J Vis (2025) 28:535–551
https://doi.org/10.1007/s12650-025-01055-8

REGULAR PAPER

Junxiu Tang · Huimin Ren · Zikun Deng · Di Weng · Tan Tang · Lingyun Yu · Jie Bao · Yu Zheng · Yingcai Wu

Towards better illegal chemical facility detection with hazardous chemicals transportation trajectories

Received: 17 October 2024 / Revised: 9 January 2025 / Accepted: 9 February 2025 / Published online: 2 March 2025

Abstract Unregistered illegal facilities that do not qualify for chemical production pose substantial threats to human lives and the environment. For human safety and environmental protection, the government needs to figure out the illegal facilities and shut them down. A new, convenient, and affordable approach to detect such facilities is to analyze the trajectories of hazardous chemicals transportation (HCT) trucks. The existing study leverages a machine learning model to predict how likely a place is illegal. However, such a model lacks interpretability and cannot provide actionable justifications required for decision-making. In this study, we collaborate with HCT experts and propose an interactive visual analytics approach to explore the suspicious stay points, analyze abnormal HCT truck behaviors, and figure out unregistered illegal chemical

J. Tang · Y. Wu
State Key Lab of CAD&CG, Zhejiang University, Hangzhou 310058, Zhejiang, China
E-mail: tangjunxiu@zju.edu.cn

Y. Wu
E-mail: ycwu@zju.edu.cn

H. Ren
Li Auto Inc, Beijing, Beijing 101318, China
E-mail: hren@wpi.edu

Z. Deng
School of Software Engineering, South China University of Technology, Guangzhou 510641, Guangdong, China
E-mail: zkdeng@scut.edu.cn

D. Weng (✉)
School of Software Technology, Zhejiang University, Ningbo 315103, Zhejiang, China
E-mail: dweng@zju.edu.cn

T. Tang
School of Art and Archeology, Zhejiang University, Hangzhou 315103, Zhejiang, China
E-mail: tangtan@zju.edu.cn

L. Yu
Department of Computing, Xi'an Jiaotong-Liverpool University, Suzhou 215123, Jiangsu, China
E-mail: Lingyun.Yu@xjtlu.edu.cn

J. Bao · Y. Zheng
JD iCity, JD Technology, Beijing, Beijing 100176, China
E-mail: baojie@jd.com

Y. Zheng
E-mail: msyuzheng@outlook.com

J. Bao · Y. Zheng
JD Intelligent Cities Research, Beijing, Beijing 100176, China

facilities. First, experts receive an initial result from the detection model for reference. Then, they are supported to check the detailed information of the suspicious places with three coordinated views. We apply a visualization that tightly encodes the geo-referred movement activities along the timeline to present the HCT truck behaviors, which can help experts finally verify their conclusions. We demonstrate the effectiveness of the system with two case studies on real-world data. We also received experts' positive feedback from an expert interview.

Keywords Stay point · Spatialtemporal analysis · Hazardous chemicals transportation · Visual analysis

1 Introduction

Chemicals are broadly used in many fields, including medicine, industry, daily necessities, etc. Proper use of chemical materials can benefit society, while illegal operations and transportation of chemicals, especially hazardous chemicals, may introduce safety risks and even endanger the environment and human lives (Xinhua 2019). Governments have devoted great efforts to finding illegal facilities without qualification and registration by supervising the production and transportation of chemicals (Xinhua 2020; Pipeline and Administration 2019; Planas et al. 2008).

One of the measures to supervise hazardous chemical transportation (HCT) trucks is to mandate the installation of GPS trackers on these trucks (Zhu et al. 2021). Based on the truck trajectories, an intuitive approach to finding illegal facilities is analyzing the stay points of these trucks because the trucks need to stop for a while when loading or unloading chemicals. Nonetheless, the prior works on stay point analysis mostly focus on the location recommendation (Zheng and Xie 2011; Yuan et al. 2011, 2013) or transportation prediction (Zhang et al. 2015; Wang et al. 2014). These studies cannot solve the illegal chemical facility detection tasks in the HCT domain. Recent research (Zhu et al. 2021) infers the possibility of illegal loading and unloading operations at suspicious stay points through a machine learning model. However, the model can only provide a classification result of a place. Reliable detection and decision-making require further human-in-the-loop analysis of the HCT trucks' behaviors and spatial-temporal contexts of the suspicious place.

When a stay point is detected, a comprehensive analysis workflow is required to fulfill experts' needs of inspecting the spatial-temporal context of the stay point. Proposing an effective approach for reliable detection and decision-making poses two challenges.

Finding suspicious facilities Traditionally, government staff needs to conduct in-person field inspections on the illegal facilities that are reported through stay point detection. Analyzing the stay points of HCT trucks considerably reduces efforts to find suspicious facilities. To determine whether a stay point may be illegal, local authorities can refer to a registered allowed list of chemical facilities to check if that stay point is on the list. Simply filtering places that are not registered can result in false positive identifications of illegality at the large volume of stay points. The allowed list and Points of Interest (POI) information on the available maps can be outdated. For example, trucks may stop at a newly opened parking lot that has not been registered on the allowed list, and such staying should be allowed. Due to these reasons, simply filtering is not reliable, and it is challenging to efficiently dismiss fake suspicious places. The map context around the stay points offers rich information for inferring the possible reasons behind the stop of trucks. An efficient approach is required for supporting analysts to quickly go over a sheer amount of stay points and determine whether a stay point is suspicious or not based on an intuitive summary of the surrounding environment and related spatiotemporal context.

Reasoning HCT trucks' behaviors The trucks' behaviors, including staying and moving activities, are important cues for discovering illegal factories. For example, the staying duration can be consulted to determine whether the truck is loading or unloading chemicals or just stopping due to traffic congestion. The trucks' movements also provide hints on whether the trucks take detours to suspicious places. Moreover, the temporal patterns of the trucks' stay points can reveal the regularity of transportation. Such behaviors demand a scalable and integrated spatiotemporal visualization to support effective decision-making in locating illegal factories and summarizing the patterns of unauthorized HCT activities.

To address the above challenges, we propose a visualization approach based on spatiotemporal stay point analysis. We collaborated closely with domain experts to compile the design requirements for HCT analysis and develop Helium, a novel visual analytic system that assists the experts in detecting illegal chemical facilities based on HCT trajectories. Helium comprises three coordinated views. In the spatial view, experts can access an overview of suspicious locations and investigate the surroundings of trajectories and stay points. In the temporal view, a series of visualizations show the POIs, staying duration distribution, and related trucks of the selected stay points. In the activity view, Helium presents the truck's transportation activities by tightly combining geo-referred movement activities and the timeline. We also report several cases based on real data, and the results were confirmed by experts.

The contributions of this work are as follows: (1) We summarize the domain requirements of detecting illegal chemical facilities by deeply collaborating with experts. (2) We design a visual analytics system named Helium that facilitates domain experts' decision-making for identifying illegal chemical facilities. The system embeds a tailored design for spatial-temporal trajectory visualization, which can reveal trucks' transportation patterns. (3) We demonstrate the effectiveness and efficiency of Helium through two case studies conducted on real-world datasets and an expert interview.

2 Related work

2.1 Urban data visualization

Urban data visualization has long been studied and applied in various urban applications (Zheng et al. 2016; Chen et al. 2015; Andrienko et al. 2017; Deng et al. 2023). Existing urban visualizations can be divided into spatial, temporal, and spatial-temporal based on which characteristics are displayed and analyzed.

Spatial visualization aims to provide a spatial context and semantics for urban data. Zheng et al. (2016) further categorized spatial visualization for urban data into point-, line-, and region-based. The point-based visualization depicts location [e.g., houses (Weng et al. 2018), stores (Weng et al. 2019), and station] with dot-like elements [e.g., circles (Deng et al. 2020; Jamonnak et al. 2022) and glyphs (Deng et al. 2022)] according to their geographic positions. Based on point-based visualizations, Kernel density estimation (KDE) is popularly used to obtain a continuous and smooth geographic distribution. The line-based visualization is suitable for plotting urban data with linear characteristics, such as trajectories and bus routes (Lorenzo et al. 2016; Weng et al. 2021). The region-based visualization is usually for space division, for example, choropleth maps (Huang et al. 2020). In our study, we leverage point-based visualizations for large-scale stay points and line-based visualizations for the trucks' behavior over urban space.

Temporal visualization enables time-oriented analyses of urban data. To do that, urban data is encoded along various types of timelines, such as radial (Liu et al. 2011) and linear (Deng et al. 2022), to reveal the temporal variation. Due to the importance of geographic context, temporal visualizations in urban analysis are oftentimes associated with spatial visualization to provide in-depth analysis in a spatiotemporal context. Such an association constitutes spatiotemporal visualization.

Spatiotemporal visualization demands the appropriate encodings of multi-dimensional heterogeneous temporal and spatial information. They can be divided into linked and integrated views (Sun et al. 2017). In *linked views*, spatial and temporal information are visualized in different views separately (Zanabria et al. 2021; Zhao et al. 2022; Yue et al. 2024; Chen et al. 2024; Zhang et al. 2023). The views are coordinated with flexible interactions to enable analysis. In this way, users can fully unfold the spatial or temporal patterns. The disadvantages of linked views are that users have to switch frequently between different views and thus suffer from the context switching costs (Baldonado et al. 2000). In *integrated view*, spatial and temporal visualizations are tightly combined in the same view such that users can analyze spatial and temporal patterns in the same context (Kraak 2003; Wu et al. 2019; Xie et al. 2021; Deng et al. 2024). **Trajectory visualization** is the most related to our study. Each trajectory is formed by a set of time-stamped locations connected in time order and thus has spatial-temporal information. Linked views are popular in revealing the temporal variation along trajectories. For example, Lu et al. (2015) coordinated a geographic map and a Sankey diagram to investigate trajectories' travel times along road segments. Ferreira et al. (2013) designed an interactive interface that enables users to query taxi trips and reason spatial and temporal patterns.

No existing integrated visualization can tightly combine time-dependent behavior and temporal variation in trajectories with geographic context. Such a combination is particularly intuitive and effective in analyzing the transport behavior of trucks. To fill the gap, we propose an integrated spatiotemporal visualization, which is a two-dimensional coordinate system, where geo-referred movement activities can be tightly encoded along the timeline.

2.2 Stay point analysis

Stay points in trajectories reflect the staying patterns of mobility. **Stay point detection** is commonly used in trajectory data preprocessing and mobility pattern mining (Zheng 2015; Jiang et al. 2013, 2017). Analyzing stay points is one of the most closely focused topics in urban computing (Zheng 2015) and visual analytics (Chen et al. 2015). Li et al. (2008) first introduced an algorithm to detect stay points in a trajectory. Then, several works (Yuan et al. 2015, 2011) utilize density clustering approaches to detect stay points more effectively. Subsequent studies adopt such detection methods for different applications in diverse scenarios.

Data mining methods on stay point analysis enable many applications to find hot spots and then properly adjust urban resources (Xiao et al. 2022; Yuan et al. 2013; Ruan et al. 2020). Moreover, stay point analysis can help model movement behaviors (Markovic et al. 2019; Damiani et al. 2014) and further infer the place semantics (Andrienko et al. 2016) or trip purpose (Li et al. 2022). Similar to finding patterns, revealing anomalies through stay point analysis also attracts researchers' attention (He et al. 2018; Hu et al. 2021; Zhu et al. 2021). For example, Hu et al. (2021) introduced SALON, a platform for universal stay point detection and presentation. They provided a case about abnormal location discovery by straightforward filtering the allowed list locations on stay points. Zhu et al. (2021) further explored the abnormal location detection problem. They proposed ICFinder to classify whether loading/unloading events occur, and conclude the illegality possibility of a suspicious location. However, these works are limited in explanation and not enough for in-depth human-in-the-loop analysis on stay point discovery of complex circumstances.

Visual analytics approaches, combined with data mining methods, empower flexible exploration on stay point analysis (Parent et al. 2013; Andrienko et al. 2007, 2016). Andrienko et al. (2007) proposed a framework that involves human perception in mobility pattern extraction. When extracting significant places, they visualized the staying time of individual places on maps and enabled users to make decisions with geographical context. Based on vehicle GPS trajectory data, many studies utilize visual analytics approaches for diverse purposes, such as trajectory query (Al-Dohuki et al. 2017; Ferreira et al. 2013), traffic jam analysis (Wang et al. 2013; Lee et al. 2020), and location placement decision-making (Liu et al. 2017; Weng et al. 2018). A series of interactive visual analytics systems (Gu et al. 2017; Buchmüller et al. 2017; Cappers 2017) were proposed in VAST challenge (Whiting et al. 2017) to support experts' visual exploration on finding vehicles' abnormal activities. In these works, the stay points are fixed sensor stations, so the place semantics and locations of stay points are known. This is different from our scenario because we need to figure out illegal facilities among all the stay points with unknown place semantics. In our work, a stay point in one trajectory is the location cluster, where a HCT truck stays in a small area for a period (see details in Sect. 3.1). Previous work (Andrienko et al. 2016) has proposed approaches to infer personal place semantics among stay points, but it is not suitable for the domain of illegal chemical facility detection, since the goals are different.

To enable experts' exploration in the detection pipeline, we propose a comprehensive visual analytics approach for reliable decision-making, which is equipped with an existing classification model (Zhu et al. 2021). We first abstract the domain problem as stay point analysis and then designed a novel visual analytics transportation system to assist experts in detecting illegal facilities. The system tightly integrates spatial and temporal analysis of stay points in transportation trajectories and enables experts to do reliable decision-making on suspicious places that are potentially illegal facilities.

3 Background

3.1 Background knowledge

Hazardous chemicals have huge potential risks to human lives and the environment (Wang et al. 2017). Thus, hazardous chemical transportation must be strictly supervised by the government. According to the government stipulation, specialized trucks that transport hazardous chemicals must be equipped with a positioning system supervised by the government, like GPS (Zhu et al. 2021; Liu et al. 2022). Through the logs of the positioning system, government supervisors can monitor HCT truck movements. These trajectory data provide a straightforward and intuitive demonstration of the hazardous chemicals supply chain. When a HCT truck loads or unloads chemicals, it stops somewhere for a while. Therefore, if a truck came to/from illegal facilities, the illegal positions should be among the stay points in the trajectories. A **stay point** is the area, where a truck stays in a small spatial range for a period. We use the centroid position of the area to represent one stay point in this work. Except for the places, where a truck loads or unloads chemicals, some of the stay points can be a parking spot, petrol station, or rest area on the highway. To clearly distinguish the identification of different positions in the problem, we use five terms to describe a place (Fig. 1):

Point of Interest (POI) the place on the map without geographical significance but other functions, such as restaurants, etc. Several stay points in HCT truck trajectories can belong to these positions.

Loading/Unloading (L/U) location the place, where one HCT truck loads or unloads chemicals. This may be a chemical production factory, processing plant, or place where chemicals are used, such as a school, scientific research institution, hospital, etc. Loading/Unloading locations can be either legal or illegal. HCT trucks usually take a U-turn in such places (Zhu et al. 2021).

Allowed list facility the place, where chemical production, processing, storage, or use is registered and approved. They are known POIs and also legal Loading/Unloading locations.

Illegal facility the place, where chemical production, processing, storage, or use is not registered or permitted. This is the illegal Loading/Unloading location. Opposite to the allowed list facility, these facilities are not qualified to handle chemicals, which, therefore, need to be checked and shut down.

Suspicious place the place is possibly an illegal facility but remains not verified. It is initially detected to be a Loading/Unloading location by the model (Zhu et al. 2021) and not near an allowed list facility. However, these two conditions are not enough to identify the place as an illegal facility, because the model result is a probability (Fig. 1). This kind of place is the target that experts need to analyze further. For instance, if the model indicates a place has suspicious L/U activities, it is still not clear whether the place is an illegal facility or just a newly opened parking lot that is not updated in the POI list.

Ideally, when experts find a suspicious place, they can check it through the satellite view and street view of a map so that they can judge the facility according to the images of the surroundings. However, these views cannot be updated in real time according to the policy (Google 2023), making it inaccurate to only use maps for illegal factory identification. Besides, some illegal facilities do not produce chemicals or own obvious devices, but they will hoard chemicals covertly, which is difficult to spot from the outside. Therefore, inferring whether a place has anomalous loading or unloading events is a better indicator of illegal activity.

3.2 Data

We preprocess the original HCT truck GPS data from Nantong, a city that has a developed chemical industry in China, for the following study and visualization. The data records 4243 trucks affiliated with Nantong in October 2020. Except for trajectory data, we also hold a POI list of Nantong for geographical range search. Moreover, we possess an allowed list of chemical facilities provided by the government, where the data includes facility names and locations.

We take apart the trajectory into segments at 0 a.m. each day, because for safety, HCT trucks should be off the road between 0 and 6 a.m. on the next day, according to the government stipulation. Each trajectory has a sequence of GPS points with longitudes, latitudes, and timestamps. We utilize the classic stay point detection method (Li et al. 2008) and density clustering method (Yuan et al. 2011) to extract stay points. The default distance and duration threshold is 100 ms and 10 min. Experts can adjust the methods and parameters on demand. Each stay point contains several original GPS points. With these original GPS points, we can access the (a) start time, (b) end time, (c) stay duration, and (d) turning angle of each stay point. We also calculate the driving duration between each two near stay points.

4 Method overview

In this section, we report the collaboration with domain experts, summarize the user requirements, and introduce the overview structure of the proposed visual analytics system.

4.1 Requirement analysis

We collaborate with two domain experts on supervising HCT. Both of them have rich experience in this domain. EA is an experienced product director who has been participating in the HCT supervision project for 3 years. EB is one leader of the project who has been participating in the project for more than 4 years.

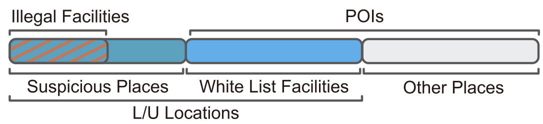

Fig. 1 The relationship of illegal facilities, POIs, suspicious places, allow list facilities, other places, and L/U locations

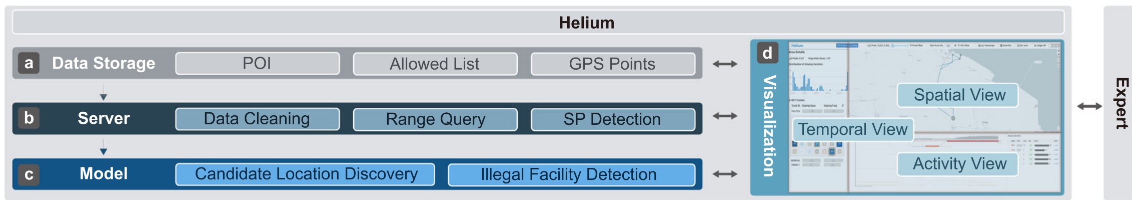

Fig. 2 The system architecture of Helium. Helium consists of four modules, namely, **a** data storage, **b** server, **c** model, and **d** visualization. Experts can visually analyze suspicious places in Helium through flexible interactions

The design study follows the 9-stage methodology proposed by Sedlmair et al. (2012). In *Winnow* and *Cast* stage, we presented the intuitive spatiotemporal visual analytics cases on trajectory data (Lu et al. 2015; Guo et al. 2011; Ferreira et al. 2013) to our collaborators, showing that the approach is more intuitive and flexible compared to only relying on model results. Then, in *Discover* stage, we learned the knowledge of HCT supervision and abstracted the problem as stay point analysis. To further explore the design requirements of the problem, we conducted regular meetings with the collaborated experts for 3 months. The experts provided typical cases in the past, showing how they found illegal facilities by filtering stay points, which are L/U locations classified by the model (Zhu et al. 2021), from the allowed list facilities.

They also introduced the difficulty in finding suspicious places among ambiguous stay point detection. We discussed iteratively and proposed a three-stage analysis workflow to achieve the illegal facility detection goals: in the first stage, experts need to locate a suspicious place for further detailed analysis; then in the second stage, they investigate the detailed information about the selected stay points and make the initial judgment; finally, they infer the illegal chain through the relevant HCT behaviors and spatial-temporal contexts. The workflow is not acyclic because experts can combine the discovery in each stage as a reference for the conclusion. We discussed iteratively and summarized the user requirements for visual analytics as follows:

R1: Obtain the geographical distribution of the potential illegal facilities. The model calculates the loading/unloading probability of each location cluster of multiple nearby stay points. Visualizing model results can provide experts with visual guides to the suspicious places that have possible illegality. Through effective visualizations, experts can first investigate those areas with high loading/unloading probability or dense suspicious places. The system should also provide the surrounding geographical context of the investigated places, like the POI information, which can help experts infer the POI category of such places. Moreover, experts expected a more flexible clustering method. The existing clustering in the model was decided by a DBSCAN algorithm with fixed distance and density parameters. Given that the distribution of chemical facilities is diverse in different city areas, experts need to first manually cluster stay points under some circumstances and then access the loading/unloading probability.

R2: Summarize temporal patterns of HCT trucks' behavior. After selecting one suspicious place, for further analysis, experts need to know the staying records of all the HCT trucks related to that place. They hope to obtain the summary of staying records per truck in an intuitive way because the same HCT truck usually has constant temporal staying patterns. From such patterns, experts can know when the suspicious place is often open for trucks and how long trucks usually stay. This is important to exclude normal POIs from suspicious places. For example, several HCT trucks stay at a place for a while during mealtime. Although the place seems suspicious according to the preliminary filtering, the historical temporal information suggests that it might be a newly opened restaurant, but the POI information is not updated.

R3: Abstract HCT trucks' spatial behaviors related to the suspicious place. Except for temporal patterns, experts also need to observe the trajectory detail of the HCT trucks that have been to the suspicious place. They require intuitive and effective visualizations for trajectories. Through such visualizations, experts should access the truck's behavior patterns and know the HCT events.

R4: Correlate locations and verify the transportation chain through historical trajectory data. With analysis being further, experts might regard a suspicious place to be an illegal facility. Under such circumstances, experts need to mine more clues with the known information and collect gist for field investigation. The HCT records of the trucks that have been to the regarded illegal facility require more checking. The system should enable experts to access all relevant milestone nodes in trajectories, including the origin, loading/unloading places, and destination.

4.2 System overview

Guided by the user requirements, we propose Helium, a visual analytics system that enables experts to analyze illegal chemical facilities. Helium consists of four modules: data storage, server, model, and visualization system (Fig. 2).

The data storage module stores all the data, including the POIs, the allowed list facilities, and the original GPS points of the HCT truck trajectories that are affiliated with Nantong. The server module includes data cleaning, geographical range query, and stay point detection. Experts can configure the processing methods and parameters on demand. The model module provides the loading/unloading probabilities of stay point clusters. Experts can adjust the input cluster of stay points to model and obtain the new inference result in real time. The visualization module contains three views, namely, spatial view (Fig. 6a1), temporal view (Fig. 6a2), and activity view (Fig. 6a3). Experts can interact with these coordinated views to explore suspicious places, analyze abnormal HCT truck behaviors, and figure out illegal chemical facilities.

The data storage module is conducted on a cloud server. Functions like geographical range query and stay point detection are partially supported by JUST (Li et al. 2020). We implement the model by Python and XGBoost (Chen and Guestrin 2016). The web-based visual analytics system is based on React.js, Mapbox GL JS, D3.js (Bostock et al. 2011) and G2.js.

5 Model

This section introduces the main idea of the loading/unloading detection model in ICFinder (Zhu et al. 2021), why we choose it, and how we leverage it in visual analytics.

The limited number of labels makes it impossible to train an effective classification model directly for identifying illegal hazardous chemical facilities. Since the HCT trucks have to stop when loading/unloading(L/U) hazardous chemicals, ICFinder (Zhu et al. 2021) proposed a detection method based on the stay points. In particular, the L/U locations that are not on the allowed list have a high potential to be illegal chemical facilities. There are two main parts in ICFinder, *Candidate Location Discovery* and *Illegal Facility Detection*.

Candidate Location Discovery generates candidate locations from HCT trajectories with three steps. First, *noise filtering* removes the outlier GPS points in the HCT trajectories, i.e., the points would be removed if the speed is higher than 130 km/h. Second, *stay point detection* extracts all stay points based on a spatial clustering algorithm (Li et al. 2008; Yuan et al. 2011). Finally, *location discovery* generates the candidate locations based on the DBSCAN clustering of the stay points (Ester et al. 1996).

Illegal Facility Detection detects illegal facilities with probability. It also comprises three steps. First, *location labeling* labels the candidate illegal locations based on the allowed list and domain knowledge. Second, *feature extraction* extracts behavior features of HCT trajectories. Both legal and illegal chemical locations have loading/unloading events, but many of their features are very different. Therefore, the extracted features should be shared at both legal and illegal chemical facilities, but at the same time, are very different from the Non-L/U locations. As a result, ICFinder extracts two features, including duration and turning angle, and a context feature, the embedding of delivering goods to detected L/U locations. Finally, *illegality ranking* employs XGBoost (Chen and Guestrin 2016) to train a classification model and then ranks the probabilities of candidate locations to detect illegal facilities. In summary, the model can output a L/U probability of one location. The higher the probability value, the more possible one place has loading/unloading events.

Combined with the known allow list, this classification model is more straightforward in identifying L/U activities at stay points than trajectory pattern mining models, like LEAD (Liu et al. 2022). Moreover, this model is quantitatively better than other related models and has been verified in real usage scenarios (Zhu et al. 2021). Thus, we use the model result and the allowed list as references for users to locate suspicious places, which is an initial input for the following visual analysis and exploration. Experts can adjust the model input through manual clustering, which depends on domain knowledge and specific geographical context, and then obtain a new result (**R1**).

6 Visual design

This section introduces the visual design of Helium. We first explain design details about the three coordinated views and then introduce the interactive analyzing workflow of Helium.

6.1 Spatial view

The spatial view provides a geographical overview of stay points, trajectories, and geographical contexts (Fig. 3).

The main component of this view is a map with several functional layers. The map contains two styles for experts to choose on demand: street style and satellite style (Fig. 3a, b). We present corresponding icons of different POI categories on the map. Experts can set the visibility of POIs in the layer configuration panel (Fig. 3a3) so that they can be aware of the surrounding environment (**R1**).

We visualize L/U probabilities of stay points in a heatmap (Fig. 3c). The darkness of color encodes the value of L/U probability. The redder one place is, the more probably L/U events occurred in the place. We overlay the heat map on the map so that experts can perceive the distribution of potential illegality and quickly locate those suspicious places from the heat map visual guide (**R1**). We also provide experts with a layer of original stay points, from which experts can manually check the clustering results of the model and adjust them if necessary. Given that L/U events can frequently occur near the allowed list facilities, we also show allowed list icons in the heatmap. Moreover, users can filter the places near the allowed list facilities from the L/U probability heat map.

By default, we present trajectories in the original GPS locations rather than map-matched locations (Fig. 3a2). The reason is that internal roads within the industrial parks cannot be well contained in the road network dataset, which results in the flawed map-matching near the stay points. Note that a "stay point" is not only a "point" but can be a small area within which the moving range of HCT trucks is smaller than the detection distance threshold (Zheng 2015). The illegal loading/unloading events are exactly hidden in such trajectories. Therefore, these trajectories near or within the stay point area are an essential reference for experts to judge the HCT trucks' behavior (e.g., turning details). Within trajectories, stay points are visualized as markers with different styles according to their types (Fig. 3a1).

6.2 Temporal view

This view presents temporal statistics about the staying events in the focused suspicious place. After experts select one suspicious place for further analysis, the temporal view will pop up from the left. This view has three parts to display temporal information (Fig. 4).

Time distribution A histogram visualizes the distribution of all the stay points' duration (**R2**) at the place (Fig. 4b). Experts can brush the distribution to filter for further analysis. For example, they can decide the filtered duration according to the specific cases, such as whether HCT trucks load or unload chemicals for a while or park for a long period. By default, all the stay points will be selected.

Trucks' statistics Helium groups filtered stay points by HCT truck IDs (**R2**). A lineup chart shows the rankings of HCT trucks' stay frequency in the selected suspicious place (Fig. 4c). Specifically, we provide two criteria for experts to rank: the quantities of stay points near the suspicious place and the days staying events occur at the suspicious place. With the ranking list, experts can select trucks for further checking.

HCT calendar After the selection of one truck, a calendar will be visible (Fig. 4d). The calendar presents the staying details by dates, as the trajectories are taken into segments at 0 a.m. (see Sect. 3.2). Considering the HCT truck's behavior contains multiple attributes, we choose the glyph-based visualization to present these data since such designs are effective (Liu et al. 2017; Weng et al. 2021). Each date unit is shown as a rectangle, with three visual element encoding (Fig. 4f).

First, the orange border visibility encodes whether the truck makes a U-turn at this place. According to the domain knowledge, when trucks load or unload chemicals in one place, most of them will take U-turns at

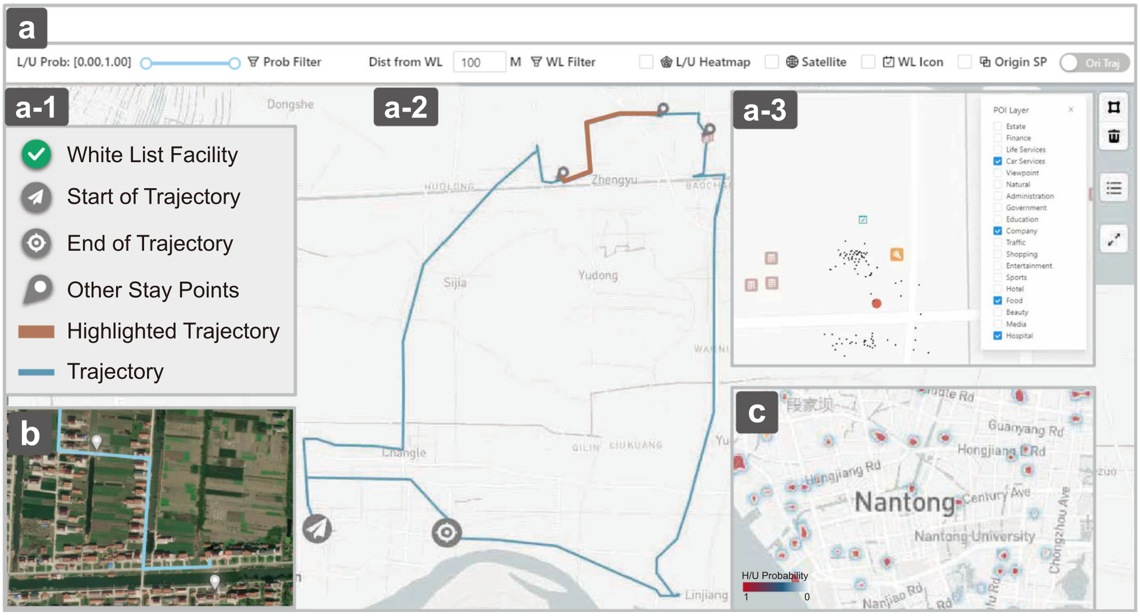

Fig. 3 The spatial view includes **a** a street map with layer configurations, **b** a satellite map, and **c** a L/U probability heatmap. In the street map, the system provides **a1** stay point markers, **a2** original GPS locations, and **a3** the POI configuration panel

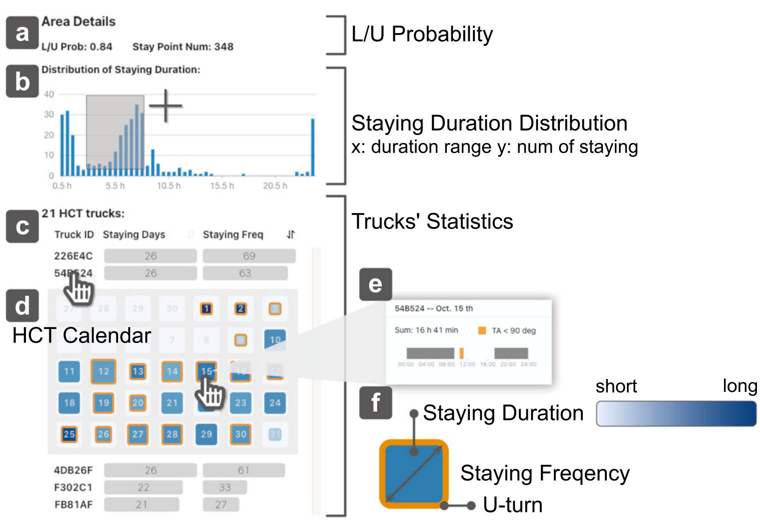

Fig. 4 The temporal view in Helium. **a** The L/U probability. **b** The distribution of stay points on staying duration. **c** Statistics about the related HCT trucks. **d** When users select one truck, a HCT calendar displays details about the truck's staying behaviors per day. **e** When users hover over one date, a card presents the staying events within that day. **f** The visual encoding of the calendar unit

the chemical facility. If the turning angle of trajectories in the suspicious place is acute, the rectangle will be bordered. If the truck stays more than once on 1 day, the rectangle will be bordered if one of the staying has an acute turning angle.

Second, the color darkness of the rectangle encodes the sum of the staying duration. The longer the duration is, the darker the rectangle is. Third, the rectangle size encodes the frequency of the truck staying at the suspicious place in 1 day. This design is motivated by the findings that some trucks will stay in one place more than once. Each unit is interactive. For example, the hover interaction will pop up a tooltip (Fig. 4e) about the staying temporal distribution chart in 1 day, where the colored blocks represent the staying time.

The orange color encodes whether there is a U-turn during the stay. Moreover, the click interaction can lead to the highlight trajectory presentation of the day in both the spatial view and the activity view.

6.3 Activity view

The activity view visualizes trucks' activities during hazardous chemical transportation (Fig. 5). It consists of two parts: the transportation activity chart (Fig. 5a) and the staying cluster panel (Fig. 5b).

The **transportation activity chart** tightly integrates the Euclidean distance from each stay point to the suspicious place and temporal interval events of trucks' moving and staying. Such spatial and temporal contexts can facilitate the understanding and reasoning of trucks' behaviors (**R2, R3**). Particularly, the tight connection of relative distance and absolute time information can intuitively demonstrate the moving and staying events, indicating whether the suspicious place is an origin, destination, or passing-by place.

Alternative designs To present such spatial-temporal contexts, we have considered several alternative designs that can be divided into linked views and integrated views.

Linked views Some studies combine temporal visualizations with maps to explore regular moving patterns (Ma et al. 2016; Andrienko and Andrienko 2008), but such linked views require users to switch frequently between multiple views, which increases the exploring costs (Baldonado et al. 2000).

Integrated views Another idea is to encode the moving and staying duration as charts or glyphs over the map. However, such overlaid visualizations suffer from overlapping at one place if there is more than one staying event. Moreover, the position of the moving duration presentation is hard to decide because, unlike staying duration, moving duration describes the time interval of a trajectory between two points rather than one single point. The Gantt chart (Jo et al. 2014; Aigner et al. 2005) or flow-like designs (Monroe et al. 2013) can clearly display the duration of moving and staying events but lack considerations on the spatial distance presentation. Storyline-like visualization (Tang et al. 2019) can present positions of different times, but such design more likely encodes categorical locations rather than the numerical distance vertical. Marey's graphs employed for bus or train schedules can address distance encoding (Palomo et al. 2016). In Marey's graph, the y-axis plots stations by distance, and the x-axis represents the timeline. The distance between the two stations is reflected by their vertical distance. We adopt an extended design of Marey's graph by considering that stay points can appear in the trajectory as stations but are more arbitrary, compared to sequential places.

Visual encodings Motivated by the domain requirements and the incompatibility of alternative designs, we design a transportation activity chart (Fig. 5a) to illustrate the spatial-temporal information of trajectories in a 2D space. First, *the horizontal axis* is a timeline of 1 day, where the start point represents 0:00, and the endpoint represents 24:00. Second, different from the distance encoding of Marey's graph, *the vertical axis* plots stay points according to their Euclidean distances to the suspicious place. Each stay point has two timestamps: start and end staying. Thus, a staying event in one stay point is visualized as a horizontal line segment. The start and end horizontal coordinates of the line segment are decided by the start and the end timestamps of staying. The slashes represent the truck's moving. To provide experts with an overview of the staying events in the suspicious place, we present the distribution of staying duration absolute time on the horizontal axis (Fig. 5a1).

The **staying cluster panel** presents the HCT milestones among trucks' trajectories. According to experts, one HCT truck usually has similar moving and staying patterns, such as frequent staying places, every day. Thus, we cluster all of the truck's stay points in geographical locations by the DBSCAN approach to summarize the HCT chain (**R4**). To access a proper number of stopovers, according to our internal test, we choose the minimum samples as 5 and the maximum distance between two samples as 100 ms in DBSCAN. We rank the places by their occurring frequency in trajectories and provide milestone labels, like origin, stopover, and destination, based on their occurring orders in trajectories (Fig. 5b). We also provide the query results of whether the place is near the allowed list facilities and the name of the nearest allowed list facility.

The transportation activity chart and the staying cluster panel share the same color encoding. In the transportation activity chart, the horizontal line segments representing the staying events are encoded as the cluster color or gray in non-clustered locations. The gray slashes represent the moving events. When experts select (hover or click) the specific segment in the activity view (Fig. 5a2), the system will highlight the relative trajectory segments in the activity view and spatial view (Fig. 5c), as well as the calendar unit in the temporal view. Besides, experts can locate the clustered place in the staying cluster panel (Fig. 5b1, c1).

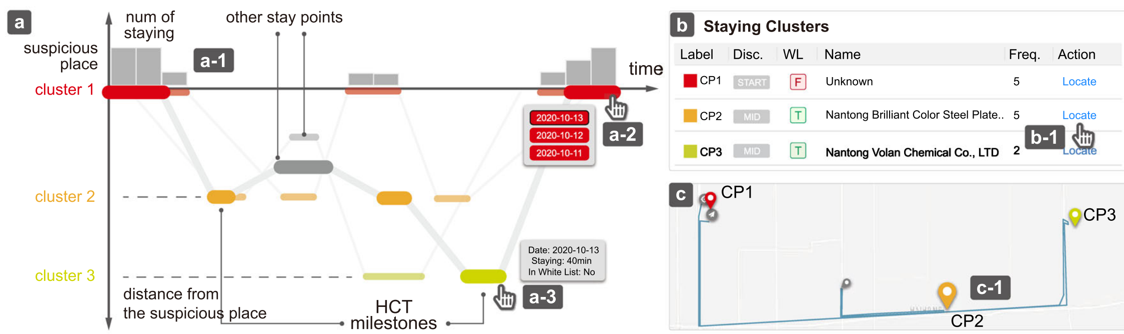

Fig. 5 **a** The transportation activity chart depicts the moving and staying activities of trucks. It encodes the spatial and temporal contexts of a trajectory in a 2D space. The polyline becomes bold when being hovered or selected. **b** The staying cluster panel lists the clustered places in the historical trajectories. **c** The activity view is coordinated with the spatial view

6.4 Interactive analysis workflow

With the three coordinated views, the whole analysis workflow follows the Visual Information Seeking Mantra: "overview first, zoom and filter, then details on demand" (Shneiderman 1996), introduced in detail below.

Locate the suspicious place First, the expert can access an overview of L/U probabilities from the heat map in the spatial view. Then, experts can zoom in on the map to see details about a target area, where all the individual stay points of this area are visualized in different clusters. If the calculated cluster area is incorrect, for example, stay points in different blocks are clustered together, experts can manually cluster stay points by drawing a polygon on the map. Moreover, experts can set the visibility of different layers (like POIs, stay points, etc.) in the spatial view on demand.

Obtain related HCT trucks' temporal patterns When experts choose one cluster, the temporal view will pop up. Experts can filter a specific range of staying duration in the bar chart. All the stay points in the cluster are selected by default. After that, Helium will show the filtered trucks in the HCT truck panel. By checking the HCT calendar of individual trucks, experts can know the number of stays, the turning angles, and the staying duration at the suspicious place. Then, they can investigate those suspicious records for further spatial-temporal analysis in the activity view.

Check HCT activities in spatial-temporal contexts Experts can interact with the activity view by clicking and hovering. When experts click one moving or staying event in the transportation activity chart, Helium will navigate the viewport of spatial view to the selected trajectory or stay point so that experts can check the shape of trajectories, the positions, and the surroundings (Fig. 5c). When they hover over one staying event, a tooltip will show details (Fig. 5a3). Moreover, experts can access the details about trucks' frequently staying places in the staying cluster panel and investigate these places (Fig. 5b1, c1).

7 Evaluation

To demonstrate the effectiveness and usability of Helium, we evaluated the system with two case studies and an expert interview. We run the system on a PC with a 1.80GHz Intel(R) Core(TM) i7-10510U CPU and a 1920*1080 screen. We use a real-world HCT truck dataset, which includes 4,243 trucks' 202,536 trajectories and 297,085 stay points [detected by Li et al. (2008), see Sect. 3.2] in 1 month.

7.1 Case study

We invited experts EA and EB in Sect. 4.1 to join the case study. First, we introduced the system including visual encodings, coordinated views, and user interactions to both experts. After ensuring that they were familiar with the system operation, we invited them to explore Helium freely. They were encouraged to speak out about their findings during the exploration. We recorded their findings in the study and summarized them as follows. We provide the demo videos for both case studies in the supplementary material.

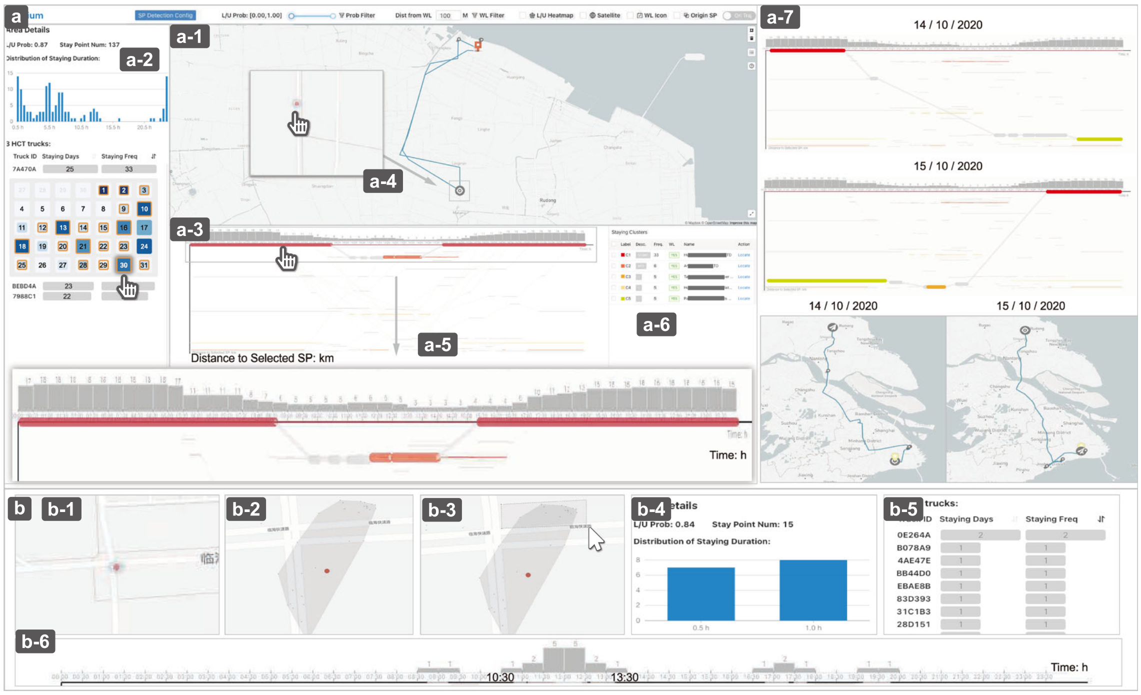

Fig. 6 Two case studies about **a** an in-depth analysis of suspicious places and **b** reconsidering model results. **a** In the first case study, EA investigated the HCT truck behaviors through the coordinated spatial and activity views and further judged the suspicious place as an illegal facility. **b** In the second case study, after exploring the system, EB laid down a fake positive place detected by the automatic model

7.1.1 In-depth analysis of suspicious places

EA first filtered the stay points by the allowed list in the spatial view (Fig. 6a1). Then, she located one red area in the remaining heatmap, which indicates the place had a high L/U probability (Fig. 6a4). *"This region is filtered by the allow list but has a high L/U probability, so it seems suspicious."* She opened the satellite layer to investigate the environment and surroundings of the suspicious place. However, there seem to be no apparent factory buildings. To learn more information about this place, she clicked the cluster of the stay points in this place. Then, Helium popped the temporal view on the left (Fig. 6a2). EA noticed that there were 137 stay points in this place, and these stay points had a distribution of diverse staying duration from very short (0.5 h) to very long (nearly 24 h). The L/U probability is 0.87. As shown in the truck's statistics panel, these stay points were mainly generated from 3 trucks. All of these trucks had more than 20 days recorded staying in the place and high staying frequencies. EA selected the first truck and investigated its transportation activities in the activity view (Fig. 6a3).

She found that the truck mainly stayed at this place during 0:00–6:00 and during 19:00–24:00. According to government law, during such time, HCT trucks were not allowed to be on the street. To check other staying activities, EA clicked one of the lines during 6:00–19:00 on the horizontal axis. The highlighted transportation activity polyline illustrated that the truck started its HCT driving from this place, and after visiting several allow list facilities (Fig. 6a5, a6), it returned to this place again. The other days also show the same pattern (Fig. 6a7). EA investigated the chemical-producing qualifications of these allow list facilities in an inner system affiliated with the supervising department and found that all these facilities produce resin or oil paint. She kept on checking other two trucks and got similar patterns. This pattern indicates all the related trucks transported similar chemicals to the suspicious place. Thus, EA inferred *"This place might be an unregistered storage warehouse for resin or oil paint. It requires a field inspection and further check."*

This case study shows that Helium can extend experts' workflow for in-depth analysis of suspicious places. After knowing the model result, the interactive system enables experts to grasp trucks' staying behaviors and transportation details through visual patterns. The human-in-the-loop analysis provides experts with more reliable and reasonable justifications for decision-making.

7.1.2 Reconsidering model results

EB was attracted by a red area on the heatmap (Fig. 6b1). After he filtered the allowed list, the red area remained in the spatial view. If EB followed the old filtering-based detection method, the detection was done, and the place needs to be reported as an urgent warning place for an in-person field survey. However, in Helium, EB could keep on deeper analysis of the place.

EB zoomed in on the map and found that the initial area clustered by the rule-based DBSCAN algorithm seemed not reasonable because some of the stay points in this cluster were on different sides of a crossroad. Given that illegal L/U activities cannot occur on different sides of the road, he investigated the distribution of stay points in the spatial view, and then he drew a polygon to manually make a reasonable cluster. With the temporal view popped up, EB found there were 11 trucks related to the place (Fig. 6b5). Among these trucks, he noticed that all of the trucks had a round 0.5–1 h duration (Fig. 6b4), and few staying frequencies (Fig. 6b5). Then, he selected several trucks and checked them in the activity view. The chart showed that these trucks usually stayed at the suspicious place for a while around 10:30 and 13:30, or during 16:30 to 18:00 (Fig. 6b6). After checking the above content, EB tended to put this place as a weaker warning for an in-person field survey, because these patterns indicated the staying events at this place might be on account of the meal time.

EB stated that he had gotten fake positive detection results from the model before. Some of those places turned out to be newly opened POIs (like restaurants, parking lots, or gas stations), or streets that sometimes had traffic jams. If he can know the related HCT trucks' trajectories to such places, he could leave these places as a lower priority for the in-person field survey. "*Although we still need to check these places in-person, for safety, with what the system reveals, we can leave more field survey opportunities and higher priorities to the places that are difficult to judge just based on the system, especially when we got lots of suspicious candidates.*" He admitted that field surveys required extra effort and working hours and they preferred to finish them as efficiently as possible, and the system could assist in multiple considerations for the model results.

This case study illustrates that Helium can assist experts in dismissing the possibly fake positive results from the classification model. The transportation activity chart can provide experts with embedded spatial-temporal information on HCT trucks' trajectories related to the suspicious place. With such information, experts can make more granular inferences about suspicious places.

7.2 Expert interview

After experts explored Helium, we interviewed them separately. We collected experts' feedback when they used the system during the case study and summarized their comments.

Effectiveness The experts both responded with positive feedback about Helium. They agreed with the system's effectiveness in finding illegal facilities because it can demonstrate the known cases and find new suspicious alternatives for further analysis. (1) EA was impressed by the visual analytics approach, "*With the help of this system, we can dig potential illegal facilities on the basis of a more explainable and reasonable reference.*" (2) EB mentioned that Helium had provided necessary controls or adjustments of automatic methods for domain experts so that they could fix some apparent drawbacks in time. For example, due to the limitation of the clustering methods, some stay points cannot be clustered correctly. Helium supports experts to manually cluster stay points based on their knowledge about the place or real-world scenarios; for example, the stay points related to one facility cannot be divided into a street or a river. (3) EA appreciated the whole workflow of Helium, "The visual analytics system provides a complete pipeline in finding and analyzing the illegal places through HCT truck trajectories." Although Helium cannot absolutely conclude one place's illegality, the coordinated views in the system provide experts with a more comprehensive decision-making context than a model inference result.

Intuitiveness Both experts appreciated the intuitiveness of the visualizations in the system. EA said, "*The visual encoding of date unit in the calendar can make me known the overview of daily staying events.*" EA also mentioned that the coordinated interactions in the spatial view and activity view could avoid the visual clutter and overlapping of the trajectory on the same road. Moreover, EB thought highly about the activity view, "*The chart in the activity view requires some learning efforts, but it is beneficial to inspect the behavior of the HCT truck in 1 day. With this chart, I can intuitively access the staying duration, timestamp of staying, and staying frequency of one trajectory.*" EA agreed that the activity view could clearly illustrate the HCT truck's transportation event sequence, "*... the staying clustering panel indicates one truck's frequent HCT milestones, helping me summarize the HCT chains.*"

Due to limited availability of domain experts for collaboration, we conducted case studies and expert interview with EA and EB. To further validate the effectiveness of Helium in real-world scenarios, we plan to conduct a long-term study in collaboration with both the experts after the system is deployed in real-world settings.

8 Discussion

Lessons learned The most important lesson we learned during the collaboration with domain experts is not to filter information by prior knowledge. As the second case illustrated, some suspicious places might be fake positives, which increases the field survey cost. If EB follows the traditional workflow, they will probably waste the efforts for that field survey. It is important for users to explore and check the original data flexibly when they access the result inferred by automatic methods. Integrated visualizations act as good assistance that can provide users the opportunity to investigate the suspicious place in detail, combined with known spatial-temporal context.

Generalizability First, Helium is a general visual analytics system without a tailored design for specific cases. Although we utilize the HCT data in Nantong during the design and development process, the proposed visualization approach for illegal chemical facility detection can also be applied in other cities. Second, the visual analytics approach for stay point analysis is general for other transportation visualization systems. Researchers can apply the transportation activity chart (Fig. 5) to visualize the moving and staying events in one trajectory, to summarize frequent staying behaviors based on the visual pattern of a horizontal line, and to check transportation events from both spatial and temporal perspective. For variable scenarios, the vertical axis can represent different meanings, like distance to the destination or a specific place. For example, to analyze whether a taxi takes a detour during the trip to the passengers' destination, the vertical axis of the transportation activity chart can represent the distance to the destination.

Limitation and future work First, Helium lacks data about trucks' affiliated companies and HCT orders. Besides, it does not combine the analysis of the nearby social media data from multiple platforms. These data can provide a more complete context to infer the related chemicals of one suspicious place. However, these data are either sensitive or go beyond of the research scope of this paper, so they are not considered for now. Second, segmenting trajectories requires more precise methods. We divide the trajectories by 0:00 a.m. (Sect. 3.2), but the trajectories can be divided at a more fine-grained level, because a truck could have many individual HCT drives in 1 day. Fine-grained trajectory segmentation by semantic meanings, for example, allow list facility or trucks' affiliated companies, is worth further study. Third, our evaluation mainly relies on qualitative data from case studies and expert interviews, as it was designed to capture expert insights on the system's real-world effectiveness. A more comprehensive quantitative assessment will be conducted when sufficient ground truth data becomes available. Moreover, future work can integrate more data if possible, like trucks' affiliation to companies and trucks' HCT orders and social media posts around the suspicious place, for more comprehensive illegal chemical facility detection and supervision. Fourth, the visual analysis system relies on the ICFinder model (Zhu et al. 2021), which is tailored for specific usage scenarios. When applying the approach to other cities, differences in data features may require retraining the model to ensure its effectiveness.

9 Conclusion

This work introduces Helium, a visual analytics system for better detection of illegal chemical facilities based on HCT truck trajectory analysis. Through Helium, experts can explore suspicious places, analyze abnormal HCT truck behaviors, and figure out illegal chemical facilities. The system consists of a data storage and processing server, a L/U inference model, and a web-based visualization interface. In Helium, we propose a novel design to visualize HCT truck trajectories, which can tightly combine the time-dependent behavior and temporal variation in trajectories with geographic contexts. We report two case studies and an expert interview to demonstrate the effectiveness of Helium and the intuitiveness of visual designs. This work provides the domain of HCT supervision with an intuitive visual exploration approach and a more reliable and reasonable decision-making workflow, compared to their traditional one.

Acknowledgements We thank all reviewers for their constructive comments. The work was supported by National Key R&D Program of China (2022YFE0137800), Key "Pioneer" R&D Projects of Zhejiang Province (2023C01120), NSFC (U22A2032, 62402421, 62402184, 62302440, 62272396), Guangdong Provincial Fund for Basic and Applied Basic Research - Regional Joint Fund Project (Key Project) (2023B1515120078) and Guangdong Provincial Natural Science Foundation for Outstanding Youth Team (2024B1515040010).

References

Xinhua: Cause of deadly chemical plant blast in east China revealed (2019). Website. http://en.people.cn/n3/2019/1116/c90000-9632825.html. Accessed 4 May 2024

Xinhua: China to implement all-around management on safe production of hazardous chemicals (2020). Website. http://english.www.gov.cn/statecouncil/ministries/202002/27/content_WS5e57b812c6d0c201c2cbd1a4.html. Accessed 4 May 2024

Pipeline Administration HMS (2019) How to comply with federal hazardous materials regulations. Website. https://www.fmcsa.dot.gov/regulations/hazardous-materials/how-comply-federal-hazardous-materials-regulations. Accessed 4 May 2024

Planas E, Pastor E, Presutto F, Tixier J (2008) Results of the MITRA project: monitoring and intervention for the transportation of dangerous goods. J Hazard Mater 152(2):516–526

Zhu Z, Ren H, Ruan S, Han B, Bao J, Li R, Li Y, Zheng Y (2021) ICFinder: a ubiquitous approach to detecting illegal hazardous chemical facilities with truck trajectories. In: Proceedings of ACM SIGSPATIAL, pp 37–40

Zheng Y, Xie X (2011) Learning travel recommendations from user-generated gps traces. ACM Trans Intell Syst Technol 2(1):1–29

Yuan J, Zheng Y, Zhang L, Xie X, Sun G (2011) Where to find my next passenger. In: Proceedings of international conference on ubiquitous computing, pp 109–118

Yuan NJ, Zheng Y, Zhang L, Xie X (2013) T-Finder: a recommender system for finding passengers and vacant taxis. IEEE Trans Knowl Data Eng 25(10):2390–2403

Zhang F, Yuan NJ, Wilkie D, Zheng Y, Xie X (2015) Sensing the pulse of urban refueling behavior: a perspective from taxi mobility. ACM Trans Intell Syst Technol 6(3):1–23

Wang Y, Zheng Y, Xue Y (2014) Travel time estimation of a path using sparse trajectories. In: Proceedings of ACM SIGKDD, pp 25–34

Zheng Y, Wu W, Chen Y, Qu H, Ni LM (2016) Visual analytics in urban computing: an overview. IEEE Trans Big Data 2(3):276–296

Chen W, Guo F, Wang F-Y (2015) A survey of traffic data visualization. IEEE Trans Intell Transp Syst 16(6):2970–2984

Andrienko GL, Andrienko NV, Chen W, Maciejewski R, Zhao Y (2017) Visual analytics of mobility and transportation: state of the art and further research directions. IEEE Trans Intell Transp Syst 18(8):2232–2249

Deng Z, Weng D, Liu S, Tian Y, Xu M, Wu Y (2023) A survey of urban visual analytics: advances and future directions. Comput Vis Media 9(1):3–39

Weng D, Zhu H, Bao J, Zheng Y, Wu Y (2018) HomeFinder revisited: finding ideal homes with reachability-centric multi-criteria decision making. In: Proceedings of the 2018 CHI conference on human factors in computing systems, pp 1–12

Weng D, Chen R, Deng Z, Wu F, Chen J, Wu Y (2019) SRVis: towards better spatial integration in ranking visualization. IEEE Trans Vis Comput Graph 25(1):459–469

Deng Z, Weng D, Chen J, Liu R, Wang Z, Bao J, Zheng Y, Wu Y (2020) AirVis: visual analytics of air pollution propagation. IEEE Trans Vis Comput Graph 26(1):800–810

Jamonnak S, Zhao Y, Huang X, Amiruzzaman M (2022) Geo-context aware study of vision-based autonomous driving models and spatial video data. IEEE Trans Vis Comput Graph 28(1):1019–1029

Deng Z, Weng D, Liang Y, Bao J, Zheng Y, Schreck T, Xu M, Wu Y (2022) Visual cascade analytics of large-scale spatiotemporal data. IEEE Trans Vis Comput Graph 28(6):2486–2499

Lorenzo GD, Sbodio ML, Calabrese F, Berlingerio M, Pinelli F, Nair R (2016) AllAboard: visual exploration of cellphone mobility data to optimise public transport. IEEE Trans Vis Comput Graph 22(2):1036–1050

Weng D, Zheng C, Deng Z, Ma M, Bao J, Zheng Y, Xu M, Wu Y (2021) Towards better bus networks: a visual analytics approach. IEEE Trans Vis Comput Graph 27(2):817–827

Huang Z, Lu Y, Mack EA, Chen W, Maciejewski R (2020) Exploring the sensitivity of choropleths under attribute uncertainty. IEEE Trans Vis Comput Graph 26(8):2576–2590

Liu H, Gao Y, Lu L, Liu S, Qu H, Ni LM (2011) Visual analysis of route diversity. In: 2011 IEEE conference on visual analytics science and technology (VAST), pp 171–180

Deng Z, Weng D, Xie X, Bao J, Zheng Y, Xu M, Chen W, Wu Y (2022) Compass: towards better causal analysis of urban time series. IEEE Trans Vis Comput Graph 28(1):1051–1061

Sun G, Liang R, Qu H, Wu Y (2017) Embedding spatio-temporal information into maps by route-zooming. IEEE Trans Vis Comput Graph 23(5):1506–1519

Zanabria GG, Silveira J, Poco J, Paiva A, Nery MB, Silva CT, Adorno S, Nonato LG (2021) CrimAnalyzer: understanding crime patterns in São Paulo. IEEE Trans Vis Comput Graph 27(4):2313–2328

Zhao W, Wang G, Wang Z, Liu L, Wei X, Wu Y (2022) A uncertainty visual analytics approach for bus travel time. Vis Inform 6(4):1–11

Yue X, Feng D, Sun D, Liu C, Qin H, Hu H (2024) Airpollutionviz: visual analytics for understanding the spatio-temporal evolution of air pollution. J Vis 27:215–233

Chen F, Yu Y, Ni L, Zhang Z, Lu Q (2024) Dstvis: toward better interactive visual analysis of drones' spatio-temporal data. J Vis 27:623–638

Zhang X, Pang X, Wen X, Wang F, Li C, Zhu M (2023) Triplan: an interactive visual analytics approach for better tourism route planning. J Vis 26:231–248
Baldonado MQW, Woodruff A, Kuchinsky A (2000) Guidelines for using multiple views in information visualization. In: Proceedings of the working conference on advanced visual interfaces, pp 110–119
Kraak M-J (2003) The space-time cube revisited from a geovisualization perspective. In: Proceedings of international cartographic conference, pp 1988–1996
Wu Y, Xie X, Wang J, Deng D, Liang H, Zhang H, Cheng S, Chen W (2019) ForVizor: visualizing spatio-temporal team formations in soccer. IEEE Trans Vis Comput Graph 25(1):65–75
Xie X, Wang J, Liang H, Deng D, Cheng S, Zhang H, Chen W, Wu Y (2021) PassVizor: toward better understanding of the dynamics of soccer passes. IEEE Trans Vis Comput Graph 27(2):1322–1331
Deng Z, Chen S, Schreck T, Deng D, Tang T, Xu M, Weng D, Wu Y (2024) Visualizing large-scale spatial time series with geochron. IEEE Trans Vis Comput Graph 30(1):1194–1204
Lu M, Wang Z, Yuan X (2015) TrajRank: Exploring travel behaviour on a route by trajectory ranking. In: Proceedings of IEEE Pacific visualization symposium, pp 311–318
Ferreira N, Poco J, Vo HT, Freire J, Silva CT (2013) Visual exploration of big spatio-temporal urban data: a study of New York city taxi trips. IEEE Trans Vis Comput Graph 19(12):2149–2158
Zheng Y (2015) Trajectory data mining: an overview. ACM Trans Intell Syst Technol 6(3):1–41
Jiang S, Fiore GA, Yang Y, Ferreira J, Frazzoli E, González MC (2013) A review of urban computing for mobile phone traces: current methods, challenges and opportunities. In: Proceedings of the 2nd ACM SIGKDD international workshop on urban computing
Jiang S, Ferreira J, Gonzalez MC (2017) Activity-based human mobility patterns inferred from mobile phone data: a case study of Singapore. IEEE Trans Big Data 3(2):208–219
Li Q, Zheng Y, Xie X, Chen Y, Liu W, Ma W-Y (2008) Mining user similarity based on location history. In: Proceedings of ACM SIGSPATIAL, p 34
Yuan NJ, Zheng Y, Xie X, Wang Y, Zheng K, Xiong H (2015) Discovering urban functional zones using latent activity trajectories. IEEE Trans Knowl Data Eng 27(3):712–725
Xiao Z, Fang H, Jiang H, Bai J, Havyarimana V, Chen H (2022) Understanding urban area attractiveness based on private car trajectory data using a deep learning approach. IEEE Trans Intell Transp Syst 23(8):12343–12352
Ruan S, Bao J, Liang Y, Li R, He T, Meng C, Li Y, Wu Y, Zheng Y (2020) Dynamic public resource allocation based on human mobility prediction. Proc ACM Interact Mob Wearable Ubiquitous Technol 4(1):1–22
Markovic N, Sekula P, Vander Laan Z, Andrienko G, Andrienko N (2019) Applications of trajectory data from the perspective of a road transportation agency: literature review and Maryland case study. IEEE Trans Intell Transp Syst 20(5):1858–1869
Damiani ML, Issa H, Cagnacci F (2014) Extracting stay regions with uncertain boundaries from gps trajectories: a case study in animal ecology. In: Proceedings of ACM SIGSPATIAL, pp 253–262
Andrienko N, Andrienko G, Fuchs G, Jankowski P (2016) Scalable and privacy-respectful interactive discovery of place semantics from human mobility traces. Inf Vis 15(2):117–153
Li Z, Xiong G, Wei Z, Zhang Y, Zheng M, Liu X, Tarkoma S, Huang M, Lv Y, Wu C (2022) Trip purposes mining from mobile signaling data. IEEE Trans Intell Transp Syst 23(8):13190–13202
He T, Bao J, Li R, Ruan S, Li Y, Tian C, Zheng Y (2018) Detecting vehicle illegal parking events using sharing bikes' trajectories. In: Proceedings of ACM SIGKDD, pp 340–349
Hu Y, Ruan S, Ni Y, He H, Bao J, Li R, Zheng Y (2021) SALON: a universal stay point-based location analysis platform. In: Proceedings of ACM SIGSPATIAL, pp 407–410
Parent C, Spaccapietra S, Renso C, Andrienko G, Andrienko N, Bogorny V, Damiani ML, Gkoulalas-Divanis A, Macedo J, Pelekis N, Theodoridis Y, Yan Z (2013) Semantic trajectories modeling and analysis. ACM Comput Surv 45(4):1–32
Andrienko GL, Andrienko NV, Wrobel S (2007) Visual analytics tools for analysis of movement data. ACM SIGKDD Explor Newsl 9(2):38–46
Al-Dohuki S, Wu Y, Kamw F, Yang J, Li X, Zhao Y, Ye X, Chen W, Ma C, Wang F (2017) SemanticTraj: a new approach to interacting with massive taxi trajectories. IEEE Trans Vis Comput Graph 23(1):11–20
Wang Z, Lu M, Yuan X, Zhang J, Van De Wetering H (2013) Visual traffic jam analysis based on trajectory data. IEEE Trans Vis Comput Graph 19(12):2159–2168
Lee C, Kim Y, Jin S, Kim D, Maciejewski R, Ebert D, Ko S (2020) A visual analytics system for exploring, monitoring, and forecasting road traffic congestion. IEEE Trans Vis Comput Graph 26(11):3133–3146
Liu D, Weng D, Li Y, Bao J, Zheng Y, Qu H, Wu Y (2017) SmartAdP: visual analytics of large-scale taxi trajectories for selecting billboard locations. IEEE Trans Vis Comput Graph 23(1):1–10
Gu Q, Yin H, Chen L, Li H, Liu C, Yue X, Qu H (2017) PreserVis, a visual analytic system for traffic and pollution patterns: multi-challenge award for compelling synthesis of information. In: 2017 IEEE conference on visual analytics science and technology (VAST), pp 183–184
Buchmüller J, Jentner W, Streeb D, Keim DA (2017) ODIX: a rapid hypotheses testing system for origin-destination data IEEE vast challenge award for excellence in spatio-temporal graph analytics. In: 2017 IEEE conference on visual analytics science and technology (VAST), pp 197–198
Cappers BCM (2017) Exploring lekagul sensor events using rules, aggregations, and selections. In: 2017 IEEE conference on visual analytics science and technology (VAST), pp 193–194
Whiting MA, Cook K, Jordan Crouser R, Fallon J, Grinstein G, Haack J, Henderson C, Liggett K, Staheli D, Strasburg J, Tagestad J, Varley C (2017) VAST challenge 2017: mystery at the wildlife preserve. In: 2017 IEEE conference on visual analytics science and technology (VAST), pp 173–178
Wang J, Chen C, Wu J, Xiong Z (2017) No longer sleeping with a bomb: a duet system for protecting urban safety from dangerous goods. In: Proceedings of ACM SIGKDD, pp 1673–1681

Liu S, Xu Z, Ren H, He T, Han B, Bao J, Zheng K, Zheng Y (2022) Detecting loaded trajectories for hazardous chemicals transportation. In: 2022 IEEE 38th international conference on data engineering (ICDE), pp 3294–3306

Google: Google-contributed street view imagery policy (2023). Website. https://www.google.com/streetview/policy/. Accessed 4 May 2024

Sedlmair M, Meyer M, Munzner T (2012) Design study methodology: reflections from the trenches and the stacks. IEEE Trans Vis Comput Graph 18(12):2431–2440

Guo H, Wang Z, Yu B, Zhao H, Yuan X (2011) TripVista: triple perspective visual trajectory analytics and its application on microscopic traffic data at a road intersection. In: Proceedings of IEEE Pacific visualization symposium, pp 163–170

Li R, He H, Wang R, Huang Y, Liu J, Ruan S, He T, Bao J, Zheng Y (2020) JUST: Jd urban spatio-temporal data engine. In: Proceedings of IEEE international conference on data engineering, pp 1558–1569

Chen T, Guestrin C (2016) XGBoost: a scalable tree boosting system. In: Proceedings of ACM SIGKDD, pp 785–794

Bostock M, Ogievetsky V, Heer J (2011) D^3 data-driven documents. IEEE Trans Vis Comput Graph 17(12):2301–2309

Ester M, Kriegel H-P, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of international conference on knowledge discovery and data mining, pp 226–231

Ma Y, Lin T, Cao Z, Li C, Wang F, Chen W (2016) Mobility viewer: an Eulerian approach for studying urban crowd flow. IEEE Trans Intell Transp Syst 17(9):2627–2636

Andrienko G, Andrienko N (2008) Spatio-temporal aggregation for visual analysis of movements. In: 2008 IEEE symposium on visual analytics science and technology, pp 51–58

Jo J, Huh J, Park J, Kim B, Seo J (2014) LiveGantt: interactively visualizing a large manufacturing schedule. IEEE Trans Vis Comput Graph 20(12):2329–2338

Aigner W, Miksch S, Thurnher B, Biffl S (2005) PlanningLines: novel glyphs for representing temporal uncertainties and their evaluation. In: Proceedings of international conference on information visualisation, pp 457–463

Monroe M, Lan R, Lee H, Plaisant C, Shneiderman B (2013) Temporal event sequence simplification. IEEE Trans Vis Comput Graph 19(12):2227–2236

Tang T, Rubab S, Lai J, Cui W, Yu L, Wu Y (2019) istoryline: effective convergence to hand-drawn storylines. IEEE Trans Vis Comput Graph 25(1):769–778

Palomo C, Guo Z, Silva CT, Freire J (2016) Visually exploring transportation schedules. IEEE Trans Vis Comput Graph 22(1):170–179

Shneiderman B (1996) The eyes have it: a task by data type taxonomy for information visualizations. In: Proceedings 1996 IEEE symposium on visual languages, pp 336–343