# CONCEPÇÃO E DESENVOLVIMENTO DE MECANISMOS PARA OBTENÇÃO DE PONTOS DE INTERESSE EM CIDADES BRASILEIRAS

William Rodrigues da Silva (PIBIC/CNPq/FA/UEM), Marcos Aurélio Domingues (Orientador), e-mail: ra99514@uem.br

Universidade Estadual de Maringá / Departamento de Informática/Maringá, PR.

**Área e subárea do conhecimento: Ciência da Computação / Sistemas de Informação.**

**Palavras-chave:** Coleta de informações da web, Pontos de interesse, *Crawler* Java

**Resumo:** O setor de turismo no Brasil vem crescendo nos últimos anos. Apesar do crescimento do setor, este ainda apresenta vários problemas, entre os quais, vale ressaltar a falta de informações para turistas. Uma possível solução para a falta de informações turísticas no que se refere a pontos de interesse (serviços e locais turísticos) consiste na proposta de uma plataforma que faça uso de sistemas de recomendação de pontos de interesse. Entretanto, para o uso da plataforma, se faz necessário à obtenção de pontos de interesse para então recomendá-los. Este projeto teve como objetivo a investigação e proposta de mecanismos para obtenção de pontos de interesse em cidades brasileiras para serem utilizados pela plataforma de recomendação de pontos de interesse.

## Introdução

O setor de turismo no Brasil cresceu a uma taxa de 13,1% no ano de 2012 (SEBRAE-SP, 2016). Eventos como a Copa das Confederações de 2013, a Copa do Mundo de 2014 e as Olimpíadas de 2016 impulsionaram este setor, que mesmo em época de crise se mantem em crescimento. Apesar do crescimento do setor de turismo, este ainda apresenta vários problemas, entre os quais, vale ressaltar a falta de informações para turistas (UOL, 2016).

Uma possível solução para a falta de informações turísticas no que se refere a pontos de interesse (serviços e locais turísticos) consiste na concepção e desenvolvimento de uma plataforma para um sistema de recomendação sensível ao contexto capaz de recomendar pontos de interesse por meio de aparelhos de telefones celulares inteligentes (*smartphones*).

Um sistema de recomendação sensível ao contexto é uma tecnologia de filtragem de informação que, além do perfil e interesse do usuário (turista), utiliza também informação contextual para poder recomendar a este uma lista ordenada de itens, por exemplo, locais turísticos, que lhe sejam potencialmente interessantes (ADOMAVICIUS; TUZHILIN, 2011). Informação contextual pode ser definida como um conjunto de fatores externos que influenciam o comportamento e interesse de um usuário.

Uma das plataformas de recomendação mais conhecida atualmente é a *Foursquare* (FOURSQUARE, 2016). Nesta plataforma, os usuários da rede social *Foursquare* marcam os lugares já visitados por eles, e as recomendações são feitas a partir dos relacionamentos dentro da rede social. Contudo, neste paradigma, as recomendações se tornam imparciais quando os amigos/relacionamentos de um dado usuário nunca estiveram no local que ele está visitando no momento ou quando o usuário possui poucos amigos/relacionamentos.
Diante destes fatos, o orientador deste projeto de iniciação científica propôs um projeto de pesquisa intitulado "Uma Plataforma Avançada para Recomendação Automática de Pontos de Interesse em Cidades Brasileiras". Para que a plataforma de recomendação de pontos de interesse, proposta no projeto de pesquisa, não necessite de mecanismo de *check-in* e produza recomendações interessantes para o usuário, se faz necessário que ela utilize mecanismos que obtenham e armazenem automaticamente as informações sobre pontos de interesse de uma cidade. Assim, o objetivo principal deste projeto de iniciação científica foi à investigação e proposta de mecanismos para obtenção de pontos de interesse em cidade brasileira para serem utilizados como entrada em sistemas de recomendação de pontos de interesse, proporcionando o desenvolvimento de uma plataforma que possa produzir recomendações de maior interesse para o usuário e evitar a necessidade de mecanismos de *check-in*.

## Materiais e métodos

*Desenvolvimento do crawler Java*

O desenvolvimento de mecanismos para a obtenção de pontos de interesse iniciou-se com a criação de um *crawler* (ferramenta para obtenção de informações) utilizando a linguagem de programação Java, que faz o levantamento de informações sobre pontos de interesse em cidades brasileiras, cria objetos com estas informações e faz a persistência dos objetos criados no banco de dados da plataforma. As informações que um objeto registrado pelo *crawler* pode possuir são: Nome, endereço formatado, número de telefone formatado, latitude, longitude, URL, website, nota de avaliação, tipo, comentário sobre o local e autor do comentário. O *crawler* realiza também uma categorização dos pontos de interesse com base no tipo do local, desta forma criando uma generalização dos tipos.
O *crawler* desenvolvido pode ser usado em dois diferentes modos de operação, sendo eles: "Modo API" e "modo WEB", e cada modo extraí informações de fontes diferentes.
No "modo API", o *crawler* realiza a coleta de informações de pontos de interesse utilizando a API Google Places[1], partindo de uma coordenada inicial e buscando os pontos de interesse em um dado raio (em metros) com base no ponto inicial especificado. Neste modo, o *crawler* faz inicialmente uma requisição a API e recebe respostas no formato JSON contendo os pontos de interesse encontrados na região determinada, e a partir desta resposta, realiza para cada ponto uma requisição específica solicitando mais informações sobre o local em questão. O *crawler* faz

[1] https://developers.google.com/places/web-service/intro

então uma validação de coerência com base no nome e categoria do local, e se aprovada, cria e persiste um objeto no banco de dados com as informações do ponto de interesse.
No "modo WEB", o *crawler* tem também como objetivo a obtenção de informações de pontos de interesse, porem, as informações obtidas provêm de outra fonte. Neste modo, o *crawler* navega automaticamente pelo kekanto[2], um site web que lista pontos de interesse em cidades brasileiras, e a partir dos parâmetros recebidos, extrai da estrutura do site as informações necessárias para formar um objeto referente a um ponto de interesse. O *crawler* então realiza este processo para todos os pontos de interesse registrados em uma determinada cidade, criando e persistindo objetos no banco de dados para cada um destes.

*Desenvolvimento do sistema web*

A segunda etapa do desenvolvimento ocorreu com a criação de um sistema web composto de dois módulos. O primeiro módulo, nomeado como "OndeTem Web", tem como objetivo permitir que proprietários de pontos de interesse realizem um cadastro com informações referentes aos seus pontos de interesse, e estas informações ficam então registradas em um banco de dados auxiliar para que possam ser avaliadas posteriormente por um administrador do sistema. O segundo módulo, nomeado como "OndeTem Admin" tem como objetivo permitir que um administrador do sistema realize a verificação e posterior aprovação dos pontos de interesse cadastrados por seus proprietários. Desta forma, o sistema viabiliza um novo mecanismo para obtenção de pontos de interesse.

*Unificação dos sistemas desenvolvidos*

Por fim, para possibilitar um acesso mais direto e unificado dos mecanismos de obtenção de pontos de interesse, as funcionalidades dos modos do *crawler* foram adicionadas ao módulo "OndeTem Admin", permitindo que um administrador além de realizar as funções já citadas, possa também executar os modos "API" e "WEB" do *crawler* Java. A Figura 1 mostra uma das telas do sistema unificado.

**Figura 1** – Tela do sistema unificado.

[2] http://kekanto.com.br

## Resultados e Discussão

Para demonstrar os dados obtidos pelo sistema *crawler* Java, este foi executado no "modo WEB" para a cidade de Maringá-PR. A Tabela 1 apresenta uma pequena amostra das informações obtidas e registradas no banco de dados da plataforma.

**Tabela 1** – Amostra das informações obtidas com o *crawler* Java no "modo WEB".

| formated_address | latitude | longitude | name |
|---|---|---|---|
| Av. Tiradentes, 500, Zona 01 - Maringá | -23.426240000 | -51.938172000 | Catedral de Maringá |
| Av. Tamandaré, 200, Zona 01 - Maringá | -23.418971000 | -51.934994000 | Café Cremoso |
| Av. São Paulo, 1099, Zona 09 - Maringá | -23.418286000 | -51.932920600 | Maringá Park Shopping |
| Av. Herval, 26, Zona 50 - Zona Central - Maringá | -23.425390000 | -51.936368000 | Hotel Deville Maringá |

## Conclusões

Com a realização deste projeto de iniciação científica foram desenvolvidos mecanismos para obtenção de pontos de interesse em cidades brasileiras que podem extrair informações de três diferentes fontes, colaborando assim para que a plataforma de recomendação de pontos de interesse tenha acesso a estas informações e possa realizar de modo adequado as suas recomendações.

## Agradecimentos

Agradeço ao programa PIBIC/CNPq/FA/UEM pelo apoio financeiro e ao professor Marcos Aurélio Domingues pela orientação.

## Referências

SEBRAE-SP. **Turismo cresce 13,1% em 2012.** Disponível em: <**http://www.sebraesp.com.br/index.php/76-noticias/multissetorial/8867-pesquisa-setor-de-turismo-cresce-13-1-em-2012**>. Acesso em: 21 ago. 2016.

UOL Mais. **Turismo no Brasil tem falta de informação.** Disponível em: <**http://mais.uol.com.br/view/s70pk4i6az2h/turismo-no-brasil-tem-falta-de-informacao-04020D9A3166E0914326**>. Acesso em: 21 ago. 2016.

ADOMAVICIUS, G.; TUZHILIN, A. **Context-Aware Recommender Systems**. In: Recommender Systems Handbook, 2011, p. 217-253.

FOURSQUARE. Disponível em: <**https://foursquare.com**> Acesso em: 21 ago. 2016.