

## **P239, Sistema de monitoramento para prevenção e predição de danos**

Caio Palumbo Abreu<sup>1</sup>, Patricia Sartori Guastella<sup>1</sup>, Ivo Pons<sup>2</sup>, Marcio Cabral<sup>2</sup>, Roberto Marcondes<sup>2</sup>

1 Comgás

2 Data Machina

**Resumo** - Este projeto de P&D foi elaborado pela empresa Data Machina - Inteligência em Análise de Dados, e apresenta o projeto do desenvolvimento de uma metodologia e uma API (*Application Programming Interface*) que complemente o IPC (Índice de Priorização de Chamado) calculado atualmente no setor GRD (Gestão de Risco de Danos e Irregularidades) da Comgás a fim de servir como um marcador de “desempate” quando houver dúvidas em relação ao atendimento de regiões distintas levando em consideração os dados da cidade e buscando refletir, no índice relativo a cada zona, o impacto populacional, social e de infraestrutura de um possível incidente naquela zona.

Palavras-chave: *Cidades Inteligentes; Dados; Prevenção e Predição de dados; Big Data.*

### **Introdução**

Com o P239 (Sistema de Monitoramento para Prevenção e Predição de Danos), iniciou-se uma série de pesquisas a fim de propor soluções para alguns problemas latentes na companhia. Atualmente o setor de Gestão de Risco de Danos e Irregularidades (GRD) recebe um grande número de chamados de intervenção na infraestrutura urbana no qual cada um deles possui um risco associado; para entender esse risco são levados em consideração informações técnicas em relação ao sistema de gás nas proximidades da intervenção, assim como o tipo de intervenção e o maquinário envolvido. Para chamados que possuem um alto grau de risco é necessário o envio de técnicos que possam acompanhar essas intervenções a fim de minimizar o risco de incidentes; no entanto, como o número de técnicos é uma variável limitada, se torna necessário um novo mecanismo para entender os arredores da região e suas características para uma melhor compreensão do impacto causado em caso de incidente. Desta forma, a priorização de chamado estaria considerando tanto o risco associado àquela região quanto o impacto causado em caso de incidente, completando, assim, sua matriz de risco. No decorrer deste trabalho veremos uma metodologia proposta e a construção de uma

API para contribuir na robustez do cálculo do índice de priorização de chamado do GRD.

### **Desenvolvimento**

Neste trabalho foram estudados diferentes objetos e informações que representam a composição e dinâmica de uma região geográfica. Com isso foi desenvolvido um índice quantitativo que possibilita a comparação de diferentes regiões em relação às suas características. Para início dos estudos limitou-se a região geográfica para o município de São Paulo por conter um grande conjunto de dados públicos abertos e compreender cerca de 37,3%<sup>1</sup> da rede Comgás.

Neste contexto, o índice quantitativo foi denominado *IPC+* que representa uma extensão do índice IPC já calculado nos dias atuais pela setor GRD da Comgás. O *IPC+* contribuirá para o melhor entendimento da região do chamado possibilitando uma ação mais assertiva na alocação de profissionais e de medidas de segurança.

#### **A. Identificação dos dados**

Inicialmente, definiu-se quais informações relevantes ao problema seriam obtidas e quais dados seriam coletados para adquirir tais informações. Desta forma, dividiu-se as

---

<sup>1</sup> Informação obtida a partir do arquivo *shapefile* da rede completa da Comgás disponibilizada pela própria Companhia.

informações em quatro categorias: Populacional e 'Pontos Notáveis'; Sociais e Econômicos; Mobilidade; e Obstáculos Técnicos e para cada categoria definiu-se os dados a serem obtidos que os compunham. Para a categoria Populacional e 'Pontos Notáveis' temos: Número de habitantes, Número de estabelecimentos de educação, Número de estabelecimentos de Saúde. Para a categoria Sociais e Econômicos: Zoneamento, IPVS (Índice Paulista de Vulnerabilidade Social), RAIS (Relação Anual de Informações Sociais), Renda média per capita. Para Mobilidade: Pessoas em trânsito, Zona azul, Linhas de metrô, Terminais de ônibus, Estações de metrô, Pontos de ônibus. E a categoria Obstáculos Técnicos é composta por: Geotécnica, Rede de transmissão elétrica, Oleoduto, Rios, Canais, Córregos.

#### **B. Coleta de dados**

Os dados foram facilitados pelo IPEA (Instituto de Pesquisa Econômica Aplicada), GeoSampa (plataforma de visualização e armazenamento de dados da prefeitura de São Paulo), IBGE (Instituto Brasileiro de Geografia e Estatística), INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira), CNES (Cadastro Nacional de Estabelecimentos de Saúde), SMDU (Secretaria Municipal de Desenvolvimento Urbano), Metrô, CPTM (Companhia Paulista de Trens Metropolitanos), SPTrans (São Paulo Transporte), CET (Companhia de Engenharia de Tráfego), SMUL (Secretaria Municipal de Urbanismo e Licenciamento), entre outros órgãos e fontes internas.

#### **C. Banco de dados**

Os dados coletados foram tratados com auxílio da linguagem de programação *Python* visando corrigir potenciais erros de georreferenciamento e de limpeza das informações não utilizadas no momento. Após o tratamento, utilizou-se o banco de dados não relacional MongoDB em seu serviço em nuvem, Mongo Atlas. Neste momento os dados já estão processados e realocados em seus respectivos hexágonos tornando as solicitações da API e do Painel de Consulta mais ágeis.

#### **D. Processamento**

Para este trabalho a malha hexagonal se torna conveniente devido às suas propriedades de vizinhança, todos os hexágonos vizinhos

possuem mesma distância, e por sofrer menor distorção ao cobrir o globo terrestre em relação à malha triangular e quadrangular. Após a obtenção da malha o próximo passo foi distribuir os dados coletados em seus respectivos hexágonos de acordo com sua geolocalização. Desta forma, para cada hexágono teremos um vetor de características que representa as informações sobre os dados coletados contidos em cada hexágono.

#### **E. Cálculo do IPC+**

Optou-se por utilizar o método AHP (Analytic Hierarchy Process) no qual é realizada uma comparação pareada entre os dados determinando o grau de importância que um tem em relação ao outro obtendo como resultado um vetor de pesos para cada característica do hexágono.

Levando em consideração que cada característica do hexágono possui uma magnitude diferente, realizamos um processo de normalização. O método utilizado foi a padronização em que é realizada, para cada dado, a subtração da média e a divisão do desvio padrão. Este processo visa eliminar o viés que as diferentes magnitudes poderiam trazer e possibilitar o tratamento dos dados de maneira igualitária. Desta forma, temos para o *IPC+* a soma ponderada entre as características.

#### **F. API**

Para possibilitar o acesso ao *IPC+* foi desenvolvido uma API (Application Programming Interface) que realiza o intermédio entre nossas bases de dados e um solicitante de modo automatizado.

#### **G. Painel de consulta**

Desenvolveu-se uma plataforma que permite o acesso facilitado às informações das características de um dado hexágono de interesse e sua primeira vizinhança no qual possibilita a análise do impacto em diferentes graus, maior no hexágono central e secundário na vizinhança.

### **Resultados**

#### **A. API**

A API responde a consultas para solicitação de informações do *IPC+* para uma determinada posição, representada por uma

coordenada geográfica no padrão WGS 84. A API funciona no seguinte endereço:

<https://api.dataexmachina.com.br/v0/>

A documentação completa com os parâmetros retornados está disponível em:

<https://app.swaggerhub.com/apis/Mac77/IPC/0.0.6>

#### B. Painel de consulta

O Painel de Consulta, no momento deste documento, está em sua versão 1.0 e possui a aparência mostrada na Figura 1 a seguir. Este pode ser acessado através do link:

<https://comgas-grd-dev.web.app/>



**Figura 1** - Painel de consulta do IPC+.

O Painel de Consulta ainda possibilita calibrar os pesos das características, isto é, o processo AHP pode ser realizado diretamente na plataforma sendo disponibilizado apenas para usuários autorizados.

### Conclusões e contribuições

É importante salientar que o *IPC+* foi construído para ser utilizado como uma ferramenta comparativa entre hexágonos e perde seu sentido se olhado individualmente, isto é, se para um dado hexágono tivermos o *IPC+* de 10, não somos capazes de dizer o significado desse valor. Da mesma forma, valores baixos de *IPC+* não querem dizer que o hexágono não é importante ou não possui relevância, este valor apenas nos informa que dentre os dados analisados e os pesos definidos, o hexágono em questão possui uma baixa combinação ponderada desses dados.

O projeto desenvolvido alcançou seu objetivo no desenvolvimento de uma metodologia e de um índice que quantifique uma região geográfica em relação às suas informações e características, sendo capaz de complementar o IPC calculado pelo GRD possibilitando o desempate e o melhor

entendimento da região em que se pretende realizar uma intervenção na infraestrutura.

### Referências

Pereira, R. H. M., Braga, C. K. V., Serra, Bernardo, & Nadalin, V. (2019). **Desigualdades socioespaciais de acesso a oportunidades nas cidades brasileiras**, 2019. *Texto para Discussão Ipea*, 2535. Instituto de Pesquisa Econômica Aplicada (Ipea). Disponível em: <<http://repositorio.ipea.gov.br/handle/11058/9586>>. Acessado em: 04/2020.

GeoSampa. **Mapa Digital da Cidade de São Paulo**. Prefeitura de São Paulo. Disponível em: <<http://geosampa.prefeitura.sp.gov.br/>>. Acessado em: 04/2020.

Metrô. **Pesquisa Origem Destino 2017**, 2017. Disponível em: [http://www.metro.sp.gov.br/pesquisa-od/arquivos/Ebook%20Pesquisa%20OD%202017\\_final\\_240719\\_versao\\_4.pdf](http://www.metro.sp.gov.br/pesquisa-od/arquivos/Ebook%20Pesquisa%20OD%202017_final_240719_versao_4.pdf). Acessado em: 04/2020.

BRODSKY, Isaac. H3: **Uber's Hexagonal Hierarchical Spatial Index**. Uber Engineering, 27 de Junho de 2018. Disponível em: <<https://eng.uber.com/h3/>>. Acessado em: 04/2020.

Vaidya, Omkarprasad S., and Sushil Kumar. **"Analytic hierarchy process: An overview of applications."** *European Journal of operational research* 169.1: 1-29., 2006