

## Visual R-MAT: uma ferramenta visual de apoio a análises sob dados representativos de trajetórias de múltiplos aspectos

Tiago Oliveira da Luz<sup>1</sup>, Vanessa Lago Machado<sup>1,2</sup>, Ronaldo dos Santos Mello<sup>1</sup>

<sup>1</sup> INE - Universidade Federal de Santa Catarina (UFSC)  
Florianópolis – SC – Brasil

<sup>2</sup>Instituto Federal Sul-Rio-Grandense (IFSul)  
Passo Fundo – RS – Brasil

{[@gmail.com](mailto:tiagoluz.grad,vanessalagomachado), [r.mello@ufsc.br](mailto:r.mello@ufsc.br)}

**Abstract.** *The phenomenon of IoT brings a surge of vast amounts of data. Present in these data is trajectory data, which can be semantically enriched to generate multiple aspect trajectories. As an alternative solution to the management of these data surges summarization approaches that generate representative trajectory data. Then, based on these data surge the need for visual tools as support to data analysts. This paper aims to propose a 2D visual tool to support data analysts in tasks of analysis and comparison between a dataset of multiple aspect trajectories and their respective representative data, allowing, for example, the analysis of the summarization algorithm.*

**Resumo.** *O fenômeno da IoT traz consigo o surgimento de vastas quantidades de dados. Presente nesses dados estão os dados de trajetória, que podem ser semanticamente enriquecidos para gerar trajetórias de múltiplos aspectos. Ainda, como alternativa no gerenciamento de tais dados surgiram abordagens de sumarização que geram dados de trajetória representativa. Com base nesses dados surge, portanto, a necessidade de ferramentas como apoio visual na análise de tais dados. Este trabalho propõe uma ferramenta visual 2D de apoio ao analista de dados nas tarefas de análise e comparação entre um grupo de trajetórias e uma trajetória representativa, permitindo, por exemplo, inferências em relação ao algoritmo de sumarização.*

### 1. Introdução

Com a diversificação de aparelhos conectados a internet surge uma quantidade crescente de dados. Dentre esses dados estão presentes os dados de mobilidade de objetos móveis, informando as sequências de registros espaciais dessa mobilidade e o instante de tempo de cada registro, os dados de trajetórias. Tais dados contêm informações vitais para o funcionamento de diversas aplicações, como sistemas de recomendação, sistemas de predição, monitoramento de tráfego, entre outras.

Os dados de trajetórias podem conter enriquecimento de diferentes aspectos semânticos, gerando dados de trajetórias de múltiplos aspectos, ou seja, dados complexos para manipulação. Ainda, os dados de trajetórias são gerados e coletados constantemente e, junto com essa complexidade dos dados, geram problemas de gerenciamento de dados.

Uma alternativa para o problema de gerenciamento, em relação ao volume de dados a ser gerenciado, é a sumarização desses dados gerando dados representativos. Contudo, a partir do dado representativo identificou-se a necessidade de ferramentas visuais que forneça ao usuário um apoio nas análises e comparações dos dados representativos em relação ao conjunto completo de trajetórias.

Desse modo, este trabalho propõe uma ferramenta visual para trajetórias de múltiplos aspectos e da trajetória representativa, como apoio ao analista de dados, permitindo a filtragem das informações, e os resultados apresentados são parte de um trabalho em andamento.

O restante do trabalho encontra-se organizado conforme segue. A seção 2 apresenta os conceitos básicos referente aos dados de trajetórias. A seção 3 é dedicada a alguns trabalhos relacionados. A seção 4 descreve nossa ferramenta proposta e, por fim, a seção 5 conclui este trabalho e apresenta algumas possibilidades de trabalhos futuros.

## 2. Fundamentação Teórica

A presente seção define alguns conceitos básicos que são relevantes para compreensão desse artigo.

### 2.1. Dados de Trajetórias

Dados de trajetória de um objeto móvel referem-se a as informações referentes a mobilidade de um indivíduo ou qualquer objeto com seu devido registro de deslocamento, como por exemplo o registro da mobilidade de um carro ou de um navio. Tal deslocamento pode ser obtido por meio de diferentes dispositivos e sensores, permitindo assim a coleta de informações sobre esse objeto, desse modo a sequência dos pontos registrados para cada objeto móvel é chamada de trajetória do objeto móvel, a qual pode envolver diferentes dimensões (espacial, temporal e/ou semântica) [Bogorny et al. 2012].

### 2.2. Trajetórias Brutas

A trajetória bruta de um objeto móvel é dada pela sequência de registro de pontos  $P = \langle p_1, p_2, \dots, p_n \rangle$  sobre o espaço geográfico ao longo do tempo, ou seja, cada registro de ponto se encontra no formato  $(x, y, t)$ , onde  $x$  e  $y$  referem-se as informações das coordenadas referente ao espaço geográfico que o objeto se encontra no instante de tempo  $t$  [Erwig et al. 1999]. Desse modo, tais informações descrevem a mobilidade de determinado objeto com as dimensões espacial e temporal, contudo, muitos sistemas hoje possibilitam a coleta de dados ricos em informações semânticas, o que facilita a análise e mineração deste tipo de dados.

### 2.3. Trajetórias de Múltiplos Aspectos (MATs)

Quando uma trajetória bruta é enriquecida com uma informação semântica ela é conhecida como *trajetória semântica*, como quando é associada a informação referente ao *ponto de interesse (PoI)* que o objeto móvel visitou, por exemplo, informando que determinado registro de ponto da trajetória refere-se a um restaurante [Alvares et al. 2007].

Contudo, mais recentemente, em 2019, [Mello et al. 2019] identificaram que determinados pontos dessa trajetória, ou a trajetória inteira, poderiam ser enriquecidos com diversas e diferentes contextos semânticos, assim propuseram as trajetórias de múltiplos

aspectos, multiple aspect trajectory (MATs). Em uma trajetória de múltiplos aspectos a sequência de registro de pontos é caracterizada de modo que cada registro de ponto possui o formato  $p_i = (x_i, y_i, t_i, S_i)$ , no qual  $S_i = \langle s_1, s_2, \dots, s_n \rangle$  é um conjunto de aspectos semânticos associados ao ponto  $p_i$  na posição  $(x_i, y_i)$  no instante de tempo  $t_i$ .

A Figura 1 mostra como exemplo uma trajetória bruta de um indivíduo enriquecida com diversas informações, como o meio de transporte utilizado, postagens nas redes sociais, condições climáticas, informações de saúde do indivíduo, entre outras. Nesse exemplo é possível verificar que uma MAT é composta por três dimensões (espacial, temporal e semântica), em que a terceira dimensão pode ser representada por múltiplos e heterogêneos aspectos. Desse modo, verifica-se que esse tipo de trajetória possui uma maior riqueza de informações, contudo apresenta também maior complexidade de gerenciamento.

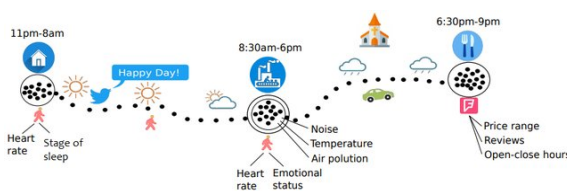


Figura 1. Um exemplo de trajetória de múltiplo aspectos [Mello et al. 2019].

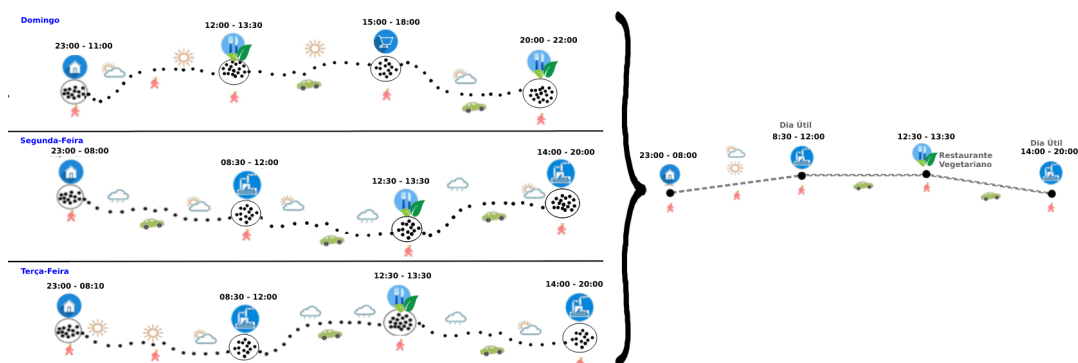
## 2.4. Trajetórias Representativas

O gerenciamento de dados de trajetórias é um grande desafio enfrentado devido ao vasto volume e variedade de dados gerados continuamente por diferentes fontes de dados. Desse modo, algumas tarefas como de armazenamento e de análise desses dados acabam tornando-se penosas. Nesse caso se torna interessante técnicas de sumarização com o objetivo de criar uma trajetória representativa dos dados. Isto é, uma trajetória que possua o comportamento padrão de um objeto móvel, por exemplo, no espaço e tempo, assim como aspectos semânticos no caso de MATs, reduzindo significativamente a quantidade de dados a ser processada [Machado et al. 2022].

Compreender padrões em trajetórias pode ajudar analistas no seu processo de decisão. Por exemplo, em um sistema de recomendação objetiva-se a compreensão dos comportamentos de cada usuário para encontrar produtos que possam interessá-los. A Figura 2 mostra as MATs de uma mulher apresentando sua trajetória em diferentes dias da semana (Domingo, Segunda-Feira e Terça-Feira), sendo que do lado direito é apresentada uma trajetória representativa, em que suas MATs são sumarizadas e apresenta as ações que são mais frequentes. Desse modo é possível verificar que ela vai ao trabalho em dias úteis nos turnos matutino e vespertino e frequentemente almoça em restaurantes vegetarianos.

Com base na trajetória representativa provido na Figura2 (lado direito), um sistema de recomendação pode aprender sobre os hábitos do indivíduo e fazer recomendações. Por exemplo, visto que ela prefere restaurantes vegetarianos, quando ela estiver em uma região diferente a habitual, próximo ao seu habitual horário de almoço, o sistema pode identificar e recomendar os restaurantes vegetarianos mais próximos.

Contudo, a maioria dos métodos propostos para geração de trajetórias representativas estão limitados a trajetórias brutas, sem considerar os aspectos semânticos. Um



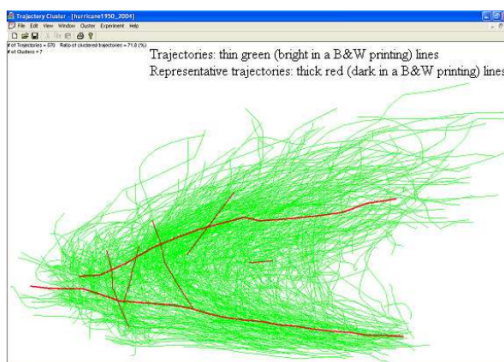
**Figura 2. Exemplo de MATs (lado esquerdo) e uma MAT representativa (lado direito). Adaptado de [Machado et al. 2022]**

método proposto para sumarização de trajetórias de múltiplos aspectos é o MAT-SG, Multiple Aspect Trajectory Summarization based on a spatial Grid [Machado et al. 2022]. O método MAT-SG é baseado em uma grade espacial que cobre um conjunto de MATs de entrada. Para todos os pontos na mesma célula da grade é gerado um ponto representativo, sumarizando todos as dimensões dos pontos contidos na célula. Desse modo, a MAT representativa é gerada como um conjunto de pontos representativos.

### 3. Trabalhos Relacionados

A maioria dos trabalhos que abordam trajetórias tem como foco principal a análise destes dados e não sua visualização e costumam trazer representações dos dados apenas como saídas de seus métodos.

O algoritmo TRA-CLUS [Lee et al. 2007], por exemplo, propõem um método para clusterização de dados de trajetórias espaciais e apresenta em seu trabalho representações visuais como saída do método de clusterização, para visualização e análise dos resultados. A Figura 3 mostra a saída do método de clusterização como análise dos resultados para um dataset avaliado, em que as linhas em verdes representam as trajetórias de entrada e as linhas em vermelho representam a trajetória central de cada cluster gerado pelo algoritmo.



**Figura 3. Visualização espacial do resultado gerado pelo algoritmo TRA-CLUS [Lee et al. 2007]**

Em 2012, Tominski et al. introduziram uma nova solução para a representação de

MATs por meio de uma visualização híbrida entre 2D/3D [Tominski et al. 2012]. Uma das visualizações possíveis, conforme representado na Figura 4, refere-se as trajetórias que são representadas por faixas empilhadas sobre a imagem de satélite do local em 2D. Tais faixas são codificadas como um gradiente de cor para representar um determinado atributo semântico. Outra visualização propõe um gráfico temporal para a exploração dos dados temporais, permitindo analisar a sobreposição de trajetórias no tempo. Contudo, essa ferramenta tem como foco apenas a visualização de conjuntos de trajetórias, não abordando o conceito de trajetórias representativas e sendo limitado a análise de um único aspecto semântico.



**Figura 4. Visualização 3D de dados de trajetórias, envolvendo as dimensões espacial e semântica, em que o gradiente de cores refere-se a informação de velocidade [Tominski et al. 2012]**

A Tabela 1 apresenta um resumo dos principais pontos de cada trabalho.

**Tabela 1. Comparação entre os trabalhos relacionados e suas características**

Trabalho	Visualização de trajetórias	Visualização de MATs	Visualização de MATs representativas
[Lee et al. 2007]	x		
[Tominski et al. 2012]	x	x	
Visual R-MAT	x	x	x

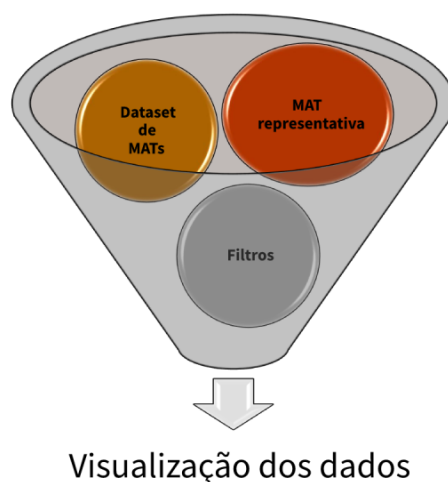
Como visto, as abordagens apresentadas geram visualizações dos dados de trajetórias espaciais [Lee et al. 2007] ou trajetórias semânticas [Tominski et al. 2012]. Ainda, embora a ferramenta proposta por [Tominski et al. 2012] traga uma gama maior de flexibilidade para as visualizações, nenhum dos trabalhos possui como foco a visualização de dados representativos. Além disso, verificamos também que há uma demanda por ferramentas que gerem visualizações para dados de trajetórias de múltiplos aspectos, utilizando as semânticas como alternativa de filtros sobre os dados.



#### 4. Visual R-MAT

Tendo em vista a demanda identificada, este trabalho propõe uma ferramenta chamada Visual R-MAT, *Visualization of Representative Multiple Aspect Trajectory*, a qual possui como o objetivo fornecer uma visualização simplificada em 2D de trajetórias de múltiplos aspectos. Mais especificamente, com a capacidade de comparar um conjunto de trajetórias de múltiplos aspectos com a de uma trajetória representativa, permitindo ainda analisar e comparar dados específicos.

A Figura 5 apresenta a arquitetura proposta para Visual R-MAT, em que a ferramenta possui como dados de entrada o dataset de MATs, a MAT representativa e os filtros sobre os múltiplos aspectos, e gera como dado de saída a visualização dos dados.



**Figura 5. Arquitetura proposta da ferramenta Visual R-MAT**

Em suma, a ferramenta Visual R-MAT, desenvolvida na linguagem Python, realiza a leitura de dois arquivos no formato “.csv”, em que o primeiro refere-se ao dataset de MATs e o segundo refere-se aos dados da MAT representativa. O arquivo referente ao dataset de trajetórias é estruturado de forma que cada linha representa o dado de um ponto de uma trajetória, o qual possui o seguinte formato: "tid;lat\_lon;date\_time;[semantics]", onde: tid é o identificador da trajetória a qual o ponto pertence; lat\_lon refere-se as informações espaciais do referente ponto (latitude, longitude); date\_time é o instante de tempo da ocorrência daquele ponto; e, [semantics] referem-se aos  $N$  aspectos semânticos relacionados ao ponto da trajetória em questão. O formato da trajetória representativa é o proposto pelo método MAT-SG [Machado et al. 2022] e encontrado na Figura 6, que apresenta as informações associadas a um ponto representativo.

A ferramenta, demonstrada na Figura 7, permite o carregamento dos arquivos de entrada, em que o usuário carrega o arquivo (Figura 7(A)) e informa o tipo do arquivo a ser carregado: o dataset (Figura 7(B)) ou uma trajetória representativa (Figura 7(C)). Uma vez carregado os arquivos na ferramenta é possível configurar o gráfico gerado, em que o usuário pode escolher entre visualizar somente as trajetórias do dataset (selecionando "Plot Dataset" na Figura 7(D)), somente a trajetória representativa (selecionando "Plot

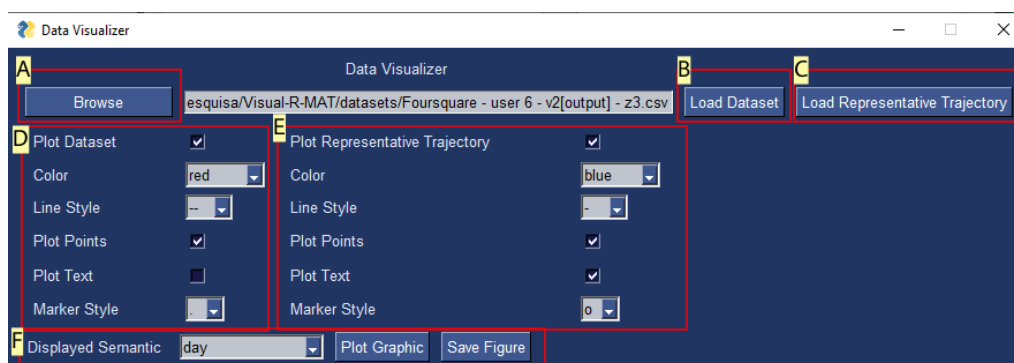
```

rt
(x,y)= (40.77, -73.87),
Ranking of DAY= [Friday -> 0.667, Thursday -> 0.333, ],
Ranking of POI= [LaGuardia Airport (LGA) (LaGuardia Airport) -> 1, ],
Ranking of TYPE= [Airport -> 1, ],
Ranking of ROOT_TYPE= [Travel & Transport -> 1, ],
Ranking of PRICE= [Unknown -> 1, ],
Ranking of RATING= [*5.3 -> 1, ],
Ranking of WEATHER= [Clouds -> 0.667, Clear -> 0.333, ],
Ranking of LABEL= [*6 -> 1, ],
Ranking of CHECKIN_ID= [*304348 -> 0.333, *310528 -> 0.333, *309049 -> 0.333, ],
Ranking of VENUE_ID= [*4ace6c89f964a52078d020e3 -> 1, ],
Ranking of Interval Temporal= [25/05/2022 19:22 - 25/05/2022 20:19 > 1.0, ],
Local mapped: 130 - 88, 131 - 113, 132 - 120,

```

**Figura 6. Formato aceito para trajetórias representativas**

Representative Trajectory”na Figura 7(E)) ou ambos. Uma vez definidas as visualizações é possível configurar o gráfico, como o tipo da linha a ser apresentada para as trajetórias e a cor dessas linhas.

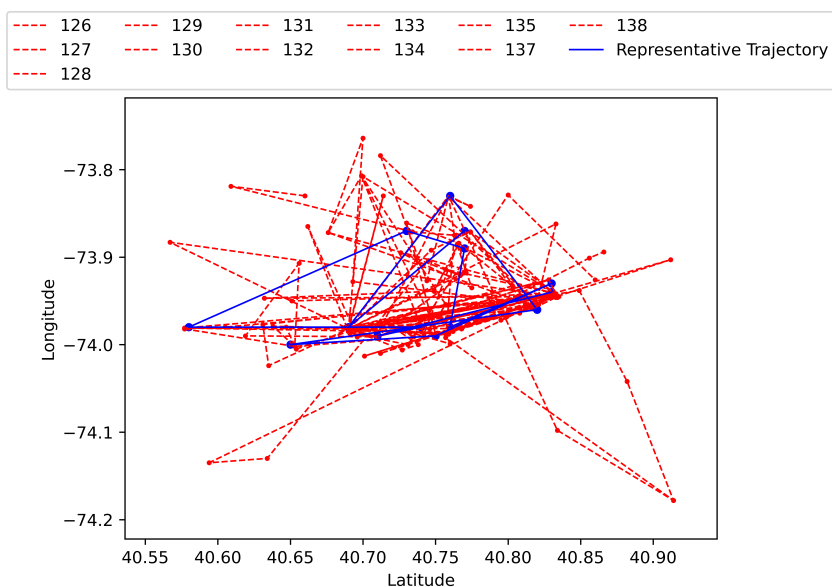


**Figura 7. Ferramenta Visual-R-MAT**

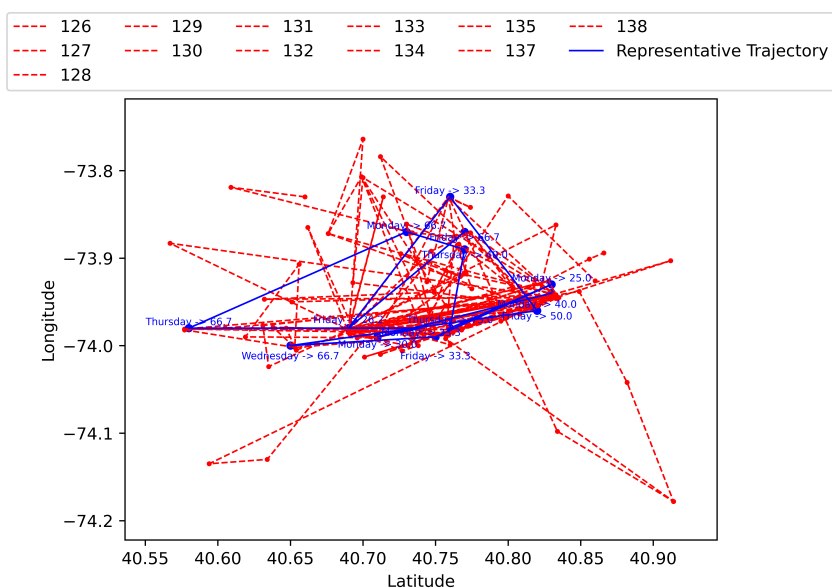
A Figura 8 mostra o resultado da visualização proposta para as seleções realizadas na 7(A-E). Nesse caso as linhas tracejadas em vermelho representam o dataset de MATs, em que cada trajetória é identificada por um identificador (trajetórias 126 a 138), e a linha azul representa a MAT representativa. Tal visualização permite analisar e comparar a trajetória representativa sob todas as trajetórias de entrada.

Também é possível configurar diversas características de visualização do gráfico. As informações semânticas podem ser apresentadas ou não, sendo possível selecionar um dos aspectos encontrados no dataset para ser apresentado ao lado dos pontos. Tal configuração é possível conforme 7(F). Assim como as demais configurações, essa visualização pode ser aplicada somente para o dataset de MATs, somente para os pontos da MAT representativa, ou para ambos os casos. Por exemplo, de acordo com as seleções realizadas na 7(A-F), a Figura 9 apresenta uma visualização do aspecto semântico ”Dia da Semana”(dado pela seleção ”day”na 7(F)), com visualização somente sob os pontos de dados da trajetória representativa (opção ”Plot Text”na Figura 7(E)).

A Visual R-MAT permite a visualização dos dados de trajetórias representativas e do conjunto de trajetórias de múltiplos aspectos, permitindo a visualização dos dados em forma textual, no caso dos dados não espaciais. Contudo, tal visualização baseada em aspectos de forma textual não se mostra eficaz para visualização de grandes quantidades de



**Figura 8. Visualização espacial das trajetórias em um dataset (linhas em vermelho) e da trajetória representativa gerada pelo método MAT-SG (linha azul)**



**Figura 9. Visualização de um dataset de MATs, sua MAT representativa e a informação do Dia da Semana identificado como padrão na MAT representativa.**

dados. Desse modo, verifica-se a possibilidade de extensão da ferramenta com a inclusão de recursos, possibilitando o destaque da trajetória representativa por meio de um filtro realizado pelo usuário. Isso possibilitaria o destaque dos pontos relevantes para a análise desejada. Por exemplo, ao selecionar os dados referentes ao dia da semana, poderia ser destacado os pontos das trajetórias que possuem o dia da semana identificado como mais frequente, por exemplo, domingo. Tal extensão permitirá maior controle na análise dos dados e extração de informações relevantes.



## 5. Conclusão

Neste artigo exploramos os conceitos envolvendo dados de trajetórias e os problemas que surgem ao tentar representar estes dados visualmente. Também, ao analisar outros trabalhos percebe-se que poucos deles abordam o problema da representação de trajetórias de múltiplos aspectos, assim como das trajetórias representativas, possibilitando assim o apoio visual na análise dos padrões identificados em um conjunto de MATs, assim como a comparação com a trajetória representativa.

Desse modo, a ferramenta de apoio Visual R-MAT<sup>1</sup> possibilita a análise e comparação visual do conjuntos de MATs em relação as suas trajetórias representativas (providas pelo método MAT-SG). Além disso, oferece também a possibilidade de configurar os aspectos visuais da representação, assim como a visualização dos aspectos das MATs, de acordo com os aspectos contidos nos dados de entrada.

Tal ferramenta se mostra promissora como apoio visual para o analista de dados, tanto na análise dos dados como para comparação das trajetórias sob os dados representativos. Adicionalmente, considerando que tal trabalho refere-se a um trabalho em andamento, nota-se a possibilidade de futuras melhorias, destinadas como próximas atividades desse projeto. Dentre essas tarefas encontra-se:

- permitir a capacidade de destacar os pontos das trajetórias identificados como relevantes, de acordo com determinado filtro realizado pelo usuário, sob alguns aspectos das trajetórias de múltiplos aspectos;
- a visualização da origem de cada ponto da MAT representativa, uma vez que o método MAT-SG possui informação do mapeamento entre a MAT representativa e a origem de cada ponto representativo, assim a Visual R-MAT poderá mostrar visualmente cada origem;
- a visualização do mapeamento pertencente aos pontos destacados na MAT representativa, permitindo identificar visualmente a origem dos pontos destacados por algum filtro, de modo a visualizar a origem desses pontos nas trajetórias MATs do dataset de entrada, possibilitando assim a visualização de quais trajetórias MATs, ou ainda quais pontos dessas trajetórias, originaram determinado filtro em relação a MAT representativa;
- por fim, outra possibilidade de melhoria refere-se a visualização dos dados aplicados sob uma imagem de satélite 2D referente a parte espacial da trajetória.

## Agradecimentos

O presente trabalho foi realizado com apoio do Programa de Educação Tutorial (PET), do Fundo Nacional de Desenvolvimento da Educação (FNDE) e da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, assim como do programa de inovação e pesquisa Horizon 2020 da União Europeia no âmbito do acordo n° 777695 (MASTER).

## Referências

Alvares, L. O., Bogorny, V., Kuijpers, B., de Macêdo, J. A. F., Moelans, B., and Vaisman, A. A. (2007). A model for enriching trajectories with semantic geographical

<sup>1</sup>disponível em <https://github.com/Tiago-O-Luz/Visual-R-MAT-V1>.

- information. In Samet, H., Shahabi, C., and Schneider, M., editors, *15th ACM International Symposium on Geographic Information Systems, ACM-GIS 2007, November 7-9, 2007, Seattle, Washington, USA, Proceedings*, page 22, New York, NY, USA. Association for Computing Machinery (ACM).
- Bogorny, V., Braz, F., Alvares, L. O. C., Siqueira, F. L., Loy, A. M., Moreno, B. N., and Times, V. C. (2012). *Introdução a trajetórias de objetos móveis*. Editora Univille.
- Erwig, M., Güting, R. H., Schneider, M., and Vazirgiannis, M. (1999). Spatio-temporal data types: An approach to modeling and querying moving objects in databases. *GeoInformatica*, 3(3):269–296.
- Lee, J.-G., Han, J., and Whang, K.-Y. (2007). Trajectory clustering: A partition-and-group framework. In *Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data, SIGMOD '07*, page 593–604, New York, NY, USA. Association for Computing Machinery.
- Machado, V. L., Mello, R. d. S., and Bogorny, V. (2022). A method for summarizing trajectories with multiple aspects. In *International Conference on Database and Expert Systems Applications*, pages 433–446. Springer.
- Mello, R. d. S., Bogorny, V., Alvares, L. O., Santana, L. H. Z., Ferrero, C. A., Frozza, A. A., Schreiner, G. A., and Renso, C. (2019). MASTER: A multiple aspect view on trajectories. *Trans. GIS*, 23(4):805–822.
- Tominski, C., Schumann, H., Andrienko, G., and Andrienko, N. (2012). Stacking-based visualization of trajectory attribute data. *IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS*, 18(12).