Research Article

Semantic Trajectory Frequent Pattern Mining Method with Fuzzy Stay Time Constraint

Aiguo Li, Jiahao Fu, Ruifang Gao, and Jie Yang

School of Computer Science and Technology, Xi’an University of Science and Technology, Xi’an 710054, China

Correspondence should be addressed to Aiguo Li; liag@xust.edu.cn and Jiahao Fu; 758633861@qq.com

Received 16 February 2022; Revised 13 June 2022; Accepted 9 July 2022; Published 31 July 2022

Academic Editor: Kashif Zia

Copyright © 2022 Aiguo Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the security system, transforming a large number of collected target trajectories into semantic trajectories with a less volume and high quality and mining their frequent patterns are helpful to analyze the target behavior patterns, identify hazard sources, and enhance the internal prevention, and control of the security system. Aiming at the limitation of semantic trace frequent pattern mining method defined by precise stay time in practical application scenarios, a fuzzy semantic trace frequent pattern mining method is proposed. Firstly, the membership function of fuzzy stay time is defined, so the stay time of the target at the stay point is fuzzified, and the fuzzy semantic trajectory is obtained. Then, a fuzzy semantic trajectory frequent pattern mining algorithm FST-FPM (fuzzy semantic trajectory frequent pattern mining) is proposed. The FST-FPM algorithm is experimentally verified on the Geolife public dataset and the self-collected RFID positioning dataset. The experimental results show that FST-FPM algorithm can mine frequent patterns of fuzzy semantic trajectories on Geolife dataset and RFID positioning dataset, and the running time is reduced by more than 10% compared with classical PrefixSpan algorithm, PrefixSpan-x algorithm, and LFFT2 algorithm.

1. Introduction

In the security system, how to improve the protection capability of key facilities has always been a research hotspot [1]. Studying the moving track and staying time of the target at the key points in the protection area in the security system plays an important role in mining the behavior characteristics of the target and evaluating its threat degree. The moving track of the target in the protection area can be obtained through RFID positioning technology [2, 3]. It is helpful to analyze the target behavior pattern by analyzing the target trajectory and mining the abnormal trajectory [4, 5]. Mining the frequent patterns of the moving track of the target is of a great significance for analyzing the target intention, identifying hazard sources, and identifying security vulnerabilities. The original target trajectory data obtained is transformed into semantic trajectory with a lesser volume and higher quality, which is conducive to the mining of target behavior patterns [6].

1.1. Related Work. As semantic trace frequent pattern mining has become a research hotspot, scholars have proposed various methods to mine frequent patterns in various scenarios. Cai et al. [7] constructed original tracks from photos with geographical location marks, enriched the original tracks through semantic annotation, and proposed a density-based similarity measurement method to mine semantic track frequent patterns, resulting in more detailed spatial frequent patterns. Lior et al. [8] proposed a new multi-objective mining algorithm, which uses a small amount of memory to quickly mine a given itemset by mining the support of multiple prespecified itemsets in the FP-Tree. Davashi et al. [9] proposed an efficient method for mining uncertain frequent patterns based on the upper bound. A new upper bound is used to constrain the expected support. By tightening the upper bound of the expected support and pruning the infrequent binomial sets and their supersets, the mining effect is improved. Han et al. [10] proposed a frequent pattern mining algorithm PrefixSpan based on projection database. By constructing projection database,
the algorithm recursively mines frequent subsets in projection database, only needs to scan the database once, and does not generate candidate itemsets, so it has higher mining efficiency. Many scholars have improved the PrefixSpan algorithm to mine frequent patterns more effectively. Tahereh et al. [11] proposed an improved PrefixSpan algorithm. When constructing the projection database, the algorithm adds a filter to each prefix to delete sub-patterns, and only outputs the largest frequent patterns. At the same time, it sets the minimum pattern length threshold to filter the short patterns in thealarm data, so that the algorithm has a high performance. Liu et al. [17] analyzed 627 articles on Pythagorean fuzzy sets (PFS) in Web of Science from 2013 to 2020, provided conclusive and comprehensive analysis on PFS.

1.2. Our Contribution. In the target behavior analysis scenario, the semantic trajectory frequent pattern mining method defined by precise stay time has great limitations. On the one hand, the behavior of the target in the same time range can represent the same pattern; on the other hand, considering the precise stay time will lead to too strong constraints and too few frequent patterns. To solve this problem, this paper proposes a fuzzy semantic trajectory frequent pattern mining method based on the moving trajectory and stay time of the target in the security environment. Firstly, based on the fuzzy set theory, the stay time is fuzzified, and the fuzzy semantic trajectory is obtained. Then, a fuzzy semantic trajectory frequent pattern mining algorithm FST-FPM is proposed. Finally, the effectiveness of FST-FPM algorithm is verified on real datasets.

1.3. Organization. The rest of this paper is organized as follows: the second chapter is the problem definition and formal description, which introduces the definition of fuzzy semantic trajectory frequent pattern mining. The third chapter is the frequent pattern method of fuzzy semantic trajectory, which introduces the definition method of fuzzy stay time membership function, the method of stay point extraction, and the method of fuzzy semantic trajectory construction. The fourth chapter describes the FST-FPM algorithm, and introduces the detailed steps and pseudo code of FST-FPM algorithm. The fifth chapter is the experimental results and analysis, which introduces the experimental results and analyzes the reasons. The sixth chapter is the conclusion, which summarizes the work and contributions of this paper.

2. Problem definition and Formal Description

After extracting the stay point from the original trajectory set traj, it is transformed into semantic trajectory ST, and the processing of fuzzy stay time is introduced to obtain the semantic trajectory with fuzzy stay time constraint FST. Then, the FST-FPM algorithm proposed in this paper is used for frequent pattern mining. The specific method is to fuzzy the stay time of personnel in a specific area into several levels through the fuzzy set theory, calculate the membership s and support P of fuzzy stay time, and then, give a minimum support threshold σ and fuzzy stay time membership threshold ρ to find out all the frequent patterns of minimum support P > σ and fuzzy stay time membership s > ρ in the FST. The semantic trajectory frequent pattern with fuzzy stay time constraint FTP can be obtained.

In this paper, the related concepts used in mining semantic trajectory frequent patterns with fuzzy stay time constraints are defined as follows:

Definition 1. the target track is marked as traj, and its form is as follows

\[ \text{traj} = \{ \text{oid}, [\text{Lat}, \text{Lon}, \text{Data}, \text{Time}] \} \]

where, oid is the person Id, Lat is the decimal latitude, Lon is the decimal longitude, Data is the date, and Time is the time when the person arrives at the point.

Definition 2. if the stay time of the target in a certain area is greater than or equal to the threshold ω, the area is called a
stay point and is recorded as $R, R = (\text{loc}, t_e, t_s)$. Where, loc represents the geographical area number, $t_s$ represents the time of entering the geographical area, and $t_e$ represents the time of leaving the geographical area.

**Definition 3.** the time difference between the target entering and leaving a certain area is called the stay time in the area, which is recorded as $\Delta t, \Delta t = t_e - t_s$.

**Definition 4.** a sequence consisting of $n$ stay points is called the target semantic trajectory, which is recorded as ST, and its data form is

$$\text{ST} = \{\text{oid}, [R_1, R_2, \ldots, R_n]\},$$  \hspace{1cm} (2)

where, oid represents the personnel ID, $R_i$ represents the $i$th stay point, $i = 1, 2, \ldots, n$.

**Definition 5.** given the domain of stay time $U$, let $A$ be a fuzzy set on $U$. For any stay time $\Delta t \in U$, a number $\mu_A(\Delta t) \in [0, 1]$ can be determined to indicate the degree to which $\Delta t$ belongs to $A$, $\mu_A: U \longrightarrow [0, 1], \Delta t \rightarrow \mu_A(\Delta t)$, where $\mu_A$ is the membership function of $A$, and $\mu_A(\Delta t)$ is the membership degree of element $\Delta t$ in $U$ to fuzzy set $A$.

**Definition 6.** let the stay time fuzzy set be $\{A_1, A_2, \ldots, A_p\}$, and the membership degree of stay time $\Delta t$ on fuzzy set $A_j$ is composed of a $p$-ary set, denoted as $S(\Delta t)$, whose form is

$$S(\Delta t) = \left\{\mu_{A_1}(\Delta t), \mu_{A_2}(\Delta t), \ldots, \mu_{A_p}(\Delta t)\right\},$$  \hspace{1cm} (3)

where, $\mu_{A_1}(\Delta t)$ represents the membership degree that $\Delta t$ belongs to the fuzzy set $A_1$, $j = 1, 2, \ldots, p$.

The stay time of the target is blurred into $p$ levels, and the fuzzy set $A_j$ represents the $j$th fuzzy stay time level. The closer the $\mu_{A_j}(\Delta t)$ is to 0, the lower the $\Delta t$ belongs to $A_j$; the closer the $\mu_{A_j}(\Delta t)$ is to 1, the higher the $\Delta t$ belongs to $A_j$.

$$0 \leq \mu_{A_j}(\Delta t) \leq 1, \sum \mu_{A_j}(\Delta t) = 1.$$

**Definition 7.** after fuzzifying the stay time $\Delta t$ in the target semantic trajectory ST, the new semantic trajectory is called fuzzy semantic trajectory. Fuzzy stay time membership set $S(\Delta t)$, denoted as $FST$, in the form of

$$\text{FST} = \{\text{oid}, [(S_1(\Delta t_1), R_1), \ldots, (S_n(\Delta t_n), R_n)]\},$$  \hspace{1cm} (4)

where, oid represents the target number and $S_i(\Delta t_i)$ represents the fuzzy stay time membership set of the $i$th stay point $R_i, i = 1, 2, \ldots, n$. In equation (4), let $K = |\text{FST}|$ be called the $K$-term fuzzy semantic locus.

**Definition 8.** given a fuzzy set $A_j \in U, \mu_{A_j}(\Delta t_i) \in S(\Delta t_i)$. Given a minimum support threshold $\sigma$ and a fuzzy stay time membership threshold $\rho$, if the fuzzy semantic trajectory FST has support $P \geq \sigma$ and fuzzy stay time membership $\mu_{A_j}(\Delta t_i) \geq \rho$, then, this fuzzy semantic trajectory FST is called $K$-term fuzzy semantic trajectory frequent pattern and is recorded as FTFPK = $|\text{FST}|$ if $i = 1, 2, \ldots, n$.

According to the above definition, the semantic trajectory frequent pattern mining under fuzzy stay time constraint is the process of adding fuzzy stay time membership to the semantic trajectory to make it become the semantic trajectory of fuzzy stay time constraint, and then, mining the semantic trajectory frequent pattern of fuzzy stay time constraint.

### 3. Fuzzy Semantic Trace Frequent Pattern Mining Method

According to the definition of fuzzy semantic locus, a method for mining frequent patterns from fuzzy semantic locus is proposed. Firstly, the target semantic trajectory ST is obtained by extracting the stay point of the target trajectory; Then, the stay time is fuzzified, and the fuzzy membership degree set $S(\Delta t)$ and support $P$ degree of the target semantic trajectory are calculated. The fuzzy semantic trajectory $FST$ is obtained by adding the fuzzy stay time membership set $S(\Delta t)$ to the target semantic trajectory. Finally, the proposed fuzzy semantic trajectory frequent pattern mining algorithm is used to mine all fuzzy semantic trajectory frequent patterns FTFP from the fuzzy semantic trajectory dataset FST.

#### 3.1. Calculation Method of Membership Degree of Fuzzy Stay Time

In the scenario of personnel behavior pattern analysis, considering the accurate stay time of personnel, it is impossible to mine effective frequent patterns. Therefore, the stay time of personnel is fuzzified by fuzzy set theory, and the stay time can be fuzzified into $p$ levels according to expert experience. In order to get the membership degree of fuzzy stay time defined in Definition 6, the membership functions of fuzzy stay time at each level need to be defined, respectively. In this paper, the membership function of fuzzy stay time is divided into three level: small, medium, and large.

The membership function of small fuzzy stay time is defined as:

$$f(\Delta t) = \begin{cases} 1, 0 < \Delta t < \Delta t_1 \\ \Delta t_1 - \Delta t, \Delta t_1 < \Delta t \leq \Delta t_2 \\ 0, \Delta t > \Delta t_2 \end{cases},$$  \hspace{1cm} (5)

The membership function of medium fuzzy stay time is defined as:

$$f(\Delta t) = \begin{cases} 0, \Delta t < \Delta t_1 \\ \Delta t - \Delta t_1, \Delta t_1 \leq \Delta t \leq \Delta t_2 \\ \Delta t_2 - \Delta t_1, \Delta t_2 < \Delta t \leq \Delta t_3 \\ \Delta t_3 - \Delta t_2, \Delta t_3 < \Delta t \leq \Delta t_4 \\ 0, \Delta t > \Delta t_4 \end{cases}. $$  \hspace{1cm} (6)
The membership function of large fuzzy stay time is defined as:

\[
    f(\Delta t) = \begin{cases} 
    0, & \Delta t < \Delta t_1 \\
    \frac{\Delta t - \Delta t_1}{\Delta t_2}, & \Delta t_1 \leq \Delta t \leq \Delta t_2 \\
    1, & \Delta t > \Delta t_2 
    \end{cases}
\]

where, \( \Delta t_1 \) satisfies \( 0 < \Delta t_1 < \Delta t_2 < \Delta t_3 < \Delta t_4, i \in [1, 4] \).

3.2. Stay Point Extraction Method. Extracting the stay point from the target trajectory, and transforming the target trajectory into the target semantic trajectory with stay time, can reduce the data volume, improve the data quality, and facilitate the mining of frequent patterns. This section extracts the stay point by the stay time interval of the target track. The pseudo code of stay point extraction algorithm is shown in algorithm 1.

The input of algorithm 1 is the target trajectory, the coordinate value range of each region and the stay time threshold, and the output is all the extracted stay points. The line 1 to the line 14 judges whether each sampling point is a stay point. The line 4 records the start time of starting a sampling point. The line 5 judges the region of the sampling point according to the coordinate value. The line 6 determines whether the region of the next sampling point has changed. The line 7 records the end time of the sampling point. The line 8 determines whether the end time minus the start time of the sampling point meets the threshold value of the stay point. If so, the output is a stay point. The line 15 returns all the stay points, where, traj, represents a track of the target U. C represents a sampling point in a track, and loc represents the number of the divided area. The stay point is detected by judging whether the coordinate value of the sampling point enters or leaves the coordinate value range of each area and whether the stay time interval meets the stay time threshold \( \omega, v = 0, 1, \ldots, n \).

3.3. Fuzzy Semantic Trajectory Construction Method. This section illustrates the fuzzy semantic trajectory construction method. The form of the target track collected in the security system is shown in Definition 1. The information of a sampling point includes the target number, the coordinate value of the sampling point, and the date and time of the sampling point, such as

\[
    \text{traj} = \{01, [3, 7, 20210401, 12: 53: 12], [3, 8, 20210401, 13: 13: 12], [4, 8, 20210401, 14: 05: 08] \}
\]

where, 01 represents the target number, 3 represents the abscissa value, 7 represents the ordinate value, 20210401 represents the date, and 12: 53: 12 represents the time. This target track contains three sampling points, which are 12: 53: 12 at [3, 7], 13:13:12 at [3, 8], and 14: 05: 08 at [4, 8].

The stay time threshold \( \omega \) of algorithm 1 is set to 5 minutes. After the target trajectory is extracted according to algorithm 1, it can be transformed into the target semantic trajectory shown in Definition 4. It includes target number, stay point number, and time of entering and leaving the stay point, such as

\[
    \text{ST} = \{01, [(3, 12, 53: 12, 13: 12), (4, 14: 05: 08, \infty)] \}
\]

where, 01 refers to the target number, <3, 12, 53:12, 13:12> refers to a stop point, meaning 12: 53: 12–13: 12 stop at stop point 3. This target semantic track contains two stops, i.e., 12: 53: 12–13: 12, stop at stop 3, and stop at stop 4 from 14: 05: 08.

Given a short stay time fuzzy set, according to formula (5), the parameter is set to \( \Delta t_1 = 10, \Delta t_2 = 40 \), and the unit is minutes. Given a fuzzy set with a long stay time, according to formula (7) the parameter is set to, and the unit is minutes. Thus, the fuzzy stay time membership set of the target semantic trajectory is calculated. Add fuzzy stay time membership to the target semantic trajectory to obtain the fuzzy semantic trajectory as shown in Definition 7, such as

\[
    \text{FST} = \{01, [(0.67, 0.33), 3, 12: 53: 12, 13: 12, 14: 05: 08, \infty)] \}
\]

where, 01 represents the target number, <(0.67, 0.33), 3, 12: 53: 12, 13: 12> represents the first stay point, 0.67 indicates that the stay time belongs to the membership of the fuzzy set with short stay time, and 0.33 indicates that the stay time belongs to the membership of the fuzzy set with long stay time. 3 indicates the stop number, 12: 53: 12 is the entry time, and 13: 12 is the departure time.

4. FST-FPM Algorithm Description

This section details the steps of FST-FPM (fuzzy semantic trajectory frequency pattern mining) algorithm. The main idea of FST-FPM algorithm is that after scanning the semantic trajectory database once, according to the given minimum support threshold and fuzzy stay time membership threshold, all frequent itemsets with length of 1 are obtained, and then, each frequent sequence with length of 1 is used as a prefix to build its projection database respectively. After scanning the projection database, all frequent itemsets are obtained, and the prefix of the projection database is added, that is, the frequent sequence with length of 2. Finally, the projection database is recursively generated, and its frequent subsets are continuously mined until a new projection database cannot be constructed. The characteristics of FST-FPM algorithm are as follows: only when the support of the itemset is greater than the given minimum support threshold and the fuzzy stay time membership of the itemset is greater than the given threshold, the support of the itemset will be increased by 1, and the itemsets that do not meet the minimum support threshold and the fuzzy stay time membership threshold will be deleted to reduce the construction of the projection database. Algorithm 2 gives the complete FST-FPM pseudo code.
Algorithm 2 inputs semantic trajectory database \( FD \), minimum support threshold \( \sigma \), fuzzy stay time membership threshold \( \rho \), and fuzzy stay time membership function, and outputs all the frequent sequences that meet the conditions. FTGFP.

Algorithm 1: Stay point extraction.

\[
\begin{aligned}
\text{i)} & \quad \text{Input: Original track } traj, \text{ stay time threshold } \omega, \\
\text{ii)} & \quad \text{longitude and latitude range of each area} \\
\text{iii) } & \quad \text{Output: Stay point } R \\
& \quad \text{for all user } U_i \text{ in original trajectory sequence } traj \\
& \quad \text{for all trajectory } in \ U_j \\
& \quad \text{for all sampling point } C_i \text{ in } traj_k \\
& \quad \text{Record start time } t_{si} \\
& \quad \text{Judge the area where } C_i \text{ is located } loc_i \\
& \quad \text{if } loc_i \neq loc_{i-1} \\
& \quad \text{Record end time } t_{ei} \\
& \quad \text{if } t_{ei} - t_{si} > \omega \\
& \quad \text{Output stay point } R \\
& \quad \text{end if} \\
& \quad \text{end if} \\
& \quad \text{end for} \\
& \quad \text{end for} \\
& \quad \text{end for} \\
& \quad \text{return } R \\
\end{aligned}
\]

Algorithm 2: FST-FPM algorithm.

\[
\begin{aligned}
\text{i) } & \quad \text{Input: fuzzy semantic trajectory database } FD, \text{ minimum support threshold } \sigma, \\
\text{ii) } & \quad \text{fuzzy stay time membership threshold } \rho, \text{ fuzzy stay time membership function} \\
\text{iii) } & \quad \text{Output: fuzzy semantic locus frequent pattern set } \text{FTFP} \\
& \quad \text{initialize variable } i = 1; \ w, k = 1, 2, \ldots n \\
& \quad \text{scan fuzzy semantic trajectory database } FD \\
& \quad \text{calculate the support } P \text{ of each sequence} \\
& \quad \text{calculate the fuzzy stay time membership } S(\Delta t) \text{ of each sequence accord to the time membership function} \\
& \quad \text{if } P < \sigma \text{ and } \mu_A(\Delta t) < \rho \text{ the sequence is less than the minimum support threshold and fuzzy stay time membership threshold} \\
& \quad \text{define the sequence as an infrequent itemset and delete it} \\
& \quad \text{create a new sequence database } Q_i \text{ and add frequent itemsets with length } i \text{ to the database} \\
& \quad \text{for each frequent itemset } T \text{ in the sequence database } Q_i \\
& \quad \text{construct the projection database } C_w \text{ of frequent itemset } T \\
& \quad \text{get the frequent itemset } T_k \text{ of } C_w \\
& \quad \text{T and } T_k \text{ are constructed as frequent sequences } T_i \text{ with length } i \\
& \quad \text{FTFP } = \text{FTFP } \cup \ T_i \\
& \quad \text{if projection database } C_w \text{ is not empty} \\
& \quad \text{repeat steps 8–12, } i = i + 1 \\
& \quad \text{else output frequent sequence } \text{FTFP} \\
& \quad \text{end for} \\
& \quad \text{end if} \\
& \quad \text{return } \text{FTFP} \\
\end{aligned}
\]

Algorithm 2 inputs semantic trajectory database \( FD \), minimum support threshold \( \sigma \), fuzzy stay time membership threshold \( \rho \), and fuzzy stay time membership function, and outputs all the frequent sequences that meet the conditions. FTGFP. The line 1 initializes each variable. The line 2 scans the fuzzy semantic track database \( FD \). The line 3 calculates the support \( P \) of each semantic track. Lines 5 to 7 delete semantic tracks that do not meet the minimum support threshold and fuzzy stay time membership threshold. Lines 8 to 16 recursively generate projection database, and continuously mine frequent subsets until the projection database cannot be constructed. The line 19 returns all the fuzzy semantic trace frequent patterns FTGFP that have been mined, where, \( Q_i \) is used to store frequent sequences, \( i \) represents the number of iterations, \( i = 0, 1, \ldots, n \). \( C_w \) represents the projection database, stores the projection database of each frequent itemset \( T \), \( w = 1, 2, \ldots, n \). \( T_k \) represents the frequent itemsets mined each time, \( k = 1, 2, \ldots, n \). \( T_i \) represents a frequent sequence of frequent itemsets \( T_k \) combined with their prefixes.
5. Analysis of Experimental Results

5.1. Introduction to Data Set and Experimental Environment. The experiments were conducted using the public dataset [18] (hereinafter referred to as the Geolife dataset) in the Geolife project of Microsoft Research Asia and the real target trajectory dataset (hereinafter referred to as the RFID positioning dataset) collected from the RFID positioning subsystem of an integrated security system.

The Geolife dataset includes trajectory data of 182 people over a five-year period. In the experiment, 50 people, 2611200 sampling points, and 1518 tracks were selected. According to the track distribution, the area was divided into 6 areas, numbered 1–6. After denoising the dataset, 2293 stay points were obtained through algorithm 1.

The RFID positioning dataset includes 10 targets, 5488 sampling points, and 916 semantic tracks. After denoising the dataset, 2744 stay points were obtained by extracting stay points through algorithm 1. The positioning area is divided into 6 areas, numbered 1–6. The semantic trace form is defined in Definition 4.

An overview of the collated dataset is shown in Table 1. The semantic track data form is shown in Table 2. List different stay points in a track of a target by line. For example, the first three lines of Table 2 record the first track of target 1, indicating that target 1 stays in area 1 from 09:00 to 11:30, area 4 from 13:00 to 18:00, and area 5 from 19:00 to 19:40.

This experiment was conducted on a computer configured with core i5-6500 (3.20 GHz dual core) and 8 GB memory. The system was win10 and the program is written in Python (3.8.0) + Pycharm.

5.2. Data Preprocessing. After sorting the original track traj into semantic track ST, the membership degree of fuzzy stay time is added to it. Firstly, the stay time of each stay point is calculated according to Definition 3, and then, the fuzzy stay time membership function defined in this paper is used to calculate the fuzzy stay time membership degree of each stay point. According to the application scenario, the stay time is divided into five levels: very short, short, medium, long, and very long. The very short stay time is the membership function of small fuzzy stay time, the short, medium, and long stay time are the membership function of medium fuzzy stay time, and the very long stay time is the membership function of large fuzzy stay time. The parameter values of fuzzy stay time membership function under each stay time are as follows:

1. In the case of very short stay time, the value of the membership function parameter of fuzzy stay time is: \( \Delta t_1 = 10, \Delta t_2 = 20 \).

2. In the case of short stay time, the value of the membership function parameter of fuzzy stay time is: \( \Delta t_1 = 10, \Delta t_2 = 15, \Delta t_3 = 25, \Delta t_4 = 40 \).

3. In the case of medium stay time, the value of the membership function parameter of fuzzy stay time is: \( \Delta t_1 = 20, \Delta t_2 = 35, \Delta t_3 = 110, \Delta t_4 = 150 \).

4. In the case of long stay time, the value of the membership function parameter of fuzzy stay time is: \( \Delta t_1 = 100, \Delta t_2 = 130, \Delta t_3 = 220, \Delta t_4 = 270 \).

5. In the case of very long stay time, the value of the membership function parameter of fuzzy stay time is: \( \Delta t_1 = 210, \Delta t_2 = 240 \).

where, \( \Delta t \) is in minutes. Thus, the membership function of fuzzy stay time is obtained, as shown in Figure 1.

Five fuzzy stay time membership degrees of each stay point are calculated through the fuzzy stay time membership function. As shown in Table 3, the fuzzy stay time membership degrees of some stay points in the case of short stay time are obtained. In this way, the fuzzy semantic trajectory FST defined in Definition 7 is obtained.

5.3. Running Time Comparison. Previous studies have shown that the operation efficiency and frequent pattern mining effect of the algorithms improved based on PrefixSpan algorithm are better than those of the same type. Therefore, this section compares the running time of FST-FPM algorithm with PrefixSpan [10] algorithm, PrefixSpan-x [12] algorithm and LFFT2 [13] algorithm on the two datasets in the same experimental environment. The running time of each algorithm on two datasets is observed by setting different minimum support thresholds. In this experiment, the running time of each algorithm is compared when the minimum support threshold is 20%, 30%, 40%, and 50%, as shown in Figure 2 and 3. Figure 2 shows the running time comparison of FST-FPM algorithm with PrefixSpan algorithm, PrefixSpan-x algorithm and LFFT2 algorithm on the two datasets when the fuzzy stay time membership threshold is 0.2. Figure 3 shows the running time comparison of FST-FPM algorithm on the two datasets when the fuzzy stay time
Figure 1: Fuzzy stay time membership function.

Table 3: Data preprocessing results.

| Personnel number | Track number | Stay point          | Stay time (min) | Time membership |
|------------------|--------------|---------------------|-----------------|-----------------|
| 1                | 1            | (1,09:00, 11:30)    | 150             | 0               |
| 1                | 1            | (4,13:00, 18:00)    | 300             | 0               |
| 1                | 1            | (5,19:00, 19:40)    | 40              | 0               |
| 1                | 2            | (2,12:00, 12:30)    | 30              | 0.67            |
| 1                | 2            | (3,14:00, 17:30)    | 210             | 0               |
| 1                | 2            | (6,18:00, 18:30)    | 30              | 0.67            |

Figure 2: Comparison of running time of each algorithm (a) Geolife dataset (b) RFID positioning dataset.
The membership threshold is 0.2, 0.4, 0.6, 0.8, and 1.0, respectively. The minimum length threshold of frequent sequence of PrefixSpan-x algorithm is set to 1, and the stay time threshold of FST-FPM algorithm is set to 5 minutes.

Figure 2(a) describes the comparison of the running time of FST-FPM algorithm with PrefixSpan algorithm, PrefixSpan-x algorithm, and LFFT2 algorithm when the fuzzy stay time membership threshold is 0.2 and the minimum support threshold is 20%, 30%, 40%, and 50%, respectively, on the Geolife dataset. It can be seen that the running time of each algorithm decreases gradually with the increase of the minimum support threshold. As the minimum support threshold increases, the number of frequent blurring decreases gradually.

Figure 2(b) describes the comparison of the running time of FST-FPM algorithm with PrefixSpan algorithm, PrefixSpan-x algorithm, and LFFT2 algorithm when the fuzzy stay time membership threshold is 0.2, 0.4, 0.6, 0.8, and 1.0 on the RFID positioning dataset. It can be seen that with the increase of the minimum support threshold, the running time of FST-FPM algorithm under various fuzzy stay time membership thresholds gradually decreases, and the higher the fuzzy stay time membership threshold, the shorter the running time. Because the higher the membership threshold of fuzzy stay time is, the higher the constraint on frequent itemsets is, the less projection databases are constructed, and the higher the operation efficiency is.

Figure 3(a) describes how the running time of FST-FPM algorithm changes with the minimum support threshold when the fuzzy stay time membership threshold is 0.2, 0.4, 0.6, 0.8, and 1.0 on the Geolife dataset. It can be seen that with the increase of the minimum support threshold, the running time of FST-FPM algorithm under various fuzzy stay time membership thresholds gradually decreases, and the higher the fuzzy stay time membership threshold, the shorter the running time. When the minimum support threshold is greater than 20%, there is little difference in the running time, because the number of mining is similar.

It can be seen from Figure 3 that the running time of FST-FPM algorithm will gradually decrease with the increase of fuzzy stay time membership threshold. When the membership threshold of fuzzy stay time increases by 0.2, the running time of FST-FPM algorithm decreases by about 15%.

Figure 3(b) describes how the running time of FST-FPM algorithm changes with the minimum support threshold when the fuzzy stay time membership threshold is 0.2, 0.4, 0.6, 0.8, and 1.0 on the RFID positioning dataset. It can be seen that with the increase of the minimum support threshold, the running time of FST-FPM algorithm under various fuzzy stay time membership thresholds gradually decreases, and the higher the fuzzy stay time membership threshold, the shorter the running time. When the minimum support threshold is greater than 20%, there is little difference in the running time, because the number of mining is similar.

5.4. Mining Fuzzy Semantic Trace Frequent Patterns. The FST-FPM algorithm proposed in this paper is used to mine frequent patterns of fuzzy semantic trajectories. In order to explore the impact of minimum support threshold and fuzzy stay time membership threshold on mining, the number of mining under different minimum support threshold and fuzzy
stay time membership threshold were compared on Geolife dataset and RFID positioning dataset, respectively. Because other benchmark algorithms cannot mine, the mining quantity cannot be compared with other benchmark algorithms. Therefore, this section explores the impact on mining from two aspects: different minimum support thresholds and different fuzzy stay time membership thresholds.

Fixed FST-FPM fuzzy stay time membership threshold unchanged, set different minimum support thresholds for multiple groups of comparative experiments, and compare the number of five fuzzy stay time cases mined under different minimum support thresholds. When FST-FPM algorithm fuzzy stay time membership threshold is set to 0.4, 0.6, 0.8, and 1.0, respectively, the mining quantity is shown in Figure 4 and 5.

The four graphs in Figure 4, respectively, describe the mining quantity of Geolife dataset under different minimum support thresholds when the fuzzy stay time membership thresholds are 0.4, 0.6, 0.8, and 1.0.

It can be seen from the four figures in Figure 4 that with the increase of the minimum support threshold, the number of mined data gradually decreases. When the minimum support threshold is 20% and 30%, more mining can be done. When the minimum support threshold is greater than 30%, the number of mining decreases sharply. When the minimum support threshold is 50%, only a small amount can be mined. The data mined in the Geolife dataset are mainly those with short and medium stay time. It can be seen from the four figures in Figure 4 that with the increase of fuzzy stay time membership threshold, the number of people...
with the same stay time level also decreases in equal proportion.

The four diagrams in Figure 5, respectively, describe the mining quantity of RFID positioning datasets under different minimum support thresholds when the fuzzy stay time membership thresholds are 0.4, 0.6, 0.8, and 1.0.

From the four figures in Figure 5, we can see that with the increase of the minimum support threshold, the number of mining data gradually decreases. When the minimum support threshold is 20% and 30%, more mining can be done. When the minimum support threshold is 40% and 50%, it is almost impossible to mine. It is mainly due to the characteristics of RFID datasets, mainly for target tracks with medium and long stay time and the support is not high. When the minimum support threshold is greater than 40%, few meet the conditions. It can be seen from the four figures in Figure 5 that with the increase of fuzzy stay time membership threshold, the number of people with the same stay time level also decreases in equal proportion. When the minimum support threshold is 20% and 30%, the data mined from RFID datasets are mainly those with medium or long stay time. When the minimum support threshold is greater than 40%, it is almost impossible to mine.

It can be seen from Figures 4 and 5 that with the increase of the minimum support threshold, the number of the same fuzzy stay time type is reduced in equal proportion, and the total number is greatly reduced, because with the increase of the minimum support threshold, the requirements for the

Figure 5: RFID dataset mining FTFP (a) $\rho = 0.4$ (b) $\rho = 0.6$ (c) $\rho = 0.8$ (d) $\rho = 1.0$. 
frequency of are higher, thus reducing the number that meets the requirements.

6. Conclusions

In this paper, a fuzzy semantic trajectory frequent pattern mining method is proposed for the moving trajectory data and stay time of targets in security environment. Based on the fuzzy set theory, the stay time of the target is divided into five grades, and the membership degree calculation method of the fuzzy stay time is defined. The FST-FPM algorithm is proposed and tested on the Geolife public dataset and the RFID positioning dataset of an integrated security system. Experimental results show that FST-FPM algorithm can effectively mine frequent patterns with fuzzy semantics, and the running time is reduced by 54% compared with PrefixSpan algorithm, 23% compared with LFFT2 algorithm, and 10% compared with PrefixSpan-x algorithm. It is proved that FST-FPM algorithm can help analyze the behavior characteristics of targets in security environment, and plays an important role in evaluating the threat degree of targets.

Data Availability

The data used to support the findings of this study can be downloaded from the website below https://www.microsoft.com/en-us/download/details.aspx?id=52367.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] B. Zou and M. Li, "Vulnerability learning of adversary paths in Physical Protection Systems using AMC/EASI," Progress in Nuclear Energy, vol. 134, Article ID 103666, 2021.

[2] C. Zhu, S. Zhao, Y. Xia, and Li. Li, "An improved three-point localization method based on RSS for transceiver separation RFID systems," Measurement, vol. 187, Article ID 110283, 2022.

[3] S. Hayward, J. Earps, R. van Lopik, J. Tribe, and A. West, "A novel inertial positioning update method, using passive RFID tags, for indoor asset localisation," CIRP Journal of Manufacturing Science and Technology, vol. 35, pp. 968–982, 2021.

[4] U. Ahmed, G. Srivastava, Y. Djienouri, C. Jerry, and L. Wei, "Knowledge graph based trajectory outlier detection in sustainable smart cities," Sustainable Cities and Society, vol. 78, Article ID 103580, 2022.

[5] U. Ahmed, G. Srivastava, Y. Djienouri, and C. Jerry, "Deviation Point Curriculum Learning for Trajectory Outlier Detection in Cooperative Intelligent Transport Systems," IEEE Transactions on Intelligent Transportation Systems, vol. 12, no. 2, pp. 1–10, 2021.

[6] C. Renso, V. Bogorny, K. Matwin, and J. Macedo, "Multiple-aspect analysis of semantic trajectories (MASTER)," International Journal of Geographical Information Science, vol. 35, no. 4, pp. 763–766, 2021.

[7] G. Cai, K. Lee, and I. Lee, "Mining mobility patterns from geotagged photos through semantic trajectory clustering," Cybernetics & Systems, vol. 49, no. 4, pp. 234–256, 2018.

[8] L. Shabtay, P. Fournier, and R. Dattner, "A guided FP-Growth algorithm for mining multitude-targeted item-sets and class association rules in imbalanced data," Information Sciences, vol. 553, no. 5, pp. 353–375, 2021.

[9] R. Davashi, "UP-tree & UP-Mine: a fast method based on upper bound for frequent pattern mining from uncertain data," Engineering Applications of Artificial Intelligence, vol. 106, no. 4, Article ID 104477, 2021.

[10] J. Han, J. Pei, and X. Yan, "From sequential pattern mining to structured pattern mining: a pattern-growth approach," Journal of Computer Science and Technology, vol. 19, no. 3, pp. 257–279, 2004.

[11] T. Niyazmand and I. Izadi, "Pattern mining in alarm flood sequences using a modified PrefixSpan algorithm," ISA Transactions, vol. 90, no. 8, pp. 287–293, 2019.

[12] F. Xue, Z. Shan, and L. Yan, "Multiple trajectory feature detection technology based on data mining," Computer Science, vol. 43, no. 5, pp. 91–95, 2016.

[13] T. Wu, J. Lin, U. Chen, G. Srivastava, and X. Lv, "An efficient algorithm for fuzzy frequent itemset mining," Journal of Intelligent and Fuzzy Systems, vol. 38, no. 5, pp. 5787–5797, 2020.

[14] C. Chen, T. Hong, and Y. Li, "Fuzzy association rule mining with type-2 membership functions," Intelligent Information and Database Systems, vol. 16, no. 8, pp. 128–134, 2015.

[15] J. Lin, Y. Zhang, P. Fournier, and T. Hong, "Efficiently updating the discovered multiple fuzzy frequent itemsets with transaction insertion," International Journal of Fuzzy Systems, vol. 20, no. 8, pp. 2440–2457, 2018.

[16] J. Li and J. Su, "Semantic trajectory frequent pattern mining model," International Journal on Semantic Web and Information Systems, vol. 18, no. 1, pp. 1–20, 2022.

[17] M. Liu, C. Zhang, and C. Yu, "Knowledge diffusion trajectories in the Pythagorean fuzzy field based on main path analysis," International Journal of Intelligent Computing and Cybernetics, vol. 8, no. 3, pp. 1–15, 2021.

[18] Y. Zheng, L. Zhang, X. Xie, and Y. Wei, "Mining interesting locations and travel sequences from GPS trajectories," in Proceedings of the 18th International Conference on World Wide Web, pp. 791–800, New York, NY, United States, April 2009.