# A Geometry-Driven Neural Topic Model for Trip Purpose Inference

**Jiaqi Zhang**
Southern University of Science and Technology
**Zipei Fan** (fanzipei@iis.u-tokyo.ac.jp)
The University of Tokyo
**Xuan Song**
Southern University of Science and Technology
**Ryosuke Shibasaki**
The University of Tokyo

**Research Article**

**Keywords:** Urban human mobility, Trip purposes inference, Neural network, Trip embedding

**Posted Date:** May 2nd, 2023

**DOI:** https://doi.org/10.21203/rs.3.rs-2861897/v1



**Additional Declarations:** No competing interests reported.

**Version of Record:** A version of this preprint was published at GeoInformatica on August 19th, 2023. See the published version at https://doi.org/10.1007/s10707-023-00504-6.

# A Geometry-Driven Neural Topic Model for Trip Purpose Inference

Jiaqi Zhang[1], Zipei Fan[1,2†], Xuan Song[1†], Ryosuke Shibasaki[1,2]

[1]SUSTech-UTokyo Joint Research Center for Super Smart Cities, Department of Computer Science and Engineering, Southern University of Science and Technology, ShenZhen, Guangdong, China.
[2]Center for Spatial Information Science, The University of Tokyo, Tokyo, Japan.

†Corresponding author:fanzipei@iis.u-tokyo.ac.jp(Z.F.);songx@sustech.edu.cn(X.S.)

**Abstract**

Understanding urban human mobility, particularly trip purposes, is essential for optimizing traffic management, personalized recommendations, and urban planning. However, in real-world scenarios, trip purposes cannot be directly extracted from trajectory data. To address this issue, we propose a geometry-driven neural topic model for trip purpose inference. We integrate trajectory data with nearby points of interest (POI) data using a geometry-driven technique to enhance the interpretability of the results. Furthermore, our model captures the semantics and relationships of the data in a high-dimensional space and identifies latent topics representing distinct trip purposes. These learned topics are analyzed using clustering algorithms to group similar trips, enabling trip purpose inference. And we evaluate our model using the trajectory data of Shenzhen and Chengdu, and compare it with baseline models. The results demonstrate that our model performs well. Furthermore, we analyze trajectory data containing trip purpose information to gain insights into human mobility patterns and the influence of trip purposes, paving the way for potential implications and future research directions.

**Keywords:** Urban human mobility, Trip purposes inference, Neural network, Trip embedding

# 1 Introduction

In recent years, with the increasing popularity of devices equipped with Global Positioning System (GPS) has led to an accumulation of extensive human mobility data. Understanding and analyzing human mobility data has become a central focus for researchers [1–4]. Among which a key aspect is the trip purposes inference, aiming to reveal the latent motivations behind the mobility. Determining trip purposes not only provide strong support for urban planning [5, 6], traffic management [7], and policy-making [8] but also further improve the quality of life for residents.

However, human mobility behaviors after leaving the taxi are not readily available, as the trajectory data lacks information on trip purposes [9, 10], making trip purpose inference a challenging task [11, 12]. Despite this, we can still observe that individuals' mobility within a city follows certain patterns. For example, in Figure 1, during the morning hours of 8 to 10 a.m. on weekdays, a large number of trips to Shenzhen Science and Technology Park and Futian CBD business work areas. In the evening, the number of trips near residential areas gradually increases.


**Fig. 1** Heat map of working day trip destinations in Shenzhen.

The various methods have been proposed to analyze spatio-temporal trajectory data for travel purpose inference. These approach include traditional statistical method [13], machine learning [14–17], and deep learning approaches [2, 18], all of which have seen widespread application in this domain [5, 10, 19]. However, these models exhibit certain limitations when handling large volumes of spatio-temporal trajectory data. Generally, they rely on bag-of-words (BOW) representation, where each dimension corresponds to a unique word. The BOW representation accounts for the presence of features within the trajectory but disregards the relationships and order among them. This constraint prevents the models from capturing the latent semantic information and relationships within the data. Considering the rich spatial and temporal characteristics of spatio-temporal trajectory data and points of interest (POI) on maps, such limitations adversely affect the performance of these models.

To solve the above problems, we propose a geometry-driven neural topic model for inferring trip purposes from trajectories. Initially, we adopt a geometry-driven approach that takes into account the starting and ending points, duration, and distribution of POI surrounding the trajectory, thus enhancing the spatio-temporal correlations within trajectory data. Next, we utilize embedding techniques to capture the latent semantic information and relationships present in high-dimensional space, effectively addressing the challenges and limitations faced by traditional models when processing spatio-temporal data. Then, we develop a neural topic model that incorporates a geometry-driven loss function, preserving the geometric relationships between the trajectory feature distributions and trip purpose distributions, and between trajectory distributions and trip purpose distributions, This approach further improves the accuracy and interpretability of trip purpose inference. Finally, we assessed the effectiveness of our proposed method on a real-world urban trajectory dataset. By comparing our model with the baseline model using coherence and perplexity, we demonstrated the validity of our model.

The main contributions of this study can be summarized as follows:

- We propose a geometry-driven neural topic model that incorporates the trip start time, trip duration, and distribution information of nearby POI categories using a geometry-driven embedding method. This approach captures potential semantic relationships within the high-dimensional trajectory data space, enhancing the model's performance and interpretability.
- We develop a geometry-driven loss function that effectively maintains the geometric relationship between the trajectory feature distribution and the topic distribution, thereby enhancing the model's accuracy and interpretability.
- We compared our model with the baseline model using real-world datasets from Shenzhen and Chengdu, demonstrating its high performance. Additionally, the practical application potential of this model for urban planning and traffic management is illustrated through a case study.

The remainder of our paper is organized as follows. Section 2, we review the related works on human mobility and trip purpose inference and compare our work with the previous. The basic concepts and problem statements are presented in Section 3. Section 4 introduces the detailed proposed method to infer the trip purpose. Section 5 we introduce the result and comprehensive analysis. Finally, we summarize the papers and discuss future work in Section 6.

# 2 Related Work

In this section, we provide an overview of the relevant literature in the fields of human mobility analysis, trip purpose inference, and neural topic modeling, highlighting the key contributions and limitations of previous works.

## 2.1 Human Mobility Analysis

Human mobility analysis has been an active research area due to its significant implications for urban planning, traffic management, and location-based services[1]. In recent years, an increasing number of studies have focused on human mobility analysis using multi sources data and methodologies[4, 20]. These studies have provided insights into the spatial and temporal properties of human mobility, including established an effective human mobility model for prediction and simulation by studying human mobility after natural disasters [21], Fine-grained COVID-19 Propagation Model of human mobility Data [22], and human activities exhibit periodic patterns corresponding to daily and weekly cycles [23]. They have also explored the impact of social and geographical factors on mobility behaviors [24, 25].

While these studies have contributed to our understanding of human mobility, they often do not explicitly focus on inferring trip purposes, which is crucial for providing more nuanced and context-aware insights into travel behavior. Moreover, most of these studies rely on aggregated mobility patterns rather than individual trip-level data, which may not accurately represent the complex relationships between individual travel choices and urban contexts. Our work aims to address these limitations by developing a novel geometry-driven neural topic model for trip purpose inference that leverages rich trajectory data and points of interest information.

### 2.2 Trip Purpose Inference

Trip purpose inference has been a topic of increasing interest, as understanding the reasons behind individual travel behaviors can lead to better traffic management and urban planning strategies. In recent years, various techniques have been proposed to infer trip purposes, leveraging diverse data sources and machine learning methods.

Traditionally, trip purpose studies relied on statistical methods, such as logistic regression and decision trees, to infer trip purposes from survey data [26–28]. With the increasing availability of large-scale trajectory data, researchers have developed data-driven approaches, including clustering algorithms [29], topic models [5], and deep learning methods [30] to extract trip purpose form GPS data. However, these methods often fail to capture the complex semantic and spatial relationships between trajectory data and points of interest, limiting their accuracy and interpretability.

Our work addresses these limitations by proposing a novel geometry-driven neural topic model for trip purpose inference. By incorporating travel start time, travel duration, and nearby POI category distribution information with geometry-driven technique, we enhance the potential temporal and spatial correlations within trajectory data, and our model employs an embedding method to capture the latent semantics and relationships within high-dimensional data. Enabling more accurate and interpretable trip purpose inference.

## 3 Preliminaries

Before delving into the details of our proposed geometry-driven neural topic model for trip purpose inference, it is essential to introduce the preliminary concepts and notations used in this study.

### 3.1 Definition

**Definition 1** (Trajectory)**.** *The trajectory data consists of a series of spatio-temporal points representing motion. Formally, a trajectory $T$ can be expressed as*

$T_i = (s_1, s_2, \ldots, s_i)$. *Where the s represent the trip point, it can be expressed as* $s_j = (x_j, y_j, t_j)$*, the x and the y represents the geographical coordination (latitude, longitude), t represent the timestamp.*
**Definition 2** (Origin-Destination (OD) Pair)**.** *Origin-Destination (OD) Pair refers to the starting point (origin) and ending point (destination) of a trip. It is denoted as a tuple,* $T_{od} = (s_o, s_d)$.
**Definition 3** (Point of Interest (POI))**.** *Points of interest (POI) data refers to the information about specific locations or establishments in a given geographic area, such as restaurants, shops, parks, or tourist attractions. Each POI is associated with geographic coordinates (latitude, longitude), a category (e.g., restaurant, hotel), and additional attributes, such as the name, address, and opening hours. So the POI point can be represented as:* $N_{POI}(x_{POI}, y_{POI}, c_{POI}, a_{POI})$*. where* $(x_{POI}, y_{POI})$ *are the geographic coordinates,* $c_{POI}$ *is the category, and* $a_{POI}$ *represents additional attributes of the POI.*

## 3.2 Problem Statement

**Given** a set of trajectories $M = (T_1, T_2, ..., T_i)$ and city POIs $V = (N_1, N_2, ..., N_{POI})$. **The goal** is to extract latent information from trajectory and POI data to facilitate trip purpose inference. By simulating human mobility integrated with trip purpose information, the results can better inform residents and city managers.

## 3.3 Topic Model

Before introducing the geometry-driven neural topic model, we provide a brief overview of traditional topic models, such as Latent Dirichlet Allocation (LDA) [31]. In this model, it is assumed that documents are generated from a mixture of topics, with each topic characterized by a distribution over words. As illustrated in Figure 2, the generative process can be described as follows:

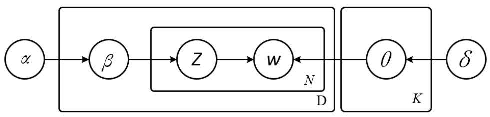


**Fig. 2** LDA graphical model.

For each topic $K$, generate a topic distribution $\theta$ from a Dirichlet distribution using parameter $\delta$. For each word $w_i$ in the document $D$, generate a word-topic distribution $Z_{w_n}$ from a Dirichlet distribution using parameter $\beta$, generate a word $w_n$ from the $p(w_n|Z_n\beta)$ a multinomial probability by $Z_n$.

The goal is to estimate the posterior distribution of latent variables, given the observed data, in order to uncover the underlying semantic structure of the documents.

We establish an analogy between trajectory data and text data. As illustrated in Table 1, we can consider trip feature attributes as words, each trip as a document, and

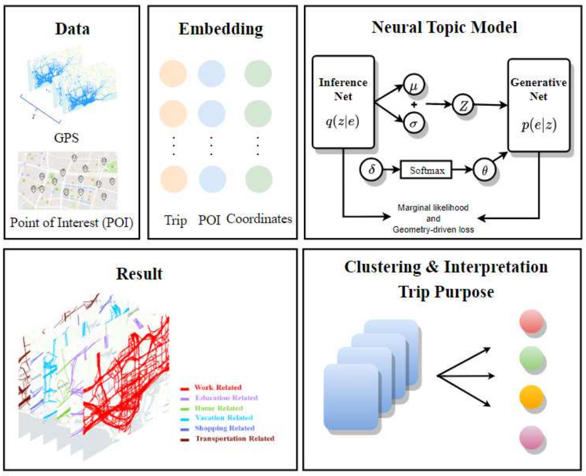


**Fig. 3** An overview of the proposed geometry-driven neural topic model for trip purpose inference.

**Table 1** Analogy trip purpose to topic model.

| | | |
|---|---|---|
| The trajectory feature attribute | $\rightarrow$ | Words |
| A trip trajectory | $\rightarrow$ | Documents |
| The trip purpose | $\rightarrow$ | Topics |

the trip purpose as the topic. This analogy enables the application of topic modeling techniques to infer trip purposes.

However, traditional topic models, such as LDA, may not be well-suited for complex spatio-temporal trajectory data, because they do not consider the spatial and temporal correlations, and their limited expressiveness. Therefore, we developed a trip purpose inference method based on neural topic model to make travel purpose inference more accurate.

# 4 Methodology

In this section, we present the geometry-driven neural topic model for trip purpose inference. Figure 3 offers an overview of the proposed geometry-driven neural topic model. Our model is an variational autoencoder, which employs an encoder-decoder based framework [32, 33]. In the next, we will describe the details of our model.

## 4.1 Feature Extraction

The spatial and temporal context of the trajectory data by integrating travel start time, travel duration, and nearby POI category distributions into the generative process. This allows the model to learn complex semantic information in high-dimensional space, effectively capturing the spatial and temporal correlations within trajectory data.

**Temporal feature:** The trip temporal Extract pertinent temporal features from the trajectory data, such as start time, end time, travel duration, day of the week (weekday or weekend), and whether the day is a holiday.

**POI feature:** The nearby POI category distribution for the point $s_i$ is an important feature to consider when analyzing the trip's spatial information, as it helps capture the characteristics of the surrounding environment form the trip.

In order to calculate the nearby POI category distribution, we first need to define a radius $r$ around the points $s_i$. This radius represents the area of interest within which we will consider the POIs.

We compute the nearby POI category distribution for the trip start point ($s_o$) and end point ($s_d$). For each point, we calculate the proportion of each POI category within a certain radius ($r$) from the point. The POI category distribution can be represented as a vector $P_s = (p_1, p_2, \ldots, p_C)$, where $n$ is the number of POI categories, and $p_c$ is the proportion of POIs belonging to category $c$ within the radius $r$ as (1):

$$p = \frac{N_c}{\sum_{j=1}^{n} N_j} \tag{1}$$

Where $N_c$ is the number of POIs of category $c$ within the radius $r$, and $\sum_{j=1}^{n} N_j$ is the total number of POIs of all categories within the radius $r$.

## 4.2 Modeling

### 4.2.1 Embedding

Traditional topic models employ the bag-of-words (BOW) method, treating each document as an unordered collection of words and disregarding syntactic or semantic information [31]. This approach results in high-dimensional feature vectors, where each dimension represents a unique word in the vocabulary. As the vocabulary grows, the model's effectiveness diminishes. To address this issue, and inspired by word embedding techniques [34], we utilize the embedding approach to represent trip features in an L-dimensional space, capturing the complex semantics and relationships in a high-dimensional space.

**feature embedding** Let $T_{features}$ represent the trip feature vector containing the trip origin point, trip destination point, destination time, day of the week, weekday or weekend indicator, and holiday indicator, POI category distribution feature vector. Then embedding the representation of the trip feature, as (2):

$$E_T = F_\phi(T_{feature}) \tag{2}$$

Where $\phi$ is the parameters of trip feature embedding.

### 4.2.2 Inference Net

Following the earlier work[35, 36], the inference net employs Gated Recurrent Unit (GRU) network to learn temporal features correlation and a Convolution Neural Network (CNN) to learn the spatial features correlation. The objective is to obtain latent representations of trip purposes by capturing spatio-temporal relationships in the data.

The output of the GRU layer is a hidden state vector that captures the temporal information in each trip, The equations for the GRU layer are as follows (3):

$$u_t = \sigma(W_u \cdot [h_{t-1}, e_t] + b_u) \tag{3}$$
$$r_t = \sigma(W_r \cdot [h_{t-1}, e_t] + b_r) \tag{4}$$
$$\tilde{h}_t = \tanh(W_h \cdot [r_t \odot ht - 1, e_t] + b_h) \tag{5}$$
$$h_t = (1 - u_t) \odot h_{t-1} + u_t \odot \tilde{h}_t \tag{6}$$

Here, $W_u$, $W_r$, and $W_h$ are weight matrices, $b_u$, $b_r$, and $b_h$ are bias terms, and $\sigma$ is the sigmoid activation function. The element-wise multiplication is denoted by $\odot$. Then, through the CNNs,it equations as (7):

$$h_{cnn} = f(W * h_T + b) \tag{7}$$

Where $h_{out}$ is the feature map, $f$ is the ReLU activation function, $W$ is the convolutional filter, $b$ is the bias term, and $*$ denotes the convolution operation. Then, following a fully connected layer, we obtain the feature representation which captures both temporal and spatial information, as (8):

$$h_{final} = \mathrm{FC}(h_{cnn}) \tag{8}$$

Fed it into an encoder part, which outputs the parameters of the approximate posterior distribution $q(z)|h_{final})$, where $z$ is the latent trip purpose representation.

$$q(z|\mathbf{h}_{final}) = \mathcal{N}(z; \mu, \sigma^2, \epsilon) \tag{9}$$

Where $\epsilon \sim \mathcal{N}(0, \boldsymbol{I})$ is a noise vector.

### 4.2.3 Generative Net

Our generative net primary goal of reconstruct the input trip embedding based on the latent trip purpose representation $z$ inferred by the inference network, while preserving the geometric properties.

For a trip purpose $z_k$, we draw the trip feature embedding distribution with $\theta_k$, where $\theta_k \sim \mathcal{LN}(0, \boldsymbol{I})$.

The generative net takes the latent trip purpose representation $z$ as input and generates a reconstructed fused feature representation $\hat{h}_{final} \sim p(h_{final}|z_k)$. Then, through a fully connected layer, inverse CNN and inverse GRU, we obtain the reconstructed embedding matrix.

### 4.2.4 Estimation

To train our model, we employ a combined loss function, the marginal likelihood function and the geometric surrogate loss function, which enables our model to learn topic distributions from trajectory feature embeddings while preserving geometric relationships between features, as shown in (10).

$$\mathcal{L}_{total} = \mathcal{L}_{GD} + \lambda \mathcal{L}_{ML} \tag{10}$$

The marginal likelihood function demands maximization, while the geometric surrogate loss function calls for minimization. To address these optimization objectives, we introduce the regularization parameter $\lambda$.

**Geometry-driven loss function** The geometry-driven loss term, $\mathcal{L}geo$, aims to minimize the distance between distributions generating POIs with identical trip purposes. This loss term encourages the model to learn a latent trip purpose representation that more effectively captures the relationships between trip purposes and the spatial distribution of POIs in the surrounding environment. Consequently, we obtain the geometry-driven loss term, $\mathcal{L}geo$.

Trip Purpose Similarity: For Two trip under the same topic, the similarity is calculated using cosine similarity, a common metric that measures the similarity between two vectors by considering the angle between them rather than their magnitudes. The cosine similarity is computed based on their latent trip purpose representations $z_{T_i}$ and $z_{T_j}$, as shown in (11):

$$S_{ij} = \frac{z_{T_i} \cdot z_{T_j}}{\|z_{T_i}\| \|z_{T_j}\|} \tag{11}$$

Spatial Distribution Similarity: The similarity between their POI distributions is computed using Euclidean distance, an effective metric for capturing the relationship between the spatial distribution of POIs in different trips, as shown in (12):

$$D_{ij} = \sum_{c=1}^{n} (p_c^{T_i} - p_c^{T_j})^2 \tag{12}$$

Where $p_c^{T_i}$ and $p_c^{T_j}$ represent the true POI category distributions for the two trips, by (1) and $n$ denotes the number of POI categories.

Thus, the geometry-driven loss term can be obtained by minimizing the difference between trip purpose similarity and spatial distribution similarity for each trajectory pair, as shown in (13):

$$\mathcal{L}_{GD} = \sum_{i=1}^{T} \sum_{j=1}^{T} \|S_{ij} - D_{ij}\|^2 \tag{13}$$

By minimizing the geometry-driven loss term, the model becomes more adept at capturing the relationship between trip purposes and the spatial distribution of POIs.

**The marginal likelihood**. We maximize the log-marginal likelihood to optimize the parameters, as (14):

$$p_\theta(h_{final}) = \int p_\theta(h_{final}|z) p(z) d\boldsymbol{z} \tag{14}$$

Since it is intractable to compute the integral directly, we use the Evidence Lower Bound (ELBO) as a surrogate objective:

$$\mathcal{L}_{ML} = \mathbb{E}q_\phi(\boldsymbol{z}|\boldsymbol{h}_{final})[\log p_\theta(\boldsymbol{h}_{final}|\boldsymbol{z})] - D_{KL}(q_\phi(\boldsymbol{z}|\boldsymbol{h}_{final})||p(\boldsymbol{z})) \tag{15}$$

Finally, we utilize the random ELBO to optimized according to model parameter and variational parameter. Set the learning rate with Adam optimizer [37], a generalized solver for neural network models, to optimize our model.

# 5 Experiment

## 5.1 Data

We utilize a real-world trajectory dataset gathered from Shenzhen taxi and Didi Chengdu. The dataset is presented in Table 2.

**Table 2** Statistics of datasets.

| City | Taxi | Occupied Trip | Effective days | Average duration(min) | POI |
|---|---|---|---|---|---|
| Shenzhen | 20898 | 1706676 | 15 | 13 | 1213656 |
| Chengdu | 33205 | 1937720 | 15 | 25 | 128376 |

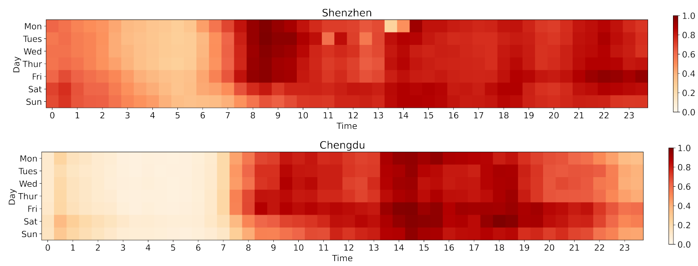


**Fig. 4** Passenger drop-off time distribution.

We set 30 minutes as the time bin. Following definition, we construct the number of visits for each city area within these time bins. Figure 4 displays the passenger drop-off time distribution for Shenzhen and Chengdu. A distinct peak is observed during morning and evening hours for residents' weekday city trips, suggesting that the trip purpose of such trajectories is likely related to work and home.

## 5.2 Baseline Model

To demonstrate the effectiveness of our proposed method, we compare its performance against baseline models:

1. **Bayesian Model**: This model represents a traditional probabilistic approach to trip purpose inference, where trip purposes are modeled as latent variables that follow a certain probabilistic distribution. We use this baseline to assess the improvement provided by our proposed method over a classical statistical method.

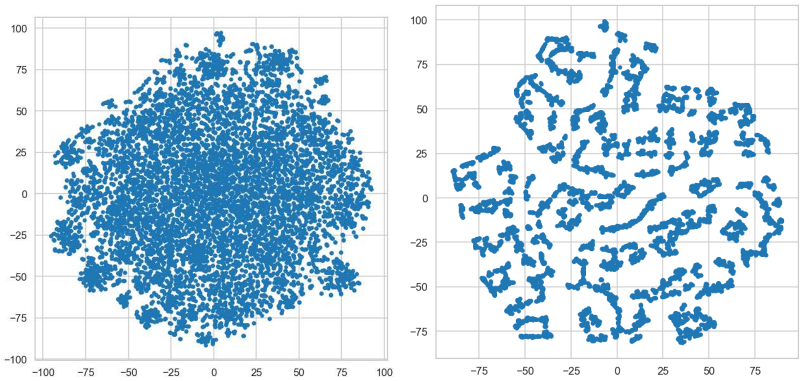


**Fig. 5** Visualization results: before (left) and after (right) applying the embedding Method.

2. **Latent Dirichlet Allocation (LDA)**: A generative probabilistic topic model widely used for text data analysis.
3. **Deep Embedded clustering (DEC)**: A deep learning-based clustering method that learns an embedding space for the data.

To compare the performance of our propose model the geometry-driven neural topic model, we use the perplexity metric and coherence, which measures how well the model generalizes to unseen data [31, 33, 38, 39].

**perplexity** Lower perplexity values indicate better performance. Perplexity is defined as the inverse of the geometric mean of the likelihood of the test data:

$$PP = \exp\left(\frac{1}{N}\sum \log \mathcal{L}\right) \tag{16}$$

Where $\mathcal{L}$ means the likelihood function, and $N$ is the number of trajectories.

**Coherence** Higher coherence score indicates that the words within a topic are more semantically related and the topic is more interpretable.

The coherence measure is calculated as follows:

$$C(T) = \frac{1}{|T|^2 - |T|} \sum_{i=1}^{|T|} \sum_{j=1, j\neq i}^{|T|} \log \frac{P\left(E_i, E_j\right) + \epsilon}{P\left(E_i\right) P\left(E_j\right)} \tag{17}$$

Where $T$ is the set of top feature for a given trip purpose, $E_i$ and $E_j$ are feature within the set, $P(E_i, E_j)$ is the joint probability of feature $E_i$ and feature $E_j$ occurring together in the same context, and $P(E_i)$ and $P(E_j)$ are the individual probabilities of feature $E_i$ and $E_j$. The term $\epsilon$ is a small constant added to avoid division by zero.

The coherence of a trip purpose is the average coherence score of all pairs of terms within the trip purpose, and the model coherence is the average coherence of all trip purpose in the model.

### 5.3 Results

#### 5.3.1 Effectiveness of Embedding

To evaluate the effectiveness of our embedding process in capturing the underlying structure and patterns of the data, we employ t-Distributed Stochastic Neighbor Embedding (t-SNE) for visualization[40]. t-SNE is a nonlinear dimensionality reduction technique particularly well-suited for visualizing high-dimensional data, as it preserves local structures and relations in the data. Figure 5 displays the visualization results, allowing us to assess how well the embedding captures the semantics and relationships between different trajectories.

#### 5.3.2 Result Comparison

Table 3 and 4 presents the results of our proposed model and all baseline models. Meanwhile, we also get the performance of our model without the geometric-driven part.

**Table 3** Comparison of the coherence with baseline model

| Model | Shenzhen | Chengdu |
|---|---|---|
| Bayesian based | 0.45 | 0.41 |
| LDA based | 0.49 | 0.46 |
| DEC based | 0.55 | 0.54 |
| **Ours - without GD** | **0.57** | **0.55** |
| **Ours** | **0.69** | **0.63** |

**Table 4** Comparison of the perplexity with baseline model and different number of trip purpose.

| | Shenzhen | | | | | |
|---|---|---|---|---|---|---|
| K | 3 | 4 | 5 | 6 | 7 | 8 |
| Bayesian based | 235.46 | 220.65 | 211.73 | 199.65 | 206.72 | 209.81 |
| LDA based | 215.62 | 200.27 | 193.12 | 182.31 | 189.55 | 200.97 |
| DEC based | 205.53 | 198.44 | 190.14 | 187.22 | 188.82 | 197.44 |
| **Ours-without GD** | **203.55** | **194.46** | **185.66** | **175.48** | **183.91** | **186.15** |
| **Ours** | **181.55** | **182.46** | **173.66** | **157.48** | **165.91** | **176.15** |
| | Chengdu | | | | | |
| K | 3 | 4 | 5 | 6 | 7 | 8 |
| Bayes Based | 240.22 | 224.82 | 210.11 | 204.53 | 207.72 | 211.53 |
| LDA Based | 234.44 | 217.92 | 203.35 | 197.12 | 205.77 | 209.54 |
| DEC based | 210.32 | 201.53 | 194.02 | 198.93 | 206.23 | 210.33 |
| **Ours-Without GD** | **205.12** | **196.33** | **185.33** | **184.02** | **188.82** | **193.52** |
| **Ours** | **184.12** | **186.33** | **179.33** | **181.02** | **186.82** | **189.52** |

The research findings indicate that our proposed model exhibits exceptional consistency and performance across various datasets when compared to baseline models.

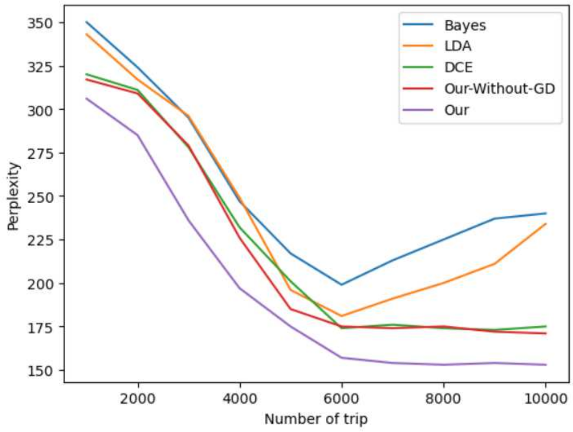


**Fig. 6** Comparison of the perplexity under various number of trips with baseline model.

In addition, we compare our model with the baseline model for various numbers of trips, as illustrated in Figure 6. We observe that our model exhibits improvement compared to traditional Bayesian and LDA topic based models. Without incorporating the geometric-driven loss, our model performs comparably to deep embedded clustering. However, as we increase the calculation of the geometric-driven loss, we find that our model's effectiveness further improves. This observation suggests that the geometric-driven loss function is important in preserving the geometric relationships between input trajectory feature embeddings and the learned trip purpose distributions.

### 5.3.3 Trip Purpose Inference Analysis

As depicted in Figure 7, the distribution of trajectory endpoints across different trip purposes reveals various patterns. For the *Work-Related* trip purpose, numerous trajectory endpoints are concentrated around the Nanshan Science and Technology District and the Futian CBD area. The *Shopping-Related* trip purpose features many trajectories ending in the Huaqiang North and Dongmen areas, which house numerous shopping malls. In the case of the *Home-Related* trip purpose, a significant number of trajectories terminate in the Meilin Zone and Baishizhou Zone, which are known as famous urban villages in Shenzhen, suggesting that the majority of these trajectories are generated by individuals returning home. Trajectory endpoints around the university town are most likely associated with the *Education-Related* trip purpose, while those in the vicinity of Dameisha and Happy Valley can be attributed to the *Vacation-Related* purpose.

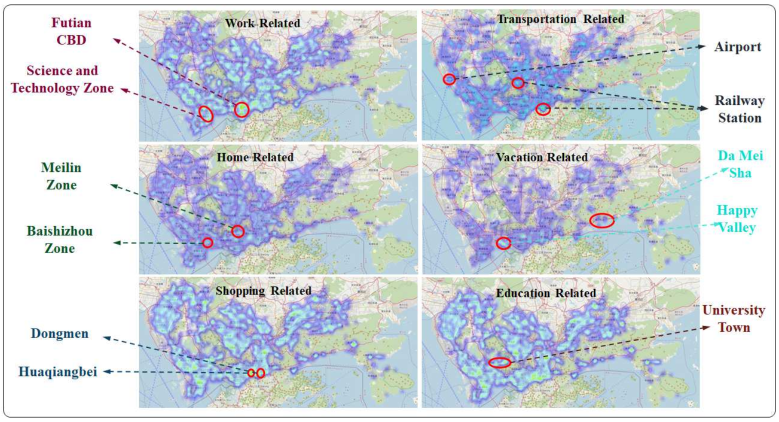


**Fig. 7** The Distribution of trajectory ending under different trip purposes in Shenzhen.

Subsequently, we examine the distribution of each trajectory's end time under various trip purposes. Additionally, we interpret our results by considering the distribution of POIs in specific regions and the proportion of different trip purposes in those areas, as depicted in Figure 8.

Figure 8 **A** illustrates the distribution of trajectory end times for different trip purposes. Topic 1 is concentrated between 8:00 and 10:00, Topic 2 between 6:00-9:00, 14:00-16:00, and 18:00-20:00, Topic 3 between 10:00 and 18:00, Topic 4 after 19:00, and Topics 5 and 6 during daytime hours. We utilize the POI feature, which helps us identify Topic 1 as *Work Related* due to the high concentration of companies and typical work start times between 9:00 and 10:00 in Shenzhen. Topic 2 is *Education Related*, as students generally attend school earlier than work hours and leave school between 14:00 and 15:00, as well as after 18:00. Topic 3 is *Shopping Related* since shopping malls typically operate from 10:00 to 22:00, attracting citizens during these hours. In Shenzhen, people generally finish work after 19:00, with many employees leaving between 22:00 and 23:00, making Topic 4 *Home Related*. Topic 5, with trajectories concentrated in vacation areas as shown in Figure 7, is *Vacation Related*. While Topic 6, with trajectories clustered around city transportation hubs, such as the railway station and airport, is *Transportation Related*.

Figure 8 **B** features the *Science and Technology Zone*, home to numerous high-tech enterprises and known as China's Silicon Valley. This area attracts many workers, resulting in a high proportion of Topic 1 trajectories.

The *University Town* area, shown in Figure 8 **C**, houses several educational institutions and has a significant distribution of educational POIs. Consequently, Topic 2 is more prevalent in this area than other trip purposes.

The *Dongmen Area*, depicted in Figure 8 **D**, is a well-known shopping district in Shenzhen's Luohu District. Numerous shopping malls are supported by the POI distribution, and the number of individuals visiting malls increases sharply between 10:00 and 19:00, declining as the malls close. This explains the area's high proportion of Topic 3 trajectories.

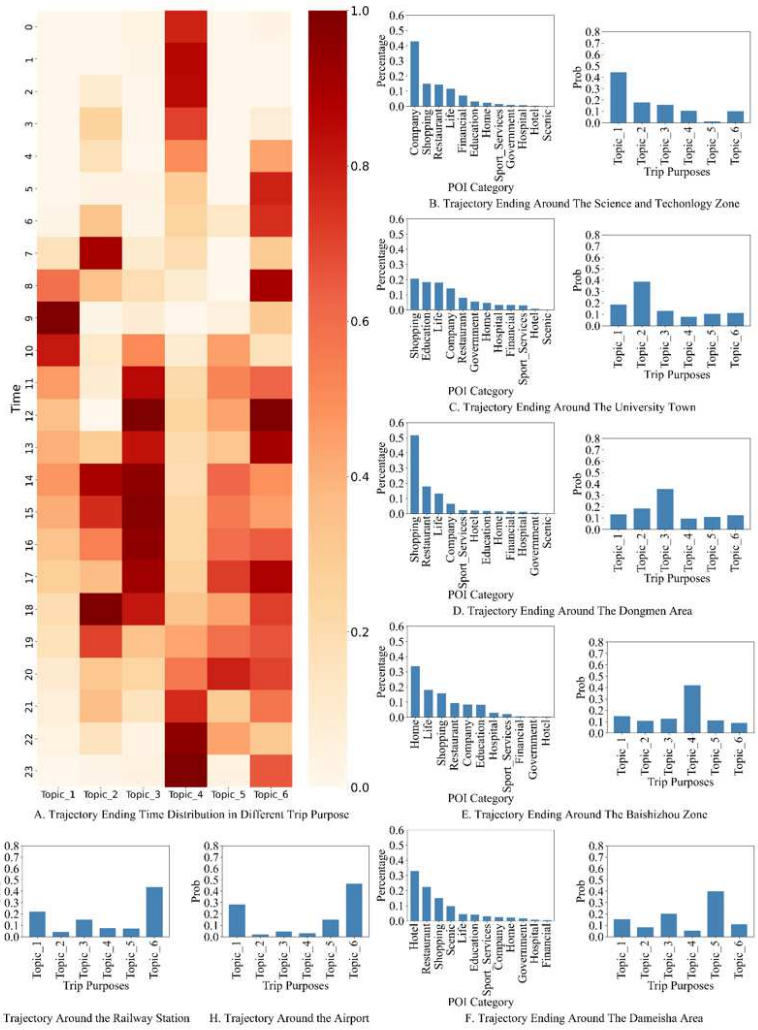


**Fig. 8** The Distribution of Trajectory Ended Time in Different Trip Purposes And Some Typical Region Cases

The *Baishizhou Zone*, shown in Figure 8 **E**, is a typical residential district in Shenzhen with many home-related POIs. Most trajectories in this area belong to Topic 4, indicating a high influx of people after 19:00.

The *Dameisha Area*, illustrated in Figure 8 **F**, is a famous vacation destination in Shenzhen, renowned for its beautiful beaches. The POI distribution reveals numerous hotels, restaurants, and scenic spots in this area, contributing to the high proportion of Topic 5 trajectories.

Figures 8 **G** and **H** display the largest and most comprehensive railway transportation hub and the only international airport in Shenzhen, respectively. As a result, most trajectories ending in these areas belong to Topic 6.

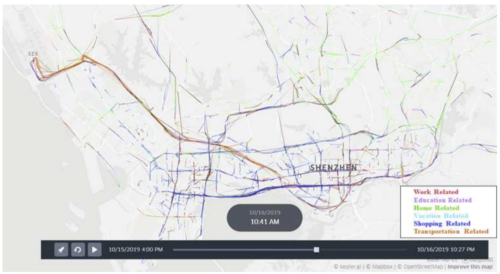


**Fig. 9** The Simulation Results of One Day in Shenzhen.

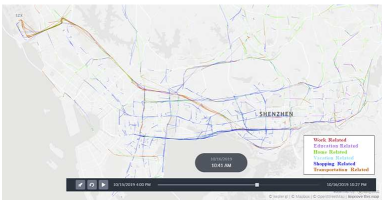


**Fig. 10** The Simulation Results After the Issuance Policy About *Telework*.

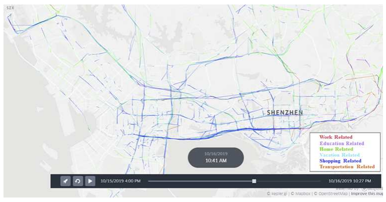


**Fig. 11** The Simulation Results After the Issuance Policy About *Do Not Leave Shenzhen*.

### 5.3.4 Case Study

We conducted simulations of human mobility in Shenzhen. After applying the geometry-driven topic model, we identified six distinct trip purposes. By using colors to differentiate these purposes, we were able to simulate human mobility patterns within the city.

Figure 9 presents a simulation of human mobility in Shenzhen, demonstrating that the majority of movement is concentrated along Shennan Avenue, Beihuan Avenue, and Binhai Avenue, the city's primary east-west arterial roads. Notably, considerable human mobility is observed near Nanshan Science and Technology Park, Futian CBD, and Luohu Dongmen Street. Our model assigns an orange color to the *Transportation Related* trip purpose, emphasizing trajectories that end around the airport and Shenzhen North railway station areas. This observation further substantiates the effectiveness of our proposed model.

By adjusting simulation parameters, our model can aid government officials in understanding urban population flow during events such as COVID-19-like infectious disease outbreaks. The model can simulate the impact of various policies, including the effects of a *Telework* policy on urban population flow, as depicted in Figure 10. Additionally, it can simulate human mobility within the city under a *Do Not Leave Shenzhen* policy, as shown in Figure 11. Consequently, this tool can assist city managers in formulating more accurate and targeted prevention and control policies.

In conclusion, our proposed geometry-driven neural topic model effectively addresses the limitations of traditional topic models in handling complex spatio-temporal trajectory data and demonstrates its superiority in trip purpose inference. The experimental results confirm the effectiveness of our method in terms of both performance and interpretability, providing valuable insights into human mobility patterns and potential implications for urban planning and traffic management.

# 6 Conclusion

we propose a geometry-driven neural topic model for trip purpose inference. Our model addresses the limitations of traditional models by incorporating geometry-driven feature extraction and trip embedding methods. Specifically, our approach combines trajectory with nearby POIs using a geometry-driven method, allowing a more comprehensive understanding of the complex relationships between trajectory data points. Simultaneously, the trip embedding method captures latent semantic and relational information between trajectory data in high-dimensional space, enabling our model to effectively capture inherent spatial and temporal correlations. This results in more accurate trip purpose inference.

The result demonstrate the effectiveness of our proposed method compared to baseline approach, exhibiting improvements in both perplexity and model coherence. The case study further underscores the practical applicability of our model in real-world urban scenarios.

In future work, we plan to extend our model in several directions. Firstly, we aim to explore advanced embedding techniques to better capture the latent semantics and relationships inherent in trajectory data. Secondly, by combining the semantic information and topic information of the trajectory, we intend to identify the anomaly trajectory. Finally, we strive to contribute valuable insights and tools for urban planning, traffic management, and location-based services.

# 7 Declarations

## 7.1 Competing Interests

The authors have no competing interests to declare that are relevant to the content of this article.

## 7.2 Authors' Contributions

Jiaqi Zhang and Zipei Fan wrote the main manuscript text. All authors reviewed the manuscript.

## 7.3 Funding

This work was partially supported by National Key Research and Development Project (2021YFB1714400) of China, Guangdong Provincial Key Laboratory (2020B121201001), and the Grant in-Aid for Scientific Research B (22H03573) of Japan Society for the Promotion of Science (JSPS).

## 7.4 Availability of Data and materials

This paper did not report any data and materials.

## References

[1] Zheng, Y., Capra, L., Wolfson, O., Yang, H.: Urban computing: concepts, methodologies, and applications. ACM Transactions on Intelligent Systems and Technology (TIST) **5**(3), 1–55 (2014)

[2] Wang, P., Fu, Y., Liu, G., Hu, W., Aggarwal, C.: Human mobility synchronization and trip purpose detection with mixture of hawkes processes. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 495–503 (2017)

[3] Faroqi, H., Mesbah, M.: Inferring trip purpose by clustering sequences of smart card records. Transportation Research Part C: Emerging Technologies **127**, 103131 (2021)

[4] Jiang, R., Cai, Z., Wang, Z., Yang, C., Fan, Z., Chen, Q., Tsubouchi, K., Song, X., Shibasaki, R.: Deepcrowd: A deep model for large-scale citywide crowd density and flow prediction. IEEE Transactions on Knowledge and Data Engineering **35**(1), 276–290 (2021)

[5] Yuan, J., Zheng, Y., Xie, X.: Discovering regions of different functions in a city using human mobility and pois. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 186–194 (2012)

[6] Zheng, Y., Liu, L., Wang, L., Xie, X.: Learning transportation mode from raw gps data for geographic applications on the web. In: Proceedings of the 17th International Conference on World Wide Web, pp. 247–256 (2008)

[7] Shao, E., Wang, H., Feng, J., Xia, T., Yang, H., Geng, L., Jin, D., Li, Y.: Deepflowgen: Intention-aware fine grained crowd flow generation via deep neural networks. IEEE Transactions on Knowledge and Data Engineering **34**(12), 5693–5707 (2021)

[8] Deng, J., Chen, X., Jiang, R., Song, X., Tsang, I.W.: St-norm: Spatial and temporal normalization for multi-variate time series forecasting. In: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, pp. 269–278 (2021)

[9] Gonzalez, M.C., Hidalgo, C.A., Barabasi, A.L.: Understanding individual human mobility patterns. Nature **453**(7196), 779–782 (2008)

[10] Wang, D., Pedreschi, D., Song, C., Giannotti, F., Barabasi, A.-L.: Human mobility, social ties, and link prediction. In: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1100–1108 (2011)

[11] Song, C., Koren, T., Wang, P., Barabási, A.-L.: Modelling the scaling properties of human mobility. Nature physics **6**(10), 818–823 (2010)

[12] Yuan, J., Zheng, Y., Xie, X., Sun, G.: Driving with knowledge from the physical world. In: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 316–324 (2011)

[13] Miller, H.J., Bridwell, S.A.: A field-based theory for time geography. Annals of the Association of American Geographers **99**(1), 49–75 (2009)

[14] Pelekis, N., Kopanakis, I., Kotsifakos, E.E., Frentzos, E., Theodoridis, Y.: Clustering uncertain trajectories. Knowledge and information systems **28**, 117–147 (2011)

[15] Zheng, Y., Xie, X., Ma, W.-Y., *et al.*: Geolife: A collaborative social networking service among user, location and trajectory. IEEE Data Eng. Bull. **33**(2), 32–39 (2010)

[16] Zhou, C., Frankowski, D., Ludford, P., Shekhar, S., Terveen, L.: Discovering personally meaningful places: An interactive clustering approach. ACM Transactions on Information Systems (TOIS) **25**(3), 12 (2007)

[17] Ying, J.J.-C., Lu, E.H.-C., Lee, W.-C., Weng, T.-C., Tseng, V.S.: Mining user similarity from semantic trajectories. In: Proceedings of the 2nd Acm Sigspatial International Workshop on Location Based Social Networks, pp. 19–26 (2010)

[18] Feng, J., Li, Y., Zhang, C., Sun, F., Meng, F., Guo, A., Jin, D.: Deepmove: Predicting human mobility with attentional recurrent networks. In: Proceedings of the 2018 World Wide Web Conference, pp. 1459–1468 (2018)

[19] Zhao, K., Musolesi, M., Hui, P., Rao, W., Tarkoma, S.: Explaining the power-law distribution of human mobility through transportationmodality decomposition. Scientific reports **5**(1), 1–7 (2015)

[20] Jiang, R., Song, X., Fan, Z., Xia, T., Wang, Z., Chen, Q., Cai, Z., Shibasaki, R.: Transfer urban human mobility via poi embedding over multiple cities. ACM Transactions on Data Science **2**(1), 1–26 (2021)

[21] Song, X., Zhang, Q., Sekimoto, Y., Shibasaki, R., Yuan, N.J., Xie, X.: Prediction and simulation of human mobility following natural disasters. ACM Trans. Intell. Syst. Technol. **8**(2) (2016). https://doi.org/10.1145/2970819

[22] Fan, Z., Song, X., Liu, Y., Zhang, Z., Yang, C., Chen, Q., Jiang, R., Shibasaki, R.: Human mobility based individual-level epidemic simulation platform. SIGSPATIAL Special **12**(1), 34–40 (2020). https://doi.org/10.1145/3404820.3404826

[23] Song, X., Zhang, Q., Sekimoto, Y., Shibasaki, R.: Prediction of human emergency

behavior and their mobility following large-scale disaster. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 5–14 (2014)

[24] Jurdak, R., Zhao, K., Liu, J., AbouJaoude, M., Cameron, M., Newth, D.: Understanding human mobility from twitter. PloS one **10**(7), 0131469 (2015)

[25] Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., Ratti, C.: Geo-located twitter as proxy for global mobility patterns. Cartography and Geographic Information Science **41**(3), 260–271 (2014)

[26] Pas, E.I., Koppelman, F.S.: An examination of the determinants of day-to-day variability in individuals' urban travel behavior. Transportation **14**(1), 3–20 (1987)

[27] Stopher, P.R., Greaves, S.P.: Household travel surveys: Where are we going? Transportation Research Part A: Policy and Practice **41**(5), 367–381 (2007)

[28] Gao, G., Wang, Z., Liu, X., Li, Q., Wang, W., Zhang, J.: Travel behavior analysis using 2016 qingdao's household traffic surveys and baidu electric map api data. Journal of Advanced Transportation **2019** (2019)

[29] Zheng, V.W., Zheng, Y., Xie, X., Yang, Q.: Collaborative location and activity recommendations with gps history data. In: Proceedings of the 19th International Conference on World Wide Web, pp. 1029–1038 (2010)

[30] Yao, H., Wu, F., Ke, J., Tang, X., Jia, Y., Lu, S., Gong, P., Ye, J., Li, Z.: Deep multi-view spatial-temporal network for taxi demand prediction. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 32 (2018)

[31] Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. Journal of machine Learning research **3**(Jan), 993–1022 (2003)

[32] Kingma, D.P., Welling, M., *et al.*: An introduction to variational autoencoders. Foundations and Trends® in Machine Learning **12**(4), 307–392 (2019)

[33] Blei, D.M., Kucukelbir, A., McAuliffe, J.D.: Variational inference: A review for statisticians. Journal of the American statistical Association **112**(518), 859–877 (2017)

[34] Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., Dean, J.: Distributed representations of words and phrases and their compositionality. Advances in neural information processing systems **26** (2013)

[35] Guo, Y., Liao, W., Wang, Q., Yu, L., Ji, T., Li, P.: Multidimensional time series anomaly detection: A gru-based gaussian mixture variational autoencoder approach. In: Asian Conference on Machine Learning, pp. 97–112 (2018). PMLR

[36] Srivastava, A., Sutton, C.: Autoencoding variational inference for topic models. arXiv preprint arXiv:1703.01488 (2017)

[37] Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)

[38] Jacobi, C., Van Atteveldt, W., Welbers, K.: Quantitative analysis of large amounts of journalistic texts using topic modelling. In: Rethinking Research Methods in an Age of Digital Journalism, pp. 89–106. Routledge, ??? (2018)

[39] Röder, M., Both, A., Hinneburg, A.: Exploring the space of topic coherence measures. In: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, pp. 399–408 (2015)

[40] Van der Maaten, L., Hinton, G.: Visualizing data using t-sne. Journal of machine learning research **9**(11) (2008)