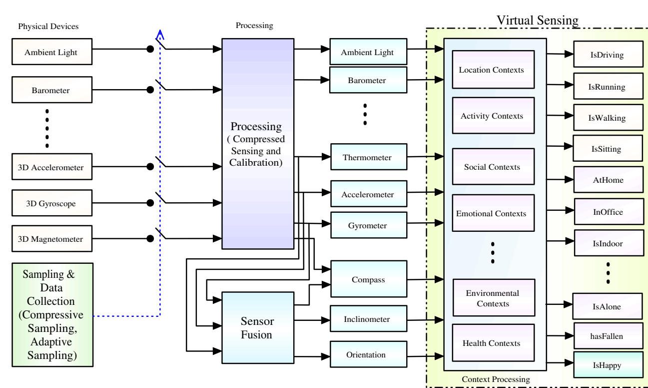

Figure 3. SenseDroid provides sensing probes for several physical sensors and ability to construct virtual sensors using and compressive context processing. The additional virtual sensing abilities and probes provide unique opportunities for collaborative applications.

4. Collaborative Compressive Sensing for Hierarchical Sense-making

SenseDroid framework supports multi-tiered data aggregation of spatio-temporal sparse fields and its reconstruction. A sparse signal is a signal that can be represented with a small number of nonzero coefficients and hence can contain most of its salient information in a relatively small number of random projections. It follows that if a signal is compressible in some orthonormal basis, then a very accurate reconstruction can be obtained from random measurement and projections. This sampling approach is used in a hierarchical architecture to show that signals and spatio-temporal sparse fields can be accurately recovered from fewer random measurements and projections contaminated with noise. In order too achieve this, the total spatial filed area is subdivided into zones and each zone is covered by the mobile local cloud (LCs). The total spatial field is then the sum of the

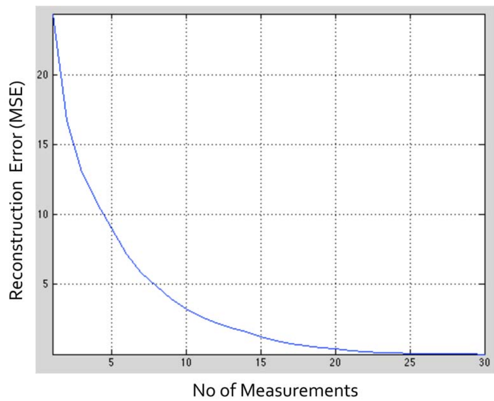

Figure 4. Accuracy of reconstruction as a function of number of measurements. As the number of measurements (or compression ratio) increases, the reconstruction error is reduced.

all the subfields computed and processed by the local cloud. The multi-tired hierarchical structured architecture enables compressive processing as different levels of granularity and accuracy. Increased emphasis, attention and resources can be directed to the areas of most impact and effects.

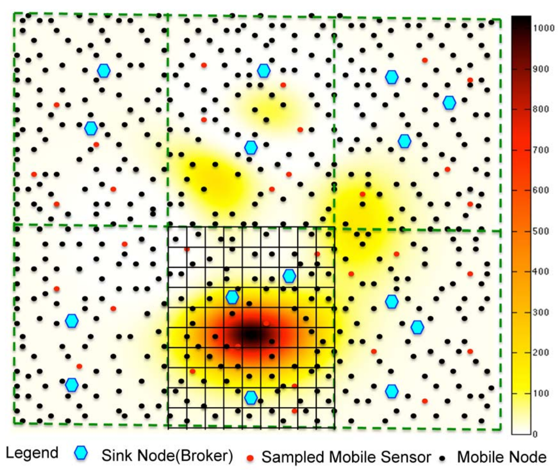

Figure 5. Distributed collaborative compressive mobile sensing of spatio-temporal sparse fields. Based on the type of sensing field, the signal sparsity, accuracy requirement, the middleware broker decides the compression ratio during data aggregation in each zone.

Let $f(i,j)$ be the two-dimensional spatial field map where i and j represents the coordinate of the location within the two-dimensional spatial field of a zone as in Fig. 5. When the spatial field map is discretized to $f[i,j]$, the coordinates $i \in \{1,2,..,W\}$ and $j \in \{1,2,..,H\}$ where W and H are the width and hight of the discretized spatial field map respectively. If we consider prior available data of a LC – a set of T spatial fields $\mathrm{F} = \{f_1[i,j], f_2[i,j], ..., f_T[i,j]\}$ taken at time instants $t_1, t_2, .., t_T$, these can be used to improve sensing by exploiting local correlation during reconstruction. Specifically, to monitor a discrete spatial field map that has $N = WH$ variable parameters (and hence N-dimensional), the number of sensors needed is less than equal to number of unknown variable or parameters i.e. $\#Sensors \leq \#Parameters$. By transforming the data, a low dimensional representation is possible with a very less error. In the transformed domain, the spatial field map can be represented using few principal components/parameters such that $K \ll N$ (with the assumption of sparsity [2, 5]).

Let the two-dimension spatial field map $f[i,j]$ can be represented using a one-dimensional vector $x[k]$ where $1 \leq k \leq N$ and $N = WH$ such that

$$x[k] = f\left[k \, mod \, H, floor\left[\frac{k}{W}\right]\right]. \tag{1}$$

In other words, stack the columns of the two-dimensional map to transform into a vector where N is the total no of grid points and $x[k]$ represent the sensor measurement at k-th grid point. The set of one dimensional spatial field traces $\Gamma = \{\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_T}\}$ can thus be represented as matrix $\mathbf{X}$ of size $T \times N$ with each row indicating a trace $\mathbf{x}$. In many scenarios , in spite of the sparsity of the field, the full spatial field $x[k], k = 1..N$ at all N locations is not available . Only M samples obtained from M spatial field sensors ($M \ll N$) located at $L = \{i_1, i_2, ..., i_M\}$ is available. Let $x(L)$ denotes the field measurement at these locations. The spatial field characterization problem is then the estimate the field at each of N points given measurement of M sensors at location L in a LC.

Any vector $\mathbf{x}$ can be represented using a basis Φ as,

$$\begin{gathered} x[k] = \sum_{n=0}^{N} \Phi[k,n]\alpha[n], \\ \mathbf{x} = \mathbf{\Phi}\alpha \end{gathered} \tag{2}$$

where $\alpha[n]$ are the coefficients of the expansion over the basis Φ. Once we define a basis Φ for the data, knowing the coefficients α is equivalent to knowing the spatial field map $\mathbf{x}$. The basis Φ is often selected as transformation matrix of FFT or DCT.

If the signal is k-sparse in the transformed domain, then the locations of the non-zero K coefficients in α can be denoted by $J = \{j_1, j_2, ..., j_K\}$. If M sensors are available at location $L = \{i_1, i_2, ..., i_M\}$, then the sensor samples in (2) can be represented as:

$$\begin{aligned} x[i_1] &= \Phi[i_1, j_1]\alpha[j_1] + \cdots + \Phi[i_1, j_K]\alpha[j_K] \\ &\vdots = \vdots \\ x[i_M] &= \Phi[i_M, j_1]\alpha[j_1] + \cdots + \Phi[i_M, j_K]\alpha[j_K] \end{aligned} \tag{3}$$

$$x(L) = \Phi(L, J)\alpha(J). \tag{4}$$

We can approximate spatial field map with $\hat{x}$ with a linear combination of K columns of Φ and K elements of α out of N as

$$\hat{\mathbf{x}} = \begin{bmatrix} \Phi[1,1] & \cdots & \Phi[1,K] \\ \vdots & \ddots & \vdots \\ \Phi[N,1] & \cdots & \Phi[N,K] \end{bmatrix} \begin{bmatrix} \alpha[1] \\ \vdots \\ \alpha[K] \end{bmatrix} = \mathbf{\Phi_K}\alpha_\mathbf{K} \tag{5}$$

where the subscript Kindicates the selection of K columns**.** The spatial field map is now defined by only K coefficients of α_K in the basis Φ_K given by:

$$\alpha_\mathbf{K} = \mathbf{\Phi_K^\dagger}\hat{\mathbf{x}} = \left[(\mathbf{\Phi_K^*}\mathbf{\Phi_K})^{-1}\,\mathbf{\Phi_K^*}\right]\hat{\mathbf{x}} \tag{6}$$

where $\Phi_K^\dagger$ is the pseudo inverse for the overdetermined system in (5). To solve (6) , we need the knowledge of the field $x[k]$ at every location $k = 1 \cdots N$. For M available sensors at location $L = \{i_1, i_2, ..., i_M\}$, we can represent $x(L)$ as:

$$\mathbf{x_S} = \begin{bmatrix} \Phi[i_1,1] & \cdots & \Phi[i_1,K] \\ \vdots & \ddots & \vdots \\ \Phi[i_M,1] & \cdots & \Phi[i_M,K] \end{bmatrix} \begin{bmatrix} \alpha[1] \\ \vdots \\ \alpha[K] \end{bmatrix} = \tilde{\mathbf{\Phi}}_\mathbf{K}\alpha_\mathbf{K}, \tag{7}$$

where $\tilde{\Phi}_K$ is a matrix formed from the rows of $\mathbf{\Phi_K}$ corresponding to the sensor location L, $\mathbf{x_S}$ is the sensor measurement (known). The coefficients $\alpha_\mathbf{K}$ are unknown in the system in (7) and can be solved for two cases: (a) underdetermined case, $M < K$ (b) overdetermined case , $M \geq K$. For the underdetermined case $M < K$ solution with infinite solutions, the coefficients are assumed to be sparse and only non-zero significant coefficients are selected. The solution of $\alpha_\mathbf{K}$ can be uniquely determined by solving the following optimization [2, 5]:

$$\begin{gathered} \underset{\alpha_\mathbf{K}}{Minimize} \quad \| \alpha_\mathbf{K} \|_0 \\ subject\,to\ \mathbf{x_S} = \tilde{\mathbf{\Phi}}_\mathbf{K}\alpha_\mathbf{K} \end{gathered} \tag{8}$$

where $\| \bullet \|_0$ stands for $\mathcal{L}_0 - norm$ of a vector i.e. the number of non-zeros in the vector. Optimization in (8) attempts to minimize the number of non-zeros in $\alpha_{\mathbf{K}}$ while satisfying the constraint for a unique solution that is as sparse as possible. The optimization problem (8) is NP-hard and hence is extremely difficult to solve [2, 5]. A more efficient technique to find a sparse solution is based on $\mathcal{L}_1 - norm$ regularization -a relaxed version of $\mathcal{L}_0 - norm$ [2, 5]:

$$\begin{array}{ll} \underset{\alpha_{\mathbf{K}}}{Minimize} & \| \alpha_{\mathbf{K}} \|_1 \\ subject\, to\; \mathbf{x_S} = \tilde{\mathbf{\Phi}}_{\mathbf{K}} \alpha_{\mathbf{K}} & \end{array} \tag{9}$$

where $\| \bullet \|_1$ stands for $L_1 - norm$ of a vector i.e., the summation of the absolute value of all elements in the vector. The $\mathcal{L}_1 - norm$ regularization in (9) can be re-formulated as an Linear Programing (LP) problem and solved efficiently. As the cost function $\| \alpha_K \|_1$ is not smooth, linear programing can not be directly applied. To apply LP, slack variables are introduced $\{\theta_i : i = 1, 2, .., K\}$ such that $-\theta_i \leq \alpha_i \leq \theta_i$ and (9) can be re-written as [2]:

$$\begin{array}{ll} \underset{\alpha_{\mathbf{K}}, \theta}{Minimize} & \theta_1 + \theta_2 + \cdots + \theta_K \\ subject\, to\; \mathbf{x_S} = \tilde{\mathbf{\Phi}}_{\mathbf{K}} \alpha_{\mathbf{K}} & \\ -\theta_i \leq \alpha_i \leq \theta_i & \end{array} \tag{10}$$

Intuitively, by minimizing the cost function in (10), all the constraints become active i.e. $| \alpha_i | = \theta_i$ and hence (9) and(10) is equivalent. Roughly speaking, if the N-dimensional vector $\alpha_{\mathbf{K}}$ contains K non-zeros and the linear equation $\mathbf{x} = \mathbf{\Phi} \alpha_{\mathbf{K}}$ is well-conditioned, the solution α_K can be almost uniquely determined (with a probability nearly equal to 1) from M sampling points, where M is in the order of $O(K * log(N))$ [2, 5, 26]. Note that M (the number of sensors or measurements) is a logarithmic function of N (the number of unknown parameters).

A convenient closed form solution of $\alpha_{\mathbf{K}}$ based on ordinary least square estimate (OLS) for well-conditioned overdetermined case [i.e. $M \geq K$ and $rank(\tilde{\Phi}_K) = K$] is given by :

$$\alpha_{\mathbf{K}} = \tilde{\mathbf{\Phi}}_{\mathbf{K}}^{\dagger} \mathbf{x_S} = \left[\left(\tilde{\mathbf{\Phi}}_{\mathbf{K}}^{*} \tilde{\mathbf{\Phi}}_{\mathbf{K}} \right)^{-1} \tilde{\mathbf{\Phi}}_{\mathbf{K}}^{*} \right] \mathbf{x_S}. \tag{11}$$

On the other hand, a generalized least square (GLS) solution considering sensor heterogeneity and noisy measurement $\mathbf{x_S} + \mathbf{w}$ where the noise $\mathbf{w}$ is having a distribution with covariance $\mathbf{V}$ and mean $\boldsymbol{\mu}$ i.e. $\mathbf{w} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{V})$ results in:

$$\boldsymbol{\alpha_K} = \tilde{\mathbf{\Phi}}_{\boldsymbol{K}}^{\dagger} \boldsymbol{x_S} = \left[\left(\tilde{\mathbf{\Phi}}_{\boldsymbol{K}}^{*} \boldsymbol{V}^{-1} \tilde{\mathbf{\Phi}}_{\boldsymbol{K}} \right)^{-1} \tilde{\mathbf{\Phi}}_{\boldsymbol{K}}^{*} \boldsymbol{V}^{-1} \right] \boldsymbol{x_S} \tag{12}$$

The same problem can be re-written for full spatial field reconstruction for M sensors located at $L = \{i_1, i_2, ..., i_M\}$ as a sparse regression problem, which can be formulated as the following optimization [27]:

$$\begin{array}{ll} \underset{\alpha}{Minimize} & \| \mathbf{x} - \mathbf{\Phi} \alpha \|_2^2 \\ Subject\, to & : \| \alpha \|_0 \leq K \end{array} . \tag{13}$$

The optimization problem in (13) can be effectively solved using the orthogonal matching pursuit (OMP) algorithm [27].The solution to (13) is affected by the sensor noise and errors, which in turn effects reconstruction accuracy. Total error ϵ introduced in the reconstruction is sum of coefficient truncations or approximation error, ϵ_a , error due to numerical ill-conditioning, ϵ_c, and measurement noise error, ϵ_m . Note that, once we have fixed M, increasing K will in general increase the reconstruction error ϵ_c (worse conditioning) and decrease the approximation error ϵ_a (better approximation). Therefore, we should pick an optimal K such that the sum ϵ is minimal. In presence of sensor noise and errors $\mathbf{w}$, sensor measurement in (7) becomes

$$\mathbf{x_s} + \mathbf{w} = \tilde{\mathbf{\Phi}}_{\mathbf{K}} \alpha_{\mathbf{K}}. \tag{14}$$

There is no exact solution to (14) for α_K but a least square estimation (LSE) problem is solved such that the error w.r.t. the measured field is minimized using the algorithm in Fig. 6. The algorithm is primarily implemented in the brokers but is also used by the nodes for context processing.

Algorithm: Compressive Heterogeneous Sensing

Input: measured vector $\mathbf{x_S}$, target sparsity M(no of sensor measurement), and size of full signal N
Output: Coefficient Indices J, Sensing matrix $\tilde{\Phi}_K$, reconstructed signal $\hat{x}$

---

1. Initialize coefficient indexes $J = \emptyset$, residual $\mathbf{e}_r = \mathbf{x_S}$, $\alpha_K = \emptyset$
2. Formation of the basis matrix Φ
3. **While** stop criteria not met **do**
   (a) $\mathbf{e_r^{new}} = \Upsilon(\mathbf{e_r})$, where Υ is an interpolation function such that $\Upsilon : \mathbb{R}^K \to \mathbb{R}^N$
   (b) $\alpha_{\mathbf{r}} = \Phi^{\dagger} \mathbf{e}_r^{new}$
   (c) Choose a subset of coefficient indices $I \subset J$ from $\alpha_{\mathbf{r}}$ for deciding the significant coefficients
   (d) Update coefficient index set $J = J \cup I$
   (e) Find the coefficients using Ordinary Least Square (OLS) or Generalized Least Square(GLS):
      i. OLS Solution for homogenous sensors: $\alpha_{\mathbf{K}} = \tilde{\mathbf{\Phi}}_{\mathbf{K}}^{\dagger} \mathbf{x_S} = \left[\left(\tilde{\mathbf{\Phi}}_{\mathbf{K}}^{*} \tilde{\mathbf{\Phi}}_{\mathbf{K}} \right)^{-1} \tilde{\mathbf{\Phi}}_{\mathbf{K}}^{*} \right] \mathbf{x_S}$
      ii. GLS Solution for heterogenous sensors: $\alpha_{\mathbf{K}} = \tilde{\mathbf{\Phi}}_{\mathbf{K}}^{\dagger} \mathbf{x_S} = \left[\left(\tilde{\mathbf{\Phi}}_{\mathbf{K}}^{*} \mathbf{V}^{-1} \tilde{\mathbf{\Phi}}_{\mathbf{K}} \right)^{-1} \tilde{\mathbf{\Phi}}_{\mathbf{K}}^{*} \mathbf{V}^{-1} \right] \mathbf{x_S}$
      where $\mathbf{V}$ is covariance matrix of sensor accuracy characteristics.
   (f) $\mathbf{e_r} = \mathbf{x_S} - \tilde{\mathbf{\Phi}}_{\mathbf{K}} \alpha_{\mathbf{K}}$

   end while
4. Reconstructed signal $\hat{x} = \Phi_K \alpha_K$

Figure 6. Collaborative Compressive Sensing in Mobile Networks.

5. Concluding Remarks and Future Research Directions

In this paper we proposed a scalable and collaborative mobile crowdsensing framework using a hierarchical distributed middleware for efficient collective understanding of users, contexts, and their environments. To the best of our knowledge, a scalable spatio-temporal compressive and collaborative mobile sensing framework for multi-tired hierarchical network configurations specifically supporting physical and virtual sensing has not been considered. In order to support emerging mobile collaborative applications, we propose the SenseDroid framework and implemented a prototype to demonstrate the features.

There are several challenges that need to addressed and we pursue them as our future research directions. Some of these research direction are briefly explained below:

Energy Efficiency : Power consumption and energy efficiency are serious issues in mobile sensing, as continuous monitoring can largely drain the battery in a short period of time. Energy-efficiency issues have been studied in the context of mobile phone sensing recently in [16, 19, 20, 24, 30]. In [24], the authors showed that collaborative sensing can achieve over 80% power savings compared to traditional sensing without collaborations. Research in the direction of sensor scheduling, adaptive sampling, and compressive sampling and their novel combinations within the framework provide new research opportunities for of energy-efficiency.

Privacy Regulation: Privacy is a major concern in collaborative sensing application. To address this we have adopted transparency, full user control, and encryption of the data that is shared. User can fully set or control their preferences, enable or disable features, control of the type of sensors and parameter that can be shared in the application to avoid any violations. In the worst case, the user can opt-out of any such applications. Research addressing privacy in mobile phone sensing is in its infancy with initial research work proposing privacy preserving incentive mechanisms [10, 29].

Incentive Mechanisms: Incentive mechanism to motivate participation and collaboration is an important aspect that need to researched to bring desirable economic properties and appropriate utility in the collaboration framework. Few recent studies in this direction looked into selecting well-suited participants for sensing services within recruitment frameworks [21], sealed-bid second-price auction to motivate user participation [4], and reverse auction based dynamic price incentive mechanisms [9]. A comparative study of different incentive mechanisms for a client to motivate the collaboration of smartphone users on both data acquisition and distributed computing applications is evaluated in [6].

Heterogeneity in Mobile Cloud: Heterogeneity manifests in several aspects (e.g. networks, sensor types, sensor accuracy) of the collaborative sensing. Our present development in primarily focused on using Wi-Fi based mobile ad-hoc networks for localize environments specifically for the local cloud. However, support for more power efficient networks like Bluetooth can considered to support the nanocloud architecture. Handling the heterogeneity in network architectures, mobile devices and sensors, as well as services is challenging area of research and need to be further studied in the context of compressive sensing [22].

Acknowledgment

This work was partially supported by the NSF award CCF-1029783 (Variability Expedition), CNS-1063596 (Cypress), and CRI-1059436 (Irvine Sensorium).

References

[1] Nadav Aharony, Wei Pan, Cory Ip, Inas Khayal, and Alex Pentland. Social fmri: Investigating and shaping social mechanisms in the real world. *Pervasive and Mobile Computing*, 7(6):643–659, 2011.

[2] E.J. Candes and M.B. Wakin. An introduction to compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):21–30, 2008.

[3] Sunny Consolvo et al. Activity sensing in the wild: a field trial of ubifit garden. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1797–1806. ACM, 2008.

[4] George Danezis, Stephen Lewis, and Ross Anderson. How much is location privacy worth? In *Online Proceedings of the Workshop on the Economics of Information Security Series (WEIS 2005*. Citeseer, 2005.

[5] D.L. Donoho. Compressed sensing. *Information Theory, IEEE Transactions on*, 52(4):1289–1306, 2006.

[6] Lingjie Duan et al. Incentive mechanisms for smartphone collaboration in data acquisition and distributed computing. In *INFOCOM, 2012 Proceedings IEEE*, pages 1701–1709, 2012.

[7] W.Z. Khan, Yang Xiang, M.Y. Aalsalem, and Q. Arshad. Mobile phone sensing systems: A survey. *Communications Surveys Tutorials, IEEE*, 15(1):402–427, 2013.

[8] N.D. Lane, E. Miluzzo, Hong Lu, D. Peebles, T. Choudhury, and A.T. Campbell. A survey of mobile phone sensing. *Communications Magazine, IEEE*, 48(9):140–150, 2010.

[9] Juong-Sik Lee and Baik Hoh. Sell your experiences: a market mechanism based incentive for participatory sensing. In *Pervasive Computing and Communications (PerCom), 2010 IEEE International Conference on*, pages 60–68, 2010.

[10] Qinghua Li and Guohong Cao. Providing privacy-aware incentives for mobile sensing. In *IEEE International Conference on Pervasive Computing and Communications (PerCom)*, volume 18, page 22, 2013.

[11] Robert LiKamWa et al. Moodscope: building a mood sensor from smartphone usage patterns. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*, MobiSys '13, pages 389–402, New York, NY, USA, 2013. ACM.

[12] Hong Lu et al. Stresssense: detecting stress in unconstrained acoustic environments using smartphones. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, pages 351–360, New York, NY, USA, 2012. ACM.

[13] Chong Luo, Feng Wu, Jun Sun, and Chang Wen Chen. Compressive data gathering for large-scale wireless sensor networks. In *Proceedings of the 15th annual international conference on Mobile computing and networking*, pages 145–156. ACM, 2009.

[14] Jun Luo, Liu Xiang, and Catherine Rosenberg. Does compressed sensing improve the throughput of wireless sensor networks? In *Communications (ICC), 2010 IEEE International Conference on*, pages 1–6. IEEE, 2010.

[15] S. Madhani, M. Tauil, and Tao Zhang. Collaborative sensing using uncontrolled mobile devices. In *Collaborative Computing: Networking, Applications and Worksharing, 2005 International Conference on*, pages 8 pp.–, 2005.

[16] Mirco Musolesi et al. Supporting energy-efficient uploading strategies for continuous sensing applications on mobile phones. In *Pervasive Computing*, pages 355–372. Springer, 2010.

[17] S. Nath. Ace: Exploiting correlation for energy-efficient and continuous context sensing. *Mobile Computing, IEEE Transactions on*, 12(8):1472–1486, 2013.

[18] NBA. Neea study on luminarie level lighting controls (lllc) -enlighted technical proof of concept study. Technical report, Northwest Energy Efficiency Alliance, July, 2013.

[19] Jeongyeup Paek, Joongheon Kim, and Ramesh Govindan. Energy-efficient rate-adaptive gps-based positioning for smartphones. In *Proceedings of the 8th international conference on Mobile systems, applications, and services*, MobiSys '10, pages 299–314, New York, NY, USA, 2010. ACM.

[20] M. Rahman, J. Gao, and Wei-Tek Tsai. Energy saving in mobile cloud computing*. In *Cloud Engineering (IC2E), 2013 IEEE International Conference on*, pages 285–291, 2013.

[21] Sasank Reddy, Deborah Estrin, and Mani Srivastava. Recruitment framework for participatory sensing data collections. In *Proceedings of the 8th international conference on Pervasive Computing*, Pervasive'10, pages 138–155, Berlin, Heidelberg, 2010. Springer-Verlag.

[22] Z. Sanaei, S. Abolfazli, A. Gani, and R. Buyya. Heterogeneity in mobile cloud computing: Taxonomy and open challenges, 2013.

[23] Santanu Sarma, Nikil Dutt, and Nalini Venkatasubramanian. Cross-layer virtual observers for embedded multiprocessor system-on-chip (mpsoc). In *Proceedings of the 11th International Workshop on Adaptive and Reflective Middleware*, ARM '12, pages 4:1–4:7, New York, NY, USA, 2012. ACM.

[24] Xiang Sheng, Jian Tang, and Weiyi Zhang. Energy-efficient collaborative sensing with mobile phones. In *INFOCOM, 2012 Proceedings IEEE*, pages 1916–1924. IEEE, 2012.

[25] Niwat Thepvilojanapong et al. Opportunistic collaboration in participatory sensing environments. In *Proceedings of the fifth ACM international workshop on Mobility in the evolving internet architecture*, MobiArch '10, pages 39–44, New York, NY, USA, 2010. ACM.

[26] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*, 58:267–288, 1994.

[27] J.A. Tropp et al. Signal recovery from random measurements via orthogonal matching pursuit. *Information Theory, IEEE Trans. on*, 53(12):4655–4666, 2007.

[28] H. Weinschrott, F. Durr, and K. Rothermel. Streamshaper: Coordination algorithms for participatory mobile urban sensing. In *Mobile Adhoc and Sensor Systems (MASS), 2010 IEEE 7th International Conference on*, pages 195–204, 2010.

[29] Dejun Yang, Guoliang Xue, Xi Fang, and Jian Tang. Crowdsourcing to smartphones: incentive mechanism design for mobile phone sensing. In *Proceedings of the 18th annual international conference on Mobile computing and networking*, pages 173–184. ACM, 2012.

[30] Weiwen Zhang, Yonggang Wen, and D.O. Wu. Energy-efficient scheduling policy for collaborative execution in mobile cloud computing. In *INFOCOM, 2013 Proceedings IEEE*, pages 190–194, 2013.

[31] Shu Zhou, Min-You Wu, and Wei Shu. Finding optimal placements for mobile sensors: wireless sensor network topology adjustment. In *Emerging Technologies: Frontiers of Mobile and Wireless Communication, 2004. Proceedings of the IEEE 6th Circuits and Systems Symposium on*, volume 2, pages 529–532 Vol.2, 2004.