
TÉCNICAS DE VISUALIZAÇÃO DE INFORMAÇÃO PARA APOIO À EXPLORAÇÃO E ANÁLISE DE DADOS COLETADOS PELO SISTEMA INTEGRADO DE MONITORAMENTO AMBIENTAL - SIMA

ALISSON FERNANDO COELHO DO CARMO (1) (2)
MILTON HIROKAZU SHIMABUKURO (1) (3)
ENNER HERENIO DE ALCÂNTARA (1) (4)

(1) Programa de Pós-Graduação em Ciências Cartográficas (PPGCC)
Faculdade de Ciências e Tecnologias (FCT)
Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP)
Presidente Prudente – SP – Brasil

(2) alisondocarmo@gmail.com

(3) miltonhs@fct.unesp.br

(4) enner@fct.unesp.br

A modernização de dispositivos sensores, impulsionada pelo crescente desenvolvimento tecnológico, tem consolidado este recurso como uma importante ferramenta para permitir o monitoramento e estudos relacionados com fenômenos e processos no meio ambiente. Estes dispositivos são fontes constantes de dados, em razão da periodicidade da leitura e coleta dos dados. Cada registro, obtido pela leitura dos sensores, pode ser integrado com os atributos referentes ao local que foi coletado, bem como o momento de realização da coleta. Tal característica permite a consideração do componente espaço temporal durante o processo de análise. Desta forma, a periodicidade temporal presente na coleta automática dos dados transforma estas leituras em grandes e crescentes séries temporais para cada variável capturada. Uma das principais características das séries temporais é a capacidade de registrar o estado histórico de cada variável, permitindo investigações focadas na observação do comportamento e relações entre as variáveis de acordo com o fator tempo, buscando a descoberta de padrões e tendências. Uma das abordagens para evidenciar a existência de padrões e facilitar a observação do comportamento das variáveis é a inspeção visual do analista sobre uma representação gráfica sintetizada do conjunto, por meio de técnicas de Visualização de Informação. Neste cenário, recursos computacionais são utilizados para facilitar a interpretação dos dados por meio da capacidade de percepção e cognição associada ao sistema visual humano. Além disso, a interatividade do analista com os dados, descrita no cenário de Visual Analytics, é um recurso que pode potencializar o processo de análise, destacando o papel fundamental do analista durante a exploração e análise de dados, definido como *sense making loop* por Keim et al (2008). Todavia, a coleta automática de dados utilizando sensores depende do correto funcionamento de dispositivos que sofrem com o processo de degradação pela exposição ao ambiente, que pode afetar a qualidade dos dados registrados. A metodologia de amostragem dos dados também pode influenciar em sua qualidade, pois está relacionada com o funcionamento integrado de diversos mecanismos, na leitura, conversão, transmissão e registro dos dados. Neste contexto, este trabalho tem o objetivo de apresentar os resultados preliminares da investigação focada na utilização de recursos computacionais para auxiliar e potencializar o processo de exploração e análise de séries temporais ambientais, tais como métodos de análise estatística, processamento/mineração de dados e técnicas de Visualização de Informação. Neste projeto, são utilizados os dados coletados pelo Sistema Integrado de Monitoramento Ambiental (SIMA), descrito por Alcântara et al (2013), o qual possui uma rede de plataformas aquáticas que realizam a coleta de variáveis diariamente, no intervalo de uma hora. Os dados são disponibilizados pelo portal SIMA (<<http://www.dsr.inpe.br/hidrosfera/sima/index.php>>) em formato de planilhas eletrônicas, que foram posteriormente convertidas e armazenadas em um Sistema Gerenciador de Banco de Dados (SGBD) para facilitar as consultas a serem realizadas. Com os dados armazenados, foi construída uma representação para permitir a inspeção visual do período de atividade de cada plataforma, possibilitando a verificação de quais ficaram ativas por um maior período de tempo, bem como quais permanecem atualmente ativas. Baseado no período de atividade destacado pela representação visual anterior, foram gerados gráficos para cada plataforma para descrever a proporção de dados capturados e dados com erros ou ausentes, fator que permite aferir sobre a qualidade dos dados de cada plataforma. A taxa de erro realçada na representação anterior reforçou a necessidade de analisar individualmente as plataformas, respeitando suas características próprias. Tendo em vista o comportamento específico das plataformas foi utilizada uma representação do tipo *point-based* em uma matriz de dados, representando o comportamento de cada atributo da plataforma em uma escala específica. Além da variação dos valores dos atributos,

nesta representação visual foi possível observar a existência simultânea de erros entre conjuntos de variáveis, levantando a hipótese de que poderia haver certa relação na leitura de algumas variáveis. Para evidenciar esta característica, uma adaptação da representação do tipo *Calendar View* foi construída para permitir a observação visual de quais variáveis possuíam erros em cada instante de tempo, por meio da conversão de uma sequência binária – capaz de identificar os erros nos sensores – para o espaço de cor RGB. Para verificar a relação entre os valores consistentes dos atributos, a representação de Coordenadas Paralelas foi utilizada. Tal visualização permite reorganizar os eixos das variáveis buscando o comportamento visual mais homogêneo dos valores. Esta representação pode ser enriquecida por meio de inspeções com técnicas estatísticas que permitem calcular a correlação entre as variáveis. De forma geral, a integração entre recursos de diferentes áreas, como sensoriamento remoto para aquisição dos dados, análise estatística para a modelagem das séries temporais, técnicas de representação de Visualização de Informação e interatividade de Visual Analytics tem se mostrado eficazes enquanto ferramentas que podem enriquecer e potencializar o raciocínio envolvido no processo de exploração e análise de dados.

Referências

ALCÂNTARA, E. et al. A system for environmental monitoring of hydroelectric reservoirs in Brazil. *Ambiente e Água - An Interdisciplinary Journal of Applied Science*, 8(1). (2013)

KEIM et al. *Information Visualization*. cap. Visual Analytics: Definition, Process, and Challenges. pp 154--175. Springer-Verlag, Berlin, Heidelberg (2008)