

DESENVOLVIMENTO DE EXTENSÃO DO ALGORITMO *T-COLOC* PARA UTILIZAÇÃO DE INFORMAÇÕES DE CONTEXTO DO AMBIENTE¹

Lucas Eduardo Gomes²; Fernando José Braz³

INTRODUÇÃO

O desenvolvimento das tecnologias de mobilidade e o gerenciamento de localização permite compor o movimento dos objetos. O conjunto de posições assumidas por objetos móveis, ordenados pelo tempo, permite representar a trajetória de um objeto. A análise dos dados de trajetórias é de fundamental importância no entendimento do comportamento de objetos em um determinado ambiente. Embora existam algumas propostas de DW espaciais (Han et al., 1998, Rivest et al., 2001, Marchant et al., 2004, Shekhar et al., 2001), algumas destas propostas (Han et al., 1998) trabalham com um modelo de cubo de dados com dimensões espaciais e não espaciais. Entretanto, nenhuma delas aborda a situação de um objeto que se move continuamente ao longo do tempo. A modelagem de dados de trajetórias pode ser resumida em propostas para consultar posições (passado, presente e futuro) de objetos móveis (R. H. Gutting, Vazirgiannis, 2000), (O. Wolfson, Yesha, 1999), (A. Sistla, Dao, 1997). Entretanto, na construção de um DW para trajetórias é absolutamente fundamental explorar a dependência entre os dados. Os trabalhos citados anteriormente não consideram esta característica, as observações espaço-temporais são tratadas como pontos não relacionados. Os trabalhos mais recentes estendem os tradicionais modelos de DW com o objetivo de modelar e armazenar dados relativos a trajetórias de objetos móveis. Como os dados gerados por dispositivos móveis são dados brutos, eles precisam ser semanticamente enriquecidos para que seja possível analisá-los e extrair conhecimento interessante (Alvares et al., 2007b). A ideia de agregar semântica às trajetórias de objetos móveis para facilitar a análise e mineração deste tipo de dado é inovadora que foi introduzida em (Alvares et al., 2007b). Este trabalho foi baseado

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

²Aluno do Instituto Federal Catarinense – Campus Videira. Curso Técnico Integrado em Informática. E-mail: luk_eduardo@hotmail.com

³Professor Orientador do Instituto Feral Catarinense – Campus Araquari. Curso Técnico Integrado em Informática. E-mail: fernando.braz@ifc-araquari.edu.br

no modelo conceitual para trajetórias elaborado pelo grupo da EPFL em 2006 e que foi publicado em 2008 (Spaccapietra et al. 2008). Uma série de outros trabalhos envolvendo a semântica na mineração de dados foram também apresentados recentemente (Bogorny et al., 2006a), (Bogorny et al., 2006c), (Bogorny et al., 2006b), (Bogorny et al., 2007), (Bogorny et al., 2008a), (Bogorny et al., 2009). Portanto, a proposta deste trabalho foi investigar a utilização das técnicas de mineração de dados para a análise dos resumos de dados das trajetórias de um conjunto de objetos móveis mantidos em ambientes de Data Warehouse (DW), fazendo uso das informações de contexto do ambiente. O algoritmo de mineração de trajetórias de objetos móveis, *T-Coloc*, foi desenvolvido pelo orientador deste projeto durante seu período de doutoramento. A versão atual do algoritmo não contempla a utilização das informações de contexto do ambiente. Objetivou-se, portanto, desenvolver uma nova versão do algoritmo de forma a englobar o tratamento destas informações.

PROCEDIMENTOS METODOLÓGICOS

O desenvolvimento do presente projeto está envolvido nas atividades de Pós Doutorado do orientador do projeto (Fernando José Braz), que já acontece desde 2011 sob a orientação da professora Vania Bogorny (UFSC/SC).

O professor Fernando desenvolveu, durante seu período de doutoramento, trabalho focado basicamente na pesquisa de tecnologias para o armazenamento de conjuntos de dados de trajetórias de objetos móveis, avançando na consequente utilização das técnicas de mineração de dados aplicadas sobre aqueles conjuntos de dados.

Já a professora Vania, desenvolve pesquisa aprofundada na área de mineração de dados e aplicação da semântica sobre ambientes de trajetórias, é autora de diversas publicações nestas áreas, que apresentaram técnicas inovadoras e inéditas no sentido de inserir as informações de contexto nas análises de comportamento/padrões de trajetórias de objetos móveis. Ambos trabalharam, em épocas distintas, no mesmo grupo de pesquisa da comunidade europeia (GEOPKDD), formatado com o objetivo de desenvolver pesquisa na área de trajetórias de objetos móveis.

No projeto de pós-doutoramento a proposta é utilizar as informações de semântica (com base no trabalho desenvolvido pela professora Vania) no modelo de TDW proposto pelo professor Fernando. Além disso, a partir do desenvolvimento de uma extensão do modelo de TDW para fazer uso da semântica do ambiente, evoluir para a adaptação e desenvolvimento de algoritmos de mineração de dados aplicados sobre aquele conjunto de dados. Vale aqui lembrar que este seria também o início de um trabalho inédito nesta área do conhecimento, não existe ainda qualquer proposta de trabalho similar acontecendo atualmente na área correlata.

RESULTADOS E DISCUSSÕES

Apesar de não estar concluída a versão do algoritmo que considere as informações de contexto com base na versão original do *T-Coloc*. A pesquisa e o estudo desenvolvido dentro deste projeto propiciou a obtenção de uma proposta de arquitetura para o armazenamento das informações de contexto em um DW. Este resultado já é de fundamental importância para a pesquisa desenvolvida na área, visto que ainda não existe um consenso sobre um modelo de armazenamento e representação do contexto de um ambiente de trajetórias em um DW. Além disso, partindo deste ambiente de armazenamento, a nova versão do algoritmo poderá ser desenvolvida com maior facilidade, visto que já existiria uma plataforma sobre a qual seriam desenvolvido e aplicado o algoritmo.

CONSIDERAÇÕES FINAIS

A revisão do algoritmo *T-Coloc* propiciou resolver o gargalo da performance do mesmo, resultando na possibilidade efetiva da aplicação do algoritmo. Por conta disto, uma primeira versão de um aplicativo que recebe, processa, armazena e disponibiliza os dados de um DW considerando informações de contexto já foi concluído. Este aplicativo envolve o acesso a dados de um ambiente de navegação pesqueira, em outro projeto paralelo, oferecendo a possibilidade de gerar conhecimento relacionado com aquelas atividades. Seguramente, uma grande gama de aplicações pode ser atingida a partir da disponibilização e domínio desta tecnologia. Os impactos da execução deste projeto envolvem desde o desenvolvimento de tecnologia até os reflexos diretos na

qualidade de vida e meio ambiente da população envolvida pelas ações das possíveis aplicações. Na questão do gerenciamento do tráfego urbano, por exemplo, a utilização desta tecnologia permite gerir de maneira racional e inteligente os eventos de movimentação encontrados em ambientes de alta concentração de objetos móveis. A possibilidade de oferecer este tipo de ferramental tem impacto direto na redução de poluentes, na economia de combustível, na redução de recursos gastos para o controle do tráfego e ainda na oferta de conhecimento interessante para o planejamento da rede de tráfego.

REFERÊNCIAS

Alvares, L. O., Bogorny, V., Kuijpers, B., de Macedo, J. A. F., Moelans, B., and Vaisman, A. (2007b). A model for enriching trajectories with semantic geographical information. In ACM-GIS, pages 162–169, New York, NY, USA. ACM

Bogorny, V., Camargo, S., Engel, P., and Alvares, L. O. (2006a). Mining frequent geographic patterns with knowledge constraints. In ACM-GIS, pages 139–146, New York, NY, USA. ACM Press.

Bogorny, V., Camargo, S., Engel, P., and Alvares, L. O. (2006b). Towards elimination of well known patterns in spatial association rule mining. In Proceeding of the IEEE International Conference on Intelligent Systems, IS, pages 532–537. IEEE Computer Society.

Bogorny, V., Engel, P., and Alvares, L. O. (2007). Enhancing the Process of Knowledge Discovery in Geographic Databases using Geo-Ontologies. Idea Group.

Bogorny, V., Kuijpers, B., and Alvares, L. O. (2008a). Reducing uninteresting spatial association rules in geographic databases using background knowledge: a summary of results. *International Journal of Geographical Information Science*, 22:361–386.

Bogorny, V., Valiati, J., Camargo, S., Engel, P., Kuijpers, B., and Alvares, L. O. (2006c). Mining maximal generalized frequent geographic patterns with knowledge constraints. In ICDM, pages 813–817. IEEE Computer Society.

Han, J., Stefanovic, N., and Kopersky, K. (1998). Selective materialization: An efficient method for spatial data cube construction. *PAKDD'98*, pages 144–158.

Marchant, P., Brisebois, A., Bedard, Y., and Edwards, G. (2004). Implementation and evaluation of a hypercube-based method for spatiotemporal exploration and analysis. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59:6–20.

O. Wolfson, A. Sistla, S. C. and Yesha, Y. (1999). Updating and querying databases that track mobile units. *Distributed and Parallel Databases*, 7(257–387).

Press.Spaccapietra, S., Parent, C., Damiani, M. L., de Macedo, J. A., Porto, F., and Vangenot, C. (2008). A conceptual view on trajectories. *Data and Knowledge Engineering*, 65(1):126–146.

R. H. Güting, M. H. Boehlen, M. E. C. S. J. N. A. L. M. S. and Vazirgiannis, M. (2000). A foundation for representing and quering moving objects. *ACM TODS*, 25(1):1–42.

Rivest, S., Bedard, Y., and Marchand., P. (2001). Towards better support for spatial decision making: Defining the characteristics of spatial on-line analytical processing(solap). *Geomatica*, 55(4):539–555.

Shekhar, S. and Huang, Y. (2001). Discovering spatial co-location patterns: A summary of results. *Lecture Notes in Computer Science*, 2121:236+.

Sistla, O. Wolfson, S. C. and Dao., S. (1997). Modeling and querying moving objects. In *ICDE'97*, pages 422–432.