

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DAINF - DEPARTAMENTO ACADÊMICO DE INFORMÁTICA
DAELN - DEPARTAMENTO ACADÊMICO DE ELETRÔNICA
CURSO DE ENGENHARIA DA COMPUTAÇÃO

HERON GOMES FERNANDEZ
PEDRO ROMANO SPLENDORE

**SISTEMA DE IDENTIFICAÇÃO AUTOMÁTICA DE RISCOS
HIDROMETEOROLÓGICOS COM RETROALIMENTAÇÃO E
REESTRUTURAÇÃO AUTÔNOMA DA INFRAESTRUTURA
DE COMUNICAÇÃO**

TRABALHO DE CONCLUSÃO DE CURSO

CURITIBA
2021

HERON GOMES FERNANDEZ
PEDRO ROMANO SPLENDORE

**SISTEMA DE IDENTIFICAÇÃO AUTOMÁTICA DE
RISCOS HIDROMETEOROLÓGICOS COM
RETROALIMENTAÇÃO E REESTRUTURAÇÃO
AUTÔNOMA DA INFRAESTRUTURA DE
COMUNICAÇÃO**

**AUTOMATIC IDENTIFICATION SYSTEM OF
CATASTROPHIC EVENTS WITH FEEDBACK AND
AUTONOMOUS REESTRUCTURING OF
COMMUNICATION INFRASTRUCTURE**

Trabalho de Conclusão de Curso apresentado ao Curso de Engenharia da Computação da Universidade Tecnológica Federal do Paraná, como requisito parcial para a obtenção do título de Bacharel.

Orientadores: Luiz Celso Gomes Jr. e Daniel Fernando Pigatto
DAINF - Departamento Acadêmico de Informática - UTFPR
DAELN - Departamento Acadêmico de Eletrônica - UTFPR

CURITIBA
2021



[4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Esta licença permite compartilhamento, remixe, adaptação e criação a partir do trabalho, mesmo para fins comerciais, desde que sejam atribuídos créditos ao(s) autor(es). Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

**HERON GOMES FERNANDEZ
PEDRO ROMANO SPLENDORE**

**SISTEMA DE IDENTIFICAÇÃO AUTOMÁTICA DE RISCOS
HIDROMETEOROLÓGICOS COM RETROALIMENTAÇÃO E REESTRUTURAÇÃO
AUTÔNOMA DA INFRAESTRUTURA DE COMUNICAÇÃO**

Trabalho de Conclusão de Curso de Graduação
apresentado como requisito para obtenção do título
de Bacharel em Engenharia da Computação da
Universidade Tecnológica Federal do Paraná
(UTFPR).

Data de aprovação: dia 9 de dezembro de 2021.

Elaiz Aparecida Mensch Buffon
Doutora em Geografia
Universidade Federal do Paraná - UFPR

Juliana de Santi
Professora Doutora em Ciência da Computação
Universidade Tecnológica Federal do Paraná - UTFPR

Luiz Celso Gomes Junior
Professor Doutor em Ciência da Computação
Universidade Tecnológica Federal do Paraná - UTFPR

**CURITIBA
2021**

Dedicamos este trabalho a todos aqueles que nos ajudaram durante nossa jornada, especialmente aos amigos e aos professores.

AGRADECIMENTOS

Dedicamos, em primeiro lugar, nossos agradecimentos aos professores orientadores, Prof. Luiz Celso Gomes Jr. e Prof. Daniel Fernando Pigatto. Muito obrigado professores, pelos ensinamentos, por toda a ajuda e por terem aceito e se disponibilizado a orientarem nosso trabalho de forma remota da melhor maneira possível.

Agradecemos também à UTFPR e seus docentes, que nos transmitiram muitos conhecimentos durante a nossa jornada acadêmica, possibilitando que nós nos tornássemos exímios profissionais.

Por fim, agradecemos a nossas famílias por todo o suporte prestado nesta fase de nossas vidas. Sem eles, não teria sido possível alcançar o nosso sonho de podermos contribuir para a sociedade utilizando-se da tecnologia. À todos os nosso amigos de Curitiba que pudemos compartilhar momentos inesquecíveis durante esse período.

Desculpe, café está fora de questão. Quando eu me mudei para a Califórnia, prometi a minha mãe que eu não começaria a usar drogas. (COOPER, Sheldon).

RESUMO

FERNANDEZ, Heron Gomes; SPLENDORE, Pedro Romano. Sistema de identificação automática de riscos hidrometeorológicos com retroalimentação e reestruturação autônoma da infraestrutura de comunicação. 2021. 44 f. Trabalho de Conclusão de Curso – Curso de Engenharia da Computação, Universidade Tecnológica Federal do Paraná. Curitiba, 2021.

A Internet se tornou essencial para a população em diversos aspectos da vida, facilitando a utilização de diversos serviços no dia-a-dia das pessoas. Por esse motivo, é fundamental contar com uma infraestrutura de comunicação confiável e resiliente. Porém, um dos fatores que contribuem para a fragilização dessa infraestrutura são as catástrofes naturais, como as inundações e os alagamentos que podem danificar meios físicos de transmissão e/ou inviabilizar seu funcionamento, podendo levar à indisponibilidade de serviços importantes. Utilizando uma modelagem baseada em detecção de *outliers*, este trabalho de conclusão de curso tem o intuito de identificar eventos catastróficos que podem causar falhas nos enlaces de comunicação e, a partir desta identificação, oferecer uma estratégia de reestruturação autônoma da infraestrutura. A proposta define uma retroalimentação entre a detecção de eventos e a reestruturação da rede, de forma a permitir que falhas identificadas na infraestrutura auxiliem na identificação de eventos catastróficos. A entrada do modelo são os dados meteorológicos, alagamentos e inundações de Curitiba, além de dados simulados de disponibilidade de acesso à Internet. Com este modelo, espera-se tornar mais robusto a infraestrutura de redes que conecta as unidades de saúde de Curitiba.

Palavras-chave: Alagamentos, inundações, detecção de *outliers*, infraestrutura de redes, Internet, smart cities.

ABSTRACT

FERNANDEZ, Heron Gomes; SPLENDORE, Pedro Romano. Automatic identification system of catastrophic events with feedback and autonomous restructuring of communication infrastructure. 2021. 44 f. Trabalho de Conclusão de Curso – Curso de Engenharia da Computação, Universidade Tecnológica Federal do Paraná. Curitiba, 2021.

The Internet has become essential for the population in different aspects of life, facilitating the use of various services in people's daily lives. For this reason, it is critical to have a reliable and resilient communication infrastructure. However, one of the factors that contribute to the weakening of this infrastructure are natural disasters, such as floods, which can damage the physical means of transmission and/or make its operation unfeasible and can lead to the unavailability of important services. Using a modeling based on *outliers* detection, this course conclusion work aims to identify catastrophic events that can cause failures in communication links and, based on this identification, offer an autonomous infrastructure restructuring strategy. The proposal defines a feedback between event detection and network restructuring, in order to allow identified infrastructure failures to help identify catastrophic events. The model input is meteorological data, flooding from Curitiba, in addition to simulated data on the availability of Internet access. Based on this model, it is expected to make the network infrastructure that connects the healthcare units in Curitiba more robust.

Keywords: Floods, outlier detection, network infrastructure, Internet, smart cities.

LISTA DE FIGURAS

Figura 1 – Diagrama do fluxo de etapas	21
Figura 2 – Mapa discretizado em hexágonos de Curitiba	23
Figura 3 – Interpolação dos dados pluviométricos em Curitiba	24
Figura 4 – Pontos de alagamento em Curitiba	26
Figura 5 – Interface do modelo da infraestrutura	28
Figura 6 – Simulação da rede com os nós e arestas	29
Figura 7 – Simulação da infraestrutura de redes em Curitiba	30
Figura 8 – Passo a passo da reestruturação da rede	33
Figura 9 – Problemas de interpolação	38
Figura 10 – Exemplificação da reestruturação	39

LISTA DE ABREVIATURAS E SIGLAS

ICARUS	<i>Integrated Crisis Awareness and Resource Utilization for Smartcities</i>
CEMADEN	Centro Nacional de Monitoramento e Alertas de Desastres Naturais
GIS	<i>Geographic Information System</i>
LOF	<i>Local Outlier Factor</i>
ISP	<i>Internet Service Provider</i>
DF	<i>DataFrame</i>
UBS	<i>Unidade Básica de Saúde</i>

LISTA DE ALGORITMOS

Algoritmo 1 – Função de reestruturação da rede	32
--	----

SUMÁRIO

1 – INTRODUÇÃO	14
1.1 Objetivos	15
1.2 Estrutura do trabalho	15
2 – FUNDAMENTAÇÃO TEÓRICA E TRABALHOS RELACIONADOS	16
2.1 Riscos hidrometeorológicos – inundações e alagamentos	16
2.2 Detecção de <i>outliers</i>	16
2.3 Interpolação de dados geográficos	17
2.4 Redes resilientes	17
2.5 <i>Smart cities</i>	18
2.6 Trabalhos Relacionados	18
3 – PLATAFORMA ICARUS	20
3.1 Diagrama de etapas e funcionamento da plataforma	20
3.2 Coleta e Análise de Dados	20
3.2.1 Dados pluviométricos	21
3.2.2 Dados de ocorrência de alagamentos	23
3.3 Modelo de detecção de <i>Outliers</i>	24
3.3.1 Definição do algoritmo de detecção	24
3.3.2 Tratamento dos dados	26
3.3.3 Validação e testes do modelo de detecção de <i>Outliers</i>	27
3.4 Simulação da infraestrutura de redes	28
3.4.1 Criação da topologia de rede a partir da representação em Grafos	28
3.4.2 ISPs - provedores de acesso à Internet	30
3.4.3 Integração do modelo de detecção de <i>outliers</i> com a simulação da infraestrutura	31
3.4.4 Reestruturação autônoma da infraestrutura de rede	32
3.4.5 Métricas para avaliar o estado da rede	34
4 – DISCUSSÃO DOS RESULTADOS	36
4.1 Resultados obtidos	36
4.1.1 Análise dos dados de precipitação e alagamentos	36
4.1.2 Retroalimentação da plataforma ICARUS	37
4.1.3 Reestruturação da rede a partir dos <i>outliers</i>	37
4.2 Discussão dos resultados	37

5 – CONCLUSÕES	40
5.1 Considerações finais	40
5.2 Trabalhos futuros	40
5.2.1 Bases de dados adicionais	40
5.2.2 Variações do modelo de detecção de <i>outliers</i>	41
5.2.3 Retroalimentação contínua dos modelos	42
5.2.4 Integração da plataforma ICARUS com sistemas automatizados .	42
5.2.5 Outras aplicações	42
REFERÊNCIAS	43

1 INTRODUÇÃO

O rápido desenvolvimento da tecnologia impõe um grande desafio à sociedade: fazer com que a conectividade seja uma realidade para os diferentes perfis sociais e regionais. Porém, ao passo em que a tecnologia evolui, barreiras comerciais e econômicas são criadas em torno dessas novas tecnologias, fazendo com que apenas uma pequena parte da sociedade possua acesso a elas. Com isso, certas áreas urbanas ficaram com uma infraestrutura de rede suscetível a falhas (MARÍN; CREUS; ZABALLOS, 2013). Essa fragilidade é evidenciada em alagamentos e inundações, que causam as falhas de comunicação (CORDOVA et al., 2020).

Outro fator que aumentou a importância e dependência de infraestrutura de redes foi a pandemia da COVID-19, fazendo com que diversas pessoas tivessem que adaptar suas atividades, anteriormente presenciais, para a realidade remota. Conforme o mostrado em CETIC.BR/NIC.BR/CGI.BR (2021), houve um grande aumento percentual do uso da Internet a partir do acesso Wi-Fi domiciliar, seguido de uma queda no uso de dados móveis celulares e do aumento do número de chamadas de voz e vídeo *on-line*. Isso causou uma sobrecarga em tais infraestruturas, tornando-as mais sensíveis em situações de falhas (CANDELA; LUCONI; VECCHIO, 2020). Assim, devido ao aumento da demanda e sobrecarga da rede, tornou-se ainda mais importante conseguir identificar com antecedência possíveis gatilhos de falhas de comunicação.

É de suma importância mantermos os serviços da rede de comunicações em funcionamento, como por exemplo, a rede de saúde da cidade. Essa foi a motivação ao longo do trabalho, reconectar primariamente as unidades de saúde que pudessem ficar sem comunicação com a Internet. Tendo como base a notícia do NIC.BR (2021), 69% das UBS possuem um sistema eletrônico para registro das informações dos pacientes e 65% para o registro de prontuários médicos. Com isso, caso uma UBS tenha uma falha em sua conexão de rede, não seria possível obter essas informações.

Este trabalho de conclusão de curso descreve a criação de um modelo computacional, baseado na identificação de anomalias ou, conforme será referenciado ao longo do trabalho, *outliers*. Utilizando de tais análises de dados e dados abertos, tal modelo auxilia na identificação de situações de catástrofes naturais que podem afetar sistemas de infraestrutura de redes. Essas informações geradas pelo modelo de detecção de *outliers* são enviadas para a simulação da infraestrutura de rede, possibilitando a mitigação dos efeitos causados por tais eventos. Por fim, é feita uma retroalimentação entre o modelo e a simulação, com o intuito de melhorar a acurácia do sistema.

1.1 Objetivos

O objetivo geral deste trabalho é desenvolver um modelo baseado em detecção de *outliers* para prever eventos de catástrofes naturais. O modelo atua em conjunto com o sistema de gerenciamento de infraestrutura de comunicação para mitigar problemas (MATISZIW; MURRAY, 2009). Com isso, é possível auxiliar na prevenção e tomar medidas de contingência para evitar futuras falhas de comunicação. A integração de todas essas partes recebe o nome de plataforma ICARUS (*Integrated Crisis Awareness and Resource Utilization for Smartcities*) e será melhor detalhada posteriormente. Os objetivos específicos desse trabalho de conclusão de curso são:

- Reunir dados coerentes para a alimentação de um modelo adequado de detecção de *outliers* e validar esse modelo com dados reais de catástrofes naturais.
- Simular um sistema de infraestrutura de redes e criar eventos de falhas de enlace;
- Validar o sistema de detecção de *outliers* com eventos de quedas de enlace simulados;
- Auxiliar nos macroprocessos da redução de riscos de desastres;
- Determinar se o sistema desenvolvido foi eficiente em manter a conectividade durante ocorrências de falhas na rede causadas por catástrofes.

1.2 Estrutura do trabalho

Este documento está dividido em 4 capítulos, incluindo esta introdução. O capítulo 2 aborda a base teórica, as tecnologias utilizadas no desenvolvimento do modelo e os trabalhos relacionados já existentes na literatura. O Capítulo 3 descreve a plataforma ICARUS (*Integrated Crisis Awareness and Resource Utilization for Smartcities*) como um todo, sendo este capítulo dividido entre o detalhamento dos dados utilizados, o modelo de detecção de *outliers* e a simulação da infraestrutura de redes, com a implementação e a validação dos resultados obtidos. Por fim, o Capítulo 4 apresenta os resultados, os trabalhos futuros que podem ser desenvolvidos a partir deste trabalho e as considerações finais.

2 FUNDAMENTAÇÃO TEÓRICA E TRABALHOS RELACIONADOS

Neste capítulo são abordados os conceitos sobre inundações e alagamentos, interpolação de dados, detecção de *outliers*, redes resilientes, *smart cities* e os trabalhos relacionados já realizados nessas áreas.

2.1 Riscos hidrometeorológicos – inundações e alagamentos

Como os termos inundações e alagamentos serão frequentemente utilizados neste trabalho, é importante diferenciá-los. As enchentes são definidas como a elevação do nível de um canal d'água de drenagem, chegando ao nível máximo sem transbordar. Já alagamentos são os acúmulos momentâneos de águas em determinados locais, geralmente ocasionados por deficiências no sistema de escoamento. Inundações se referem a eventos nos quais ocorre o transbordamento das águas de um fluxo de água, acarretando na inundação de uma planície ou área de várzea (NORONHA, 2021). O foco deste trabalho estará na análise de inundações e alagamentos, pois esses são alguns dos eventos que mais ocorrem em Curitiba, de acordo com a Avaliação de Riscos Climáticos da Cidade de Curitiba (CURITIBA, 2021). Nele, há o seguinte trecho:

[...] buscou identificar as principais ameaças climáticas registradas na cidade, considerando dados de ocorrência de eventos extremos entre 2012 e 2020. A partir de uma análise probabilística desses dados e considerando as avaliações históricas e futuras das variáveis climáticas, além da disponibilidade de dados que pudessem ser espacialmente explícitos, fez-se uma priorização das principais ameaças climáticas: alagamento, inundação, deslizamentos e suscetibilidade às ondas de calor.

2.2 Detecção de *outliers*

Em Chandola, Banerjee e Kumar (2009), o termo detecção de anomalia é definido como: encontrar padrões de dados que não entram em conformidade com o comportamento esperado em um certo domínio. Termos como anomalia e *outlier* são comumente usados para apresentar tais distúrbios. Um desafio comum para a modelagem de detecção de *outliers* é a constante mudança do que é considerado um comportamento normal em um dado domínio. As técnicas de detecção de anomalias consistem de três modos: supervisionado, semissupervisionado, não supervisionado. O modo supervisionado depende de um *dataset* de treinamento com dados considerados normais e anômalos pré-definidos, que torna a aplicação mais limitada, pois indicar os comportamentos de dados discrepantes não é trivial.

Além disso, surge espaço para um problema já conhecido no aprendizado de máquina: uma distribuição imprecisa do conjunto de dados pode gerar detecção enviesada, que pode ser acentuada com o fato de que as anomalias podem mudar de forma ao longo do tempo. A detecção semissupervisionada de *outliers* presume que os dados de treinamento possuem somente o comportamento normal. Por conta disso, eles são mais difundidos para diversos problemas. Por fim a detecção não supervisionada de *outliers* se utiliza de técnicas que não requerem dados de treinamento classificados e presumem que o conjunto de dados normais são consideravelmente maiores que o conjunto de dados anômalos. Esse modo é o mais versátil devido a tal característica.

2.3 Interpolação de dados geográficos

Para alimentar corretamente o modelo de detecção de *outliers*, é necessário ter os dados meteorológicos em todos os pontos geográficos do local em que serão identificados os alagamentos. Por conta disso, é necessário um método de interpolação dos pontos conhecidos com dado meteorológico.

Tendo em vista que os principais dados usados pelo modelo foram os dados de precipitação de Curitiba, o método considerado mais adequado para dados de precipitação é chamado de *Kriging* ou Krigagem. De acordo com o livro "Geoestatística: Conceitos e Aplicações" (YAMAMOTO; LANDIM, 2013, p. 54):

Krigagem é um processo geoestatístico de estimativa de valores de variáveis distribuídas no espaço e/ou tempo, com base em valores adjacentes quando considerados interdependentes pela análise variográfica. [...] a krigagem apresenta estimativas não tendenciosas e a mínima variância associada ao valor estimado.

2.4 Redes resilientes

Resiliência é a capacidade que algo ou alguém possui de responder com consistência e flexibilidade frente às dificuldades encontradas, conseguindo se recuperar diante de um contexto adverso (TABOADA; LEGAL; MACHADO, 2006). Logo, uma rede resiliente é aquela que possui a capacidade de se adaptar para manter níveis estáveis e aceitáveis de operação quando colocada frente a situações não favoráveis, como as inundações e alagamentos que podem afetar os equipamentos de rede. Com isso, faz-se necessário criar métodos e meios para que os administradores de redes tornem a infraestrutura gerenciada resiliente (VASCONCELOS, 2012).

Para alcançar tal objetivo, existem diversas estratégias para tornar uma rede resiliente, como a proteção por replicação (*backups*) e redundância de serviços e as estratégias de restauração de uma rede utilizando os recursos que estão disponíveis (NATALINO;

ROSTAMI; MONTI, 2021). Este trabalho de conclusão de curso irá utilizar as estratégias de restauração de uma rede utilizando os recursos adjacentes disponíveis.

2.5 *Smart cities*

Este trabalho parte do pressuposto de uma implementação de solução para uma *smart city*. *Smart cities* são cidades inteligentes no qual o objetivo é melhorar a qualidade de vida da população no dia a dia, utilizando-se das tecnologias disponíveis que causem um menor impacto ambiental (MADRID, 2021). Com isso, gera-se um aumento da eficiência do trabalho, da comunicação, do atendimento às pessoas e também na rentabilidade, devido a grande produtividade. Porém, para que uma *smart city* surja, é necessário grande esforço em planejamento e gestão, com o esforço contínuo de melhoria e implementação de avanços tecnológicos.

A infraestrutura de redes é um ponto crucial, visto que é a partir dela que todos os dados e informações que interligam uma *smart city* são conduzidos. Esse é um grande desafio, visto que *smart cities* tendem a se desenvolver rapidamente, fazendo com que essa infraestrutura não consiga acompanhar ao mesmo passo (KASZNAR et al., 2021). Logo, durante uma catástrofe, existe a chance de que uma fibra óptica dessa infraestrutura, que provê o sinal de banda larga para estabelecimentos críticos (como hospitais, pronto-socorros e bancos), seja afetada, fazendo com que esses percam a conexão com a Internet.

2.6 Trabalhos Relacionados

Esta subseção do trabalho tem como objetivo descrever alguns trabalhos científicos publicados que são relacionados a este trabalho de conclusão de curso.

O trabalho desenvolvido por Hazra et al. (2020) propõe a criação de uma infraestrutura de rede de comunicação temporária para eventos de pós-catástrofe em cenários com recursos limitados. A criação da infraestrutura de rede temporária é possível a partir dos *smartphones*, torres de comunicação e drones, fazendo com que exista um meio de comunicação entre voluntários, abrigos e o centro de coordenação. A proposta se assemelha à análise realizada neste trabalho, visto que ambos têm como finalidade mitigar ou recuperar o acesso à Internet, realizando ações na infraestrutura de redes antes ou depois de um evento catastrófico. Entretanto, no trabalho citado, o objetivo é criar uma rede local para a comunicação, já neste trabalho, o objetivo é religar as UBS na Internet como um todo, e não somente em uma rede local.

Já no trabalho de Natalino, Rostami e Monti (2021), propõe-se elaborar uma nova estratégia, intitulada de SCORE (*Storage protection with COnnectivity and processing REstoration*), para a sobrevivência dos serviços em *cloud*. Para alcançar tal objetivo, são combinadas as estratégias de proteção – criação de *backups* e redundâncias – e de restauração – utilização de recursos disponíveis na rede para recuperar serviços em caso

de falha – com o intuito de obter uma alta eficiência e uma alta disponibilidade na rede. Além disso, outro grande objetivo desse trabalho é reduzir o tempo de indisponibilidade dos serviços. Para isso foi criado um orquestrador de redes, sendo seu dever realizar a interface e o mapeamento dos serviços com a infraestrutura de redes e seus recursos. O objetivo desse trabalho vai ao encontro do proposto neste trabalho de conclusão de curso, visto que a finalidade de ambos é tentar garantir a disponibilidade dos recursos e serviços em uma rede, seja ela física ou em nuvem. Porém, no trabalho mencionado, a restauração da rede é feita somente via software por se tratar de SDN (*Software Defined Networking*), já neste trabalho, a restauração é realizada utilizando-se de mudanças físicas e de *hardware* na rede.

Em Buffon e Paz (2017), utiliza-se dados disponíveis sobre alagamentos e precipitações pluviométricas em Curitiba, estudando tais dados para verificar as possíveis correlações entre eles. Nele, há uma explicação extensa sobre os problemas enfrentados em relação ao pré-processamento dos dados, como forma de integração de dados e situações de confiabilidade de certos registros. Sendo assim, esse estudo será utilizado como base para este trabalho e auxiliará na melhor forma de integração de dados. Nos resultados obtidos de tal trabalho, os autores indicaram que há padrões e tendências nesses dados e recomendaram para trabalhos futuros a “utilização de *softwares* para a criação de procedimentos automáticos no processo de avaliação dos dados”, que é o proposto neste Trabalho de Conclusão de Curso utilizando a modelagem de detecção de anomalias. Além disso, é recomendado a análise de dados secundários para o enriquecimento da modelagem. Sendo assim, para otimizar o modelo de detecção de *outliers*, é possível acrescentar as análises e dados disponíveis em Buffon e Sousa (2018), o qual apresenta as áreas em Curitiba com maiores riscos de inundação.

Por fim, o trabalho apresentado em Wu, Liu e Chawla (2010) demonstra uma abordagem de detecção de anomalias em dados espaciais e de precipitação. Os autores propõem um algoritmo chamado *Outstretch* que é capaz de identificar os *top-k outliers*, utilizando a estatística espacial de varredura (KULLDORFF, 1997), que segundo Wu, Liu e Chawla (2010) “é capaz de calcular uma discrepância de valor válida que permite a descoberta de todos os *outliers*, não afetado pelas regiões vizinhas que possam conter valores faltantes”. A abordagem dos autores vai ao encontro do trabalho aqui proposto considerando o escopo da detecção de *outliers*, visto que serão feitas análises em dados meteorológicos, identificando e prevendo situações atípicas em tal domínio. Contudo, o trabalho diverge do proposto pois avalia a assertividade do algoritmo com dados do comportamento do fenômeno ENOS (El Niño Oscilação Sul), que é um fenômeno muito mais atípico e persistente do que os tipos de catástrofes naturais que serão avaliados neste trabalho de conclusão de curso.

3 PLATAFORMA ICARUS

Neste capítulo a plataforma ICARUS (*Integrated Crisis Awareness and Resource Utilization for Smartcities*) é introduzida e explicada em detalhes. O nome dado a essa plataforma busca representar a consciência que uma *smart city* possui a respeito das crises que a acometem, estando integrada a um módulo de gestão e utilização de recursos. Fazem parte dessa plataforma o modelo de detecção de *outliers* e o simulador de infraestrutura de redes, sendo a entrada dessa plataforma os dados pluviométricos, histórico de inundações e alagamentos e dados geográficos de latitude e longitude, e como saída uma ação tomada para mitigar uma falha de conexão na rede.

O projeto foi dividido em cinco etapas principais: coleta e análise de dados, criação de um modelo de detecção de *outliers* para alagamentos, validação do modelo criado, simulação de uma infraestrutura de redes e integração do modelo de detecção de *outliers* com a simulação para a comunicação de dados dos modelos criados.

3.1 Diagrama de etapas e funcionamento da plataforma

O diagrama do fluxo de etapas da plataforma pode ser visto na Figura 1. Com ele, é possível identificar as etapas desse trabalho e como é realizada a interação e a integração entre ambos os modelos. A primeira etapa exibida nesse diagrama mostra o gerenciamento e aquisição de todos os dados utilizados. A segunda e a terceira etapa descrevem como o modelo de detecção de *outliers* interage com o modelo de simulação da infraestrutura e vice-versa.

3.2 Coleta e Análise de Dados

A integração de dados envolveu o uso da linguagem *Python*¹, juntamente do pacote *Pandas*², pois as fontes de dados possuem formatos desconexos e distintos.

Foram utilizados os dados climáticos do Cemaden (2021) que disponibiliza dados de precipitação por hora e possui onze estações pluviométricas automáticas na região de Curitiba. A partir disso, foi possível realizar a interpolação dos dados para obter as condições climáticas dos pontos médios, usando o pacote *OpenTurns*³.

Os dados utilizados nesta pesquisa, registros pontuais dos casos de alagamentos, foram coletados junto à Coordenadoria Municipal de Defesa Civil do município de Curitiba⁴, que foram compilados pelo Instituto de Pesquisa e Planejamento Urbano de Curitiba

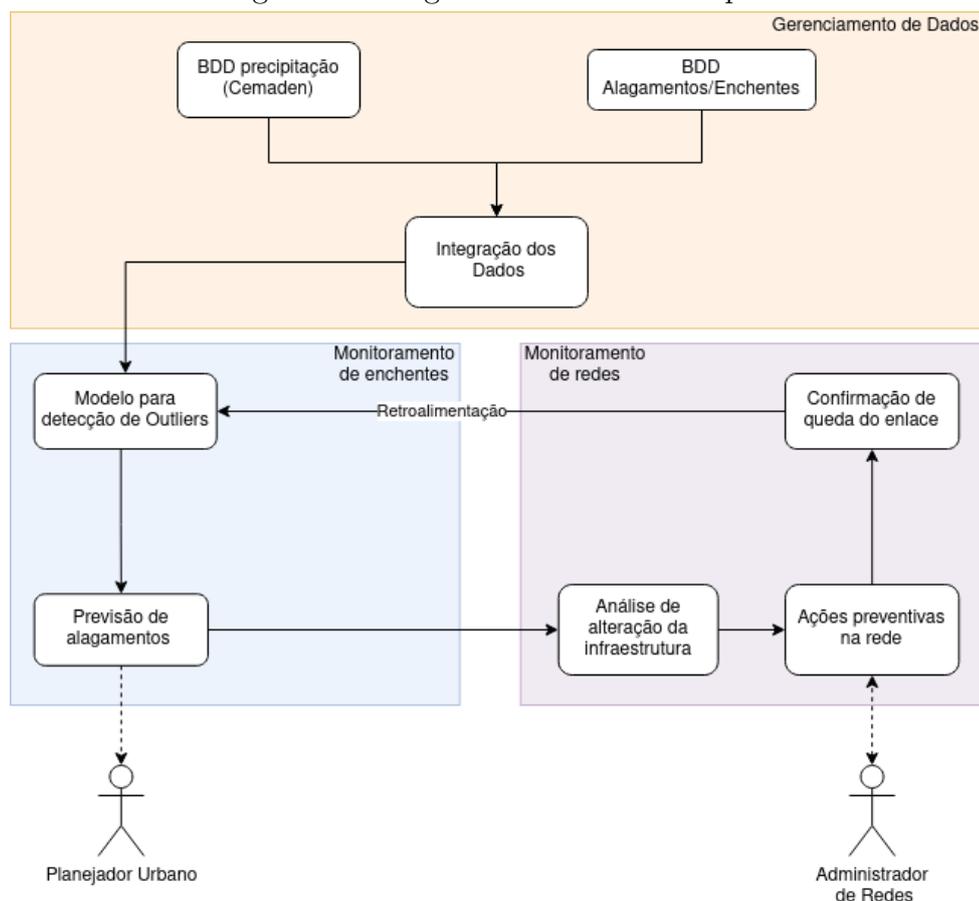
¹ *Python* Documentation: <<https://www.python.org/doc/>>

² *Pandas* Tutorial: <https://pandas.pydata.org/docs/getting_started/index.html>

³ *OpenTurns* Documentation: <<https://openturns.github.io/openturns/latest/contents.html>>

⁴ Dados utilizados na pesquisa de Buffon e Sousa (2018)

Figura 1 – Diagrama do fluxo de etapas



Fonte: Autoria Própria

(IPPUC). De acordo com Lohmann (2013), as informações contidas nos registros são armazenadas por um sistema chamado “SISGESGUARDA” (Sistema de Gerenciamento da Guarda Municipal), sendo esse sistema alimentado por ligações telefônicas recebidas na central de atendimentos e informações da Prefeitura Municipal de Curitiba. As informações armazenadas para cada uma das ocorrências, após o processo de compilação dos registros, são as seguintes: 1) Natureza (somente alagamento); 2) Data de ocorrência (dia/mês/ano/hora); 3) Data de atendimento (dia/mês/ano/hora); 4) Secretaria; 5) Logradouro; 6) Bairro; 7) Regional e, 8) Coordenadas.

3.2.1 Dados pluviométricos

O primeiro passo foi coletar os dados do CEMADEN (Centro Nacional de Monitoramento e Alerta de Desastres Naturais), acessando o mapa interativo disponibilizado. A tabela 1 mostra um trecho dos dados obtidos, sendo que a coluna *valor* possui os dados de valores medidos de precipitação de chuva em milímetros. Em seguida, conforme explicado no site do Cemaden (2021), os dados pluviométricos não são tratados e estão suscetíveis a erros, como por exemplo: ausência de medição quando não há chuva, preenchimento dos

dados com informações incoerentes ou mudança no formato dos dados.

Tabela 1 – Dados originais do CEMADEN.

codEstacao	uf	nomeEstacao	latitude	longitude	datahora	valor
410690203A	PR	Vista Alegre	-49,301	-25,413	1/1/15 00:00	0
410690203A	PR	Vista Alegre	-49,301	-25,413	1/1/15 12:20	0,2
410690203A	PR	Vista Alegre	-49,301	-25,413	1/1/15 11:10	0
410690203A	PR	Vista Alegre	-49,301	-25,413	1/1/15 11:20	0
410690203A	PR	Vista Alegre	-49,301	-25,413	1/1/15 11:30	0

Para a verificação de mudança nos formatos, as linhas que tinham os tipos de dados identificados corretamente eram salvam em outra tabela provisória para então tratar os dados não identificados. Avaliando esse outro formato, foi possível observar que esses valores variavam conforme os padrões de localidade americanos e europeus.

Também, a taxa de atualização das estações é de 10 minutos. Como esse nível de precisão não é necessário, os dados foram agrupados por hora e estação, fazendo a somatória dos valores medidos. Após esse tratamento, as linhas com os horários ausentes de cada estação pluviométrica foram preenchidas indicando que não ocorreu chuva. Para definir quais eram os horários que faltavam, foram encontradas as datas iniciais e finais de anotação de valores para cada estação. A motivação para essa escolha foi que nem todas as estações estavam disponíveis ao mesmo tempo.

Para a geolocalização dos dados, o mapa de Curitiba foi discretizado em hexágonos com 0,7 km de distância entre os pontos mais distantes. Para isso, utilizamos os dados presentes no repositório do *GitLab*⁵. Essa discretização pode ser vista na figura 2.

O novo mapa foi utilizado para fazer a interpolação dos dados obtidos pelas estações pluviométricas com o intuito de obter um valor estimado de chuva em cada um dos hexágonos. Para alcançar esse objetivo, todos os dados de estações pluviométricas foram usados para alimentar o método de interpolação geoespacial conhecido como Kriging ou Krigagem.

Para o modelo de covariância da Krigagem, foi usado um modelo definido como *SquaredExponential* pelo *OpenTurns* para conseguir uma distribuição não linear dos dados. O número de dimensões definido para o modelo base foi 2, sendo um para latitude e outro para longitude de cada estação. Foi utilizado o valor medido por cada estação para fazer o *fit* dos dados de entrada. Além disso, o treinamento é feito por hora de medição, ou seja, o modelo passa pelo processo de ajuste/*fit* a cada hora de forma independente da hora anterior.

A fim de obter os resultados de saída do modelo de interpolação (ou seja, os dados de precipitação estimados em cada hexágono de Curitiba), são obtidas as centroides desses hexágonos e finalmente os valores de interpolação são retornados. Com isso, é

⁵ *GitLab* Repository: <<https://gitlab.com/luizcelso/mobilidade-curitiba-covid19/-/tree/master/data/geo>>

Figura 2 – Mapa discretizado em hexágonos de Curitiba



Fonte: Autoria Própria

possível calcular o número de linhas retornadas por esse modelo, porque são 617 hexágonos distintos, com dados de precipitação por hora durante um ano, ou seja, são $617 \times 24 \times 365 = 5404920$ linhas. A tabela 3 mostra um trecho dos resultados dessa interpolação junto com os dados de predição do modelo (que está detalhado na próxima subseção). A figura 3 mostra o resultado da interpolação com os dados do maior caso de chuva em uma hora, se referindo às 16 horas do dia 21 de outubro de 2015.

3.2.2 Dados de ocorrência de alagamentos

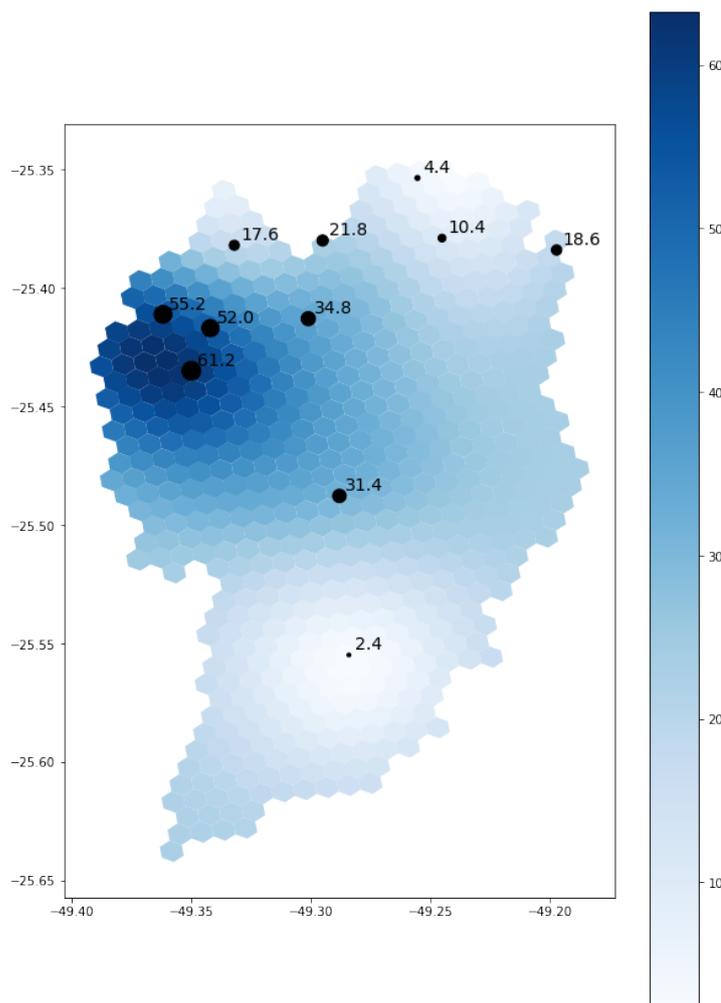
Os dados coletados de alagamentos em Curitiba são da Secretaria de Defesa Civil⁶ para a validação do modelo. A tabela 2 mostra as principais colunas desses dados, sem os pontos geométricos. É importante pontuar que esses dados sofrem com questões de confiabilidade, pois nem todos os casos de alagamentos são informados para a Defesa Civil, visto que houve quedas nos números oficiais de alagamentos, o que não reflete a realidade. Além disso, há questões políticas de falta de investimento, o que faz com que esse setor não seja priorizado. Esses problemas com os dados também fazem com que haja uma imprecisão no que é definido como inundação e alagamento.

Do ponto de vista da tecnologia, tais dados são disponibilizados em um formato do tipo GIS, sendo, portanto, trabalhado no *Geopandas*⁷. Por esse motivo, cada hexágono recebeu um ID único e foi feita a junção espacial dos hexágonos com os pontos de latitude e longitude dos casos de alagamento. A figura 4 mostra o resultado dessa junção com os pontos de alagamento em azul, os hexágonos com esses pontos de alagamento em amarelo

⁶Dados derivados da pesquisa de Buffon e Paz (2017)

⁷*Geopandas* Tutorial: <https://geopandas.org/en/stable/getting_started.html>

Figura 3 – Interpolação dos dados pluviométricos em Curitiba



Fonte: Autorial Própria

e os hexágonos que representam o resto de Curitiba em verde. Em seguida, foi feita a junção desses IDs únicos e datas de alagamento com os dados interpolados para obter uma forma de calcular a assertividade do modelo levantado.

3.3 Modelo de detecção de *Outliers*

Conforme foi apresentado na seção 2, é necessário fazer o levantamento dos dados conhecidos, o tipo de supervisionamento dos dados, o conhecimento prévio da equipe e o tempo disponível para fazer a implementação para definir o modelo. Sendo assim, foi considerada como a melhor abordagem usar o algoritmo *Local Outlier Factor* (LOF).

3.3.1 Definição do algoritmo de detecção

O algoritmo *Local Outlier Factor*, apresentado em Breunig et al. (2000), tem como finalidade encontrar *outliers* localmente, ou seja, esse modelo analisa os vizinhos para

Tabela 2 – Dados originais de alagamento da Defesa Civil.

ENDERECO	BAIRRO	CIDADE	DATA	HORA	FLAGRA
R. SANTA ANASTÁSIA, 14	CAJURU	CURITIBA	2015-12-26	16:41:00	NÃO
AV. ANITA GARIBALDI, 4895	BOA VISTA	CURITIBA	2015-12-25	21:28:00	NÃO
R. PEDRO GUSSO, 60	NOVO MUNDO	CURITIBA	2015-12-25	21:25:00	NÃO
R. CAPITÃO JOÃO ZALESKI, 211	PINHEIRINHO	CURITIBA	2015-12-25	21:19:00	NÃO
R. INÁCIO WOLSKI, 612	CIDADE INDUSTRIAL	CURITIBA	2015-12-25	20:55:00	NÃO
R. JORDÂNIA, 620	CAJURU	CURITIBA	2015-12-25	20:52:00	NÃO
R. ALBERTO MULLER SOBRINHO, 120	PORTÃO	CURITIBA	2015-12-25	20:09:00	NÃO

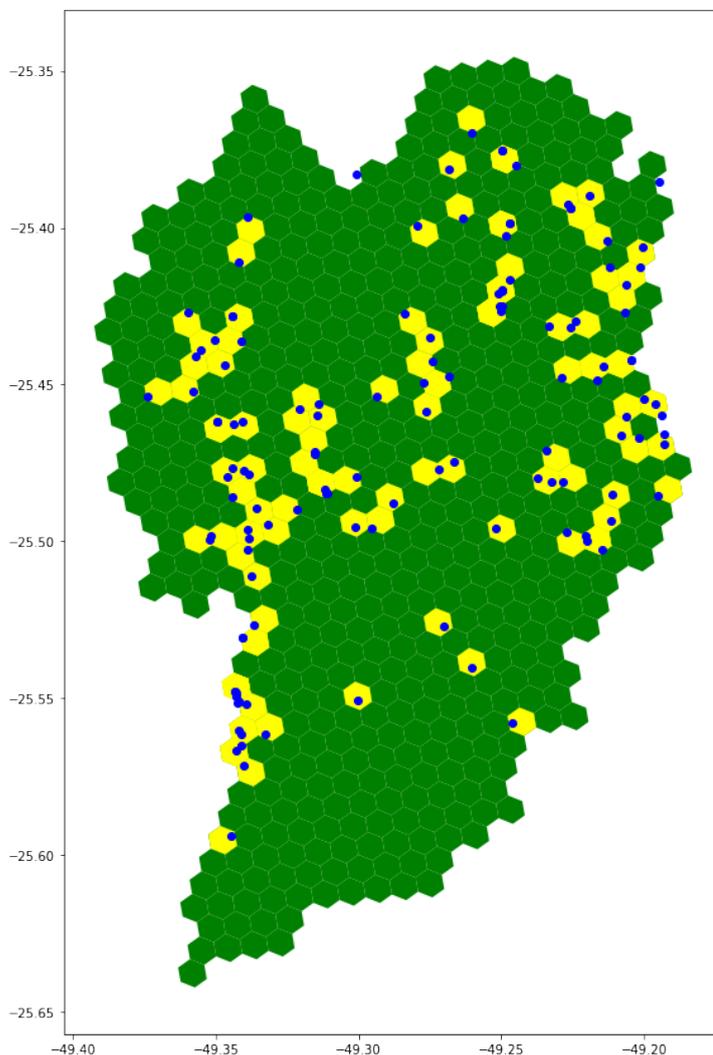
verificar quanto esses valores destoam em relação aos dados mais próximos. Devido a esse funcionamento, esse algoritmo é classificado como não supervisionado, não dependendo da sinalização dos casos que desejamos encontrar.

Outro fator importante é que esse algoritmo se comporta de forma adequada para dados com maiores dimensões e é suficientemente rápido para um grande número de dados, sendo essas duas situações presentes nesse problema.

Para a implementação, foi utilizado o pacote *scikit-learn*, pois nele já existe uma implementação otimizada do modelo, precisando definir quais dados podem ajudar a identificar o problema. Feito isso, os dados selecionados foram colocados como entrada do algoritmo junto ao número de vizinhos a ser avaliado e taxa de contaminação, sendo essa a porcentagem dos dados que são considerados como *outliers*.

Por fim, é essencial notar que outras abordagens poderiam ser usadas para lidar com esse problema. É possível utilizar outros algoritmos de aprendizado de máquina, como por exemplo: outros tipos de algoritmos de detecção de *outliers*, algoritmos de classificação ou modelos de regressão logística. Inclusive, esses dois últimos exemplos possuem a propriedade de ser um modelo de aprendizado de máquina supervisionado. Ou seja, há uma grande gama de abordagens que se baseiam no uso de inteligência artificial que podem ser usados para a análise do problema de identificação de alagamentos e inundações.

Figura 4 – Pontos de alagamento em Curitiba



Fonte: Aatoria Própria

3.3.2 Tratamento dos dados

Após os dados de precipitação serem tratados, como apresentado na seção anterior, eles foram normalizados usando a função `preprocessing.normalize()` do `scikit-learn` que coluna todos os valores de cada coluna para uma escala entre 0 e 1. A motivação da normalização se dá pois, como o LOF usa a distância de cada ponto entre os seus vizinhos, os valores dos dados devem ter a mesma escala para não terem pesos distintos e influenciar o algoritmo.

O LOF teve como entrada os dados de latitude, longitude, data e hora (enumerados em `timestamp`) e valor pluviométrico medido. Além disso, o parâmetro de número de vizinhos foi definido como 15 e a contaminação foi calculada para que a quantidade de *outliers* fosse 4 vezes maior a quantidade de alagamentos no ano de 2015. Assim, um exemplo de retorno do modelo pode ser visto na tabela 3, sendo que a coluna *valor*

representa os valores medidos pelos pluviômetros, o *LOF_score* quantifica quão anômalo o dado é em relação aos seus pares (quanto menor, maior o grau de *outlier*) e o *LOF_pred* indica se os dados são considerados *outliers* (-1 indica que é *outlier* e 1 indica que não é).

Tabela 3 – Dados interpolados do CEMADEN com os resultados do LOF.

valor	latitude	longitude	dataHora	LOF_score	LOF_pred
0	-49,38760	-25,43321	2015-01-01 00:00:00	-1,295563	1
0	-49,38503	-25,44681	2015-01-01 00:00:00	-1,294875	1
0	-49,38431	-25,42597	2015-01-01 00:00:00	-1,294029	1
0	-49,38319	-25,48124	2015-01-01 00:00:00	-1,298531	1
0	-49,38246	-25,46041	2015-01-01 00:00:00	-1,294029	1
0	-49,38174	-25,43957	2015-01-01 00:00:00	-1,298531	1
0	-49,38102	-25,41872	2015-01-01 00:00:00	-1,296409	1

3.3.3 Validação e testes do modelo de detecção de *Outliers*

Para avaliar o modelo de detecção de *outliers*, a principal medida foi o *F-score*, que é uma medida estatística para avaliar a acurácia. A motivação para tal escolha foi porque ele combina os valores de precisão e revocação. A precisão analisa a porcentagem de dados que foram corretamente previstos dentro de todos os dados previstos. Já a revocação analisa a porcentagem de dados corretamente identificados dentre todos os dados corretos. Dado essas variáveis, é possível calcular o *F-score* conforme a equação 1.

$$F\text{-score} = \frac{2}{\text{precisão}^{-1} + \text{revocação}^{-1}} = 2 * \frac{\text{precisão} * \text{revocação}}{\text{precisão} + \text{revocação}} \quad (1)$$

Conforme apresentado na seção 3.2.2, os dados de alagamentos apontados pela Defesa Civil serviram para fazer a validação dos *outliers* identificados pelo LOF. Após feita a junção dos dados, eles foram agrupados por dia e hexágono de Curitiba, fazendo a soma dos valores medidos, e sinalizando caso encontrasse qualquer dado de alagamento ou *outlier*. Feito isso, houveram 487 situações únicas consideradas como *outliers* e 111 casos de alagamentos em hexágonos distintos por dia.

Contudo, não houve qualquer intersecção entre os dados apontados como *outliers* e os casos reais. Por esse motivo, foi feito um agrupamento somente pelos dias, e assim, comparado os resultados para saber se os dados estavam minimamente coerentes. Feito isso, o número de dias com alagamentos em Curitiba foi 28, e 128 foi o número de dias que o LOF encontrou algum valor anômalo a cada hora em Curitiba. Sendo assim, o *F-score* calculado desse modelo nessas condições foi de 0,143.

É necessário frisar que esses resultados são preliminares e outras informações, como as de retroalimentação, serão agregadas no modelo final. Para o modelo ficar mais preciso, é necessário diminuir o nível de contaminação do algoritmo LOF para um número

mais próximo aos casos de alagamento. Outra questão é que a detecção de alagamentos e inundações somente com dados de precipitação de chuva é frágil, pois o que causa esses alagamentos envolve outras variáveis do ambiente, como entupimento de bueiros e outras questões relacionadas a ação humana.

3.4 Simulação da infraestrutura de redes

O objetivo desta etapa foi criar um modelo que simule uma infraestrutura de redes essencial para uma cidade inteligente. No caso, usamos dados de unidades de saúde de Curitiba que estariam conectadas à Internet. Esta conectividade é considerada essencial e é o objetivo da plataforma mantê-la. Com isso, foi possível realizar os testes em um cenário mais próximo à realidade. A interface desse modelo pode ser vista na Figura 5.

Figura 5 – Interface do modelo da infraestrutura

```
Digite a opção:
1 para visualizar o mapa no estado atual
2 para gerar o arquivo com os outliers detectados
3 para ler o arquivo de outliers para remoção de arestas
4 para remover uma aresta manualmente
5 para realizar a reestruturação da rede
6 para verificar as métricas de conectividade e eficiência
7 para sair
Opção escolhida:
```

Fonte: Autoria Própria

3.4.1 Criação da topologia de rede a partir da representação em Grafos

A simulação da rede foi feita utilizando uma representação em grafos, sendo as arestas as fibras ópticas e os nós as unidades de saúde. Foi utilizado o *Python*, juntamente com as bibliotecas *NetworkX*⁸ e *Geopandas*⁹ para realizar essa construção e simulação.

Tomando como base o mapa de Curitiba, mostrado na figura 2, e suas coordenadas geográficas, foram inseridos os nós (sendo os nós as unidades de saúde), usando as coordenadas de latitude e longitude. Esses nós foram classificados como pertencentes a um ISP (*Internet Service Provider*) específico, que será melhor detalhado nas próximas seções.

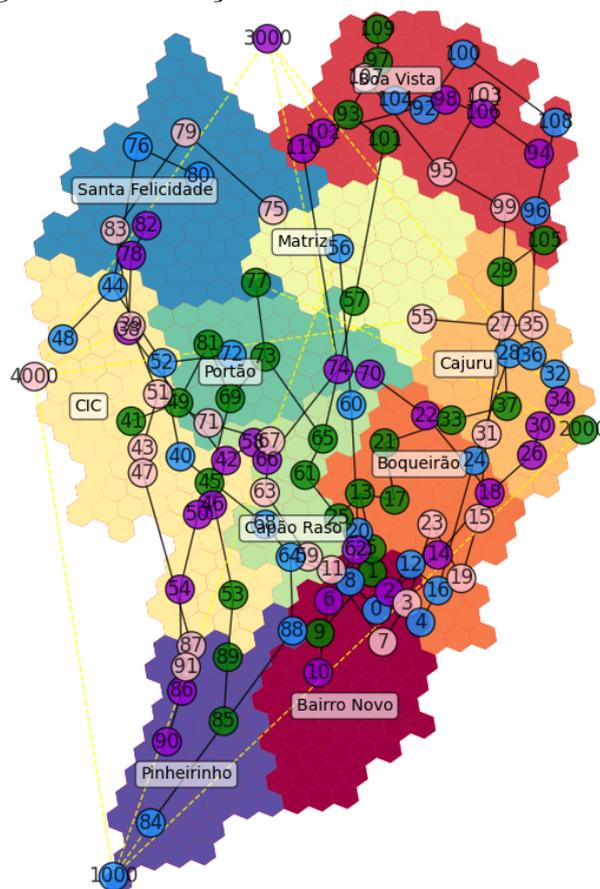
Para criar as conexões dos nós, foram criadas arestas que ligam esses nós aos nós mais próximos deles, sendo possível conectar somente nós que pertençam ao mesmo ISP. Com isso, houveram casos, por exemplo, em que o nó 1 e o nó 2 eram reciprocamente os mais próximos entre eles, criando apenas um subgrafo de dois nós. Logo, algumas

⁸*NetworkX* Documentation: <<https://networkx.org/documentation/stable/>>

⁹*Geopandas* Tutorial: <https://geopandas.org/en/stable/getting_started.html>

partes do grafo ficaram desconexas, possuindo apenas uma conectividade local. Realizamos então a ligação dos subgrafos, utilizando a menor distância entre dois pontos de subgrafos diferentes. Assim, finalizando a etapa de simulação da infraestrutura de redes que conecta essas unidades de saúde simuladas, que pode ser vista na figura 6.

Figura 6 – Simulação da rede com os nós e arestas



Fonte: Autoria Própria

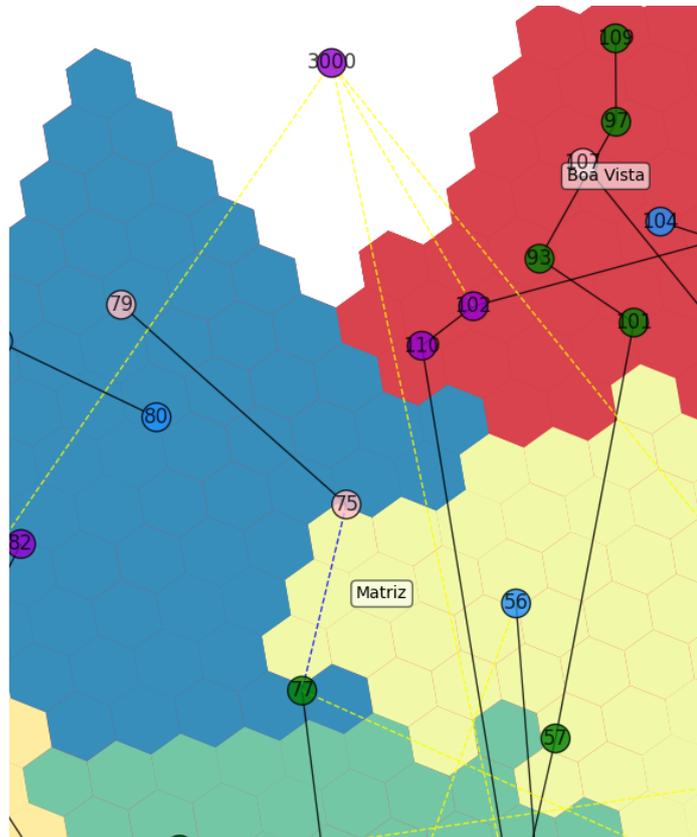
Para realizar os cálculos da reestruturação e das métricas da rede, foi adicionado às arestas do grafo um peso representando custo de comunicação entre os pontos. Quanto maior for o valor do peso, pior será a conexão entre dois pontos, sendo que, um peso zero significa largura de banda infinita e um peso infinito significa não ser possível trafegar nenhum dado por essa aresta. A fórmula utilizada para calcular o peso entre dois nós pode ser vista na equação 2. Nessa equação, $popA$ e $popB$ representam as populações referentes aos bairros nos quais cada nó está, $popMax$ é igual à população do bairro com o maior número de pessoas e $dist$ representa a distância entre o nó A e o nó B. Quanto mais populosa for a área na qual os nós se encontram, maior será o valor normalizado que multiplica o peso da aresta criada, visto que são áreas mais críticas e com uma maior demanda de dados. Utilizamos a distância como parte do cálculo do peso, pois não foi possível encontrar informações a respeito do tráfego de dados na rede. Dividimos os valores

das populações por dois para encontrar a média e dividimos pela população máxima para normalizar esse valor, fazendo com que fique entre zero e um. Finalmente, esse valor normalizado é multiplicado pela distância (em quilômetros), o que resultará no peso que é atribuído a essa aresta.

$$Peso = \frac{\frac{popA+popB}{2}}{popMax} * dist \quad (2)$$

Para representar cada conexão suscetível a falhas foi utilizada uma linha preta contínua. Para representar cada conexão referente aos *ISP*, por definição imunes a quedas de enlace, foi utilizada uma linha pontilhada amarela. Por fim, para simular uma reconexão realizada no grafo, foi utilizada uma linha pontilhada azul. Essas arestas estão ilustradas na Figura 7.

Figura 7 – Simulação da infraestrutura de redes em Curitiba



Fonte: Autorial Própria

3.4.2 ISPs - provedores de acesso à Internet

Com o intuito de simular uma conexão de rede real, foram criados quatro pontos que representam os ISPs no mapa de Curitiba, para que assim, pudéssemos ter um nó destino para realizar a reestruturação da rede. Esses pontos foram distribuídos, um ao norte

de Curitiba, um ao leste, um a oeste e um ao sul e os números dos nós que representam os ISPs são: 1000, 2000, 3000 e 4000. Cada um desses nós possui uma cor específica, sendo a mesma cor utilizada nos nós clientes desse ISP para a identificação no grafo. A cada nó cliente foi atribuído um ISP/cor, seguindo a fórmula 3,

$$\text{NúmeroISP} = (\text{Resto}(u/4) + 1) * 1000 \quad (3)$$

em que u é o número do nó e NúmeroISP representa o ISP ao qual o nó u fará parte, sendo que, caso o valor do resto seja zero, o ISP será o de número 1000, caso seja 1 o ISP será o 2000 e assim por diante. Com isso, distribuimos os nós para os ISPs de forma homogênea no mapa de Curitiba, pois os nós são numerados primeiramente pelo bairro em que estão, por exemplo, nó 0 até o nó 10 pertencem ao bairro CIC. A finalidade da criação foi simular o provedor de serviços de Internet ao qual o nó deve possuir uma conexão física estabelecida para conseguir alcançar a Internet.

Foi criado uma aresta entre esses ISPs, ligando os quatro entre si com um peso zero (considerando que não existam restrições de tráfego e banda entre eles), com a finalidade de, futuramente, podermos nos reconectar a qualquer ISP, visto que em uma situação de catástrofe, os provedores iriam se auxiliar nessa missão de reconexão das UBS. Foram também criadas duas conexões de cada ISP com as unidades de saúde do grafo, sendo uma conexão em um ponto central e uma conexão no ponto mais próximo a ele. Vale ressaltar que esse valor zero poderia facilmente ser alterado para um valor significativo, caso houvessem limitações físicas ou lógicas entre esses ISPs. Uma das premissas utilizadas foi que as arestas dos ISPs são imunes aos eventos catastróficos, como alagamentos e inundações, supondo que tais conexões são conexões subterrâneas e não seriam influenciadas por tais eventos. A motivação por trás dessa escolha foi para deixar a reestruturação possível em todos os casos.

3.4.3 Integração do modelo de detecção de *outliers* com a simulação da infraestrutura

A integração da detecção de *outliers* com a simulação da infraestrutura consiste em enviar para a simulação todos os *outliers* encontrados. A partir disso, a simulação decide se é relevante ou não alterar as conexões da rede. Foi envolvido nessa integração o envio para o detector de *outliers* os dados referentes às quedas de enlace detectadas pela simulação, sendo esses dados acrescentados no modelo com o intuito de melhorar a acurácia do modelo.

A integração da simulação de infraestrutura de redes envolveu enviar manualmente os arquivos com os *outliers* detectados pelo modelo para receber os alertas de possíveis quedas de enlace. Esses alertas são enviados no formato .geojson, possuindo as coordenadas dos hexágonos afetados em Curitiba. Então, a simulação da infraestrutura realiza a identificação do hexágono e faz uma chamada à função de intersecção. O objetivo dessa

função é verificar quais arestas estão passando por aquele ou aqueles hexágonos indicados pelo detector de *outliers* e com isso é feita a deleção dessa aresta. Então, a simulação gera um outro arquivo .geojson que é enviado manualmente para o modelo, para a realização da retroalimentação. Os hexágonos afetados, detectados pelo modelo, são marcados em vermelho no mapa de Curitiba.

3.4.4 Reestruturação autônoma da infraestrutura de rede

Com o intuito de realizar a reestruturação da infraestrutura de redes simulada foram implementadas funções com o objetivo de religar os nós que perderam conectividade com a Internet. Para realizar essa reconexão, é feita uma busca na rede a fim de localizar um nó destino, sendo este nó o que possui o menor custo até um ISP, maximizando assim a largura de banda disponível para o nó religado. Um exemplo desse processo pode ser visto na figura 8.

Diante de algum evento catastrófico, é possível que surjam múltiplos subgrafos que não possuam conectividade com a Internet, sendo necessário classificar e priorizar o subgrafo com o maior número de nós desconectados. Para isso, foi implementada uma fila de prioridades para a execução da recuperação de cada nó, levando em consideração o número de nós pertencentes a um subgrafo desconectado dos ISPs. Um pseudocódigo dessa reestruturação pode ser visto no algoritmo 1.

Algoritmo 1: Função de reestruturação da rede

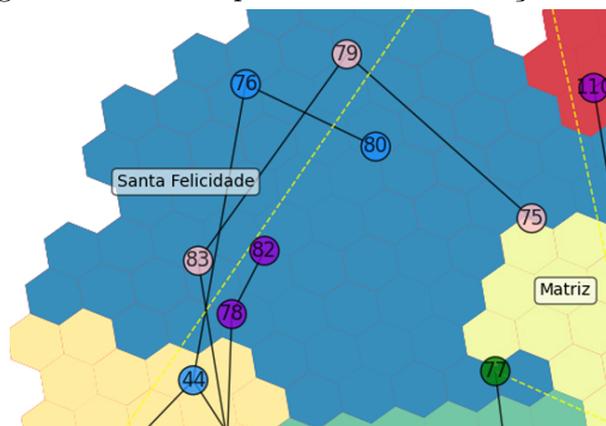
```

Input: o grafo  $G$  com todos os nós e arestas existentes
Output: o grafo  $G$  conexo
grafosdesconexos  $\leftarrow$  listaDeSubgrafos( $G$ )
while grafosdesconexos  $>$  1 do
    menorPeso  $\leftarrow$   $\infty$ 
     $G' \leftarrow$  buscaMaiorPrioridade( $G$ )
    for  $v \in G$  do
        for  $u \in G'$  do
            pesoAtual = calculoPesoAtéISP( $u, v$ )
            if pesoAtual  $<$  menorPeso then
                menorPeso  $\leftarrow$  pesoAtual
                 $u' \leftarrow u$ 
                 $v' \leftarrow v$ 
            end
        end
    end
    criaAresta( $G, u', v',$  menorPeso)
    grafosdesconexos  $\leftarrow$  listaDeSubgrafos( $G$ )
end

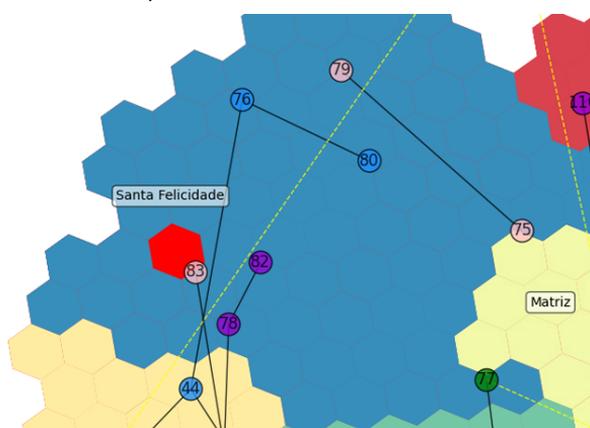
```

Portanto, o subgrafo desconectado da rede será religado a ela utilizando uma função que minimize o peso do caminho do nó até o ISP mais próximo. Esse cálculo é

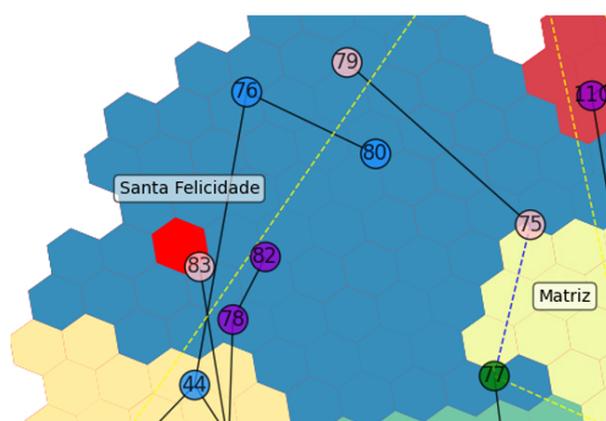
Figura 8 – Passo a passo da reestruturação da rede



1) Nó 79 conectado ao nó 83



2) Nó 79 desconectado do nó 83 devido a um outlier



3) Reestruturação realizada conectando o nó 75 ao nó 77

Fonte: Autoria Própria

realizado fazendo uma simulação de qual seria o peso do caminho do nó até o ISP se aquela aresta fosse criada, fazendo assim com que o menor peso encontrado seja selecionado.

Ao criar uma nova aresta com o nó selecionado, foi levado em consideração que esse seria um enlace provisório e, portanto, de baixo custo e com baixa capacidade de

largura de banda. Logo, utilizamos a fórmula 2, porém multiplicamos o resultado dela por 10, valor usado para simular um peso maior entre esses nós.

Na prática, para realizar essa reestruturação temos algumas opções, sendo as principais: realizar o re-roteamento da rede se for possível ou acionar algum elemento externo a rede que irá realizar a recuperação. Esses elementos externos podem ser: enviar um *drone* ao local afetado para conectar o nó a um link via satélite, criar uma conexão com outro nó adjacente utilizando uma tecnologia sem fio, etc. A implementação da reestruturação supôs uma reconexão por adjacência com os nós vizinhos.

3.4.5 Métricas para avaliar o estado da rede

Para avaliar o estado da rede após uma catástrofe e após a reestruturação da rede, foram criadas 3 métricas. Com isso, foi possível verificar se houveram melhorias ou pioras na rede após a escolha da religação dos nós, comparado com o estado inicial da rede. As seguintes funções de métricas foram implementadas: cálculo da conectividade média global dos nós, cálculo do peso total acumulado de todos os nós até os ISPs e o cálculo da porcentagem de nós desconectados do grafo. Iremos representar o grafo como sendo $G = (V, E)$, sendo que V representa o conjunto de nós do grafo e E representa os pares de arestas existentes nesse grafo.

Para o cálculo da conectividade global da rede, foi utilizada a função `nx.average_node_connectivity(G)` disponível na biblioteca `networkx`. Essa função obtém a média da conectividade local de todos os pares de nós do grafo especificado, e é calculada através da fórmula 4,

$$\bar{K}(G) = \frac{\sum_0^E * K_G(E)}{\binom{n}{2}} \quad (4)$$

em que K é o valor da conectividade local, \bar{K} é a média da conectividade local, E são todos os pares de arestas do grafo e G é o grafo propriamente dito.

Já no cálculo do peso total acumulado, o primeiro passo é somar o peso de todas as arestas com a rede em seu estado inicial e salvar esse valor. A medida que as catástrofes acontecerem e os nós forem reconectados, o peso das arestas do grafo serão alterados. Com isso, podemos utilizar a fórmula 5,

$$PesoAcumulado(t) = \frac{\sum_{u \in V} G(Peso_t(u, ISP))}{\sum_{u \in V} G(Peso_{t-1}(u, ISP))} \quad (5)$$

em que u representa um nó, V representa todos os nós do grafo, ISP é o nó destino, t representa o momento após a reestruturação e $t-1$ representa o estado inicial da rede. Com isso, obtemos a razão em que a rede ficou mais rápida ou mais lenta.

Por fim, temos a última métrica que é responsável por medir a quantidade de nós em um dado momento que estão sem conectividade com a Internet. Para isso, é utilizado

a fórmula 6,

$$RazãoNósDesconectados = \frac{G.has_edge(V,ISP)}{V} \quad (6)$$

na qual V representa o número total de nós do grafo e a função $G.has_edge(V,ISP)$ verifica quantos desses nós possuem uma aresta até o nó ISP .

4 DISCUSSÃO DOS RESULTADOS

4.1 Resultados obtidos

Para a melhor compreensão dos resultados, esta seção foi dividida em 3 partes: análise dos dados de precipitação e alagamentos, retroalimentação da plataforma ICARUS e a reestruturação da rede a partir dos *outliers*.

4.1.1 Análise dos dados de precipitação e alagamentos

A avaliação da qualidade do algoritmo de detecção é um ponto muito importante porém difícil. Para o treinamento do modelo, dispomos apenas de poucos pontos de medição de chuvas, o que é muito impreciso. Para a avaliação da ocorrência de alagamentos no passado, temos apenas os dados de registro da Defesa Civil, que apresenta problemas de precisão e subnotificação. Portanto, já esperávamos uma dificuldade no processo de avaliação, como descrito a seguir. Propostas para melhorar a qualidade dos dados e do detector são descritas na seção 5.2.

Inicialmente, o modelo de detecção de *outliers* foi executado somente com os dados de precipitação interpolados, conforme mostrado na seção 3.3.3. Porém, ao utilizar uma contaminação que fizesse com que o algoritmo retornasse um valor de *outliers* similar ao número de alagamentos, os pontos de alagamento por dia não tiveram qualquer intersecção com os pontos de *outliers* detectados. Ao agrupar esses dados por dia, obteve-se 28 dias com alagamentos e 53 dias com *outliers* em Curitiba. A intersecção entre estes dois dados foi de 3 dias, retornando um *F-score* de 0,074.

Utilizando os dados da Defesa Civil de Curitiba, foi feita a discretização dos pontos de alagamentos de 2015 em hexágonos, para que pudessem ser enviados para a simulação de rede. Com isso, todos os pontos de alagamentos de 2015 foram utilizados para simular falhas de enlace na rede. No ano de 2015, ocorreram 125 ocorrências de alagamentos dentro de Curitiba. Dentre essas 125 ocorrências de alagamentos, houve 157 intersecções de arestas com os pontos discretizados de alagamentos. Dentre essas 157 intersecções, foram removidas 62 arestas da rede. O número de remoções de arestas da rede foi menor que o número de intersecções porque existem arestas no grafo que passam por mais de um hexágono afetado ao mesmo tempo. Vale ressaltar que, dos 125 pontos de alagamentos, apenas 71 impactaram a rede.

Feito isso, os pontos que afetaram a rede foram retornados como uma nova dimensão para o modelo de detecção de *outliers* na expectativa de aumentar a precisão dos resultados desse. Vale ressaltar que a simulação também envia os dados de quedas de enlace não relacionados a alagamentos para o modelo.

4.1.2 Retroalimentação da plataforma ICARUS

Ao utilizar essa nova dimensão, o modelo de detecção de *outliers* foi retroalimentado e executado novamente de forma similar à seção anterior. Feito isso, encontrou-se 2 pontos de intersecção no mesmo dia. Ao agrupar os dados por dia, houve 56 dias de *outliers* identificados e 28 dias de alagamentos, sendo que 5 desses se interseccionaram. Com esses resultados, o *F-score* ficou igual a 0,12. É importante notar que este é um teste inicial e que demanda mais testes dessa retroalimentação para demonstrar sua efetividade, conforme apontado na seção 5.2 de trabalhos futuros.

Finalmente, esses resultados foram enviados para a simulação de infraestrutura de redes.

4.1.3 Reestruturação da rede a partir dos *outliers*

Para verificar a reestruturação da rede, foi escolhido o dia 12/04/2015 dos 56 dias de *outliers* detectados, pois esse foi o dia que mais tinha *outliers*, simulando então as quedas de enlace na infraestrutura.

No dia escolhido, foram detectadas 17 intersecções com as arestas da rede em relação aos 19 hexágonos afetados. Dentre essas 17 intersecções, foram removidas 16 arestas. Com isso a rede ficou 23,47% desconectada e com uma métrica de conectividade de 60,88%.

Ao realizar a reestruturação da rede, foram criadas 13 arestas para religar os enlaces que não possuíam conectividade com a Internet. Foram feitas apenas 13 religações pois, considerando um subgrafo que possua dois caminhos para a Internet e ambos caminhos foram afetados, houve a religação de apenas um caminho. Após a reestruturação, pôde-se observar que a conectividade foi para 101,92% e a rede ficou 6,47% mais rápida do que era originalmente.

4.2 Discussão dos resultados

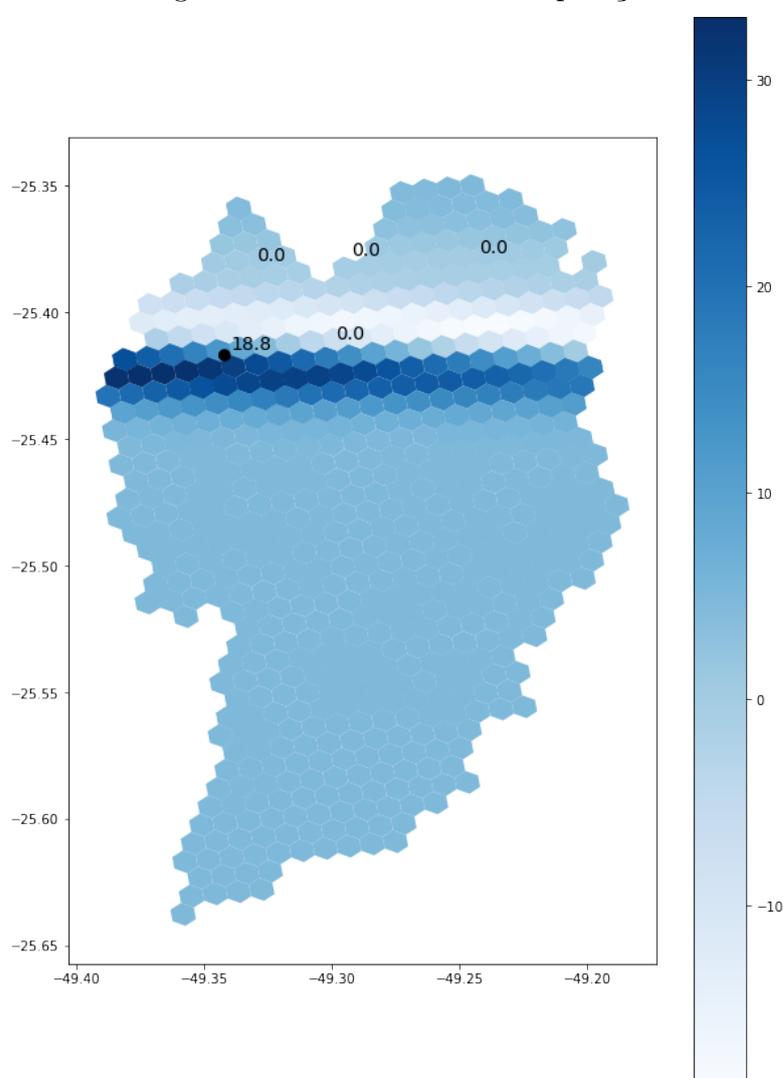
Nesta seção serão avaliados os resultados obtidos para cada subsistema dentro da plataforma ICARUS, e serão apresentados possíveis pontos fracos do sistema, juntamente com algumas mudanças para otimizar tal plataforma.

O modelo de detecção teve uma baixa acurácia na predição dos resultados esperados e há diversos pontos que justificam tal fragilidade. Ainda assim, é interessante notar que a retroalimentação de dados ao modelo melhorou o *F-score* em mais de 60%. Ainda assim, os resultados ficaram aquém do esperado, mesmo considerando que esse é um problema complexo.

Do ponto de vista de fontes de dados, houve um fator que indicava que esses dados nem sempre se comportavam da forma esperada, sendo a falta dos dados um ponto importante a se ressaltar.

A figura 9 mostra um caso exemplo de uma hora na qual só havia uma estação pluviométrica anotando dados, e com o preenchimento de dados das estações que já deveriam estar funcionando, a Krigagem desses dados retornou um valor negativo em pontos próximos a alta variação dos dados pluviométricos. A solução tomada para esse problema foi substituir os valores negativos da interpolação por 0, pois um pluviômetro não mede valores negativos. Essa solução não é ideal, pois isso compromete a confiança do algoritmo de interpolação. Além disso, esse exemplo mostra que são necessários mais estações pluviométricas funcionando de forma contínua para que seja possível que tal algoritmo consiga executar de forma esperada.

Figura 9 – Problemas de interpolação



Fonte: Autoria Própria

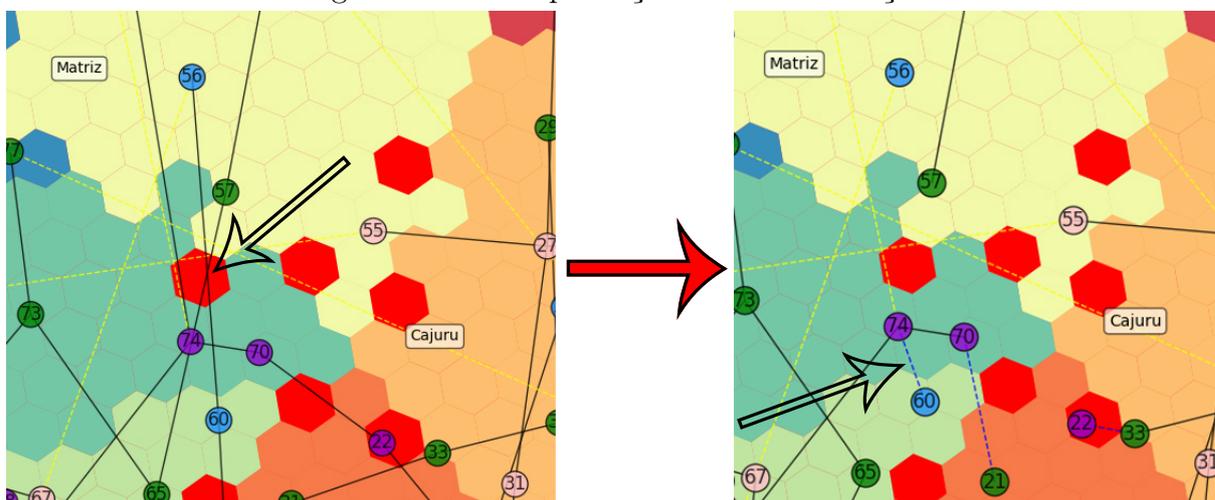
Outro problema conhecido é a falta de consistência e precisão dos dados de ocorrências de alagamentos, pois a variação de tempo em que um caso é anotado pode variar. Inclusive, houve queda consistente no número de anotação dos casos de alagamento

ao longo do tempo, sendo que isso não está de acordo com o apresentado pela Avaliação de Riscos Climáticos da Cidade de Curitiba (CURITIBA, 2021).

De qualquer forma, é possível fazer melhorias a esse modelo ao se utilizar de mais dados que possam aumentar as chances de identificar os alagamentos. Além disso, é possível utilizar métodos estatísticos para calcular os pesos de cada dimensão dentro do algoritmo LOF, pois é razoável avaliar, por exemplo, que ao normalizar o tempo de 1 ano em um espaço de 0 a 1, esse valor está com um peso muito atenuado. Vale ressaltar que o modelo pode ser completamente alterado, sendo possível usar outro algoritmo de detecção de *outliers*, ou algoritmos supervisionados, como algoritmos de classificação e regressão.

Como mencionado na seção anterior, a infraestrutura de redes ficou 6,47% mais rápida do que no seu estado original. Tal resultado pode ser explicado pelo fato de que utilizamos um fator multiplicativo de 10, conforme explicado na seção 3.4.4. Utilizando como exemplo a figura 10, podemos observar que originalmente a rede possuía uma aresta entre os nós 56 e 60 com uma alta distância. Porém, houve um evento catastrófico no caminho entre esses nós e, com isso, foi feita uma reestruturação do nó 60, religando-o ao nó 74 com uma distância muito menor do que a original, logo, com um menor peso. Vale ressaltar que a distância entre os nós foi utilizada no cálculo do peso por se tratar de um dado facilmente obtido e por não possuir dados mais complexos e elaborados a respeito do tráfego de dados na rede.

Figura 10 – Exemplificação da reestruturação



Fonte: Autoria Própria

5 CONCLUSÕES

5.1 Considerações finais

Este trabalho se propôs a realizar a reestruturação autônoma da infraestrutura de comunicação a partir da detecção de *outliers* utilizando dados de precipitação e alagamento. Para isso, foi criada a plataforma ICARUS, a qual consiste de uma colaboração entre dois sistemas computacionais – o modelo de detecção de *outliers* e a simulação da infraestrutura de redes. Como caso de uso, foi simulada a reestruturação da infraestrutura das UBS de Curitiba utilizando-se os dados de *outliers* detectados.

Referente aos dados, é essencial destacar a importância de dados abertos para que trabalhos similares a esse sejam realizados. Além disso, é indispensável que os problemas nos dados já apresentados na seção 4.2 de discussão de resultados sejam mitigados, para que o retorno científico de tais trabalhos possam ser ainda maior.

Referente ao modelo de detecção de *outliers*, a correlação entre os *outliers* e os alagamentos ficaram aquém da expectativa. Isso se justifica por ser um problema de alta complexidade e dependente de mais variáveis que o abordado neste trabalho de conclusão de curso. De qualquer forma, é um ponto inicial importante que possibilita o entendimento de outras abordagens para esse problema. Além disso, indicou que um sistema com retroalimentação nessa situação terá um desempenho otimizado a longo prazo.

Em relação a infraestrutura de rede, foi possível cumprir sua função de realizar a reestruturação autônoma da rede, tendo como base os *outliers* detectados pelo modelo. Vale ressaltar que a reestruturação de redes ainda seria válida se tivessem sido utilizados dados de outras entidades para alimentar a simulação do mapa de Curitiba. Pode-se notar também que o meio físico utilizado para realizar essa reestruturação não foi explorado.

Com isso, podemos concluir, com as ressalvas apontadas, que a plataforma ICARUS cumpriu seu objetivo de manter a conectividade das UBS de Curitiba em momentos de catástrofes a partir dos dados de precipitação e de alagamentos.

5.2 Trabalhos futuros

Foram identificadas diversas oportunidades de trabalhos tendo como base este trabalho de conclusão de curso. Esta seção apresenta sugestões dos itens que foram considerados mais relevantes.

5.2.1 Bases de dados adicionais

Devido à complexidade do tema, existem outros fatores que não foram considerados no modelo de detecção de *outliers*, sendo possível buscar por bases de dados que contenham

esses outros fatores. Dentre essas bases, podemos citar:

- o mapa de riscos do IPPUC (Instituto de Pesquisa e Planejamento Urbano de Curitiba), que avalia as áreas com riscos acentuados de desastres naturais;
- o modelo de riscos de inundação do plano diretor de drenagem do instituto Águas Paraná que possui informações sobre as áreas de Curitiba que estão mais suscetíveis a inundações e alagamentos;
- dados de altimetria da cidade de Curitiba.

Além de utilizar essas bases, também poderíamos acrescentar no modelo de detecção de *outliers* colunas referentes a temperatura, a previsão do tempo e a saturação do solo, sendo essa a média de precipitação acumulada em um certo número de dias anteriores. Com essas informações, poderíamos enviar antecipadamente as informações dos enlaces que possuem altas chances de sofrer uma queda para a infraestrutura e com isso seria possível deixar a rede mais robusta, não sendo necessário tomar medidas somente após a queda do enlace.

Para melhorar a simulação de topologia de rede, é possível considerar variáveis sociais e ambientais, como os dados disponibilizados pelo IBGE por setor censitário, como dados de criticidade e fragilidade nessas regiões.

Por fim, como os dados obtidos da cidade de Curitiba referentes a acusações de alagamentos não são precisos, seria possível usar os dados de outras cidades e municípios que possuam tais dados de maneira mais consistente e coesa. Com isso, seria possível identificar de uma forma mais assertiva os casos de inundações e melhorar a confiança da validação se os *outliers* encontrados fazem parte ou não dos casos reais.

5.2.2 Variações do modelo de detecção de *outliers*

O algoritmo de LOF escolhido neste trabalho tem um funcionamento que depende fortemente dos valores de distância entre os números de cada dimensão. Isso fez, por exemplo, com que as datas de um ano fossem comprimidas em um alcance de 0 a 1 no processo de normalização. Por esse motivo, seria possível utilizar diferentes métodos para calcular os pesos que cada uma dessas dimensões deveriam ter para melhorar a acurácia do modelo. Também, seria interessante avaliar a qualidade do modelo utilizando de diferentes taxas de contaminação.

Outra possibilidade seria utilizar de outros modelos de predição que não seja baseado exclusivamente em anomalia, como por exemplo: usar um modelo de regressão logística multivariada ou um modelo de classificação usando os dados de alagamentos.

É interessante notar que é possível que o modelo de identificação de alagamentos tenha sua acurácia aumentada usando esses e outros métodos de aprendizado de máquina que se baseiam em formato supervisionado dos dados. Contudo, isso pode levar a uma predição enviesada. Com isso, seria interessante também fazer uma comparação desses diferentes modelos e os ganhos e perdas entre eles.

5.2.3 Retroalimentação contínua dos modelos

A etapa de retroalimentação do modelo de detecção de *outliers* com a simulação da infraestrutura de redes deu-se apenas parcialmente neste trabalho de conclusão de curso. É possível alterar a plataforma ICARUS para que execute de maneira contínua no tempo o monitoramento dos dados meteorológicos e do status da rede, sendo possível haver uma retroalimentação entre ambos modelos de maneira assídua.

Com essa retroalimentação seria possível para o modelo de detecção de *outliers* aumentar a acurácia cada vez mais em relação ao tempo e alimentar os DFs com informações adicionais providas pela rede. Já para a simulação da infraestrutura de redes, seria possível intervir antecipadamente nos enlaces existentes para, por exemplo, re-rotear previamente os dados antes de uma catástrofe natural, e também criar um *ranking* dos enlaces mais propensos a estarem em uma condição de falha e prover soluções possivelmente definitivas a essas regiões da rede.

5.2.4 Integração da plataforma ICARUS com sistemas automatizados

É possível integrar a reestruturação sugerida da rede, feita pela plataforma ICARUS, com sistemas automatizados, por exemplo enxames de drones ou sistemas gerenciadores de conexão via satélite. A partir disso, quando a plataforma ICARUS detectar uma queda de enlace da rede, poderá fazer uma solicitação para o sistema mais apropriado, dependendo da região e da quantidade de nós afetados, para religar esses nós à Internet.

Além disso, é possível implementar uma hierarquização da necessidade de reconexão em casos de quedas. Junto a isso, disponibilizar às UBSs a alteração de sua prioridade de reconexão em relação às outras.

5.2.5 Outras aplicações

Por fim, é possível elencar outras aplicações nas quais seria possível aplicar este trabalho de conclusão de curso, sendo o modelo de detecção de *outliers* ou o modelo de reestruturação da rede. Dentre eles, pode-se destacar:

- Re-roteamento do trânsito;
- Logística, *e.g.*: *delivery* de comidas;
- Utilização de dados de terremotos, maremotos, deslizamentos, acidente de trânsito, entre outros;
- Utilização do modelo em outras localidades geográficas;
- Infraestrutura da rede elétrica.

REFERÊNCIAS

- BREUNIG, M. M. et al. Lof: Identifying density-based local outliers. **SIGMOD Rec.**, Association for Computing Machinery, New York, NY, USA, v. 29, n. 2, p. 93–104, may 2000. ISSN 0163-5808. Disponível em: <<https://doi.org/10.1145/335191.335388>>. Citado na página 24.
- BUFFON, E. A. M.; PAZ, O. L. d. S. da. Avaliação das áreas com risco de inundações em curitiba, paraná brasil, a partir do cruzamento dos mapeamentos de cobertura da terra e de vulnerabilidade social. **ANAIS DO SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO**, p. 5552–5559, 2017. Disponível em: <<http://urlib.net/rep/8JMKD3MGP6W34M/3PSMBAH>>. Citado 2 vezes nas páginas 19 e 23.
- BUFFON, E. A. M.; SOUSA, M. S. de. Proposta metodológica para avaliação dos registros secundários de alagamentos: Uma abordagem a partir de curitiba-paraná, brasil. **Caminhos de Geografia**, v. 19, n. 67, p. 188–204, 2018. ISSN 1678-6343. Citado 2 vezes nas páginas 19 e 20.
- CANDELA, M.; LUCONI, V.; VECCHIO, A. Impact of the covid-19 pandemic on the internet latency: A large-scale study. **Computer Networks**, v. 182, p. 107495, 2020. ISSN 1389-1286. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1389128620311622>>. Citado na página 14.
- CEMADEN, C. N. de Monitoramento e Alerta de D. N. **Pluviômetros Automáticos – Cemaden**. 2021. Disponível em: <<http://www2.cemaden.gov.br/pluviometros-automatico/>>. Acesso em: 20 de abril de 2021. Citado 2 vezes nas páginas 20 e 21.
- CETIC.BR/NIC.BR/CGI.BR. **PAINEL TIC COVID-19**. 2021. Disponível em: <https://cetic.br/media/docs/publicacoes/2/20210426095323/painel_tic_covid19_livro_eletronico.pdf>. Acesso em: 10 novembro 2021. Citado na página 14.
- CHANDOLA, V.; BANERJEE, A.; KUMAR, V. Anomaly detection: A survey. Association for Computing Machinery, New York, NY, USA, v. 41, n. 3, jul. 2009. ISSN 0360-0300. Disponível em: <<https://doi.org/10.1145/1541880.1541882>>. Citado na página 16.
- CORDOVA, A. et al. **Recovery Plan for the Communications and Information Technology Sector After Hurricanes Irma and Maria: Laying the Foundation for the Digital Transformation of Puerto Rico**. Santa Monica, CA: RAND Corporation, 2020. Citado na página 14.
- CURITIBA, P. de Curitiba; Secretaria Municipal de Meio Ambiente; Instituto de Pesquisa e Planejamento Urbano de C. **Avaliação de Riscos Climáticos na Cidade de Curitiba**. 2021. Disponível em: <<https://mid.curitiba.pr.gov.br/2020/00305799.pdf>>. Acesso em: 25 de novembro de 2021. Citado 2 vezes nas páginas 16 e 39.
- HAZRA, K. et al. Designing efficient communication infrastructure in post-disaster situations with limited availability of network resources. **Computer Communications**, Elsevier, v. 164, p. 54–68, 2020. Citado na página 18.

- KASZNAR, A. P. P. et al. Multiple dimensions of smart cities' infrastructure: A review. **Buildings**, v. 11, n. 2, 2021. ISSN 2075-5309. Disponível em: <<https://www.mdpi.com/2075-5309/11/2/73>>. Citado na página 18.
- KULLDORFF, M. A spatial scan statistic. **Communications in Statistics-Theory and methods**, Taylor & Francis, v. 26, n. 6, p. 1481–1496, 1997. Citado na página 19.
- LOHMANN, M. **Regressão logística e redes neurais aplicadas à previsão probabilística de alagamentos no Município de Curitiba, Pr.** Tese (Doutorado) — Federal University of Paraná, Curitiba, Brazil, 2013. Citado na página 21.
- MADRID, L. **O que é uma cidade inteligente?** 2021. Disponível em: <<https://www.scbamerica.com/o-que-e-uma-cidade-inteligente>>. Acesso em: 26 novembro 2021. Citado na página 18.
- MARÍN, J.; CREUS, R. G.; ZABALLOS, A. G. **Relatorio da situação da conectividade à internet e banda larga no Brasil.** [S.l.], 2013. Citado na página 14.
- MATISZIW, T. C.; MURRAY, A. T. Modeling s–t path availability to support disaster vulnerability assessment of network infrastructure. **Computers & Operations Research**, Elsevier, v. 36, n. 1, p. 16–26, 2009. Citado na página 15.
- NATALINO, C.; ROSTAMI, A.; MONTI, P. Storage protection with connectivity and processing restoration for survivable cloud services. In: **2021 International Conference on Computer Communications and Networks (ICCCN)**. [S.l.: s.n.], 2021. p. 1–9. Citado na página 18.
- NIC.BR. **Pesquisa do Cetic.br aponta que 69% das Unidades Básicas de Saúde possuem um sistema eletrônico para registro das informações dos pacientes.** 2021. Disponível em: <<https://nic.br/noticia/releases/pesquisa-do-cetic-br-aponta-que-69-das-unidades-basicas-de-saude-possuem-um-sistema-eletronico-para-registro-das-informacoes-dos-pacientes>>. Acesso em: 26 novembro 2021. Citado na página 14.
- NORONHA, G. **Enchentes – O que são, características, causas e impacto urbano.** 2021. Disponível em: <<https://conhecimentocientifico.com/enchentes/>>. Acesso em: 26 novembro 2021. Citado na página 16.
- TABOADA, N. G.; LEGAL, E. J.; MACHADO, N. Resiliência: em busca de um conceito. **Journal of Human Growth and Development**, scielopepsic, v. 16, p. 104 – 113, 12 2006. ISSN 0104-1282. Disponível em: <http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S0104-12822006000300012&nrm=iso>. Citado na página 17.
- VASCONCELOS, M. F. de. Emprego de resiliência na gerência de redes de computadores. 2012. Citado na página 17.
- WU, E.; LIU, W.; CHAWLA, S. Spatio-temporal outlier detection in precipitation data. In: GABER, M. M. et al. (Ed.). **Knowledge Discovery from Sensor Data**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010. p. 115–133. ISBN 978-3-642-12519-5. Citado na página 19.
- YAMAMOTO, J. K.; LANDIM, P. M. B. **Geoestatística: conceitos e aplicações.** [S.l.]: Oficina de Textos, 2013. Citado na página 17.