

---

# GIS não geográfico ?

---

Por **MundoGEO** | 0h00, 29 de Setembro de 2005

## *O uso de análises espaciais para dados não geográficos*

Que tal pensarmos no GIS para usos "não geográficos" ? Parece estranho ? Pois bem, existem hoje várias áreas do conhecimento que estão usando e abusando de técnicas de análise espacial em dados essencialmente não geográficos. Vamos dar uma olhada em algumas delas. E não estou falando de aplicações em Medicina e áreas correlatas – essas são "geográficas" por essência, mas não no espaço da superfície terrestre, e sim no mapeamento de organismos.

Estamos falando das Ciências Sociais e Ciências da Informação. Somente para citar algumas dessas áreas: sistemas de recomendação em mineração de dados de varejo, análise de relevância de artigos científicos, análise de citações e análise de estilo literário em Cienciometria e Informetria. Mas que diabos é tudo isso ? E o que isso tem a ver com o GIS ?

Basicamente, o que essas áreas fazem é utilizar técnicas estatísticas multidimensionais sofisticadas, como a análise fatorial (geralmente por componentes principais), para obter medidas de proximidade e distância entre os objetos que estão sendo estudados. Essa proximidade ou distância pode representar similaridade, relevância, aparência ou correlação entre esses objetos. Com isso, dois objetos próximos significam que são mais parecidos entre si do que dois objetos distantes.

Daí, técnicas como o escalonamento multidimensional (MDS) pegam esses objetos e colocam em um espaço n-dimensional (geralmente bi ou tri) a partir das distâncias entre eles. Basicamente, os objetos passam a ter coordenadas X,Y ou X,Y,Z associadas a eles. E aí já estamos falando de uma natureza que o GIS entende, e técnicas como análise de rede, modelos digitais de terreno, agrupamentos espaciais e outras entram no leque de ferramentas disponíveis.

Vejam um exemplo na área de Ciência da Informação – a Cienciometria e a Informetria, que trabalham com as formas de avaliação do conhecimento científico e de medição dos fluxos de informação. Trata-se de um trabalho que desenvolvi com um grupo de pesquisadores da EAESP-FGV para a avaliação de relevância de artigos científicos na área de marketing a partir da análise das citações que um artigo/autor faz de outros artigos/autores.

**"As áreas não geográficas utilizam técnicas estatísticas multidimensionais sofisticadas para obter medidas de proximidade e distância entre os objetos estudados"**

Com o avanço da informática e das telecomunicações, o volume de informações disponíveis sobre todos os aspectos da atividade humana e em todos os ramos da ciência, passou a crescer a taxas vertiginosas. A troca ineficiente de informações entre instituições e centros de pesquisa é agravada pela incapacidade humana de lidar com grandes volumes de informação. O total de publicações e informações geradas amplia-se de forma muito mais acelerada que a proporção da literatura pessoalmente conhecida e assimilada por um pesquisador ou profissional de negócios.

Utilizando técnicas de mineração de texto (Text Mining), a partir do conjunto de 50 artigos selecionados, pudemos obter uma rede de citações, na qual cada nó da rede (grafo) é um artigo e cada aresta uma citação DE um artigo PARA outro. Com o uso do GIS, essa rede pode ser representada em um espaço tridimensional, conforme figura.

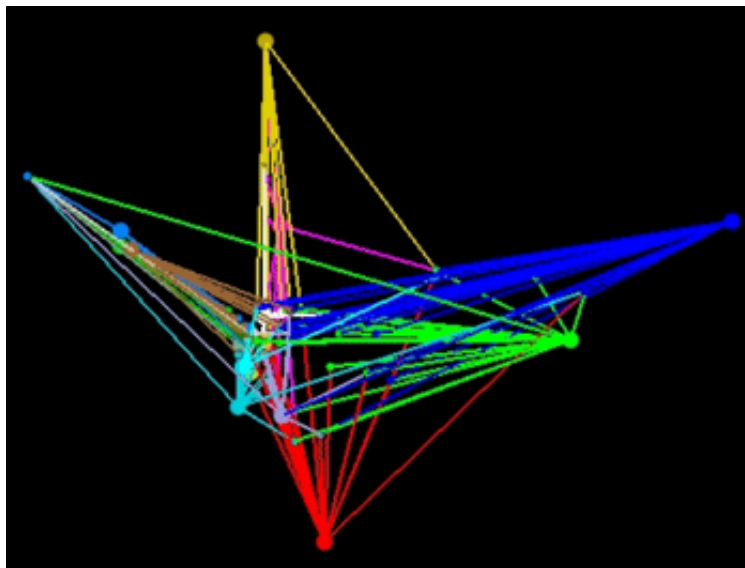


Figura 1: Artigos e Citações

Grupos de artigos "próximos" (clusters espaciais) podem indicar uma tendência de pensamento. Podemos, ainda, percorrer "temporalmente" a rede de citações, para chegarmos aos artigos de citação original (mais antigos) e verificarmos se tais artigos são mais citados diretamente (muitas arestas) ou indiretamente.

A partir daí, podemos montar uma rede de triângulos irregulares (TIN), de forma que as "montanhas" da superfície gerada indicam os artigos ou autores mais citados, e os "vales" os de menor relevância.





Figura 2: Superfície de Relevância de Artigos Científicos

Uma outra área de exploração das técnicas espaciais são os Sistemas de Recomendação, que consistem de técnicas para sugerir a usuários (ou clientes) itens de consumo, como por exemplo ocorre em sites de livrarias virtuais, como a Amazon. Para minorar os efeitos da sobrecarga de dados, existem tecnologias como o Filtro Colaborativo, que analisa e classifica o comportamento dos usuários, buscando localizar usuários parecidos entre si, que são utilizadas de forma combinada a outras técnicas de Data Mining.

Como alternativa à explosão combinatória que essas técnicas induzem, algumas estratégias de coesão e análise de grafos podem ser utilizadas, e é aí que o GIS tem muito a contribuir. No estabelecimento dos relacionamentos e das semelhanças entre os indivíduos para o Sistema de Recomendações, os algoritmos calculam distâncias temáticas entre os indivíduos e objetos, baseado em espaços cartesianos em que a geografia física é o que menos importa – as dimensões desse espaço representam outras variáveis de interesse. Nesses termos, a localização dos elementos significa descrição nas dimensões, e a vizinhança entre eles significa comportamento parecido. Assim, as técnicas de geração de superfícies, cálculo de declividade, e todos os conceitos de topologia de rede, como pesquisa de conectividade e melhor caminho, podem ser amplamente aplicados e o "GIS não geográfico" passa a fazer sentido.

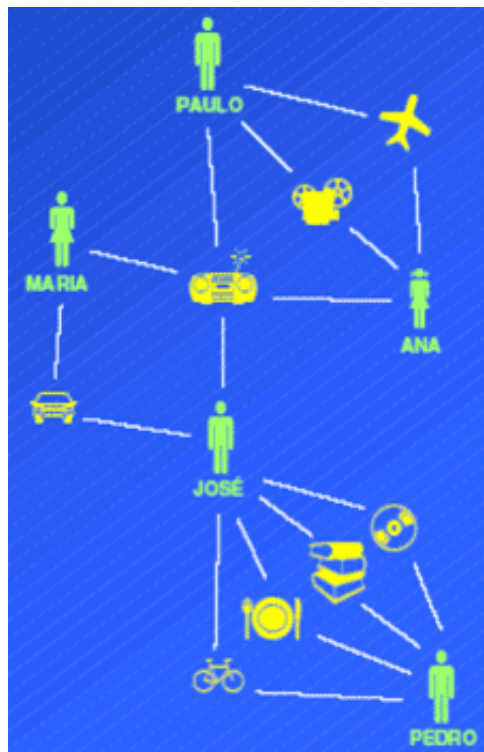


Figura 3: Rede de Pessoas e Itens em Sistemas de Recomendação

Dá para sonharmos mais ? Ou talvez estejamos exagerando no valor percebido do futuro potencial do GIS ? Bem, outros exemplos interessantes ficam para uma próxima oportunidade. É importante termos em mente que o compartilhamento de uma visão holística do problema passa pela quebra de paradigmas e pela abertura da mente para técnicas inovadoras ou apropriadas de outras áreas do conhecimento. Dessa forma, conseguiremos os diferenciais competitivos sustentáveis que o mercado nos exige, ou teremos condições de conhecer melhor o mundo que vivemos – geográfico e social.

Essa foi uma breve viagem pela evolução do GIS em praias não tradicionais. Sabemos que esse assunto não se esgota aqui e que as técnicas de "dividir para conquistar" rechearão cada vez mais de idéias nosso pensamento e de ações nossa prática, transformando o GIS no tempero que faltava para nossa refeição de sucesso. O céu é o limite.

### Referências Bibliográficas

ARANHA, Francisco. Data Mining em Grandes Redes: Superfícies de Coesão sobre Base Multidimensionalmente Escalonada, IME-USP, São Paulo, 2003.

ARANHA, Francisco; FRANCISCO, Eduardo; PONCHIO, Mateus; MOURA, Ana. Evaluating Relevance and Influence in Scientific Articles, BALAS Conference, IT Management Track, Madri, Espanha, 2005.

OLD, L. J. Using Spatial Analysis For Non-Spatial Data, ESRI Users Conference, San Diego, EUA, 2000.

### Eduardo de Rezende Francisco

Bacharel em Ciência da Computação pelo IME-USP, Mestrando em Administração (Métodos Quantitativos) pela EAESP-FGV, atua em GIS, Business Intelligence, Pesquisas de Mercado e

Estratégias de Marketing na AES Eletropaulo, é Consultor em integração Geomarketing & Data Mining e sócio-fundador da GITA Brasil.

[eduardo.francisco@aes.com](mailto:eduardo.francisco@aes.com)

[erfrancisco@hotmail.com](mailto:erfrancisco@hotmail.com)