

Conectando Dados de Movimento Textualmente Anotados a Dados Ligados

Cleto May¹, Renato Fileto¹

¹Departamento de Informática e Estatística,
Universidade Federal de Santa Catarina (UFSC), Florianópolis, SC, Brasil

`cleto.may@inf.ufsc.br, r.fileto@ufsc.br`

Abstract. *The recent progress in movement analysis mostly considers spatio-temporal data (e.g. trajectories). However, other information (e.g., tags and comments associated to the spatiotemporal data) could also help to better understand movements (e.g., identify classes of frequented places and events). This article introduces a method for semantic enriching textually annotated movement data (e.g., trajectories, user's trails in social media) by connecting them to Linked Data. The current version of our method uses spatial and textual similarity to find and rank possible connections. Real data from Flickr and LinkedGeoData have been used to test the method implementation.*

Resumo. *O progresso recente na análise de movimento considera principalmente dados espaço-temporais (e.g. trajetórias). Contudo, outras informações (e.g., tags e comentários associados a dados espaço-temporais) também poderiam ajudar a explicar movimentos (e.g., identificar classes de locais e eventos frequentados). Este artigo introduz um método para enriquecer semanticamente dados de movimento textualmente anotados (e.g., trajetórias, trilhas de usuários em mídias sociais) conectando-os a dados ligados. A versão atual do nosso método usa similaridade espacial e textual para encontrar e ranquear possíveis conexões. Dados reais do Flickr e LinkedGeoData foram utilizados para testar a implementação do método.*

1. Introdução

A última década contou com grande popularização de dispositivos móveis (e.g., smartphones) e outros equipamentos dotados de sensores (e.g., GPS, GSM, RFID, câmeras) capazes de registrar movimento. Diversas aplicações podem ser suportadas pela enorme quantidade de dados coletados pelo uso de tais tecnologias. Todavia, isso requer técnicas apropriadas para extrair informação desses dados.

Dados de movimento são sequências temporalmente ordenadas de posições ocupadas por objetos que se movem. Cada amostra de posição de um objeto móvel inclui coordenadas geográficas, o momento em que o objeto ocupa tal posição e possivelmente anotações associadas aos dados espaço-temporais (e.g., tags, comentários). Este trabalho utiliza o termo *dados de movimento* como uma generalização para trajetórias [Parent et al. 2013] e trilhas de usuários de mídias sociais (e.g., Twitter, Facebook, FourSquare). As primeiras usualmente são colhidas por sensores e aplicativos específicos para tal finalidade, enquanto as últimas são sequências de postagens de um usuário em uma mídia social. Devido a própria forma de coleta de dados, trajetórias costumam ter boa

precisão espaço-temporal, ao passo que trilhas de redes sociais costumam ser esparsas, devido a característica assíncrona das postagens dos usuários em redes sociais. Por outro lado, trilhas de redes sociais costumam ser ricas em informações textuais, enquanto trajetórias raramente possuem anotações.

O desenvolvimento de sistemas que capturam diversos aspectos semânticos do movimento (e.g., locais visitados, razões de movimentos) ainda é um desafio, mesmo quando a entrada de tais sistemas inclui dados anotados textualmente. Este trabalho visa contribuir para preencher tal lacuna, propondo técnicas para conectar dados de movimento anotados textualmente com recursos de coleções de dados ligados disponíveis na Web Semântica, para auxiliar a análise do movimento segundo diversos aspectos semânticos.

[Fileto et al. 2013] mostra os benefícios da utilização de grandes coleções de dados ligados atualmente disponíveis, aderentes a padrões e com semântica bem definida, na análise de dados de movimento. Porém, não resolve adequadamente o problema da conexão entre os dados de movimento e os dados ligados.

As principais contribuições deste artigo são: (i) proposição de um método para conectar dados de movimento a recursos de dados ligados, utilizando informações espaciais e textuais; (ii) ordenação (*ranking*); e (iii) teste da implementação do método proposto em experimentos com dados reais.

As próximas seções deste artigo estão organizadas da seguinte forma. A Seção 2 apresenta alguns fundamentos. A Seção 3 descreve o método proposto. A Seção 4 discute os experimentos realizados. A Seção 5 estuda o estado da arte de problemas relacionados à conexão entre dados de movimento e dados ligados. Finalmente, a Seção 6 sumariza os resultados obtidos até aqui e enumera trabalhos futuros.

2. Fundamentos

Esta seção apresenta os fundamentos necessários ao entendimento do problema tratado e do método de solução proposto.

2.1. Dados Ligados

Dados ligados abertos, do inglês *Linked Open Data*¹ (LOD), surgem da necessidade de interligar com semântica bem definida e tornar acessível dados disponíveis na Web. Coleções de LOD são estruturadas como triplas RDF² da forma recurso-propriedade-valor [Antoniou and Harmelen 2008]. Um recurso é identificado unicamente por uma URI a partir da qual podem ser acessadas suas propriedades. Um recurso pode ter diversas propriedades (e.g., tipo, nome, rótulo, descrição). As possíveis propriedades de um recurso variam de acordo com o tipo (e.g., uma pessoa pode ter data e local de nascimento, um local pode ter coordenadas geográficas). O valor de uma propriedade pode ser outro recurso (e.g., o valor da propriedade local de nascimento é a URI do respectivo local) ou literal (conjunto de caracteres ou número).

2.2. Dados de movimento

Uma sequência bruta de dados de movimento, em inglês *Raw Movement Data*, é uma sequência de amostras de posições de um objeto móvel, cada qual coletada em um

¹<http://linkeddata.org/>

²<http://www.w3.org/TR/rdf-primer/>

instante no tempo. A definição 1 formaliza tal conceito.

Definição 1. Sequência bruta de dados de movimento *RawMD* é uma sequência temporalmente ordenada de amostras de posições $p_1, \dots, p_n$ de objetos móveis. Cada posição p_i tem a forma $p_i((x_i, y_i), t_i, S_i)$ onde:

- (x_i, y_i) são coordenadas geográficas;
- t_i é um instante de tempo; e
- $S_i = \{s_1, \dots, s_l\}$ é uma coleção de valores textuais de atributos associados à amostra de posição espaço-temporal (e.g., palavras-chave, *tags*).

A Figura 1 ilustra uma sequência bruta de dados de movimento. Cada amostra de posição está representada por ponto preto.

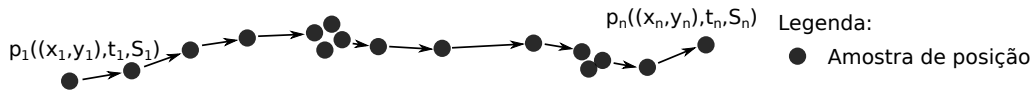


Figura 1. Sequência bruta de dados de movimento.

Uma sequência estruturada de dados de movimento, em inglês *Structured Movement Data*, é uma sequência temporalmente ordenada de episódios. Consideramos cada episódio como um segmento maximal não aninhado de amostras do movimento que cumprem um predicado (e.g., não se mover mais que uma dada distância durante um dado período de tempo) [Buchin et al. 2010]. Os predicados usados para identificar um episódio variam de acordo com a classe de episódio e a aplicação. A Definição 2 formaliza o conceito de sequência estruturada de dado de movimento.

Definição 2. Sequência estruturada de dados de movimento é uma sequência temporalmente ordenada de episódios $E_1, \dots, E_m$. Cada episódio E_i tem a forma $E_i(c_i, RawMD_i)$ onde:

- c_i é a classe do episódio (e.g., *stop*, *move* [Alvares et al. 2007]); e
- $RawMD_i$ é um segmento maximal de uma sequência bruta de dados de movimento que cumprem um predicado que determina o episódio (e.g., limite de velocidade, intervalo de tempo).

A Figura 2 ilustra uma sequência estruturada de dados de movimento. Episódios das classes *AltaVelocidade* e *BaixaVelocidade* estão representados por balões vermelhos e azuis, respectivamente.

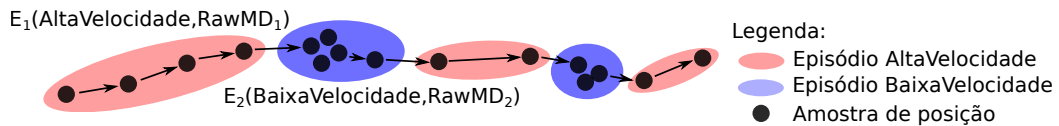


Figura 2. Sequência estruturada de dados de movimento.

Existem casos em que sequências de dados de movimento não são suficientes para responder consultas tais como: “Quais objetos móveis estiveram em locais turísticos?”. Para responder tal tipo de consulta é necessário que cada episódio precisamente esteja anotado semanticamente com as classes dos locais visitados. Preferencialmente, os

dados de movimento precisam ser conectados a itens de informação (e.g., um hotel, um restaurante, um local de interesse turístico) com definições precisas e com semântica bem definida, tais como recursos presentes em coleções de dados ligados [Fileto et al. 2013]. A Definição 3 formaliza o conceito de sequência semântica de dados de movimento.

Definição 3. Sequência semântica de dados de movimento é uma sequência temporalmente ordenada de episódios semanticamente anotados (i.e., conectados a recursos específicos presentes em coleções de dados ligados) $SE_1, \dots, SE_m$. Cada episódio semanticamente anotado SE_i tem a forma $SE_i(E_i, A_i)$, onde:

- E_i é um episódio da forma descrita na Definição 2; e
- A_i é uma coleção de anotações semânticas da forma $a(p, v)$, onde p é a propriedade que conecta o episódio E_i ao recurso v de uma coleção de LOD.

A Figura 3 ilustra uma sequência de episódios semanticamente anotados. As anotações estão representadas por balões verdes.

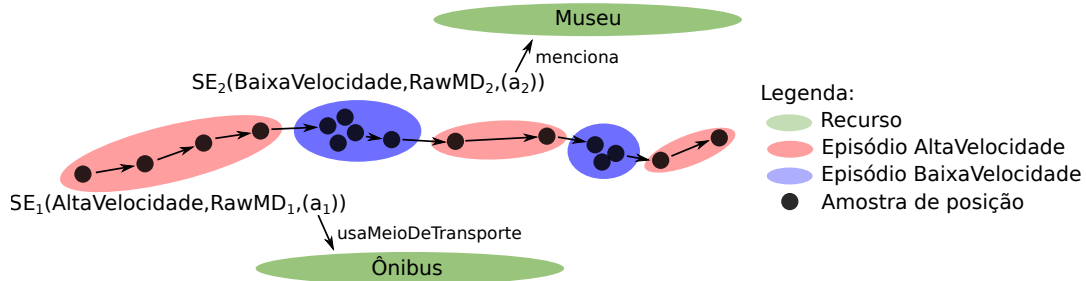


Figura 3. Sequência semântica de dados de movimento.

2.3. Problema Abordado

A Figura 4 exemplifica o problema de conexão entre dados de movimento e LOD tratado neste artigo. O canto inferior direito mostra uma imagem de satélite da região e o restante da figura uma foto daquela região, ambas obtidas do Google Maps³. O balão sobre a imagem de satélite representa a posição de uma postagem no Flickr⁴ que foi associada às palavras-chave apresentadas acima do balão amarelo (wheel, violet, toy, roda, etc.).

Nota-se que a posição do objeto móvel encontra-se em uma área densa da cidade. Além disso, a precisão da posição não é suficiente para inferir o local visitado pelo objeto móvel. A descrição textual auxilia tal tarefa. Considerando os rótulos e as descrições dos locais dentro de um certo raio de distância do ponto da postagem, pode-se concluir que o estabelecimento denominado Bike Dream é um forte candidato. O título e a descrição do recurso denominado Bike Dream estão léxicamente e semanticamente mais relacionados com as palavras-chave associadas a tal postagem, que aqueles de outros estabelecimentos ao redor e descritos na mesma coleção de LOD.

O objetivo deste trabalho é realizar a conexão entre dados de movimento e LOD de forma automática. Pode-se utilizar para isso informações espaço-temporais e textuais, assim como similaridade semântica e informações de contexto (e.g. características do objeto móvel), para encontrar conexões de menções a entidades nomeadas associadas aos dados de movimento com recursos de LOD (e.g., estabelecimentos, eventos).

³<https://maps.google.com/>

⁴<https://www.flickr.com/>



Figura 4. Conexão em área densa na cidade de Florianópolis

3. Método Proposto

O método proposto neste artigo almeja o enriquecimento semântico de dados de movimento textualmente anotados através da conexão a recursos de LOD. Sua versão atual utiliza proximidade espacial e textual, e é formalmente descrita pelo Algoritmo 1. As entradas para o método proposto são a sequência bruta de dados de movimento, uma coleção de dados ligados referente à mesma região e período de tempo que os dados de movimento e os limiares de proximidade espacial e similaridade léxica para efetuar as ligações. O resultado é a sequência semântica de dados de movimento, anotada com recursos da coleção de LOD fornecida como entrada.

Algoritmo 1: Função conecta

Entrada: *RawMD* //Sequência bruta de dados de movimento
LOD //Coleção de LOD
 τ_s //Limiar espacial
 τ_t //Limiar textual
Saída: *SemMD* //Sequência semântica de dados de movimento

```

1  início
2      StrMD ← estruturaDadosDeMovimento(RawMD)
3      para i = 0 até |StrMD| – 1 faça
4          e ← StrMD[i]
5          rProximos ← filtraEspacialmente(e, LOD,  $\tau_s$ )
6          eliminaCaracteresEspeciais(e)
7          eliminaCaracteresEspeciais(rProximos)
8          rankE ← filtraLexicamente(e, rProximos,  $\tau_t$ )
9          A ←  $\emptyset$ 
10         para cada r ∈ rankE faça
11             A ← A ∪ menciona(r.recurso, r.proximidade)
12         fim
13         SemMD[i] ← SE(e, A)
14     fim
15     retorna SemMD
16 fim

```

Inicialmente, estruturamos a sequência bruta de dados de movimento e submete-

mos os episódios resultantes ao enriquecimento semântico. A estruturação é flexível, de acordo com a aplicação. Por exemplo, aplicações que desejam identificar pontos turísticos visitados por turistas podem utilizar a velocidade baixa para detectar episódios relevantes (e.g., admirar uma construção).

A função **filtraEspacialmente**, linha 5, seleciona todos os recursos da coleção *LOD* fornecida que estejam localizados a uma distância igual ou inferior a τ_s do episódio *e*. Utilizamos o centroide do episódio para medir a distância a recursos e simplificar o processamento, mas pode-se também utilizar outros critérios (e.g., centro de massa). A Figura 5 ilustra um exemplo de aplicação do filtro espacial em que os recursos r_1 e r_2 são selecionados por estarem a uma distância inferior a τ_s do centroide c_2 do episódio e_2 . Para simplificar, abstraímos as anotações textuais na imagem. Uma extensão para o filtro espacial seria a inclusão da compatibilidade temporal (e.g., momento do episódio compatível com o horário de funcionamento de um local ou evento), tornando-o então um filtro espaço-temporal.

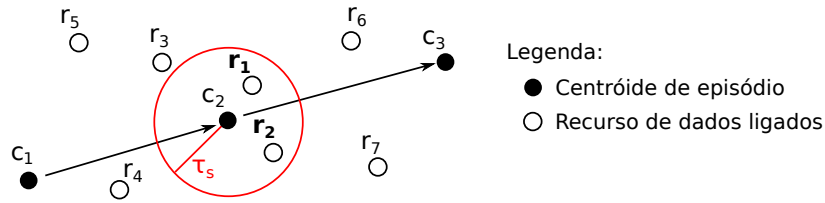


Figura 5. Filtro espacial

Nas linhas 6 e 7, a função **eliminaCaracteresEspeciais**, elimina das anotações textuais caracteres como asteriscos, parênteses, chave e barras, pois estes podem prejudicar a aplicação de funções de similaridade textual.

A função **filtraLexicamente**, presente na linha 8, analisa o conjunto de termos de cada par episódio-recurso para encontrar os recursos que são similares textualmente utilizando a função SoftTFIDF [Cohen et al. 2003]. Ela é adequada para o método pois combina funções de similaridade textual entre conjuntos de palavras (e.g., modelo vetorial) e palavras individuais (e.g., Jaro-Winkler). Os recursos que possuem similaridade textual superior ao limiar τ_t são selecionados. A Figura 6 ilustra a aplicação da função de similaridade textual entre um episódio (e_1) e dois recursos (r_1 e r_2) de coleções de dados ligados. A função é aplicada às anotações textuais de cada par episódio-recurso. Pode-se perceber que no exemplo o recurso r_1 possui a anotação textual “Bike” muito semelhante a anotação “bike” presente no episódio e_1 . Então r_1 é selecionado. Não podemos dizer o mesmo do recurso r_2 . Isso depende do limiar textual utilizado.

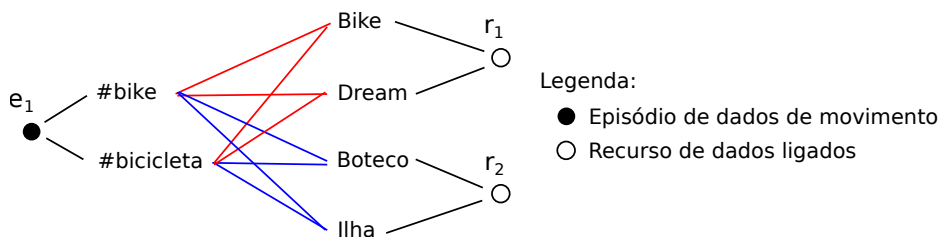


Figura 6. Investigação de correspondências léxicas

Finalmente, após encontrar os recursos próximos espacial e textualmente, na linha 11 são instanciadas todas as anotações semânticas com o respectivo recurso encontrado. Em seguida, na linha 13, o novo episódio conectado é acrescido a sequência semântica de dados de movimento. Novamente, uma extensão do método proposto seria a identificação da melhor propriedade para a conexão realizada, diferente do atual que adiciona a propriedade *menção* a todas as conexões.

4. Experimentos

As sequências brutas de dados de movimento utilizadas na experimentação foram extraídas do CoPhIR⁵. Tais dados referem-se a fotos e dados associados (posição, momento, tags, etc.) publicadas na mídia social Flickr⁶ contidas no Brasil nos anos de 2005 a 2007. Utilizaremos tags associadas a cada ponto amostrado do movimento como sua descrição textual. Consideramos cada episódio descrito por todas as tags associadas a ponto de amostragem espaço-temporal que dele façam parte.

Para enriquecer semanticamente os dados de movimento oriundos do Flickr utilizamos os dados ligados do LinkedGeoData⁷ os quais foram triplificados do OpenStreetMap⁸ (ferramenta da coleta colaborativa de dados geográficos). Utilizamos os rótulos dos recursos para investigar as conexões léxicas com tags associadas aos dados de movimento.

Os experimentos práticos foram realizados em uma máquina com processador Intel(R) Core(TM) 2 Quad 2.40GHz, com 4Gb de memória RAM e um disco rígido de 500 Gb 7200 RPM. O algoritmo para efetuar a conexão semântica foi desenvolvido utilizando a linguagem Java além dos dados espaço-temporais coletados e armazenados em uma base de dados PostgreSQL, utilizando a extensão geográfica PostGIS.

4.1. Resultados

Submetemos 36.476 pegadas de usuários do Flickr ao algoritmo. O método foi capaz de realizar 9.598 conexões utilizando os parâmetros $\tau_s = 1000$ metros e $\tau_t = 0,2$. Parâmetros estes que foram escolhidos após um estudo simples da base dados utilizada. Em trabalhos futuros pretendemos utilizar diferentes parâmetros e comparar os resultados, além de avaliar a qualidade das conexões realizadas. A Tabela 1 ilustra algumas conexões realizadas. O tempo de execução médio foi de 1,3 segundos para cada episódio analisado. O ganho de performance é acrescido quando utiliza-se uma base de dados ligados local, pois cada requisição a uma coleção de dados ligados é custosa.

4.2. Discussão

Após analisar manualmente as conexões realizadas, percebemos relevantes conexões realizadas. Por exemplo, a Figura 7 ilustra uma amostra de dado de movimento (balão amarelo) e a sua conexão com um recurso (balão azul). Nota-se que o recurso conectado não precisa necessariamente estar próximo ao dado de movimento. No entanto, a sua conexão ainda é verdadeira, inclusive comum na base de dados utilizada, pois como se trata de fotos o objeto móvel pode estar localizado a uma certa distância do alvo fotografado. No exemplo, a fotografia foi tirada próxima ao morro do Corcovado, mas o objeto

⁵<http://cophir.isti.cnr.it>

⁶<http://www.flickr.com>

⁷<http://www.linkedgeodata.org>

⁸<http://www.openstreetmap.org>

Tabela 1. Experimentos - Visão textual

Identificador do episódio	Palavras-chave do dado de movimento	URI do recurso	Rótulos
50344933	riodejaneiro, corcovado, brazil, brasil	lgd:/triplify/node1551149888	Corcovado
28567054	lovelyphotos, fortetacabana, forte, copacabana	lgd:/triplify/node2308298808	Forte de Copacabana
107131876	rio, jardim, janeiro, garden, de, coutinho, by, bothanical, cotanico	lgd:/triplify/node1613993266	Jardim Botânico
40774808	viradacultural, virada, teatromunicipaldesãopaulo, teatromunicipal, teatro, são paulo, são paulo, municipal, cultural	lgd:/triplify/way46927931	Teatro Municipal de São Paulo
72964391	pezinho, foot, copacabana, brazil, beach	lgd:/triplify/way179496798	Praia de Copacabana
37508938	voador, rio, janeiro, dish, deep, de, circo	lgd:/triplify/node331382172	Circo Voador

móvel não estava no local. A conexão ilustrada pela Figura 4 também foi realizada com um recurso relacionado com as tags associadas a amostra de dado de movimento, mesmo estando tal recurso em uma região densa de pontos de interesse.



Figura 7. Experimentos - Visão geográfica

As conexões realizadas sugerem que informações espaciais e textuais são relevantes. No entanto, somente essas informações não são suficientes em todos os casos. Outras técnicas (e.g., aprendizado de máquina) e outras informações (e.g., informações temporais e de contexto) poderiam ser utilizadas para potencializar a qualidade das conexões.

5. Trabalhos Relacionados

O problema de conectar dados de movimento a dados ligados foi inicialmente definido em [Fileto et al. 2013]. Tal trabalho apontou os benefícios desta abordagem para o enriquecimento semântico de dados de movimento e a necessidade de desenvolver técnicas eficazes e eficientes para efetuar tal conexão.

Uma linha de pesquisa recente que ajuda na definição formal e solução eficiente do problema aqui tratado propõe maneiras de fazer a junção por similaridade entre bases de dados espaciais textualmente anotadas. [Ballesteros et al. 2011] propõe o cálculo para junção espaço-textual como sendo a razão entre similaridade textual (coeficiente de Jaccard) e espacial (distância Ortodromia). [Bouros et al. 2012] e [Liu et al. 2012] calculam as similaridades independentemente e consideram similares elementos que possuem valores de similaridade espacial e textual superiores aos seus respectivos limiares.

Outra linha de pesquisa, conhecida por *Entity Linking*, consiste em detectar em um texto menções a entidades de uma base de conhecimento. Um recente trabalho que define o problema e propõe sua solução é [Ceccarelli et al. 2013]. No entanto, em tal trabalho não são consideradas as informações espaço-temporais.

O diferencial do trabalho submetido quando comparado a trabalhos relacionados está na utilização de conhecimentos de junção por similaridade de bases de dados espaciais textualmente anotadas na análise de dados de movimento, identificando menções a entidades presentes em coleções de dados ligados.

6. Conclusões e Trabalhos Futuros

O enriquecimento semântico de dados de movimento é uma questão bastante discutida atualmente na literatura [Parent et al. 2013][Yan et al. 2013]. O potencial do enriquecimento semântico utilizando conexões a coleções de dados ligados merece atenção devido a estruturação com semântica bem definida que tais dados possuem. Este artigo propõe um método para realizar o enriquecimento semântico de dados de movimento textualmente anotados através da conexão a recursos provenientes de coleções de dados ligados. Suas principais contribuições são: (i) critérios para filtragem e ranqueamento de recursos de dados ligados com dados de movimento, baseados na distância espacial e similaridade léxica entre anotações do movimento e atributos textuais de dados ligados usando a medida SoftTFIDF; (ii) implementação e teste do método proposto com dados de movimento reais obtidos de postagens em mídia social e dados ligados. Os resultados obtidos mostram que a proposta é viável e promissora, embora muitas pesquisas ainda sejam necessárias para aprimorar e validar métodos de enriquecimento semântico de dados de movimento através da conexão com recursos de coleções de dados ligados.

Em trabalhos futuros planeja-se: (i) estender o método proposto com o uso de informações de contexto, similaridade semântica e técnicas como aprendizado de máquina, para ranquear as conexões candidatas identificadas pelo uso de similaridade espaço-temporal e textual; (ii) analisar a qualidade dos resultados gerados pelo método proposto com diferentes coleções de dados de movimento e dados ligados; (iii) estender o método proposto para determinar propriedades de conexões realizadas (e.g., *meio de transporte, local, evento*); e (iv) aprimorar o método proposto para execução mais eficiente sem comprometer a qualidade dos resultados.

Agradecimentos. Este trabalho contou com o apoio do projeto European Union's IRSES-SEEK (concessão 295179), do CNPq (concessão 478634/2011-0) e da FEESC.

Referências

- Alvares, L. O., Bogorny, V., Kuijpers, B., de Macedo, J. A. F., Moelans, B., and Vaisman, A. (2007). A model for enriching trajectories with semantic geographical information. In *Proceedings of the 15th annual ACM Intl. Symp. on Advances in geographic information systems*, GIS '07, pages 22:1–22:8, New York, NY, USA. ACM.
- Antoniou, G. and Harmelen, F. v. (2008). *A Semantic Web Primer, 2Nd Edition (Cooperative Information Systems)*. The MIT Press, 2 edition.
- Ballesteros, J., Cary, A., and Rishe, N. (2011). Spsjoin: Parallel spatial similarity joins. In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, GIS '11, pages 481–484, New York, NY, USA. ACM.
- Bouros, P., Ge, S., and Mamoulis, N. (2012). Spatio-textual similarity joins. *Proc. VLDB Endow.*, 6(1):1–12.
- Buchin, M., Driemel, A., van Kreveld, M., and Sacristán, V. (2010). An algorithmic framework for segmenting trajectories based on spatio-temporal criteria. In *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, GIS '10, pages 202–211, New York, NY, USA. ACM.
- Ceccarelli, D., Lucchese, C., Orlando, S., Perego, R., and Trani, S. (2013). Learning relatedness measures for entity linking. In *Proceedings of the 22Nd ACM International Conference on Conference on Information & Knowledge Management*, CIKM '13, pages 139–148, New York, NY, USA. ACM.
- Cohen, W. W., Ravikumar, P., and Fienberg, S. E. (2003). A comparison of string metrics for matching names and records. In *Proceedings of the KDD-2003 Workshop on Data*, pages 13–18, Washington, DC.
- Fileto, R., Kruger, M., Pelekis, N., Theodoridis, Y., and Renso, C. (2013). Baquara: A holistic ontological framework for movement analysis using linked data. In Ng, W., Storey, V., and Trujillo, J., editors, *Conceptual Modeling*, volume 8217 of *Lecture Notes in Computer Science*, pages 342–355. Springer Berlin Heidelberg.
- Liu, S., Li, G., and Feng, J. (2012). Star-join: Spatio-textual similarity join. In *Proceedings of the 21st ACM Intl. Conf. on Information and Knowledge Management*, CIKM '12, pages 2194–2198, New York, NY, USA. ACM.
- Parent, C., Spaccapietra, S., Renso, C., Andrienko, G., Andrienko, N., Bogorny, V., Damiani, M. L., Gkoulalas-Divanis, A., Macedo, J., Pelekis, N., Theodoridis, Y., and Yan, Z. (2013). Semantic trajectories modeling and analysis. *ACM Comput. Surv.*, 45(4):42:1–42:32.
- Yan, Z., Chakraborty, D., Parent, C., Spaccapietra, S., and Aberer, K. (2013). Semantic trajectories: Mobility data computation and annotation. *ACM Trans. Intell. Syst. Technol.*, 4(3):49:1–49:38.