XXV Congresso de Iniciação Científica da Unicamp

18 a 20 Outubro Campinas | Brasil

Estudo de Explosões Solares via Agrupamento de Dados

João Victor Ignácio*, Guilherme Palermo Coelho, Ana Estela Antunes da Silva

Resumo

Explosões Solares, que correspondem à liberação de grandes quantidades de energia da atmosfera solar em um curto espaço de tempo, podem afetar significativamente a vida na Terra, tendo impacto direto em espaçonaves, satélites, redes de distribuição de energia e redes de comunicação, dentre outros. Neste sentido, a presente pesquisa buscou identificar padrões e extrair informações a partir de conjuntos de dados coletados por diversos equipamentos que realizam o monitoramento de explosões solares. Para isso, foi adotado um algoritmo de agrupamento de dados (K-Medoids) combinado a um coeficiente de similaridade que contempla atributos de tipos diferentes, conhecido como Coeficiente de Gower. Os resultados de agrupamento foram avaliados pela métrica de Silhueta, que indicou que a melhor partição para os dados estudados é a que os divide em 10 grupos.

Palavras-chave: Explosões Solares, Clima Espacial, Agrupamento de Dados

Introdução

As explosões solares, que são liberações abruptas de energia magnética acumulada na atmosfera solar, podem influenciar não apenas os seres vivos, mas também os diferentes sistemas e dispositivos relacionado à vida na Terra¹, como, por exemplo, os sistemas de comunicação.

Visto o impacto desses eventos, equipamentos como os satélites GOES (*Geostationary Operational Environmental Satellite*), do *Space Weather Prediction Center (SWPC)*, foram instalados visando monitorar as atividades solares e coletar dados, o que possibilitou diversos estudos.

Para a análise dos dados disponibilizados pelo SWPC, nessa pesquisa foi utilizado o algoritmo K-Medoids, uma técnica de agrupamento particional, que busca dividir a base de dados em grupos, de forma que dados com características semelhantes sejam alocados em um mesmo grupo. Para avaliar o quão diferentes os dados são é necessário uma métrica de dissimilaridade, sendo que aqui foi adotado o Coeficiente de Dissimilaridade de Gower, dado que a base de dados utilizada possui atributos de diferentes tipos. Mais especificamente, tal base possui 14 atributos do tipo binário, 3 atributos do tipo inteiro e 4 atributos do tipo real. Para identificar o número de grupos que melhor representa os dados estudados, foi utilizado o Coeficiente de Silhueta. Identificado este número, foi feita uma análise dos dados alocados a cada grupo.

Resultados e Discussão

O algoritmo K-Medoids foi aplicado a uma base de dados com 21 atributos coletados diariamente no período entre 04 de Novembro de 1997 e 25 de Dezembro de 2013. Para utilização do algoritmo de agrupamento, tais dados foram normalizados entre 0 e 1.

Na Figura 1 são apresentados os valores do coeficiente de Silhueta obtidos para cada agrupamento retornado pelo K-Medoids, com número de grupos (K) variando entre 7 e 10. Dado que quanto maior o valor do coeficiente de Silhueta melhor o particionamento dos dados, para os experimentos feitos aqui o melhor valor foi K=10 grupos.

Para a melhor partição dos dados (K=10 grupos), foram avaliados os atributos dos dados de cada grupo, sendo

destacada aqui a Média do Fluxo de Raio X para cada grupo, dada na Figura 2.

Figura 1. Valor de Silhueta para agrupamentos com K entre 7 e 10 grupos.

Figura 2. Média do Fluxo de Raio X em cada um dos K=10 grupos.

Conclusões

Como pode ser visto na Figura 2, existem grupos com diferentes características dentre os obtidos, cabendo destacar aqui o grupo 4, que tem o maior Fluxo de Raio X médio. Isto indica um conjunto de dados associado a um maior número de ocorrências de explosões solares.

Agradecimentos

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro.

¹ROYAL ACADEMY OF ENGINEERING. Extreme space weather : impacts on engineered systems and infrastructure. London: Royal Academy of Engineering, 2013.