# Toward Privacy by Design in Spatial Crowdsourcing in Emergency and Disaster Response

Pouyan Fotouhi Tehrani, Hannes Restel, Michael Jendreck,
Stefan Pfennigschmidt, Markus Hardt and Ulrich Meissen
ESPRI – Electronic Safety and Security Systems for the Public and Industries
Fraunhofer-Institut für offene Kommunikationssysteme FOKUS
Berlin, Germany
{dot.separated.names}@fokus.fraunhofer.de

***Abstract*—Disaster and emergency situations require a timely and well-coordinated response. In recent years information and communication technologies (ICT) have been used to engage volunteers as increasingly invaluable resources which complement professional response teams (e.g. medical, police, fire fighter). With a focus on privacy as an integral enabler and precondition for volunteers to take part in such ICT systems, we introduce a privacy-conserving spatial crowdsourcing approach tailored for emergency and disaster response.**

***Index Terms*—spatial crowdsourcing, privacy, privacy by design, emergency services, disaster response, community first responder**

## I. Introduction

Between 2011 and 2016 the Berlin Fire Department registered an increase of emergency calls by 30% [6]. in 50 to 75 percent of cases, depending on an incident's severity, an emergency call is legally required to be attended within 8 minutes after it has been reported [7]. In the first quarter of 2018 alone, the Fire Department reported a total of 27 cases (i.e. every third day in average) of *emergency state* during which 90% of all ambulances were being deployed [18]. Time and human resources have become scarce goods bringing fire departments, disaster management teams, and alike to appeal to volunteers to overcome (human) resource shortages and to improve response times. Consider the following scenario to see how volunteering can be leveraged for emergency response: during the rush hour on a busy conjunction, a driver gets a heart attack and causes an accident. The incident is reported automatically by the integrated emergency system in the car, *eCall*. As every minute counts in increasing the survival chance of the victims, first aid or community responders could play a vital role in saving the victims while emergency team arrives.

The point of departure for this paper are ICT-based solutions which aim to localize and contact volunteers in the direct vicinity of an incident who could timely attend the victims and provide first aid while professional care arrives. Such solutions aim to improve the overall quality and reduce the response time of emergency and disaster response by tapping the "latent talent of the crowd", that is by *crowdsourcing* [22] tasks to volunteers based on their current position. This special type of crowdsourcing which takes the position of volunteers into account is referred to as *spatial crowdsourcing*. Different approaches have been implemented and deployed in various countries such as the mobile community first responder (CFR) proposed by Yonekawa et al. [38] (JP), *Hands2Help* [21] (DE), *Mobile Retter* [33] (DE), *FirstAED* [20] (DK), and *CrowdTasker* [27] (AT) just to name a few.

In this paper we contemplate on the privacy aspects of such a spatial crowdsourcing system. As private information have become a commodity in the digital age and previous data spills/scandals have shaken the trust in information systems, we regard privacy as an integral enabler in ICT systems specifically in emergency and disaster management where personal information plays a vital role. Our main contributions in this work are as follows:

- In collaboration with the Berlin Fire Department we have developed a concept for a spatial crowdsourcing for emergency and disaster situations (see section II). This concept extends our previous work in disaster response crowdsourcing in [16].
- We summarize the principles of privacy by design and examine to which extent and how our work can be adapted to these principles (see section III).
- We introduce our production-ready implementation, *KATRETTER* and discuss its future potential and research direction (see section IV and section VI).

In order to present a concise reading experience, the theoretical and mathematical concepts of *location-based services* and *spatial crowdsourcing* will be presented in detail towards the end of this article in section V.

## II. Concept

Spatial crowdsourcing (SC) refers to assigning "location-specific tasks that require people to physically be at specific locations to complete them" [39]. A spatial crowdsourcing approach can be described through its task model, worker model, and optimization goal. A task is an assignment that is to be carried out by workers rewarded by some form of incentive. Figure 1 provides a taxonomy summarizing these factors.

In this section a *confirmation based* [3] spatial crowdsourcing solution for emergency and disaster response is conceptualized which can be characterized using the taxonomy given in Figure 1 as follows:

- **Task Model**:
  - **Worker Count**: Bounded
  - **Task Area**: Point
  - **Task assignment**: Server Assigned
- **Worker Model**:
  - **Reward Model**: Self-incentivised
  - **Constraints**: Server Imposed
- **Optimization Goal**: Maximize Task Coverage, Minimize Overdue Tasks

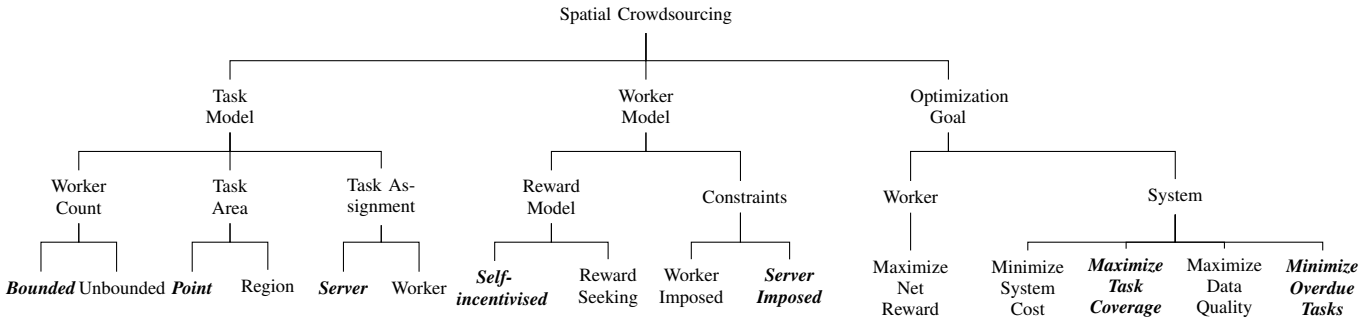


Fig. 1. Taxonomy of spatial crowdsourcing (adapted from [39, 36])

In our model tasks and corresponding constraints, i.e. when, where, and how a task is to be carried out, are dictated by the SC-server. A specific number of volunteers are assigned with spatial tasks and the goal is to maximize the number of workers that accept the tasks while trying to minimize the number of tasks which miss the deadlines.

This paper extends the definition of spatial task (ST) by To, Shahabi, and Kazemi as a tuple $< l, q, e, s, \delta, c >$ where $l$ denotes the location where query $q$ of task type $e$ is to be executed by a worker within the temporal interval of $[s, s+\delta]$. A worker is a person volunteering to accept tasks [36]. Here, a positive integer $c$ is added to the original definition to denote the upper limit on the worker count.

In contrast to common SC approaches, in emergency and disaster response timeliness and worker count are of integral importance as time to start and carry out tasks is strictly constrained. The number of workers must also be controlled, as too few or too many workers might disturb the professional work. We define two tasks in this work: *Aid* and *Assist* respectively for emergency and disaster response. Aid tasks are time-critical skill-based tasks and require prompt attention such as providing first aid during a heart attack. Assist tasks target unskilled workers, are not time-critical, and take place in some future time such as helping with securing roadways in flood season.

It is required for workers to register with the SC-server and provide regular updates about their positions. The location data must be just as precise as required by the SC-server to maintain a minimum functionality. Assuming that the location of volunteered workers is given and up-to-date, task assignment starts with building a spatial plane around $l$ which satisfies two (contradictory) requirements: i) it is large enough to accommodate an adequate number of workers, and ii) it is small enough so that workers can reach $l$ before the task expires $(s + \delta)$. Building this plane demands local knowledge about $l$ and its vicinity, such as population density, reachable roads, etc. to be able to adapt the plane's size to fulfill the aforementioned requirements.

After that, a list of workers within that plane is constructed and sorted (e.g. in regard to data staleness and euclidean distance to $l$) and finally, the first $c$ workers in the list are notified. Tasks can be accepted, rejected or ignored. Due to their critical nature, Aid tasks are handled slightly different from Assist tasks. Aid tasks are assigned in a number of iterations until a sufficient number of workers have accepted the task.

It is evident that in such a concept where system functionalities depend on private data collection it is necessary to integrate transparent mechanisms to ensure stakeholders that data cannot be misused for surveillance, sanctions, etc.

## III. Privacy by design

The search for a universal and commonly accepted definition of privacy remains futile. Many national and international legal regulations observe privacy as a *fundamental right* [14] which requires protection through legal means. Privacy, however, is not merely a legal and/or a moral matter. Privacy has been regarded as a "psychological and anthropological necessity" integral to democratic societies [9] and crucial for development of autonomous individuals [23, 9]. At the same time, opponents of privacy (or proponents of *post-privacy*) argue for its disposal for reasons such as convenience, security, greater social good, etc. (for an in-depth discussion see [23, 25, 26, 9]). Beyond arguments for and against privacy, the fact is that privacy awareness is on the rise. A study by the Pew Research Center shows that "Some 74% say it is 'very important' to them that they be in control of who can get information about them, and 65% say it is 'very important' to them to control what information is collected about them" [32]. Concerns about the protection of one's private sphere are amplified as data collection, propagation, and storage are becoming invisible and also possible "from a distance that had previously constituted the realm of communication and information privacy" [25]: in the age of information systems and data-driven services there seems to be no rear-view mirror to see if your are being followed or not. Even if one is aware about the nature of data being collected about one's self, the potential information-substance of collected data might go beyond the intended collection purpose so that "the models built with the data can have predictive power beyond the context" [26], for example, an individual's movement profile can be linked with publicly available data to deduce facts about a person's personal preferences, social circles, etc. [26].

Protecting privacy is always a *best-effort* enterprise and perpetually relative: information about individuals, even anonymized or pseudonimized, always reveals to some extent facts about those individuals. Privacy, thus, cannot be understood in absolute terms. This, however, should not mislead us to the slippery slope argument of "since there is no absolute in privacy, there is no privacy at all", rather it should sensitize us to the context in which privacy emerges so that it can be understood as "a continuum" and "a matter of judgment" [23]. In the context of spatial crowdsourcing, specifically in emergency and disaster response, we face a complicated dilemma: on the one hand the basic functionality of the system relies on having access to up-to-date information (e.g. location data)

TABLE I
OVERVIEW OF SELECT PRIVACY PRINCIPLES

| | |
|---|---|
| Langheinrich [25] | notice, choice and consent, proximity and locality, anonymity and pseudonymity, security, and access and recourse. |
| Cavoukian – Privacy by design [10] | proactive not reactive; preventative not remedial, privacy as the default setting, privacy embedded into design, full functionality, end-to-end life-cycle protection, visibility and transparency, respect for user privacy |
| FTC guideline [15] | data security, reasonable collection limits, sound retention practices, and data accuracy |
| OECD privacy guideline [29] | collection limitation, data quality, purpose specification, use limitation, openness, individual participation, and accountability |
| Regulation (EU) 2016/679 [30] | data protection (e.g. pseudonymization), collection limitation, specific processing, limited storage and accessibility, notice and consent, security and safeguarding |

of workers; on the other hand, there is little to no incentive for (unpaid) volunteers to sacrifice their private data except maybe for higher altruistic reasons. The results of a study from our previous work show that under eligible volunteers (i.e. willing to use a mobile application) around 15% are critical to automatic location data collection and about 54% against signing up to the system with their personal information (a total of 887 survey participants with 495 eligible volunteers) [8]. It can be seen that guaranteeing privacy is an important incentive to build up trust and bind workers.

Guaranteeing privacy, however, cannot be regarded as an add-on which is imposed on an existing system: privacy should be pursued by design through embedding data protection and privacy principles "throughout the entire life cycle of technologies, from the early design stage to their deployment, use and ultimate disposal" [14].

In the scope of this work we address and adapt principles of privacy by design proposed by Langheinrich [25], Cavoukian [10], FTC guidelines [15], OECD privacy guidelines [29], and the Regulation 2016/679 of the European Parliament and of the Council [30]. It is noteworthy that in contrast to the others, Cavoukian [10] takes a rather descriptive than imperative/normative approach to the privacy by emphasizing on *what* and not *how*. A summary of key factors are presented in Table I. For our purpose, we combined these principles into three categories and adapted them to the concept introduced in section II as follows.

### A. Consent and choice

Workers must transparently be able to understand and decide the purpose and the reach of data collection. Before explicit consent is given, no information shall be collected by abiding to the principle of *data protection by default* [30, Article 25]. All stakeholders must be able to follow the logical data flow and its content. It should also be possible to retract previously granted consent at any time.

### B. Collection, usage and storage limitation

Even if consent to specific types of data collection is collected, it should be made sure that those data are only used for purposes to which user consented and not more. For example, if someone grants access to one's location in context of an early warning system, it cannot be used for targeted advertisement.

In spatial crowdsourcing regular location data collection is required for successful task assignment. However, location collection should be limited to the worker's latest position. Location data are ought to be updated if and only if a worker has roamed at least more than a predefined distance from its previous registered location. Position update rate should also be temporally limited (e.g. maximum every $5 \times 60s$). If a worker leaves the area in which task assigners operate, location collection should be deactivated. In time-critical tasks where the exact position of a volunteer is required to guarantee proper functionality of the system, the SC-Server should only query a subset of volunteers with the highest probability of being in the vicinity of the task's location for their precise position and not all workers.

Data, e.g. task details and volunteer responses, moving through the system are to be removed from caches and long-term databases. Log entries should be kept only as long as necessary for accountability and functional debugging and are to be purged regularly.

### C. Data protection and security

Personal data is to be protected at all stages of collection, processing, and storing. Communication channels between volunteers and the SC-server are to be secured using cryptographic measures to avoid eavesdropping. Authentication and authorization (e.g. to prove a specific skill) should be done through separate services. By the same token, all data that is not strictly required for the purpose of task assignment and is only relevant for workers and/or task assigners (e.g. name of the victim) should be encrypted so that unauthorized access is mitigated. Workers' real IDs are to be replaced with pseudonyms so that an adversary cannot access personal data even if the SC-server is compromised.

Location data should be obfuscated by any of the measures presented in section V to make sure that even if the SC-server is compromised, the adversary cannot figure out the exact location of workers.

## IV. IMPLEMENTATION

In cooperation with the *Berliner Feuerwehr* (Berlin Fire Department) an elaborate spatial crowdsourcing project for emergency and disaster response, called *KATRETTER*[1], has been realized. In sequel, it is shown how privacy considerations of section III are materialized in a productive system consisting of an *Operator*, a *Core*, and a *Worker* plane.

### A. Components

A simplified overview of components and the respective message flow among them is given in Figure 2. Dashed boxes denote user interfaces. The components can be regarded as nodes within a distributed system where communication is limited to message passing. In addition to task assignment, a simple information dissemination mechanism (e.g. news) is integrated within the system which is out of the scope of this work. Implemented components are categorized into three planes:

[1] http://s.fhg.de/YTG

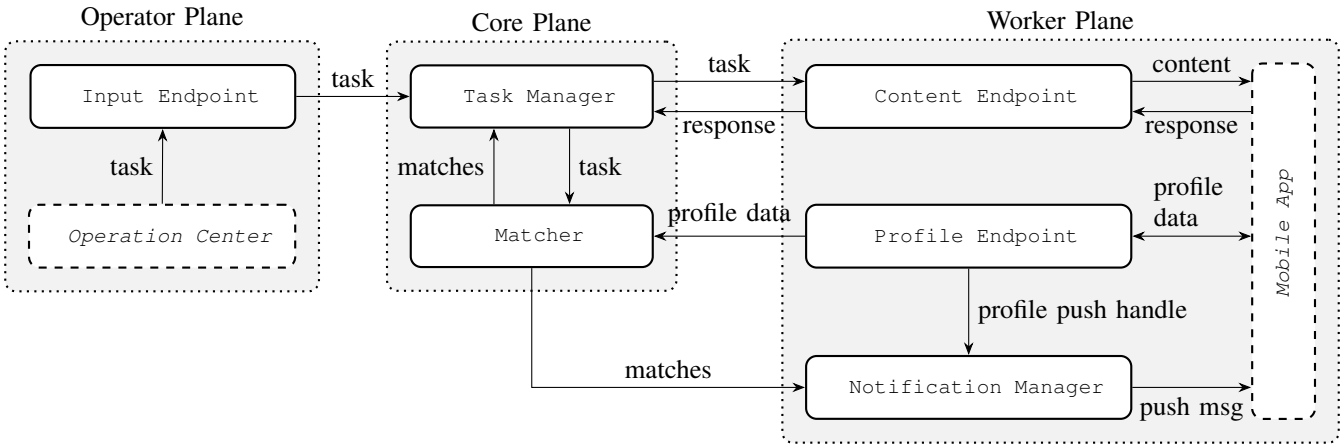


Fig. 2. Components overview

- **Operator plane**: through this plane tasks can be submitted to the system through the Operation Center interface. The Input Endpoint receives tasks from the Operation Center and is responsible for validation and forwarding to the Core plane. It should be noted that multiple operation centers (e.g. police department, emergency agencies) can simultaneously be attached to the system.
- **Core plane**: this plane is responsible for task assignment. The Task Manager and the Matcher cooperate with components from the Operator and the Worker planes to guarantee successful task assignment within given time constraints. The Task Manager monitors and controls the life-cycle of tasks; the Matcher uses workers' profile data to match suitable workers to given tasks.
- **Worker plane**: worker affairs are addressed in this plane. Workers use a mobile application (see Figure 3) to register using pseudonyms. Workers' locations are regularly registered with the Profile Endpoint which in turn are used by the Matcher and the Notification Manager for task assignment. The Profile Endpoint also maps push handles (i.e. push addresses of worker) to worker IDs. The handles are used by the Notification Manager to dispatch push messages. The Content Endpoint caters for querying task details and responding to assignments.

### B. Workflow

*1) Registration:* prior to registration, workers are engaged in a step-by-step process of inform-and-consent: each step informs the user on a specific type of data collection and how the data is processed and used (Figure 3 (a)). The step-by-step inform-and-consent is an antithesis to the common tedious *read-all/accept-all* terms-of-use, EULA, privacy-policy, and other type of digital agreements. At this point the worker has consented only on basis data collection and processing required to be part of the system; no location information is yet collected. For registration no personal information is required and each worker, or more precisely each app instance, becomes a randomly generated pseudonym.

A worker can then decide to opt-in to be assigned for Aid or Assist tasks (Figure 3 (b)). As Aid tasks require specific skills (e.g. CPR certification) and require worker verification, the sign-up requires workers to authenticate themselves through a trusted third party (TPP) via a separate channel and not through the Core or the Operator plane. The Core plane is then only informed by TPP if an app instance and its assigned pseudonym have specific attributes or not without revealing the real identities. Assist tasks are open to public can be signed-up to without any further restrictions.

*2) Worker Tracking:* after opting-in, the mobile application transmits the worker's location through the Profile Endpoint once and updates it if the worker moves further than a predefined distance (e.g. $500m$) but not more than once within a given temporal interval (e.g. at most once every $5 \times 60s$). The worker's position is obfuscated by the mobile app by shifting its center according to a default privacy preference (Equation 1) (same for all users) as described in [2] (see section V).

*3) Task Assignment:* all three of Operator, Core, and Worker planes are involved with task assignment. A task $t =< l, q, e, s, \delta, c >$ is created either by an operator or is submitted automatically through computer-aided dispatch (CAD) through the Operation Center. After being validated by the Input Manager and depending on the task type $e$ (Aid or Assist), the Task Manager initiates the actual assignment by creating an isochrone of equal travel time $\delta$ around $l$ using related cartographic data. Due to noise in stored workers' positions, the isochrone is enlarged by a predefined factor (proportional to the privacy preference used for obfuscation) before being handed over to the Matcher. The Task Manager instructs the Matcher to notify all workers which are known to be within the enlarged isochrone (using worker profile information from the Profile Endpoint) to transmit their precise and unobfuscated current location. This step is necessary prior to task assignment since only obfuscated locations are stored by the Profile Manager and the data might be out-dated since last update. After a predefined waiting interval, it is assumed that positions of online (i.e. reachable) workers within the isochrone are updated and the Task Manager instructs the Matcher to notify the first best $c$ workers within the respective area with partial contents of $q$ (the complete description of $q$ is only accessible to workers who have accepted the task). The Matcher can also be triggered by the Profile Endpoint upon a worker's location update if that worker enters the previously created isochrone. The Notification Manager uses an external push service to send push messages containing only the necessary information for workers to decide whether they want to accept a task or not. As an example, Figure 4 depics this procedure: the white marker denotes the position


(a) Inform and consent (b) Aid/Assist opt-in (c) Task (summary/detailed)

Fig. 3. Mobile application for workers

Fig. 4. Example of position inquiry during task assignment

of an incident. As the task is issued, the isochrone of the area reachable by pedestrians under three minutes is calculated (the irregular area shaded in blue). To account for location imprecision of the workers, the isochrone is extended by a predefined factor (for the sake of simplicity the isochrone is transformed into a circle) and all workers estimated to be within this area (marked in orange as moving) are requested to provide their respective precise location (marked in green as standing). As it can be seen three of the workers are within the isochrone, while one has left the depicted map area. Subsequently, depending on how many workers are required, some or all of the workers within the isochrone are assigned with the task.

As previously mentioned, Aid tasks are assigned iteratively. For example, in case of a heart attack, the task assigner could submit the following task:

$$t = <l, q : \texttt{"Heart attack (...)"},\\ e : \texttt{Aid}, s : \texttt{CURRENT\_TIMESTAMP},\\ \delta : 10 \times 60s, c : 3 >$$

Assuming that task assignment is configured to succeed in $i$ iterations and $n > c$ workers have been matched within the respective isochrone of $l$, in the first iteration first 3 out of $n$ workers are notified. If all accept the task within a predefined interval (e.g. $30s$), the task is closed (but not removed). Otherwise, the procedure is repeated in predefined intervals in no more than another $i - 1$ iterations as long as $s + \delta$ is not elapsed until enough workers have accepted the task.

*4) Task Confirmation/Rejection:* Upon receiving a task the workers can confirm the task using the mobile app (Figure 3 (c)) through the Content Endpoint and consequently pull complete task details. The worker can also update the task's status over the Content Endpoint. This process is the same for both Aid and Assist tasks.

Aid tasks are time-critical and need to be accepted in a short time interval. To prevent possibility of sanctions on behalf of the task assigner (e.g. fire department) for those workers who do not accept Aid tasks, the possibility of rejection is not given and the task assigner cannot find out which workers have been informed; Aid tasks can only be accepted or ignored.

Workers who are notified for or accept an Aid task are temporarily marked as busy not to be informed for parallel tasks and to be able to regenerate after completing the task. Workers' personal data are encrypted using an ephemeral public key, generated only for the task, and are transmitted back to the task assigner. This way it is made sure that none of the involved planes have access to this information and the worker's pseudonymity is maintained.

## V. Related Work

This section is limited to the literature review of privacy in terms of location anonymity within the context of location-based services and spatial crowdsourcing. The most prominent approaches are location *k-anonymity*, location obfuscation, and differential-privacy.

### A. Location-based Services

The basic model of location-based services is querying a location-based service (LBS) and provide own location for tailored results.

Beresford and Stajano introduce the concept of *mix zone* as a closed spatial plane aiming to "prevent tracking of long-term user movements, but still permit the operation of many short-term location-aware applications" [4]. This model divides the

space into *application* and mix zones. Users can register for specific services, thus, registering for specific application zones. A user is considered to be in a mix zone, if he is not within any of his registered application zones. Within an application zone, the corresponding application has access to user's location through a TPP in charge of pseudonomyzing users' identifications acting as a communication middleware between service providers and users. New IDs are assigned by the middleware to users upon moving into a new zone or after some predefined amount of time is elapsed. Without having access to the trusted middleware an application cannot figure out a user's real identity as she enters the respective zone as she could be any of $n$ users within the mix zone just before she enters the application zone [5].

Location k-anonymity was first introduced by Gruteser and Grunwald. It follows the goal of generating an *anonymity set* [11] of positions within which "the location information presented is indistinguishable from the location information of at least $k-1$ other subjects" [19]. $k$ is then used as a metric to quantify privacy. In this model a subject's position is denoted by a tuple $([x_1, x_2], [y_1, y_2], [t_1, t_2])$ where $x$ and $y$ are used to denote the spatial and $t$ the temporal range. To reach k-anonymity, queries to an LBS first go through a centralized *location anonymizer* with global knowledge of subjects and their location which adaptively (with regard to population density) reaches k-anonymity by reducing spatiotemporal accuracy of information by increasing spatial and/or temporal ranges so that at least the location of $\kappa \geq k$ are indistinguishable by the LBS. This, however, requires users to trust the anonymizer and introduces a single point of failure, which can reveal users' actual locations if compromised by an adversary. This has been addressed, among others, by Peng, Liu, and Wang by first transforming users' 2-D locations into a 1-D Hilbert space and then forward them to the anonymizer. A *Function Generator* is designated to distribute transformation parameters between service provider and users. An attacker would require to have both access to anonymizer and Function Generator to acquire users' location.

A similar approach without reliance on centralized anonymizers is proposed by Kido, Yanagisawa, and Satoh. Here an algorithm locally generates a set containing user's actual position and a number of dummy locations. The basic idea is to query service provider on the whole set and filter irrelevant results (from dummy positions) after the query succeeds. The algorithm presupposes grid building, i.e. dividing the spatial plane into (rectangular) regions. The scale of the regions and the number of people within each region is then used to quantify the *location anonymity*: the higher the scale of regions and the more people within a region, the higher the location anonymity. Grid building, however, introduces extra overhead and is not sensitive in its proposed form to population density so that by default urban areas would guarantee a higher privacy than rural areas.

Under the assumption that a position is expressed as a circular plane within which the actual position is normally distributed, Ardagna et al. propose a *relative privacy preference* metric composed of the subject's privacy preference (in terms of minimum distance $r_{min}$) and original location measurements $r_{meas}$ (including errors):

$$\lambda = \frac{max(r_{meas}, r_{min})}{r_{meas}^2} - 1 \tag{1}$$

Prior to sending queries to service providers, a trusted middleware (i.e. TTP) uses given $\lambda$ to obfuscate user's location by either i) enlarging the radius, ii) shifting the center, iii) reducing the radius (or a combination of those) of the circular plane representing the subject's position. *Relevance* (related to accuracy) is defined inversely proportional to $\lambda$ and can be used by service providers to ensure a minimum QoS [2].

Finally, Differential privacy (DP) [12] has also been adapted to reach location privacy. In its original form, DP aims for privacy preserving statistical databases and is defined as follows: "A randomized function $\mathcal{K}$ gives $\epsilon$-differential privacy if for all data sets $D_1$ and $D_2$ differing on at most one element, and all $S \subseteq \text{Range}(\mathcal{K})$" the following holds:

$$Pr[\mathcal{K}(D_1) \in S] \leq \exp(\epsilon) \times Pr[\mathcal{K}(D_2) \in S] \tag{2}$$

where $Pr$ denotes the probability, and $\epsilon$ the "*leakage*" [13]. In other words, Equation 2 states that by removing an item from a dataset "no outputs (and thus consequences of outputs) would become significantly more or less likely" [12] and an adversary with background information cannot infer the existence (or the absence) of an item within the database. To achieve DP, when queries act on the whole database (e.g. `SUM`, `COUNT`), the author provides a method of adding random noise to the query results.

Andrés et al. adapt DP to location and states that a mechanism $K$ guaranties $\epsilon$-geo-indistinguishably for all points $x_1, x_2 \in \mathcal{X}$ if and only if the following holds:

$$\sup_{x \subseteq \mathcal{X}} \left| \ln \frac{K(x_1)}{K(x_2)} \right| \leq \epsilon d(x_1, x_2) \tag{3}$$

where $d(x_1, x_2)$ denotes the euclidean distance between $x_1$ and $x_2$. The authors define $\epsilon$ proportional to $r$, the radius within which a subject's privacy is to be protected, and a privacy level $\ell$. In this sense *$\ell$-geo-indistinguishably* is simply reached by replacing $\epsilon$ with $\ell/r$ in Equation 3. In other words, *$\ell$-geo-indistinguishably* within $r$ states that "any two locations at distance at most $r$ produce observations with 'similar' distributions, where the 'level of similarity' depends on $\ell$". The authors describe a method to cloak subjects' actual positions using $\ell$-geo-indistinguishably. This approach does not require any TPP [1]. ...

### B. Spatial Crowdsourcing

Whereas in location-based services users initiate the communication with server, in (server assigned) spatial crowdsourcing (SC) the SC-server triggers the communication. Meaningful task assignment requires a priori knowledge of subjects' positions.

To, Ghinita, and Shahabi claim to be the first to introduce an approach to privacy in spatial crowdsourcing during the task assignment phase based on differential privacy (cf. [1]). This model uses *cellular service providers* (CSP) as TTPs for tasks such as grid construction, data sanitation, and *geocasting* ("sending a message selectively only to specific subareas defined by latitude and longitude" [28]). CSPs build adaptive (with regard to population density) grids and add fake workers to conform to requirements of DP within a predefined *privacy-budget* $\epsilon$. To assign a task, the SC-server first determines appropriate region for which workers may come into consideration, the geocast region, and initiates the a *geocast* communication either with the help of CSPs or

through a combination of CSP and hop-by-hop propagation in an ad-hoc mobile network [34]. Geocasting prevents the SC-server to figure out which workers are fake and which are real by using proxies to deliver tasks from the server to workers. In a separate publication, the authors introduce *PrivGeoCrowd* that "helps system designers investigate the effect of parameters such as privacy budget and allocation strategy, GR [geocast] construction heuristics, dataset density, etc., on the effectiveness of private SC task matching" [35].

## VI. Future Research

The most notable weakness of the proposed approach lies within its centralized design which has a negative impact on privacy as any type of adversarial attack might lead to information leakage for *all* workers. A promising solution to overcome this might be *mobile edge networking* through the basic idea of moving "network functions, contents and resources closer to end users, i.e., the network edge" [37]. Respectively, the worker and core planes could be moved to the edge of the network and the Input Endpoint could then select to which core plane a task is forwarded to. The advantage of such an approach is twofold: on the one hand, critical data such as workers' positions are stored and processed only locally in edge nodes. On the other hand, network latency is reduced notably as round trip times are decreased leading to an improvement in response times.

Another issue which must be addressed is the lack of or the damaged state of the infrastructure which is typical to disaster situations. In our implementation, we used push notifications which rely on external services operating over the traditional IP suite. New methods must be investigated on how to integrate spatial crowdsourcing with networking solutions in infrastructureless settings such as (mobile) ad hoc networks.

Moreover, the feasibility of task encryption should be studied as Aid tasks usually carry personal information such as name and address of the person in need, e.g. suffering a heart attack. The method we used for encrypting task confirmations is not applicable for this case as it uses a single key, namely that of the task assigner, to encrypt multiple confirmations, whereas public-key task encryption would require encryption for each and every worker which is to be notified. The feasibility of existing cryptographic approaches such as *proxy re-encryption* [17] should be further investigated.

Finally, it would be advantageous to engage users in the data processing mechanisms of the system beyond the initial notice and consent by utilizing feedback and quantitative metrics and allowing them to reconsider their initial choices. Utilizing methods which quantify the privacy level (see section V) and enable users to choose appropriate levels of privacy can also reinforce the mutual trust.

## VII. Conclusion

In this work we conceptualized a privacy-preserving spatial crowdsourcing framework for emergency and disaster response. We discussed how privacy by design can be an enabler for voluntary participation and showed how a system, which maintains a balance between holistic privacy, and functional and real-time constraints, can be realized.

A proof-of-concept, KATRETTER, has already been implemented in cooperation with Berliner Feuerwehr and is to be deployed for production for the purpose of emergency and disaster response in Germany.

## References


[1] Miguel E. Andrés et al. "Geo-Indistinguishability: Differential Privacy for Location-Based Systems". In: *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security - CCS '13*. New York, New York, USA: ACM Press, 2013, pp. 901–914. ISBN: 9781450324779. DOI: 10.1145/2508859.2516735. arXiv: 1212.1984.

[2] C. A. Ardagna et al. "Location Privacy Protection Through Obfuscation-Based Techniques". In: 2007, pp. 47–60. ISBN: 978-3-540-73533-5. DOI: 10.1007/978-3-540-73538-0_4.

[3] Christine Bauer, Andreas Mladenow, and Christine Strauss. "Fostering Collaboration by Location-Based Crowdsourcing". In: 2014, pp. 88–95. ISBN: 978-3-319-10830-8. DOI: 10.1007/978-3-319-10831-5_13.

[4] A.R. Beresford and Frank Stajano. "Mix zones: user privacy in location-aware services". In: *IEEE Annual Conference on Pervasive Computing and Communications Workshops, 2004. Proceedings of the Second*. IEEE, 2004, pp. 127–131. ISBN: 0-7695-2106-1. DOI: 10.1109/PERCOMW.2004.1276918.

[5] a.R. Beresford and Frank Stajano. "Location privacy in pervasive computing". In: *IEEE Pervasive Computing* 2.1 (Jan. 2003), pp. 46–55. ISSN: 1536-1268. DOI: 10.1109/MPRV.2003.1186725.

[6] Berliner Feuerwehr. *Berliner Feuerwehr in Zahlen*. 2016. URL: https://web.archive.org/web/20180531122410/https://www.berliner-feuerwehr.de/ueber-uns/berufsfeuerwehr/berliner-feuerwehr-in-zahlen-2016/ (visited on 05/31/2018).

[7] Berliner Feuerwehr. *Neues Einsatzkonzept der Berliner Feuerwehr*. 2003. URL: https://web.archive.org/web/20180531130043/https://www.berliner-feuerwehr.de/neues-einsatzkonzept-der-berliner-feuerwehr/ (visited on 05/31/2018).

[8] Berliner Feuerwehr et al. *Forschungsprojekt ENSURE*. Tech. rep. Berlin, 2016.

[9] Volker Boehme-Neßler. "Privacy: a matter of democracy. Why democracy needs privacy and data protection". In: *International Data Privacy Law* 6.3 (Aug. 2016), pp. 222–229. ISSN: 2044-3994. DOI: 10.1093/idpl/ipw007.

[10] Ann Cavoukian. "Privacy by design: the definitive workshop. A foreword by Ann Cavoukian, Ph.D". In: *Identity in the Information Society* 3.2 (Aug. 2010), pp. 247–251. ISSN: 1876-0678. DOI: 10.1007/s12394-010-0062-y.

[11] David Chaum. "The dining cryptographers problem: Unconditional sender and recipient untraceability". In: *Journal of Cryptology* 1.1 (1988), pp. 65–75. ISSN: 0933-2790. DOI: 10.1007/BF00206326.

[12] Cynthia Dwork. "Differential Privacy". In: *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming*. 2006, pp. 1–12. ISBN: 3-540-35907-9. DOI: 10.1007/11787006_1.

[13] Cynthia Dwork et al. "Calibrating noise to sensitivity in private data analysis". In: *Third Theory of Cryptography Conference, TCC*. Ed. by Halevi Shai and Tal Rabin Rabin. Vol. 3876. New York, New York, USA: Springer

Science & Business Media, Aug. 2006, pp. 265–284. ISBN: 9783540327318.

[14] European Commission. *Communication from the Commission to the European Parliament, the Council, the Economic and social Committee and the Committee of the Regions: A comprehensive approach on personal data protection in the European Union*. Tech. rep. European Commission, 2010, pp. 1–20.

[15] Federal Trade Commision (FTC). *Protecting Consumer in an Era of Rapid Change: Recommendations for businesses and policymakers*. Tech. rep. Federal Trade Commision (FTC), 2012, pp. 1–112. URL: https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf.

[16] Frank Fuchs-Kittowski et al. "ENSURE - Integration of Volunteers in Disaster Management". In: *IFIP Advances in Information and Communication Technology*. Springer, Cham, May 2017, pp. 247–262. ISBN: 978-3-319-89934-3. DOI: 10.1007/978-3-319-89935-0_21.

[17] Matthew Green and Giuseppe Ateniese. "Identity-Based Proxy Re-encryption". In: *Applied Cryptography and Network Security*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 288–306. ISBN: 978-3-540-72737-8. DOI: 10.1007/978-3-540-72738-5_19.

[18] Félice Gritti. *Rettungsdienst: Zahl der Einsätze steigt von Jahr zu Jahr*. 2018. URL: https://web.archive.org/web/20180501180547/http://www.spiegel.de/panorama/gesellschaft/rettungsdienst-zahl-der-einsaetze-steigt-von-jahr-zu-jahr-a-1203793.html.

[19] Marco Gruteser and Dirk Grunwald. "Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking". In: *Proceedings of the 1st international conference on Mobile systems, applications and services - MobiSys '03*. New York, New York, USA: ACM Press, 2003, pp. 31–42. ISBN: 1470-949X. DOI: 10.1145/1066116.1189037.

[20] Finn Lund Henriksen et al. "FirstAED emergency dispatch, global positioning of community first responders with distinct roles - a solution to reduce the response times and ensuring an AED to early defibrillation in the rural area Langeland". In: *International Journal of Networking and Virtual Organisations* 16.1 (2016), pp. 86–102. DOI: 10.1504/IJNVO.2016.075131.

[21] Marlen Hofmann, Hans Betke, and Stefan Sackmann. "Hands2Help – Ein App-basiertes Konzept zur Koordination Freiwilliger Helfer". In: *i-com* 13.1 (Jan. 2014), pp. 36–45. ISSN: 2196-6826. DOI: 10.1515/icom-2014-0005.

[22] Jeff Howe. *The Rise of Crowdsourcing*. 2006. URL: https://web.archive.org/web/20180330102002/https://www.wired.com/2006/06/crowds/ (visited on 05/23/2018).

[23] Lucas D. Introna. "Privacy and the Computer: Why We Need Privacy in the Information Society". In: *Metaphilosophy* 28.3 (July 1997), pp. 259–275. ISSN: 0026-1068. DOI: 10.1111/1467-9973.00055.

[24] Hidetoshi Kido, Yutaka Yanagisawa, and Tetsuji Satoh. "Protection of Location Privacy using Dummies for Location-based Services". In: *21st International Conference on Data Engineering Workshops (ICDEW'05)*. Vol. 2005. IEEE, 2005, pp. 1248–1248. ISBN: 0-7695-2657-8. DOI: 10.1109/ICDE.2005.269.

[25] Marc Langheinrich. "Privacy by Design — Principles of Privacy-Aware Ubiquitous Systems". In: *Ubicomp 2001: Ubiquitous Computing*. 2001, pp. 273–291. ISBN: 3540426140. DOI: 10.1007/3-540-45427-6_23.

[26] Tobias Matzner. "Why privacy is not enough privacy in the context of "ubiquitous computing" and "big data"". In: *Journal of Information, Communication and Ethics in Society* 12.2 (May 2014), pp. 93–106. ISSN: 1477-996X. DOI: 10.1108/JICES-08-2013-0030.

[27] Michael Middelhoff et al. "Crowdsourcing and Crowdtasking in Crisis Management Lessons Learned From a Field Experiment Simulating a Flooding in the City of the Hague". In: *Proceedings of the 2016 3rd International Conference on Information and Communication Technologies for Disaster Management, ICT-DM 2016*. IEEE, Dec. 2017, pp. 1–8. ISBN: 9781509052349. DOI: 10.1109/ICT-DM.2016.7857212.

[28] Julio C. Navas and Tomasz Imielinski. "GeoCast - geographic addressing and routing". In: *Proceedings of the 3rd annual ACM/IEEE international conference on Mobile computing and networking - MobiCom '97*. New York, New York, USA: ACM Press, 1997, pp. 66–76. ISBN: 0897919882. DOI: 10.1145/262116.262132.

[29] OECD. *OECD Privacy Guidelines*. Tech. rep. 2013, pp. 1–154, 9–18. DOI: 10.1787/5kgf09z90c31-en. URL: https://www.oecd.org/internet/ieconomy/privacy-guidelines.htm.

[30] European Parliament and Council of the European Union. "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Da". In: *Official Journal of the European Union* 59.4.5.2016 (2016), pp. 1–88. URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj.

[31] Tao Peng, Qin Liu, and Guojun Wang. "Enhanced Location Privacy Preserving Scheme in Location-Based Services". In: *IEEE Systems Journal* 11.1 (Mar. 2017), pp. 219–230. ISSN: 1932-8184. DOI: 10.1109/JSYST.2014.2354235.

[32] Pew Research Center. *The state of privacy in America*. 2016. URL: https://web.archive.org/web/20181001185251/http://www.pewresearch.org/fact-tank/2016/09/21/the-state-of-privacy-in-america/ (visited on 10/01/2018).

[33] R. Stroop et al. "Smartphone-basierte First-Responder-Alarmierung Mobile Retter". In: *Der Notarzt* 31.05 (Oct. 2015), pp. 239–245. ISSN: 0177-2309. DOI: 10.1055/s-0035-1552700.

[34] Hien To, Gabriel Ghinita, and Cyrus Shahabi. "A framework for protecting worker location privacy in spatial crowdsourcing". In: *Proceedings of the VLDB Endowment* 7.10 (June 2014), pp. 919–930. ISSN: 21508097. DOI: 10.14778/2732951.2732966.

[35] Hien To, Gabriel Ghinita, and Cyrus Shahabi. "PrivGeoCrowd: A toolbox for studying private spatial Crowdsourcing". In: *2015 IEEE 31st International Con-*

*ference on Data Engineering*. Vol. 2015-May. IEEE, Apr. 2015, pp. 1404–1407. ISBN: 978-1-4799-7964-6. DOI: 10.1109/ICDE.2015.7113387.

[36] Hien To, Cyrus Shahabi, and Leyla Kazemi. "A Server-Assigned Spatial Crowdsourcing Framework". In: *ACM Transactions on Spatial Algorithms and Systems* 1.1 (July 2015), pp. 1–28. ISSN: 23740353. DOI: 10.1145/2729713.

[37] Shuo Wang et al. "A Survey on Mobile Edge Networks: Convergence of Computing, Caching and Communications". In: *IEEE Access* 5 (2017), pp. 6757–6779. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2017.2685434.

[38] Chikara Yonekawa et al. "Development of a first-responder dispatch system using a smartphone". In: *Journal of Telemedicine and Telecare* 20.2 (Mar. 2014), pp. 75–81. ISSN: 1357-633X. DOI: 10.1177/1357633X14524152.

[39] Yongjian Zhao and Qi Han. "Spatial crowdsourcing: current state and future directions". In: *IEEE Communications Magazine* 54.7 (July 2016), pp. 102–107. ISSN: 0163-6804. DOI: 10.1109/MCOM.2016.7509386.