



UNIVERSIDADE ESTADUAL DO CEARÁ
CENTRO DE CIÊNCIAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO
MESTRADO PROFISSIONAL EM COMPUTAÇÃO APLICADA

BRUNO BEZERRA CHAVES

MÉTODOS COMBINATORIAIS PARA PROBLEMAS EM REDES DINÂMICAS:
ALGORITMOS DE AGRUPAMENTO E PREVISÃO DINÂMICOS

FORTALEZA – CEARÁ

2018

BRUNO BEZERRA CHAVES

MÉTODOS COMBINATORIAIS PARA PROBLEMAS EM REDES DINÂMICAS:
ALGORITMOS DE AGRUPAMENTO E PREVISÃO DINÂMICOS

Dissertação apresentada ao Curso de Mestrado Profissional em Computação Aplicada do Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências e Tecnologia da Universidade Estadual do Ceará, como requisito parcial à obtenção do título de mestre em Ciência da Computação. Área de Concentração: Ciência da Computação

Orientador: Prof. Dr. Marcos José Negreiros Gomes

FORTALEZA – CEARÁ

2018

RESUMO

Este trabalho apresenta uma abordagem para solucionar o problema da previsão de agrupamentos dinâmicos espaço-temporal. Foram implementados 2 métodos para visualização dos grupos formados. O primeiro é uma biblioteca que cria e gerencia grupos de acordo com o nível de zoom. O segundo, chamado algoritmo Convex Hull, consiste em gerar o menor polígono que englobe um determinado conjunto de pontos. O algoritmo ST-DBSCAN foi implementado como base para alcançar o método proposto. Por último, foram desenvolvidos 2 aprimoramentos no software DYNAGRAPH para possibilitar a avaliação dos métodos. Um deles foi a integração com um Editor de Características, que permite alterar os atributos visuais dos vértices e arestas de um grafo dinâmico. Outro novo recurso permitiu a visualização da formação de novos grupos dinâmicos.

Palavras-chave: Agrupamento Dinâmicos, Grafos dinâmicos.

ABSTRACT

TODO abstract **Keywords:**

LISTA DE ALGORITMOS

SUMÁRIO

1	INTRODUÇÃO	6
1.1	OBJETIVOS	7
1.1.1	Objetivo Geral	7
1.1.2	Objetivos Específicos	7
1.2	HIPÓTESES	7
1.3	JUSTIFICATIVA	8
1.4	METODOLOGIA	8
1.4.1	Etapas metodológicas do projeto	8
1.4.1.1	Revisão da literatura e soluções existentes	8
1.4.1.2	Análise de requisitos	9
1.4.1.3	Arquitetura do software	9
1.4.1.4	Modelo e desenvolvimento de software	9
1.4.1.5	Ferramentas e Materiais	9
1.5	ORGANIZAÇÃO DO TEXTO	10
2	CONCEITOS E REVISÃO BIBLIOGRÁFICA	11
2.1	AGRUPAMENTOS	11
2.2	AGRUPAMENTO POR DENSIDADE	12
2.2.1	Método DBSCAN	12
2.2.2	Método ST-DBSCAN	12
2.3	REDES DINÂMICAS	12
2.3.1	O modelo Dynagraph	12
2.3.2	Editor de características	12
2.4	TRABALHOS RELACIONADOS	12
3	FUNDAMENTAÇÃO TEÓRICA	14
4	AVALIAÇÃO DO MÉTODO PROPOSTO	18
5	CONCLUSÕES E TRABALHOS FUTUROS	19
5.1	CONSIDERAÇÕES FINAIS	19
5.2	LIMITAÇÕES	19
5.3	TRABALHOS FUTUROS	19
	REFERÊNCIAS	20

1 INTRODUÇÃO

Grandes quantidades de dados estão disponíveis para análise em organizações hoje em dia. Estas enfrentam vários desafios quando se tenta analisar dados gerados com o objetivo de extrair informações úteis. Esta capacidade analítica precisa ser reforçada com ferramentas capazes de lidar com grandes conjuntos de dados sem tornar o processo de análise uma tarefa árdua. Agrupamento de dados normalmente são usados no processo de análise de dados, pois esta técnica não exige qualquer conhecimento prévio dos dados. Contudo, os algoritmos de agrupamento geralmente requerem um ou mais parâmetros de entrada que influenciam o processo de agrupamento e os resultados que podem ser obtidos.

Nos últimos anos, o problema de agrupamento dinâmico tem atraído o interesse de pesquisas, impulsionado pelo aumento da disponibilidade de grandes conjuntos de dados contendo elementos espaciais e temporais. Este problema pode ser analisado como um problema de otimização. Seu objetivo principal é maximizar as diferenças das características dos indivíduos de grupos distintos, e minimizar as diferenças das características dos indivíduos de um mesmo grupo.

Agrupamento de dados ganhou uso muito difundido, especialmente para dados estáticos. No entanto, o rápido crescimento de dados espaço-temporais de inúmeros instrumentos, como os satélites em órbita terrestre, criou uma necessidade de métodos de agrupamento espaço-temporais para extrair e monitorar clusters dinâmicos. O agrupamento espaço-temporal dinâmico enfrenta dois grandes desafios: primeiro, os clusters são dinâmicos e podem mudar de tamanho, forma e propriedades estatísticas ao longo do tempo. Em segundo lugar, vários dados espaço-temporais são incompletos, ruidosos, heterogêneos e altamente variáveis sobre espaço e tempo.

O problema de agrupamento dinâmico com a componente de previsão divide-se em passos. A primeira etapa é obtenção das informações espaço-temporais mapeáveis e características do indivíduo. Neste passo, segue-se três estratégias para resolução do problema: os dados são analisados como um só grupo (Agrupamento Estático); trata-se os dados por intervalos pré-definidos; e mapeamento das evoluções entre intervalos observados. Sendo assim, pretende-se indicar o conjunto de grupos espacialmente correlacionados também no tempo.

Já o problema de previsão de grupos dinâmicos introduz o conceito de indicar os possíveis grupos que serão formados no tempo após um conjunto de eventos serem observados previamente.

O algoritmo proposto para previsão de agrupamentos é uma importante contribuição

deste trabalho, uma vez que poderá ser usado na obtenção de informações, na previsão de movimentação dos grupos e recomendação para o combate a endemias. O serviço proposto se baseia na localização passada dos casos de dengue.

1.1 OBJETIVOS

A seguir, são expostos os objetivos desta dissertação, definindo o produto final a ser obtido.

1.1.1 Objetivo Geral

Estudo e aplicação de métodos existentes e proposta de um método para resolver o Problema de Agrupamento em Grafos Dinâmicos e previsão de evolução destes agrupamentos.

1.1.2 Objetivos Específicos

Para que se alcance o objetivo geral, as seguintes metas foram estabelecidas:

- a) Utilizar o software Dynagraph como ambiente de suporte à visualização e interação com os resultados dos métodos de agrupamento espaço-temporal utilizados.
- b) Extração de características de previsão espaço-temporal sobre a evolução dos agrupamentos dinâmicos.
- c) Avaliação dos resultados sobre bases de dados reais ligadas a evolução de casos de Dengue e Chikungunya e outras bases dinâmicas.

1.2 HIPÓTESES

As hipóteses a seguir conduziram a elaboração desta dissertação:

- a) É possível a criação de um algoritmo capaz de sugerir novos agrupamentos geolocalizados baseados no tempo.
- b) É exequível a integração de um editor de características ao DYNAGRAPH, que é um software extensível.
- c) É realizável a utilização do modelo proposto de agrupamentos em grafos dinâmicos em um ambiente Web.

1.3 JUSTIFICATIVA

Esta pesquisa justifica-se por perceber-se a necessidade de ferramentas e estudos relacionando os assuntos abordados: agrupamento, previsão em dados dinâmicos espaço-temporais, grafos dinâmicos e sistemas web de forma integrada. E também, acelerar técnicas de agrupamento em grafos dinâmicos para tomada de decisão.

A relevância da pesquisa está em permitir uma análise dos dados extraídos para apoio à tomada de decisão, onde concentra-se na avaliação dos resultados sobre bases de dados dinâmicas relativas a casos de Dengue. A pesquisa toma como base as características de evolução dos casos da doença observados entre 2012 e 2018 em Fortaleza. Os dados foram tomados a partir de (SIMDA, 2018), onde um estado é definido como o período de uma semana.

1.4 METODOLOGIA

A seguir, são descritas as etapas metodológicas para o desenvolvimento da dissertação de acordo com o cronograma.

1.4.1 Etapas metodológicas do projeto

A presente pesquisa pode ser caracterizada quanto ao procedimento quantitativa e comparativa. A pesquisa quantitativa prioriza apontar numericamente a frequência e a intensidade dos comportamentos dos indivíduos de um determinado grupo, ou população. O método comparativo constitui-se em investigar coisas ou fatos e explicá-los de acordo com suas semelhanças e suas diferenças. Possibilita a análise de dados concretos e a dedução de semelhanças e divergências de elementos contínuos, abstratos e gerais, facilitando investigações de caráter indireto (FACHIN, 2001).

1.4.1.1 Revisão da literatura e soluções existentes

As fontes principais de pesquisa foram sites especializados em pesquisas científicas, por exemplo, o portal de periódicos da CAPES, IEEE e outros sites de referências em que possuem livros, periódicos e dissertações disponíveis. Os temas essenciais abordados na pesquisa foram:

- Estrutura de dados em grafos dinâmicos
- Modelos de previsão espaço-temporais

- Algoritmos de agrupamentos dinâmicos

1.4.1.2 Análise de requisitos

Detectou-se as necessidades de informações baseadas na previsão de agrupamentos dinâmicos em grafos. Assim sendo, e após obter os dados a partir de (SIMDA, 2018), houve a necessidade de um software para representação e tratamentos de grafos dinâmicos. Com isso foi escolhido o DYNAGRAPH, que tem como característica a extensibilidade.

1.4.1.3 Arquitetura do software

A arquitetura do software desenvolvido é apresentada, assim como o diagrama de caso de uso (UML).

1.4.1.4 Modelo e desenvolvimento de software

Foi desenvolvido um modelo capaz de representar agrupamentos e previsão dinâmicos. Em seguida, foi desenvolvido a partir do software DYNAGRAPH executar o modelo proposto e apresentar os resultados para validar a aplicação da ferramenta.

1.4.1.5 Ferramentas e Materiais

O desenvolvimento do trabalho foi realizado a partir dos seguintes equipamentos e materiais:

- Macbook Pro 13" modelo 2015 / macOS High Sierra 10.13.2:
Processador Intel Core i5 2.7 GHz;
Memória de 8GB 1867 MHzs DDR3;
HD SSD 128GB;
- Ambiente de desenvolvimento Webstorm;
- Navegador de internet Google Chrome 63.0.3239.132.
- Controle de versão Git (Software e dissertação);
- \LaTeX para produção da dissertação;
- Portal de periódicos CAPES;

1.5 ORGANIZAÇÃO DO TEXTO

Esta dissertação está organizada em 5 capítulos. O capítulo 1 apresenta uma introdução à necessidade da representação e manipulação do agrupamento espaço-temporal dinâmico, assim como a previsão da formação de novos grupos dinâmicos. Em seguida são apresentados os objetivos, hipóteses, a metodologia utilizada e contribuições. O capítulo 2 apresenta a revisão bibliográfica em modelagem com grafos dinâmicos, métodos de agrupamento por densidade, redes dinâmicas, o DYNAGRAPH, um editor de características e um conjunto de trabalhos relacionados a esta área do conhecimento. O capítulo 3 apresenta o modelo de agrupamento e previsão em redes dinâmicas. O capítulo 4 apresenta os resultados e comparação dos algoritmos apresentados. O capítulo 5 apresenta as considerações finais e propostas de trabalhos futuros.

2 CONCEITOS E REVISÃO BIBLIOGRÁFICA

2.1 AGRUPAMENTOS

Os dados espaço-temporais estão rapidamente se tornando onipresentes graças a sensores e armazenamento acessíveis. Estes dados ricos em informações têm o potencial de revolucionar diversos campos, como as ciências sociais, terrestres e médicas, onde há necessidade de extrair e compreender fenômenos espaço-temporais complexos e suas dinâmicas. Além disso, os dados em tais domínios científicos tendem a ser grandes e não marcados. Isso destaca a importância de métodos não supervisionados no monitoramento de dinâmicas espaço-temporais com pouca ou nenhuma supervisão humana.

Clustering é uma das técnicas de mineração de dados não supervisionadas mais comuns. Tem um enorme sucesso, especialmente para dados estáticos. No entanto, há pouco trabalho na configuração espaço-temporal onde os dados estão na forma de campos espaço-temporais contínuos e os clusters são dinâmicos. Além disso, os dados espaço-temporais originados por satélites em órbita terrestre, telefones celulares e outros sensores tendem a ser ruidosos, incompletos e heterogêneos, tornando sua análise especialmente desafiadora. (FAGHMOUS; KUMAR, 2013) [Faghmous and Kumar, 2013] James H Faghmous and Vipin Kumar. Spatio-temporal data mining for climate data: Advances, challenges, and opportunities. In W. Chu, editor, *Data Mining and Knowledge Discovery for Big Data: Methodologies, Challenges, and Opportunities*, pages 83–116. Springer, 2013.

Neste artigo, propomos um novo paradigma de agrupamento espaço-temporal para identificar clusters em um campo espaço-temporal contínuo onde os clusters são dinâmicos e podem mudar seu tamanho, forma, localização e propriedades estatísticas de um único passo para o próximo. O nosso paradigma decorre da observação de que, em inúmeras configurações dinâmicas, embora os agrupamentos possam se mover ou mudar de forma, existem vários pontos que não altere as associações de grupos para um período de tempo significativo. Esta observação nos permite extrair de forma autônoma clusters dinâmicos em dados espaço-temporais contínuos que podem conter valores, ruídos ou características muito variáveis.

2.2 AGRUPAMENTO POR DENSIDADE

2.2.1 Método DBSCAN

2.2.2 Método ST-DBSCAN

2.3 REDES DINÂMICAS

2.3.1 O modelo Dynagraph

2.3.2 Editor de características

2.4 TRABALHOS RELACIONADOS

Como há uma carência de estudos relacionando os assuntos abordados: agrupamento, previsão em dados dinâmicos espaço-temporais, grafos dinâmicos e sistemas web de forma integrada, foi necessário dividir o problema de agrupamentos e previsões dinâmicos em três etapas:

- Estrutura de dados em grafos dinâmicos
- Modelos de previsão espaço-temporais
- Algoritmos de agrupamentos dinâmicos

A pesquisa aborda estrutura de dados em grafos dinâmicos usando passos já descritos na literatura, principalmente o modelo Dynagraph (CALIXTO; NEGREIROS, 2013), que é baseado na primeira proposta em (CALIXTO; NEGREIROS, 2012), onde o Dynagraph usa sequências temporais para vértices, arestas, características modificáveis dos vértices e arestas e o relacionamento entre suas características. Com isso, é formado um grafo com as informações necessárias para qualquer instante no tempo. O Dynagraph é capaz de visualizar o comportamento do grafo ao longo de um período de tempo, e editá-lo.

A ideia central de (KIM; ANDERSON, 2012) é modelar uma rede dinâmica como digrafos orientados ao tempo (*time-ordered graph*), que é gerada através da ligação de instantes temporais com arestas direcionadas que unem cada nó ao seu sucessor no tempo. Com isso, transformar uma rede dinâmica em um grafo maior, mas facilmente analisável. Isto permite não só a utilização dos algoritmos desenvolvidos para grafos estáticos, mas também para melhor definir métricas para grafos dinâmicos. Segundo (KIM; ANDERSON, 2012) um sistema de grafos dinâmicos é um objeto de representação visual que pode descrever melhor o comportamento

dinâmico de objetos relacionados a eventos dinâmicos e introduzir novas formas de enxergar ou descrever a evolução de eventos dinâmicos na natureza.

(KOSTAKOS, 2009) considera a estrutura de grafos temporais como grafos estáticos, no entanto avança sobre as métricas introduzindo conceitos como disponibilidade temporal, proximidade temporal e geodésica, e estuda os seus grafos sobre redes reais.

Segundo (ESTER *et al.*, 1996), o algoritmo DBScan(*Density-Based Spatial Clustering of Applications With Noise*) calcula a densidade de uma região contando quantos pontos existem em uma determinada área seguindo uma determinada métrica. Ele permite a redução de pontos não pertencentes a nenhum padrão, assim como possibilita a formação de grupos de diferentes formas. Seu objetivo principal é dividir os pontos em grupos através da densidade de cada região.

(LAHIRI; BERGER-WOLF, 2007) apresentam um algoritmo de predição em redes temporais, e que usa a ideia de que certas interações sinalizam a ocorrência de outros em algum momento no futuro. Através de análises estatísticas o algoritmo mede o atraso entre as interações, e com isso pode-se prever quando certas interações vão ocorrer com base em observações passadas e atuais. Propõe-se a utilização de subgrafos frequentes e discute como identificar subgrafos que são persistidos em redes temporais. (LAHIRI; BERGER-WOLF, 2008) em seguida propõe um novo problema de mineração de dados para redes dinâmicas: detecção de todos os padrões de interação que ocorrem em intervalos de tempo regulares.

3 FUNDAMENTAÇÃO TEÓRICA

Referência para DBSCAN Ester, M., Kriegel, H.-P., Sander, J., and Xu X. (1996), A density-based algorithm for discovering clusters in large spatial data sets with noise. Proc. 2nd Int. Conf. on Knowledge Discovery and Data Mining. Portland, OR, pp.226–231.

Os algoritmos de agrupamento baseados em densidade tentam encontrar clusters com base na densidade de pontos de dados em uma região. A idéia-chave do agrupamento baseado em densidade é que, para cada instância de um cluster, a vizinhança de um determinado raio (Eps) deve conter pelo menos um número mínimo de instâncias (MinPts). Um dos mais conhecidos algoritmos de cluster baseados em densidade é o DBSCAN. DBSCAN separa pontos de dados em três classes (Fig. 2):

- Pontos principais. Estes são pontos que estão no interior de um cluster. Um ponto é um ponto interior se houver pontos suficientes em sua vizinhança.
- Pontos de fronteira. Um ponto de fronteira é um ponto que não é um ponto central, ou seja, não há pontos suficientes em sua vizinhança, mas ele está dentro da vizinhança de um ponto central.
- Pontos de ruído. Um ponto de ruído é qualquer ponto que não é um ponto central ou um ponto de fronteira.

Para encontrar um cluster, o DBSCAN começa com uma instância arbitrária (p) no conjunto de dados (D) e recupera todas as instâncias de D em relação a Eps e MinPts

A Survey on Density Based Clustering Algorithms for Mining Large Spatial Databases

O algoritmo de clustering baseado em densidade é um dos métodos principais para o agrupamento em mineração de dados. Os clusters que são formados com base na densidade são fáceis de entender e não se limitam às formas dos clusters. Este artigo fornece uma pesquisa detalhada dos algoritmos baseados em densidade existente, nomeadamente DBSCAN, VDBSCAN, DVBSAN, ST-DBSCAN e DBCLASD, com base nos parâmetros essenciais necessários para um bom algoritmo de agrupamento. Analisamos os algoritmos em termos dos parâmetros essenciais para a criação de clusters significativos.

Agrupamentos(introdução) A mineração de dados é um passo no processo de descoberta de conhecimento em bancos de dados (KDD) consistindo na aplicação de análise de dados e na descoberta de algoritmos que, em limitações de eficiência computacional, produzem uma enumeração particular dos dados [2]. O Gerenciamento de banco de dados espacial (SDBS) [3] são sistemas de banco de dados para o gerenciamento de dados espaciais, ou seja, objetos pontuais ou objetos espacialmente estendidos em um espaço 2D ou 3D ou algum espaço dimen-

sional elevado. A descoberta do conhecimento torna-se cada vez mais importante em termos espaciais, uma vez que quantidades crescentes de dados obtidos a partir de imagens de satélite ou outros equipamentos automáticos são armazenados em bases de dados espaciais. Spatial Data Mining [4] é o processo de encontrar padrões interessantes e anteriormente desconhecidos, mas potencialmente úteis, de grandes conjuntos de dados espaciais. Extrair padrões de conjuntos de dados espaciais interessantes e úteis são mais difíceis do que extrair o padrão correspondente de dados numéricos e categóricos devido à complexidade de relações espaciais, espaciais e autocorrelação espacial. Há um crescimento desenfreado de dados espaciais e uma série de necessidades surgem como dados espaciais técnicas de mineração, modelando propriedades espaciais semânticas ricas, como topologia, modelos de interpretação estatística para padrões espaciais, melhorando a eficiência computacional e modelo, pré-processamento de dados espaciais e muitos outros. Existem muitas técnicas, como classificação, árvore de decisão, lógica difusa, redes neurais aplicadas para mineração de dados espaciais. A maioria dos trabalhos recentes sobre dados espaciais tem técnicas de agrupamento devido à natureza dos dados. Clustering, isto é, agrupando os objetos de um banco de dados em subclasses significativas, é um dos principais métodos de mineração de dados [6]. Entre muitos tipos de algoritmos de densidade baseados em cluster. O algoritmo é mais eficiente na detecção de clusters com densidade variada. Houve muita pesquisa sobre algoritmos de agrupamento por décadas, contudo a aplicação a grandes bancos de dados espaciais apresenta os seguintes requisitos: (i) Número mínimo de parâmetros de entrada. Porque, para grandes bancos de dados espaciais, é muito difícil identificar antecipadamente os parâmetros iniciais, como o número de clusters, forma e densidade. (ii) Descoberta de clusters com forma arbitrária. Porque a forma dos clusters pode estar em qualquer forma aleatória. (iii) Uma boa eficiência deve ser alcançada em bancos de dados muito grandes.

Density-Based Algorithms for Discovering Clusters in Large Spatial Databases with Noise (DBSCAN) DBSCAN [1] é um algoritmo baseado em densidade que descobre clusters com forma arbitrária e com um número mínimo de parâmetros de entrada. Os parâmetros de entrada necessários para este algoritmo são o raio do cluster (Eps) e os pontos mínimos necessários dentro do cluster (Minpts).

2.2. Descrição do Algoritmo Nesta seção, o algoritmo DBSCAN [7] Clustering espacial baseado em densidade de aplicativos com ruído é projetado para descobrir os clusters de dados espaciais com ruído. As etapas envolvidas neste algoritmo são as seguintes, ... (i) Selecione um ponto arbitrário p (ii) Recuperar todos os pontos de densidade-reachable de p w.r.t. Eps e Minpts. (iii) Se p é um ponto central, um cluster é formado. (iv) Se p é um ponto de borda,

nenhum ponto é densidade acessível de p e DBSCAN visita o próximo ponto do banco de dados.
 (v) Continue o processo até que todos os pontos tenham sido processados.

2.3 Impacto do Algoritmo DBSCAN requer dois parâmetros de entrada (pontos mínimos e raio) e suporta o usuário ao encontrar um valor aproximado para ele usando o gráfico k -dist [7]. Ele descobre grupos de forma arbitrária. Ele é válido para grandes bancos de dados espaciais. ...

2.4 Trabalho futuro O algoritmo DBSCAN aqui considera [1] apenas objetos de ponto, mas pode ser estendido para outros objetos espaciais, como polígonos. As aplicações do DBSCAN para espaços de recursos de alta dimensão devem ser investigadas e a geração de raio para esses dados de alta dimensão também precisa ser explorada. Também não consegue detectar agrupamentos com densidade variada.

Spatial- Temporal Density Based Clustering (ST-DBSCAN)

6.1. Introdução ... O algoritmo ST-DBSCAN é construído modificando o algoritmo DBSCAN [7]. Em contraste com o algoritmo de agrupamento baseado em densidade existente, o algoritmo ST-DBSCAN [12] tem a capacidade de descobrir clusters em relação aos valores não espaciais, espaciais e temporais dos objetos. As três modificações feitas no algoritmo DBSCAN são as seguintes,

(i) O algoritmo ST-DBSCAN pode agrupar dados espaciais-temporais de acordo com atributos não espaciais, espaciais e temporais. (ii) DBSCAN não detecta pontos de ruído quando é de densidade variada, mas isso o algoritmo supera esse problema ao atribuir o fator de densidade a cada cluster. (iii) Para resolver os conflitos em objetos de borda, ele compara o valor médio de um cluster com o novo valor que vem.

6.2. Descrição do Algoritmo O algoritmo começa com o primeiro ponto p no banco de dados D . (i) Este ponto p é processado de acordo com o algoritmo DBSCAN e o próximo ponto é tomado. (ii) A função `RetrieveNeighbors` (objeto, E_{p1} , E_{p2}) recupera todos os objetos densidade-acessível do objeto selecionado em relação a E_{p1} , E_{p2} e $Minpts$. Se os pontos devolvidos no Eps -neighborhood são menores do que $Minpts$, o objeto é atribuído como ruído. (iii) Os pontos marcados como ruído podem ser alterados posteriormente, e os pontos não são diretamente acessíveis, mas serão densidade-acessível. ... (iv) Se o ponto selecionado for um objeto central, um novo cluster será construído. Então, todos os vizinhos de densidade direta de este núcleo de objetos também estão incluídos. (v) Então, o algoritmo coleta de forma iterativa objetos atingidos pela densidade do objeto do núcleo usando a pilha. (vi) Se o objeto não estiver marcado como ruído ou não estiver em um cluster e a diferença entre o valor médio do cluster e

o novo valor é menor do que ΔE , ele é colocado no cluster atual.

4 AVALIAÇÃO DO MÉTODO PROPOSTO

5 CONCLUSÕES E TRABALHOS FUTUROS

5.1 CONSIDERAÇÕES FINAIS

Criação de um método para resolver o Problema de Agrupamento em Grafos Dinâmicos e previsão de evolução destes agrupamentos. A expectativa é de que ao final desta pesquisa, tenha-se uma extensão do Dynagraph para visualização de agrupamentos dinâmicos e previsão dinâmica. Utilizar a ferramenta para outros tipos de doenças como: Chikungunya e Zika Vírus. Finalmente, espera-se que o produto final e os resultados obtidos possibilitem a previsão e prevenção de novos casos de dengue para um combate efetivo à doença.

5.2 LIMITAÇÕES

Dentre as dificuldades que podem interferir na execução deste projeto de pesquisa, as seguintes podem ser citadas: 1. A escassez de estudos relacionando os assuntos abordados: agrupamento, previsão em dados dinâmicos espaço-temporais, grafos dinâmicos e sistemas web de forma integrada; 2. Obtenção das informações dos focos e casos de dengue geolocalizadas e tempos das ocorrências. A extração dos dados semanais requer um processo manual em (SIMDA, 2018), pois é necessário o usuário selecionar o ano e a semana correspondente; 3. Visualização dos agrupamentos dinâmicos. Para contornar as dificuldades apresentadas pretende-se: 1. Automatizar a forma de obtenção dos dados; 2. Utilizar dados sintéticos para facilitar a validação da aplicação; 3. Utilizar o software Dynagraph como ambiente de suporte à validação e interação com os resultados dos métodos de agrupamento espaço-temporal utilizados.

5.3 TRABALHOS FUTUROS

??

REFERÊNCIAS

- CALIXTO, A.; NEGREIROS, M. DYNAGRAPH: Um Modelo de Edição e Representação de Grafos Dinâmicos. 1 ed. CLAIO/SBPO, p. 8, 2012.
- CALIXTO, A.; NEGREIROS, M. **DYNAGRAPH: Um Modelo de Edição e Representação de Grafos Dinâmicos**. Dissertação (Mestrado) — Mestrado Profissional em Computação Aplicada (MPCOMP), Universidade Estadual do Ceará, 2013.
- ESTER, M.; KRIEGL, H.-P.; SANDER, J.; XU, X. A density-based algorithm for discovering clusters in large spatial databases with noise. KDD, v. 96, n. 34, p. 226–231, 1996.
- FACHIN, O. **Fundamentos de metodologia**. [S.l.]: 3a Edição, Editora Saraiva, 2001.
- FAGHMOUS, J. H.; KUMAR, V. **Spatio-Temporal Data Mining for Climate Data: Advances, Challenges, and Opportunities**. [S.l.: s.n.], 2013. 83-116 p.
- KIM, H.; ANDERSON, R. Temporal node centrality in complex networks. 1 ed. PHYSICAL REVIEW, p. 8, 2012.
- KOSTAKOS, V. Temporal graphs. **Physica A**, p. 1007–1023, 2009.
- LAHIRI, M.; BERGER-WOLF, T. Structure prediction in temporal networks using frequent subgraphs. IEEE Symposium on Computational Intelligence and Data Mining, p. 35–42, 2007.
- LAHIRI, M.; BERGER-WOLF, T. Mining periodic behavior in dynamic social networks. Eighth IEEE International Conference on Data Mining, 2008.
- SIMDA. **Sistema de Monitoramento Diário de Agravos**. 2018. Disponível em: <<http://tc1.sms.fortaleza.ce.gov.br/simda/index>>.