

## *Análise de Trajetória de Veículos baseado em Itinerários*

Duílio Santa Bárbara das Virgens Andrade  
Jorge Alberto Prado de Campos

Universidade de Salvador - UNIFACS  
Alameda das Espatódeas, 915, Caminho das Arvores, CEP 41820-460, Salvador, BA, Brasil  
duilio.andrade@gmail.com, jorge@unifacs.br

**Abstract.** The widespread use of mobile devices equipped with GPS has enabled large-scale capture of moving objects motion in many different fields. The massive dataset generated by these devices, however, has little or no utility in the way they are presented. In order to be used by a large number of applications, raw trajectory data must be mapped onto more abstract entities that carry some kind of semantic information from the application domain. This paper proposes a conceptual model to represent the trajectory of a moving object supposed to follow a given itinerary. The proposed model uses information about the itinerary and the moving object's trajectory to produce entities with some semantic information about the characteristics of the movement. We expected that the model improves queries processing time and facilitates the analysis and presentation of such kind of movement. As a case of study, we run an experiment implemented over a module of Weka data mining toolkit. Weka-STPM is a free and open source tool for semantic trajectory data mining and visualization.

**Palavras-chave:** Moving Objects, Semantic Trajectory Annotation, Itinerary, Weka-STPM.

### **1. Introdução**

A popularização dos dispositivos de posicionamento ampliou a demanda por aplicações capazes de monitorar e analisar a movimentação de animais, veículos, bens e pessoas para os mais diversos fins. A análise da trajetória de objetos móveis é um tema que tem despertado interesse crescente da comunidade de geoprocessamento. No que concerne ao estudo da movimentação de veículos, destacam-se os estudos para a concepção de modelos e enriquecimento semântico de trajetórias (LEAL *et al.*, 2011; ROCHA *et al.*, 2010; SPACCAPIETRA *et al.*, 2008; YAN *et al.*, 2011), identificação de padrões baseado em similaridade de trajetórias (HWANG *et al.*, 2006; TIAKAS *et al.*, 2009), mineração de dados e descoberta de conhecimento (BOGORNY *et al.*, 2011; PANG *et al.*, 2013; ZHENG *et al.*, 2013) e estudos de mobilidade urbana (VELOSO *et al.*, 2011; VIEIRA *et al.*, 2012).

Um problema central na maioria dos trabalhos que tratam da análise de trajetória de objetos móveis diz respeito ao modelo de representação da informação. Os dados coletados com os dispositivos de posicionamento são normalmente armazenados em uma sequência de pontos espaço-temporais do tipo (id,x,y,t), onde o "id" identifica o objeto móvel, "x" e "y" são as coordenadas geográficas e "t" o instante no qual o objeto ocupava a referida posição. Esta sequência de registros, denominada de trajetória bruta, é de pouca utilidade para aplicações interessadas em analisar esta formidável massa de dados. Os dados brutos de trajetórias possuem um baixo nível de abstração, diversos registros redundantes ou inconsistentes e carregam pouca ou nenhuma informação semântica sobre o domínio da aplicação.

A concepção de um modelo de dados com entidades mais abstratas e que associe algum tipo de semântica ao movimento dos objetos é indispensável para subsidiar diversos tipos de análises envolvendo os objetos móveis. Esta concepção depende fortemente do domínio da aplicação, da granularidade espaço-temporal dos dados coletados, do tipo de movimento e do tipo do espaço onde o movimento é desenvolvido (restrito ou sem restrições).

No contexto da análise da movimentação de veículos em um ambiente urbano, os dados coletados são oriundos de um mecanismo de posicionamento por satélites, o que impõe certa precisão na localização e uma frequência de registro bastante elevada, usualmente um registro

a cada 30 segundos. Além disso, o espaço geográfico onde o objeto desenvolve o seu movimento é restrito, isto é, os veículos somente podem se deslocar ao longo da rede de transportes, observando o sentido permitido para o fluxo e respeitando ligações e interseções.

Considerando o tipo de movimento, podemos classificar a trajetória dos veículos em dois grandes grupos: casual e programada. A trajetória casual não segue um roteiro ou cronograma conhecido ou fixo. Os veículos particulares desenvolvem trajetórias casuais, isto é, estes veículos não estão, em princípio, obrigados a seguir um determinado caminho. Embora seja razoável pensar que os proprietários dos veículos particulares possuam uma rotina diária, as trajetórias podem ser alteradas por compromissos eventuais ou condições de trafegabilidade das vias. Outro exemplo de movimentos casuais é o realizado pelos veículos que compõem o sistema de transporte de táxi. Táxis se movem em trajetórias aleatórias e a menos que o táxi esteja sempre vazio, a sua posição é altamente imprevisível e depende unicamente do desejo dos passageiros. Uma trajetória programada, por outro lado, segue um caminho ou itinerário bem definido. Os ônibus e trens urbanos, por exemplo, realizam trajetórias predefinidas, se movimentam em espaços ainda mais restritos, têm que obedecer a uma escala de horários rigorosa e possuem lugares específicos para parar, permitindo o embarque e desembarque de passageiros.

Embalados pelo barateamento da tecnologia de posicionamento e de comunicação de dados, diversas iniciativas públicas e privadas têm sido implantadas no sistema de transporte público. Virtualmente, todos os grandes centros urbanos utilizam GPS nos ônibus e usam esta informação para disponibilizar a posição instantânea do veículo ou em aplicações desenvolvidas para auxiliar a identificação de linhas, rotas e pontos de ônibus próximos que atendam a necessidade do usuário. A despeito do armazenamento de todas as posições coletadas dos veículos serem registradas na base de dados das empresas, as aplicações que utilizam informações sobre o posicionamento dos veículos só estão interessadas na informação em tempo real, ou seja, oferecem serviços apenas no nível operacional. Para o oferecimento de serviços e análises no nível estratégico ou tático, faz-se necessário o processamento dos registros históricos das trajetórias.

Objetivando a utilização de dados históricos das trajetórias de objetos móveis com movimentos condicionados a uma escala de serviço e o cumprimento de um determinado itinerário, este artigo propõe o modelo de dados *IB-Trajectory*. O referido modelo visa facilitar as consultas, extrações de conhecimento e desenvolvimento de novos algoritmos para aplicação que lidam com este tipo de movimento. Incluem-se no rol de potenciais usuários do modelo as agências reguladoras dos transportes públicos, secretarias de transportes e seus órgãos gestores, empresas de transporte de logística, de entrega de encomendas, de serviços de manutenção, entre outras.

O restante deste artigo está estruturado da seguinte forma: a próxima seção apresenta detalhes do modelo *IB-Trajectory* e discute a ferramenta desenvolvida para converter dados brutos de trajetórias em entidades do modelo. A seção 3 apresenta um estudo de caso de possíveis análises realizadas com entidades do modelo. Neste caso, utilizamos dados da trajetória de ônibus de um grande centro urbano. A seção 4 apresenta as conclusões e reflexões para trabalhos futuros.

## 2. Enriquecimento Semântico de Trajetórias Baseadas em Itinerário

O primeiro modelo que trata trajetórias de objetos móveis como movimentos que correspondem a deslocamentos semanticamente significativos e, por definição, um conceito espaço-temporal, foi proposto por Spaccapietra *et al.* (2008). Semanticamente, Spaccapietra considera uma trajetória como uma lista ordenada de *Stops* e *Moves*. Um *Stop* é uma parte da trajetória que é relevante para a aplicação e onde é considerado que o objeto não se

movimentou efetivamente, ou seja, permaneceu parado por um determinado período de tempo. Um *Move* é uma parte da trajetória onde o objeto encontra-se em movimento. A representação espacial de um *Stop* é um único ponto, enquanto um *Move* é representado por uma função de deslocamento ou uma polilinha construída com base nos pontos da trajetória.

O modelo seminal de *Stops* e *Moves*, entretanto, é muito abstrato para subsidiar a análise de trajetórias de objetos móveis baseadas em itinerários. Para esta finalidade, propomos uma extensão do referido modelo denominada de *Itinerary-Based Trajectory (IB-Trajectory)*. O modelo *IB-Trajectory* utiliza as informações do banco de dados da rede de transporte, dos itinerários e dos dados brutos coletados dos veículos durante a operação para compor entidades que capturem a semântica do movimento deste tipo de objeto (Figura 1).

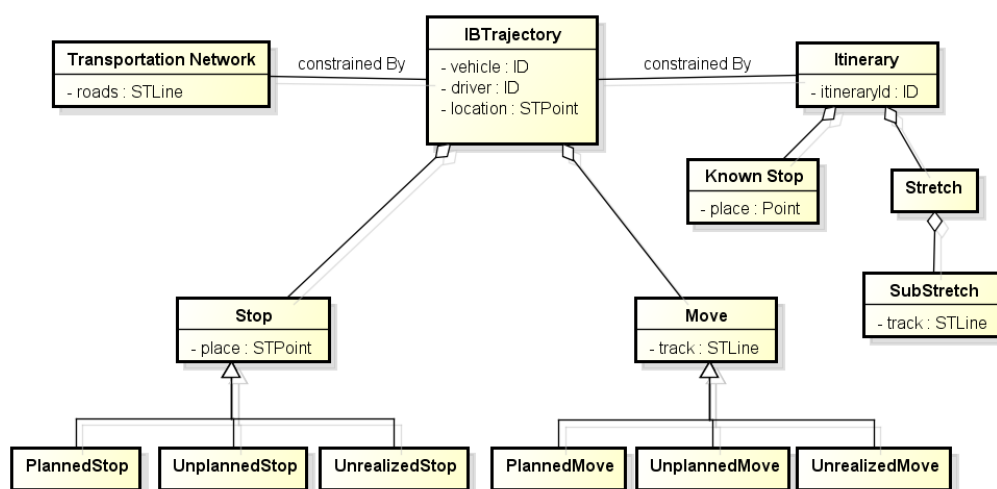


Figura 1. Modelo Conceitual *IB-Trajectory*.

A entidade básica do modelo é *IBTrajectory*. Esta entidade representa uma única viagem de um veículo, a fim de cumprir com um determinado itinerário. *IBTrajectory* é uma composição de tipos de dados abstratos (ou seja, *Stop* e *Move*). Os tipos de dados abstratos *Stop* e *Move* representam partes distintas da trajetória. A entidade *Stop* representa um ponto de parada do veículo para algum propósito. A entidade *Move* representa segmentos da viagem no qual o veículo está em movimento. Estas entidades, entretanto, não carregam nenhuma informação semântica sobre os aspectos do movimento. Assim, é impossível distinguir, por exemplo, se um veículo parou em um local programado ou em algum outro lugar por algum motivo desconhecido; ou ainda, se o veículo viajou ao longo do itinerário programado ou seguiu uma rota alternativa. Com o objetivo de capturar a essência de uma trajetória do veículo no contexto de movimentos pré-definidos, as entidades *Stops* e *Moves* são especializadas por três subtipos semanticamente relevantes. Subtipos da entidade *Stops* são *PlannedStop*, *UnplannedStop* e *UnrealizedStop*. Da mesma forma, a entidade *Move* é especializada através dos subtipos *PlannedMove*, *UnplannedMove* e *UnrealizedMove*.

Entidades *PlannedStop* representam todas as paradas detectadas dos veículos que coincidem com os pontos de parada descritos na base de itinerário. Já as entidades *UnplannedStop* representam paradas detectadas que não fazem parte dos pontos de parada definidos na base de itinerário. Finalmente, as entidades *UnrealizedStop* representam pontos de parada da base de itinerário que não foram detectadas na trajetória coletada.

Quando os veículos estão trafegando, as características do movimento são representadas por subtipos de *Moves* (ou seja, *PlannedMove*, *UnplannedMove* e *UnrealizedMove*). Estas entidades são semanticamente equivalentes às partes estacionárias da trajetória. Uma *PlannedMove*, por exemplo, representa trechos da trajetória definida na base de itinerário que

efetivamente foram percorridos pelo veículo. Entidades *UnplannedMove* representam trechos da trajetória realizada pelo veículo que não correspondem a seguimentos de trechos mapeados na base de itinerários, ou seja, são trechos que não estão previstos no itinerário, mas que o veículo realizou. As entidades *UnrealizedMove* representam trechos que não foram percorridos na trajetória mas que fazem parte dos segmentos de trecho mapeados na base de itinerários, ou seja, são trechos do itinerário que não foram percorridos.

O objetivo deste trabalho é demonstrar o potencial do modelo *IB-Trajectory* como base para consulta e análise de trajetórias de veículos que devem seguir um determinado itinerário. O processo de conversão de dados brutos da trajetória em entidades do modelo foge do escopo deste artigo. Gostaríamos de registrar, entretanto, que este processo está completamente automatizado através da criação de um novo módulo na ferramenta Weka-STPM (ALVARES *et al.*, 2010; BOGORNY *et al.*, 2011). O Weka-STPM é uma ferramenta de código aberto para a mineração de dados, análise semântica e visualização de trajetórias. Foi desenvolvido e incorporado à ferramenta Weka-STPM o módulo IBTGM (*Itinerary-Based Trajectory Generation Module*). IBTGM é carregado com os dados brutos da trajetória, a base viária da região e com a base de itinerários e paradas conhecidas e produz como saída as tabelas com as entidades *IBTrajectory* e suas versões especializadas dos *Stops* e *Moves* (Figura 2)

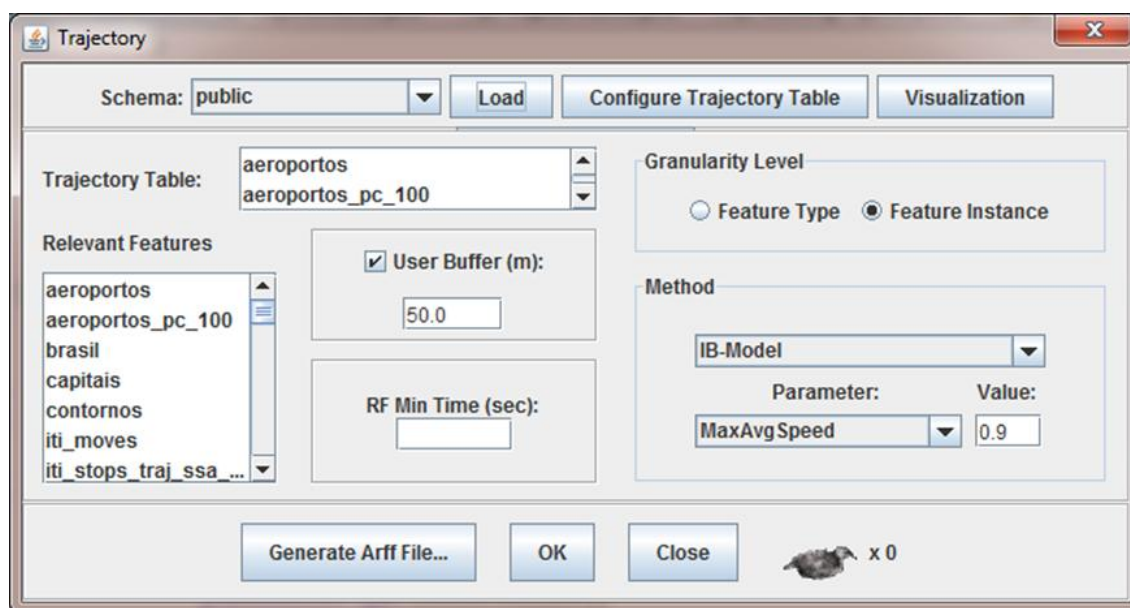


Figure 2. Interface com o usuário do Módulo IBTGM implementado no *Weka data mining toolkit*.

Na seção seguinte são apresentadas algumas análises realizadas sobre uma base de dados de trajetórias do sistema público de transporte da cidade do Recife-PE.

### 3. Análise Trajetórias Baseadas em Itinerário

De forma a ilustrar a potencialidade do Modelo *IB-Trajectory*, escolhemos como estudo de caso, dados reais obtidos do sistema de transporte público de um grande centro urbano. Não se pretende aqui apresentar nenhum sistema especialista na análise e apresentação do sistema de transporte público por ônibus. Queremos demonstrar que a simples transformação de dados brutos em entidades mais abstratas e semanticamente significativas já permite a execução de consultas que evidenciam alguns padrões na movimentação dos veículos.

Acreditamos que o desenvolvimento de sistemas especialistas baseados no modelo *IB-Trajectory* pode disponibilizar funcionalidades não abordadas neste artigo.

Os dados que nos referimos foram obtidos do projeto UbiBus. Este projeto explora a utilização de contexto e computação ubíqua para construir aplicações que auxiliem o passageiro do sistema público de transporte (VELOSO *et al.*, 2011; VIEIRA *et al.*, 2012). A base de dados disponibilizada pelo projeto UbiBus corresponde a aproximadamente dez dias de trajetórias brutas de 384 linhas de ônibus com aproximadamente 40 milhões de registros.

O nosso primeiro exemplo é uma viagem de um ônibus da linha 914 (Terminal Integrador PE-15/ Terminal Integrador Afogados) realizada no dia 21/08/2012. A Figura 3.a apresenta uma visão geral do itinerário. Um itinerário de uma linha de ônibus é o caminho de ida e volta entre os dois pontos extremos. Em nossa base de dados, itinerários são armazenados como uma polilinha, desta forma, a apresentação de um itinerário no mapa é feita com uma linha contínua. Em certos trechos da figura, o itinerário parece uma única linha. Este efeito é causado pelo nível de zoom e pelo fato de em alguns trechos os itinerários passarem pela mesma via, mas em mãos opostas.

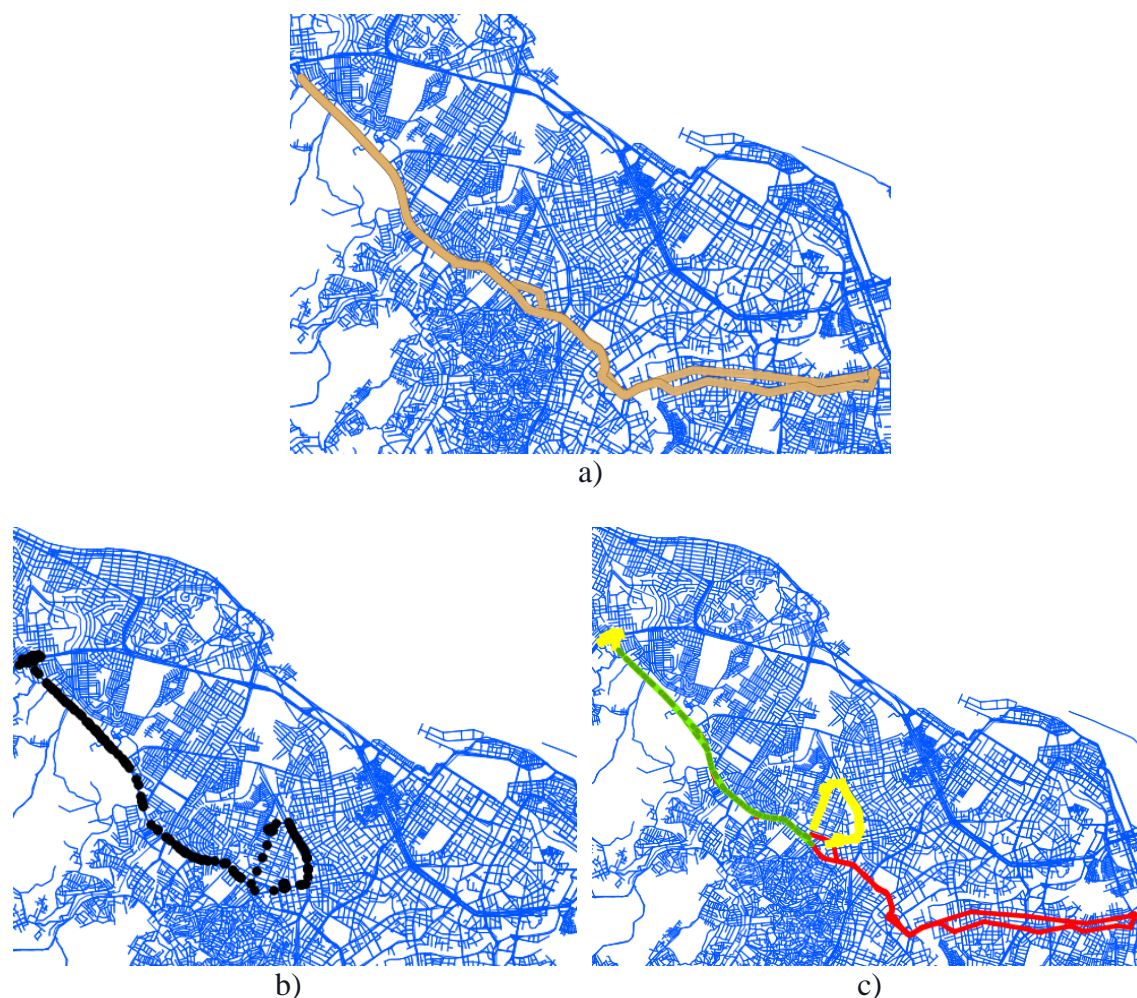


Figura 3 Dados da movimentação incompleta de um ônibus da linha 914: a) apresentação do itinerário da linha 914; b) dados da trajetória bruta; e c) *Moves* do modelo *IB-Trajectory* para esta trajetória.



No processo de conversão de dados brutos de trajetória em entidades do modelo *IB-Trajectory*, identifica-se, em um primeiro passo, todos os pontos da trajetória bruta que pertencem a uma entidade *IBTrajectory*. No contexto dos ônibus do sistema público de transporte, uma *IBTrajectory* são todos os pontos percorridos (ida e volta) com o objetivo de cumprir um itinerário. Para efeito de ilustração, escolhemos duas trajetórias do ônibus da linha 914: a primeira trajetória com um problema no cumprimento do itinerário e outra que respeitou integralmente o que era previsto.

Os dados brutos da primeira trajetória (Figura 3.b) mostra que o ônibus, por algum motivo desconhecido, não cumpriu todo o itinerário. Esta é uma situação bastante usual no sistema de transporte público. Os passageiros denunciam esta prática e alegam que em certos horários, o número de passageiros em alguns trechos é deficitário e os motoristas desviam ou deixam de cumprir parte do itinerário. A situação mais comum é a de retornar sem ter atingido um dos pontos extremos do itinerário. As empresas, por sua vez, alegam que todas as viagens são cumpridas integralmente, independentes do número de passageiros no veículo.

Para os dados brutos de trajetória apresentados na figura 3.b, o módulo IBTGM adicionado ao Weka-STPM gera entidades do tipo *PlannedMove*, *UnplannedMove* e *UnrealizedMove*, as linhas verde, amarela e vermelhas na Figura 3.c, respectivamente. Analisando a imagem da Figura 3.c pode-se inferir que o ônibus iniciou a viagem no canto superior esquerdo e cumpriu o itinerário até certo ponto. No meio do itinerário, aproximadamente, o ônibus realiza um pequeno contorno (linha amarela no centro da imagem), e retorna ao início do itinerário, deixando assim de cumprir a última metade do seu itinerário (linha vermelha no canto inferior direito).

Observa-se uma ocorrência estranha de uma *UnplannedMove* no início da trajetória (linha amarela no canto superior esquerdo). Verificamos que esta ocorrência ocorre em virtualmente todas as trajetórias da linha 914. Neste caso identificamos um problema na base de dados de itinerários. O itinerário da linha 914 não contempla uma pequena volta que o veículo tem que fazer para se reposicionar no início do itinerário. Este fato gera uma ocorrência de um *UnplannedMove* no final de cada trajetória.

Os dados brutos da segunda trajetória do ônibus da linha 914 representa uma trajetória sem maiores problemas. O ônibus nesta trajetória cumpriu rigorosamente o itinerário. Desta forma, o módulo IBTGM gera somente instâncias *PlannedMove*, ou seja, linhas verdes na figura 4.b. A exceção fica por conta *UnplannedMove* no início da trajetória já discutido anteriormente.

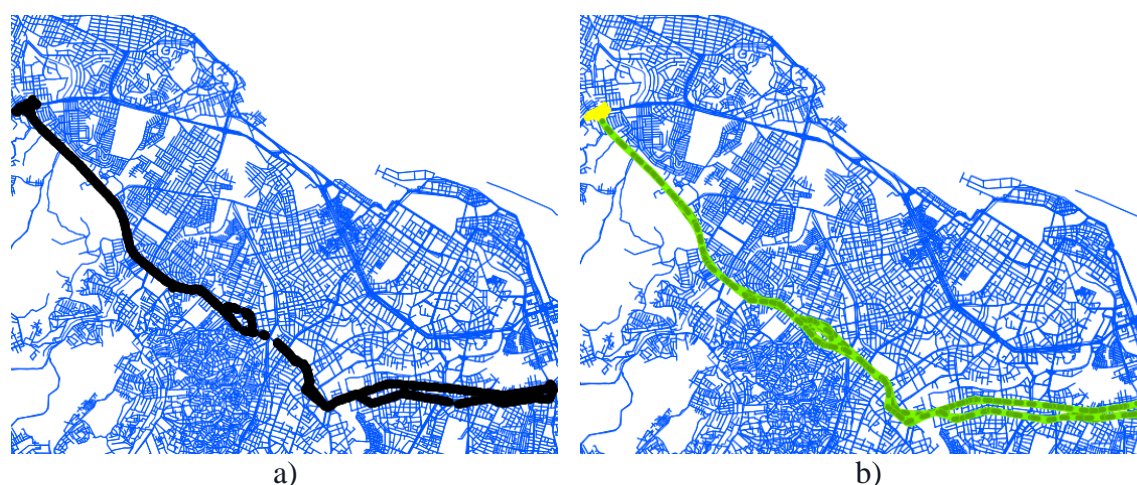


Figura 4. Dados da movimentação completa de um ônibus da linha 914: a) dados da trajetória bruta e b) *Moves* do modelo *IB-Trajectory* para esta trajetória.

Para efeito da discussão da parte estacionária das trajetórias, escolhemos estudar uma linha mais curta e alimentadora do sistema. A linha em questão é a 929-Alameda Paulista / Terminal Integrador Rio doce (Maranguape I).

A Figura 5.a apresenta o itinerário da linha 929 e a Figura 5.b os dados brutos de uma trajetória realizada no dia 21/08/2012. Considerando os dados brutos da trajetória e o registro da localização de todos os pontos de ônibus ao longo do itinerário, o módulo IBTGM gera as instâncias das versões especializadas dos *Stops* (i.e., *PlannedStop*, *UnplannedStop* e *UnrealizedStop*), representados por regiões verdes, amarelas ou pontos vermelhos (Figura 5.c).

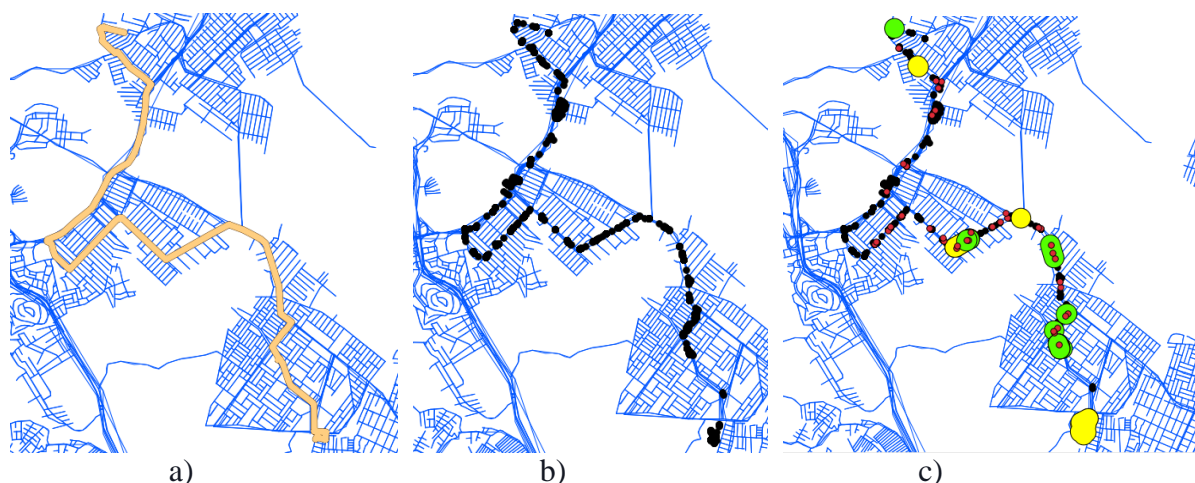


Figura 5. Dados das paradas de um ônibus da linha 929: a) apresentação do itinerário da linha 929, b) dados da trajetória bruta e c) *Stops* do modelo *IB-Trajectory* para esta trajetória.

As entidades *PlannedStop* e *UnplannedStop* são representados por um região pois equivalem ao fecho convexo dos conjuntos de pontos que foram utilizados na geração do *Stop* durante o processo de conversão de dados brutos em entidades do modelo. *UnrealizedStop* por sua vez é representado por um ponto, que é a localização da parada que deveria ter sido feita, mas não o foi.

Consultas envolvendo as diversas especializações de *Stops* podem ser realizadas para determinar se os veículos de certa linha para regularmente em locais não previstos ou se ignora as paradas que deveriam ter sido feitas. As análises envolvendo trechos estacionários irão requerer algumas informações de contexto. Determinar se um veículo está cheio, por exemplo, pode ser uma informação relevante na análise das paradas não realizadas. As condições de trafegabilidade das vias, por outro lado, pode ajudar na eliminação de falsos positivos de paradas não planejadas.

#### 4. Conclusão

Este trabalho apresentou um modelo conceitual de dados, denominado de *IB-Trajectory*, para representar trajetórias de objetos móveis que podem ou devem seguir um determinado itinerário. *IB-Trajectory* possui entidades mais abstratas e semanticamente enriquecidas para capturar a essência do movimento deste tipo de objetos móveis. O referido modelo tem como objetivo criar uma infraestrutura de dados para suportar uma ampla gama de análises sobre as trajetórias dos veículos que devem seguir determinado itinerário.

O modelo *IB-Trajectory* fornece entidades de primeira classe para representar os lugares onde o veículo fez paradas programadas, fez paradas inesperadas e lugares onde o veículo deveria ter parado, mas não o fez. Da mesma forma, especializações dos movimentos do

veículo são utilizadas na representação de segmentos percorridos ou ignorados pelo veículo; e segmentos da rede rodoviária fora do itinerário, mas percorrida pelo veículo.

## Referências

- ALVARES, L.; PALMA, A.; OLIVEIRA, G.; BOGORNY, V. Weka-STPM: from trajectory samples to semantic trajectories. **Proceedings of the Workshop on Open Source Code 1**, v. 1, p. 1–6, 2010.
- BOGORNY, V.; AVANCINI, H.; PAULA, B. DE. Weka-STPM: a Software Architecture and Prototype for Semantic Trajectory Data Mining and Visualization. **Transactions in GIS 15**, p. 227–248, 2011.
- HWANG, J.-R.; KANG, H.-Y.; LI, K.-J. Searching for similar trajectories on road networks using spatio-temporal similarity. **Advances in Databases and Information Systems (ADBIS)**, p. 282–295, 2006.
- LEAL, B.; MACEDO, J. DE; TIMES, V.; *et al.* From Conceptual Modeling to Logical Representation of Trajectories in DBMS-OR and DW Systems. **October**, 2(3), p. 463–478, 2011.
- PANG, L.; CHAWLA, S.; LIU, W.; ZHENG, Y. On detection of emerging anomalous traffic patterns using GPS data. **Data & Knowledge Engineering (DKE)**, 87, ACM press, p. 357–373, 2013.
- ROCHA, J.; OLIVEIRA, G.; ALVARES, L.; BOGORNY, V.; TIMES, V. DB-SMoT: A direction-based spatio-temporal clustering method. **Intelligent Systems (IS), 2010 5th IEEE International Conference**, p. 114–119, 2010.
- SPACCAPIETRA, S.; PARENT, C.; DAMIANI, M.; MACEDO, J. DE; PORTO, F. A conceptual view on trajectories. **Data and Knowledge Engineering**, v. 65, n. 1, p. 126–146, abr 2008.
- TIAKAS, E.; PAPADOPOULOS, A. N.; NANOPOULOS, A.; *et al.* Searching for similar trajectories in spatial networks. **Journal of Systems and Software**, v. 82, n. 5, p. 772–788, maio 2009.
- VELOSO, M.; PHITHAKKITNUKON, S.; BENTO, C. Urban mobility study using taxi traces. **Proceedings of the 2011 international workshop on Trajectory data mining and analysis TDMA 11**, p. 23–30, 2011.
- VIEIRA, V.; SALGADO, A.; TEDESCO, P.; *et al.* The UbiBus Project: Using Context and Ubiquitous Computing to build Advanced Public Transportation Systems to Support Bus Passengers. **Anais do VIII Simpósio Brasileiro de Sistemas de Informação**, p. 55–60, 2012.
- YAN, Z.; CHAKRABORTY, D.; PARENT, C.; SPACCAPIETRA, S.; ABERER, K. SeMiTri: a framework for semantic annotation of heterogeneous trajectories. **Proceedings of the 14th International Conference on Extending Database Technology**, p. 259–270, 2011.
- ZHENG, K.; ZHENG, Y.; YUAN, J.; SHANG, S. On discovery of gathering patterns from trajectories. **IEEE 29th International Conference on Data Engineering (ICDE)**, p. 242–253, 2013.