Opportunistic privacy preserving monitoring*

Luca Becchetti
Sapienza University of Rome
Via Ariosto 25
Rome, Italy
becchetti@dis.uniroma1.it

Luca Filipponi
Sapienza University of Rome
Via Ariosto 25
Rome, Italy
filipponi@dis.uniroma1.it

Andrea Vitaletti
Sapienza University of Rome
Via Ariosto 25
Rome, Italy
vitale@dis.uniroma1.it

ABSTRACT
In this paper we present an approach to support privacy in opportunistic sensing. In particular, we use compact and privacy preserving representations of data (or sketches), that allow us to compute relevant statistics over data without disclosing users' sensitive information (e.g. locations). We exemplify our approach by referring to the important application of noise pollution monitoring. We present preliminary experimental results showing that sketches can actually be employed to produce accurate environmental maps, at the same time preserving users' privacy.

1. INTRODUCTION
Opportunistic people-centric sensing [12] has been gaining popularity, with several systems and applications being proposed to leverage users' mobile devices to collectively measure environmental data. In such systems, nodes report sensor data through opportunistic network connections, such as third-party access points. However, people centric sensing generally suffers from privacy related issues, namely the need to share data provided by users without disclosing any sensitive information about user's privacy (e.g., locations).

In this paper we present an approach to support privacy in opportunistic sensing. In particular we use sketches, namely compact and privacy preserving representations of data, that allow us to compute relevant statistics over the data without disclosing users' sensitive information. The proposed techniques can be exploited in several application domains such as environmental monitoring, analysis of social patterns, traffic maps etc. Here we exemplify our approach referring to the relevant application of noise pollution monitoring.

The Directive 2002/49/EC of the European Parliament has made the avoidance, prevention, and reduction of environmental noise a primary issue in European policy and the Commission required Member States to provide accurate noise pollution maps. Today's noise measurements are mainly carried out by designated officers that collect data in a location of interest. Even if this assessment procedure is still compliant with European regulations [9], it often fails to provide scalable and accurate estimations of the real noise pollution levels. Nowadays, the applicability of fixed wireless sensor networks [10] to wide area long-term monitoring, is still limited due to its high installation and maintenance costs.

*Partially supported by EU STREP Projects ICT-215270 FRONTS, ICT-257245 VITRO and by PRIN 2008 research project COGENT.

2. RELATED WORK
People-centric sensing [7][8] has leveraged the use of human carried devices (such as smart-phones) to sense information directly or indirectly related to human activity or environment, in an opportunistic or participatory way. The MetroSense project [14] is working with industry and agencies to develop new applications, classification techniques, privacy approaches, and sensing paradigms for mobile phones enabling a global mobile sensor network capable of societal-scale sensing. Most related to our reference application, the NoiseTube [13] project has developed a novel platform for the monitoring of urban noise pollution, based on mobile phones. The same approach has been followed by the NoiseSpy [11] project in which noise maps are built on the basis of data coming from users' devices in a participatory way. Both these projects require users to agree and share informations regarding noise levels measurements, together with their position in order to allow geotagging to an external system. While both these works demonstrate the feasibility of the use of smartphones/cellphones as sound meters, their platforms suffer of a major lack of privacy for involved users, thus allowing an attacker to trace users' movements. Privacy preservation in location based services has already been addressed by [16, 15]. In [15], accurate traffic speed maps in a small campus town are build from shared GPS data of participating vehicles, where the individual vehicles are allowed to "lie" about their actual location and speed at all times. In our approach, data are always correct but represented in a compact and privacy preserving way (i.e sketches). Differently from [16], where data are available in clear to the intended receiver, in our work sketches allow a central authority to select relevant traces to reconstruct an accurate map, but without revealing to anybody (central authority included) relevant information on users' positions.

3. SYSTEM OVERVIEW
Taking inspiration from the NoiseTube [13] set-up, we aim to leverage user's smartphones to sample the environment and provide collected data to a central authority. Nevertheless contrary to NoiseTube, our solution is based on *opportunistic monitoring* and explicitly considers *privacy preservation* a primary concern to actually encourage users' participation. Nowadays mobile phones are personal devices, primarily intended to serve the user with telephony, messaging and other functionalities. Additional services like noise monitor-

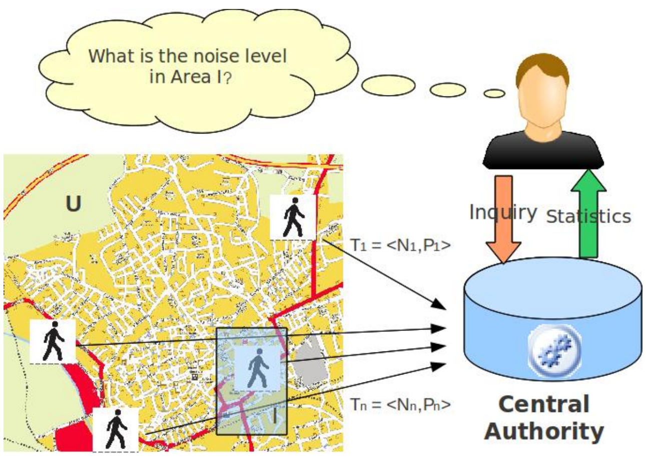


Figure 1: System Overview

ing cannot be considered of primary importance, and thus people would likely not support them if they will negatively affect the behavior of the "primary" services. For this reason, we think a mobile phone application should be transparent and communicate opportunistically whenever users get connected for their own purposes. This approach strongly limits the availability of real-time data, because data are only transmitted whenever a connection is available. However, for the statistical purposes of the noise monitoring application we are discussing here, this is not a major problem. Instead, we do believe that a mandatory requirement to achieve a significant users' participation is the preservation of their privacy. Sampled data must be geo-referenced to be of some utility, as a consequence users' movements could be easily traced with a serious loss of privacy. Thus, a privacy preserving representation of data is needed, but at the same time, operations on sampled data should be still possible to the central authority, in order to select relevant samples among the whole amount of received data for each inquiry. The noise monitoring service we envision, foresees the participation of three main actors: mobile users, central authority and system users. **Mobile users** are responsible for noise data collection. They participate to the service by running a noise monitoring application on their phones (see section 3.1). This application exploits the on board microphone to sample environmental data while the phone is idle, so as not to affect phone's normal usage. Whenever the user establishes a connection, environmental data are opportunistically sent to the **central authority** which is in charge to elaborate statistics and to answer queries issued by **system users** on the average noise level in an area of interest.

More in detail, we consider n users moving in an area U and collecting traces on environmental noise. The trace generated by user i, denoted by T_i consists of a set of pairs $< n_i^t, p_i^t >$ where n_i^t and p_i^t are respectively the observed noise level and the location of the user in U at time t. Let $P_i = \{p_i^t\}, \forall t$ and $N_i = \{n_i^t\}, \forall t$, be respectively the set of positions (i's *position set*) occupied by the user over time and the corresponding set of noise levels (i's *noise set*). We stress that, in order to guarantee users' privacy, P_i should not be disclosed to third parties, including the central authority, and thus it should be represented in a suitable privacy preserving format. Instead, the set of noise values must be publicly available, since it is used for the estimation of the average noise level. To achieve this, user i sends to the central authority the pair $(\mathbf{Sk}(P_i), N_i)$, where $\mathbf{Sk}(P_i)$ is a *sketch* (i.e., a suitably generated compact summary) of P_i. In Subsection 3.2 we show how to generate $\mathbf{Sk}(P_i)$ so that it has the following properties: i) $\mathbf{Sk}(P_i)$ represents P_i implicitly in small space (in the order of 10^2 bytes at most) and does not allow to infer P_i; ii) considered any area I of interest, $\mathbf{Sk}(P_i)$ allows the central authority to estimate the extent to which the set P_i of i's positions covers I. Note that this is achieved using only $\mathbf{Sk}(P_i)$, so that P_i is never explicitly disclosed to third parties. All users' traces $\mathcal{T} = T_1, T_2, .., T_n$ are made available to the central authority according to the mechanism described above. This, in turn collects the traces and performs elaboration to answer queries issued by system users concerning the average noise level in any area $I \subseteq U$ of interest. Let's denote by $\mathcal{Q}(I)$ the query asking for the average noise in area I. Upon reception of $\mathcal{Q}(I)$, the central authority selects the minimum number of P_i's guaranteeing coverage of I and then calculates the average noise level over time in that area. This is the classic set cover problem , but contrary to classic set cover we have to enforce some kind of privacy preserving technique to encourage users participation to the monitoring activity. In order to select the minimal subset of user traces that cover I, the central authority uses the techniques described in Section 4. In performing this computation, it only uses the sketches $\mathbf{Sk}(P_i)$ of users' position sets and not the position sets themselves.

3.1 Mobile phones as sound meters

The feasibility of using Mobile phones as sound meters has already been discussed in [13, 11]. Both these works, presented a mobile phone application that, requiring user participation, logs sound pressure values using the onboard microphone and some correction algorithms, thus obtaining a limited estimation error. Shifting from participatory to opportunistic monitoring, additional challenges have to be discussed. In opportunistic contexts indeed, sampled values can be affected by errors generated by external and unpredictable noise sources. As an example, consider a user carrying the mobile phone in her pocket: noise values could suffer from attenuation effects or spikes due to the noise generated by objects in the pocket (e.g., coins, keychains, voices). In such a scenario, an additional filtering technique is needed in order to obtain more accurate data. We defer the study of these aspects to future work, and here we focus on the evaluation of our system from an algorithmic point of view.

3.2 Privacy preserving data representation

As discussed before, in the application we envision, a user only sends a compact summary of her position set, from which it is hard to recover the original set. In this section we present a class of sketches [6, 5, 4] that, while compact and addressing the privacy issues mentioned above, allow the (approximate) implementation of some basic primitives on sets (such as union and intersection) that are required to implement the algorithms presented in section 4. In the rest of this subsection we present techniques used by mobile users' terminals to produce compact summaries of their respective position sets.

Compact representation of sets: We only briefly outline the principles underlying the technique we propose, leaving out many theoretical aspects for the sake of brevity. The interested reader can refer to [6, 5, 4]. In the remainder of this subsection, we consider without loss of generality subsets of $[n] = \{0, \ldots, n-1\}$, for a suitable integer n. We briefly note that standard techniques allow us to reduce to this situation in all practical cases[1].

Assume we have a family $\mathcal{H}$ of hash functions such that: i) every $h \in \mathcal{H}$ produces a permutation of $[n]$; ii) if h is chosen uniformly at random from $\mathcal{H}$ the following holds: for every set $S \subseteq [n]$:

$$\mathbf{P}[x = \arg\min(\pi(S))] = 1/|S|, \forall x \in S.$$

Such a family is said *minwise independent* [4]. In practice, minwise independent hash functions are hard to generate, since they require a high number of truly random bits. In this paper, we use functions of the form $h(x) = ((ax+b) \mod c) \mod n$ [3], that excellently approximate minwise independent families. Here, c is a large prime (e.g., the well-known Mersenne prime $2^{32}-1$) and n is the number of possible locations in U. Finally, $a \in \{1, \ldots, c-1\}$ and $b \in \{0, \ldots, c-1\}$.

Sketch generation and maintenance: Considered any subset S of $[n]$, we construct its sketch as follows: for m times, we choose, independently, uniformly and at random, a hash function from a minwise independent family. Let $H_i(x)$ the i-th function chosen and let $\min_i(S) = \min_{x \in S} H_i(x)$. Then $\mathbf{Sk}(S) = \{\min_i(S), \ldots, \min_m(S)\}$. In our case, we consider hash functions of the form $h(x) = ((ax+b) \mod c) \mod n$. In practice, generating such a hash function means generating a and b uniformly at random from $\{1, \ldots, c-1\}$ and $\{0, \ldots, c-1\}$ respectively.

Sketch properties: Given sets S_1 and S_2, the sketch of $S_1 \cup S_2$ can be immediately obtained from $\mathbf{Sk}(S_1)$ and $\mathbf{Sk}(S_2)$ as follows: $\mathbf{Sk}(S_1 \cup S_2) = \{M_1, \ldots, M_m\}$, where $M_i = \min\{\min_i(X_1), \min_i(X_2)\}$. Another interesting property of these sketches is that they allow to easily and accurately estimate the Jaccard coefficient of two sets, a standard measure of the similarity between sets, widely used in information retrieval. Given two subsets S_1 and S_2 of $[n]$, their Jaccard coefficient is defined as $J(L(S_1), L(S_2)) = \frac{|L(S_1) \cap L(S_2)|}{|L(S_1) \cup L(S_2)|}$. It can be shown [4] that for every S_1, $S_2 \subseteq [n]$: $\mathbf{P}[\min(\pi(S_1)) = \min(\pi(S_2))] = J(S_1, S_2)$. This suggests a simple statistical estimator of the Jaccard coefficient of two sets, which we discuss in the next paragraph. To estimate $J(S_1, S_2)$, we simply consider their sketches $\mathbf{Sk}(S_1)$ and $\mathbf{Sk}(S_2)$ and let $C_m = |\{i : \min_i(S_1) = \min_i(S_2)\}|$. Then, a simple probability argument allows to show that C_m/m is an increasingly accurate estimate of $J(S_1, S_2)$.

Compact representation of position sets: All mobile users will use the same set $H_1(\cdot), \ldots, H_m(\cdot)$ of minwise independent hash functions. These will be generated by the central authority and then sent to each mobile user once, i.e., the first time she joins the application. Note also that, in practice, the linear functions we use are represented in terms of a small set of parameters. For example, if we use 100 hash functions, each mobile user will need to receive 202 integer values (the coefficients a and b of each hash function plus c and n), for a total of less than 1 KByte, if we represent integers using 4 bytes. Then, mobile user i will generate sketches of her position sets as follows: her sketch $\mathbf{Sk}(P_i)$ is initially set to $\{0, \ldots, 0\}$. Let $\{M_1, \ldots, M_m\}$ be i's sketch at some point. If she moves to a new position p (e.g., identified by the GPS coordinates of a new base station she connects to), then $\mathbf{Sk}(P_i)$ is updated as follows: $M_j = \min\{M_j, h_j(p)\}, \forall j = 1, \ldots, m$. This sketch update corresponds to updating i's position set as follows: $P_i = P_i \cup \{p\}$. This representation of position sets would require an attacker willing to recover P_i knowing $\mathbf{Sk}(P_i)$ to generate a sketch for all the possible P_i in the world U (even with varying size) and estimate the Jaccard coefficient between them. The more is the size of U, the more an attack is unfeasible.

4. PROBLEM STATEMENT

The problem of finding the minimum number of traces covering the area I of interest for a system user, can be formulated as an instance of the $NP-complete$ set cover problem. In the classical set cover problem we are given a set I, taken from a universe U, and a collection $\mathcal{T} = T_1, T_2, .., T_n$ of subset of U. The pair $(U, \mathcal{T})$ is sometimes called a set system. The aim is to compute a sub-collection $\mathcal{T}' \subseteq \mathcal{T}$ which covers I with minimum cost, namely using the smallest number of sets in $\mathcal{T}$.

In this section we first recall the Greedy Algorithm for set cover and then, after providing a few useful remarks, we introduce our algorithm that implements the greedy set cover using sketches.

4.1 Greedy Algorithm for Set Cover with Unitary Costs

This algorithm is given in figure 2[2].

Algorithm `Standard-Greedy`

`Require:` `set system` $(\mathcal{T}, U)$

1: $C = \emptyset$ `(`C `contains identifiers of sets in set cover)`
2: $\hat{\mathcal{T}} = \mathcal{T}$
3: $\hat{S} = \arg\max_{S \in \hat{\mathcal{T}}} |S \cap (U - C)|$
4: **while** $|S \cap (U - C)| > 0$ **do**
5: $\quad \hat{\mathcal{T}} = \hat{\mathcal{T}} - \{\hat{S}\}$
6: $\quad C = C \cup \hat{S}$
7: $\quad \hat{S} = \arg\max_{S \in \hat{\mathcal{T}}} |S \cap (U - C)|$
8: **end while**
9: `return` C

Figure 2: Greedy Algorithm for Set Cover.

Since it seems hard to give the sketch of the difference of two

[1]In our case, the position set P_i of a user i is a finite set of geographical positions (e.g., GPS coordinates). As such, it can be put in correspondence with a subset of the integers using standard techniques. E.g., [4] shows how to achieve this for Web documents using Rabin's fingerprinting method.

[2]In order to make the pseudo-code more readable, we slightly abuse notation, since we regard C as both a set of sets (the set cover) in line 9 and as the union of the sets that form the cover in lines 3, 4, 6 and 7. Analogous considerations hold for Algorithm 3.

sets given the sketches of the two sets, we slightly modify the algorithm above, by replacing $|S \cap (U-C)|$ with $|(S \cup C) \cap U|$ in steps 3, 4 and 7 of Algorithm 2. Maximizing the former quantity is equivalent to maximizing the latter.The proof is trivial, and is omitted due to space constraints.

4.2 Greedy Set Cover Algorithm using sketches.

We next describe the algorithm `PP-Greedy`. This algorithm is an implementation of the standard Greedy Set Cover algorithm for the case of unitary set costs. The novelty is that it is "rephrased" in terms of operations on set sketches instead of the sets themselves. Algorithm 3 is the sketch-based counterpart of Algorithm 2.

Essentially, in lines 5 and 11, instead of considering the maximization of $|(S \cup C) \cap U|$, we choose the set S, such that the (estimated) Jaccard coefficient between $S \cup C$ and U is maximized. This, up to approximations, is the set S such that $\mathbf{Sk}(S \cup C)$ and $\mathbf{Sk}(U)$ share the largest number of equal minima, i.e., the set that maximizes $Eq(U, C \cup S)$, where $Eq(U, C \cup S) = \sum_{i=1}^{m}(\min_i(U) == \min_i(C \cup S))$, namely the number of times the minima of U and $C \cup S$ agree.

Algorithm `PP-Greedy`

`Require: Sketch` $\mathbf{Sk}(S_i)$, `for` $i = 1, \ldots, |\mathcal{T}|$, $\mathbf{Sk}(U)$

1: $E = 0$
2: $C = \emptyset$ `(`C `contains identifiers of sets in set cover)`
3: $\mathbf{Sk}(C) = \{\infty\}_{i=1,\ldots,m}$
4: $\hat{\mathcal{T}} = \mathcal{T}$
5: $\hat{S} = \arg\max_{S \in \hat{\mathcal{T}}} Eq(U, C \cup S)$
6: $\hat{E} = Eq(U, C \cup \hat{S})$
7: **while** $\hat{E} > E$ **do**
8: $\quad E = \hat{E}$
9: $\quad \hat{\mathcal{T}} = \hat{\mathcal{T}} - \{\hat{S}\}$
10: $\quad C = C \cup \hat{S}$
11: $\quad \hat{S} = \arg\max_{S \in \hat{\mathcal{T}}} Eq(U, C \cup S)$
12: **end while**
13: `return` C

Figure 3: Privacy Preserving Greedy Algorithm for Set Cover.

5. PRELIMINARY RESULTS

In this section we discuss the results of a preliminary experimental activity performed to verify the performance of PP-Greedy vs Greedy on synthetic traces.

Generating traces: Our experiments are based on mobility traces generated through the Global Mobility Simulation Framework (GMSF) [1] developed at ETH Zurich. It allows to generate mobility traces according to different models, such as Random Waypoint, Manhattan, or GIS Based. In our experiments, 10, 50 and 100 mobile users move in a square U of side length $d = 1000$ units, according to the Manhattan model; each user generates a trace made of a total of 35000 positions . We adopted the Manhattan model, because it is the most suited to describe the mobility patterns of users in a urban scenario. Moreover, in order to make the simulation more realistic for an opportunistic scenario, whenever a user stops in the simulated environment, it generates a sketch of the set of positions P_i she went through so far, and opportunistically sends it to the central authority. In other words, each user sends to the central authority a set of sketches $Sk(P_i)$; each of them representing a subset of the 35000 positions sampled by the user.Sketches are generated starting from the set of positions P_i applying the technique described in section 3.2 with $m = 100$ hash functions; this number of hash functions provides the best trade-off between accuracy and size of the sketches [2]. The resulting sketch is an array of 100 integers, with a total size of 400 bytes.

The area of interest I considered for the experiments are the crossroads in the Manhattan topology, each one centered on the diagonal of U (from the center to the top-left corner of the square) made of 200 positions.

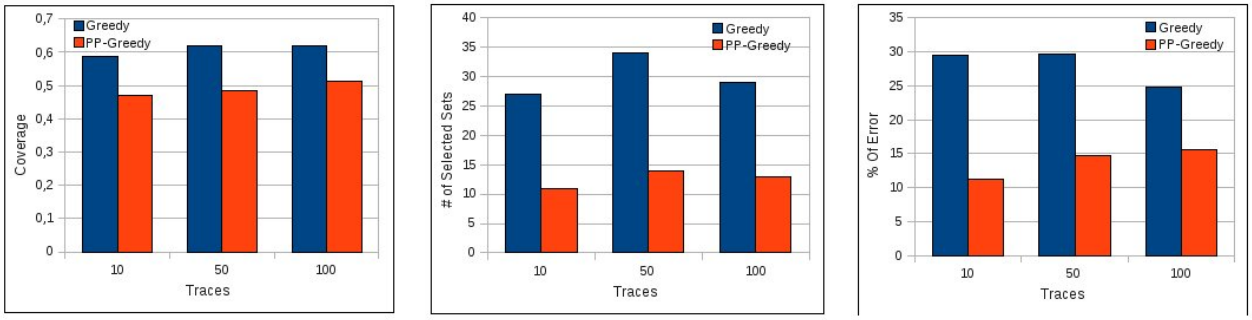
Metrics: The Greedy algorithm Gr receives in input the area of interest I and the set of sets of positions sampled by the users $P = P_1, .., P_n$, and provides as output the set $P_{Gr} \subseteq P$ that approximate the set cover of I. The PP-Greedy algorithm $PP - Gr$, instead of the set P, receives in input $Sk_P = (Sk(P_1), .., Sk(P_n))$, and provides as output the set $P_{PP-Gr} \subseteq Sk_P$ that approximate the set cover of I. We evaluate the performance of Gr and $PP-Gr$ considering three metrics. The *cardinality* of output i.e., the number of position sets used to cover the area of interest, the *coverage* of the output, intended as the fraction of positions in the area of interest that are covered by the output.[3]. The *error*, defined as the fraction of positions in the output which are not in I. As an example consider the following sets $I = 1, 3$, $P_{Gr} = 1, 2, 3, 4$. In this case the cardinality is 2, the coverage is 100% and the error is 50%.

Results: As expected and as figure 4(a) shows, the coverage achieved by both algorithms slightly increases with the number of participating mobile users (i.e. number of traces). The behaviors of both algorithms are similar, but PP-Greedy always achieves 10% less coverage than Greedy; this is the consequence of the loss of information due to the use of sketches instead of the explicit position sets. More surprisingly, the cardinality of PP-Greedy outputs is remarkably lower, approximately half the cardinality of Greedy (see figure 4(b)). The higher cardinality of the Greedy solution results in an increased error of Greedy, as can be observed in figure 4(c). This is due to the fact that each new set added to the solution contributes with a minimum number of positions . Thus, increasing the cardinality of the solution in general improves the coverage, but when the coverage is already high, each new set added to the solution will have an increasingly higher chance of contributing with new positions that do not belong to the area of interest, thus increasing error. This explains why the PP-Greedy's error is always lower than 15%, while Greedy's error is always higher than 25%. There seems to exist a "breaking-point" for the solution, beyond which the addition of more P_i's to the solution slightly increases coverage, but at the same time it significantly increases error.

6. CONCLUSIONS AND FUTURE WORK

The lower accuracy of PP-Greedy is fairly compensated by the lower error and cardinality of its outputs; as suggested

[3] We calculate the coverage of P_{PP-Gr} considering the set of corresponding positions



(a) Coverage (b) Cardinality (c) Error

Figure 4: For every area of interest and for every performance metric, we averaged the results over 10 **runs of both algorithms.**

by our results, there is an interesting trade-off between accuracy on one-side and cardinality and error on the other. This observation seems to support the conclusion that PP-Greedy algorithm is a good privacy preserving approximation of Greedy but at the same time this preliminary results deserve future investigations about the effects of tuning the granularity of P_i's with respect to the dimension of I that would directly impact on coverage and error of both algorithms, possibly reducing the actual performance differences. Moreover, tuning the number m of hash functions could outline a better trade-off between the accuracy of sketches and their size. Finally, we plan to extend our approach to different application domains.

7. REFERENCES

[1] R. Baumann, F. Legendre, and P. Sommer. Generic mobility simulation framework (gmsf). In *MobilityModels '08: Proceeding of the 1st ACM SIGMOBILE workshop on Mobility models*, pages 49–56, New York, NY, USA, 2008. ACM.

[2] L. Becchetti, P. Boldi, C. Castillo, and A. Gionis. Efficient semi-streaming algorithms for local triangle counting in massive graphs. In *KDD '08: Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 16–24, New York, NY, USA, 2008. ACM.

[3] T. Bohman, C. Cooper, and A. M. Frieze. Min-wise independent linear permutations. *The Electronic Journal of Combinatorics,*, 7, 2000.

[4] A. Z. Broder. Identifying and filtering near-duplicate documents. In *Combinatorial Pattern Matching, 11th Annual Symposium, CPM 2000, Montreal, Canada, June 21-23, 2000, Proceedings*, volume 1848 of *Lecture Notes in Computer Science*, pages 1–10. Springer, 2000.

[5] A. Z. Broder, M. Charikar, A. M. Frieze, and M. Mitzenmacher. Min-wise independent permutations. In *ACM Symposium on the Theory of Computing*, New York, NY, USA, 1998.

[6] A. Z. Broder, S. C. Glassman, M. S. Manasse, and G. Zweig. Syntactic clustering of the web. In *Proc. of the World Wide Web Conference*, 1997.

[7] A. T. Campbell, S. B. Eisenman, N. D. Lane, E. Miluzzo, R. A. Peterson, H. Lu, X. Zheng, M. Musolesi, K. Fodor, and G.-S. Ahn. The rise of people-centric sensing. *IEEE Internet Computing*, 12:12–21, 2008.

[8] C. Cornelius, A. Kapadia, D. Kotz, D. Peebles, M. Shin, and N. Triandopoulos. Anonysense: privacy-aware people-centric sensing. In *MobiSys*, pages 211–224, 2008.

[9] European Commission Working Group Assessment of Exposure to Noise (WG-AEN). Good Practice Guide for Strategic Noise Mapping and the Production of Associated Data on Noise Exposure, January 2006.

[10] L. Filipponi, S. Santini, and A. Vitaletti. Data collection in wireless sensor networks for noise pollution monitoring. In *DCOSS '08: Proceedings of the 4th IEEE international conference on Distributed Computing in Sensor Systems*, pages 492–497, Berlin, Heidelberg, 2008. Springer-Verlag.

[11] E. Kanjo. Noisespy: A real-time mobile phone platform for urban noise monitoring and mapping. *Mobile Networks and Applications*.

[12] N. D. Lane, S. B. Eisenman, M. Musolesi, E. Miluzzo, and A. T. Campbell. Urban sensing systems: Opportunistic or participatory. In *In Proc. ACM 9th Workshop on Mobile Computing Systems and Applications (HOTMOBILE '08)*, 2008.

[13] N. Maisonneuve, M. Stevens, M. E. Niessen, P. Hanappe, and L. Steels. Citizen noise pollution monitoring. In *dg.o '09: Proceedings of the 10th Annual International Conference on Digital Government Research*, pages 96–103. Digital Government Society of North America, 2009.

[14] Metrosense - secure people-centric sensing at scale. http://metrosense.cs.dartmouth.edu/.

[15] N. Pham, R. Ganti, Y. Uddin, S. Nath, and T. Abdelzaher. Privacy-preserving reconstruction of multidimensional data maps in vehicular participatory sensing. In J. Silva, B. Krishnamachari, and F. Boavida, editors, *Wireless Sensor Networks*, volume 5970 of *Lecture Notes in Computer Science*, pages 114–130. Springer Berlin / Heidelberg, 2010.

[16] S. Zhong, L. (erran Li, Y. G. Liu, and Y. R. Yang. Privacy-preserving locationbased services for mobile users in wireless networks. Technical report, 2004.