

TerrainIntelGis: A Geographic Intelligence System to Support Decision-Making in Environmental and Natural Disaster Scenarios

Carlos H. T. Brumatti, Mariana Schaefer, Maria Lúcia B. Villela, Jugurta Lisboa-Filho, Julio C. S. Reis
Department of Informatics, Universidade Federal de Viçosa, Brazil
{carlos.h.tavares,mariana.schaefer,maria.villela,jugurta,jreis}@ufv.br

Abstract

Environmental and natural disasters can inflict catastrophic harm upon both urban and rural landscapes. Consequently, the imperative emerges for a streamlined administration of resources allocated to rehabilitate these afflicted areas. In this context, the application of data mining techniques emerges as a pivotal tool in fortifying this endeavor. In this paper, we present the TERRAININTELGIS, a Geographical Intelligence System (GIS) data-mining based, developed to serve as an additional tool for supporting decision-making by public institutions and private companies in allocating these resources used in environmental and natural disaster scenarios.

Keywords: Sistema de Inteligência Geográfica; Mineração de Dados; Desastres Naturais ou Ambientais

1 Introdução

Desastres ambientais ou naturais, apesar de terem origens diferentes, impactam a sociedade em diversas proporções [6]. Para fins de ilustração, a recuperação das áreas afetadas pelo rompimento da barragem de Fundão, em Mariana, considerado o maior desastre ambiental do Brasil, irá custar cerca de 150 milhões às prefeituras envolvidas, além da criação de um fundo de 20 bilhões de dólares pela empresa envolvida [2]. Tal desastre atingiu 39 municípios, em dois estados distintos, afetando cerca de 1,2 milhões de pessoas. Estimativas do Banco Mundial apontam que anualmente há um prejuízo de cerca de 520 bilhões de dólares associados aos desastres naturais¹. Além disso, existem atualmente cerca de 13,297 áreas de risco no país².

Diante deste cenário, uma das principais dificuldades enfrentadas por empresas e entidades governamentais está

¹<https://www.worldbank.org/en/news/press-release/2016/11/14/natural-disasters-force-26-million-people-into-poverty-and-cost-520bn-in-losses-every-year-new-world-bank-analysis-finds>

²<https://geoportal.cprm.gov.br/portal/apps/opsdashboard/index.html>
<http://www.cprm.gov.br/>

In: XXII Workshop de Ferramentas e Aplicações (WFA 2023), Ribeirão Preto, Brasil. Anais Estendidos do Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia). Porto Alegre: Sociedade Brasileira de Computação, 2023.

© 2023 SBC – Sociedade Brasileira de Computação.
ISSN 2596-1683

relacionada à alocação eficiente de recursos financeiros destinados à recuperação das áreas afetadas por esses desastres naturais e/ou ambientais, em que é necessário considerar as várias particularidades de cada município. Essa alocação eficiente é importante principalmente por identificar e considerar áreas mais vulneráveis que, por exemplo, devem ser priorizadas. Logo, é extremamente importante que sejam construídos mecanismos que possam ser úteis neste processo, e é neste contexto que se insere o objetivo deste trabalho.

Especificamente, apresentamos o TERRAININTELGIS, um Sistema de Inteligência Geográfica (SIG) proposto para suportar o processo de tomada de decisão em cenários de desastres ambientais. De forma geral, o sistema implementa uma metodologia descrita em um trabalho anterior [3] que explora técnicas de Mineração de Dados (MD) [4], para agrupamento de áreas e/ou cidades semelhantes, dado algum padrão oculto *a priori*. É importante mencionar que o TERRAININTELGIS é ajustável para diferentes tipos de dados, especialmente, os públicos. Esperamos que, ao utilizar o sistema, o usuário (especialista) responsável pela tomada de decisão dentro de uma organização possa identificar subsídios importantes que sejam úteis no seu processo de tomada de decisão relacionado, por exemplo, aos investimentos destinados à recuperação das áreas atingidas por um determinado desastre.

O artigo está organizado na seguinte format: na Seção 2 apresentamos brevemente esforços relacionados. Em seguida, na Seção 3, descrevemos a ferramenta, destacando a arquitetura e cada um dos módulos que a compõem. Detalhes relativos à disponibilização e apresentação do sistema, incluindo um vídeo de demonstração explorando dados reais são apresentados nas Seções 4 e 5, respectivamente. Depois, na Seção 6, relacionamos potenciais impactos sociais da ferramenta proposta. Por fim, na Seção 6, concluímos o trabalho, destacando direções de pesquisas futuras neste contexto.

2 Esforços Relacionados

Existem alguns trabalhos anteriores que exploraram a aplicação de técnicas de MD no contexto ambiental, com foco na tarefa de agrupamento de áreas específicas. Em [9], por exemplo, técnicas de agrupamento são exploradas com o objetivo de identificação de áreas mais propensas a desastres ambientais na Indonésia. Tal tarefa foi realizada com o uso de dados referentes aos tipos de desastres mais comuns na

região, tais como registros de enchentes, deslizamentos e terremotos. Da maneira similar, o estudo apresentado em [8] agrupa áreas deste mesmo país (i.e., Indonésia), com o objetivo de classificá-las em graus de riscos previamente definidos: muito alto, alto, moderado, baixo e muito baixo. Por fim, em [1], os autores calculam o *Roots Mean Square Deviation* (RMSD) para determinar o valor ideal do número de grupos e assim utilizar o *KMeans* nos dados de desastres naturais coletados, com a finalidade de prever a ocorrência de novos desastres.

Em suma, nosso esforço é complementar a esses trabalhos no sentido de propor a aplicação de técnicas de MD no contexto ambiental. No entanto, diferente dos esforços anteriores, que em sua maioria estão focados em entender causas de desastres naturais com o intuito, por exemplo, de prever eventos futuros, nossa proposta se baseia na disponibilização de arcabouço computacional funcional para apoio ao processo de tomada de decisão de instituições públicas e empresas privadas na alocação desses recursos utilizados em cenários de desastres ambientais e naturais. Além de vantagens relacionadas ao uso direto de técnicas de MD, propomos a automação de várias etapas intermediárias (e.g., definição do número de agrupamentos), o que pode beneficiar e potencializar a utilização do TERRAININTELGIS por usuários (especialistas) de diferentes áreas do conhecimento.

3 TerrainIntelGis

Conforme supracitado, por meio deste sistema, é possível que o usuário especialista, responsável pelo processo de tomada de decisão em uma organização (pública ou privada), possa interagir e utilizar, mesmo sem conhecimentos prévios, técnicas de MD, em especial as relacionadas ao Agrupamento de Dados. Ademais, a proposta aqui descrita busca ser genérica em relação à limitação geográfica do problema. Em outras palavras, a ideia é que possam ser utilizados quaisquer dados de interesse do usuário, desde que sejam relacionados à localização geográfica do desastre. Para isso, ele pode realizar o *upload* de um arquivo no formato *csv*, contendo as informações de interesse. Uma visão geral do SIG proposto é apresentada na Figura 1, que de forma geral pode ser dividido em dois módulos principais (i.e., Geoprocessamento e Sistema Web) que serão detalhados a seguir.

Módulo 1: Geoprocessamento. Os dados tratados nessa etapa são os destinados principalmente para compor o mapa principal do sistema. Sendo assim, dados geográficos da região de interesse podem ser coletados de diversas fontes e formatos, passando por um tratamento inicial via algum *software GIS* (*Geographic Information System*) como, por exemplo, o *QuantumGIS* (QGIS)³. Uma vez tratados, os dados

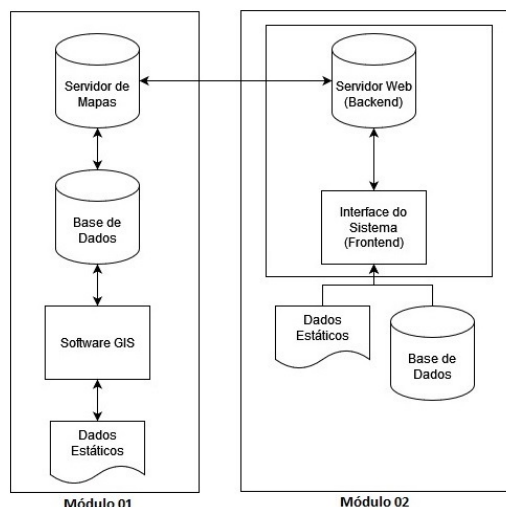


Figura 1. Arquitetura geral do sistema proposta.

são armazenados em um sistema de banco de dados geográficos, como o PostgreSQL⁴ com a extensão espacial PostGIS⁵, o que permite o armazenamento estruturado da informação e a realização de consultas simples ou complexas, relacionando as camadas entre si [5].

Por fim, o banco de dados é conectado a um servidor de mapas. Essa etapa é de grande importância por garantir assim a interoperabilidade dos dados coletados no primeiro momento através dos padrões definidos pela *Open Consortium Geospatial* (OGC)⁶, como o *Web Map Service* (WMS) e o *Web Feature Service* (WFS).

Módulo 2: Sistema Web. Esse módulo pode ser subdividido em dois componentes principais: *frontend* e *backend*. No *frontend*, foi empregado um fluxo linear de páginas, do tipo Aplicação de Múltiplas Páginas (do inglês *Multiple Page Application* - MPA⁷), conforme apresentado na Figura 2. Utilizamos para isso a Linguagem de Marcação de Hipertexto (do inglês *HyperText Markup Language* - HTML), a linguagem de estilo Folhas de Estilo em Cascata (do inglês *Cascading Style Sheet* - CSS), e a linguagem de programação Javascript. Já no *backend* do sistema, foi utilizado a linguagem de programação Python, versão 3.6. Aqui, vale destacar que exploramos os algoritmos de MD disponíveis na biblioteca *scikit-learn*⁸. No entanto, a metodologia empregada é genérica para acomodar qualquer outra estratégia de agrupamento de interesse.

Funcionalidades Implementadas. Ao acessar o sistema, o usuário tem duas possibilidades: (1) informar um valor pré-definido para o número de agrupamentos que serão realizados pelo TERRAININTELGIS, ou (2) explorar a funcionalidade

⁴<https://www.postgresql.org/>

⁵<http://postgis.net/>

⁶<https://www.ogc.org/standards/>

⁷<https://imasters.com.br/front-end/por-que-utilizamos-single-page-applications-spa>

⁸<https://scikit-learn.org/stable/modules/generated/sklearn.cluster.KMeans.html>

³<https://qgis.org/en/site/>

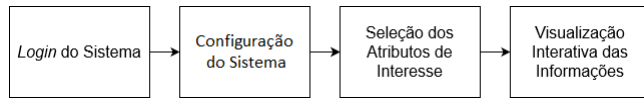


Figura 2. Sequência linear das telas implementadas no *front-end*.

de definição automática de número (ideal) de agrupamentos disponibilizada na ferramenta. Para esta versão da ferramenta, todo o agrupamento de dados está sendo realizado com o algoritmo *KMeans* [7], por entender que ele é bastante explorado na literatura e de fácil compreensão [3].

Além disso, o usuário pode ainda definir quais serão os atributos de interesse para usar os respectivos valores como entrada do algoritmo. Após a aplicação dos algoritmos pelo *backend* da ferramenta, é disponibilizado um mapa para visualização da informação resultante. Nesse mapa, há uma camada contendo toda a região de interesse, com todas as cidades contornadas. Cada cidade é preenchida com a cor do seu respectivo agrupamento, que pode ser verificada pela legenda correspondente gerada. Ao clicar sobre uma determinada cidade, é exibido um *popup* contendo informações relacionadas, como nome e agrupamento correspondente à cidade. Ademais, há a possibilidade de ativar e desativar mapas complementares, por meio de um controle de mapas disponíveis. Um exemplo de interação com esta funcionalidade do sistema é apresentado por meio da Figura 3(a).

Ao lado do mapa principal, também são apresentadas informações sobre cada um dos agrupamentos, como a média dos valores que os atributos possuem em um determinado agrupamento, assim como a relação das cidades que estão contidas nelas, conforme exemplificado na Figura 3(b). Há ainda a possibilidade de realizar o *download* desse conjunto de informações no formato de um arquivo PDF, para armazenamento ou uso posterior.

Por fim, há ainda um mapa de calor com a relação de cada um dos agrupamentos gerados e a relevância de cada um dos atributos escolhidos pelo usuário dentro daquele agrupamento (ver Figura 4 que exemplifica a escolha de um único atributo pelo usuário). Por meio de uma escala auxiliar é possível identificar que, quanto mais perto de vermelho, mais relevante são os atributos e quanto mais perto de azul, menos relevante são os determinados atributos.

4 Disponibilização do Sistema

O TERRAININTELGIS foi implementado utilizando o conceito de containerização possibilitado através da tecnologia *Docker*⁹. Dessa forma, a distribuição e implantação da aplicação ocorre mais facilmente. Apenas o módulo do geoprocessamento necessita de ser instalado e configurado localmente na rede do usuário. No entanto, todas as informações necessárias estão disponíveis no repositório do sistema, no seguinte link: <https://github.com/TavaresCarlos/TerrainIntelGIS>.

⁹<https://www.docker.com/>

5 Demonstração do Sistema

Para fins de demonstração, com o uso de dados públicos, o TERRAININTELGIS foi ajustado para um caso de estudo real. Especificamente, a aplicação de interesse foi configurada considerando a região da Bacia Hidrográfica do Rio Paraopeba (BHRP). Esta foi a bacia atingida pelos resíduos de mineração após o rompimento da Barragem I, da mina Córrego do Feijão, em 2019, na cidade de Brumadinho¹⁰. O vídeo está disponível no seguinte link: <https://youtu.be/7seAYuxYfqk>. No exemplo, foram considerados os dados (públicos) do Cadastro Ambiental Rural (CAR)¹¹ em relação aos 48 municípios que compõem essa bacia. Além disso, optou-se pela definição automática do valor ideal de agrupamentos (i.e., K), que foi igual a 4.

6 Potenciais Impactos Sociais

O TERRAININTELGIS é um SIG proposto com o uso exclusivo de tecnologias *opensource*. Dessa forma, um dos benefícios proporcionados está relacionado à economia de gastos relacionados a licenças ou planos. Além disso, buscou-se abstrair os conceitos de MD por meio do sistema, por entender que o usuário alvo pode não conhecer profundamente tais técnicas, o que pode favorecer aspectos de usabilidade.

Por fim, a aplicação aqui descrita é interessante por estar alinhada com os Objetivos de Desenvolvimento Sustentável (ODSs)¹² da Organização das Nações Unidas (ONU), em relação aos objetivos 6, 8, 9, 10, 11, 13, 14, 15 e 17. Vale destacar ainda que, dos ODSs citados, o de número 13 (i.e., “Ação Contra a Mudança Global do Clima”)¹³ merece destaque aqui, pois, entende-se que as mudanças climáticas que vêm ocorrendo levam ao desequilíbrio do meio ambiente e, conseqüentemente, ao aumento do número de desastres naturais. Ademais, sabe-se que uma das causas das mudanças climática é a exploração excessiva dos recursos naturais, o que contribui ainda mais para a ocorrência dos desastres ambientais.

7 Considerações Finais

Neste trabalho propomos o TERRAININTELGIS, um sistema para apoiar o usuário responsável (especialista) no processo de tomada de decisão em relação à alocação de recursos financeiros em cenários de desastres ambientais ou naturais. É importante mencionar que, além de dados públicos disponíveis, o usuário pode ainda utilizar bases de dados internas sigilosas para o uso dos algoritmos de MD, respeitando apenas o formato e o padrão do arquivo.

Buscando ainda avaliar aspectos de aceitação por parte dos possíveis usuários do sistema, ele foi apresentado de forma preliminar a membros de um grupo de pesquisa do

¹⁰<https://g1.globo.com/mg/minas-gerais/noticia/2023/01/25/quatro-anos-da-tragedia-em-brumadinho-270-mortes-tres-desaparecidos-e-nenhuma-punicao.ghtml>

¹¹<https://www.car.gov.br/#/>

¹²<https://brasil.un.org/pt-br/sdgs>

¹³<https://brasil.un.org/pt-br/sdgs/13>

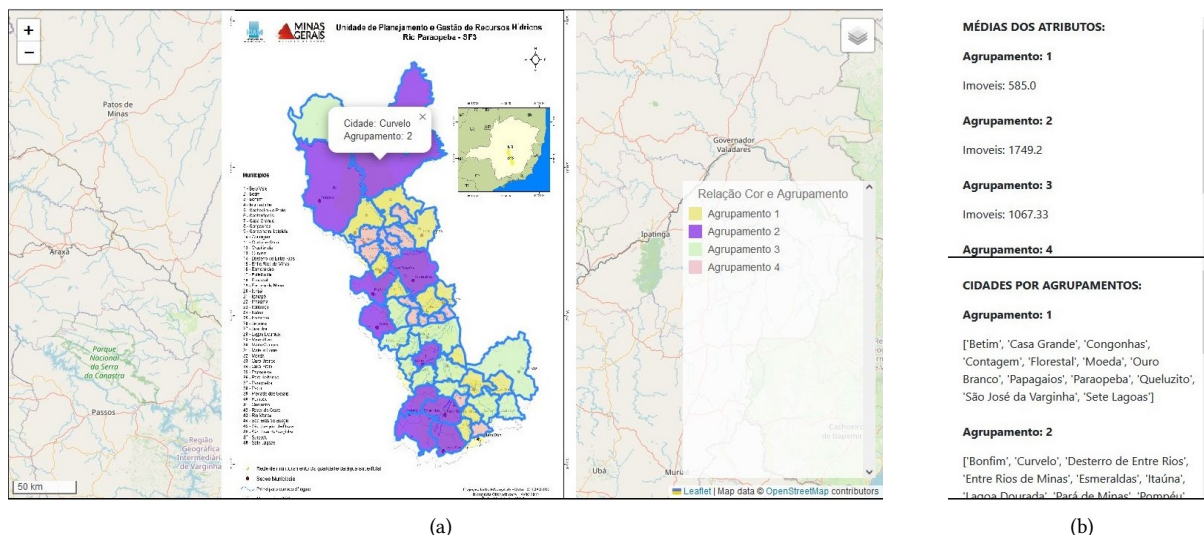


Figura 3. (a) Mapa customizado com camadas WMS e WFS resultantes do Módulo 01; (b) Informações complementares de cada um dos agrupamentos.

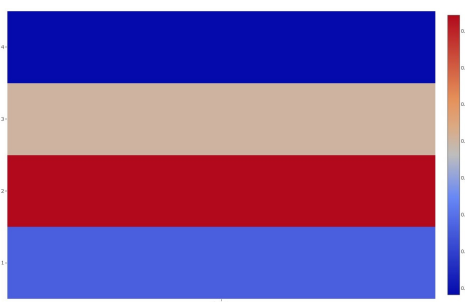


Figura 4. Mapa de calor com as propriedades escolhidas pelo usuário.

Departamento de Solos, da Universidade Federal de Viçosa (UFV-MG), que então avaliaram o seu uso. Tal grupo trabalha em cenários semelhantes ao do estudo de caso da demonstração do sistema (Seção 5). Em suma, os *feedbacks* obtidos foram bastante positivos. Potenciais usuários destacaram que o sistema poderá melhorar o desempenho e a eficácia em seus trabalhos relacionados a agrupamentos de espaços geográficos, bem como ser útil para apoiá-los no processo de tomada de decisão em cenários ambientais. Por fim, eles também manifestaram que pretendem utilizar a ferramenta, caso tenham acesso a ela.

Como trabalhos futuros, planejamos investigar alternativas que possam facilitar ainda mais a implantação do sistema por usuários finais interessados, bem como permitir a leitura e mineração de dados em formatos distintos como imagens e arquivos do tipo *geojson*. Além disso, é interessante ainda explorar o uso de outras técnicas de MD e aprendizado de máquina neste cenário, como previsão, por exemplo, para determinar os valores ideais a serem alcançados com os investimentos realizados. Finalmente, pretendemos realizar a avaliação do TERRAININTELGIS com possíveis usuários, para identificação e implementação de potenciais melhorias.

Agradecimentos. CAPES, CNPq e FAPEMIG e aos voluntários que participaram da avaliação preliminar do sistema.

References

- [1] Suwardi Annas and Zulkifli Rais. 2020. k-Means and GIS for mapping natural disaster prone areas in Indonesia. In *Proce. of the Mathematics, Science, and Computer Science Education Int'l Seminar (MSCEIS)*.
- [2] Francisco A. R. Barbosa, Paulina M. Maia-Barbosa, Andrea Maria A. Nascimento, et al. 2015. O desastre de Mariana e suas consequências sociais, econômicas, políticas e ambientais: porque evoluir da abordagem de Gestão dos recursos naturais para Governança dos recursos naturais? *Arquivos do Museu de História Natural e Jardim Botânico da UFMG* 24, 1-2 (2015).
- [3] Carlos H. T. Brumatti, Mariana A. R. Schaefer, Jugurta Lisboa Filho, and Julio C. S. Reis. 2022. Uma metodologia para apoio à tomada de decisão em cenários de desastres ambientais utilizando dados públicos. In *Proc. of the Brazilian Symposium on Geoinformatics (GEOINFO)*. 329–334. <http://urlib.net/ibi/8JMKD3MGPDW34P/487MDA5>
- [4] Cássio Oliveira Camilo and João Carlos da Silva. 2009. Mineração de dados: Conceitos, tarefas, métodos e ferramentas. *Universidade Federal de Goiás (UFG)* 1, 1 (2009), 1–29.
- [5] Jugurta Lisboa Filho and Cirano Iochpe. 1999. Um estudo sobre modelos conceituais de dados para projeto de bancos de dados geográficos. *Revista IP-Informática Pública* 1, 2 (1999), 37–90.
- [6] Mateus de O. Fornasier and Ana Lara Tondo. 2020. A responsabilidade social empresarial entre o direito, a economia e a política da sociedade global: desastres ambientais e reflexividade. *Caderno CRH* 32 (2020), 591–608.
- [7] James MacQueen et al. 1967. Some methods for classification and analysis of multivariate observations. In *Proc. of the Berkeley Symposium on mathematical statistics and probability*. 281–297.
- [8] Rienna Oktarina et al. 2021. Determine the clustering of cities in Indonesia for disaster management using K-Means by excel and RapidMiner. In *IOP Conference Series: Earth and Environmental Science*, Vol. 794. IOP Publishing, 012094.
- [9] Bambang Supriyadi, Agus Perdana Windarto, Triyuni Soemartono, et al. 2018. Classification of natural disaster prone areas in Indonesia using K-means. *Int'l Journal of Grid and Distrib.Comput.* 11, 8 (2018), 87–98.