# Introduction

Background:

In urban traffic, road users are often found moving in groups. These groups can be formed for different reasons. For instance, social connections (e.g. friends, couples, families) between pedestrians; mixed groups formed by traffic regulations, i.e. road users who follow the same phase of traffic lights, etc. The members of the same group interact differently to other road users in comparison to individuals [1], and they tend to keep similar speed and appropriate distance.[4]

An obvious benefit that comes from grouping is safety. Being in a group creates a buddy system where people can look after one another on the streets. \cite{jacobsen2015safety} found that people walking and bicycling in larger groups are less likely to be injured by motorists because the motorists are more cautious with groups. It will also have a beneficial effect on traffic planning: if groups are formed, this leads to an reduction in the number of road users that have to be included in computations, thus leading to a decrease in computational complexity for later applications, e.g. for traffic simulation and pedestrian navigation.

Motivation: ???

Although forming groups are common and natural for road users in normal traffic senerios, there is only very few researches on how to form groups in the shared space, where the traffic features (e.g. curbs, zebralines and traffic lights) are removed to minimizes demarcations between vehicles and pedestrians. Some studies suggested an increased risk at higher traffic volumes in shared spaces \cite{quimby2010review}, \cite{reid2009dft}. Later, \cite{holmes2015accident} launched a survey to find out about people’s experiences of using shared spaces in towns and cities. Pedestrians felt strongly that drivers did not recognize a shared space and were not slowing down to allow people to cross. Problems were pronounced in areas with high volumes of traffic or through traffic. Apart from safety aspects, currently shared spaces have efficiency problems as well: the bottleneck effect happens when traffic density is high. With the backgroud above, we can conclude that the formation of road user groups before and during crossing will improve the safety and efficiency in shared spaces.

However, how to form groups is not a trivial task in shared spaces. Firstly, the traffic signs are removed from the road surface, therefore, the location and number of groups should be decided, which is different from most area traffic management methods[]. Secondly, the area traffic regulation systems, like SCTOOT and GJ are not really online or dynamic. It learns regulations from historical data, and keep it as fixed periods. However, in shared spaces, the coming and leaving spots could be very different in different time periods, therefore, a dynamic algorithm is needed for road users here. Last but not least, even a single pedestrian group can gain or lose members on the fly due to interactions with other single road users or groups [13], the so-called splitting and merging phenomenon. The methods should take the coexistance of users trajectories into condideration.

State-of-the-art: (可以与literature review合并) 5-7个就足够了

Group behaviors have been studied in computer graphics [HLLO10], [CSM12], robotics[KLB12], pedestrian dynamics [GBS14], and social psychology [S.73]. Early techniques have been mainly used to simulate static or fixed-sized groups and perform group-based collision avoidance [SC14], [Lv13], [KG15]. Group initialization has also been addressed. [SKM17] simply used a threshold based on the distance between the team leader and members to group the pedestrians with similar OD. However, this approach is sensitive to the order of the input data because the algorithm is greedy - once the first possible solution is accepted, other solutions will never be reconsidered. [HPN + 16] clustered the original groups by the pairwise similarity metric defined over agents based on their starting positions and velocities. This works for the simulation application because the agents who are together at the beginning will keep coherent until the end of the experiment. However, the traffic scenario is more complicated. E.g. the road users who have the same origin and velocity at the beginning may split and reach different goals later. [HKBK14] considered the dynamic group behaviors via specifying the group shape as a queue and give deformation penalty, which is effective but cannot be generalized to other group shapes. However, none of these methods can efficiently simulate heterogeneous groups with dynamic behaviors in arbitrary environments.

Objectives

The aim of this paper is to find an dynamic clustering algorithm for the road users in shared spaces, considering their origins, destinations, and even the trajectory shapes.

Definitions

Here, a group is a formation of road users moving in a coordinated manner. A group can split, merge, avoid collisions while moving (\cite{mihaylova2014overview}). We concentrate on the following application scenario: Road users can appear from random locations around the shared space, then pass through, finally leave to their destinations. We are searching for a clustering algorithm to assign those road users to several groups according to their origins, destinations (OD data) and time.

# Methodology

In the following we present a definition suitable for a shared space. Assuming a stream of road users who come continuously and independently from all the directions of a shared space. For each road user, the coordinates of its origin (ox,oy), destination (dx,dy) and appearance time (t) are known. All road users have a threshold of tolerance time for waiting before crossing (w). The goal is to find an dynamic algorithm to cluster the road users who has similar oringin and destination before they are tired of waiting.

The problem above can be formulated as a series of facility location problem. In a basic formulation, the facility location problem is the following: giving a set of demand points and a set of candidate facility sites with costs of building facilities at each of them, the goal is to select a subset of sites where facilities should be built. Each demand point is then assigned to the closest facility, incurring a service cost equal to the distance to its assigned facility. The objective is to minimize the sum of facility costs and the sum of the service costs for the demand points \cite{charikar1999improved}. In our application, where the incoming road users need to form a group, the group center can be seen as the facility, and all road users are customers. However, this series of problems is NP-hard, so the best hope is to use a algorithm with provable approximation of the best solution.

Our algorithm is based on the framework proposed by [], which clusters dynamic and consistent points with a sliding window. The framework is suitable for points other than trajectories. Moveover, it continuously takes single point each time regardless of the restriction in time period. Therefore, the framework is modified to adaptive to our application (see Alorithm 1).

Therefore, our main algorithm works as follows: start with a solution given by Meyerson’s algorithm of cost Θ, use Lemma 2.3 to maintain a solution during Θ/4αf updates, and then recomputed from scratch. We call the intervals between consecutive recomputations periods, and note that they are random objects: the length of a period is determined by the cost of its initial solution, which is a random variable.

Assume the data is continuously come and ordered with coming moments. A window with size w is sliding from the start of data and each time insert a new point/delete the oldest point from the current data. Before the sliding window runs, w points are taken to calculate a course solution with repeated Meyerson algorithm, which is an approximate result of optimized solutions. Then, when each new point x is inserted, the transport cost from x to all current centers are calculated. If the minimum distance d is larger than the facility construction cost f, there has a possibility of d/f to build a new center, otherwise the new point will be assigned to the closet centers. The cost of each update is added in a round, until it reaches the threshold of update cost. Besides the cost threshold, to include the temporal info into consideration, we import a threshold called waiting time. In a cross, the road users has a tolerance waiting time, if they have to wait too long, they might go regards any rules. []. Therefore, there are two rules to control the update: cost update and waiting recomputation.

The pseudo code of current algorithms can be summarized as below:

Input: A list of road users, with their appearing time ot, appearing location (ox,oy), destination location (dx,dy), the sliding window size w, waiting time th\_waiting.

Output: A list of center locations Facils, and the facils update time mutation.

Let Facils = [] and take the first w points from data.

1: coarse solution Meyerson():

1. generate initial solution by Meyerson(list of x0 to xw);
2. repeat Meyerson 5 times
3. Lowerest cost theta and initial centers

2: while taken the last user:

1. Waiting <- x\_w+t[‘of’] – x\_w[‘of’]
2. If waiting > th\_waiting:
3. Delete x0:xw <- previous center have gone
4. Recompute solution for ne period by Meyerson()
5. If waiting <= th\_waiting:
6. If t – lasttime > cost threshold: # reach the recompuation cost criteria
7. Recompute solution by Meyerson()
8. Else
9. If OD\_similarity < f:
10. Assign xt to nearest center
11. Else
12. Create new center at location xt
13. take a coming user, calculate its distance d <- OD\_similarity to the nearest center vi
14. if d

Algorithm Meyerson()

Distance metric OD\_similarity()

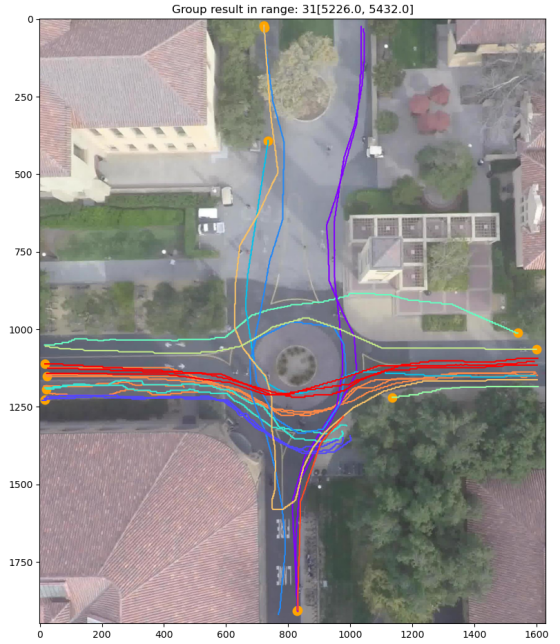
Return

data preprocessing

online facility location with OD data

explain the parameters

# Result



# Discussion

5.2 compare with state-of-art methods

# Conclusion

Datasets provided by \cite{pascucci2017discrete} or \cite{robicquet2016learning} will be used. The paper will describe the approach and the experiment in detail. The source code and data will be made publicly available.

\section{Outlook}

In the paper, the basic grouping strategy of road users in shared space will be presented. Future work will investigate more interaction effects, e.g.: the correlation between a variety of road users (e.g. pedestrians, cyclists, vehicles, etc.), the different possibilities to form groups, as well as forming groups which are allowed to split and merge.