



CEPSys – *Crowdsourcing* colaborativo no apoio a situações de emergência

Paulo NUNES¹, Bruno INÁCIO¹, Carlos SANTOS*¹ e Ana BAPTISTA¹

¹ Faculdade de Ciências, Universidade de Lisboa (Portugal)

(paulo.jorge.anunes@gmail.com; inacio.bruno@gmail.com; csalexandre@gmail.com; wipabaptista@gmail.com)

Palavras-chave: *Crowdsourcing*, Redes Sociais, Sistemas Colaborativos, Sistemas de Apoio à Decisão, Gestão de Emergências

Resumo:

As situações de emergência ou calamidade são complexas, normalmente caracterizadas pela ocorrência múltipla de eventos distintos dispersos no espaço. O sucesso na gestão e distribuição dos meios num ambiente de catástrofe natural ou de origem antropogénica depende do fluxo de informação.

As redes sociais e os dispositivos móveis cativam atualmente milhões de pessoas em todo o mundo, são a ligação do cidadão à sociedade da informação em que vivemos. Esta utilização massiva de equipamentos móveis de comunicação, dotados com modernos sensores deu origem a um novo fenómeno de geração e partilha de informação útil para os mais diversos fins: *marketing*, comércio eletrónico, análise de tendências, gestão colaborativa, etc.. Vivemos numa sociedade tecnológica, marcada pela troca de informação em tempo real e pela mobilidade. Esta realidade traz novas oportunidades na recolha de informação colaborativa que pode ser explorada em prol do benefício coletivo. Por outro lado, verifica-se uma elevada adesão das pessoas às redes sociais tornando-as num veículo de excelência para fazer chegar avisos às pessoas e mantê-las informadas.

Os parágrafos anteriores foram o mote para o desenvolvimento do Sistema: *Collaborative Emergency Processing System* (CEPSys). Este sistema explora uma fonte de informação colaborativa baseada na Internet, designado: *social media crowdsourcing*, como complemento para os meios convencionais de recolha de informação usados na gestão das situações de calamidade ou emergência.

O CEPSys foi desenvolvido por alunos de mestrado no âmbito da disciplina da Aplicações na Web da Faculdade de Ciências da Universidade de Lisboa. O protótipo criado perscruta a informação livre disponibilizada pelos serviços web da rede social Twitter (Twitter, 2015) e identifica mensagens prioritárias de emergência geradas de forma colaborativa pelos utilizadores e fá-las chegar aos organismos coordenadores de emergência apresentando a localização do utilizador, o relato e informação auxiliar como imagens e vídeo.

Os potenciais utilizadores do sistema são as entidades com papel ativo na estrutura nacional de Proteção Civil. Contudo, pode ser utilizado noutros sectores com interesse na informação colaborativa (ex.: relatos de crimes às autoridades policiais, relatos de situações de fragilidade social aos serviços sociais e de apoio à vítima, relato de necessidade de intervenção de infraestruturas aos serviços das câmaras e restantes agentes públicos, etc.).

A comunicação está estruturada numa breve descrição do protótipo do sistema, a sua arquitetura e componentes desenvolvidos para cumprir os objetivos propostos e por fim uma conclusão onde são apresentados os desenvolvimentos futuros que podem ser feitos para melhorar o protótipo e potenciais resultados que se podem atingir com o processo de desenvolvimento.



1. Introdução

As situações de emergência ou calamidade de origem natural ou antropogénica são situações complexas com elevado impacto na vida quotidiana das sociedades. São caracterizadas pela ocorrência múltipla de eventos com características muito distintas numa grande extensão espacial. O sucesso na gestão e distribuição dos meios, muitas vezes escassos num ambiente de catástrofe depende do fluxo de informação que chega e parte dos organismos coordenadores. Esta necessidade despertou o interesse no desenvolvimento de um sistema colaborativo que utilizasse a informação proveniente do *crowdsourcing* para gerar um mecanismo complementar de recolha de informação útil para identificar situações de calamidade. No âmbito da cadeira de Aplicações de Web da Faculdade de Ciências da Universidade de Lisboa foi desenvolvido um protótipo do CEPSys (Colaborative Emergency Processing System). O sistema faz a recolha automática de informação através de serviços web das redes sociais. De seguida, são desencadeados processos de tratamento da informação e armazenamento numa base de dados relacional e é feita a sua apresentação num interface web com suporte geoespacial, onde os eventos são visualizados quase em tempo real.

O protótipo do CEPSys, criado para efeitos de demonstração, identifica através de *hashtags* as mensagens prioritárias de emergência geradas pela população na rede social Twitter, trata a informação e centraliza-a num portal acessível pelos organismos coordenadores da Proteção Civil e respetivos colaboradores. Os potenciais utilizadores do sistema serão as entidades com papel ativo na estrutura nacional de Proteção Civil.

Este projeto trata só a questão de recolha de informação colaborativa para efeitos de gestão de emergência, no entanto, o *crowdsourcing* pode ser utilizado para recolher informação de apoio a diversos campos: segurança (relato de furtos e incidentes, monitorização de criminalidade, etc.), administração local e pública (sistemas colaborativos de decisão para gestão dos municípios e cidades), estudos estatísticos socio económicos, monitorização de eventos públicos, etc..

2. Arquitetura e descrição do Sistema e Componentes

2.1 Arquitetura

O CEPSys apresenta a arquitetura esquematizada na figura 1. O sistema baseia-se numa componente de *Back-end*, constituído por uma base de dados relacional que armazena a informação recolhida pelos programas. O *Crawler* e o *Semantic Analyser* são dois programas desenvolvidos em Java (Gosling & McGilton, 1996) e são ativados de forma cíclica para conduzir duas tarefas principais: recolha e tratamento de informação.

O primeiro recolhe na base de dados as configurações necessárias para aceder aos serviços web das redes sociais e recolhe a informação que corresponde aos critérios de busca configurados pelo utilizador. O segundo faz o tratamento da informação recolhida e organiza-a na base de dados. O sistema apresenta dois portais web de acesso. Um portal de administração que permite interagir com a base de dados, adicionar configuração de critérios de busca, configuração de credenciais de acesso aos serviços web das redes sociais e visualizar resultados. O portal do cliente apresenta a informação organizada de uma forma orientada ao cliente num suporte geoespacial que permite identificar a localização dos eventos e relatórios gráficos com a distribuição temporal das ocorrências. O portal de cliente e outros clientes externos são alimentados por um serviço web numa arquitetura baseada em REST (*Representational State Transfer*) proposto por Fielding (2000), garantindo que o sistema é facilmente escalável e é interoperável com outros clientes e aplicações.

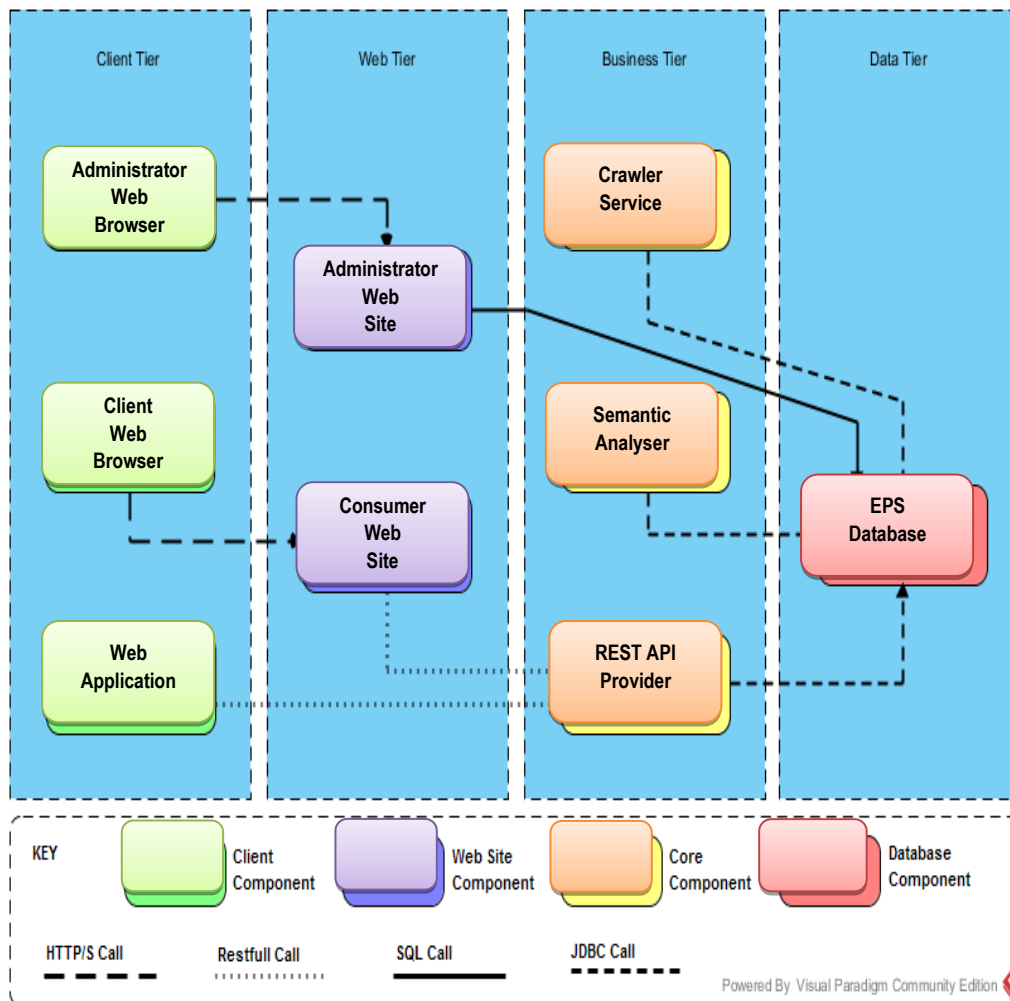


Figura 1 – Vista C&C multi-tier do CEPSSys

2.1.1 – Componentes de Back-End

No *back-end* do sistema são desenvolvidas as tarefas de recolha e processamento de informação. Como componentes principais podemos identificar dois programas desenvolvidos em Java que realizam a recolha de dados, o *Crawler* e o processamento através do *Semantic Analyser*.

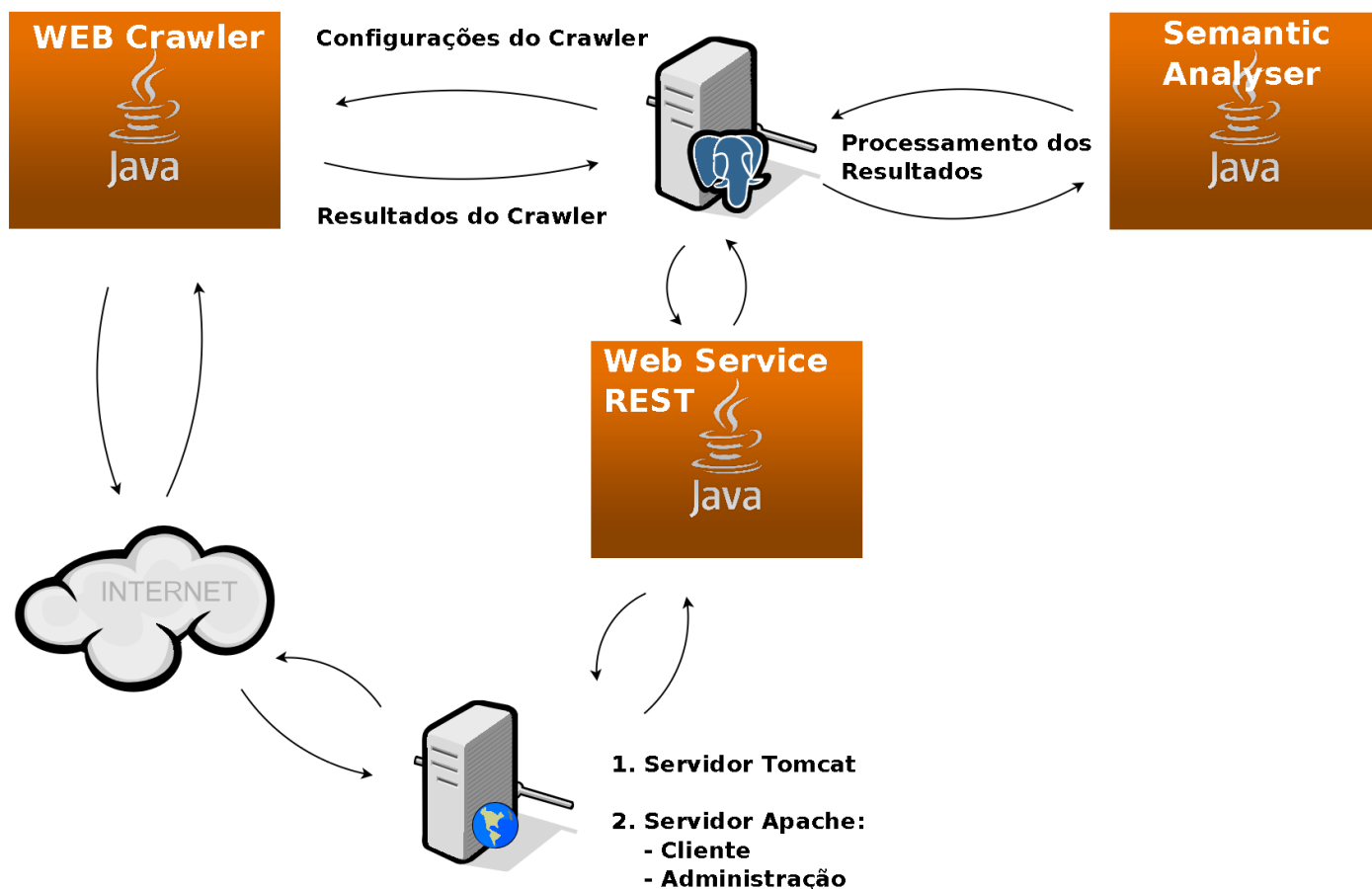


Figura 2 – Esquema das interações entre os programas de recolha de informação (*Crawler*), armazenamento e processamento de informação (*Semantic Analyser*)

Na figura 2 é apresentado o esquema simples do workflow com as tarefas de recolha e processamento da informação. O *Crawler* vai ler na base de dados as configurações dos critérios de busca e credenciais de acesso aos serviços web das redes sociais para construir um conjunto de acessos com um conjunto de parâmetros de pesquisa. Para explicar este conceito em termos simples podemos dizer que o *Crawler* vai criar um conjunto de tarefas. Associada a cada tarefa temos a identificação e credenciais de acesso a uma rede social com uma pesquisa única. Os resultados das várias tarefas são armazenados na tabela de resultados. O *Semantic Analyser* vai atuar sobre a tabela de resultados identificando a origem dos dados e o formato da informação para organizá-la e determinar os campos da base de dados que são preenchidos com os eventos.

Alguma informação proveniente das redes sociais encontra-se identificada com as coordenadas do dispositivo móvel que lhe deu origem. Todavia, durante o desenvolvimento do sistema verificou-se que a informação na sua maioria não apresenta referência espacial fundamental para o sistema. Esta dificuldade foi ultrapassada dotando o *Semantic Analyser* com a capacidade de identificação de moradas devidamente identificadas no corpo da mensagem. Caso o colaborador não tenha um dispositivo com capacidade de adicionar diretamente a localização aos *Tweets*, o colaborador pode adicionar a morada entre os identificadores @loc, da seguinte forma : @loc morada @loc. O sistema filtra a morada e utiliza os serviços de geocoding da API do GoogleMaps para obter as coordenadas geográficas associadas à morada.



2.1.2 – Modelos de dados

O núcleo do sistema é uma base de dados relacional onde é armazenada a informação. A base de dados alimenta o serviço *web REST* que fornece informação aos clientes. Para base de dados foi escolhida uma opção *open source* : *PostgreSQL* (The PostgreSQL Global Development Group, 1996). Esta base de dados apresenta capacidade de armazenamento de informação espacial em estruturas vectoriais e *raster* através da extensão *PostGIS* (PostGIS Project Steering Committee, 2014), propriedade de muita utilidade para adicionar informação geográfica de base (redes viárias, locais de abrigo, localização de meios de salvamento, etc.). Contudo, neste trabalho só foi armazenada informação alfanumérica.

O modelo de dados foi projetado para ser adaptável e tanto quanto possível flexível. Estas características foram pensadas para que os utilizadores possam configurar a alterar os critérios de pesquisa e acesso às redes sociais de forma dinâmica, garantindo que o processo de reconfiguração seja um processo de fácil execução.

O modelo de dados prevê tabelas para o armazenamento dos critérios da pesquisa, necessárias à construção das procuras nos serviços *web* das redes sociais. Armazena as credenciais de acesso às redes sociais. Os resultados recolhidos pelo programa são armazenados numa tabela em formato *JSON* (ECMA International, 2013). O *Semantic Analyser* age sobre a tabela de resultados com os registos recolhidos e reorganiza-os na tabela de eventos e tabelas associadas.

2.1.3 – Portal de Administração do Sistema

O portal de administração do *CEPSys* foi desenvolvido através da *framework CakePHP* (Cake Software Foundation, 2015) e corre a partir de um servidor *Web Apache*. O portal foi desenvolvido para interagir com a base de dados, apresenta formulários de consulta e alteração da informação de configuração dos critérios de pesquisa e credenciais de acesso aos serviços *web* externos. Permite visualizar os resultados das pesquisas do *Crawler* e dos resultados do processamento do *Semantic Analyser*. O portal de administração possibilita ainda a ativação manual dos programas *Crawler* e *Semantic Analyser*.

2.1.4 – Portal do Cliente

O portal de cliente foi projetado com o objetivo de facilitar a atividade num centro de coordenação de emergências. O interface encontra-se dividido em dois painéis simples e foi desenhado por forma a facilitar o acesso a informação das ocorrências detetadas pelos colaboradores da aplicação. Apresenta uma componente espacial com a informação georreferenciada. Um painel onde as mensagens dos colaboradores são apresentadas e onde a autoridade coordenadora pode publicar avisos à população. O interface apresenta também uma componente estatística que admite a visualização e consulta do histórico de informação. O portal foi desenvolvido utilizando a *framework CakePHP* e as linguagens de desenvolvimento *web* : *HTML5* (W3C, 2014), *CSS* (W3C, 2015), *PHP* (The PHP Group, 2015), *D3* (Mike Bostock, 2015.), *Javascript* (ECMA International, 2011) e *AJAX* (The jQuery Foundation, 2015).

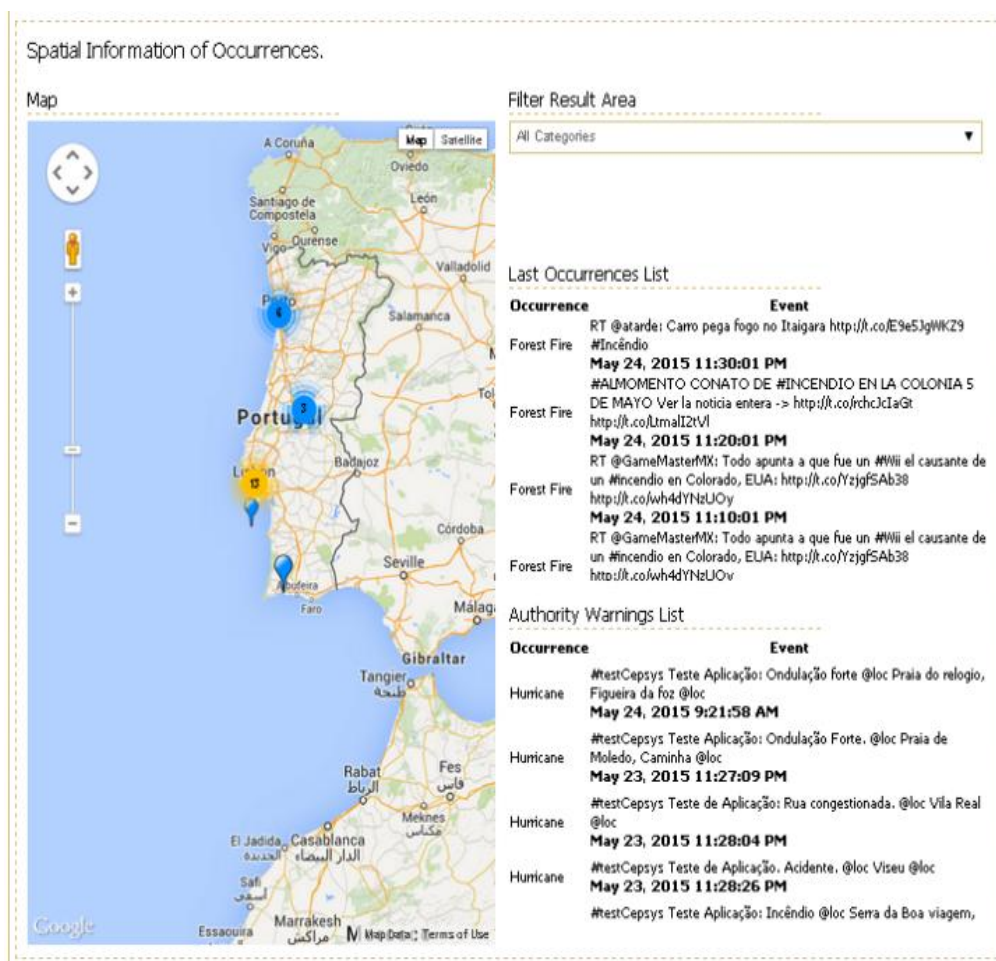


Figura 3 – Painel de informação geoespacial do portal de cliente do Cepsys

O painel de informação geoespacial (figura 3) foi projetado para a visualização das ocorrências e a sua localização. Para o desenvolvimento desse painel utiliza-se a API (*Application Programming Interface*) da Google Maps (Google, 2015) para inserir a base cartográfica sobre a qual são inseridos os pontos correspondentes aos eventos gerados pelos colaboradores. O painel atualiza-se automaticamente adicionando os novos eventos recolhidos pelo sistema nas redes sociais. Ao lado da informação geoespacial é introduzida a informação textual em dois painéis verticais. O painel superior é reservado à introdução da informação de coordenação gerada pela Autoridade Nacional de Proteção Civil. O painel inferior é reservado à publicação de informação gerada pela população.

O painel de informação estatística (figura 4) foi projetado para a visualização das estatísticas relativas aos eventos armazenados na base de dados. Neste painel pretende-se visualizar informação estatística sobre as séries temporais relativas à geração dos eventos. Este bloco foi projetado para dar informação sobre os meses, e horas do dia em que são gerados mais eventos e a relevância e duração dos eventos. A informação textual associada à estatística pode ser consultada de forma dinâmica filtrando os períodos ao nível dos meses e das horas do dia.

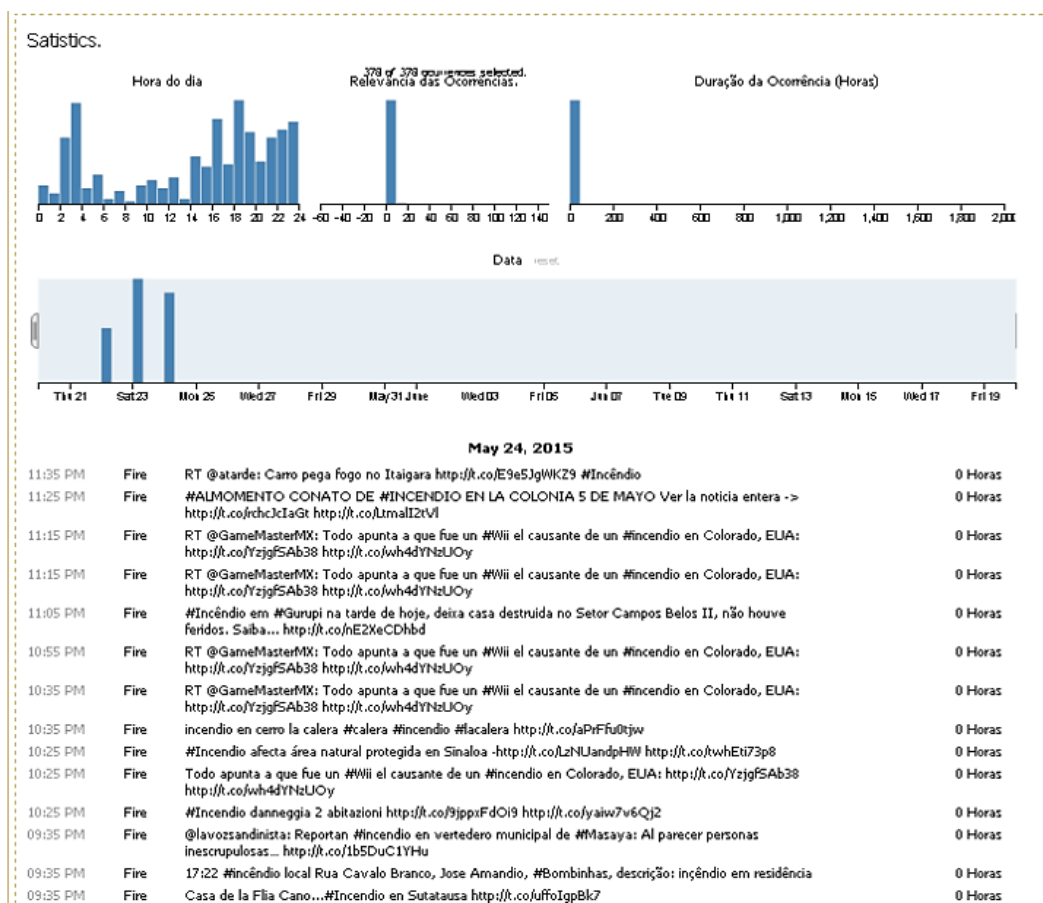


Figura 4 – Painel de informação estatística do portal de cliente do Cepsys

2.1.5 – Serviço Web REST

O acesso de terceiros à informação recolhida ocorre através do serviço web RESTfull desenvolvido em Java utilizando a *framework Jersey*. Este tipo de implementação de serviços web permite criar uma aplicação que possibilita o acesso à informação através dos métodos HTTP : PUT, GET, POST e DELETE. Neste tipo de arquitetura o cliente acede à informação através do URI. A forma de construção do URI por parte do cliente determina qual a ação a ser executada : acesso, adição, alteração ou apagar a informação. As respostas aos pedidos dos clientes são geradas em formato JSON.

A aplicação que fornece os serviços REST corre num servidor *Tomcat* que recebe pedidos reencaminhados a partir do servidor Apache onde correm os portais de administração e de cliente.

3. Conclusões

O protótipo do CEPSSys demonstrou que é possível utilizar o *crowdsourcing*, as redes sociais e dispositivos móveis como mecanismos complementares aos sistemas convencionais de aviso prévio e coordenação de ações coletivas, potenciando o fluxo de informação de e para a população. O sistema apresenta uma fase de maturação que impede a sua implementação operacional imediata. Todavia, lança a semente para o uso das redes sociais e da apetência humana por tecnologia para criar sistemas eficazes de troca de informação. O sistema trará valor acrescido através da recolha de informação em tempo real, complementada



com imagens e vídeo publicadas pelos colaboradores. Auxiliará a clarificar a situação nos centros de comando e controlo, apoiará a decisão dos coordenadores e melhorará a distribuição dos recursos.

O protótipo foi desenvolvido em ambiente universitário sem a consulta dos potenciais clientes. Contudo, considera-se que com as devidas adaptações (desenvolvimento das políticas de segurança da informação, desenvolvimento de algoritmos de interpretação e categorização da informação através de aprendizagem automática servidores de processamento de mensagens sms e email, etc.) poderá ser uma ferramenta importante de apoio à decisão e gestão de incidentes. Como resultado final é esperado que o sistema possa contribuir para melhorar a segurança coletiva, devolvendo à sociedade a sua colaboração.

O princípio subjacente ao sistema com base na utilização colaborativa de informação pode ser facilmente escalado para outras aplicações que visam a recolha colaborativa de informação. Como exemplos de futuras aplicações podem-se indicar: sinalização e monitorização de furtos e criminalidade, relato e monitorização de deficiências e anomalias nas infraestruturas (relato de deficiências nas infraestruturas aos municípios, etc.), sistemas colaborativos de apoio às decisões nos investimentos dos municípios e a monitorização dos resultados, gestão das redes de transporte e distribuição de serviços (relato de anomalias, etc.), relatos de situações de vulnerabilidade social e monitorização do turismo.

Referências Bibliográficas

- Cake Software Foundation, 2015, *CakePHP*, <http://cakephp.org/>, consultada em junho de 2015.
- ECMA International, 2011. *ECMAScript Language Specification*, <http://www.ecma-international.org/ecma-262/5.1/Ecma-262.pdf>, consultada em junho de 2015.
- ECMA International, 2013. *The JSON Data Interchange Format*, <http://www.ecma-international.org/publications/files/ECMA-ST/ECMA-404.pdf>, consultada em junho de 2015.
- Fielding, Roy (2000). *Representational State Transfer (REST)*, http://www.ics.uci.edu/~fielding/pubs/dissertation/rest_archj_style.htm, consultada em junho de 2015.
- Google, 2015. *Google Maps Javascript API V3 Reference*, <https://developers.google.com/maps/documentation/javascript/>, consultada em junho de 2015.
- Gosling, J. & McGilton, H., 1996, *The Java Language Environment*, <http://www.oracle.com/technetwork/java/langenv-140151>, consultada em junho de 2015.
- Mike Bostock, 2015. *D3 Data Driven Documents*, <http://d3js.org/>, consultada em junho de 2015.
- Oracle Corporation, 2015, *Jersey RESTful Web Services in Java*, <https://jersey.java.net/>, consultada em junho de 2015.
- PostGIS Project Steering Committee, 2015, <http://postgis.net/>, consultada em junho de 2015.
- The Apache Software Foundation, 2015, <http://www.apache.org/>, consultada em junho de 2015.
- The jQuery Foundation, 2015, <http://api.jquery.com/jquery.ajax/>, consultada em junho de 2015.
- The PHP Group, 2015, <https://secure.php.net/>, consultada em junho de 2015.
- The PostgreSQL Global Development Group, 1996, <http://www.postgresql.org/>, consultada em junho de 2015.
- Twitter, 2015, <https://dev.twitter.com/>, consultada em junho de 2015.
- W3C, 2014, *W3C Recommendation 28 October 2014*, <http://www.w3.org/TR/2014/REC-html5-20141028/>, consultada em junho de 2015.
- W3C, 2015, <http://www.w3.org/Style/CSS/>, consultada em junho de 2015.