

Um sistema webGIS para planejamento de leitura e agrupamento de unidades consumidoras de energia

A webGIS measurement and clustering planning system for electric power consumer units

Rodrigo Fumihito de Azevedo Kanehisa¹, Victor Henrique Bezerra de Lemos¹, Anderson Silva Fonseca¹, Daniel Lima Gomes Júnior², Geraldo Braz Júnior¹, Cáudio de Souza Baptista³
Núcleo de Computação Aplicada, Brasil

¹Universidade Federal do Maranhão (UFMA)

²Instituto Federal de Educação, Ciência e Tecnologia do Maranhão (IFMA)

³Universidade Federal de Campina Grande (UFCG)
{rkanehisa, victorhbl12, andsfonseca, geraldo, daniel}@nca.ufma.br, baptista@dsc.ufcg.edu.br

Nayara Carvalho Guedes⁴, Eliana Márcia Garros Monteiro⁵

⁴Centrais Elétricas do Pará (CELPA), Brasil,

⁵Companhia de Energia do Maranhão (CEMAR), Brasil,
Equatorial Energia,

nayara.guedes@celpa.com.br, eliana.monteiro@cemar-ma.com.br

Resumo — Os sistemas webGIS proporcionam uma solução completa contendo conceitos relacionados aos sistemas de informações geográficas (GIS) utilizando a infraestrutura da internet. Além disso, os sistemas webGIS possibilitam a integração entre diferentes tecnologias para tratamento de problemas que necessitam da utilização de métodos computacionais específicos, como o agrupamento de dados. No cenário das empresas de energia elétrica, faz-se necessário a utilização de heurísticas de agrupamento para o planejamento otimizado das leituras do consumo de unidades consumidoras. Este trabalho apresenta um sistema webGIS para construção de simulações de agrupamento utilizando os métodos K-Means e BIRCH, utilizando o cenário real de companhias de distribuição de energia elétrica.

Palavras Chave – webGIS; clusterização; planejamento; K-Means; BIRCH.

Abstract — The webGIS systems provides a complete solution containing Geographic Information Systems (GIS) related concepts using the internet infrastructure. Moreover, the webGIS systems enables the integration with different technologies for problem treatment that need to use specific computing methods, as data clustering. In the scenario of the energy companies, it is necessary to use clustering heuristics for optimized measurement planning of the electric power consumer units. This work presents a webGIS system for creating clustering simulations using K-Means and BIRCH methods, using the real scenario of energy distribution companies.

Keywords - webGIS; clustering; planning; K-Means; BIRCH.

I. INTRODUÇÃO

Com o avanço das tecnologias de dispositivos móveis e a popularização de tais dispositivos no mercado, em especial os

tablets e smartphones, foi adicionada a vantagem da flexibilidade aos mapas digitais, antes restritos ao uso com computadores pessoais (PCs). Isso torna possível a cada usuário, individualmente, criar e manipular mapas sob demanda, de acordo com suas necessidades.

A Web é um meio de distribuição de informações rico por garantir toda a infraestrutura de rede que a Internet provê. Diante desse ambiente, os mapas desempenham um papel fundamental, possibilitando o uso de funções tradicionais, como a interpretação sobre relacionamentos geoespaciais, até função de aumentar a realidade do usuário podendo conectar informações como fotos e descrições a uma localização específica num ambiente com grande riqueza de interação [1].

De acordo com [2], as aplicações baseadas na web permitem a visualização a diferentes usuários através de um navegador. Além disso, possibilita o uso de aplicações baseadas em computação, integradas de forma simples ao usuário, tendo em vista que o sistema não precisa ser configurado, apenas acessado e utilizado para a finalidade projetada.

Sistemas que utilizam mapas não são mais usados apenas em sua funcionalidade natural, de exibição, mas possibilitam a exploração de padrões espaciais e navegação sobre o espaço geográfico, além de novas funcionalidades como a facilitação do entendimento de problemas a serem tratados paralelamente com a análise geográfica do problema real. A utilização da Web possui um potencial enorme em relação à sua capacidade de atingir grande quantidade de usuários com custos reduzidos e facilidade de atualização dos dados.

Sistemas webGIS são o produto da tecnologia da Internet combinada à tecnologia GIS (*Geographic Information System*)

para melhor compartilhamento de dados através do browser [3]. Os sistemas GIS adaptados à web devem responder às requisições dos clientes de maneira eficiente, com baixo tempo de resposta e alta usabilidade, ou seja, facilitar ao máximo a utilização e compreensão do sistema pelo usuário. Assim, o principal objetivo dessa modalidade de aplicações é ser simples o suficiente para que usuários não-especialistas possam tomar decisões e conclusões diante dos dados [4].

Segundo [5], a tecnologia webGIS associada a outros serviços de informação possibilita a colaboração interativa e o compartilhamento de informações para uma construção eficiente de entendimento para sistemas de suporte a decisão.

Dessa forma, os sistemas webGIS podem ser aplicados em diferentes cenários. Essa abordagem baseada em mapas proporciona aos usuários uma visão holística do ambiente analisado e permite, por exemplo, integrar dados de planejamento das leituras de unidades consumidoras nas distribuidoras de energia elétrica. A informação espacial é importante tanto para a etapa de organização de leitura quanto para o faturamento dessas companhias.

Para resolver esse problema, faz-se necessário integrar ao sistema webGIS estratégias e algoritmos de clusterização para geração dos grupos de leitura. Para compreender melhor este domínio do problema, é apresentada na seção II uma descrição do cenário de planejamento de leituras do consumo de energia no Brasil.

II. CENÁRIO DE PLANEJAMENTO DE LEITURA EM EMPRESAS DE ENERGIA ELÉTRICA NO BRASIL

O processo de faturamento de uma empresa de distribuição de energia elétrica envolve dois processos principais que são a leitura do consumo registrado pelos medidores e a entrega das faturas aos consumidores.

Embora existam no mercado medidores eletrônicos com possibilidade de telemedição, dispensando os leituristas, a efetiva implantação desses medidores é um processo gradual, que depende da sua viabilidade econômica. No Brasil, por exemplo, a substituição dos mais de 63 milhões de medidores convencionais por medidores eletrônicos com telemedição demandará vários anos e, portanto, o cálculo do agrupamento e roteamento de leituristas – que demanda criar os grupos de leitura permanecerá como um real problema por muito tempo.

A Agência Nacional de Energia Elétrica (ANEEL), regulamenta as condições gerais de fornecimento de energia elétrica [6], visando aprimorar o relacionamento entre os agentes responsáveis pela prestação do serviço público de energia elétrica e os consumidores. A agência estabelece, entre outras coisas, que uma concessionária se obriga a fazer as leituras em intervalos de 30 dias, com variações entre 27 e 33 dias. Além disso, precisa apresentar a fatura aos consumidores com prazo mínimo de 5 dias para que o seu pagamento seja efetuado. É facultado aos consumidores escolher o dia do pagamento entre 6 diferentes datas distribuídas ao longo do mês.

Tipicamente as distribuidoras de energia elétrica precisam organizar a leitura das medições e entrega das faturas, através da definição de grupos de clientes que cada leiturista ou

entregador deverá visitar e as rotas que deverão seguir. Normalmente esses critérios são definidos baseados na criação de regiões de atendimento.

Entre os critérios para definição do tamanho dos grupos e seleção das unidades consumidoras que compõem cada grupo existem dois critérios básicos:

a) Homogeneidade: Os grupos devem ser os mais homogêneos possíveis quanto à carga de trabalho da equipe de leituristas. Isto leva à minimização dos custos operacionais de mão de obra.

b) Compacidade: Forma geográfica dos grupos deve ser a mais compacta possível. Isto contribui para que as rotas de leitura sejam as mais eficientes possíveis.

Outras atividades relacionadas à gestão comercial também podem se beneficiar da definição desses grupos, com redução de deslocamentos e tempo para realização de serviços, como corte, religamento e inspeção de fraudes. A definição dessas regiões de atendimento recai em um problema de agrupamento.

Estes aspectos destacam a importância de desenvolvimento de métodos computacionais para a definição de maneira eficiente dos grupos de leitura com a possibilidade de otimizar recursos e deslocamentos para realizar a leitura de consumo de energia das unidades consumidoras e entrega de faturas, melhorando o processo de faturamento da concessionária. Além disso, pode proporcionar uma melhoria da qualidade do serviço com consequente incremento no índice de satisfação dos clientes.

O problema em questão pode se apresentar e ser resolvido do ponto de vista de modelagem e simulação computacional sobre diversos enfoques.

Apresentaremos a seguir os algoritmos de agrupamento utilizados e seus respectivos resultados em um processo de aquisição de conhecimento acerca do problema.

III. HEURÍSTICAS DE AGRUPAMENTO

A criação automática dos grupos de leitura envolve um estudo dos parâmetros das empresas de energia e inserção de restrições aos grupos. Neste trabalho apresentamos duas abordagens de clusterização que auxiliam a tarefa manual de criação dos grupos de leitura.

A. K-Means

O K-Means [7] é um dos algoritmos mais simples de aprendizagem não supervisionado para solucionar o problema de clusterização em problemas conhecidos.

A clusterização de um conjunto de dados é realizada considerando usando um número k de clusters, sendo o objetivo a definição da mesma quantidade de centroides, um para cada agrupamento de dados, sendo que os centroides devem ser colocados distantes um dos outros.

A clusterização com K-Means pode ser definida em fases.

a) Fase 1: definição dos centroides e associação de cada ponto do conjunto de dados ao centroide mais próximo até que não existam mais pontos restantes. Nesse estágio um agrupamento inicial é obtido.

b) Fase 2: re-calcular novos k centróides como baricentros dos clusters resultantes da fase anterior. Dessa forma são gerados novos k centróides, sendo que cada ponto do conjunto de dados é novamente calculado para ser conectado ao centróide mais próximo. Esta fase se repete até que os centróides não se movimentem mais. Pode-se afirmar que o objetivo é a minimização de uma função objetivo apresentada na Equação 1.

$$J = \sum_{j=1}^k \sum_{i=1}^n \|x_i^{(j)} - c_j\|^2 \quad (1)$$

A Figura 1 demonstra a criação de três. Neste ponto, o algoritmo K-Means encerra sua execução, quando os centróides convergirem a estabilidade, retornando os dados agrupados. A quantidade de grupos vai depender da quantidade de clusters definidos no processamento.

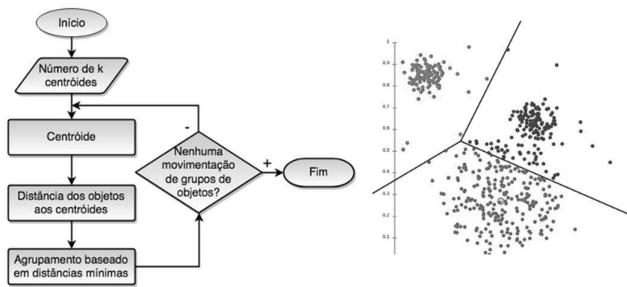


Figure 1. Resultado da clusterização K-Means com $k=3$. Adaptado de [8]

Um aspecto importante para a escolha do K-Means é a possibilidade de usar grande quantidade de pontos nos conjuntos de dados gerados, que é o cenário das companhias de energia elétrica.

Em geral, são realizadas clusterizações em dados reais mantendo as restrições das companhias e utilizando grande quantidade de pontos inseridos aos grupos de dados.

B. BIRCH

O algoritmo de clusterização e mineração de dados não supervisionado BIRCH (*Balanced Iterative Reducing and Clustering using Hierarchies*) proposto em [9] realiza o agrupamento hierárquico em conjuntos de dados particularmente grandes, que é o caso das empresas de distribuição de energia elétrica em grandes países como o Brasil. O algoritmo recebe como entrada um conjunto de N pontos a serem agrupados, um número K de grupos a serem formados.

O BIRCH propõe o conceito de *Clustering Feature* (CF) e opera em quatro fases, sendo que as fases 2 e 4 são opcionais. O fluxo do algoritmo pode ser visualizado na Figura 2.

a) Fase 1: constrói uma árvore CF, sendo \vec{X}_i o conjunto de pontos em um cluster d -dimensional. A árvore CF é uma estrutura de dados que contém informações sobre o conjunto de dados. O recurso de agrupamento CF é definido como o triplo $CF = (N, LS, SS)$ onde N é o conjunto de dados, LS é a soma linear dos dados (Equação 2) e SS é a soma quadrática dos pontos dos dados (Equação 3). Os recursos de agrupamento são organizados na árvore CF com dois parâmetros B e T , onde B é o fator de ramificação que determina a precisão necessária e T é o limiar que define o tamanho da folha da árvore.

b) Fase 2: o algoritmo verifica todas as entradas de folhas na árvore CF inicial para reconstruir uma árvore CF menor, ao mesmo tempo que remove *outliers* e agrupa subgrupos aglomerados em grandes grupos.

c) Fase 3: um algoritmo de agrupamento existente é usado para agrupar todas as entradas de folhas.

d) Fase 4: os centróides produzidos na fase 3 são usados como sementes e redistribuem os pontos de dados para as sementes mais próximas para obter um novo conjunto de clusters.

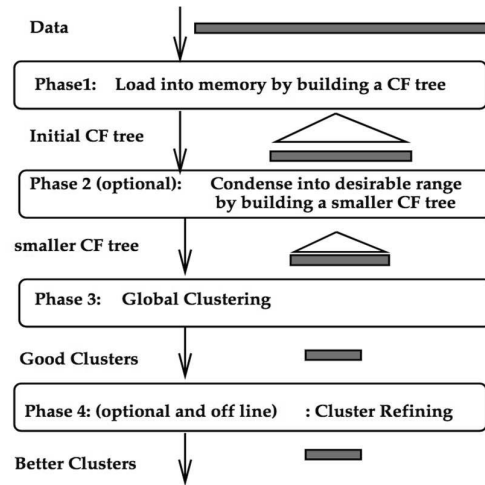


Figure 2. Visão geral do algoritmo BIRCH [9].

$$\sum_{i=1}^N \vec{X}_i \quad (2)$$

$$\sum_{i=1}^N \vec{X}_i^2 \quad (3)$$

Por se tratar de um algoritmo de agrupamento que pode ser aplicado a um grande conjunto de dados, o BIRCH torna-se interessante para testes com cenários reais, que no caso do setor de energia brasileiro envolve milhões de usuários em determinadas regiões de leitura. A saída do algoritmo realiza o agrupamento dos dados gerando etapas e subdividindo cada etapa em unidades de leitura.

IV. ARQUITETURA E TECNOLOGIAS UTILIZADAS

A. Arquitetura

As restrições associadas aos leituristas são fundamentais para as propostas de otimização do sistema. O fluxograma apresentado na Figura 3 mostra as atividades para geração dos clusters e visualização webGIS no sistema. Essas atividades foram desenvolvidas para adaptar as atividades às restrições do processo de leitura.

Além disso, a Figura 4, apresenta a visão física arquitetural do sistema em termos de nós conectados. Os nós são unidades de hardware (servidores, computadores, etc.) que executam módulos ou componentes de software identificados no projeto. A alocação de componentes aos nós leva em consideração toda a capacidade dos nós (características dos hardwares).

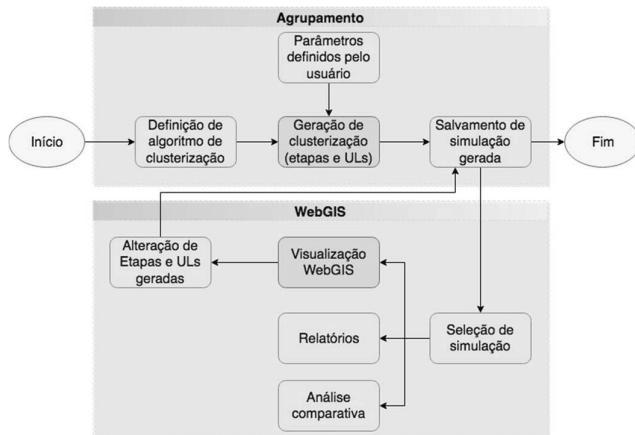


Figure 3. Fluxograma para geração dos clusters e visualização webGIS.

Os nós e componentes principais identificados são:

- Servidor Oracle 12c (Nó): responsável por toda a estruturação do sistema, possibilitando consultas espaciais;
- Servidor IIS (Nó): executa todos os componentes do sistema (representados pelas camadas Lógica de apresentação, Lógica de aplicação e Lógica de aquisição de dados), além da camada de processamento do agrupamento, Lógica de clusterização;
- Lógica de apresentação (Componente): executa todas as interfaces do sistema em ASP.NET e realiza o tratamento primário da consistência dos dados;
- Lógica de aplicação (Componente): executa todas as funcionalidades que conectam a interface, o módulo de clusterização e os dados do sistema. Possui as regras de autenticação, funcionalidades de negócio e controle de interações;
- Lógica de dados (Componente): executa todas as ações referente ao tratamento e recuperação de informação em .NET usando camada de persistência ADO.NET;
- Lógica de clusterização (Componente): executa todas as ações de agrupamento e clusterização na aplicação.

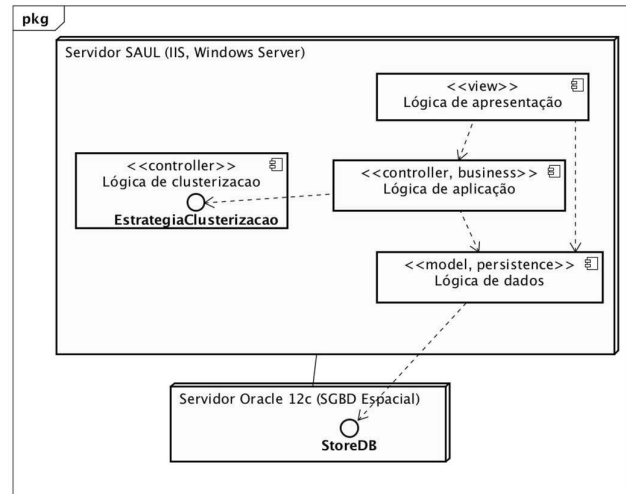


Figure 4. Visão deployment do sistema.

É importante observar que toda a interação entre os usuários e o sistema é realizado diretamente através do nó Servidor Web IIS. Assim, as informações geradas a partir dos dados fornecidos pelos nós, que armazenam os sistemas que já existem na companhia, serão acessados através de um único ponto de acesso, promovendo uma melhor integração, simplificação e manutenção do sistema. O módulo de clusterização funcionará inteiramente no mesmo servidor, sob a forma de serviço, possibilitando sua chamada a qualquer momento.

B. Tecnologias utilizadas

A plataforma .NET [10] é um conjunto de componentes desenvolvidos pela Microsoft para suportar aplicações e serviços, com forte enfoque na facilidade de desenvolvimento de aplicações para os mais diversos tipos de dispositivos. Como elementos importantes desta plataforma temos o Framework .NET e o ambiente de desenvolvimento Microsoft Visual Studio.

O Framework .NET é a base da plataforma. Ele pode ser entendido como uma máquina virtual, tendo em vista que representa uma camada entre o sistema operacional e código das aplicações desenvolvidas para a plataforma. Neste ambiente o código escrito na linguagem de programação escolhida não é compilado em linguagem de máquina, mas sim em uma linguagem intermediária, chamada de MSIL (*Microsoft Intermediate Language*). Esta linguagem é interpretada pelo Framework em momento da execução.

No entanto, o Framework .NET possui um conjunto completo de bibliotecas de componentes, automatizando diversas tarefas e possuindo características que o tornam um pouco mais que uma máquina virtual. Ele abrange uma estrutura de objetos, classes e ferramentas que se integram ao sistema operacional para fornecer suporte ao desenvolvimento.

O componente central do Framework .NET é o CLR (Common Language Runtime), responsável pela comunicação

com o sistema operacional, pelo gerenciamento de memória, requisição da execução de instruções na CPU, etc. O CLR interpreta a MSIL durante a execução dos programas .NET, traduzindo em linguagem de máquina.

Esta estratégia arquitetural garante que a portabilidade na plataforma não é responsabilidade do compilador de cada linguagem e sim do CLR. Além disso, existem projetos de migração, como o Mono, para sistemas não suportados pela Microsoft, como Linux e Solaris, que garantem grande compatibilidade com o Framework .NET original.

A linguagem utilizada é C#, linguagem de programação inspirada em C++, orientada a objetos, lançada juntamente com a plataforma .NET. Esta linguagem tem como principal foco o desenvolvimento para Web e dispositivos móveis.

Para tratamento e armazenamento das informações espaciais utiliza-se o Oracle habilitado para trabalhar com dados georreferenciados. Trata-se de um SGBD objeto-relacional que propicia uma abordagem aberta, ampla e integrada para o gerenciamento de informações, consistente em um banco de dados Oracle e uma instância do servidor Oracle.

Por fim, utiliza-se a biblioteca Accord.NET Framework [11], que é um framework de aprendizagem de máquina .NET disponível na licença Gnu Lesser Public License, utilizada para aprendizado de máquina e computação científica disponível para a aplicações .NET escritas em C#. Possui algoritmos para audição e visão computacional, processamento de sinais, aprendizado de máquina e aplicações de estatísticas, estando também disponível para uso comercial.

V. RESULTADOS OBTIDOS

Este trabalho apresenta o resultado dos agrupamentos realizados com informações da CEMAR e CELPA (empresas de energia elétrica da Equatorial Energia no Brasil), durante o projeto de P&D para a criação de um sistema webGIS com geração de clusterizações de regiões de leitura.

De forma mais específica, foi realizada a criação de geometrias de localidades a partir dos pontos das localidades disponíveis, cerca de 70 mil pontos e aproximadamente 128 localidades, que podem ser visualizadas na Figura 5.

Na interface de criação da região de leitura (RL), o usuário seleciona os polígonos marcando as localidades e, após finalizar a marcação de localidades, uma região de leitura é criada. As localidades que foram selecionadas em uma região de leitura não podem fazer parte de outra região de leitura e ficam desabilitadas para seleção. Com a definição da RL, é possível definir um escopo para a criação dos clusters nos algoritmos de agrupamento.

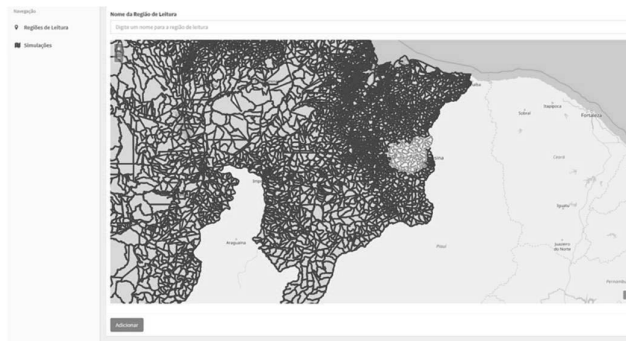


Figure 5. Interface de visualização e seleção de localidades para criação de uma região de leitura.

Neste estudo, foram testados os algoritmos K-Means e BIRCH. Ambos foram integrados ao sistema webGIS de agrupamento.

Foi realizada, ainda, análise comparativa entre os resultados das duas clusterizações e verificou-se que os clusters gerados com BIRCH produziram agrupamentos com melhor divisão espacial, no entanto, em casos específicos, geraram saídas consideradas ruins.

A Figura 6 apresenta a comparação entre os agrupamentos realizados com o algoritmo K-Means e o BIRCH. Comparando as imagens, observou-se a diferença entre a saída dos dois algoritmos. O BIRCH, em alguns casos como o destacado na imagem gerou dois grupos onde o K-Means gerou apenas um, essa divisão em dois grupos realizada pelo BIRCH não foi considerada viável, tendo em vista que dessa forma seria necessária uma etapa para ser efetuada a leitura de apenas duas instalações.

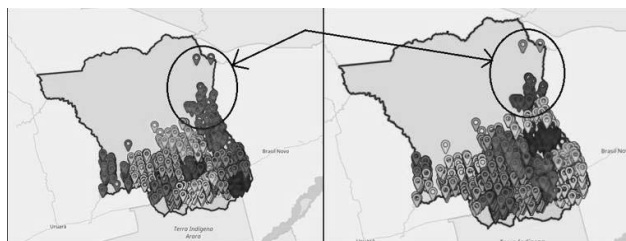


Figure 6. Resultado de comparação das saídas geradas pelo K-Means (esquerda) e BIRCH (direita), com simulações geradas com dados georreferenciados das unidades consumidoras.

A visualização webGIS permite ao usuário selecionar as etapas (que são os dias de leitura) e visualizar as instalações pertencentes ao grupo de leitura. Cada etapa recebe uma cor que é mapeada à interface para seleção de etapas. O sistema possibilita visualizar todas as etapas ou uma etapa específica. A Figura 7 apresenta a tela de visualização com as opções de alteração pelo usuário.

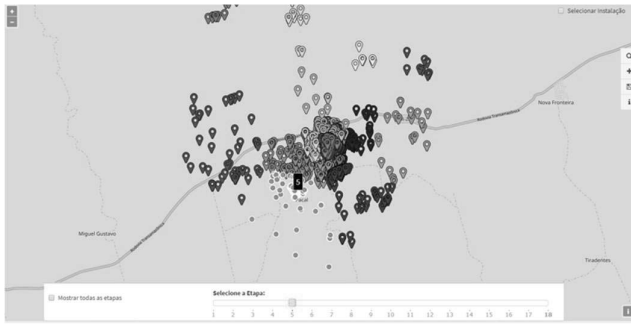


Figure 7. Visualização das etapas de leituras gerados com o algoritmo K-Means, que permite a manipulação das pelo usuário (parte inferior).

A Tabela 1 mostra a comparação entre os algoritmos utilizados e os seus respectivos tempos de processamento aplicados em grande quantidade de dados das empresas de energia elétrica.

As estratégias de clusterização devem possibilitar o processamento de grande quantidade de dados, visto que o cenário real consiste em diversas situações de geração de grupos de leituras em cerca de 4 milhões de unidades consumidoras.

TABLE I. RESULTADOS DE CLUSTERIZAÇÃO COM K-MEANS E BIRCH COM K=18 (QUANTIDADE USUAL DE ETAPAS DE MEDIÇÃO).

Cidade	Nº de Instalações	Método	Tempo Médio
Uruará	14.071	K-Means	1,18750 s
		BIRCH	0,89063 s
Medicilândia	9.607	K-Means	0,35000 s
		BIRCH	0,78923 s

Os resultados obtidos possibilitaram a utilização da ferramenta pelos operadores das companhias para validação e testes dos dados reais. Anteriormente à ferramenta, a clusterização era realizada manualmente, já com o uso dos algoritmos implementados, boa parte da definição manual das etapas poderá ser auxiliada pelo sistema de clusterização com dados georeferenciados.

VI. CONCLUSÕES

O problema de clusterização para auxiliar a tarefa de planejamento de leituras de unidades consumidoras mostra-se uma tarefa de alto nível de complexidade, devido às diversas restrições que podem ser atribuídas à criação dos grupos de leitura, e de alto custo computacional, já que deverá agrupar milhares de unidades consumidoras.

Apesar disso, este trabalho apresenta um resultado que permite aos operadores e analistas utilizarem as clusterizações com K-Means e BIRCH de forma a auxiliar o trabalho de planejamento, além de agilizar o processo de criação de grupos.

Os resultados obtidos e a sua integração ao sistema no formato webGIS permite ainda a visualização dos operadores através de um sistema de informações geográficos na web interação a partir de diversos dispositivos.

Apesar dos resultados já alcançados pretende-se realizar novos estudos e novas abordagens com algoritmos diferentes, como o GRASP (*Greedy Randomized Adaptive Search Procedure*) [12], um algoritmo aplicado a soluções de problemas de otimização combinatória, que possibilita uma análise com uma modelagem mais complexa do problema e possibilita a inclusão de maior quantidade de restrições para a clusterização.

Além disso, pretende-se modificar a abordagem com K-Means para inserir internamente ao algoritmo novas restrições, além das distâncias entre as unidades consumidoras. Essa abordagem possibilitaria ao sistema sugerir e comparar simulações às simulações atuais em execução na empresa de distribuição de energia elétrica.

AGRADECIMENTOS

Os autores agradecem a UFMA, IFMA, UFCG e Equatorial Energia pelo apoio financeiro durante o projeto de P&D intitulado “Construção Auto Organizável de Agrupamentos de Leitura sujeito a Restrições através de Métodos de Inteligência Computacional e Sistemas de Informação Geográfica”.

REFERÊNCIAS BIBLIOGRÁFICA

- [1] M. J. Kraak and A. Brown, “Web Cartography: Development and Prospects”. London: Taylor and Francis. 2001.
- [2] F. Rosa and M. Casquilho, “Web-based scientific computing, a little explored field an engineering problem: Pneumatic transport”, 2017 12th Iberian Conference on Information Systems and Technologies (CISTI), Lisbon, 2017, pp. 1-6.
- [3] G. B. Shi, “The Design of Campus Monitoring and Managing System for WaterSaving Based on Webgis,” 2017 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Exeter, United Kingdom, 2017, pp. 951-954.
- [4] TORUN, A.; KÖBBEN, B.; LEMMENS, R. Processing Spatial Data on the Internet. International Archives of Photogrammetry and Remote Sensing. Vol. XXXIII, Part B6. Amsterdam 2000.
- [5] C. Yu, “Research on Emergency Logistics Decision Support System Design under Data Ming & WebGIS Technology,” 2016 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), Changsha, 2016, pp. 375-378.
- [6] ANEEL, “Resolução N. 456, de 29 de novembro de 2000”. Brasil: Agência Nacional de Energia Elétrica. 2000.
- [7] J. B. MacQueen, “Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability”. Berkeley: University of California Press, 1:281-297. 1967.
- [8] S. Mukherjee, “Improving generalization of k-means clustering based probabilistic neural network using noise injection”. 2017 International Conference on Soft Computing and its Engineering Applications (icSoftComp), Changa, Anand, India, 2017, pp. 1-5.
- [9] T. Zhang, R. Ramakrishnan and M. Livny, “BIRCH: A New Data Clustering Algorithm and Its Applications”. Data Mining and Knowledge Discovery, 1(2), 141—182. 1997.
- [10] Microsoft, “Introdução à plataforma .NET da Microsoft”. 2017. Disponível em <<https://msdn.microsoft.com/pt-br/aa702903.aspx>>. Acesso em 21 de novembro de 2017.
- [11] Accord Framework, “Accord - Machine Learning .NET Framework”. 2008. Disponível em <<http://accord-framework.net/>>. Acesso em 06 de dezembro de 2017.
- [12] R. Z. Rios-Mercado and E. Fernández, “A reactive grasp for a commercial territory design problem with multiple balancing requirements”. Computer & Operations Research, pp. 1–43. 2007.