
O Geomarketing com outros olhos

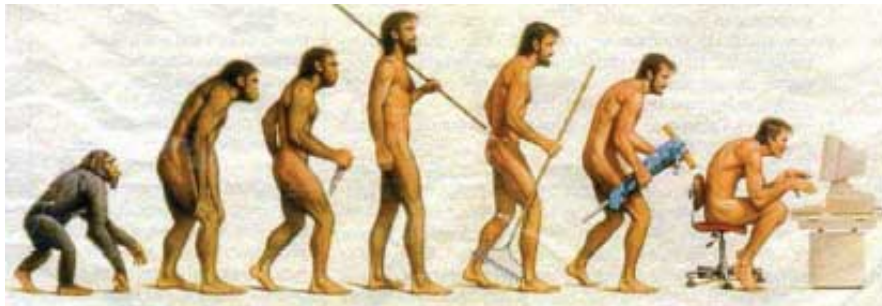
Por **MundoGEO** | 0100, 14 de Março de 2004

A Inteligência Competitiva está tornando indispensável o uso do GIS integrado à Gestão do Conhecimento – é a revanche do Spatial Data Mining

Qual a melhor região da cidade para se abrir um determinado negócio ou uma nova filial? Muitas empresas se propõem a responder com precisão esse tipo de pergunta, necessária para o sucesso dos empreendimentos comerciais de hoje. Áreas como logística e expansão das grandes corporações precisam dessa informação para planejar seu crescimento. Para as empresas de consultoria que abordam esse problema, é fundamental o conhecimento da área geográfica de atuação das lojas, seus concorrentes e aspectos do negócio em si, entre outras coisas.

E como essa pergunta é endereçada? De que maneira esse conhecimento se torna disponível? Através do tratamento adequado dos dados, que se tornam informações que, se integrados ao processo de gestão, apóiam a tomada de decisão, estabelecendo ações... E se tornam conhecimento. Essa boa prática nós já conhecemos. Mas ela tem funcionado?

Desde o início dos tempos o Ser Humano aprimora suas estratégias de sobrevivência. O que muda do Homo erectus para o Homo digitalis é a forma como ele gerencia seu conhecimento para liderar ou conquistar posições estratégicas e diferenciais competitivos.



No mundo atual, de acordo com a IBM, apenas 1% dos dados corporativos é utilizado efetivamente para análise nas grandes corporações. Não é à toa que cerca de 90% das estratégias corporativas fracassam, não por causa das estratégias em si, mas porque a execução dessas estratégias tem sido mal sucedida. Olhar para os dados corporativos de uma forma mais ampla, com alternativas e integrações, é questão fundamental. Um outro número importante: 75% dos custos de desenvolvimento de aplicações analíticas está na Camada de Integração.

Torna-se cada vez mais comum o uso do GIS aplicado ao mundo dos negócios. O Geomarketing está aí para usar e abusar das ferramentas geoespaciais, aliando informações (primárias e secundárias) à experiência dos profissionais na lida com os

mapas digitais. O mercado ainda não tem uma percepção ampla do uso integrado do Geomarketing a ferramentas mais tradicionais de gestão do conhecimento, mas a tendência de integração está aumentando.

Lembram-se do Spatial Data Mining? Pois é... Tivemos a oportunidade de conversar sobre ele nas edições 24 e 26 da InfoGEO. O SDM utilizado de forma complementar ao uso comum do Geomarketing nos negócios pode ser esse diferencial que as empresas estão buscando.

Esse processo de integração pode ser complexo, por envolver diferentes estruturas de dados, tipos, relacionamentos, métodos de acesso e mecanismos de inferência. O processo de descobrimento de padrões interessantes, previamente desconhecidos e potencialmente úteis, advindos de bancos que contenham dados espaciais, se dá através da segmentação por concentração espacial, da busca de padrões por proximidade, do aprendizado de relacionamentos posicionais não triviais e da utilização de GeoEstatística, entre outras técnicas.

O principal desafio deste processo é conectar o domínio do conhecimento de alto nível, geográfico, com as ferramentas usadas para buscar padrões de anomalias ou regularidades potencialmente interessantes dentro dos dados. Obviamente, dados incompletos e/ou inconclusivos dificultam a modelagem; a suposição de que há independência estatística entre os dados espacialmente distribuídos pode causar problemas para dados realmente causar problemas para dados realmente interrelacionados. Além disso, o custo computacional é elevado.

De qualquer forma, os desenvolvimentos atuais na área de modelos semânticos para objetos geoespaciais (aliás, muito bem abordados no último simpósio GEOINFO, em Campos de Jordão) ajudam a minimizar os problemas de modelagem.

Esse processo utiliza métodos estatísticos e de inteligência artificial para endereçar a solução. Alguns desses métodos estão na figura abaixo.



Métodos Utilizados pelo Spatial Data Mining

Para ilustrar como esses conceitos podem ser explorados na prática, vejamos algumas abordagens para responder a pergunta inicial (a abertura da nova loja). No exemplo, estamos partindo de uma base georreferenciada de estabelecimentos comerciais, com

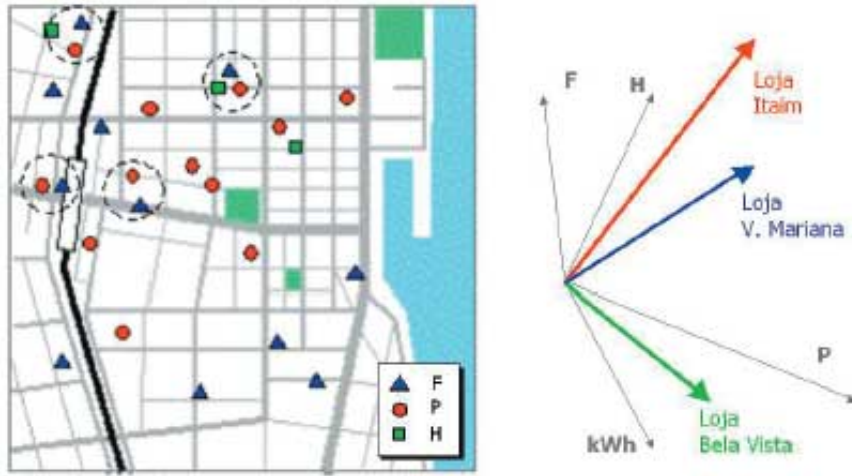
informações cadastradas de ramo de atividade e consumo (kWh) de energia elétrica (que pode ser substituído por qualquer variável quantitativa representativa – faturamento, por exemplo), e também de uma lista de endereços candidatos a serem a nova loja. Digamos que a proximidade do estabelecimento a farmácias [F], padarias [P] e hospitais [H] seja o mais relevante (essa informação pode ser advinda do próprio sistema de SDM, inclusive). Além disso, sabemos que a filial do bairro Itaim, por exemplo, é a mais rentável.

Uma técnica interessante é a aplicação de funções do tipo Neighboring Class Set. Nelas, podemos encontrar conjuntos de classes em que os objetos espaciais estejam próximos entre si (de acordo com um limite máximo de distância fornecido ou inferido). Assim, podemos verificar se existe algum padrão de proximidade (entre F e P, ou entre F, P e H, por exemplo) e como esse padrão se mantém nas regiões de interesse e ao longo do tempo.

Outra abordagem é o Vector Space Model. A idéia básica é a seguinte: começamos estabelecendo as dimensões (vetores) que queremos utilizar, e seus significados. Elas podem ser o número de F, P e H nas proximidades ou a distância efetiva da F, P e H mais próxima (para cada endereço candidato a nova loja). Em ambos os casos, uma quarta dimensão será o consumo médio mensal de energia (kWh) na região. Assim, os relacionamentos espaciais entre os elementos e as qualificações adicionais (consumo de energia) caracterizaram nossos objetos em quatro dimensões (F, P, H, kWh). As nossas lojas candidatas passaram a ter essa característica. Assim, por exemplo, Itaim passou a ser (8, 6, 4, 800), Vila Mariana (7, 8, 3, 1000), Bela Vista (3, 8, 1, 1200) e assim por diante. As distâncias calculadas podem ser lineares ou baseadas na malha de logradouros.

A partir daí, a técnica do Vector Space Model consiste na identificação da semelhança entre esses vetores n-dimensionais. Os vetores que estiverem próximos (em tamanho e distância angular) serão considerados parecidos e, além disso, os mais parecidos com a loja do Itaim serão considerados os melhores candidatos a se tornarem o local da nova loja.

Outras técnicas podem ser utilizadas. O mesmo princípio dos vetores n-dimensionais pode ser utilizado associado a algoritmos de agrupamento, caracterizando a Clusterização geoespacial. Já os Algoritmos Genéticos descobrem relações através do cruzamento de regras pré-existentes e da adaptação (seleção natural) das melhores. Eles podem trazer regras que nunca seriam descobertas, através do uso de "mutações" das relações estabelecidas. Problemas semelhantes podem ser facilmente abordados por essas técnicas de integração, que trazem o melhor dos dois mundos para um endereçamento mais adequado da solução.



Neighboring Class Set e Vector Space Model

Em suma, o grande objetivo dessa integração é encontrar soluções que tragam diferencial competitivo. Vale a pena prosseguir? A experiência nos mostra que sim. Cada vez mais as empresas buscam novas estratégias para a melhoria do desempenho nos negócios. Alternativas às implantações, cada vez mais comuns, de sistemas de BI, CRM, ERP, Call Center ativo, etc, são a palavra de ordem. A governança corporativa e os modelos de CMM (Capacity Maturity Model), em moda para a gestão de projetos, têm na informação (e sua manutenção) um aliado mais que necessário.

E o SDM está aí para isso. A gestão do conhecimento e a análise espacial trazem, de forma relacionada, o diferencial competitivo buscado. Caracterizar espacialmente as próprias informações que temos em casa é o ponto de partida dessa integração. Não se trata de buscar o novo, e sim de olhar para o "velho" (os próprios dados dos sistemas legados das companhias) com "outros olhos".

Eduardo de Rezende Francisco

Bacharel em Ciência da Computação pelo IME-USP e Mestrando em Administração (Métodos Quantitativos) pela EAESP – FGV

Atua em GIS, Business Intelligence e Estratégias de Marketing na AES Eletropaulo
Consultor em Integração Geomarketing & Data Mining

sócio-fundador da GITA Brasil

erfrancisco@hotmail.com

eduardo.francisco@aes.com

Colaborador

João Marcelo Arcoverde e a equipe da Match Simile, no suporte e parceria para o amadurecimento das idéias e conteúdos de SDM expostos.