

# Health participatory sensing networks

Andrew Clarke\* and Robert Steele

*Discipline of Health Informatics, The University of Sydney, Sydney, Australia*

**Abstract.** The use of participatory sensing in relation to the capture of health-related data is rapidly becoming a possibility due to the widespread consumer adoption of emerging mobile computing technologies and sensing platforms. This has the potential to revolutionize data collection for population health, aspects of epidemiology, and health-related e-Science applications and as we will describe, provide new public health intervention capabilities, with the classifications and capabilities of such participatory sensing platforms only just beginning to be explored. Such a development will have important benefits for access to near real-time, large-scale, up to population-scale data collection. However, there are also numerous issues to be addressed first: provision of stringent anonymity and privacy within these methodologies, user interface issues, and the related issue of how to incentivize participants and address barriers/concerns over participation. To provide a step towards describing these aspects, in this paper we present a first classification of health participatory sensing models, a novel contribution to the literature, and provide a conceptual reference architecture for health participatory sensing networks (HPSNs) and user interaction example case study.

Keywords: Participatory sensing, public health, epidemiology, mobile health

## 1. Introduction

The use of health participatory sensing as a data collection methodology is rapidly becoming a reality that will revolutionize the scale and types of data that can be aggregated for a number of population health, epidemiological, statistical and data analysis purposes. Participatory sensing is the act of using mobile devices to allow public and professional users to collect, analyse and submit or share local knowledge to a larger interactive participatory sensing network [1]. The range of possibilities for participatory sensing is large [2], however previous work in participatory sensing has not considered in detail the different participatory models that are likely to occur in the health context. As these models have the potential to scale to millions or nation-wide levels, both the potential and complexities of these data systems are also substantial. This raises the need for a categorization of health participatory sensing models and a description and analysis of the basic architectures possible. Interestingly, we also suggest in this work that beyond sensing alone, the possible models of health participatory sensing may also often include elements of public health intervention or two-way interaction.

The growth in the potential for participatory sensing has been largely accelerated through the high levels of smartphone adoption in many countries [3], leading to the proliferation of powerful sensing platforms that are highly human-centric, making them ideal as the centre-points for health participatory sensing models [4–6]. The potential capabilities are further extended with the addition of ubiquitous external sensing components such as activity monitoring [6] and other wearable consumer health sensors.

---

\*Corresponding author: Andrew Clarke, Discipline of Health Informatics, The University of Sydney, Sydney, Australia.  
E-mail: andrew.clarke@sydney.edu.au.

Contemporary commercial implementations such as Nike Fuel and Jawbone Up [6] demonstrate the achievability and potential for continuous physical activity sensing. Jawbone Up extends beyond physical activity monitoring to include sleep pattern and quality, and a nutritional diary. Other initiatives such as Cykelscore [4], Riderlog [5] and the Copenhagen Wheel [7] are moving towards participatory sensing for specific usage groups. Cykelscore acts as an active transportation data collection tool and incentive framework for cycling, while Riderlog acts as just a collection tool. Alternatively, the Copenhagen wheel goes beyond physical activity sensing to urban environmental monitoring with air quality and noise sensors included in the implementation to provide additional data beyond just the activity of the individual.

In the research domain, attempts to make data collection more automatic or less invasive are being explored [8] and there are also efforts to automate nutrition and dietary intake information capture [9], with proposals including in relation to acoustical dietary intake [10], specialized hardware [11,12], and food image analysis [13]. Additionally, there are attempts to make sensors more unobtrusively wearable, such as with projects like Heartphones [14] – a coupling of a heart rate monitor with headphones and a mobile device.

Overall this has led to a great improvement in potential capability, but with most research focused on the use of such data collection by the individual, there has been less attention given to the potential for and challenges to usage to provide population wellness measures or for wider population health or population epidemiological usage. With the growth in sensor capabilities now and in the near future, an analysis of health participatory sensing capabilities, models and architectures is timely. In Section Two we provide a classification of health participatory sensing models, in Section Three we describe a conceptual reference architecture for HPSNs, in Section Four we address user interaction and Section Five is the Conclusion.

## 2. Classification of health participatory sensing models

The classification of health participatory sensing models has not previously been formalized, with the types of interaction not systematically identified or described. In this section we classify health participatory sensing models according to five distinct categories and briefly discuss the potential privacy, security, interface and incentivization aspects.

### 2.1. Incidental participatory sensing

Incidental participatory sensing requires the lowest levels of participation by an individual. It is defined as contributing sensor information for population health measures, which the individual would have already collected for their own use or benefit [15]. An example of incidental participatory sensing is physical activity self-tracking that has become increasingly popular in recent years. Due to the nature of physical activity level as a risk or preventative factor for a number of lifestyle-related diseases, it is also an important item for population health and epidemiological data collection. In our previous work we have discussed in greater depth the possibilities for such data collection for numerous secondary uses [16].

Demonstrating the low perceived privacy risk of sharing this data at least for some individuals, is the recent practice of sharing such physical activity or fitness data, often in real-time, publicly via social media. Additionally, it's reasonable to predict that as mobile device sensors evolve, the collection and

utilization of data by individuals will also expand to quantify further details of their activities, leading to increasingly rich and useful incidental participatory data collection.

Due to the intrinsic self motivation that users already have to collect some types of sensor data, this category of participatory sensing benefits from the lowest 'barrier to entry', since individuals already contribute their own resources such as effort, time, CPU processing, network bandwidth and device battery usage to achieve data capture. This likely indicates that incentives to participate would be unneeded or minimal, if privacy concerns/barriers can be avoided or minimized as indicated in previous work [17]. Sensors and devices that are already acceptable in terms of user experience are already evidenced via numerous commercially available sensor products.

## 2.2. *Passive participatory sensing*

Passive participatory sensing is defined as sensor usage that requires explicit additional effort to enable data collection, which an individual would not have done unless they were explicitly participating in population health data capture, but does not attempt to and does not require any change of the day-to-day behaviour of the individual. For example, this could include the individual using additional sensors that they would not have otherwise used, that collect data, for example, relating to physical activity, diet, heart rate, sleep cycles or environment, specifically to contribute to population data capture. This allows for a potentially more complete data collection in comparison to that which is possible through purely incidental data collection.

This brings to the fore the question of user motivation, once an individual is required to contribute any significant additional effort to collect sensor data. Previous work has discussed the idea of incentive schemes that use a market-based system [18] or the preservation of social translucence [19]. Inherently incentives will continue to pose a challenge to participatory sensing systems. In the health domain there is also the possibility of further incentives – greater self knowledge of health, information and risk assessment, improved care, diagnostics, and possibly even an interest to contribute in a 'citizen scientist' capacity.

However passive participatory sensing also increases privacy concerns. Since the data collection is intrinsically something that would not otherwise have occurred, a higher level of privacy responsibility rests with the collector. In a purely incidental approach it is necessary that any data that leaves the mobile device either be kept strictly de-identified (with re-identification avoided) or securely and privately transmitted and stored. However, in a passive participatory model, responsibility for the secure storage on the device is also of import. This indicates that the type and characteristics of the HPSN architecture will vary depending on the category of health participatory sensing that is deployed.

## 2.3. *Passive participatory sensing with subjective human-sensing and feedback*

This model combines the potential sensing advantages of passive participatory sensing with 'human-sensing' capabilities, allowing for the large amounts of objective sensing data to be complemented with subjective human-generated data and feedback. By human-sensing we refer to manual information inputs or responses provided by individuals.

This could easily be implemented through the addition of context-sensitive micro-surveys that are displayed to the user and attached to relevant collected sensor data. This would allow for both collection of data that is difficult to record through sensors alone, and also allow data that may have been missed via sensor collection to be added to the overall collection. Additionally, where anomalous data has been collected by sensors, human-sensing and feedback allows for validation to be performed by requesting

subjective details or clarification of the data collected. As discussed below, the load on the user of providing such manual input would obviously need to be low.

Human-sensing has potential as a complement to sensor-based participatory sensing, able to bridge the gaps or some limitations of sensing technology [20]. While it would intrinsically increase the effort required to participate and would present a further challenge to motivation, incentives and potential participation levels, it doesn't pose additional security concerns. This is apparent as it would not require direct one-to-one communication, rather generic context rules and micro-surveys could be broadcast to all participants with the processing of context and hence triggering of the additional data collection occurring locally. This is described in further detail in Section 3.

#### *2.4. Active participatory sensing*

Active participatory sensing provides inputs to the individual to alter the actions they would have taken whilst participating in the HPSN. Active participatory sensing in the health context has a somewhat different goal to that of many other active participatory sensing contexts [21]. While an active participatory model for typical sensing might focus on affecting individuals to collect a more complete data set in terms of spatial/temporal range, health and epidemiological-related active participatory sensing would be more concerned with affecting a health-related action and hence have a component equating to a public health intervention. The instigation to carry out 'active' sensing activities could essentially constitute a public health intervention action. As such, the behaviour change would be to firstly attempt to improve the sensing data captured in terms of risk and preventative factors. Additionally for public health goals, this allows for immediate and continuous feedback of the effectiveness of campaigns on recipient groups. It is assumed that active participatory sensing would have similar levels of technical sensor capabilities to passive, with the focus shifted to the potential two-way communication that can be built on sensing data and an inherent feedback loop.

This has the potential to be both a powerful data collection tool as well as a novel public health intervention platform. Its potential scope includes the ability, in a timely and accurate manner, to quantify precisely the effectiveness of public health interventions.

#### *2.5. Active participatory sensing with subjective human sensing and feedback*

This final category incorporates the goals and framework of active participatory sensing but adds subjective human-sensing and feedback – hence incorporating the most complete level of data collection with public health intervention.

As the most all-encompassing level of HPSN system that we have considered it may correspondingly have greater challenges in relation to motivation. However, it should be noted that security and privacy provisions would be required to be no more stringent than for that of either of its component parts. The interaction of these two components adds the capability to give human-sensing or feedback related to a specific intervention. This is a higher level of capability than that available in the other models. It also allows for more complete and useful information to be collected by enhancing what can be detected through sensors.

While this model represents the most complete functionality, in many cases it is likely that not all its capabilities would be required or that the potential motivation challenges may lead to a less comprehensive participatory sensing model being utilized for a particular goal. This is expected, and suggests the material distinctions existing between the different categories of health participatory sensing networks.

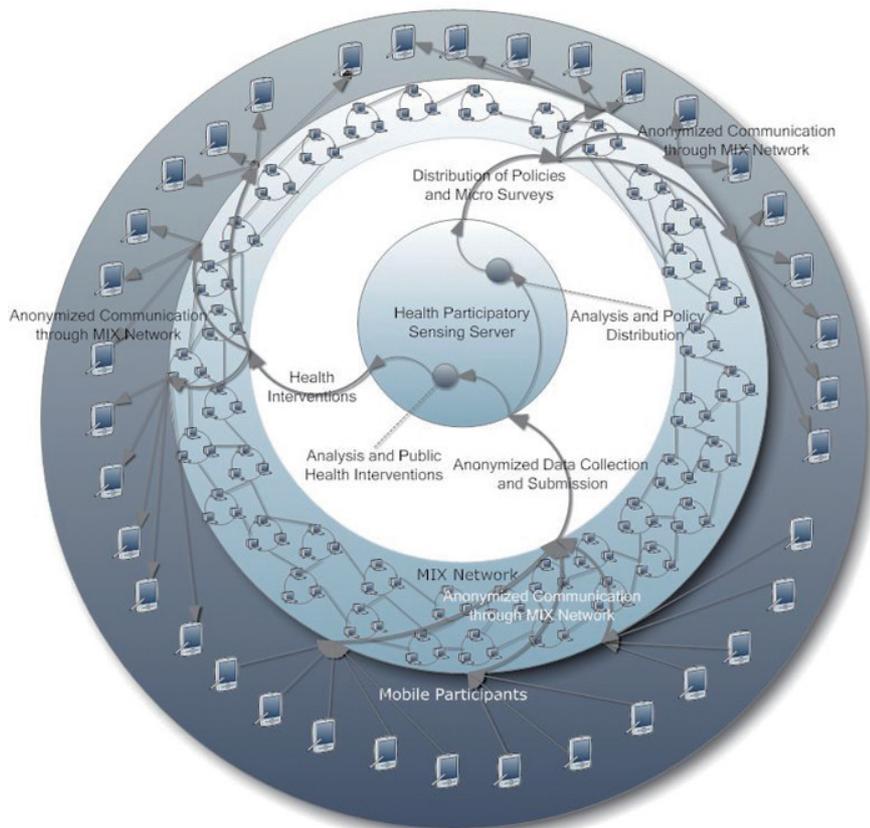


Fig. 1. Health participatory sensing network reference architecture.

### 3. Health participatory sensing network reference architecture

#### 3.1. Reference architecture

The provided reference architecture is able to support all models of health participatory sensing described in Section 2. However, the emphasis in discussion is on the most comprehensive models of HPSN such as described in Sections 2.4 and 2.5. In the discussion of this reference architecture, we also describe and emphasize how HPSNs can achieve strict privacy and anonymity for all individual participants, via such technologies and techniques as MIX networks [22], k-anonymity, de-identification and submission of only ‘aggregate’ data.

The HPSN as a whole allows for the collection of various types of health-related data and various user interactions depending on the category of HPSN being adopted. The user will have granular control of the three types of interactions within the HPSN, as displayed in Fig. 1 which are: (1) data collection and submission, (2) micro surveys and (3) health interventions. At the coarsest level, each of these three functionalities can be disabled or enabled, with more specific options also being provided in each area. For example, a HPSN of the types described in Sections 2.1 and 2.2 would require no micro-survey or health intervention capabilities. An active HPSN as described in Section 2.4 would require data collection and health intervention capabilities, but not micro-surveys. An active HPSN with human-sensing as described in Section 2.5 would require all three capabilities.

Figure 1 shows a diagram of a conceptual reference architecture that would allow the higher levels of privacy required to support health participatory sensing, along with the distributed and multi-party capabilities of the overall platform. This architecture would also include a number of different public health organizations utilizing and sharing a single distributed HPSN server group that communicates to participants via MIX networks to preserve the anonymity of the participants. A public health organization/health organization would be defined as an organization involved in population data collection and/or public health intervention.

Data collection is the process of automated collection and submission of anonymized data. For the user there is the ability to customize what data submission policy providers (health organizations) they are subscribed to, as well as what types of data may be submitted. An example would be of opting-in to submission of physical activity data to one health organization, but physical activity and sleep patterns to a second organization. Additionally, this can be adjusted on an individual policy level-basis rather than on specific categories. In a similar way micro surveys and health interventions can be opted into, with specific categories of interventions and types of micro surveys being customizable. The distribution of data policies, micro surveys and health interventions takes this granular nature into account, only distributing the selected categories/policies to an individual's device.

Data policies, health interventions and micro surveys are distributed to an individual's device based on the customizations discussed above. This is achieved by the user mobile device periodically polling for updated policy sets based on the specific customization. However, beyond broad user customization for health interventions and micro surveys, once on the device only the most personalized/suited of those distributed is enacted on the mobile device, based on local processing so as to further maintain privacy (see Section 3.4). The best fit surveys/interventions are displayed/completed and responses then stored for data submission/analysis.

The functionality introduces a feedback loop – not for the individual but for a generalized group. Once data is submitted to the server, including the automated data collection, micro surveys and specific details of health interventions deployed, this can then be used to adapt the later policies, survey and health interventions distributed. This can have both automated and manually supervised aspects to it.

### *3.2. Distribution of data collection policy rules and micro surveys whilst protecting recipient anonymity and privacy*

The HPSN data collection methodology has a number of advantages due to being able to collect near real time, population-wide data and adjust the data collected when needed. As such, this requires a robust process for distribution of updated data collection policies, micro-surveys and health interventions, with the key requirements that updates be distributed in a timely and accurate manner. An addition to this, that is specific to anonymous data collection, is that of the critical need to preserve individual user privacy. The conventional approach would be to distribute personalized encrypted data updates through a MIX network. While this would meet the requirement for anonymous distribution, as no detail of the individual is known, it also provides the data policy owner with specific details of the policies distributed to individual participants. To allow for greater efficiencies and provide a higher level of privacy, a generic distribution approach using the MIX nodes themselves as distributed servers would provide an elegant solution.

To support the collection of anonymized data, distribution should be achieved through transmission of general policies packaged and distributed broadly. Additionally, the use of a multi-purpose HPSN with multiple public health organizations involved also raises the issue of quality assurance of the messages

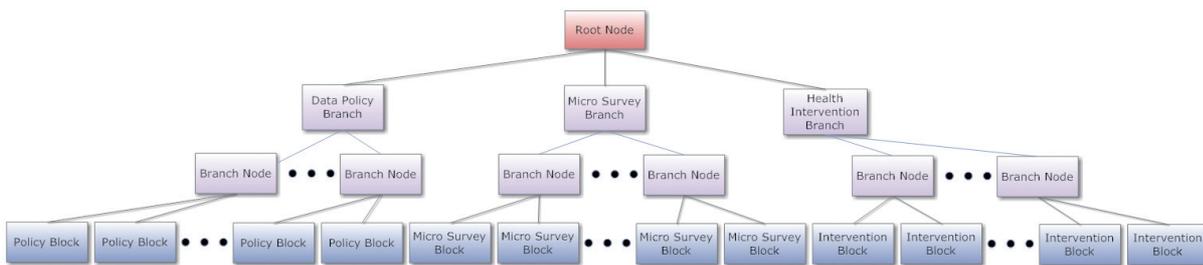


Fig. 2. Policy distribution verification tree structure.

throughout the network that may have come from various different senders and sent through a distributed multi-user network. As such, the completeness (all the requested data is retrieved), correctness (the data returned is accurate and has not been modified) and freshness (the data is the most up to date version available) need to be assured.

A granular approach utilizing verification objects constructed by hashing and digital signing of distributed content can assure the completeness and correctness of data. Additionally, including timestamps and expiry rates can assure the freshness of distributed data without direct communication to the associated public health organization. Our previous work found through implementation that user CPU and data overheads of this type of approach can be quite minimal [23], without significant additional overheads for the data owners/distributor.

This approach would additionally allow for dissemination and retrieval of data through the anonymous communications network with users retrieving policy updates and interventions relevant to them only, without breaching their anonymity.

To do this efficiently while improving privacy, we have considered the use of an adapted query assurance process suitable for distributed data sources [23]. This would allow for the nodes in the MIX network to cache general data and just distribute the requested subset to requests – without the policy owner being aware or involved in this efficient distribution.

Going into further detail, this approach uses signed verification objects distributed with policy data that allow the policy content to be verified in a granular way – that is, even if only a small portion of the policy data is delivered to an individual user, verification can be assured at reasonable levels of overhead. This is achieved by the use of sorted and signed merkle hash verification trees. Each policy block will be hashed and stored as a leaf node in the merkle hash tree with the branches to the root comprised of a hash of the branches/leaves below. Finally, the root is signed by the policy owner with a digital timestamp attached and included in the signed value. Due to the signed nature of the verification tree, the policy data distributed can be authenticated as correct. As the tree is sorted completeness can be assured by attaching border values with any policy distribution. Finally, due to the expiring timestamp, the distribution of old/expired policies can be detected and discarded.

In Fig. 2 the organisation of a policy distribution verification tree is displayed. It allows for many layers of branches depending on the size of the data set. Further, based on the data requested, a subset of this tree would be packaged as a verification object and distributed with the policy data to allow for verification. The verification tree contains three specific types of nodes -root, branch and leaf.

The verification tree structure is shown in Fig. 3 and contains four primary elements as follows:

$Root_{node}(S(T_s, H((Child_{node}1), \dots, (Child_{node}n-1), (Child_{node}n))), T_s)$ : Given S means the contents are digitally-signed,  $T_s$  is a time stamp with expiry duration, H means the contents are hashed and  $Child_{node}$  refers to the H value stored in  $Branch_{node}$  or  $Leaf_{node}$  directly below the root.

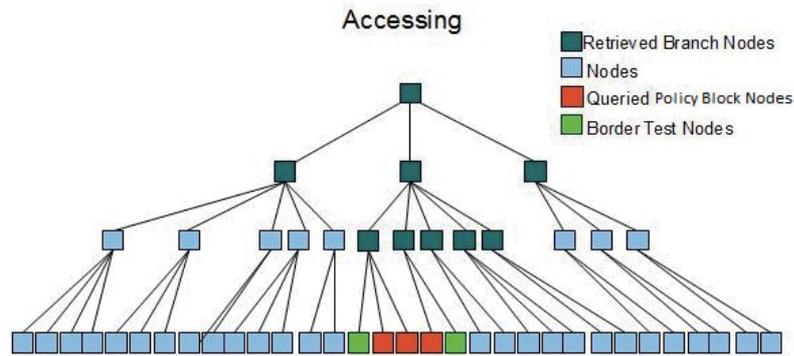


Fig. 3. Verification tree accessing.

$\text{Branch}_{\text{node}}(\text{H}(\text{Child}_{\text{node}1}, \dots, (\text{Child}_{\text{node}n-1}), (\text{Child}_{\text{node}n})), \text{Index}_{\text{value}})$ : Given H means the contents are hashed, Childnode refers to the H value stored in  $\text{Branch}_{\text{node}}$  or  $\text{Leaf}_{\text{node}}$  directly below the current branch and  $\text{Index}_{\text{value}}$  is the sort value of the branch.

$\text{BranchS}_{\text{node}}(\text{S}(\text{T}_s, \text{H}(\text{Child}_{\text{node}1}, \dots, (\text{Child}_{\text{node}n-1}), (\text{Child}_{\text{node}n}))), \text{T}_s, \text{Index}_{\text{value}})$ : Similar to the  $\text{Root}_{\text{node}}$ , the  $\text{BranchS}_{\text{node}}$  contains a digital signature and an expiring time stamp. Additionally, it also contains the  $\text{Index}_{\text{value}}$  as the sort value of the branch. The intention is that by providing signed branch nodes throughout the verification tree in proximity to often retrieved leaves or clusters of leaves, the overall efficiency of the verification scheme can be improved overtime based on usage.

$\text{Leaf}_{\text{node}}(\text{H}(\text{M}, \text{RecordPath}), \text{RecordPath}, \text{ReadCount}, \text{WriteCount})$ : Given H is the hash of its contents and M is the stored record. RecordPath is the direct path to the record.

### 3.3. Anonymous data collection and submission

The utilization of a methodology for anonymous submission of collected data is necessary in all implementations that do not incorporate a trusted server, and this capability alone should remove much of the potential and perceived privacy risks of a health participatory sensor network. Two such possible extant examples are MIX networks [22] (Fig. 1) and onion routing networks [24]. Either would be capable of allowing for anonymous submission with onion routing having some additional advanced capabilities. With the addition of an anonymous submission network the remaining primary privacy concern relates to the data submitted.

Secure de-identified approaches have gained specific coverage in the health context [25] due to the often highly sensitive nature of the data details. Conversely, such data has great importance as a source for research and epidemiology. Typically, de-identification works by removing or making less specific identifiers and quasi-identifiers. This leads to a k-anonymity type approach where as long as k number of individuals are indistinguishable from each other and do not have sensitive details in common, data privacy is considered to be preserved. This has been extended into the sensing context with spatial k-anonymity explored in [26]. However, assured k-anonymity approaches in most cases require a trusted server or aggregator to perform this analysis and decision making based on data received. This is hard to achieve in an anonymous submission distributed network that attempts to reduce the need for highly secure trusted components. The alternative is to make de-identification decisions locally without external knowledge of the potential k value of collected values, though this is not as assuredly secure as a trusted model. The advantage of reducing the sensitivity at the initial level of the mobile device before submission is promising.

A novel infrastructure approach is required to address the inherent property of detailed sensor data that even if de-identified could still act as a privacy risk through later re-identification of the individual with other known data. This risk can be reduced by submitting only less detailed and aggregate data [27]. In the public health domain this raises the question of ‘what is sufficient data for public health uses?’. It seems likely that more detailed information than is currently collected by traditional methodologies could be submitted without significant privacy risks, due to the very broad nature of population wide health measures as potential preventative and risk factors.

This is possible through storing detailed sensor data locally then performing anonymizing/aggregation on the data before submission. This includes removing any overly rare and identifying demographic information, making any temporal or location details more general and generating calculated aggregate measures based on the detailed data. This allows potentially long term analysis to be performed on the device and submitted without the server needing to know the sensitive detailed history.

The types of data that could be collected through this approach are quite varied, especially with the potential capabilities of extending the sensing component through micro surveys and feedback, and through future advancements of health sensor technologies. Some examples that are often of interest in previous large scale health data collection are:

- Physical Activity Patterns and Intensity – Due to its significance as a preventative factor in a number of lifestyle diseases, it is of high importance when considering a population-wide health participatory sensing model. The physical activity can usually be split between the following three categories.
  - \* Work Related Activity: The amount and intensity of physical activity completed during work.
  - \* Recreational Activity: Activity outside that associated with work or transportation.
  - \* Transportation Activity: Active transportation (walking, cycling or similar) as a form of physical activity that coincides with travel. Active transportation is a focus of public policy in many regions [28].
- Nutritional, Caloric Burn and Caloric Intake – This type of data could provide more detailed information on overall energy expenditure and nutritional intake of individuals, segments of the population and the population overall [29].
- Body Mass Index (BMI) and change over time – This would allow for a current snapshot of BMI, with the potential for trend analysis and based on the individual change over time.
- Sleep Patterns and Regularity – Sleep patterns are both an indicator and a preventative/risk factor for a number of conditions.

As indicated, all these types of data if submitted are submitted using the fully privacy-preserving and anonymization mechanisms described, and only population aggregate measures are stored at the public health organizations for population health purposes.

### 3.4. Public health interventions

A major area of potential usefulness of HPSNs is the ability to distribute targeted or personalized public health interventions to individuals (see Sections 2.4 and 2.5). It may appear that to do so implicitly requires potentially identifying details to be known by the server, which may affect willingness and motivation to participate in such a network.

However, the sensitivity of health related information and the capability of the proposed reference architecture suggests a novel approach to non-identifying targeted public health interventions.

Additionally, it seems likely that there will be a number of different public health organizations that would be interested in participating in these types of networks with individuals able to subscribe or opt-in to partake in passive or active participation with each individual organization.

The novel approach introduced, involves broadcasting larger generalized public health intervention packages to the entirety of the participants or subsets and then based on local processing, the correct information or intervention information is displayed or actioned on individual devices. This would allow for communication with individuals that could be meaningful and personalized without risk of re-identification of the individual. This approach could also be used for the dissemination of micro-surveys to individuals for data collection for the applicable models of HPSN.

### 3.5. HPSN and participant device interaction

The participant device will have the following capabilities to interact with the network (Fig. 4):

- De-identification – This component manages the de-identification of data prior to submission by making specifics of the data less sensitive and individually identifying if necessary. This is actioned by a consideration of the organization subscription policies which will identify what data is to be submitted and at what particular level of detail.
- Micro surveys – An opt-in additional component (see Sections 2.3 and 2.5) that would allow the public health organizations to enact additional data collection by context-aware micro surveys to complement or enhance the sensor data collection.
- Local aggregate processing – Calculates aggregate measures [27] according to rules delivered through the participatory network. This also takes into account local rules and is conscious of potential re-identification risks when allowing data submission.
- Intervention Delivery – An opt-in additional component that retrieves general interventions targeted at particular demographic groups (see Sections 2.4 and 2.5), and these interventions are locally processed according to the individual's data and present the correct targeted intervention to the end user.
- Subscription policies – This component will allow an individual to control which public health organizations they interact with and the opt-in and opt-out of capabilities and collection offered. For example, an individual may choose to collect data and receive interventions from an organization X, while only collecting data and opting out of interventions from organization Y. This could be provided to a detailed granular level where individuals can collect for only a specific measure or receive communication on a very narrow topic.
- Sensing policies – Local policies defined by the user that set limitations as to what and how they will participate in the HPSN. This would include which sensors are used, battery management policy, time and detail limitations, amongst others.

## 4. User interaction

For our example purposes, let us consider user David who is a participant in a HPSN, having subscribed to the Department of Health, and to the Active Transport Initiative Inc data collection policies. Additionally, he has subscribed to micro surveys from Active Transport Inc and health interventions from the Department of Health.

The subscribed to policies, micro-surveys and health interventions are updated periodically. This is controlled by either the client checking for updates, with the maximum valid period of a set of policies

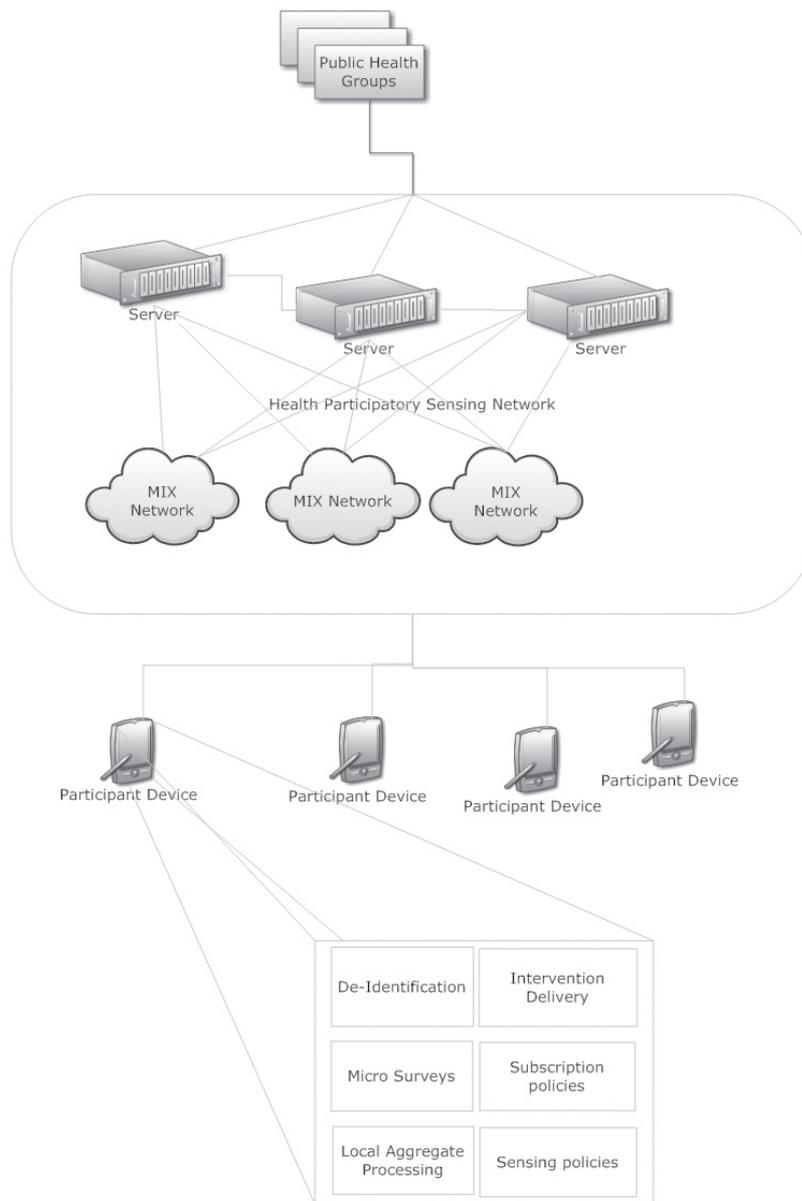


Fig. 4. Network and participant device communication.

set by an expiring timestamp, one referring to the distribution process (maximum time before the data should no longer be distributed) and a second referring to the expected validity period (maximum period before the policy needs to be updated).

Throughout his daily schedule his mobile device automatically collects physical activity data and relevant data to each organization is submitted intermittently throughout the day utilizing the privacy-preserving mechanisms of the HPSN. Again this is totally anonymous and does not allow re-identification – it can just provide a valuable input when combined with the mass of other individual’s data, for population health measures.

More specifically the Department of Health is interested in overall physical activity in a day with age bracket and coarse location information also submitted. Alternatively, Active Transport Initiative Inc is only concerned with physical activity related to transportation (e.g. walking/cycling commuting) with the additional data of age bracket and start coarse location and end coarse location submitted. Occasionally during the day/very infrequently, David is prompted to complete a micro-survey related to active transport. This is a 30 second survey for example asking for a ranking of the five most significant factors as to whether on a specific day he would cycle/walk to work. This micro-survey was presented to David based on the locally stored data relating to his travel habits – although it would have been sent to a much larger group. As an average two day per week cycle commuter, he is a prime target for increase in active transport modal share. David's preferences also restrict how often micro-surveys from individual organizations and the overall HPSN system can request micro-surveys.

Based on his personal preferences and previous trend data, David is also prompted by a health intervention from the Department of Health, suggesting the health benefits of cycling. This suggestion is computed on his local device based on the typical time of day that David exercises, current weather patterns including UV rating and potential vitamin D intake levels. David chooses not to take up the suggestion of this intervention, triggering a wait-time (set by David's preferences) as to how soon a new intervention can be received.

David is also interested to track his own health-related data and finds this a benefit that also assists his motivation to participate in the HPSN. The data he displays to himself for this purpose is kept securely on his own device, and is never transmitted to the HPSN – for this reason this data for self-use can be more detailed and can be viewed by David without aggregation processing [27] having first occurred if desired.

## 5. Conclusion

The utilization of health participatory sensing networks for population health and epidemiological data collection is a development that can greatly increase the capability and scale of health data collection.

The challenges to this approach however are not insignificant, with privacy and incentives to participate, core concerns affecting the level of participation and therefore potential data collection. In this paper, we have introduced the first classification of health participatory sensing networks, and discussed these classifications with particular reference given to the required level of effort for individuals, privacy concerns and potential realizable benefits and incentives.

In this work we have also provided a conceptual reference architecture that supports the capabilities required for the more highly detailed health participatory sensing categories, that includes anonymous submission and secure de-identification – this would have significant advantages that may positively affect participation, as compared to a more traditional trusted server methodology. We have also described a user interaction case study.

## References

- [1] J.A. Burke, D. Estrin, M. Hansen, A. Parker, N. Ramanathan, S. Reddy and S.M. B., Participatory sensing, presented at the World Sensor Web Workshop, ACM Sensys, Boulder, Colorado, 2006.
- [2] R. Kwok, (2009, 20 April). *Personal technology: Phoning in data*. Available: <http://www.nature.com/news/2009/090422/full/458959a.html>.

- [3] Gartner, (2011, April 20). *Gartner Says Worldwide Smartphone Sales Soared in Fourth Quarter of 2011 With 47 Percent Growth*. Available: <http://www.gartner.com/it/page.jsp?id=1924314>.
- [4] ECF. (2012, July 10). *Track that Bike: How Can Electronic Chips Boost Cycling?* Available: <http://www.ecf.com/news/track-that-bike-how-can-electronic-chips-boost-cycling/>.
- [5] Bicycle Network. (2011). *Riderlog*. Available: <http://www.bv.com.au/general/ride-to-work/91481/>.
- [6] Jawbone. (2011, April 20). *UP® by Jawbone® with MotionX® Technology Empowers You to Live a Healthier Life*. Available: <http://content.jawbone.com/static/www/pdf/press-releases/up-press-release-110311.pdf>.
- [7] C. Outram, C. Ratti and A. Biderman, The Copenhagen Wheel: An innovative electric bicycle system that harnesses the power of real-time information and crowd sourcing, presented at the EVER Monaco International Exhibition and Conference on Ecologic Vehicles and Renewable Energies, 2010.
- [8] K. Yatani and T.N. Khai, BodyScope: A Wearable Acoustic Sensor for Activity Recognition, presented at the Proceedings of the 14th international conference on Ubiquitous computing, Pittsburgh, USA, 2012.
- [9] R. Steele, An Overview of the State of the Art of Automated Capture of Dietary Intake Information, *Critical Reviews in Food Science and Nutrition*, 2013.
- [10] S. Passler and W.J. Fischer, Acoustical method for objective food intake monitoring using a wearable sensor system, in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2011 5th International Conference on*, 2011, pp. 266–269.
- [11] J. Lester, D. Tan, S. Patel and A.J. Brush, Automatic classification of daily fluid intake, presented at the Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2010 4th International Conference on, 2010.
- [12] K.-H. Chang, S.-Y. Liu, H.-H. Chu, J.Y.-J. Hsu, C. Chen, T.-Y. Lin, C.-Y. Chen and P. Huang, The diet-aware dining table: observing dietary behaviors over a tabletop surface, presented at the Proceedings of the 4th international conference on Pervasive Computing, Dublin, Ireland, 2006.
- [13] G. Villalobos, R. Almaghrabi, B. Hariri and S. Shirmohammadi, A personal assistive system for nutrient intake monitoring, presented at the Proceedings of the 2011 international ACM workshop on Ubiquitous meta user interfaces, Scottsdale, Arizona, USA, 2011.
- [14] P. Ming-Zher, K. Kyunghee, A.D. Goessling, N.C. Swenson and R.W. Picard, Heartphones: Sensor Earphones and Mobile Application for Non-obtrusive Health Monitoring, in *Wearable Computers, 2009. ISWC '09. International Symposium on*, 2009, pp. 153–154.
- [15] P. Klasnja, S. Consolvo, D.W. McDonald, J.A. Landay and W. Pratt, Using mobile and personal sensing technologies to support health behavior change in everyday life: lessons learned, *AMIA Annu Symp Proc*, 2009, pp. 338–342.
- [16] A. Clarke and R. Steele, How personal fitness data can be re-used by smart cities, presented at the Seventh International Conference on Intelligent Sensors, Sensor Networks and Information Processing, Adelaide, Australia, 2011.
- [17] P. Klasnja, S. Consolvo, T. Choudhury, R. Beckwith and J. Hightower, Exploring Privacy Concerns about Personal Sensing, presented at the Proceedings of the 7th International Conference on Pervasive Computing, Nara, Japan, 2009.
- [18] L. Juong-Sik and H. Baik, Sell your experiences: A market mechanism based incentive for participatory sensing, in *Pervasive Computing and Communications (PerCom), IEEE International Conference on*, 2010, pp. 60–68.
- [19] I. Krontiris and N. Maisonneuve, Participatory Sensing: The Tension Between Social Translucence and Privacy, in *Trustworthy Internet*, L. Salgarelli, G. Bianchi and N. Blefari-Melazzi, eds, ed: Springer Milan, 2011, pp. 159–170.
- [20] R. Steele and A. Clarke, A Real-time, Composite Healthy Building Measurement Architecture Drawing Upon Occupant Smartphone-collected Data, presented at the 10th International Healthy Buildings Conference, 2012.
- [21] J. Rula, Fabi, #225 and N.E. Bustamante, Crowd (soft) control: moving beyond the opportunistic, presented at the Proceedings of the Twelfth Workshop on Mobile Computing Systems San Diego, California, 2012.
- [22] K. Sampigethaya and R. Poovendran, A Survey on Mix Networks and Their Secure Applications, *Proceedings of the IEEE* **94** (2006), 2142–2181.
- [23] A. Clarke and R. Steele, Secure and Reliable Distributed Health Records: Achieving Query Assurance across Repositories of Encrypted Health Data, in *System Science (HICSS), 2012 45th Hawaii International Conference on*, 2012, pp. 3021–3029.
- [24] S. Mauw, J.H.S. Verschuren and E.P. Vink, A Formalization of Anonymity and Onion Routing, in: *Computer Security – ESORICS 2004*, (vol. 3193), P. Samarati, P. Ryan, D. Gollmann and R. Molva, eds, Springer Berlin Heidelberg, 2004, pp. 109–124.
- [25] K. El Emam, Risk-Based De-Identification of Health Data, *Security and Privacy, IEEE* **8** (2010), 64–67.
- [26] P. Kalnis and G. Ghinita, Spatial k-Anonymity, in: *Encyclopedia of Database Systems*, L. Liu and M.T. Özsu, eds, Springer US, 2009, pp. 2714–2714.
- [27] A. Clarke and R. Steele, Summarized data to achieve population-wide anonymized wellness measures, in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, 2012, pp. 2158–2161.
- [28] F. Racioppi, C. Dora, R. Krech and O.V. Ehrenstein, (2002). *A physically active life through everyday transport* Available: [http://www.euro.who.int/\\_\\_data/assets/pdf\\_file/0011/87572/E75662.pdf](http://www.euro.who.int/__data/assets/pdf_file/0011/87572/E75662.pdf).

- [29] R. Steele, An Overview of the Role of Informatics-based Systems in Furthering an Integrated Paddock to Plate Food Supply System, presented at the Progress in Industrial Ecology, 2013.
- 

**Andrew Clarke** is currently pursuing a PhD in Health Informatics at the University of Sydney, having previously completed a Bachelor of Computer Network Engineering at La Trobe University.

**Professor Robert Steele** is the Head of Discipline and Chair of the Discipline of Health Informatics at The University of Sydney and the Director of the Health Informatics, Computation and Innovation Lab. He is the author of over 100 refereed articles and his professional service includes the current role of Vice Chair of ACM SIGMOBILE.



# Hindawi

Submit your manuscripts at  
<http://www.hindawi.com>

