

## DEFINIÇÃO DE UM MODELO DE DATA WAREHOUSE DE TRAJETÓRIAS PARA A INCORPORAÇÃO DAS INFORMAÇÕES SEMÂNTICAS DO AMBIENTE<sup>1</sup>

Cesar Lima Caetano<sup>2</sup>; Fernando José Braz<sup>3</sup>

### INTRODUÇÃO

Data warehouses tradicionais podem ser definidos como uma coleção de dados orientados a assuntos originários de diversas fontes de dados distintas. Entretanto, em um ambiente de objetos móveis, um data warehouse deve considerar informações não espaciais, espaciais e temporais, tornando-se um problema de grande complexidade. Além disso, tratando-se de trajetórias, os dados não podem ser agrupados por assunto, pois, em geral, eles constituem um conjunto de pontos localizados no tempo e no espaço. Portanto, os modelos tradicionais de data warehouse não suportam análises de dados espaçotemporais de trajetórias.

Ainda não existe um modelo padrão para a utilização da tecnologia de DW no armazenamento e na análise de dados de trajetórias de objetos móveis, mas há algumas propostas que trabalham com o armazenamento das informações das trajetórias com e sem a utilização de estruturas espaciais. As mais recentes propostas podem ser encontradas em Braz e Orlando (2007), Pelekis *et al.* (2008) e Almeida (2010). No trabalho de Braz *et al.* (2007) foi apresentado um modelo para data warehouse de trajetórias baseado em um conjunto de posições assumidas por um objeto ao longo do tempo representa a sua trajetória e é caracterizado individualmente por pontos  $(tid, x, y, t)$ , em que  $tid$  identifica a trajetória, o par  $x, y$  significa a posição espacial e  $t$  expressa o instante de tempo. No modelo multidimensional, cada célula (*cuboid*) representa as características de um conjunto de trajetórias que atravessa a célula.

A proposta de Pelekis *et al.* (2008) é bastante similar ao trabalho descrito anteriormente, até mesmo recorre ao modelo estrela. A estrutura da tabela de fatos mantida pela data warehouse é muito semelhante à proposta anterior. A principal

---

<sup>1</sup>O presente trabalho foi realizado com apoio do CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brasil

<sup>2</sup>Aluno do Instituto Federal Catarinense – Campus Araquari. Bacharelado em Sistemas de Informação. E-mail: cesar.trabalho7@gmail.com

<sup>3</sup>Professor Orientador do Instituto Federal Catarinense – Campus Araquari. E-mail: Fernando.braz@ifc-araquari.edu.br

diferença entre ambas reside nas dimensões utilizadas. A proposta de Pelekis também permite representar um conjunto de trajetórias como uma trajetória e reconstruir trajetórias com falhas. O modelo usa uma tabela de fatos e três tabelas de dimensões: espaço, tempo e perfil de objeto.

Almeida (2010) propôs um modelo semântico de data warehouse de trajetórias, bem mais rico em termos semânticos, se comparado com os anteriores. Este modelo de data warehouse é baseado no modelo conceitual de stops e moves, que representam, respectivamente, áreas de parada e movimentação de trajetórias. O modelo permite responder a questões do tipo: (i) Quantos objetos fazem paradas e deslocamentos entre dois bairros; (ii) Qual a velocidade média dos objetos; (iii) Qual a distância percorrida por objetos de um determinado tipo, dentro de um dado bairro.

Fica claro que ainda não existe um modelo de data warehouse que possa ser definido como padrão para o armazenamento de dados de trajetórias. Existem iniciativas isoladas que abordam características específicas de um ambiente de trajetórias. Por conta disto, a proposta principal deste projeto de pesquisa é desenvolver um modelo de data warehouse de trajetórias que consiga agregar todos os modelos apresentados anteriormente e que, além disso, apresente novos operadores ou funcionalidades que permitam incrementar a capacidade de análise de dados. Entre os novos operadores a serem desenvolvidos consideramos um para identificar e contar trajetórias com alto grau de similaridade (entre elas) dentro de uma determinada área espaço temporal, e outro para identificar, a partir da definição de um formato de trajetória, identificar e contar as trajetórias com alto nível de similaridade com o esboço apresentado. A manutenção do tipo de trajetória (passeio, trabalho, transporte público etc) também é um dos objetivos do novo modelo. Em resumo, o armazenamento da semântica tanto das trajetórias quanto do ambiente de movimentação é um dos principais objetivos deste projeto. Importante ressaltar que este data warehouse servirá como repositório para os dados produzidos em projetos que estão sendo desenvolvidos em outro campus.

## **PROCEDIMENTOS METODOLÓGICOS**

O projeto de pesquisa foi desenvolvido nas instalações do campus Araquari, fazendo uso dos equipamentos instalados nos laboratórios do curso de

Sistema de Informação. Foram três as fases do projeto: (1) a revisão das principais propostas de Data Warehouse de Trajetórias, (2) a definição de um modelo de data warehouse que englobe as principais propostas e (3) a definição das novas medidas e dimensões que permitam incrementar a capacidade de análise de dados fazendo uso da semântica das trajetórias e ambiente.

Na **primeira fase** aconteceu a revisão das principais propostas de data warehouse de trajetórias, visando identificar os modelos mais completos, partindo dos trabalhos apresentados na fundamentação teórica deste projeto. Como resultado da execução da fase anterior, as principais propostas (modelos) de data warehouse de trajetórias foram conhecidas. Na segunda fase aconteceu a efetiva incorporação de todos estes modelos em um único ambiente de data warehouse de trajetórias. Finalmente, na última fase do projeto, era previsto o efetivo desenvolvimento dos operadores que permitiriam identificar e contar as trajetórias com alto grau de similaridade, seja entre elas, seja comparando com um padrão fornecido pelo emitente da consulta. Além disso, nesta fase também seriam incorporadas as novas estruturas de dados que procuram representar a semântica das trajetórias e do ambiente em estudo.

## RESULTADOS E DISCUSSÕES

O progresso do projeto de pesquisa foi prejudicado pela saída antecipada do bolsista anterior, visto que uma boa parte das atividades que estavam previstas não puderam ser cumpridas a contento. Naquele momento, visto que este projeto é complementar a outro que também é desenvolvido na sala de pesquisa em Informática do campus Araquari, e que também aborda a questão de trajetórias, foi decidido que os dois projetos caminhariam em parceria, visando recuperar o tempo perdido. Desta forma, a meta de apresentar uma nova proposta de trajectory data warehouse não foi atingida, justamente por conta da alteração do rumo deste projeto. Os alunos envolvidos neste projeto tiveram a principal responsabilidade de estudar e entender os sistemas de informação geográficas (SIGs), visto que seriam amplamente utilizados na certificação dos demais projetos que aconteciam na mesma área.

## CONSIDERAÇÕES FINAIS

Visto que as metas não foram totalmente atingidas, justamente por conta da alteração do foco do projeto, os principais resultados até a presente etapa do projeto foram a disponibilização de um ambiente de SIG para a representação e identificação de conjuntos de trajetórias de objetos móveis. Este ambiente possibilitará a identificação e confirmação de trajetórias similares dentro de um conjunto de dados.

## REFERÊNCIAS

- ALMEIDA, C. A. *Data warehouse* de trajetórias: um modelo semântico com suporte à agregação por direção dos movimentos. 2010. Dissertação (Mestrado)–Universidade Federal de Campina Grande, 2010.
- BRAZ, F. J. Warehousing and mining aggregate measures concerning trajectories of moving objects. 2009. Tese (Doutorado)–Università Ca' Foscari, 2009.
- BRAZ, F. J. *et al.* Approximate aggregations in trajectory data warehouse. *In: INTERNATIONAL CONFERENCE ON DATA ENGINEERING, WORKSHOP ON SPATIO-TEMPORAL DATA MINING*, 2007, Istambul. Anais... Istambul, 2007. p. 536-545.
- BRAZ, F.; ORLANDO, S. Trajectory data warehouses: proposal of design and application to exploit data. *In: BRAZILIAN SYMPOSIUM ON GEOINFORMATICS*, 9., 2007, Campos do Jordão. Anais... Campos do Jordão, 2007. p. 61-72.
- PELEKIS, N. *et al.* Hands-on in Hermes MOD engine. 2011. Disponível em: <[http://infolab.cs.unipi.gr/hermes/Downloads/Hands\\_On\\_In\\_Hermes\\_MOD\\_Engine.pdf](http://infolab.cs.unipi.gr/hermes/Downloads/Hands_On_In_Hermes_MOD_Engine.pdf)>. Acesso em: maio 2012.
- PELEKIS, N. *et al.* Hermes: aggregative LBS via a trajectory DB engine. *In: SIGMOD CONFERENCE*, 2008, Vancouver. Anais... Vancouver, 2008. p. 1.255-1.258.
- SPACCAPIETRA, S. *et al.* A conceptual view on trajectories. *Data & Knowledge Engineering*, v. 65, p. 126-145, 2008.