Estratégias crowdsourcing para aplicativos de cidades Alternative Title: Crowdsourcing strategies for smart cities applications

Wancharle S. Quirino
Universidade Federal do
Espirito Santo - UFES
Av. Fernando Ferrari, S/N
- 29060-970 - ES - Brasil
wsquirino@inf.ufes.br

Celso A. S. Santos Universidade Federal do Espirito Santo - UFES Av. Fernando Ferrari, S/N - 29060-970 - ES - Brasil saibel@inf.ufes.br Juan X. E. A. Calles Universidade Federal do Espirito Santo - UFES Av. Fernando Ferrari, S/N - 29060-970 - ES - Brasil callesjuan@gmail.com Fernando F. Tinelli Universidade Federal do Espirito Santo - UFES Av. Fernando Ferrari, S/N - 29060-970 - ES - Brasil fftinelli@gmail.com

RESUMO

Crowdsourcing é um modelo de resolução de problemas através da contribuição de um grande número de pessoas e tem entre suas principais vantagens o baixo custo. Por outro lado, smart cities compreendem hoje um desafio multidisciplinar, onde têm-se por objetivo o desenvolvimento sustentável e o aumento da qualidade de vida dos seus habitantes. Assim sendo, vemos no crowdsourcing um recurso capaz de contribuir para construção de smart cities. Este artigo busca investigar a intersecção existente entre os domínios de smart cities e crowdsourcing e descobrir as lacunas e desafios que caracterizam este contexto de aplicações. Como resultado, propomos algumas estratégias para facilitar o desenvolvimento de aplicações crowdsourcing. Estas estratégias são então aplicadas para a construção de diversas aplicações deste tipo, sendo duas delas discutidas no fim do artigo.

Palavras-Chave

Crowdsourcing, smart cities, aplicações, plataforma.

ABSTRACT

Crowdsourcing is a problem-solving model through the contribution of a large number of people and has a low cost among its main advantages. On the other hand, smart cities today comprise a multidisciplinary challenge, where they have the objective of sustainable development and improving the quality of life of its inhabitants. Thus, we see in crowdsourcing a resource capable of contributing to building smart cities. This paper investigates the existing intersection between the fields of smart cities and crowdsourcing and discover the gaps and challenges that characterize this applications context. As a result, we propose some strategies to facilitate the development of crowdsourcing applications. These strategies are then applied to the construction of several applications of this type, two of which are discussed at the end of the article.

Categories and Subject Descriptors

[**Information Systems**]: World Wide Web – Web application – *crowdsourcing*.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SBSI 2016, May 17–20, 2016, Florianópolis, Santa Catarina, Brazil. Copyright SBC 2016.

General Terms

Design, Experimentation.

Keywords

Crowdsourcing, smart cities, applications, platform.

1. INTRODUÇÃO

Crowdsourcing é um modelo de resolução de problemas por meio da contribuição de um grande número de pessoas, chamando de sistema/serviço crowdsourcing àqueles que seguem esse modelo. Dentre alguns serviços atualmente ofertados que seguem esse modelo, ou parte dele, podemos citar Amazon Mechanical Turk, Waze, Foursquare, Wikipédia, YouTube e muitos outros. Duas das características que tornam atrativa a ideia do crowdsourcing são o baixo custo computacional e financeiro na geração de soluções e a facilidade de resolução de problemas que demandam elevado grau de cognição, pois as soluções são produto do fator humano [23].

De forma geral, as entidades envolvidas no modelo são (i) o dono do problema, ou beneficiado, é quem propõe e descreve o chamado do problema a ser resolvido; (ii) a plataforma de crowdsourcing, responsável por gerenciar/organizar o problema e as soluções geradas e; (iii) a crowd propriamente dita, geradora das soluções, incentivada por algum tipo de recompensa. Apesar de ser possível identificar cada uma dessas partes nos exemplos citados acima, identificação das entidades envolvidas torna-se mais subjetiva, pois elas variam de acordo com a perspectiva com a qual um determinado serviço/problema é analisado. Por exemplo, a organização Wikipédia pode ser vista como dona do problema, criar a maior enciclopédia jamais vista, onde recorre à crowd para criar todo o seu conteúdo [11,23].

Além das entidades, o modelo *crowdsourcing* pode ser caracterizado por muitas outras dimensões, como a natureza de colaboração (explícita, implícita ou mista); o tipo de chamado (aberto para todo mundo, limitado a certo grupo ou simplesmente não possui chamado algum); o tipo de *crowd* (composta por agentes internos ou externos à uma organização); o nível de conhecimento exigido da *crowd* (são necessários colaboradores com conhecimentos gerais, situacionais ou especializados sobre algum domínio); a arquitetura da plataforma (do tipo independente ou *standalone*, do tipo *piggyback*, isto é, depende de outras plataformas para seu funcionamento); a forma como é processado o retorno (manual ou automático); o incentivo para os participantes entre outras dimensões [8,10,11,15,16,23].

O conceito de *smart city* (cidade inteligente) está relacionado à construção de cidades em que se prioriza o uso eficiente dos recursos existentes ao invés do investimento e esforço na sua ampliação ou modificação que é mais custosa e sofre de

limitações físicas. De fato, busca-se o crescimento sustentável junto à melhoria na qualidade de vida de seus habitantes. Assim, as *smart cities* estão diretamente relacionadas a investimentos em TICs e, principalmente, no capital humano e social, pois esses são os fatores para promover a sustentabilidade. De forma concreta, experimentos em *smart cities* têm contribuído nas dimensões de governança, meio ambiente, mobilidade, saúde, segurança, transporte e outros, onde um dos elementos chave é a participação dos cidadãos [5,21,27].

Tendo os cidadãos um papel fundamental na construção de *smart citites*, é possível contextualizar os serviços *crowdsourcing* como ferramentas capazes de promover a realização destas "novas" cidades. Neste caso, identifica-se a *crowd* na população, considerando que o dono do problema poderia ser tanto um órgão que governa uma cidade (uma secretaria ou a prefeitura), quanto o próprio cidadão (partindo de uma iniciativa pública).

Assim sendo, neste estudo pretendemos investigar a interseção entre os domínios de *crowdsourcing* e *smart cities*, especificamente identificando sistemas de *crowdsourcing* cujos problemas abordados estão relacionados com *smart cities*. Nosso objetivo, portanto, é descobrir os desafios que existem no desenvolvimento de plataformas *crowdsourcing*, dando preferência aos desafios das plataformas utilizadas *smart cities* e *eGov*. E ao identificar algumas demandas, iremos propor algumas estratégias para resolve-las.

A organização das próximas seções é feita da seguinte forma: na seção 2 descrevemos o roteiro segundo o qual foi elaborada a pesquisa, uma análise quantitativa e qualitativa dos trabalhos encontrados e as demandas encontradas; na seção 3 apresentamos algumas estratégias que podem ser utilizadas para solucionar as demandas encontradas e algumas aplicações criadas para validar essas estratégias; finalmente expomos as nossas conclusões na seção 4.

2. CROWDSOURCING E SMART CITIES

Para tentar estabelecer um panorama do uso do modelo *crowdsourcing* em *smart cities*, utilizamos um processo de revisão de artigos como método de pesquisa. O processo foi dividido em quatro etapas: planejamento, pesquisa, extração de dados e análise dos resultados. Cada uma das etapas da revisão é detalhada nas seções seguintes.

2.1 Planejamento

Nesta etapa definimos as fontes de pesquisa, a questão de pesquisa utilizada para composição da expressão de busca e os critérios para inclusão e exclusão de trabalhos na classificação.

As bibliotecas digitais ACM Digital Library e IEEE Xplorer foram selecionadas como motores de busca por satisfazerem os critérios da pesquisa. Como critério, foram utilizadas apenas as bases com maior afinidade com áreas de estudos pesquisadas, com resultados menos redundantes que motores de buscas genéricos, com acesso irrestrito e com suporte ao uso de operadores lógicos na expressão de busca. Além disso, após a realização de buscas preliminares no Google Scholar, constatamos que parte significativa dos resultados relevantes para o trabalho eram provenientes das bases ACM e IEEE Xplorer.

O processo de elaboração da expressão de busca foi realizado de forma a identificar artigos que pudessem responder à seguinte questão de pesquisa: Como os sistemas *crowdsourcing* estão sendo utilizados no desenvolvimento de *smart cities*?

Apesar de não ser o objetivo principal, espera-se que a expressão de busca também traga resultados que respondam perguntas mais

específicas como: Quais são as vantagens (ou desvantagens) para se utilizar *crowdsourcing* em *smart cities*? Qual perspectiva (técnica, organizacional, processo, humana) apresenta mais dificuldades? E qual apresenta menos?

Portanto, foram testadas algumas combinações de expressão de busca observando a existência de conteúdo relevante nos resultados da primeira página dos buscadores. Por fim, eliminaram-se termos que se comportavam como sinônimos e que não influenciavam nos resultados de busca o que resultou na expressão de busca a seguir:

(volunteer OR crowd* OR citizen OR collaborative OR cooperative OR "co-create" OR participat*) AND ("smart cities" OR "smart city" OR "intelligent city" OR "intelligent cities") AND (system OR framework OR application)

Os critérios de inclusão e exclusão foram definidos com a expectativa de ter ao fim do estudo uma visão geral de que aplicações já estão em produção atualmente. Dessa forma, foram incluídos trabalhos que atendiam os seguintes critérios de inclusão:

- Implementa alguma aplicação ou sistema crowdsourcing com objetivo de promover as smart cities.
- II. Realiza algum experimento, ou caso de estudo, com objetivo de promover as *smart cities*.

Os seguintes critérios de exclusão foram utilizados:

- Trabalhos que n\u00e3o est\u00e3o relacionados com crowdsourcing e smart cities simultaneamente.
- II. A crowd não faz parte da solução, isto é, trabalhos em que apenas a tecnologia faz parte da solução.

A justificativa para utilização do critério de exclusão II é o fato de que as soluções produzidas pelo processo de *crowdsourcing* são, em geral, tão boas quanto os resultados produzidos através de outros processos [23]. Além disso, apesar da tecnologia ser uma alavanca fundamental, os maiores desafios estão, em última análise, nas áreas de governança e gestão [7]. Por isso, trabalhos que apresentem apenas soluções tecnológicas, ou de infraestrutura de uma *smart city*, não serão considerados.

2.2 Execução

Nesta etapa, foi realizada a validação da expressão de busca. Em seguida, aplicou-se a expressão nas bibliotecas digitais. O total de resultados obtidos, referentes à pesquisa que foi realizada durante o primeiro semestre de 2014, foi de 310 artigos, distribuídos 93 no IEEE Xplorer e 217 na ACM.

Além disso, usando os critérios de inclusão e exclusão, foram analisados os títulos e resumos dos 310 artigos anteriores. Os trabalhos que atenderam os nossos critérios foram armazenados para a próxima etapa. O total de artigos inclusos foi de 129, com 45 no IEEE Xplorer e 84 na ACM.

Na última etapa os 129 artigos filtrados foram lidos e aqueles que responderam à questão da pesquisa foram usados para alimentar o questionário formulado de extração de dados. Os artigos que não responderam à questão da pesquisa foram desconsiderados. Embora, alguns deles, apresentassem ideias interessantes que ajudassem a compreender os domínios de *crowdsourcing* e *smart cities*, de forma isolada, não eram o foco principal deste trabalho.

2.3 Extração de dados

A fase de extração de dados foi realizada por meio de um formulário eletrônico¹ composto por 23 questões relativas aos artigos. O objetivo desta fase é classificar e categorizar os artigos de forma a permitir uma análise quantitativa. As perguntas do formulário podem ser agrupadas da seguinte forma:

- Perguntas de controle: titulo, autor, DOI, data de publicação, perspectiva do artigo, etc.
- Perguntas sobre a crowd: possui incentivo? Qual retorno?
 Qual tipo, tarefa e conhecimento s\u00e3o utilizados?
- Perguntas sobre o Sistema: qual arquitetura, tipo de chamado, tipo de iniciativa e tipo de colaboração são utilizados?
- Perguntas sobre o problema: descrição do problema, domínio do problema etc.

As repostas para estas perguntas foram então compiladas e analisadas de forma quantitativa e subjetiva.

2.4 Análise dos resultados

A seguir faz-se uma análise quantitativa dos 14 artigos finalistas encontrados neste estudo. Inicialmente, observou-se que poucos destes trabalhos relataram sistemas já implementados. De fato, apenas 4 dos sistemas estavam em produção e os restantes, apresentavam casos de estudo ou ainda estavam na fase de testes.

A maioria dos artigos possuía uma arquitetura *crowdsourcing* independente/*standalone*. Além disso, a maior parte dos trabalhos expôs *crowdsourcing* com uma perspectiva humana, levantando questões relacionadas à inteligência coletiva fornecida pela *crowd*. Outros quatro trabalhos deram maior ênfase a sistemas *crowdsourcing* como mediadores de tarefas e outros três a aspectos tecnológicos ou de implementação.

Em relação ao domínio das aplicações, na sua maioria as aplicações lidavam com questões de segurança, lazer e transporte das *smart cities*. Consequentemente o conhecimento desejado da *crowd* foi predominantemente o situacional, com 11 entre os 14 artigos. Em apenas um trabalho houve necessidade da participação de pessoas com algum conhecimento especializado. Outro aspecto relacionado ao domínio de problemas, e ao conhecimento requerido, foi o tipo de tarefa realizada pela *crowd*, onde em 13 artigos predominou-se a captura/fornecimento de dados, havendo apenas três trabalhos que consistiam em analisar dados.

Em relação à participação da *crowd*, para 11 artigos ela ocorreu de maneira explícita, isto é, os participantes estavam cientes de estar inseridos numa iniciativa colaborativa de resolução de problemas. Os chamados foram do tipo aberto (habilitado para qualquer pessoa) em seis iniciativas e limitado em apenas um. Em 12 trabalhos a *crowd* é composta de agentes externos e em apenas um de agentes internos. Além disso, os recrutamentos foram feitos em apenas quatro dos trabalhos analisados.

A forma de coleta de dados foi feita de forma manual em 7 casos e automática em 5. Por outro lado, o processamento dos resultados gerados pela *crowd* foi feito em sua totalidade de forma automática. Em relação a incentivos, o maior critério adotado para promover a participação das pessoas foi o social. Em segundo lugar, aspectos lúdicos também foram abordados para estimular a continuidade e maior envolvimento dos participantes. Em 4 trabalhos não houve qualquer consideração com incentivos. Finalmente, 12 artigos não abordaram questões ligadas à

satisfação de *stakeholders* (*crowd* e dono de problema), isto é, não incluíram mecanismos que fornecessem algum retorno dos participantes.

Em seguida analisamos subjetivamente cada um dos trabalhos, destacando também as principais contribuições e desafios encontrados. Dividimos os trabalhos nos temas tecnologia, validação dos dados e tipo de incentivo. Frisamos, porém, que muitos deles possuem contribuições em outros temas e essa divisão visa apenas facilitar a compreensão do presente texto.

2.4.1 Tecnologia

O trabalho de Botero e Saad-Sulonen [2] discute o desafio de criar parcerias entre a cidade e seus cidadãos no intuito de criar novos e melhores serviços públicos. O sistema Urban Mediator é apresentado com objetivo de ser um ambiente mediador entre os serviços oferecidos pela cidade e os sistemas mantidos pelos cidadãos como: sites de comunidades, fóruns e blogs. A integração dos serviços com o Urban Mediator pode ser feita através de diversos padrões de sindicação (RSS, GEORSS etc..). Essa forma de integração facilitou a comunicação entre os envolvidos e não produziu acoplamento de tecnologias, pois cada um (cidade, cidadãos e Urban Mediator) utilizou sua própria tecnologia. Além disso, alguns cidadãos apreciaram o fato do sistema disponibilizar publicamente a informação que eles enviaram, ao contrário do que é geralmente feito com as pesquisas, questionários, ou outras ferramentas de relatórios criados até momento pela cidade.

Goodspeed et al. [14] seguem outra direção e comentam o desafio da colaboração coletiva entre as próprias agências de serviços públicos. Eles apresentam o caso da cidade de Boston durante o desenvolvimento de um sistema de crowdsourcing para o departamento de planejamento urbano da cidade. Um dos principais obstáculos encontrados foi a hesitação de agências governamentais em compartilhar informação. Organizações governamentais evitam compartilhar informação, pois isso resulta em perda de autonomia e aumento de interdependência [14]. Para contornar esses obstáculos é preciso fazer negociações com benefícios mútuos, cooperação e coordenação. Além disso, pode ser interessante utilizar uma arquitetura desacoplada em que as agências não fiquem dependentes entre si ao compartilhar dados, e que através do uso de padrões de sindicância e/ou dados ligados, as agências possam cooperar e obter benefícios indiretos, como aconteceu com o sistema Urban Mediator, mas sem perder autonomia.

Um dos poucos trabalhos encontrados que usam uma arquitetura piggyback é apresentado por Silva et al. [28]. No trabalho, as informações capturadas através de outras aplicações existentes (e.g. Foursquare) são reaproveitadas e exploradas, através de análise estatística, para obter um estudo sobre as dinâmicas de uma cidade. Algumas conclusões que essa abordagem fornece são: classificação de regiões de uma cidade pela sua semântica (e.g. regiões de lazer, residenciais, industriais) e; mapas de circulação que mostram os padrões de movimentação de pessoas na cidade. Além disso, argumenta-se que a captura de informações de cidadãos por meio de crowdsourcing é feita em tempo real e de forma menos custosa que as abordagens tradicionais de censo.

Asimakopoulou e Bessis [1] apresentam a possibilidade de usar crowdsourcing em smart cities para o gerenciamento de catástrofes. O trabalho defende que (i) se basear apenas em infraestruturas de redes de sensores não é suficiente e que (ii) o uso de referências cruzadas entre dados fornecidos por pessoas e

¹ Formulário para extração de dados: https://goo.gl/cTdqP4

redes de sensores convencionais podem contribuir com informações mais relevantes.

Vakali et al. [29] apresentam um uso prático para *crowdsourcing* em *smart cities*. No trabalho, o *crowdsourcing* é explorado para a validação de informações fornecidas por sensores. Para os autores, as redes de sensores convencionais são apropriadas para captura de informações quantitativas ou mensuráveis, e que a *crowd* é mais adequada para capturar informações qualitativas. Assim, as medições feitas são sumarizadas usando técnicas de processamento de linguagem natural e em seguida publicadas em redes sociais. As interações entre publicações de sensores e usuários são monitoradas. Como resultado as medições realizadas podem ser validadas ou informações complementares incluídas.

Por fim, Zeiger e Huber [31] apresentam o *Unified Sensing* Platform (USP) que permite o sensoriamento participativo do meio ambiente. Neste sistema, os participantes (cidadãos e servidores) são convidados a monitorar a qualidade do ar e condições do clima através de um sensor portátil acoplado a um aplicativo mobile. Através desse aplicativo os dados são captados do sensor e enviados para o USP. Através do USP é possível ter uma visão aproximada das condições ambientais de *smart cities*, detectar locais poluídos e prever a dispersão da fumaça nos casos de incêndios de edifícios.

Observando os trabalhos citados, destacamos as seguintes questões tecnológicas:

- Integração de dados: uma aplicação crowdsourcing poderia facilitar sua integração com outras aplicações e obter vantagens dessa colaboracão sem perder sua autonomia.
- 2. Reutilização de Tecnologias: poucos trabalhos utilizaram uma arquitetura do tipo piggyback, sendo que a maioria criou sua própria aplicação standalone. Isto talvez explique porque a maioria dos trabalhos não estavam em produção já que o esforço para colocar em produção pode ser maior numa aplicação standalone.
- Crowdsensing: redes de sensores podem ser utilizadas em combinação com redes de pessoas para obter benefícios complementares num sistema crowdsourcing.

Assim sendo, consideramos que algumas tarefas (coleta, processamento e visualização) de uma aplicação *crowdsourcing* precisam ser generalizadas e desacopladas entre si, tendo como objetivo facilitar a integração de dados com outras aplicações e a reutilização tecnológica. Por exemplo, permitir a utilização de dados tanto por redes de sensores como por redes de pessoas.

2.4.2 Validação de dados

Roitman et al. [26] apresentam um sistema de crowdsensing desenvolvido pela IBM para smart cities. Dois desafios foram levantados: o primeiro é que alguns eventos em uma cidade podem ser observados por poucas testemunhas oculares, ao contrário do que ocorre em meios tradicionais onde a detecção de um evento pode ser reportada por centenas e até milhares de sensores sociais; o segundo é que alguns sensores sociais podem reportar eventos de pessoas que não estão fisicamente presentes na cidade, por exemplo, um tweet de um evento ocorrido em New York feito por um usuário localizado em Los-Angeles. Devido a isso, percebe-se que a validação contextual dos dados fornecidos pelos sensores sociais é de extrema importância. O sistema apresentado resolveu esses problemas marcando automaticamente os dados coletados, através do uso de diversas heurísticas e modelos de processamento, com geo-tags que permitiam validar a proximidade física do sensor com o evento relatado.

A validação dos dados coletados é um processo importante de uma aplicação *crowdsourcing*, mas nem sempre é algo necessário ou que deve ser rigidamente implementado. Por exemplo, Doran et al. [9] mostra que cidadãos podem servir como sensores humanos e fornecer informações suplementares e alternativas para as *smart cities* mesmo na ocorrência de relatos inválidos. Por exemplo, uma percepção equivocada que determinada área da cidade é insegura (em contraste com as taxas oficiais e apenas devido a fatores como falta de patrulhamento, ambiente pouco iluminado e edifícios abandonados) deve ser encarada pelos gestores de *smart cities* como relatos válidos de problemas que foram incorretamente direcionados. Nesse caso, existe um problema real de percepção de segurança dos moradores que pode impactar também na percepção dos turistas que visitem o local.

Por outro lado, Jollivet [19] argumenta o valor de *crowdsourcing* como forma de resolução de problemas sócio-técnicos complexos, os quais incluem pessoas como partes inerentes do sistema. Nele propõe-se o uso de *crowdsourcing* como método de atribuição de níveis de confiabilidade, sendo necessário que os participantes de um sistema forneçam a forma como realizam as suas tarefas. Um experimento foi conduzido no qual se detectou um possível risco de propagação de doenças devido ao uso de temperaturas inapropriadas para a limpeza de materiais em um hospital.

De forma similar, Manzoor et al. [20] apresentam um framework para aplicações em smart cities que promove a participação de cidadãos na captura e divulgação de conteúdo. Além de fornecer um *middleware* que gerencia a comunicação entre *crowd* e demais componentes (ex. redes de sensores e sinks) a principal contribuição do trabalho está no uso de técnicas de ludificação para incentivar a participação das pessoas. O incentivo é dado por meio de pontuações atribuídas às anotações sobre problemas, ameacas, e outras ocorrências. Assim, são pontuados usuários que avaliam anotações dos demais participantes (inclusive sobre sensores) e, reciprocamente, usuários cujas anotações são bem avaliadas, ou seja, consideradas como informações confiáveis. Nesse sentido, a crowd participa no processo de duas maneiras: novas informações validando fornecendo/capturando e informações fornecidas por outras partes.

Percebe-se, através dos trabalhos citados, a importância da validação de dados numa aplicação *crowdsourcing*. Além disso, consideramos que essa seja uma questão de tratamento específico por aplicação, pois cada aplicação possui um nível de exigência próprio para validar os dados. Assim, encaramos como vantajosa a separação do processo de validação dos demais processos da aplicação ou até mesmo que esse seja um processo externo à aplicação. Essa separação pode permitir, por exemplo, o uso de várias abordagens para validar os dados.

A validação dos dados de forma externa à aplicação permite o reuso de tecnologias, heurísticas e modelos de processamento, tais como os utilizados por Roitman et al. [26], ou utilizar técnicas de ludificação, como fez Manzoor et al [20], para desenvolver um game com foco em validar dados, mas com a vantagem de poder utilizar um grupo de pessoas diferente do grupo que coletou os dados e com graus de confiança também diferentes.

2.4.3 Incentivo

As cidades de Corfu (Grécia) e Oulu (Finlândia) implantaram o sistema CLIO de *crowdsourcing* para formação e interação com a memória coletiva da cidade através da troca de textos, vídeos e fotos sobre lugares da cidade [25]. Esse sistema fornece várias interfaces de acesso desde navegadores *Web* e *smartphones* até navegadores de realidade aumentada. A interação do sistema se mostrou eficiente, pois ao testar o sistema com estudantes de

algumas escolas, e seus membros familiares, percebeu-se que o sistema reforçou os laços familiares e aproximou gerações. Segue abaixo duas respostas obtidas dos estudantes que confirmam isso:

"me and my grandfather used to play in the same square, but the surrounding shops are not the same"

"until now I felt that I had nothing in common to talk with my grandparents"

Na cidade de Oulu implantou-se também o sistema *RunWithUs*, de Gil-Castiñeira et al. [13], que tem como objetivo ajudar os usuários na prática de esportes e registrar suas conquistas de modo a fazer com que a atividade seja mais social. Por meio do *RunWithUs* os usuários podem conhecer novos parceiros de corrida, montar grupos de amigos, agendar atividades e saber se o nível de condicionamento físico de um possível parceiro é similar ao seu. Ao fazer uma pesquisa com os cidadãos, sobre o *RunWithUs*, concluiu-se que eles estão mais dispostos ao registro de suas performances do que à competição ou à interação social. Os entrevistados relataram, por exemplo, que não gostariam de correr com um desconhecido.

Percebe-se, portanto, que enquanto no sistema CLIO a interação social foi um incentivo importante, e recompensador, o mesmo não ocorreu com o sistema *RunWithUs* que foi implementado na mesma cidade. Por si só, o incentivo social não produz bons resultados para todos os tipos de *crowd*. Portanto, aplicações *crowdsourcing* devem procurar usar outros tipos de incentivos além do social e tentar encontrar mecanismos para promover a interação social entre usuários relutantes.

A ludificação de uma aplicação *crowdsourcing* é uma alternativa ao incentivo social, principalmente para aplicações com objetivos que geram pouco prestígio social aos usuários. Por exemplo, Quercia [24] relata a elaboração de dois projetos *crowdsourcing*, *Urbanopticon* e *UrbanGems*, ambos usando conceitos de ludificação para promover a participação da *crowd*. Em *Urbanopticon* mostra-se uma foto de um lugar aleatório em uma cidade, em seguida o usuário deve tentar localizá-lo em um mapa. Já em *UrbanGems* mostram-se duas fotos de dois lugares aleatórios em uma cidade dentre as quais o usuário deve escolher aquela que achar mais agradável. Fazendo isso é possível capturar uma imagem mental que os cidadãos possuem sobre as diferentes regiões de uma cidade, informação que pode ser de utilidade no planejamento urbano.

Por outro lado, consideramos que os incentivos podem ser complementares e que uma abordagem combinada pode gerar melhores resultados. Navarro et al. [22] apresenta uma aplicação destinada a incentivar o uso de bicicletas nas cidades. Assim, para atingir o seu objetivo, e utilizando-se de crowdsourcing, a aplicação oferece os seguintes recursos: captura e visualização de informações sobre rotas e localização/disponibilidade de pontos de interesse (e.g. localização de bicicletários, oficinas, chuveiros e outros); acesso a redes sociais e possibilidade de interação com outros ciclistas (e.g. identificação de amigos percorrendo numa localização próxima), e; técnicas de ludificação que tornem o ciclismo mais atraente (e.g. competições de quilômetros percorridos, incluindo pontuação, atribuição de níveis e inclusive estabelecimento de missões). Ou seja, para incentivar o uso da aplicação, foi realizado um conjunto combinado de incentivos: rede social, ludificação e a disponibilização de informações de interesse do usuário.

3. Estratégias Crowdsourcing

Em crowdsourcing, 4 estratégias primárias podem ser indicadas: crowd voting, crowd wisdom, crowd creation, e crowdfunding

[17]. Além disso, uma aplicação *crowdsourcing* pode utilizar uma ou mais dessas estratégias para atingir seu objetivo.

Durante a pesquisa, percebemos as seguintes demandas gerais para o desenvolvimento de aplicações *crowdsourcing*: integração de dados; reutilização tecnológica; e desacoplamento das camadas de coleta, processamento e visualização. Consideramos que para resolver essas demandas deve-se não somente usar as estratégias *crowdsourcing* na aplicação final, mas também usá-las no contexto de desenvolvimento da aplicação.

Sendo assim, o *crowd creation*, mais comumente associado com criação colaborativa de conteúdo multimídia, poderia ser usado para criar múltiplos protótipos de software e hardware. Para selecionar o melhor protótipo para determinado contexto podemos usar *crowd voting*. Além disso, os protótipos que não foram selecionados poderiam ser selecionados no futuro num outro contexto favorável, pois eles agora fazem parte da *crowd wisdom*. E a versão final do software, ou hardware, pode ser produzida com recursos obtidos por meio de *crowdfunding*.

Por outro lado, quando se desenvolve uma aplicação que não possui controle total de uma "crowd" de pessoas" (mas sim, de uma parte da "crowd" global") e que ela mesma faz parte de uma "crowd" de aplicações", pode-se considerar que a integração de dados é um requisito de grande prioridade e relevância, já que a aplicação tem que se comunicar com as outras aplicações para solucionar um problema maior. Além disso, a análise do crowd wisdom indicou a recorrência de soluções para problemas comuns, assim como a grande possibilidade de reuso de tecnologias.

3.1 Aplicações Genéricas

Uma estratégia importante para o desenvolvimento de aplicações é a de dividir para conquistar. Quando dividimos um problema em subproblemas menores e as soluções encontradas para esses últimos são bem definidas e testadas, podemos generalizar a solução para outros problemas similares. Foi o que aconteceu, por exemplo, com os programas de banco de dados que evoluíram de simples programas de manipulação de arquivos para sistemas genéricos de banco de dados que possuem até mesmo uma linguagem própria (SQL).

Algumas aplicações generalizam a arquitetura *crowdsourcing* e assim conseguem fazer reuso da tecnologia. Como exemplo, temos a plataforma *Wikimapps* [3] que é uma plataforma para criação, manutenção e hospedagem de *Wikimapps* e é uma generalização de uma aplicação de contexto específico, *Wikicrimes* [12]. *Wikimapps* é uma solução do tipo tudo-em-um, possui seus próprios mecanismos de exibição, coleta e processamento de informação.

Google forms é outro exemplo de aplicação genérica. Essa aplicação generalizou a criação de questionários, sendo usada nos mais diversos contextos que envolvam questionário, como no caso da educação colaborativa [6]. Esse conceito de questionário dinâmico também foi explorado pela aplicação MoCoMapps [18], que fornece um serviço mobile para usuários ordinários usarem e criarem aplicações baseadas em mapas colaborativos. O objetivo é atender tanto os usuários não programadores como os usuários finais e criadores de aplicações.

Tanto o *Google forms* quanto o *MoCoMapps* têm foco no processo de coleta explícita de informação do usuário. O projeto *ManyEyes*, por outro lado, foca na visualização de informações previamente coletadas pelo usuário. Esse portal fornece vários tipos de visualização de dados customizáveis, onde os usuários podem fazer o envio dos dados e criar suas próprias visualizações interativas. O objetivo do *ManyEyes* é apoiar a colaboração em

torno de visualizações em grande escala, promovendo assim um estilo social da análise de dados [30].

Percebe-se assim que é possível generalizar as camadas de coleta e visualização de dados de aplicações colaborativas. Embora algumas aplicações sejam definidas pela camada de processamento de dados, essa camada, dependendo do contexto da aplicação, também pode ser generalizada. Além disso, consideramos que pode ser interessante, em alguns casos, mesclar o processamento local, com um processamento genérico e externo. Como exemplo, uma aplicação de processamento OCR poderia fazer uso de algoritmos próprios para reconhecer caracteres complexos e utilizar algoritmos genéricos e de terceiros para os caracteres comuns restantes.

Observando os resultados positivos produzidos pelo processo de generalização das aplicações colaborativas citadas, consideramos que a generalização, e consequente desacoplamento, das camadas de coleta, processamento e visualização de aplicações colaborativas é uma boa estratégia para responder a algumas das demandas da nossa pesquisa e também para facilitar o desenvolvimento de aplicações *crowdsourcing* para *smart cities*. Para validar essa estratégia criamos uma plataforma para o desenvolvimento de protótipos das aplicações *crowdsourcing* de forma desacoplada. A plataforma criada é apresentada em detalhes na próxima seção.

3.2 A Plataforma

Nossa plataforma, denominada *Searchlight*, tem como objetivo facilitar o desenvolvimento e prototipação de aplicações *crowdsourcing*. Sendo assim, ela fornece uma base de como essas aplicações podem ser desenvolvidas de forma rápida, descentralizada e manutenível.

O primeiro princípio adotado na Searchlight é a divisão de responsabilidades da aplicação crowdsourcing em três camadas básicas e independentes: coleta, processamento e visualização. Para cada camada existe um elemento responsável por resolver, de forma isolada, as suas categorias de problemas. Por exemplo, para as atividades de coleta de informações, a plataforma propõe a criação de um aplicativo coletor (Searchlight Collector - SLC) que irá requisitar os dados que se deseja coletar. Para o processamento das informações, bem como seu armazenamento formal, propõe-se a criação de um webservice (Searchlight Service - SLS) com banco de dados que possua modelo de dados flexível (NoSQL, MongoDB) que permita sua reutilização para diversas estruturas em um mesmo banco. A visualização das informações coletadas por SLC e processadas por SLS é realizada por uma web widget que possui duas visualizações predefinidas: Map View e List View. Além disso, essa web widget pode ser estendida com visualizações criadas por terceiros.

Uma visão geral dos três elementos principais da plataforma é apresentada Figura 1, onde também é possível observar a comunicação entre eles e com a Internet:



Figura 1. Comunicação entre os componentes da plataforma

O segundo princípio adotado pela plataforma é utilização de um modelo de dados flexível. Isto significa que a plataforma não utiliza uma estrutura predefinida para armazenar as propriedades das anotações no banco de dados. Ela deixa a cargo do SLC ou do SLV escolher a melhor estrutura para o seu contexto. O objetivo por traz desse princípio é facilitar a reutilização e publicação dos dados coletados. Essa reutilização pode ser feita de duas formas: indireta, em um ambiente de dados ligados, e direta, por outras aplicações do mesmo domínio (que por serem semelhantes tendem a possuir os mesmos requisitos de estruturas de dados). Consideramos que através desses dois princípios é possível propor uma plataforma genérica para produção de aplicações *crowdsourcing*.

Os componentes SLV e SLC possuem objetivos respectivamente semelhantes aos aplicativos ManyEyes e MoCoMapps. O principal diferencial é o fato de serem abertos ao público para extensão, modificação e utilização, considerando que o projeto ManyEyes teve seu acesso gratuito fechado recentemente pela IBM². Além disso, esses componentes estão sendo projetados para trabalhar juntos, mas de maneira desacoplada sendo opcional a utilização de um ou de outro. Atualmente existe uma diferença significativa, em escala, entre SLV e ManyEyes, pois este último possui uma variedade de visualizações enquanto o SLV possui apenas duas visualizações criadas inicialmente para testar a plataforma. Por outro lado, entre o SLC e MoCoMapps a diferença é sutil, pois enquanto no MoCoMapps a aplicação de coleta de dados é projetada no próprio smartphone, no SLC o modelo da interface é construído numa aplicação Web externa e o aplicativo coletor posteriormente faz o download e instalação do modelo para o smartphone.

3.2.1 Protótipos e Aplicações Finais

Oito protótipos (que podem ser acessados no site do projeto) foram criados para avaliar a validade da plataforma em relação a estratégia de desacoplamento das camadas de visualização, processamento e coleta de dados. Os mais relevantes, porém, são os que geraram as aplicações finais Universidade Inteligente e Placa Fácil.

A aplicação Universidade Inteligente tem como objetivo coletar reclamações, sugestões e indicações de pontos de interesse na Universidade. Basicamente, é uma aplicação de coleta de dados que permite o envio de comentários, fotos e vídeos sobre algum tema por alunos. Sendo assim, um protótipo foi criado e utilizado por um grupo de 7 estudantes que percorreram a universidade relatando os problemas e pontos de interesse da universidade o que permitiu gerar um mapa³ sumarizando essas informações. Perguntados sobre o protótipo os estudantes relataram problemas de usabilidade com a interface, mas aprovaram a ideia de criar um aplicativo onde eles poderiam relatar os problemas da Universidade.

A Figura 2 exibe as relações entre os componentes da plataforma com o protótipo Universidade Inteligente. O protótipo foi criado com os componentes genéricos da plataforma, sendo necessário apenas definir o modelo de interface do aplicativo coletor, ou seja, definir um questionário que aceite campos de texto, fotos e vídeos e posição geográfica. Percebe-se, nos itens b e c da figura, que o protótipo criado é basicamente uma instância do aplicativo genérico de coleta SLC. Os outros componentes se comunicam com ele, mas não possuem conhecimento algum sobre a lógica de

² ManyEyes foi fechado pela IBM http://goo.gl/xg4SR3

³ Mapa gerado pelo protótipo http://goo.gl/GvskES

negócio do aplicativo coletor e, através de um protocolo genérico, o SLS consegue armazenar os dados que são genericamente exibidos pelo SLV em um mapa online.

Visualização	stv	SLV
Processamento	SLS	SLS
Coleta	SLC (App Genérico)	App Universidade Inteligente

Figura 2. Relacionamento entre as camadas de uma plataforma *crowdsourcing* com os componentes da nossa plataforma e *App* Universidade inteligente

O aplicativo Placa Fácil tem como objetivo capturar fotos de placas de obras públicas e através de OCR identificar os dados da obra para um posterior cruzamento entre dados oficiais com dados e fotos coletadas do local. Aqui, novamente o foco é a etapa de coleta de informações, mas ao contrário do Universidade Inteligente, a coleta não é feita por um aplicativo genérico. Detalhes da interface com o usuário, como o processo de marcação, e recorte, dos campos textuais da imagem são requisitos não atendidos pelo aplicativo genérico SLC e, por isso, precisou-se de um aplicativo dedicado. O processamento do OCR, bem como o armazenamento das informações, ainda é de responsabilidade do SLS, sendo a exibição e o cruzamento dos dados coletados realizados através do SLV.

Um esquema do processo de funcionamento do aplicativo pode ser visto na Figura 3. O processo seguido é o seguinte: o usuário fotografa uma imagem; marca os trechos onde serão realizados o processamento OCR, ou seleciona uma máscara de modelos conhecidos; alguns filtros e tratamentos são realizados na imagem; imagem e metadados são enviados para o SLS como uma anotação genérica; o *middleware* OCR presente no SLS realiza o processamento OCR e armazena o resultado; por fim a aplicação recebe o resultado do processamento OCR. Assim, com exceção da instância do SLS, o OCR Service, todos os outros itens são tarefas realizadas apenas pelo aplicativo mobile.

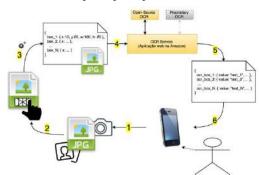


Figura 3. Processo de funcionamento do app Placa Fácil.

Percebe-se que apesar de não termos generalizado completamente a aplicação Placa Fácil ainda foi possível ter reuso tecnológico nas camadas de processamento e visualização. A Figura 4 mostra, no item a, que o aplicativo Placa fácil é uma instância da camada de coleta e que ele se comunica com as instâncias padrão do SLS, acrescida de um middleware OCR, e do SLV pertencentes as outras camadas. Consideramos que essa generalização das camadas de uma aplicação *crowdsourcing* pode ser uma estratégia útil para agilizar o desenvolvimento de aplicações *crowdsourcing*.

Além disso, nessa aplicação, o responsável pelo desenvolvimento do aplicativo é totalmente independente do responsável pelo serviço OCR, ou seja, o aplicativo apenas envia uma imagem com

metadados e espera um resultado do serviço. Isto fornece mais liberdade ao responsável pelo desenvolvimento do middleware OCR do SLS, de forma que o processamento OCR pode ser feito usando algoritmos abertos, e gratuitos, para a maioria dos textos e um serviço pago quando os algoritmos gratuitos falharem.



Figura 4. a) divisão de camadas do Placa Fácil. b) Aplicação hipotética.

Em uma situação hipotética, poderíamos querer substituir a forma algorítmica do processamento OCR por um processamento no *Amazon Mechanical Turk* para ser feito por pessoas em *crowdsourcing*. Desconsiderando-se os requisitos iniciais da interface e usando os componentes da nossa plataforma, ou os de outras plataformas privadas, o Placa Fácil poderia ser considerado como uma aplicação totalmente genérica. Por exemplo, poderíamos ter o MoCoMapps na camada de coleta, o *Amazon Mechanical Turk* na camada de processamento e o IBM ManyEyes na camada de visualização, conforme mostra o item b da Figura 4.

4. Conclusões

Neste estudo avaliamos 129 artigos em relação ao tema pesquisado dos quais selecionamos para análise 14 trabalhos que demostraram como sistemas *crowdsourcing* estão sendo utilizados no desenvolvimento de *smart cities*. Ao analisar as demandas e as soluções apresentadas por esses trabalhos identificamos três grupos importantes: tecnologia, validação de dados e incentivos.

Em relação à questão tecnológica, percebemos que poucos trabalhos utilizaram uma arquitetura *piggyback*. De fato, a maioria deles possui uma arquitetura *standalone* com soluções próprias e autossuficientes, mas similares em funcionalidade. Dessa forma, percebe-se que há espaço para criação de componentes genéricos da arquitetura *crowdsourcing* que permitam o reuso tecnológico. Além disso, uma aplicação *crowdsourcing* pode facilitar sua integração com outras aplicações e obter vantagens dessa colaboração sem perder sua autonomia ao compartilhar dados.

Afirmamos ainda que a validação de dados em *crowdsourcing* deve ser tratada com soluções especializadas por aplicação, pois cada aplicação possui seu próprio nível de exigência em relação a validade dos dados. No entanto, isso não impede o uso de soluções com validações externas ao aplicativo que também podem trazer benefícios similares.

Percebemos, também, que muitos trabalhos são baseados em incentivos sociais e/ou lúdicos que, embora sejam interessantes, não têm se mostrado suficientes em alguns casos. Consideramos que uma abordagem combinada de incentivos possa trazer melhores resultados e que mais estudos na área sejam necessários.

Identificamos que utilizar as estratégias de *crowdsourcing* não somente no aplicativo, mas também no processo de desenvolvimento do aplicativo *crowdsourcing* pode facilitar esse processo. Além disso, se considerarmos o aplicativo como membro de uma "*crowd* de aplicações", a integração de dados obtém maior prioridade e relevância.

Observamos que a generalização (e consequente desacoplamento) das camadas de coleta, processamento e visualização de

aplicações colaborativas seja uma boa estratégia para solucionar as demandas encontradas na nossa pesquisa e assim facilitar o desenvolvimento de aplicações *crowdsourcing*. Essa estratégia, foi utilizada com êxito em 8 protótipos de nossa plataforma, e permitiu a criação das aplicações finais Universidade Inteligente e Placa fácil. Desta forma, consideramos que ela também pode ser útil em outras plataformas e situações.

5. REFERÊNCIAS

- [1] Asimakopoulou, E.; Bessis, N. Buildings and Crowds: Forming Smart Cities for More Effective Disaster Management. *Proc. Fifth Int. Conf. on Innovative Mobile and Internet Services in Ubiquitous Computing*, IEEE (2011), 229–234.
- [2] Botero, A.; Saad-Sulonen, J. Co-designing for New Citycitizen Interaction Possibilities: Weaving Prototypes and Interventions in the Design and Development of Urban Mediator. Indiana University (2008), 266–269.
- [3] Caminha, C., Furtado, V., Vasconcelos, E.; Ayres, L. Uma ferramenta de autoria para criação de mapas colaborativos para aplicações em egov 2.0. *Proc. XXX CSBC*, (2010), 20– 23.
- [4] Catania, B., Guerrini, G., Belussi, A., Mandreoli, F., Martoglia, R.; Penzo, W. Wearable Queries: Adapting Common Retrieval Needs to Data and Users. *Proc. 7th Int.* Work. on Ranking in Databases, ACM (2013), 7:1–7:3.
- [5] Chourabi, H. et al. Understanding Smart Cities: An Integrative Framework. *IEEE Computer Society* (2012), 2289–2297.
- [6] Denton, D.W. Enhancing instruction through constructivism, cooperative learning, and cloud computing. *TechTrends* 56, 4 (2012), 34–41.
- [7] DiGiammarino, F.; Trudeau, L. Virtual networks. *THE PUBLIC* 5, (2008).
- [8] Doan, A., Ramakrishnan, R.; Halevy, A.Y. Crowdsourcing systems on the World-Wide Web. *Communications of the* ACM 54, 4 (2011), 86.
- [9] Doran, D., Gokhale, S.; Dagnino, A. Human Sensing for Smart Cities. Proc. 2013 Int. Conf. on Advances in Social Networks Analysis and Mining, ACM (2013), 1323–1330.
- [10] Erickson, L., Petrick, I., Trauth, E.; Erickson, L.B. Hanging with the right crowd: Matching crowdsourcing need to crowd characteristics. *Proc. AMCIS* 2012, (2012), 1–9.
- [11] Estelles-Arolas, E.; Gonzalez-Ladron-de-Guevara, F. Towards an integrated crowdsourcing definition. *Journal of Information* ... 38, 2 (2012), 189–200.
- [12] Furtado, V., Caminha, C., Ayres, L.; Santos, H. Open government and citizen participation in law enforcement via crowd mapping. *IEEE Intelligent Systems* 27, 4 (2012), 63– 69.
- [13] Gil-Castiñeira, F. et al. RunWithUs: A Social Sports Application in the Ubiquitous Oulu Environment. *Proc. 10th Int. Conf. on Mobile and Ubiquitous Multimedia*, ACM (2011), 195–204.
- [14] Goodspeed, R., Spanring, C.; Reardon, T. Crowdsourcing As Data Sharing: A Regional Web-based Real Estate Development Database. Proc. 6th Int. Conf. on Theory and Practice of Electronic Governance, ACM (2012), 460–463.

- [15] Harris, C.G. You're Hired! An Examination of Crowdsourcing Incentive Models in Human Resource Tasks. Proc. Workshop on Crowdsourcing for Search and Data Mining, CSDM (2011), 15–18.
- [16] Hetmank, L. Components and Functions of Crowdsourcing Systems – A Systematic Literature Review. *Proc. 11th Int. Conf. on Wirtschaftsinformatik*, March (2013), 55–69.
- [17] Howe, J. Crowdsourcing: How the power of the crowd is driving the future of business. Random House, 2008.
- [18] Hupfer, S. et al. MoCoMapps: mobile collaborative mapbased applications. *Proc. Conf. Computer Supported Cooperative Work Companion*, (2012), 43–44.
- [19] Jollivet, P. Crowd sourced security, trust & cooperation for learning digital megacities: valuing social intangible assets for competitive advantage and harmonious development. *Proc. Int. Conf. on Smart and Sustainable City*, IET (2011), 52–52.
- [20] Manzoor, A. et al. Data Sensing and Dissemination Framework for Smart Cities. Proc. Int. Conf. on MOBILe Wireless MiddleWARE, Operating Systems, and Applications, EAI (2013), 156–165.
- [21] Nam, T.; Pardo, T.A. Conceptualizing Smart City with Dimensions of Technology, People, and Institutions. ACM (2011), 282–291.
- [22] Navarro, K.F. et al. SocialCycle what can a mobile app do to encourage cycling? Proc. 38th Annual IEEE Conf. on Local Computer Networks - Workshops, IEEE (2013), 24–30.
- [23] Pedersen, J. et al. Conceptual foundations of crowdsourcing: a review of IS research. Proc. 46th Hawaii Int. Conf. on System Sciences, (2013), 579–588.
- [24] Quercia, D. Urban: Crowdsourcing for the Good of London. Proc. Int. World Wide Web Conferences Steering Committee (2013), 591–592.
- [25] Ringas, D.; Christopoulou, E. Collective City Memory: Field Experience on the Effect of Urban Computing on Community. ACM (2013), 157–165.
- [26] Roitman, H. et al. Harnessing the Crowds for Smart City Sensing. Proc 1st Int. Workshop on Multimodal Crowd Sensing, ACM (2012), 17–18.
- [27] Shapiro, J.M. Smart cities: quality of life, productivity, and the growth effects of human capital. *The review of economics and statistics* 88, 2 (2006), 324–335.
- [28] Silva, T.H. et al. Challenges and opportunities on the large scale study of city dynamics using participatory sensing. *Proc. IEEE Symp on Computers and Communications* (ISCC), (2013), 528–534.
- [29] Vakali, A., Angelis, L.; Giatsoglou, M. Sensors talk and humans sense Towards a reciprocal collective awareness smart city framework. *Proc. 2013 IEEE Int. Conf. on Communications Workshops (ICC)*, IEEE (2013), 189–193.
- [30] Viegas, F.B. et al. Manyeyes: a site for visualization at internet scale. *Visualization and Computer Graphics, IEEE Transactions on 13*, 6 (2007), 1121–1128.
- [31] Zeiger, F.; Huber, M. Demonstration Abstract: Participatory Sensing Enabled Environmental Monitoring in Smart Cities. Proc. 13th Int. Symp. on Information Processing in Sensor Networks, IEEE Press (2014), 337–338.