Capítulo 2

Segurança em Mobile Crowd Sensing

Joélisson Joaquim de V. Laurido e Eduardo Luzeiro Feitosa

Abstract

Mobile Crowd Sensing (MCS) applications allow that participants or users to collect (using different sensors on your mobile devices - microphone, camera, GPS, etc.) and share information in order to assist other users in decision-making (based on third-party opinion) or inform about different events. Examples of MCS applications include coverage of instantly news, traffic congestion, weather disasters, air pollution, parking spaces, among others. Once the operation of MCS applications involves human or mechanical participation (cars, drones, sensors, etc.) to collect data that will be used by end users, the privacy and safety of the participants, not to mention the reliability of generated and received information are important issues that need investigation. In this context, this chapter aims to provide understanding of the environment of MCS applications, focusing mainly on issues related to privacy, security and reliability of the information.

Resumo

Aplicações de Mobile Crowd Sensing (MCS) são aquelas onde os participantes ou usuários coletam (através dos diferentes sensores em seu dispositivo móvel - microfone, câmera, GPS, entre outros) e compartilham informações com o intuito de auxiliar outros usuários na tomada de decisões (baseada na opinião de terceiros) ou informar sobre os mais diversos acontecimentos. Entre exemplos comuns de aplicações de MCS pode-se citar a cobertura de notícias instantaneamente, informações sobre pontos de congestionamento, desastres climáticos, poluição do ar, buracos em vias públicas, vagas em estacionamentos, entre outros. Uma vez que o funcionamento de aplicações MCS envolve a participação humana ou mecânica (automóveis, drones, sensores, entre outros) para coletar os dados que serão usados por usuários finais, a privacidade e a segurança dos participantes, sem falar na confiabilidade das informações geradas e recebidas são questões importantes que precisam investigadas. Neste contexto, este Capítulo objetiva fornecer entendimento sobre o ambiente de aplicações de Mobile Crowd Sensing, focando principalmente nas questões relacionados a privacidade, a segurança e a confiabilidade das informações.

2.1. Introdução

Há algum tempo, o processo de sensoriamento passou a ser utilizado na gestão de cidades, através da monitorização de áreas urbanas e da observação da dinâmica das comunidades, visando a fornecer aos gestores informações essenciais para a tomada de decisão sobre os mais variados assuntos. Atualmente, esse processo de gestão é parte essencial da vida nas cidades. Por exemplo, o sensoriamento de fatores ambientais permite que autoridades ou agências obtenham dados e informem a população sobre condições de tráfego, poluição sonora, poluição do ar, qualidade da água, segurança pública, entre outros assuntos, informando o que acontece, quando acontece e o que fazer quando algo acontecer. Mas como fazer esse sensoriamento?

As atuais e tradicionais técnicas de sensoriamento, tais como as redes de sensores sem fio (RSSF), vêm sendo muito aproveitadas para adquirir as "condições" do mundo real. Contudo, as redes de sensores comerciais nunca foram implantadas com sucesso no mundo real, devido a problemas como a cobertura insuficiente, falta de escalabilidade e, principalmente, o custo de instalação e manutenção.

Graças ao exponencial crescimento do poder computacional e da quantidade e disponibilidade de sensores embutidos nos dispositivos móveis usados diariamente, é possível,nos dias atuais, que o cidadão comum possa monitorar aspectos sobre a cidade, sua comunidade ou uma região e, assim, contribuir de alguma forma para melhorar e/ou auxiliar a vida das outras pessoas. Entre os exemplos óbvios estão o nível de ruído na cidade e o fluxo de tráfego. Também é possível incluir situações que interferem na qualidade de vida, como buracos em vias públicas e estradas, os quais podem ser detectados usando dispositivos móveis.

Recentemente, um novo paradigma de sensoriamento vem tirando proveito dessa vasta gama de recursos para monitorar e compartilhar informações de interesses comum, coletadas através dos sensores embutidos nos dispositivos móveis, o que acaba por auxiliar as pessoas na tomada de decisões. Denominado de *Mobile Crowd Sensing*, ou simplesmente MCS, esse paradigma parte do princípio de que é possível (e até fácil em certo ponto) utilizar os mais variados sensores (câmeras, acelerômetros, sistemas de posicionamento global - GPS, sensores de temperatura, entre muitos outros) presentes em *smartphones*, dispositivos de IoT (I*nternet of Thing*), tocadores de músicas, consoles de jogos (Wii e XboX Kinect, por exemplo) e veículos (GPS, computadores de bordo) para coletar informações relevantes para uma cidade ou uma comunidade.

As aplicações MCS vêm sendo utilizadas para o monitoramento do ruído nas cidades [Maisonneuve et al. 2010, Rana et al. 2010], poluição ambiental [IBM 2010] e fenômenos climáticos [Thepvilojanapong et al. 2010], medição da densidade demográfica [Weppner and Lukowicz 2013], cenários de emergências [Ludwig et al. 2015], medicina [Lane et al. 2010], anomalias no tráfego [Pan et al. 2013, Ganti et al. 2011] e até mesmo detectar terremotos [Minson et al. 2015].

Um aspecto importante é que as aplicações MCS vêm ganhando popularidade com a criação de diversos sistemas e aplicativos e, consequentemente, conquistando o envolvimento de mais e mais pessoas, redes e grupo de colaboradores. Em conjunto com outras tecnologias e plataformas, como rede sociais e armazenamento em nuvem,

as MCS têm desempenhado um papel imprescindível na integração entre pessoas que buscam informações ajustadas para tomar decisões que possam ajudá-las.

O diferencial de MCS em relação a outras abordagens de sensoriamento com participação de usuários reside no fato de que toda ação de registro (coleta de dados e informações) parte do dispositivo móvel do usuário participante do serviço, conectado a qualquer rede de acesso à Internet, muitas vezes sem a necessidade de registro em tempo real [Wang et al. 2013]. Assim, o custo de implantação e utilização torna-se bastante reduzido, uma vez que não se faz necessário construir uma infraestrutura específica, como ocorre em sensores convencionas. Serviços e aplicações de MCS podem usar a infraestrutura das próprias operadoras de telefonia móveis bem como redes domésticas e/ou empresarias através de WiFi e 3G Além disso, quanto maior o número de participantes mais tarefas podem ser disponibilizadas e mais áreas pode ser cobertas [Tuncay et al. 2012], o que acaba por aumentar a confiabilidade das informações, permitindo que que elas sejam mais facilmente aceitas.

Por outro lado, o uso de MCS também traz seus riscos. O principal deles é a privacidade [Ganti et al. 2011, Lane et al. 2010, Kong et al. 2015]. O cenário a seguir exemplifica melhor o problema. Considerando um sistema MCS cujo objetivo é recolher imagens ou vídeos curtos sobre irregularidades no trânsito em diferentes partes de uma cidade, qualquer pessoa que deseja participar precisa, primeiramente, enviar uma consulta ao servidor de aplicativos para descobrir o conjunto de locais próximos a ela que ainda necessitam de coleta de dados. Se o participante não estiver disposto a divulgar sua identidade, o servidor pode até remover a identificação (ID) da consulta, mas ainda precisa saber as informações de localização do participante, a fim de responder à consulta. Assim, devido a essa forte correlação entre os participantes e seus movimentos, um servidor malicioso pode identificar um participante através de sua informação de localização.

Diante do exposto, o objetivo deste Capítulo é introduzir o leitor no paradigma de *Mobile Crowd Sensing*. A ideia é apresentar os principais conceitos e definições sobre o tema. Contudo, o foco principal é voltado para a questão de segurança (privacidade e confiabilidade) das informações coletadas e transmitidas pelos participantes de MCS. Além de possibilitar a compreensão dos problemas de segurança em MCS e apresentar as soluções existentes, este Capítulo também identifica oportunidades de estudos na área, bem como aponta alguns trabalhos futuros.

Para alcançar os objetivos propostos, o restante deste Capítulo está organizado da seguinte forma. A Seção 2.2 caracteriza *Mobile Crowd Sensing*, apresentando conceitos, definições e características únicas, bem como discutindo a evolução dos esquemas de sensoriamento participativo até chegar às MCS e apresentando exemplos de aplicações reais. A Seção 2.3 trata dos aspectos de segurança em MCS, onde são discutidos os problemas de privacidade dos participantes e de confiabilidade dos dados. Os problemas causados e as técnicas empregadas como contramedidas a esses problemas são relatadas. A Seção 2.4 apresenta algumas soluções (arquiteturas, aplicações, bibliotecas e ferramentas) existentes que levam em consideração a segurança em MCS ou permitem que sejam empregadas nos aspectos de segurança em MCS. A Seção 2.5, por sua vez, aponta algumas questões em aberto. Por fim, a Seção 2.6 apresenta os comentários finais sobre o tema.

2.2. Mobile Crowd Sensing

Como já mencionado na Seção anterior, *Mobile Crowd Sensing* (MCS) apresenta um novo paradigma de sensoriamento baseado no poder dos dispositivos móveis. Ele tira proveito do grande número de dispositivos que "acompanham" um usuário (telefones celulares, dispositivos portáteis e veículos inteligentes) e de sua mobilidade para adquirir conhecimento local e compartilhar esse conhecimento dentro de uma esfera social.

Esta Seção discute o processo evolutivo do sensoriamento, traz definições sobre tais aplicações, enumera os elementos arquiteturais de sistemas MCS, apresenta as características únicas que o paradigma proporciona e termina exemplificando algumas aplicações MCS.

2.2.1. Da Sabedoria das Multidões a MCS

Para entender o paradigma de MCS, é preciso, primeiramente, compreender a ideia de resolver problemas usando populações como fonte de informação (definido em inglês como *crowd-powered problem-solving*).

Em seu livro "The Wisdom of Crowd" (Sabedoria das Multidões), Surowiecki [Surowiecki 2005] revela que a agregação de dados ou informações a partir de um grupo de pessoas resulta, muitas vezes, em decisões melhores do que as feitas por um único indivíduo . Na mesma linha de raciocínio da sabedoria das multidões, existe o conceito de inteligência coletiva, um tipo de "saber compartilhado" oriundo da colaboração de muitas pessoas em suas diversidades. Segundo Levy [Levy and da Costa 1993], a inteligência coletiva é aquela distribuída por toda parte, na qual todo o saber está na humanidade, já que ninguém sabe tudo, porém todos sabem alguma coisa.

Com base nos conceitos de sabedoria das multidões e inteligência coletiva, duas formas de aproveitar a participação de multidões, intrinsecamente ligados a MCS. A primeira delas é *Crowdsourcing*. Definido por Jeff Howe e Mark Robinson em [Howe 2006], *Crowdsourcing é o ato de uma empresa ou instituição, tendo uma função uma vez realizada por funcionários, terceirizá-la a uma rede indefinida (e geralmente grande) de pessoas sob a forma de um convite aberto*. *Crowdsourcing* é um modelo de produção que utiliza a inteligência coletiva, a cultura colaborativa e a formação de comunidades para solucionar problemas, criar conteúdo ou buscar inovação. Um exemplo típico é a Wikipedia, a maior enciclopédia do mundo, criada por milhares de contribuidores, de forma colaborativa.

A segunda forma é o *Sensoriamento Participativo* [Burke et al. 2006], que sustenta a ideia que cidadãos e dispositivos móveis podem formar redes de sensores participativas para aquisição e compartilhamento de conhecimento local, enfatizando a participação explícita do usuário. *Sensoriamento Participativo* requer um ativo envolvimento dos indivíduos para contribuir com dados sensoriados (uma foto, um relato sobre uma via pública interditada) de acordo com o fenômeno monitorado.

Embora as diferenças entre *Crowdsourcing*, *Sensoriamento Participativo* e MCS possam não ser facilmente percebidas, elas existem. Todas as três formas de sensoriamento assumem que existe a participação de multidões, uma vez que quanto maior número de participantes, mais dados podem ser sensoriados, maiores áreas podem ser

cobertas e o que aumenta a confiabilidade das informações coletadas.

Contudo, além da coleta de dados, MCS também permite (e algumas vezes exige) que o participante atue no processamento das informações coletadas. Além disso, MCS aceita a participação oportunista [Lane et al. 2010], onde é necessário apenas que o aplicativo capture os dados através dos sensores disponíveis, e a participação mecânica (carros inteligentes, drones, entre outros) na atividade de coleta [Konidala et al. 2013].

A Figura 2.1 ilustra a abrangência de MCS em comparação aos conceitos de sabedoria das multidões, *crowdsourcing* e sensoriamento participativo.

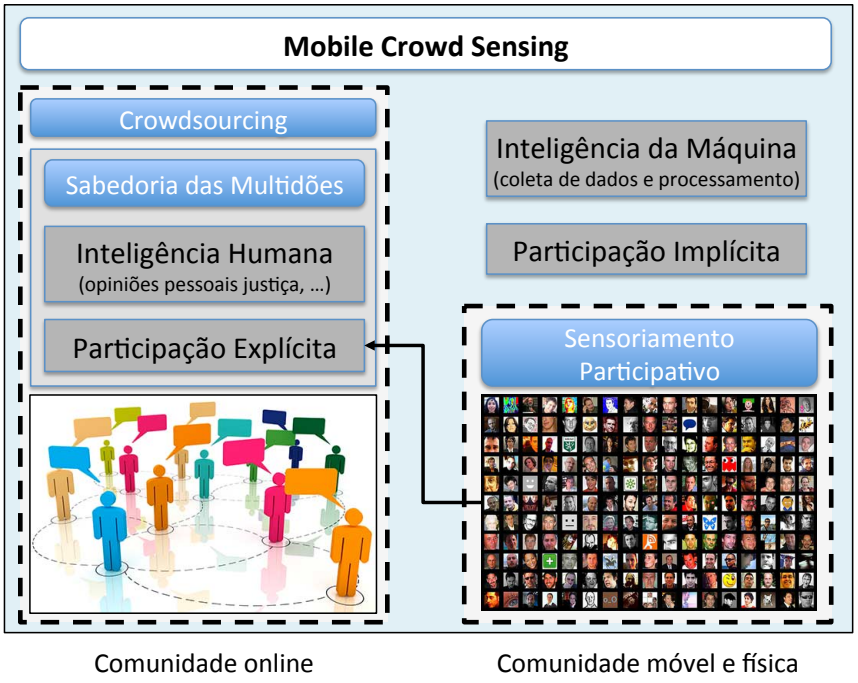

Figura 2.1. Comparação entre MCS e os conceitos relacionados. Adaptado de [Guo et al. 2015]

Percebe-se na Figura 2.1 que enquanto a sabedoria das multidões e *crowdsourcing* contam apenas com a inteligência humana, o sensoriamento participativo e MCS exploram uma fusão da inteligência humana e da máquina.

Segundo Hu et al. [Hu et al. 2013], em comparação com sistemas e redes de sensores fixos, com grande complexidade estrutural e logística, MCS possui vantagens como:

- **Generalidade**, uma vez que atendem a diversidade de sistemas operacionais e hardwares, ignorando nuances e generalizando o desenvolvimento de aplicações necessárias e populares;
- **Escalabilidade**, pois não limitam a quantidade de usuários;
- **Mobilidade de conteúdo**, já que as informações coletadas podem ser passadas por diversos canais (redes sociais, canais de rádio, jornais online e televisão, por exemplo), antes de serem acessadas/vistas, não havendo barreiras para as informações transmitidas.

2.2.2. Definições

Como já mencionado neste Capítulo, *Mobile Crowd Sensing* (MCS) apresenta um novo paradigma de sensoriamento baseado no poder dos dispositivos móveis.

Para Ganti et al. [Ganti et al. 2011], os pesquisadores responsáveis por cunharem o termo, *Mobile Crowd Sensing se refere a uma variada e ampla gama de modelos de sensoriamento no qual indivíduos com dispositivos computacionais e sensores são capazes de coletar e contribuir com dados valiosos para diferentes aplicações*. Em linhas gerais, MCS tira proveito do grande número de dispositivos que "acompanham" um usuário (telefones celulares, dispositivos portáteis e veículos inteligentes) e de sua mobilidade para adquirir conhecimento local e compartilhar esse conhecimento dentro de uma esfera social. A informação recolhida mais a fusão dos dados coletados (que acontece com o apoio da computação em nuvem), torna MCS uma plataforma versátil que muitas vezes pode substituir as infraestruturas de sensoriamento estáticos, permitindo uma ampla variedade de aplicações.

Guo et al. [Guo et al. 2015] definem MCS como "*um novo paradigma de detecção que capacita os cidadãos comuns à contribuir com dados sensoriados ou gerados a partir de seus dispositivos móveis, agregados e fundidos na nuvem para a extração de inteligência da multidão, prestando serviços centrados nas pessoas*". Já autores como [Dimov 2014] definem MCS como sendo um novo modelo de negócios do conceito de *crowdsourcing*, uma vez que permite que um grande número de dispositivos móveis sejam usados não somente para troca de informações entre seus usuários, mas também para atividades que podem ter um grande impacto social.

Quanto ao que pode ser sensoriado com MCS, Ganti et al. [Ganti et al. 2011] afirmam que tudo depende do fenômeno que se está medindo. Eles sugerem a seguinte classificação:

- **Ambiental**, onde fenômenos de natureza ambiental são os alvos. Exemplos incluem o nível e qualidade da água em rios [IBM 2010, Minkman et al. 2015], a medição dos níveis de poluição em cidades [Dutta et al. 2009, Leonardi et al. 2014] e o monitoramento de vida selvagem em seu habitat [Mediated Spaces, Inc 2015].

- **Infraestrutura** envolve a medição em larga escala de fenômenos relacionados a infraestrutura pública. Exemplos incluem a disponibilidade de vagas em estacionamentos [Mathur et al. 2010], a condição de rodovias e estradas, problemas com aparelhos públicos (hidrantes, iluminação, entre outros) e o monitoramento do trânsito em tempo real [Mohan et al. 2008, Google 2015].

- **Social**, onde os participantes coletam e compartilham informações entre si. Exemplos incluem dados sobre os exercícios [Reddy et al. 2007] que realizaram durante o dia ou quanto andaram de bicicleta [Eisenman et al. 2007].

2.2.3. Componentes Arquiteturais

Embora não exista uma arquitetura formal, visto que ela depende do que se quer sensoriar, a Figura 2.2 mostra uma arquitetura genérica de MCS.

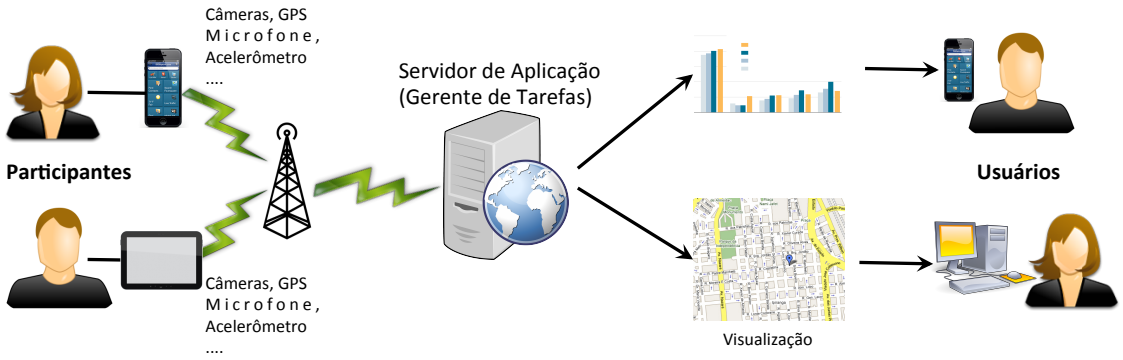

Figura 2.2. Exemplo de arquitetura geral em MCS. Adaptado de [Pournajaf et al. 2014]

Entretanto, vários autores [Pournajaf et al. 2014, Ganti et al. 2011] concordam que qualquer sistema MCS deve possuir pelos menos três componentes chave:

- **Participantes**: São as entidades, pessoas, que usam algum tipo de sensor para obter ou medir a informação requisitada sobre um elemento de interesse;
- **Aplicações ou Usuários Finais**: São as entidades que requisitam dados através de tarefas e então utilizam a informação coletada pelos participantes;
- **Servidor de Aplicação**: Também chamado de gerentes de tarefas, são as entidades responsáveis pela distribuição de tarefas aos participantes que conseguem atender aos requisitos das aplicações. O modo de distribuição de tarefas do servidor de aplicação varia de acordo com as tarefas e os participantes, e pode ser categorizado em:
  - **Centralizado**: Um servidor central fornece aos participantes diferentes tarefas para executar. Por exemplo, em um aplicativo que mede o entusiasmo de festas (*party thermometer*), um servidor central pode escolher um conjunto de participantes, pedindo que eles classificam uma determinada festa. Uma questão importante do modelo centralizado é ter um único ponto de falha para interações entre participantes e aplicações;
  - **Descentralizado**: Cada participante pode se tornar um gerente de tarefa e decidir ou executar uma tarefa ou passá-la para outros participantes que possam estar melhor adaptados para cumprí-la. Esta decisão pode ser baseada em certos atributos de outros participantes, como localização, habilidades ou hardware disponível no dispositivo. Um bom exemplo é o modelo de recrutamento descentralizado proposto em [Tuncay et al. 2012], que notifica os participantes qualificados de uma atividade de sensoriamento próxima;
  - **Híbrido**: Neste esquema, um servidor central e um conjunto de participantes atuam na construção do núcleo de gerenciamento de tarefas. Um bom exemplo é o esquema de bolha [Lu et al. 2010], que requer um servidor central para manter o controle das tarefas de sensoriamento, que são alocadas principalmente de forma descentralizada.

2.2.4. Ciclo de Vida de MCS

De acordo com Zhang et al. [Zhang et al. 2014a], o ciclo de vida em MCS consiste: (1) na criação de aplicativos de acordo com os requisitos; (2) na atribuição de tarefas de sensoriamento para os participantes; (3) na execução da tarefa (sensoriamento, computação e upload) no dispositivo móvel do participante, e (4) coleta e processamento dos dados enviados pelos participantes (Figura 2.3).

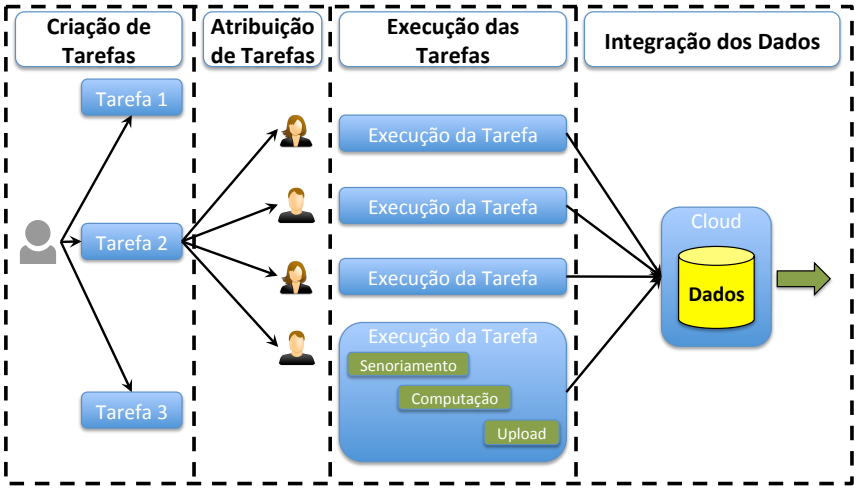

Figura 2.3. Ciclo de vida em MCS. Adaptado de [Zhang et al. 2014a]

As funcionalidades fundamentais de cada fase são descritos como:

- **Criação de tarefas**: O organizador/gerenciador MCS define uma tarefa de acordo com as necessidades estabelecidas e elabora uma aplicação MCS que deve ser fornecida aos participantes.

- **Atribuição de tarefas**: Após a tarefa MCS e a aplicação móvel serem criadas, a próxima fase é atribuir a tarefa, ou seja, recrutar participantes e lhes atribuir tarefas de sensoriamento individuais que são supostamente para serem executadas no dispositivo móvel de cada participante. Encontrar participantes suficientes e adequados para o sensoriamento é a questão central neste estágio.

- **Execução de tarefas individuais**: Depois de receber a tarefa, o participante deve tentar terminá-la dentro do período pré-definido. Esta fase pode ser dividida em 3 sub-fases - sensoriamento, computação e *upload* dos dados.

- **Integração de dados**: Esta etapa recebe os fluxos de dados coletados de todos os participantes como entrada, agrega-os e fornece aos usuários finais no formato adequado. Para algumas aplicações MCS [Sherchan et al. 2012], o processamento de dados nesta fase é bastante simples. Um servidor armazena os dados centrais e fornece interfaces para usuários consultarem e partilharem os dados. Enquanto outros aplicativos MCS [Mun et al. 2009, Rachuri et al. 2011] empregam algoritmos complicados para integrar dados e extrair alto nível de inteligência coletiva a partir dos dados brutos.

2.2.5. Características Únicas

De acordo com Guo et al. [Guo et al. 2015], MCS possui quatro características chave que a diferenciam das outras abordagens. São elas.

Sensoriamento baseado em Multidões

Em comparação com redes de sensores tradicionais, a diferença fundamental de MCS é o envolvimento de multidões para o sensoriamento em grande escala, o que oferece duas grandes vantagens. A primeira é aproveitar a capacidade de sensoriamento nos dispositivos dos participantes bem como a infraestrutura de comunicação existente, facilitando a implantação e reduzindo custos. A segunda é aproveitar a mobilidade inerente dos usuários de dispositivos móveis, o que oferece uma cobertura espaço-temporal sem precedentes comparada às implementações de redes de sensores estáticas.

Além disso, o sensoriamento através de multidões de MCS oferece novos recursos. O primeiro deles é o **modo de geração de dados**. MCS pode reunir (sensoriar) dados de dois modos diferentes:

1. **Sensoriamento Móvel**, onde o processo de coleta é baseado no contexto individual, ou seja, o sensoriamento é feito a partir dos dispositivos dos indivíduos [Lane et al. 2010];

2. **Redes Sociais Móveis** (*Mobile Social Networks* ou MSN), onde os dados postados pelos usuários em serviços de redes sociais contribuem, em larga escala, como outra fonte de dados.

A combinação desses dois modos é uma característica única de MCS.

O segundo é o **estilo de sensoriamento**. Ganti et al. [Ganti et al. 2011] afirmam que o sensoriamento em MCS pode ser classificado em **participativo** e **oportunista**, dependendo dos requisitos das aplicações e dos recursos do dispositivo. O sensoriamento é participativo, também chamado de **explicito**, quando a tarefa de sensoriamento é informada pelo usuários. Em outras palavras, a geração de dados (sensoriamento) é feita a partir do dispositivo do usuário de forma individual ou baseada no contexto. Já o sensoriamento oportunista, como o próprio nome diz, tira proveito das interações online envolvendo elementos físicos (por exemplo, *check-in* em lugares) ocorridas em redes sociais móveis como fonte de dados. Assim, os dados coletados são enviados pelos usuários de serviços de redes sociais móveis. No entanto, os dados são usados para um segundo propósito (o objetivo principal é a interação social online), e, portanto, ela é realizada de modo **implícito**. Devido essa caracterização de dois estilos de sensoriamento, Guo et al. [Guo et al. 2015] afirmam que MCS apresenta uma nova dimensão: a consciência do usuário na realização do sensoriamento.

O terceiro é a **organização de voluntários**. Os participantes podem ser cidadãos auto-organizados com diferentes níveis de envolvimento organizacional, variando de totais estranhos, grupos pouco organizados de vizinhos que enfrentam um problema comum até grupos bem-organizados de ativistas previamente existentes [Maisonneuve et al. 2010]. Os voluntários podem ser organizados em:

1. *Grupos*, pessoas organizadas de forma oportunista (tipicamente vizinhos) que desejam abordar de forma colaborativa um problema enfrentado por todos;

2. *Comunidades*, onde as pessoas que vivem em uma determinada área se unem por causa de seus interesses comuns;

3. *Urbanos*, onde qualquer cidadão (na sua maioria estranhos) pode participar da atividade de sensoriamento em escala urbana.

Sistema Centrado no Usuário

Como já mencionado neste Capítulo, MCS destaca-se pela capacidade de integrar a inteligência humana e da máquina. Contudo, para motivar a participação plena dos usuários bem como melhorar sua experiência de uso, sistemas MCS são desenvolvidos centrados no usuário (*user-centric*).

Embora vital, tal fato traz alguns problemas. O primeiro deles é como **motivar os usuários a participarem**? A primeira vista, a promessa de ganho financeiro é um método importante e bastante usado como incentivo. Contudo, o simples fato do entretenimento pode ser motivador em muitas situações, mesmo quando não há perspectiva de ganhos financeiros [Ganti et al. 2011]. As pessoas também podem ser motivadas a participar por razões éticas e sociais, tais como a socialização com outros ou o reconhecimento. Vale lembra que os usuários tem noção de que ao usar seus dispositivos e sensores estarão consumindo seus próprios recursos (energia e dados móveis, por exemplo).

Outro problema é a **segurança e privacidade do usuário**. O compartilhamento de dados pessoais em sistemas MCS pode levantar preocupações quanto a privacidade. Para motivar a participação do usuário, é essencial que novas técnicas para proteção da privacidade do usuário sejam elaboradas, permitindo que seus dispositivos possam contribuir de forma confiável. De forma particular, a definição de segurança e privacidade pode e precisa evoluir nos sistemas MCS, já que informações pessoais podem não ser obtidas diretamente, mas sim inferidas a partir de dados agregados. Por exemplo, o fato de que um objeto com uma etiqueta RFID poder ser identificado exclusivamente e rastreado de volta até seu utilizador pode trazer muitos problemas de privacidade [Acampora et al. 2013].

Requisitos de Comunicação

O sucesso de qualquer solução MCS depende de recursos de comunicação e de conexões heterogêneas temporárias que permitam a coleta eficiente dos dados sensoriados. Embora aplicações e sistemas MCS possam ter arquiteturas de comunicação diferentes, grande parte delas baseai-se em quatro pontos.

O primeiro é a **conexão de rede heterogênea**. Os atuais dispositivos móveis são geralmente equipados com múltiplas interfaces e tecnologias de comunicação sem fio (por exemplo, GSM, Wi-Fi e Bluetooth). Enquanto as interfaces GSM e WiFi podem fornecer conectividade com uma infraestrutura de comunicação pré-existente, Bluetooth ou Wi-Fi podem fornecer conexão de curto alcance entre dispositivos móveis e redes oportunistas auto-organizadas para compartilhar dados [Conti and Kumar 2010, Guo et al. 2013].

O segundo é a **topologia de rede e mobilidade humana**. A mobilidade dos dispositivos móveis e seus donos não só fornece uma boa cobertura de sensoriamento para

MCS, mas também traz desafios para as comunicações:

- A topologia da rede evolui ao longo do tempo, o que torna difícil encontrar rotas estáveis entre dispositivos móveis;
- Os protocolos de roteamento tradicionais, projetados para redes sem fio estáticas, não conseguem lidar com topologias altamente dinâmicas para cumprir as tarefas básicas de comunicação em MCS (especialmente para implementações ad hoc puras);
- Uma vez que a mobilidade humana desempenha um papel importante na dinâmica e no comportamento de sistemas MCS, os esforços em pesquisa sobre mobilidade humana [Karamshuk et al. 2011, Clementi et al. 2013] precisam avançar.

O terceiro é o **serviço tolerante a interrupções**. Em algumas aplicações MCS, os dados sensoriados não precisam ser transmitidos em tempo real ou com garantias de completude e precisão. Portanto, esses sistemas podem tirar proveito das redes tolerantes a interrupção [Gao and Cao 2011], que só dependem de conectividade de rede intermitente e têm custo de implantação muito menor. Vale lembrar que muitos dispositivos móveis não possuem garantias de conexão o tempo todo devido à fraca cobertura de rede (por exemplo, a força fraca de sinal devido a interferências ou nenhum sinal em uma área rural), restrição de energia (por exemplo, bateria fraca), ou preferências do usuário (por exemplo, telefone desligado em uma reunião) [Ma et al. 2014].

O quarto e último é a **exigência de alta escalabilidade**. Uma vez que MCS baseia-se em dados de sensoriamento de um grande volume de usuários móveis, a escalabilidade é claramente um requisito básico e um desafio fundamental para os sistemas de comunicação subjacentes. Para alcançar a escalabilidade suficiente, os protocolos de comunicação MCS e arquiteturas de rede são geralmente altamente distribuídas e descentralizadas. Tais soluções podem também melhorar a robustez do sistema global MCS. Além disso, o projeto eficiente de energia tem de ser considerado, devido aos recursos limitados de energia de cada dispositivo individual e do grande número de dispositivos no sistema.

Processamento de Dados e Inteligência na Extração

Uma vez que um dos objectivo de MCS é extrair inteligência de alto nível a partir de um grande volume de entradas, a participação humana no processo de sensoriamento traz algumas incertezas para os sistemas MCS.

Uma delas é a **baixa qualidade dos dados**, geralmente definida como o grau de como se encaixam os dados para utilização em operações, tomada de decisão e planejamento. Por exemplo, os participantes anônimos podem enviar dados incorretos, de baixa qualidade ou até mesmo falsos [Zhang et al. 2014b, Wang et al. 2011a]. Além disso, dados oriundos de pessoas diferentes podem ser redundantes ou inconsistentes ou o mesmo sensor pode sentir o mesmo evento sob diferentes condições (por exemplo, detecção de ruídos em ambiente com o celular no bolso ou na mão). Portanto, a seleção de dados é muitas vezes necessária para melhorar a qualidade dos dados, devendo ser explorados métodos de filtragem de faltas, estimativa de qualidade, incentivo ao participante especialista, entre outros.

Outra questão é a **mineração de dados heterogêneos**. Os dados em MCS são obtidos em comunidades físicas e virtuais tanto de forma offline quanto online. Diferentes comunidades representam maneiras distintas de interação (por exemplo, comentários, transferências online, localização offline) e conter conhecimento diferente (por exemplo, a amizade em comunidades online, os padrões de movimento em comunidades offline). Portanto, o problema é como associar, de forma eficaz, dados com diferenças espaciais.

2.2.6. Exemplos de Aplicações

Para melhor apresentar o paradigma, esta Seção descreve algumas aplicações voltadas para MCS.

CreekWatch

Desenvolvido pela IBM Almaden Research Center, o Creek Watch [IBM 2010] visa monitorar o níveis das águas de riachos, baseando-se em imagens coletadas por participantes. O aplicativo permite o uso do smartphone (somente versões para IOS estão disponíveis) para marcar uma posição (usando GPS), tirar fotos daquele ponto e informar observações perceptíveis pelo usuário como o nível da água, o fluxo e a presença de lixo. A Figura 2.4 ilustra alguns passos para o uso do Creek Watch.

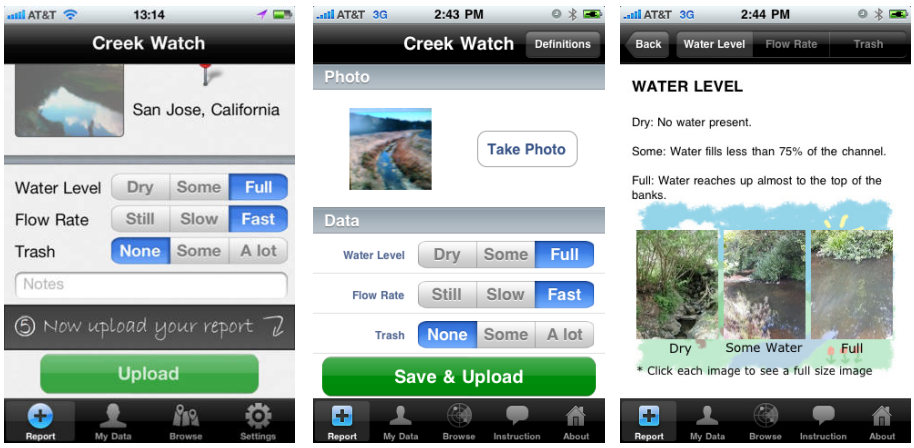

Figura 2.4. Exemplo de uso do Creek Watch

Os dados coletados são automaticamente enviados para o servidor central ou armazenados no próprio dispositivo se não houver acesso a Internet. Todas as informações podem ser acessadas e baixadas em `http://creekwatch.org`. De acordo com a IBM, o Creek Watch já conta com mais de 4.000 usuários em 25 países.

NoiseTube

NoiseTube é uma nova abordagem para a avaliação da poluição sonora que envolve o público em geral [Maisonneuve et al. 2009]. O objetivo é transformar telefones celulares equipados com GPS em sensores de ruído que permitam aos cidadãos medir sua exposição pessoal ao ruído no seu ambiente quotidiano.

Cada usuário pode contribuir compartilhando suas medições geolocalizadas bem como anotações pessoais para produzir um mapa de ruído coletivo. Uma vez que os dados coletados são enviados ao servidor, qualquer usuário pode ver suas próprias contribuições

e a de outros através do Web site ou visualizá-las usando o Google Earth. Essa maior possibilidade de exibição faz com que o NoiseTube forneca dados para convencer e auxiliar autoridades na tomada de decisão sobre o problema da poluição sonora. NoiseTube usa o conceito de translucidez social, que consiste em fazer os participantes e suas atividades se tornarem visíveis um para o outro.

A versão inicial do NoiseTube foi escrita em Java, focada em smartphones com sistema operacional Symbian/S60 e testada em um dispositivo Nokia N95 com 8GB. Além disso, um receptor de GPS externo (conectado via Bluetooth) foi utilizado. Atualmente, está disponível para plataformas iOS, Android e smartphones com Java ME. O servidor é implementado using Ruby on Rails, MySQL, Google Maps e Google Earth.

A Figura 2.5 ilustra algumas interfaces do NoiseTube, que representam a coleta e a visualização das medições.

Figura 2.5. Exemplo de interfaces do NoiseTube

Nericell

O Nericell é uma aplicação MCS voltada para monitorar o tráfego em ruas, avenidas e rodovias em tempo real, mas de forma oportunista [Mohan et al. 2008]. Desenvolvida pela Microsoft Research, a ideia é que quando um usuário participante coloca seu telefone no bolso e começa a dirigir, a aplicação irá automaticamente monitorar as condições da estrada e do trânsito, transmitindo certas informações para um serviço na nuvem para a agregação e geração de relatórios.

Um aspecto interessante da Nericell é o uso de vários sensores presentes em dispositivos móveis, tais como bluetooth, rádio celular, microfone, acelerómetro e GPS. Segundo os criadores, um sensoriamento rico e diversificado é fundamental no contexto das cidades, onde as condições das estradas tendem a ser variáveis (buracos), existem vários tipos de veículos (2 rodas, 3 rodas, carros, caminhões) e o fluxo de tráfego pode ser caótico (por exemplo, vários motorista buzinando em um congestionamento). Assim, por exemplo, o acelerômetro pode ser usado para detectar buracos e o microfone para detectar buzinas, ajudando a determinar a condição de trafegabilidade.

Além da questão do sensoriamento, Nericell também resolve certos aspectos técni-

cos como consumo de energia. Sobre consumo de energia, a aplicação emprega o conceito de sensoriamento por gatilho, onde um sensor relativamente barato em termos de energia é usado para desencadear a operação de um sensor de maior consumo energético, quando necessário. Por exemplo, o acelerômetro é utilizado para detectar uma alta incidência de frenagens, o que provoca então o sensoriamento baseado no microfone para verificar se há buzinas.

Waze

Waze é um popular sistema de navegação que usa *Crowdsensing* para oferecer informações de tráfego em tempo quase real. Criado em 2008, o Waze registrava aproximadamente mais de 50 milhões de usuários em 2013, ano em que foi comprado pelo Google. Waze recolhe periodicamente dados do GPS em dispositivos móveis e os utiliza para calcular a velocidade do veículo. Assim, pode fornecer informações úteis sobre as condições de tráfego em diferentes áreas. O sistema também oferece aos seus usuários alertas pré-definidos informando incidentes como engarrafamentos e pontos de verificação da polícia, bem como informações sobre as condições de tráfego. Também é possível usar subcategorias de incidentes para melhor especificá-los, por exemplo, "engarrafamento pesado" em vez de apenas engarrafamento.

A Figura 2.6 ilustra algumas interfaces do Waze.

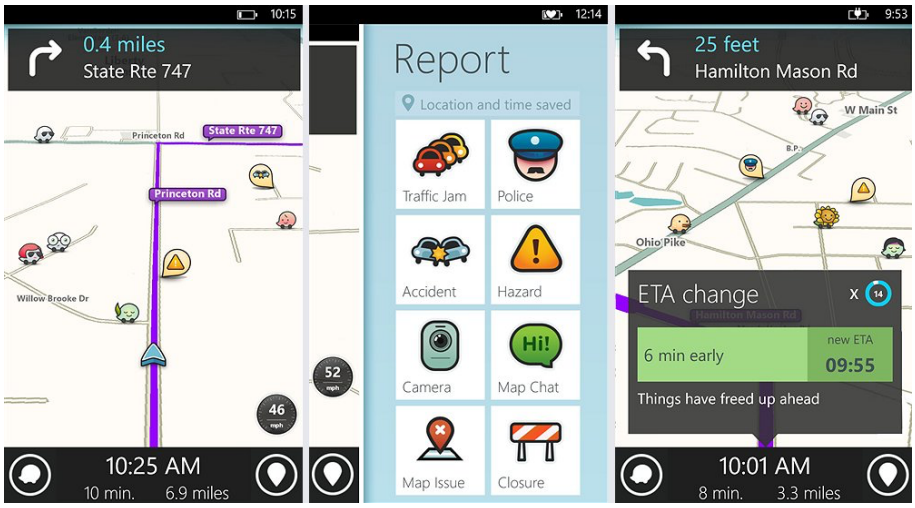

Figura 2.6. Exemplo de interfaces do Waze

mCrowd

mCrowd é uma plataforma de *Crowdsourcing*, desenvolvida em [Yan et al. 2009], que permite aos usuários móveis postar e trabalhar em tarefas relacionadas com o sensoriamento móvel distribuído. Ela permite que os usuários de iPhones possam utilizar, de forma plena, os diferentes sensores para participar e realizar tarefas de *Crowdsourcing*, incluindo a coleta de imagens com ciência de localização (geolocalização), marcação de imagens, monitoramento do tráfego rodoviário, entre outros. Obviamente, para incentivar a participação, é oferecida uma compensação financeira para cada usuário. Além de facilitar o *Crowdsourcing*, mCrowd torna viável que os usuários usem serviços populares como

Amazon Mechanical Turk (`https://www.mturk.com/mturk/welcome`) (Mturk) e ChaCha (`http://www.chacha.com`), através de uma interface única, simplificando assim sua participação no *Crowdsourcing*.

Por ser um aplicativo interativo e de fácil manuseio, o tratamento com as tarefas é um processo simples. Primeiro, as tarefas de interesse são postadas e em seguida os usuários capazes de executá-las são convocados para atuar. Diferentes serviços de *Crowndsourcing* podem ser usados, como Mtrurk ou ChaCha, para postar uma imagem ou texto. Ao acessar o mCrowd, o usuário pode escolher uma categoria de tarefas e postá-la no Mturk através de um mecanismo de busca. Após enviar a tarefa, uma lista com as tarefas é montada e a recompensa associada a tarefa.

A Figura 2.7 mostra a interface de usuário do mCrowd.

(a) Visão do Solicitante (b) Exibição de Tarefas (c) Resultados

Figura 2.7. Interfaces do mCrowd. Os usuários que desejam iniciar uma tarefa postam a partir da (a) visão do solicitante, as tarefas postadas são mostrados na exibição de tarefas (b) e os resultados apresentados pelos participantes são mostradas em resultados (c). Fonte: [Yan et al. 2009].

mCrowd foi estruturado para suportar quatro tipo de tarefas: marcação de imagem, captura de Imagem, consultas textuais e consultas baseadas em localização. Na marcação de imagem, os objetos de interesse são etiquetados ou marcados para serem consumidos pelos usuários, por exemplo, marcar um determinado objeto em uma imagem. Já na captura de imagens, a imagem é capturada quando há interesse em um determinado local, como parques de diversões. Consultas textuais são destinadas a obter as opiniões dos usuários como, por exemplo, quais os melhores restaurantes, pontos turísticos, entre outros. Consultas baseadas em localização buscam informações sobre locais mais próximos de um determinado ponto (por exemplo, restaurantes, floriculturas, entre outros).

2.3. Privacidade e Confiabilidade em Mobile Crowd Sensing

Esta Seção discute os dois principais problemas de segurança enfrentados por soluções, sistemas e aplicações MCS: a privacidade do usuário e a confiabilidade dos dados.

Em relação a privacidade, o problema é a possibilidade da divulgação de informações privadas (identidades, endereços físicos e na Internet, rotas, estilo de vida, entre muitas outras), uma vez que aplicações MCS gerenciam grandes volumes de informação que tipicamente são e devem ser tornadas públicas.

Já quanto a confiabilidade dos dados, o problema é que não existem garantias de que ou os dados inseridos pelos participantes são reais ou que as informações mantidas nos servidores ainda são reais.

2.3.1. Privacidade

O que é privacidade? Na área do Direito Civil, conceitua-se privacidade como a "faculdade que tem cada indivíduo de obstar a intromissão de estranhos em sua vida privada e familiar, assim como de impedir-lhes o acesso a informações sobre a privacidade de cada um, e também impedir que sejam divulgadas informações sobre esta área da manifestação existencial do ser humano" [Vieira and Alves 2014]. Em outras palavras, privacidade é o direito de cada indivíduo de manter e controlar o conjunto de informações que o cerca, podendo decidir se, quando, por que e por quem essas informações podem ser obtidas e usadas. Esse conjunto de informações incluem desde o seu modo de vida, as relações familiares e afetivas, os segredos, os pensamentos, os hábitos, os fatos e até mesmo os planos de futuro.

No paradigma MCS, devido suas características únicas, a privacidade envolve o direito do usuário e/ou participante de permanecer livre de intrusos e autônomo. A privacidade em MCS traz preocupações com a divulgação direta da identidade dos participantes bem como com a divulgação de atributos sensíveis, incluindo locais (por exemplo, endereço residencial ou do trabalho) e outras informações privadas como atividades pessoais ou condições (por exemplo, estilo de vida ou doença) que permitam inferir sobre a identidade dos participantes [Pournajaf et al. 2014].

Do ponto de vista dos participantes, as ameaças à privacidade podem ocorrer quando:

1. O participante recebe uma tarefa específica e compartilha suas preferências durante a atribuição dessa tarefa ou notifica o servidor que aceitou a tarefa. Neste momento, alguns atributos como a localização, o tempo, os tipos de tarefas que o participante está interessado ou alguns atributos do sensor que ele dispõe podem ser revelados. Embora seja possível argumentar que estas informações por si só podem não violar a privacidade, o fato é que elas podem permitir que alguém (atacante, adversário) possa rastrear as tarefas selecionadas pelo participante e, consequentemente, revelar sua identidade ou outros atributos sensíveis [Shin et al. 2011]. Alguns dos atributos que podem ser usados para rastrear os participantes são seus IDs, endereços IP ou outras informações de rede.

2. O participante realiza ou participa de tarefas espaciais com frequência. Essa repetição pode levar à divulgação de atributos sensíveis, como endereço residencial ou, eventualmente, sua identificação através de ataques de inferência [Krumm 2007]. Nestes tipos de tarefas, mesmo que o participante esteja usando o aplicativo de forma anônima, sua trajetória pode revelar locais sensíveis.

3. No processo de definição de tarefas, uma entidade maliciosa (aplicativo ou uma entidade separada) cria tarefas que impõem limitações estritas sobre os atributos do participante ou o dispositivo que está carregando (por exemplo, exigindo um estilo de vida especial ou um tipo de sensor raro para se qualificar para a tarefa). Este ataque pode resultar na divulgação da identidade ou outros atributos sensíveis do participante que aceita uma tarefa tão rigorosa [Shin et al. 2011].

4. No processo de distribuição de tarefas, uma entidade maliciosa (aplicativo, uma entidade separada ou outro participante) pode compartilhar tarefas para um conjunto limitado de participantes para ser capaz de aprender seus atributos ou rastreá-los [Shin et al. 2011] (por exemplo, empurrar ou atribuir uma tarefa a apenas um participante).

5. Várias aplicações ou usuários finais podem conspirar para conectar-se a informação dos participantes, com a finalidade de desanonimização. O usuário final malintencionado pode criar várias aplicações em uma tentativa de coletar mais dados privados.

De acordo com Christin et al. [Christin et al. 2011], as principais informações reveladas quando ocorrem quebras na privacidade dos participantes são:

- **Tempo e Localização**: A divulgação desses dois tipos de dados tem si mostrado altamente sensível à privacidade dos participantes, uma vez que permitem inferir ou de fato obter endereços residências e do local de trabalho, bem como as rotinas e hábitos dos participantes [Shilton 2009]. Por exemplo, visitas frequentes a hospitais podem permitir que os empregadores infiram sobre a condição médica de seus empregados, assim como a participação em eventos políticos pode fornecer informações sobre os pontos de vista políticos dos utilizadores [Liu 2007]. Em resumo, sem qualquer mecanismo de protecção, a divulgação de informações de localização pode levar à graves consequências sociais [Shilton 2009]. Além disso, as ameaças resultantes do rastreamento de localização e tempo não se limitam a aplicações onde é necessária autenticação. Mesmo no caso de contribuições anônimas, registros de localização podem ser analisados para deduzir a identidade dos participantes com base na localização de sua residência e consultas as listas telefônicas [Mun et al. 2009].

- **Amostras de Som**: Além inferir identidades e preferências a partir de dados espaço-temporais (tempo e localização), a "imagem" dos usuários pode ser refinada com amostras de outras modalidades de sensoriamento. Um bom exemplo são as amostras de som registradas intencionalmente pelos participantes ou capturadas automaticamente pelos dispositivos móveis. Enquanto os participantes podem facilmente preservar a sua privacidade somente gravando eventos não-sensíveis no primeiro caso, os dispositivos móveis, de forma eficaz, se comportam como espiões inteligentes no caso de gravações automatizados. Mesmo em locais públicos, o reconhecimento de padrões sonoros característicos, que são exclusivos para determinados eventos e locais, podem permitir que adversários determinem o contexto atual de um participante.

- **Fotos e Vídeos**: O conteúdo de imagens e vídeos gravados também é susceptível a revelar informações pessoais sobre os participantes e seu ambiente. Por exemplo, a aplicação SensoDiet [Reddy et al. 2007] marca as fotos tiradas das refeições consumidas e não existe qualquer contramedida para esconder os rostos das pessoas que compartilham a refeição com os participantes. Em todos os cenários em que a câmera é orientada longe do participante, rostos de outras pessoas na vizinhança são possivelmente capturados nas imagens, e, portanto, as conclusões sobre o número e a identidade das relações sociais do participante podem ser feitas. A publicação de imagens capturadas pode levar a consequências semelhantes como as em redes sociais online, como o Facebook. Semelhante a gravações de som, o contexto do utilizador atual e o ambiente em volta também podem ser extraídos a partir dos dados do sensor. Por exemplo, imagens que mostram pontos de interesse podem facilmente estabelecer a presença do participante nesses locais.

- **Acelerômetro**: É fácil pensar que as leituras de acelerômetros presentes em dispositivos móveis são os dados menos importantes no que diz respeito a revelar informações pessoais sobre os participantes. No entanto, essa hipótese nem sempre é verdadeira e muitas vezes pode servir como uma falsa sensação de segurança. Por exemplo, se o celular está na cintura, informações sobre a marcha, a forma de caminhar, correr, entre outras, podem gerar possíveis indicações sobre a identidade de um usuário [Derawi et al. 2010]. Assim, empregadores podem querer verificar se seus funcionários estão realmente trabalhando durante suas horas de trabalho.

- **Dados ambientais**: Dados ambientais (registro de partículas, das concentrações de gás ou pressão barométrica) podem não representar uma ameaça direta a privacidade dos participantes. No entanto, esses registros, quando combinados com informações secundárias, tais como temperatura, podem revelar a localização dos participantes a um nível de granularidade melhor (um quarto dentro de um edifício) do que as obtidas via GPS ou outros serviços de localização.

- **Dados biométricos**: Da mesma forma que dados de um sensor biométrico podem ser usados para o diagnóstico do estado fisiológico do usuário, atacantes e adversários podem identificar anomalias de saúde ou doenças com base nos dados capturados. Informações médicas vazadas podem ser utilizadas por companhias de seguros de saúde ou empregadores para revogar contratos, se um agravamento das condições fisiológicas dos participantes é identificado.

Como forma de manter a privacidade dos usuários, e ao mesmo tempo disponibilizar as informações coletadas, várias técnicas e métodos de segurança são empregados. São elas [Christin et al. 2011]:

2.3.1.1. Preferências do Usuário

Permitir aos participantes expressarem suas preferências de privacidade é uma técnica que visa controlar o processo de coleta de dados junto ao dono do dispositivo sensor, evitando assim danos à privacidade dos participantes. Além de simples, também permite a aplicação em diferentes graus.

Das et al. [Das et al. 2010] propuseram um esquema binário (acesso completo aos dados do sensor ou nenhum acesso) que pode ser estendido através da introdução de níveis intermediários adicionais. Por exemplo, os participantes podem decidir ativar seletivamente as medições do sensor, dependendo de uma variedade de fatores, como a presença em locais sensíveis (casa ou escritório), e de pessoas de seu convívio social (presença de amigos ou familiares). A seleção desses fatores permite aos participantes expressamente indicar quais tipos de informação eles estão aptos e felizes a coletar em diferentes contextos.

Os autores argumentam que um esquema com granularidade mais fina nas preferências de sensoriamento pode permitir aos participantes ajustar tanto a atividade quanto o tempo de atuação. Por exemplo, amostras podem ser recolhidas a cada hora ao invés de a cada 15s ou as informações de localização podem ser capturadas em diferentes graus de granularidade.

Embora essa técnica não forneça claramente um *trade-off* entre privacidade e dados divulgados, ela obviamente melhora a aceitação global de uma solução ao oferecer granularidade de configurações de privacidade para os participantes. Vale ressaltar que embora algumas soluções permitem que os usuários desativem totalmente sua função sensora [Miluzzo et al. 2008, Shilton et al. 2008], isso é de pouca utilidade para MCS.

2.3.1.2. Distribuição de Tarefas Anonimamente

A coleta de dados de um sensor é geralmente desencadeada por meio de tarefas que especificam as modalidades de sensoriamento (por exemplo, regiões de interesse, critérios a cumprir para iniciar a captura, sensores usados e frequência de amostragem) com base nos requisitos da aplicação [Reddy et al. 2009]. Assim, as tarefas são distribuídas para os dispositivos que satisfaçam os requisitos das tarefas.

Como explicado na Seção 2.2.2, a distribuição de tarefas pode ocorrer de forma centralizada - através de um servidor de tarefas ou aplicativos, que seleciona os dispositivos apropriados com base em critérios pré-determinados para otimizar o processo de sensoriamento, de forma descentralizada, onde os próprios dispositivos, de forma autônoma, gerenciam e distribuem tarefas ou de forma híbrida. Em todas essas três abordagens, os processos de contato e envio das tarefas podem colocar a privacidade dos participantes em perigo de várias maneiras, uma vez que baixar tarefas fornece informações ao servidor de tarefas sobre a localização dos participantes, com data e hora precisas. Além disso, a natureza das tarefas fornece dicas sobre os dispositivos utilizados.

Como forma de proteger os participantes, foram propostos alguns mecanismos para assegurar o anonimato e a privacidade de localização:

- Transferência das tarefas somente em locais densamente povoadas [Shin et al. 2011], já que a alta densidade de pessoas presentes nesses locais torna a identificação dos participantes pelo servidor difícil e, portanto, esconde suas identidades.
- Autenticação baseada em atributos [Kapadia et al. 2009], onde ao invés de usar a sua identidade, os participantes podem usar credenciais baseadas em criptografia

que comprovam sua participação em um determinado grupo (por exemplo, os alunos matriculados no clube de ciclismo da universidade);

- Esquemas de roteamento para preservação da privacidade de localização, que visam esconder a localização dos participantes [Kapadia et al. 2009] utilizando rotas específicas.

2.3.1.3. Anonimato

Segundo o Dicionário Aurélio, anonimato significa: (1) qualidade do que é anônimo; (2) sistema de escrever anonimamente, sem se identificar; e (3) o mesmo que anonimado. De forma geral, anonimato é a qualidade ou condição de permanecer anônimo. Em segurança da informação, anonimato é o que permite ocultar qualquer informação que possa identificar um usuários antes de compartilhar informações.

Embora o uso de técnicas de anonimato seja muito abrangente, em MCS a intenção é remover qualquer informação que possa identificar os usuários/participantes ou outras entidade durante a distribuição e realização de tarefas [Pournajaf et al. 2014]. Os métodos de anonimato mais usuais para proteção da privacidade em MCS são:

Pseudônimos

Pseudônimo é um mecanismo comum e até certo ponto simples para proteção do anonimato e da privacidade dos participantes [Christin et al. 2011], visto que ao invés de transmitir nomes em texto simples, toda a interação com o aplicativo é executado sob um pseudônimo. O uso de pseudônimos em tarefas de sensoriamento já é bem conhecido e é empregado nos trabalhos de Shilton et al. [Shilton et al. 2008], Shilton [Shilton 2009] e Deng e Cox [Deng and Cox 2009] e mais recentemente em Konidala et al. [Konidala et al. 2013] e Gisdakis et al. [Gisdakis et al. 2014].

Vale ressaltar que o uso de pseudônimos não garante necessariamente a privacidade, especialmente em aplicativos baseados em localização. Todos autores que empregam esse mecanismo são unanimes em afirmar que quando usado em conjunto com outros esquemas de autenticação, soluções à base de pseudônimos garantem anonimato e confidencialidade para os participantes.

K-Anonimato

Além dos esquemas de roteamento, mecanismos baseados no *k-anonimato* podem ser aplicados para proteger a privacidade localização dos participantes quando recebendo tarefas ou enviando os dados. A ideia chave por trás *k-anonimato* [Sweeney 2002] é construir grupos de k participantes que partilham um atributo comum (por exemplo, os k participantes localizados no mesmo distrito), tornando-os indistinguíveis uns dos outros.

Diferentes métodos vêm sendo utilizados para encontrar um atributo comum apropriado e assim construir grupos de k usuários. Estes métodos podem ser classificados em

duas categorias principais de generalização e perturbação [Huang et al. 2010b]. No primeiro caso, o valor original do atributo é generalizado por um valor com menor grau de detalhe. Por exemplo, as coordenadas exatas dos k participantes são substituídas pelo nome do distrito da localização atual.

Em Shin et al. [Shin et al. 2011], os autores usaram uma forma de generalização baseada na divisão de uma área geográfica em várias regiões, chamada de *Tessellation*, para mapear os pontos de acesso. O ponto de acesso no centro de cada região mantém um registo do número médio de dispositivos ligados, que é igual ao valor máximo K que pode ser conseguida dentro da região. Para garantir *k-anonimato* em toda a rede, regiões com $K < k$ são combinados em células com um valor igual à soma de cada região individual. Uma vez que as células tenham sido definidas, os participantes marcam seus dados com o limite geográfico da sua célula atual ao invés de fornecer suas coordenadas exatas.

Por outro lado, a perturbação é baseada na substituição dos dados dos sensores originais por um novo valor resultante de uma função aplicada às leituras dos sensores de k membros do grupo. Por exemplo, a localização de cada membro do grupo pode ser substituída pela localização média de todos os membros. Domingo-Ferrer e Mateo-Sanz [Domingo-Ferrer and Mateo-Sanz 2002] usam micro-agregação para substituir o local real pela localização mais próxima dos k participantes.

Um risco a soluções de *k-anonimato* é a possibilidade de ataques de homogeneidade, como apresentado por Machanavajjhala et al. [Machanavajjhala et al. 2007]. Esses ataques exploram a monotonia de determinados atributos para identificar indivíduos a partir do conjunto de k participantes.

Ocultação Seletiva

Locais sensíveis podem ser selecionados pelos participantes e protegidos usando a ocultação seletiva de localização [Mun et al. 2009]. A ideia é que quando um usuário se aproxima de um local que foi definido previamente como sensível, a aplicação gera registros fictícios de localização para evitar que o local seja selecionado. Obviamente, os registros gerados são e precisam ser realistas, ou seja, devem realmente representar estradas e ruas existentes. Em geral, um algoritmo de ocultação seleciona, primeiro, os lugares mais próximos e, em seguida, refina a seleção levando em conta o histórico de resultados dos participantes (por exemplo, suas capacidades físicas com base em suas experiências anteriores). Além disso, o algoritmo ainda altera as atividades e modifica sua duração para manter a consistência dos resultados da aplicação.

Quando comparado à *k-anonimato*, o esquema de ocultação seletiva melhora a privacidade de localização sem impactar os resultados da aplicação.

Agregação de dados

Em MCS, os dados precisam ser agregados e isso é normalmente feito pelo servidor de aplicação ou algum serviço próprio na nuvem. Contudo, a abordagem de agregação

de dados para preservação da privacidade proposta em Shi et al. [Shi et al. 2010] não depende de uma entidade central para proteger a privacidade de dados, mas sim da proteção mútua entre participantes. Antes de transmitir os dados para o servidor, os dispositivos móveis distribuem parcialmente seus dados entre os vizinhos. Em seguida, fazem o upload dos dados de sensoriamento vindo de seus vizinhos e os restantes de seus próprios dados. Esta distribuição diminui a probabilidade de atribuir, com sucesso, cada leitura do sensor para o dispositivo móvel que realmente a capturou. Por exemplo, se os dois dispositivos móveis (A e B) trocam metade de seus dados, a probabilidade de que os dados reportados por A ser realmente capturado por ele mesmo é só de 50% e o mesmo para B.

Dependendo da natureza das funções de agregação, dois regimes distintos podem ser aplicados. Para funções aditivas, cada dispositivo móvel particiona seus dados em $n+1$ fatias e envia uma fatia para cada um de n nós selecionados. Uma vez que cada nó distribuiu suas fatias para seus vizinhos, as fatias trocados e própria fatia do nó são combinadas e enviadas para o servidor de agregação que é então capaz de calcular o resultado agregado. Para as funções de agregação não-aditivas, tais como percentuais e histogramas, um método que combina corte, consulta de contagem e busca binária pode ser aplicado. No entanto, esta abordagem só assegura a proteção da privacidade de dados se os nós e o servidor não conspirarem para violar a privacidade de alvos potenciais.

2.3.1.4. Processamento de Dados

Em aplicações típicas de MCS e sensoriamento participativo, o processamento de dados é compartilhado entre os dispositivos móveis e servidores de aplicativos. No entanto, devido às limitações de recursos em plataformas móveis, a distribuição das tarefas de processamento entre ambas as partes é tipicamente feita em direção ao servidor. Embora o pré-processamento seja geralmente levado a cabo sobre os dispositivos para reduzir a quantidade de dados para transformar, a fim de economizar largura de banda e energia de transmissão, processamento de tarefas complexas pode exceder o poder computacional de dispositivos móveis, obrigando a execução no servidor.

O processamento de dados no dispositivo móvel consiste, principalmente, na extração de características, a partir dos dados brutos, a fim de remover informações sensíveis (por exemplo, vozes humanas gravadas ou pessoas fotografadas) que possam pôr em perigo a privacidade dos participantes e também para economia de recursos. Por exemplo, um classificador de áudio pode analisar as amostras de som para determinar se vozes humanas foram registradas [Miluzzo et al. 2008]. Além disso, o nível de intensidade das amostras de áudio pode ser determinado localmente através da execução de algoritmos de processamento de sinal para minimizar os dados a serem transferidos para o servidor [Maisonneuve et al. 2009]. Após processamento, os dados brutos podem ser excluídos do armazenamento local e os resumos processados são relatados ao servidor central.

No lado do servidor, os dados relatados podem então ser processado para eliminar informações sensíveis sobre a privacidade, tais como as características de identidade ou dados que ameaçam o anonimato/privacidade dos participantes. Por exemplo, os dados capturados podem ser agregados entre vários participantes para torná-los indistinguíveis ou publicados sob a forma de estatísticas [Ganti et al. 2011] e mapas [Dong et al. 2008].

Ao fazer isso, os dados confidenciais não estão diretamente relacionados aos usuários finais, o que evita a identificação direta dos participantes. No entanto, os participantes devem contar com o aplicativo para: anonimizar eficientemente os dados, proteger suficientemente sua privacidade, e não divulgar informações sensíveis à privacidade em seus dados comunicados a terceiros.

2.3.1.5. Controle de Acesso e Auditoria

Dependendo do cenário de aplicação, os resultados de um sensoriamento podem não ser só de interesse dos participantes, mas também de outras pessoas como, por exemplo, pesquisadores, pessoal médico, amigos, familiares, conselhos municipais ou até mesmo um público maior. No entanto, os participantes podem não estar dispostos a partilhar seus dados com todos os tipos de pessoas dentro do grupo de pessoas interessados com a mesma granularidade. Além de abordar questões de privacidade antes do processo de sensoriamento e da liberação de dados para a aplicação, os participantes podem definir o público-alvo que está autorizado a acessar seus dados a partir da interface de usuário de muitas aplicações.

Os participantes podem definir grupos [Grosky et al. 2007, Gaonkar et al. 2008], selecionar certas pessoas [Reddy et al. 2007, Gaonkar et al. 2008, Shilton 2009] ou autorizar todos [Gaonkar et al. 2008]. Podem também refinar sua seleção, especificando a natureza dos dados que compartilham, bem como definir subconjuntos específicos de dados acessíveis [Reddy et al. 2007, Miluzzo et al. 2008, Shilton 2009]. Além disso, eles podem definir condições precisas de liberação do acesso aos dados [Shilton et al. 2008].

Para destacar as implicações de privacidade de partilhar dados, ferramentas gráficas, incluindo mapas ou imagens, são usadas para visualizar os dados sendo liberados e aumentar a consciência dos participantes [Shilton et al. 2008]. Depois que os dados foram publicados, os participantes também podem monitorar o acesso aos dados, consultando arquivos de log da aplicação. Esses logs registram a natureza dos dados acessados, a frequência desses acessos, bem como a identidade das pessoas a acessá-los [Shilton et al. 2008, Shilton 2009, Mun et al. 2010]. Com base nos resultados dessas auditorias, os participantes controlam a distribuição de suas informações e podem julgar sua adequabilidade. Se necessário, eles podem atualizar suas políticas de controle de acesso para restringir ou ampliar as condições de autorização, a fim de corresponder às suas preferências de privacidade.

2.3.2. Confiabilidade dos Dados

Embora, num primeiro momento possa ser difícil estabelecer uma relação entre confiabilidade dos dados e segurança em MCS, é fácil afirmar que o simples fato de haver participação humana no processo já é o problema. Em outras palavras, os dados coletados e enviados pelos participantes nem sempre podem ser tratados ou considerados como confiáveis. Os motivos são:

1. **Contribuições erradas**: Infelizmente, aplicações MCS podem permitir que qualquer participante possa contribuir com dados, o que significa que as aplicações

podem receber dados errados e mal intencionados. Os participantes podem propositalmente registrar medições incorretas, colocando seus dispositivos em posições inadequadas. Por exemplo, gravações de áudio incorretas podem ser obtidas caso o dispositivo móvel do participante esteja indevidamente colocado em locais que atrapalham seu funcionamento, como no bolso durante uma tarefa de detecção de ruído.

2. **Conluio**: Nos atuais sistemas de reputação que usam MCS, cada participante pode publicar falsos relatórios para o servidor de aplicativo e fornecer informações arbitrárias sobre os valores durante o processo de votação de reputação. Além disso, atacantes podem conspirar para inutilizar aplicativos MCS.

3. **Contribuições mal-intencionadas**: Participantes mal-intencionados, para seus próprios benefícios, podem forjar informações para obter um maior ganho antes de finalizar tarefas que envolvem algum tipo de ganho monetário [Tuncay et al. 2012].

O problema da confiabilidade dos dados é geralmente discutido como diretamente relacionado a privacidade dos usuários [Pournajaf et al. 2014, Ganti et al. 2011] . Por exemplo, para alcançar a confiabilidade dos dados de localização dos participantes, algumas soluções baseiam-se em verificar a localização segura do dispositivos móveis em tempo real [Capkun et al. 2006] enquanto outras tentar estimar a confiabilidade dos dados coletados.

Contudo, pesquisas recentes envolvendo sistemas confiáveis (*trusted system*) aplicados a MCS vêm sendo realizadas. Em linhas gerais, estes sistemas se concentram em como avaliar a confiabilidade dos dados compartilhados e como manter a reputação de entidades da rede de processamento de dados. Huang et al. [Huang et al. 2010a] propuseram um sistema de reputação com base na função de Gompertz para calcular scores de pontuações de dispositivos para medir a confiabilidade dos dados coletados. No entanto, ele não leva em conta a preservação da privacidade.

Mais recentemente, vários esquemas de reputação que estão têm sido propostos [Dua et al. 2009, Christin et al. 2012, Li et al. 2014, Shin et al. 2011]. Algumas das abordagens [Dua et al. 2009, Christin et al. 2012] invocam a existência de uma terceira parte confiável, mas o estabelecimento e manutenção desta entidade em um ambiente distribuído não é trivial. No esquema de [Dua et al. 2009], vários pseudônimos são atribuídos para cada participante. Um servidor confiável é necessário para gerenciar o mapeamento entre a verdadeira identidade um participante e seus pseudônimos, e os valores de reputação entre os diferentes pseudônimos. Em comparação com outros métodos, este método não requer operações criptográficas caras e tem baixo custo de comunicação. Além disso, Dua et ai. [Christin et al. 2012] propôs e implementou um **Trusted Platform Module** (TPM), que é um microcontrolador embutido dentro de cada dispositivo móvel, para atestar a integridade de leituras dos sensores. No entanto, os chips TPM ainda não são amplamente adotado em dispositivos móveis.

Alguns métodos que não dependem da existência de uma terceira parte de confiança foram propostos. Em [Li et al. 2014], foi proposta uma solução de preservação da reputação de anonimato chamado IncogniSense. Ele gera pseudônimos periódicos usando

assinaturas cegas e disfarça valores exatos de reputação dinamicamente em grupos de reputação. Por exemplo, a solução depende de uma redundante número de participantes e incorre em despesas gerais de comunicação pesados. Shin et al. [Shin et al. 2011] propôs outra solução baseada em assinatura cega. Os autores consideraram o problema do ponto de vista do incentivo e tem como objetivo permitir aos participantes ganharem créditos através de contribuições. Assim, a necessidade de penalizar participantes maliciosos não precisa ser considerada.

Além disso, fazendo uso do fato de que múltiplas fontes de informação estão geralmente disponíveis em uma rede, Wang et al. [Wang et al. 2011b] propuseram avaliar a semelhança de múltiplas informações de diferentes fontes sobre o mesmo acontecimento e, em seguida, ajustar a pontuação de confiança de cada parte de informação. Com base na avaliação de confiança, as pontuações de confiança de nós também pode ser ajustada dinamicamente.

2.4. Implementações Seguras e de Segurança para MCS

Esta Seção apresenta algumas arquiteturas de aplicações MCS projetadas para garantir a segurança e privacidade dos usuários. Também apresenta ferramentas e bibliotecas voltados à segurança em sistemas MCS. Primeiramente serão discutidas as arquiteturas e posteriormente as bibliotecas. No fim, uma discussão sumariza os trabalhos apresentados.

2.4.1. Arquiteturas

Anonymous Authentication of Visitors for Mobile Crowd Sensing at Amusement Parks

Konidala et al. [Konidala et al. 2013] elaboraram um aplicativo que emprega pseudônimos certificados e um esquema de assinaturas parcialmente cegas para garantir a autenticidade e privacidade de usuários em parques de diversão. A primeira técnica utiliza criptografia para tornar anônima a presença do visitante no parque. Já a assinatura parcialmente cega (*Partially Blind Signature*) permite ao visitante ocultar a mensagem a ser assinada bem como explicitamente incorporar informações necessárias como, data da emissão, data da validade, a identidade do assinante.

Em linhas gerais, a solução funciona da seguinte forma: ao chegar no parque, o visitante solicita o aplicativo, instala e informa dados como idade, sexo, nacionalidade, altura (para recomendação de atrações), restrições alimentares, além de problemas de saúde e preferências de passeio. Para manter a privacidade do usuário, alguns dados são omitidos e não são solicitados, como nome, endereço, entre outros. Após essa coleta, o aplicativo envia as preferências e localização do usuário para o servidor de aplicativos, gerenciado pelo parque, o que permite ao servidor ser capaz de produzir, de forma dinâmica, um itinerário personalizado para cada visitante. Como incentivo para uso do aplicativo e eliminação dos tickets físicos, o visitante concorre a prêmios e participa de promoções.

A solução também permite ao parque gerir e implantar seus recursos de forma eficiente, melhorando o fluxo de acesso às atrações e analisando vários índices de desempenho (filas, idade, tempo de espera de cada atração, atrações mas requisitadas, entre

outras). Com a localização dos visitantes, o parque também pode inferir sobre o comportamento de grupos e visitantes individuais. Os autores acreditam que com a solução proposta pode alcançar um nível satisfatório de segurança, porém não asseguram que o visitante tenha sua privacidade totalmente preservada, tendo em vista que algumas informação são armazenadas e existem fatores alheios ao escopo do trabalho como, por exemplo, uso de WiFi dentro do parque.

No que diz respeito a segurança, os autores desenvolveram um protocolo para autenticação anônima dos visitantes (AAV). O protocolo utiliza pseudônimos para dissociar os dados dos visitantes de suas verdadeiras identidades e um procedimento de ofuscação, implementado através de um esquema de assinatura parcialmente cega (fase de emissão de pseudônimos certificados). Assim, nenhum pseudônimo de visitante é revelado para o servidor de aplicativo. Como o servidor de aplicativo não tem nenhum papel na geração dos pseudônimos, ele não pode vincular o pseudônimo de qualquer visitante com qualquer ticket ou cartão de crédito utilizado para a compra de bilhetes. Segundo os autores, o protocolo alcança uma autenticação anônima, através do qual o servidor de aplicativo aceita dados somente de pseudônimos que foram certificados.

Para lidar com ataques de falsa localização, o protocolo usa a abordagem de [He et al. 2011], onde o servidor formula e poem em prática certas heurísticas, tais como calcular o tempo decorrido entre o local anterior dos visitantes e a localização atual. Se esse tempo coincide com o tempo médio gasto por outros visitantes entre os mesmos dois locais, os dados são considerados legítimos. Sempre que o servidor de aplicativo identifica que os dados de um determinado pseudônimo não corresponde a estas heurísticas e os valores limite, o servidor pode revogar o pseudônimo certificado e negar todas as comunicações futuras. O protocolo também usa HTTPS, por padrão, no canal de comunicação entre o aplicativo e o servidor.

SPPEAR: Security Privacy-preserving Architecture for Participatory-sensing Applications

Com o foco direcionado em preservar a privacidade dos participantes e permitir a concessão de incentivos aos participantes, Gisdakis et al. [Gisdakis et al. 2014] propuseram um arquitetura para Sensoriamento Participativo que faz uso de pseudônimos. Denominada SPPEAR, a arquitetura é abrangente e segura, e capaz de: (i) ser escalável, confiável e aplicável a qualquer tipo de aplicação de sensoriamento participativo e MCS; (ii) garantir a não identificação dos usuários, oferecendo forte proteção a privacidade; (iii) limitar a participação de usuários legítimos de forma totalmente responsável; (iv) evitar de forma eficiente que as identidades dos usuários sejam reveladas; (v) ser resiliente as entidades participantes; e (vi) poder suportar vários mecanismos de incentivo de forma a preservar a privacidade.

A arquitetura SPPEAR é formada por:

- **Usuários**: Atuam tanto como produtores de informação (ou seja, enviando dados) quanto consumidores de informação (ou seja, solicitando informações do sistema). Os dispositivos de usuários com capacidades de sensoriamento participam das tare-

fas submetendo amostras autenticadas ou por meio de consulta aos dados coletados.

- **Serviço de Tarefas - TS**: Esta entidade inicia tarefas e campanhas de sensoriamento. Também, define e fornece as recompensas que os participantes receberão por suas contribuições.

- **Gerente de Grupo - GM**: É responsável por registrar os dispositivos do usuário e emitir as credenciais anônimas para eles. Além disso, o GM autoriza a participação de dispositivos em várias tarefas de forma indiferente, usando tokens de autorização.

- **Provedor de identidade - IdP**: Oferece serviços de gerenciamento e identificação de credenciais (por exemplo, autenticação de usuário e controle de acesso, entre outros) para o sistema.

- **Certificação pseudônimo Authority - PCA**: Fornece credenciais anônimos efêmeras, denominadas pseudônimos, para os dispositivos. Um pseudônimo é um certificado X.509, que se liga a uma identidade anônima com uma chave pública e que possui uma validade, medida em tempo. PCA gera a quantidade desejada de pares de chaves e gera o mesmo número de assinaturas de pedidos de certificado.

- **Serviços de Agregação de Amostras - SAS**: Os dispositivos de usuário mandam amostras para esta entidade que é responsável por armazenar e processar os dados coletados. Para cada amostra submetida autêntico, o SAS emite um recibo para o dispositivo, que depois o envia para reivindicar créditos para a tarefa de detecção. O SAS possui interfaces que permitem a qualquer usuário autenticado e autorizado consultar os resultados das tarefas de sensoriamento e campanhas.

- **Autoridade de Resolução - RA**: É a entidade responsável pela revogação do anonimato dos dispositivos (por exemplo, dispositivos que perturbam o sistema ou poluem o processo de recolha de dados).

A Figura 2.8 apresenta a arquitetura SPPEAR.

Uma vez que separa processos e funções através de entidades, de acordo com o princípio de separação de direitos: "a cada entidade é dado o mínimo de informações necessárias para executar a tarefa desejada", SPPEAR atinge os objetivos de segurança e confiabilidade para qualquer aplicação.

Participatory Privacy: Enabling Privacy in Participatory Sensing

Em [De Cristofaro and Soriente 2013], Cristofaro e Soriente propuseram uma arquitetura de reforço a privacidade para aplicações MCS. A arquitetura proposta, denominada PEPSI (Figura 2.9) faz uso de um provedor de serviços de aplicações que gerencia as campanhas MCS. Os participantes (chamados de nós móveis) enviam relatórios dos dados de sensoriamento para o provedor de serviços, que após processar os dados, encaminha os relatórios aos usuários finais. PEPSI envolve uma quarta entidade, designada

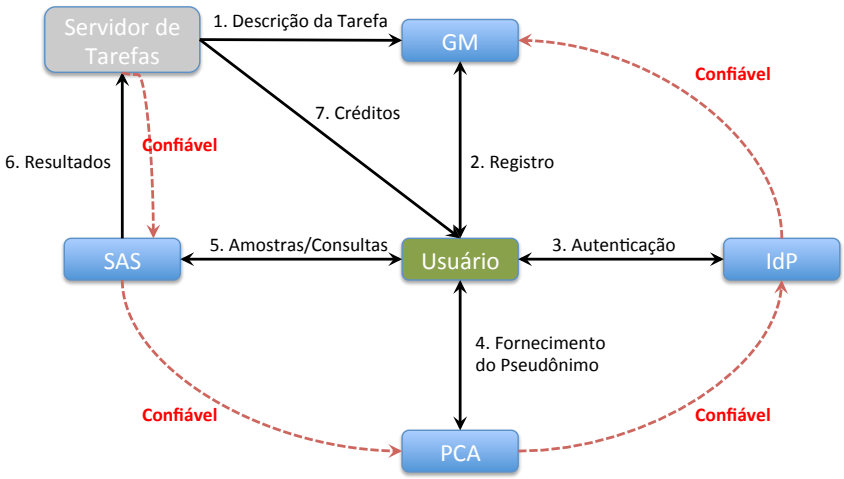

Figura 2.8. Visão geral da arquitetura SPPEAR

como autoridade de registro. Ela é responsável pela definição inicial dos parâmetros do sistema e pelo registro dos nós móveis e usuários finais.

Uma vez que um dos principais objetivos da PEPSI é ocultar relatórios de dados e consultas de pessoal não autorizado, todos os dados, relatórios e consultas são criptografadas. PEPSI usa Identidade baseada em Criptografia (*Identity-base Encryption* - IBE). Cada relatório é identificado por um conjunto de rótulos (etiquetas), que são usados ??como identidades. Assim, nós móveis podem derivar uma chave de criptografia pública única a partir desses rótulos e usá-la para criptografar os relatórios a serem transmitidos ao provedor de serviço. Durante o registro, nós móveis registram esses rótulos para a autoridade de registro, que, em seguida, atua como o gerador de chave privada do sistema IBE. Ele gera a chave de decifração privada correspondente a esses rótulos e passa a chave privada para os usuários finais interessante sobre seu registro.

Outra característica de PEPSI é que o provedor de serviços não apenas retransmiti todos os relatórios a todos os usuários finais. Isto geraria uma carga de processamento pesada para os usuários finais, porque eles precisariam tentar todas as suas chaves privadas para descriptografar um relatório. Ao invés disso, o provedor de serviços impõe um mecanismo de marcação através do qual ele pode combinar, de forma eficiente, os relatórios com as consultas. O mecanismo exige que os nós móveis etiquetem cada relatório com um token criptográfico que identifica o tipo de relatório apenas para usuários finais autorizados, sem vazamento de nenhuma informação do relatório. A marca é calculada a partir das mesmas etiquetas utilizadas para derivar a chave pública. Ao mesmo tempo, devido às propriedades de mapeamento bilinear inerente a IBE, usuários finais podem computar a mesma marca para o mesmo relatório usando suas chaves de criptografia privadas e fornecer a etiqueta para o provedor de serviços quando fizer uma subscrição de consulta. Em seguida, o provedor de serviços apenas encaminha um relatório marcado aos usuários finais que forneceram a mesma marca.

Towards a Practical Deployment of Privacy-preserving Crowd-sensing Tasks

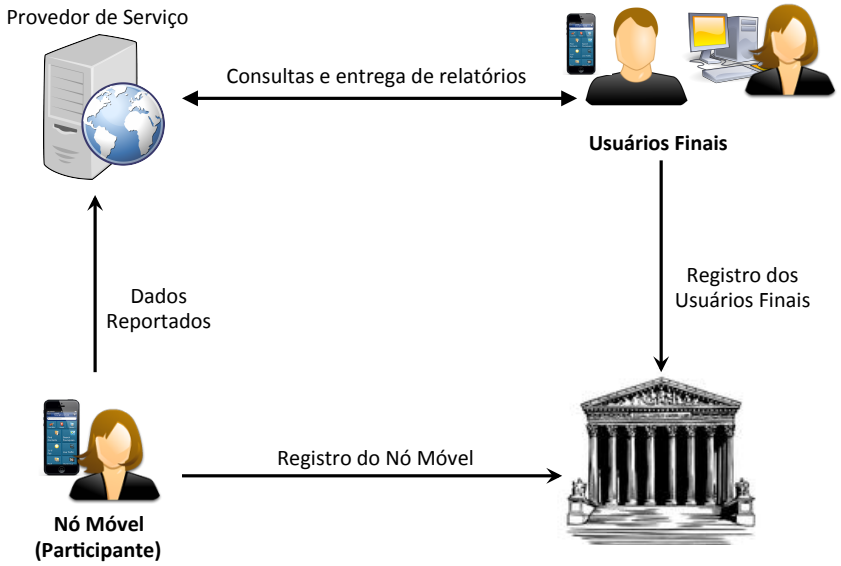

Figura 2.9. Arquitetura PEPSI. Fonte: [De Cristofaro and Soriente 2013]

Haderer et al. [Haderer et al. 2014] apresentaram uma nova plataforma de *Crowdsourcing* capaz de preservar a privacidade através da união de dois middlewares conhecidos: APISENSE e PRIVAPI.

APISENSE é uma plataforma de *Crowdsourcing* móvel que facilita a implantação de experiências MCS por cuidar dos desafios críticos neste domínio [Haderer et al. 2013]. Ela fornece uma plataforma de *Software-as-a-Service*, onde experimentos são descritos como scripts que serão enviados para os dispositivos móveis, a fim de recolher dados. A arquitetura distribuída do APISENSE é formada pelo: (i) serviço Hive, responsável por gerenciar a comunidade de usuários móveis e publicar as tarefas de sensoriamento; (ii) Honeypot, utilizado para descrever as tarefas como scripts (com base em uma extensão de JavaScript) e enviá-los para (iii) os dispositivos móveis, que recebem e executam as tarefas. A plataforma APISENSE apoia a implementação de diferentes estratégias de incentivo, incluindo *feedback* do usuário, *ranking* do usuário e recompensas. A seleção de estratégias de incentivo depende da natureza das experiências *Crowdsourcing*. No que diz respeito a privacidade, uma camada no dispositivo móvel implementa vários algoritmos para filtrar e "perturbar" informações sensíveis (por exemplo, catálogo de endereços, localização) dependendo das preferências do usuário. O usuário mantém o controle do seu dispositivo móvel para seleccionar os sensores, bem como quando e onde estes sensores pode ser utilizados pela plataforma.

PRIVAPI é um middleware de preservação da privacidade que pode ser facilmente integrado no topo da APISENSE. Seu objetivo é pré-processar os dados recolhidos de mobilidade antes de ser liberado. Graças ao seu conhecimento sobre o conjunto de dados inteiro, pode-se usar uma estratégia de anonimização ideal nos dados de mobilidade enquanto ainda oferece um nível satisfatório de utilidade.

A Figura 2.10 apresenta a arquitetura proposta.

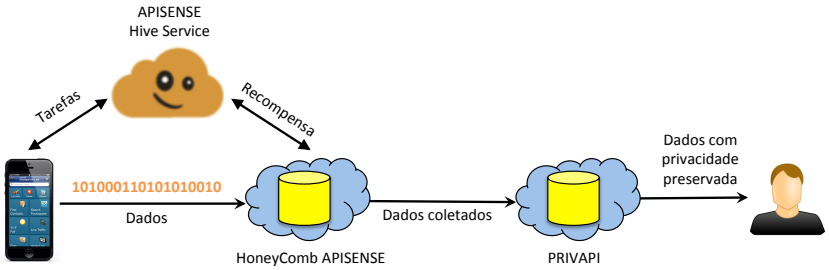

Figura 2.10. Visão geral da arquitetura proposta

2.4.2. Bibliotecas

Caché: Caching Location-Enhanced Content to Improve User Privacy

Amini et al. [Amini et al. 2011] propuseram a criação de uma biblioteca capaz que atender a uma quantidade significativa de aplicações que realizam a busca de dados de localização dos usuários. Denominada Caché, o objetivo da biblioteca, além de dar suporte o outros aplicativos, é oferecer aos usuários um nível de privacidade aceitável, tendo em vista a crescente preocupação com a exposição de informações sensíveis. Arquitetada para trabalhar como um repositório de localização, aplicada a privacidade, Caché atua realizando uma pré-busca das informações solicitadas antes que o usuário necessite. Como a busca é feita localmente, o usuário evita que sua localização seja exposta no momento que realmente esteja utilizado-as. Assim, ao invés de compartilhar a localização atual em cada pedido de informação, o usuário só precisa compartilhar o período de tempo.

A biblioteca funciona da seguinte maneira (Figura 2.11). Em primeiro lugar, existe a necessidade do aplicativo se adequar ao Caché. Por exemplo, o desenvolvedor do aplicativo deve fornecer algumas informações relevantes como a forma de fazer o download do conteúdo, URL de acesso e intervalo de atualização do conteúdo. Depois, o usuário instala o aplicativo habilitando o Caché e seleciona as regiões de seu interesse. Em seguida, o Caché realiza o download e atualiza o conteúdo com base na taxa de atualização que o desenvolvedor especificou. Como o Caché trabalha com bloco de informações relativamente grandes, é esperado o momento mais oportuno para baixar o conteúdo e atualiza-lo (por exemplo, quando o dispositivo móvel estiver conectado a uma conexão WiFi). A última etapa acontece quando a aplicação exige um conteúdo, que é recuperado do Caché ao invés de fazer uma consulta externa. Desta forma, o conteúdo a ser utilizado offline (sem necessidade de conexão de dados). Segundo o autor, sua arquitetura é semelhante a um proxy Internet (não transparente).

Embora não tenha sido criado especificamente para MCS, Caché, através do uso da pré-busca, permite o uso de conteúdo com capacidade de localização, garantindo a privacidade do usuário, sem contar com os benefícios relacionados ao download de dados. Caché também incentiva os desenvolvedores de aplicativos a elaborarem ideias para melhorar a privacidade do usuário como, por exemplo, trabalhar com esquemas melhores de amostragem para que os usuários não necessitem acessar conteúdo online e sem a necessidade de de infraestrutura de terceiros para garantir a privacidade dos usuários.

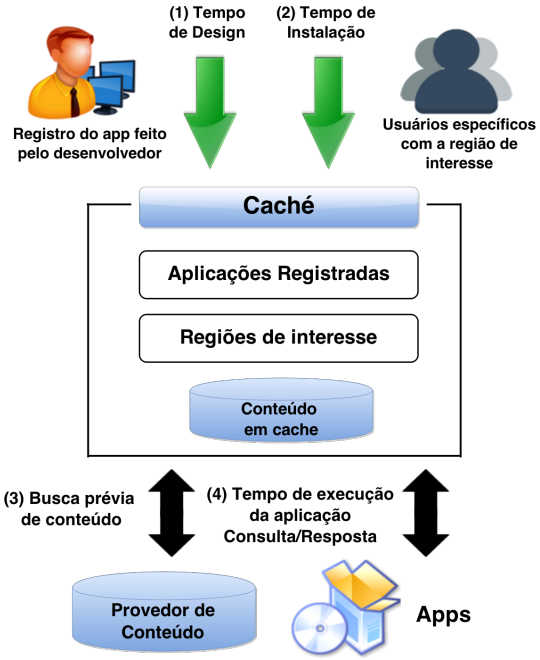

Figura 2.11. Funcionamento do Caché. Uma vez que o desenvolvedor tenha registrado a aplicação (1) e o usuário ter especificado as regiões para quais conteúdos devem ser armazenados em cache (2), o Caché faz o pedido e armazena o conteúdo (3) para uso futuro pela aplicação (4). Fonte: [Amini et al. 2011].

2.5. Pesquisas em Aberto

Neste Capítulo já foram discutidas as ameaças à segurança em MCS, bem como foram apresentados algumas soluções, ferramentas e bibliotecas seguras ou que fornecem segurança para MCS. O leitor deve ter percebido que as soluções são normalmente sob medida (para atender uma determinada situação) e que práticas para lidar com a privacidade são escassas. Assim, ainda existe uma ampla gama de desafios de pesquisa sem solução. Esta Seção destaca algumas pontos de pesquisa ainda em aberto.

2.5.1. Como inserir os participantes na questão da privacidade

Um dos principais desafios para novas gerações de sistemas e aplicações MCS é a inclusão dos participantes na questão da privacidade. Para tanto, os seguintes aspectos precisam ser estudados:

- **Privacidade sob medida**: Nas atuais soluções, a noção de privacidade é altamente individual e depende do ponto de vista e da opinião do usuário. Assim, é fundamental para criar consciência para as ameaças à privacidade ao usuário, e ajudá-lo, tornando a configuração complexa de aplicações de sensoriamento participativos de fácil compreensão. No entanto, a maioria das contramedidas de preservação da privacidade discutidos não apresentam uma interface de usuário que pode sensibilizar e facilitar a compreensão dos mecanismos complexos em uso. Interfaces de usuário personalizadas, que consideram as características originais de dispositivos móveis, são, portanto, muito procurados. Além disso, a existência de tais interfaces podem incentivar a aceitação e, mais tarde, a adoção dos mecanismos de preserva-

ção da privacidade por parte dos participantes, como eles serão capazes de entender melhor (através das interfaces) os meandros dos mecanismos.

- **Facilidade de uso**: A usabilidade das aplicações e suas configurações de privacidade precisam ser levadas em consideração. Toda vez que um participante é obrigado a realizar uma extensa e manual configuração de suas preferências, frequentemente sua paciência termina e ela acaba ou por deixar as configurações no modo padrão ou por marca quaisquer opções sem entender as implicações de suas escolhas. O estudo de Gross e Acquisti [Gross and Acquisti 2005] demonstra essa situação no caso das configurações de privacidade em redes sociais online.

- **A transparência dos níveis de protecção da privacidade**: Para avaliar se a protecção da privacidade oferecida é adequada, os usuários precisam ser capazes de comparar o nível oferecido de proteção contra seus requisitos de proteção individual. Embora a maioria das soluções pesquisadas baseiam suas avaliações em métricas matemáticos e verificáveis, a percepção do usuário e seu nível de satisfação com as soluções existentes não tem sido explicitamente considerado.

- **Incorporação de *feedback* do usuário**: Além de fornecer interfaces de usuário para configurar os níveis de privacidade, insights sobre como a proteção é percebida e para quais extensões os usuários estão envolvidos na configuração de suas configurações de privacidade são obrigados a apresentar a usabilidade.

Em linhas gerais, os usuários ainda são considerados no processo de avaliação da usabilidade e utilidade das soluções de privacidade. Os estudos existentes na área apenas correlacionam preocupações com a privacidade com as modalidades de sensoriamento usadas (Klasnja et al., 2009), ou analisam a forma como os participantes entendem, selecionam e se comportam com os métodos, por exemplo, de ocultação da localização (Brush et al. , 2010).

2.5.2. Adaptabilidade das Soluções de Privacidade

Toda vez que uma nova modalidades de sensoriamento é incorporada as plataformas de dispositivos móveis, surgem novas famílias de aplicações e surgem também novos desafios a privacidade. A capacidade de lidar com essa ampla gama de cenários é extremamente necessária. Assim, as futuras soluções de segurança e privacidade em MCS precisam ser combináveis e adaptáveis.

Entre as questões estão:

- **Aplicações independentes X soluções sob medida**: Algumas das soluções apresentadas neste Capítulo são ou foram adaptadas para cenários de específicos. Por exemplo, a ocultação de locais sensíveis, através da criação de registros falsos de localização para evitar correlações entre usuários e locais só foi avaliada no cenário aplicação PEIR [Mun et al. 2009]. Outros cenários precisam ser investigados para determinar os limites potenciais e os inconvenientes das soluções propostas em função das especificidades de aplicação. Esta investigação irá destacar as mudanças necessárias para proceder a partir de soluções de privacidade adaptados aos conceitos de privacidade de aplicação agnóstica.

- **Abordagem sistêmica**: Não existe, ou pelo menos os autores deste Capítulo não encontraram, uma arquitetura de privacidade flexível que aborde o problema do ponto de vista do sistema. Existem várias contramedidas onde os aspectos de privacidade são abordados.

- **Cenários evolutivos de sensoriamento**: Em cenários nos quais as características dos dados dos sensores são conhecidas antecipadamente, soluções de privacidade podem ser adaptadas em conformidade, por exemplo, pela adição de ruído com propriedades correspondentes. No entanto, em caso de sensoriamento de cenários dinâmicos e/ou imprevisíveis, em que as características dos dados do sensor não podem ser determinadas com antecedência, novos conceitos de privacidade precisam ser inventados.

2.5.3. *Trade-offs* entre Privacidade, Desempenho e Fidelidade dos Dados

Mecanismos robustos para proteção da privacidade (remoção ou ofuscação de leituras do sensor, por exemplo) podem influenciar a fidelidade de dados, o atraso no sensoriamento ou a integridade dos dados. No entanto, proteger a integridade dos dados do sensor neutraliza mecanismos de preservação da privacidade. Consequentemente, existe um *trade-off* entre as garantias de privacidade e fidelidade.

Já está claro que a participação do usuário precisa ser incentivada através da garantia da sua privacidade. Por outro lado, sistemas vulneráveis ou defeituosos podem contribuir para que dados corrompidos ou errados sejam utilizados pelas aplicações. Percebe-se então que existe um ***trade-off* entre anonimato e qualidade/integridade dos dados**. Para evitar que dados degradem a precisão dos resultados das aplicações, os dispositivos ou os dados em questão precisam ser identificados e eliminados a partir do conjunto de dispositivos encarregados do sensoriamento. Assim, a investigação sobre sistemas de reputação que servem tanto para o anonimato quanto para as exigências e especificidades dos cenários de sensoriamento é necessária.

Outro ponto que precisa de análise é a **proteção da privacidade de outras pessoas**. O trabalho de Tang et ai. (2010) demonstra que os participantes valorizam a privacidade da localização de seus amigos, mas a maioria dos mecanismos de preservação da privacidade atuais se concentram na proteção apenas próprios participantes (DietSense [] prova que os rostos de pessoas não envolvidas podem aparecer nas imagens da aplicação). O fato que é que os sistemas atuais, a função de proteger a privacidade dos outros é do usuário participante. Soluções automatizadas para minimizar os dados capturados de forma que ele não viole a privacidade dos outros é de grande interesse.

Por fim, embora a proteção de dados sensíveis seja altamente valorizado, em certas situações, como em cenários de emergência, podem ser necessários meios para **substituir ou sobrescrever as configurações de privacidade** especificadas pelos participantes. Esta questão pode ser comparada a encontrada no cenário de saúde, onde os médicos podem ser capazes de substituir o controle de acesso de sensores corporais para obter acesso a dados críticos de saúde.

2.5.4. Como medir privacidade?

Diferentes métodos, critérios ou métricas estão sendo usados para avaliar o desempenho das soluções propostas em termos de protecção da privacidade. Embora possa ser difícil ou mesmo impossível chegar a métricas universais para quantificar a privacidade, a necessidade de definir métricas generalizadas é amplamente reconhecida. Capturando o nível de proteção de privacidade, independentemente do cenário de aplicação particular, pode ser visto como uma meta de pesquisa de longo prazo, mas a definição dessas métricas é obrigatória para alcançar uma base comum para a comparação de mecanismos. Mas como obtê-las?

As métricas de privacidade empregadas atualmente precisam ser pesquisadas para determinar quais parâmetros de entrada (por exemplo, a quantidade de participantes na mesma região) são considerados necessários para calcular o grau de privacidade e qual é a natureza dos parâmetros de saída (por exemplo, a distância euclidiana entre os dados reais e os ocultados/perturbados), considerando os cenários de aplicação. Adicionalmente, as métricas de privacidade de outros domínios de aplicação devem ser analisadas em relação à sua aplicabilidade em MCS.

Além disso, analisando certas soluções, percebe-se a existência de uma entidade central para proteger a privacidade e anonimato do participante. No entanto, tais soluções não demonstram garantias ou provas de que o grau prometido de privacidade é respeitado, uma vez que detalhes de implementação dificilmente estão disponíveis e até mesmo a abordagem empregada para proteger a privacidade é normalmente desconhecida. A investigação sobre viabilidade dos mecanismos de privacidade ainda permanece como um campo de pesquisa em aberto.

2.5.5. Normas para Investigar Privacidade

Devido sua natureza sensível, conjuntos de dados públicos do mundo real para aplicações MCS são escassos. Por isso, a investigação da privacidade ocorre geralmente com conjuntos de dados privados ou sintéticos. Como resultado, a base de dados não é bem aceita para a avaliação de novos mecanismos, o que dificulta sua aferição em relação à outros.

Para superar essa limitação, a comunidade de pesquisa deve fornecer para conjuntos de dados abertos que podem servir como uma base para avaliações de desempenho e segurança. Isso inclui conjuntos de dados do mundo real, bem como conjuntos de dados sintéticos representativos para várias modalidades de sensoriamentos diferentes.

Além disso, como as implementações de mecanismos de privacidade estão quase sempre indisponíveis para o público em geral, torna-se difícil ou mesmo impossível referência-las contra mecanismos propostos. Tornar a descrição técnica e detalhada da implementação ou sua própria execução disponível para a comunidade de pesquisa permite validar os resultados e as soluções individuais de referência.

2.6. Considerações Finais

Este Capítulo apresentou o paradigma de *Mobile Crowd Sensing* (MCS) e suas aplicações para o dia a dia. A Seção 2.2 tratou da explicar MCS. Primeiro, a evolução das ideais sobre sobre sensoriamento foi contextualizada. Em seguida, definições sobre MCS foram

feitas e os componentes e o ciclo de vida explicados. Depois, as características únicas de MCS foram enumeradas. Por fim, algumas aplicações MCS foram apresentadas.

A Seção 2.3 tratou os aspectos de segurança em MCS. Os dois principais problemas de segurança (privacidade do usuário e confiabilidade dos dados) foram bem discutidos. Os problemas e também as técnicas para resolvê-los foram apresentadas. Perto do fim, as soluções existentes (Seção 2.4) baseadas em segurança para MCS foram discutidas. A Seção 2.5 apontou várias questões em aberto

2.6.1. Observações Finais

A longo prazo, MCS é capaz de estimular a pesquisa em uma série de domínios.

Em primeiro lugar, o mundo hoje consiste de espaços físicos e virtuais, onde qualquer objeto sensoriado está entre esses espaços. Assim, é importante para explorar abordagens para agregação e fusão dos dados complementares para o melhor entendimento.

Em segundo lugar, ainda é preciso estudar a fusão da inteligência humana e da máquina em todo o ciclo de vida de MCS, desde o sensoriamento, transmissão de dados e processamento de dados.

Em terceiro lugar, alguns dos fatores éticos, como a inventividade e privacidade do usuário, devem ser os blocos de construção fundamentais de arquiteturas MCS.

Finalmente, o sucesso de MCS baseia-se na utilização de conhecimentos multidisciplinares, incluindo ciências sociais, ciência cognitiva, economia, ciência de computação, e assim por diante. Todas essas áreas devem ser considerada no desenho de técnicas e sistemas MCS.

Referências

[Acampora et al. 2013] Acampora, G., Cook, D., Rashidi, P., and Vasilakos, A. (2013). A survey on ambient intelligence in healthcare. *Proceedings of the IEEE*, 101(12):2470–2494.

[Amini et al. 2011] Amini, S., Lindqvist, J., Hong, J., Lin, J., Toch, E., and Sadeh, N. (2011). Caché: caching location-enhanced content to improve user privacy. In *Proceedings of the 9th international conference on Mobile systems, applications, and services*, pages 197–210. ACM.

[Burke et al. 2006] Burke, J., Estrin, D., Hansen, M., Parker, A., Ramanathan, N., Reddy, S., and Srivastava, M. B. (2006). Participatory sensing. In *In: Workshop on World-Sensor-Web (WSW'06): Mobile Device Centric Sensor Networks and Applications*, pages 117–134.

[Capkun et al. 2006] Capkun, S., Cagalj, M., and Srivastava, M. (2006). Secure localization with hidden and mobile base stations. In *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, pages 1–10.

[Christin et al. 2011] Christin, D., Reinhardt, A., Kanhere, S. S., and Hollick, M. (2011). A survey on privacy in mobile participatory sensing applications. *Journal of Systems and Software*, 84(11):1928 – 1946. Mobile Applications: Status and Trends.

[Christin et al. 2012] Christin, D., Rosskopf, C., Hollick, M., Martucci, L., and Kanhere, S. (2012). Incognisense: An anonymity-preserving reputation framework for participatory sensing applications. In *Pervasive Computing and Communications (PerCom), 2012 IEEE International Conference on*, pages 135–143.

[Clementi et al. 2013] Clementi, A., Pasquale, F., and Silvestri, R. (2013). Opportunistic manets: Mobility can make up for low transmission power. *IEEE/ACM Trans. Netw.*, 21(2):610–620.

[Conti and Kumar 2010] Conti, M. and Kumar, M. (2010). Opportunities in opportunistic computing. *Computer*, 43(1):42–50.

[Das et al. 2010] Das, T., Mohan, P., Padmanabhan, V. N., Ramjee, R., and Sharma, A. (2010). Prism: Platform for remote sensing using smartphones. In *Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services*, MobiSys '10, pages 63–76, New York, NY, USA. ACM.

[De Cristofaro and Soriente 2013] De Cristofaro, E. and Soriente, C. (2013). Participatory privacy: Enabling privacy in participatory sensing. *Network, IEEE*, 27(1):32–36.

[Deng and Cox 2009] Deng, L. and Cox, L. P. (2009). Livecompare: Grocery bargain hunting through participatory sensing. In *Proceedings of the 10th Workshop on Mobile Computing Systems and Applications*, HotMobile '09, pages 4:1–4:6, New York, NY, USA. ACM.

[Derawi et al. 2010] Derawi, M. O., Nickel, C., Bours, P., and Busch, C. (2010). Unobtrusive user-authentication on mobile phones using biometric gait recognition. In *Proceedings of the 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, IIH-MSP '10, pages 306–311, Washington, DC, USA. IEEE Computer Society.

[Dimov 2014] Dimov, D. (2014). Crowdsensing: State of the art and privacy aspects. `http://resources.infosecinstitute.com/crowdsensing-state-art-privacy-aspects/`.

[Domingo-Ferrer and Mateo-Sanz 2002] Domingo-Ferrer, J. and Mateo-Sanz, J. M. (2002). Practical data-oriented microaggregation for statistical disclosure control. *IEEE Trans. on Knowl. and Data Eng.*, 14(1):189–201.

[Dong et al. 2008] Dong, Y. F., Kanhere, S., Chou, C. T., and Bulusu, N. (2008). Automatic collection of fuel prices from a network of mobile cameras. In *Proceedings of the 4th IEEE International Conference on Distributed Computing in Sensor Systems*, DCOSS '08, pages 140–156, Berlin, Heidelberg. Springer-Verlag.

[Dua et al. 2009] Dua, A., Bulusu, N., Feng, W.-C., and Hu, W. (2009). Towards trustworthy participatory sensing. In *Proceedings of the 4th USENIX Conference on Hot Topics in Security*, HotSec'09, pages 8–8, Berkeley, CA, USA. USENIX Association.

[Dutta et al. 2009] Dutta, P., Aoki, P. M., Kumar, N., Mainwaring, A., Myers, C., Willett, W., and Woodruff, A. (2009). Common sense: Participatory urban sensing using a network of handheld air quality monitors. In *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*, SenSys '09, pages 349–350, New York, NY, USA. ACM.

[Eisenman et al. 2007] Eisenman, S. B., Miluzzo, E., Lane, N. D., Peterson, R. A., Ahn, G.-S., and Campbell, A. T. (2007). The bikenet mobile sensing system for cyclist experience mapping. In *Proceedings of the 5th International Conference on Embedded Networked Sensor Systems*, SenSys '07, pages 87–101, New York, NY, USA. ACM.

[Ganti et al. 2011] Ganti, R., Ye, F., and Lei, H. (2011). Mobile crowdsensing: current state and future challenges. *Communications Magazine, IEEE*, 49(11):32–39.

[Gao and Cao 2011] Gao, W. and Cao, G. (2011). User-centric data dissemination in disruption tolerant networks. In *INFOCOM, 2011 Proceedings IEEE*, pages 3119–3127.

[Gaonkar et al. 2008] Gaonkar, S., Li, J., Choudhury, R. R., Cox, L., and Schmidt, A. (2008). Micro-blog: Sharing and querying content through mobile phones and social participation. In *Proceedings of the 6th International Conference on Mobile Systems, Applications, and Services*, MobiSys '08, pages 174–186, New York, NY, USA. ACM.

[Gisdakis et al. 2014] Gisdakis, S., Giannetsos, T., and Papadimitratos, P. (2014). Sppear: Security & privacy-preserving architecture for participatory-sensing applications. In *Proceedings of the 2014 ACM Conference on Security and Privacy in Wireless & Mobile Networks*, WiSec '14, pages 39–50, New York, NY, USA. ACM.

[Google 2015] Google (2015). Waze. `http://waze.com`.

[Grosky et al. 2007] Grosky, W., Kansal, A., Nath, S., Liu, J., and Zhao, F. (2007). Senseweb: An infrastructure for shared sensing. *MultiMedia, IEEE*, 14(4):8–13.

[Gross and Acquisti 2005] Gross, R. and Acquisti, A. (2005). Information revelation and privacy in online social networks. In *Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society*, WPES '05, pages 71–80, New York, NY, USA. ACM.

[Guo et al. 2015] Guo, B., Wang, Z., Yu, Z., Wang, Y., Yen, N. Y., Huang, R., and Zhou, X. (2015). Mobile crowd sensing and computing: The review of an emerging human-powered sensing paradigm. *ACM Comput. Surv.*, 48(1):7:1–7:31.

[Guo et al. 2013] Guo, B., Zhang, D., Wang, Z., Yu, Z., and Zhou, X. (2013). Opportunistic iot: Exploring the harmonious interaction between human and the internet of things. *J. Netw. Comput. Appl.*, 36(6):1531–1539.

[Haderer et al. 2014] Haderer, N., Primault, V., Raveneau, P., Ribeiro, C., Rouvoy, R., and Ben Mokhtar, S. (2014). Towards a Practical Deployment of Privacy-preserving Crowd-sensing Tasks. In *Middleware Posters and Demos '14*, Bordeaux, France.

[Haderer et al. 2013] Haderer, N., Rouvoy, R., and Seinturier, L. (2013). Dynamic deployment of sensing experiments in the wild using smartphones. In Dowling, J. and TaÃ¯ani, F., editors, *Distributed Applications and Interoperable Systems*, volume 7891 of *Lecture Notes in Computer Science*, pages 43–56. Springer Berlin Heidelberg.

[He et al. 2011] He, W., Liu, X., and Ren, M. (2011). Location cheating: A security challenge to location-based social network services. In *Distributed Computing Systems (ICDCS), 2011 31st International Conference on*, pages 740–749.

[Howe 2006] Howe, J. (2006). Crowdsourcing: A definition. *Crowdsourcing: Tracking the rise of the amateur*.

[Hu et al. 2013] Hu, X., Liu, Q., Zhu, C., Leung, V. C. M., Chu, T. H. S., and Chan, H. C. B. (2013). A mobile crowdsensing system enhanced by cloud-based social networking services. In *Proceedings of the First International Workshop on Middleware for Cloud-enabled Sensing*, MCS '13, pages 3:1–3:6, New York, NY, USA. ACM.

[Huang et al. 2010a] Huang, K. L., Kanhere, S. S., and Hu, W. (2010a). Are you contributing trustworthy data?: The case for a reputation system in participatory sensing. In *Proceedings of the 13th ACM International Conference on Modeling, Analysis, and Simulation of Wireless and Mobile Systems*, MSWIM '10, pages 14–22, New York, NY, USA. ACM.

[Huang et al. 2010b] Huang, K. L., Kanhere, S. S., and Hu, W. (2010b). Preserving privacy in participatory sensing systems. *Computer Communications*, 33(11):1266 – 1280.

[IBM 2010] IBM (2010). Creekwatch: Explore your watershed. `http://creekwatch.researchlabs.ibm.com`.

[Kapadia et al. 2009] Kapadia, A., Kotz, D., and Triandopoulos, N. (2009). Opportunistic sensing: Security challenges for the new paradigm. In *Communication Systems and Networks and Workshops, 2009. COMSNETS 2009. First International*, pages 1–10.

[Karamshuk et al. 2011] Karamshuk, D., Boldrini, C., Conti, M., and Passarella, A. (2011). Human mobility models for opportunistic networks. *Communications Magazine, IEEE*, 49(12):157–165.

[Kong et al. 2015] Kong, L., He, L., Liu, X.-Y., Gu, Y., Wu, M.-Y., and Liu, X. (2015). Privacy-preserving compressive sensing for crowdsensing based trajectory recovery. In *Distributed Computing Systems (ICDCS), 2015 IEEE 35th International Conference on*, pages 31–40.

[Konidala et al. 2013] Konidala, D., Deng, R., Li, Y., Lau, H., and Fienberg, S. (2013). Anonymous authentication of visitors for mobile crowd sensing at amusement parks. In Deng, R. and Feng, T., editors, *Information Security Practice and Experience*, volume 7863 of *Lecture Notes in Computer Science*, pages 174–188. Springer Berlin Heidelberg.

[Krumm 2007] Krumm, J. (2007). Inference attacks on location tracks. In *Proceedings of the 5th International Conference on Pervasive Computing*, PERVASIVE'07, pages 127–143, Berlin, Heidelberg. Springer-Verlag.

[Lane et al. 2010] Lane, N. D., Miluzzo, E., Lu, H., Peebles, D., Choudhury, T., and Campbell, A. T. (2010). A survey of mobile phone sensing. *Comm. Mag.*, 48(9):140–150.

[Leonardi et al. 2014] Leonardi, C., Cappellotto, A., Caraviello, M., Lepri, B., and Antonelli, F. (2014). Secondnose: An air quality mobile crowdsensing system. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*, NordiCHI '14, pages 1051–1054, New York, NY, USA. ACM.

[Levy and da Costa 1993] Levy, P. and da Costa, C. I. (1993). *tecnologias da inteligência, As*. Editora 34.

[Li et al. 2014] Li, Q., Cao, G., and La Porta, T. (2014). Efficient and privacy-aware data aggregation in mobile sensing. *Dependable and Secure Computing, IEEE Transactions on*, 11(2):115–129.

[Liu 2007] Liu, L. (2007). From data privacy to location privacy: Models and algorithms. In *Proceedings of the 33rd International Conference on Very Large Data Bases*, VLDB '07, pages 1429–1430. VLDB Endowment.

[Lu et al. 2010] Lu, H., Lane, N. D., Eisenman, S. B., and Campbell, A. T. (2010). Fast track article: Bubble-sensing: Binding sensing tasks to the physical world. *Pervasive Mob. Comput.*, 6(1):58–71.

[Ludwig et al. 2015] Ludwig, T., Reuter, C., Siebigteroth, T., and Pipek, V. (2015). Crowdmonitor: Mobile crowd sensing for assessing physical and digital activities of citizens during emergencies. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 4083–4092, New York, NY, USA. ACM.

[Ma et al. 2014] Ma, H., Zhao, D., and Yuan, P. (2014). Opportunities in mobile crowd sensing. *Communications Magazine, IEEE*, 52(8):29–35.

[Machanavajjhala et al. 2007] Machanavajjhala, A., Kifer, D., Gehrke, J., and Venkitasubramaniam, M. (2007). L-diversity: Privacy beyond k-anonymity. *ACM Trans. Knowl. Discov. Data*, 1(1).

[Maisonneuve et al. 2009] Maisonneuve, N., Stevens, M., Niessen, M., and Steels, L. (2009). Noisetube: Measuring and mapping noise pollution with mobile phones. In Athanasiadis, I. N., Rizzoli, A. E., Mitkas, P. A., and Gomez, J. M., editors, *Information Technologies in Environmental Engineering*, Environmental Science and Engineering, pages 215–228. Springer Berlin Heidelberg.

[Maisonneuve et al. 2010] Maisonneuve, N., Stevens, M., and Ochab, B. (2010). Participatory noise pollution monitoring using mobile phones. *Info. Pol.*, 15(1,2):51–71.

[Mathur et al. 2010] Mathur, S., Jin, T., Kasturirangan, N., Chandrasekaran, J., Xue, W., Gruteser, M., and Trappe, W. (2010). Parknet: Drive-by sensing of road-side parking statistics. In *Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services*, MobiSys '10, pages 123–136, New York, NY, USA. ACM.

[Mediated Spaces, Inc 2015] Mediated Spaces, Inc (2015). The wildlab: Use mobile technology to explore, discovery, and share the natural world. `http://thewildlab.org`.

[Miluzzo et al. 2008] Miluzzo, E., Lane, N. D., Fodor, K., Peterson, R., Lu, H., Musolesi, M., Eisenman, S. B., Zheng, X., and Campbell, A. T. (2008). Sensing meets mobile social networks: The design, implementation and evaluation of the cenceme application. In *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems*, SenSys '08, pages 337–350, New York, NY, USA. ACM.

[Minkman et al. 2015] Minkman, E., van Overloop, P., and van der Sanden, M. (2015). Citizen science in water quality monitoring: Mobile crowd sensing for water management in the netherlands. In *World Environmental and Water Resources Congress 2015*, pages 1399–1408.

[Minson et al. 2015] Minson, S. E., Brooks, B. A., Glennie, C. L., Murray, J. R., Langbein, J. O., Owen, S. E., Heaton, T. H., Iannucci, R. A., and Hauser, D. L. (2015). Crowdsourced earthquake early warning. *Science Advances*, 1(3):e1500036+.

[Mohan et al. 2008] Mohan, P., Padmanabhan, V., and Ramjee, R. (2008). Nericell: Rich monitoring of road and traffic conditions using mobile smartphones. In *ACM Sensys*. Association for Computing Machinery, Inc. Raleigh, NC, USA.

[Mun et al. 2010] Mun, M., Hao, S., Mishra, N., Shilton, K., Burke, J., Estrin, D., Hansen, M., and Govindan, R. (2010). Personal data vaults: A locus of control for personal data streams. In *Proceedings of the 6th International COnference*, Co-NEXT '10, pages 17:1–17:12, New York, NY, USA. ACM.

[Mun et al. 2009] Mun, M., Reddy, S., Shilton, K., Yau, N., Burke, J., Estrin, D., Hansen, M., Howard, E., West, R., and Boda, P. (2009). Peir, the personal environmental impact report, as a platform for participatory sensing systems research. In *Proceedings of the 7th International Conference on Mobile Systems, Applications, and Services*, MobiSys '09, pages 55–68, New York, NY, USA. ACM.

[Pan et al. 2013] Pan, B., Zheng, Y., Wilkie, D., and Shahabi, C. (2013). Crowd sensing of traffic anomalies based on human mobility and social media. In *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, SIGSPATIAL'13, pages 344–353, New York, NY, USA. ACM.

[Pournajaf et al. 2014] Pournajaf, L., Xiong, L., Garcia-Ulloa, D. A., and Sunderam, V. (2014). A survey on privacy in mobile crowd sensing task management. Technical report, Technical Report TR-2014-002, Department of Mathematics and Computer Science, Emory University.

[Rachuri et al. 2011] Rachuri, K. K., Mascolo, C., Musolesi, M., and Rentfrow, P. J. (2011). Sociablesense: Exploring the trade-offs of adaptive sampling and computation offloading for social sensing. In *Proceedings of the 17th Annual International Conference on Mobile Computing and Networking*, MobiCom '11, pages 73–84, New York, NY, USA. ACM.

[Rana et al. 2010] Rana, R. K., Chou, C. T., Kanhere, S. S., Bulusu, N., and Hu, W. (2010). Ear-phone: An end-to-end participatory urban noise mapping system. In *Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks*, IPSN '10, pages 105–116, New York, NY, USA. ACM.

[Reddy et al. 2007] Reddy, S., Parker, A., Hyman, J., Burke, J., Estrin, D., and Hansen, M. (2007). Image browsing, processing, and clustering for participatory sensing: Lessons from a dietsense prototype. In *Proceedings of the 4th Workshop on Embedded Networked Sensors*, EmNets '07, pages 13–17, New York, NY, USA. ACM.

[Reddy et al. 2009] Reddy, S., Samanta, V., Burke, J., Estrin, D., Hansen, M., and Srivastava, M. (2009). Mobisense: mobile network services for coordinated participatory sensing. In *Autonomous Decentralized Systems, 2009. ISADS '09. International Symposium on*, pages 1–6.

[Sherchan et al. 2012] Sherchan, W., Jayaraman, P., Krishnaswamy, S., Zaslavsky, A., Loke, S., and Sinha, A. (2012). Using on-the-move mining for mobile crowdsensing. In *Mobile Data Management (MDM), 2012 IEEE 13th International Conference on*, pages 115–124.

[Shi et al. 2010] Shi, J., Zhang, R., Liu, Y., and Zhang, Y. (2010). Prisense: Privacy-preserving data aggregation in people-centric urban sensing systems. In *Proceedings of the 29th Conference on Information Communications*, INFOCOM'10, pages 758–766, Piscataway, NJ, USA. IEEE Press.

[Shilton 2009] Shilton, K. (2009). Four billion little brothers?: Privacy, mobile phones, and ubiquitous data collection. *Commun. ACM*, 52(11):48–53.

[Shilton et al. 2008] Shilton, K., Burke, J., Estrin, D., Hansen, M., and Srivastava, M. (2008). Participatory privacy in urban sensing. In *Proceedings of the International Workshop on Mobile Devices and Urban Sensing*, MODUS, pages 1–7.

[Shin et al. 2011] Shin, M., Cornelius, C., Peebles, D., Kapadia, A., Kotz, D., and Triandopoulos, N. (2011). Anonysense: A system for anonymous opportunistic sensing. *Pervasive and Mobile Computing*, 7(1):16 – 30.

[Surowiecki 2005] Surowiecki, J. (2005). *The Wisdom of Crowds*. Anchor.

[Sweeney 2002] Sweeney, L. (2002). K-anonymity: A model for protecting privacy. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, 10(5):557–570.

[Thepvilojanapong et al. 2010] Thepvilojanapong, N., Ono, T., and Tobe, Y. (2010). A deployment of fine-grained sensor network and empirical analysis of urban temperature. *Sensors*, 10(3):2217.

[Tuncay et al. 2012] Tuncay, G. S., Benincasa, G., and Helmy, A. (2012). Autonomous and distributed recruitment and data collection framework for opportunistic sensing. In *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*, Mobicom '12, pages 407–410, New York, NY, USA. ACM.

[Vieira and Alves 2014] Vieira, A. P. and Alves, J. C. R. (2014). Direito à privacidade na sociedade da informação. *Revista Jus Navigandi*, (3979).

[Wang et al. 2011a] Wang, H., Uddin, M., Qi, G.-J., Huang, T., Abdelzaher, T., and Cao, G. (2011a). Photonet: A similarity-aware image delivery service for situation awareness. In *Information Processing in Sensor Networks (IPSN), 2011 10th International Conference on*, pages 135–136.

[Wang et al. 2013] Wang, L., Zhang, D., and Xiong, H. (2013). effsense: Energy-efficient and cost-effective data uploading in mobile crowdsensing. In *Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication*, UbiComp '13 Adjunct, pages 1075–1086, New York, NY, USA. ACM.

[Wang et al. 2011b] Wang, X., Govindan, K., and Mohapatra, P. (2011b). Collusion-resilient quality of information evaluation based on information provenance. In *Sensor, Mesh and Ad Hoc Communications and Networks (SECON), 2011 8th Annual IEEE Communications Society Conference on*, pages 395–403.

[Weppner and Lukowicz 2013] Weppner, J. and Lukowicz, P. (2013). Bluetooth based collaborative crowd density estimation with mobile phones. In *Pervasive Computing and Communications (PerCom), 2013 IEEE International Conference on*, pages 193–200.

[Yan et al. 2009] Yan, T., Marzilli, M., Holmes, R., Ganesan, D., and Corner, M. (2009). mcrowd: A platform for mobile crowdsourcing. In *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*, SenSys '09, pages 347–348, New York, NY, USA. ACM.

[Zhang et al. 2014a] Zhang, D., Wang, L., Xiong, H., and Guo, B. (2014a). 4w1h in mobile crowd sensing. *Communications Magazine, IEEE*, 52(8):42–48.

[Zhang et al. 2014b] Zhang, X., Yang, Z., Wu, C., Sun, W., Liu, Y., and Liu, K. (2014b). Robust trajectory estimation for crowdsourcing-based mobile applications. *Parallel and Distributed Systems, IEEE Transactions on*, 25(7):1876–1885.