Analysis of Shunting Yard Usage and Train Unit Clustering

Ivo Yordanov[1]

Supervisor(s): Prof. Dr. M.M. de Weerdt[1]**, I.K. Hanou**[1]

[1]**EEMCS, Delft University of Technology, The Netherlands**

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 23, 2024

Name of the student: Ivo Yordanov
Final project course: CSE3000 Research Project
Thesis committee: Prof. Dr. M. M. de Weerdt, I. K. Hanou, Prof. J. Sun

Abstract

This research aims to find patterns in the live position data of trains within shunting yards. These patterns can be converted to heuristics and applied in algorithms developed by railway operators in the Netherlands to tackle the Train Unit Shunting Problem. The usage patterns were extracted from real-world location data of train units, Specifically, this research focuses on finding patterns in the paths taken in both temporal and spatial metrics. These investigations identified the busiest times in the shunting yards according to various metrics, such as the total number of trains and the number of serviced trains. These metrics are extracted from the type of train movement at a given moment, which was provided with the dataset. The accuracy of the provided train movement category has been analyzed and shown to be inaccurate for uses within shunting yards compared to a proposed approach. Finally, an algorithm for classifying train units that belong to the same train has been shown to have initial success.

1 Introduction

In the Netherlands, only daily train schedules are generated automatically, while unused trains have traditionally been shunted manually. Shunting yards are where trains are positioned when unused or serviced. The Train Unit Shunting Problem with Service Scheduling (TUSPwSS) is about designing a schedule to allocate trains between a station and a shunting yard, factoring in special requirements, such as servicing and maintenance, and the position of other trains [1]. Unfortunately, this is an NP-Hard problem, which has historically been solved by operators manually creating schedules based on certain rules of thumb. Nowadays, various algorithms have been examined with partial success. The most successful model takes hours to compute for large instances and is not consistent or robust to small delays [2]. An examination is needed into whether and how human insight could be incorporated into the decision-making of solvers for the TUSPwSS to improve explainability and performance [3]. If these metrics improve, automation may become a larger part of creating shunting plans, leaving planners in a supervisory role. Additionally, the operational capacity of shunting yards may increase as the performance improves, which would reduce the already existing pressure on the expansion of the infrastructure [1].

In preparation for this and other future research, live GPS data of train units within major stations and shunting yards have been collected, analyzed, and prepared. The real paths that trains took within shunting yards are a solution to the TUSPwSS because they adhere to all physical and organizational limitations. Additionally, analyzing this data allows establishing patterns over long periods of how the shunting yards are used. Analyzing multiple shunting yards and incorporating insight is listed as a potential direction of future research in the literature [3]. Therefore, the main focus of this paper is to establish patterns in the usage of shunting yards as well as examine methods that can be utilized in future research to establish such patterns.

The dataset analyzed by this research has been sourced, checked, and processed by ProRail. The GPS locations of NS train units over 10 months in 7 areas across the Netherlands are presented. The dataset also includes descriptions of the tracks and their positions, including some labeling of shunting yards. No documentation was included, so the assumptions made about the data and relevant columns are documented. The shunting yards and periods utilized are also noted next to the different examinations conducted in this research. In general, part of the experiments focus only on the shunting yard at Amersfoort Bokkeduinen and another part uses data from all shunting yards. The number of locations and the long period of documented locations allow researchers great flexibility in examining how applicable patterns are.

As the need for accommodating larger passenger volumes arises, the Dutch railway system has to better utilize existing infrastructure due to space limitations [2]. The first step in incorporating extracting human insight is analyzing how shunting yards have been used spatially and temporally. Timewise, if algorithms assign more operations during times when the yards are less busy, that would necessarily increase the operational capacity during peak hours. To do so, they may incorporate historic usage statistics that factor in the specific day or month, which may result in faster convergence to a solution if similarities exist. While seasonal patterns may not be accounted for in daily schedules, they may reveal structural differences in train usage, which can be incorporated into all aspects of train planning. Additionally, operators have historically created rules for using the tracks within a shutting yard. If an algorithm rewards following those rules, explainability will be improved as the solutions are more spatially similar to existing schedules. For these reasons, analysis of the usage of shunting yards is a first step towards any pattern investigation.

In operating shunting yards, three main tasks are achieved - matching of train units to combine, parking, and routing [2]. Patterns can be derived in all three categories. However, to do so train paths have to be correctly classified according to the current movement of trains. The dataset contained a column called "ActivityType", which is assumed to classify a movement point according to the current train movement direction. This information can be used to derive conclusions on when trains are entering and exiting the shunting yards, are being serviced, are stopped, and much more. Therefore, because of its importance, this paper examines the accuracy of this column based on a proposed novel approach. This analysis may highlight the need to rethink the approach used in generating this column to aid future research on finding patterns using it.

To analyze how train units are combined in the matching task in shunting yards, the dataset first has to include which train units belong to the same train. Answering this question unlocks establishing patterns on when and where train units are attached and detached. More importantly, any research that matches trains from multiple days would be able to factor in the fact that the train arriving at a similar time may have a different number of train units. Visually, when examining the animation of train movements it is very easy to spot the train units that move together. However, establishing such moving clusters is a difficult problem in Data Science with limited research on the topic. An experiment was conducted to determine whether a modified existing algorithm could be applied in the context of trains.

The main body of this report is organized into several key sections for clarity and thoroughness. Section 2 reviews the existing literature, providing detailed explanations of the algorithms employed in the subsequent experiments. Section 3 delves into the data processing methods, outlining the specific parameters used to achieve the results. Section 4 presents the experimental results, highlighting the key findings. Section 5 addresses the ethical considerations relevant to this research. Finally, Sections 6 and 7 offer a comprehensive discussion of the results and identify the main conclusions, synthesizing the insights gained from the study.

2 Background

The most promising approach for generating solutions to the TUSPwSS has been a local search algorithm [2]. The algorithm proposed relaxes some constraints but induces penalties for breaking them - thus transforming the problem into a local optimization [2]. Despite its success, the local search algorithm is still computationally heavy and needs improvement to become more robust and consistent [3]. Therefore, this research proposes that patterns established in this and future studies be incorporated into the penalties and reward system that the algorithm already utilizes to test whether performance and explainability are improved with minimal overhead.

Inspiration in finding patterns in the movement trajectories of trains can be drawn from the analysis of the GPS movement of other objects. In animal trajectory analysis, path segmentation is used to determine changes in the movement [4]. Path segmentation involves splitting a path into subcomponents based on a perceived difference in some path characteristic. The same paper analyzes the effectiveness of a plethora of characteristics based on individual steps of a trajectory or across multiple steps. The authors showcase the effectiveness of this method in identifying movement patterns, from which they conclude on ecological processes. In the case of trains, there are many more restrictions, such as trains moving on specific tracks and rarely stopping. This research segments paths based on train stops as it assumes a train stop occurs because of a change of destination or purpose of movement.

To examine train units that belong to the same train, the problem of identifying moving clusters has been researched. The Coherent Moving Cluster (CMC) algorithm builds upon the drawbacks of other methods that do not capture clusters in their entirety [5]. The algorithm applies density-based clustering over multiple time frames and examines how the clusters change in time to determine which points move together. Its main drawback is performance, as even the authors present multiple optimizations. However, CMC has been chosen because this research wants to establish whether this family of algorithms can even be applied to the dataset at hand. The CMC algorithm requires multiple parameters - a minimum number of points in a cluster as well as a minimum number of consecutive timestamps to identify a moving cluster in addition to the other parameters required by the clustering method. Overall, since the train paths are a sequence of GPS coordinates, CMC algorithm is applicable to identify train units that belong to the same train.

3 Methodology

All sub-questions necessitated a common data structure that captures the path that a train took. In the data, the "GroupIdHash" column is assumed to represent the area, train number, and time of entering the simulation concatenated in a string format. Therefore, the individual locations were grouped according to that column to produce a path structure consisting of locations and the times when they happened. Thus, the "GroupIdHash" column effectively facilitated the organization of the data into coherent train paths, enabling a comprehensive analysis of train movements

Splitting the tracks of a shunting yard into multiple categories allows for spatial analysis of yards. The usage of service yards - which are often a part of the shunting yard - can also be analyzed. In the tracks dataset, the "Geocode" column is assumed to identify the general area that the track is in. Sometimes, the general area marks the shunting yard, and other times the station and yard. Additionally, no service yards are identified. For this reason, wherever analysis of the movement within a service yard is conducted, the Amersfoort Bokkeduinen yard is used. Based on satellite imagery and open source maps, the service yard is manually identified as shown in Figure 1.

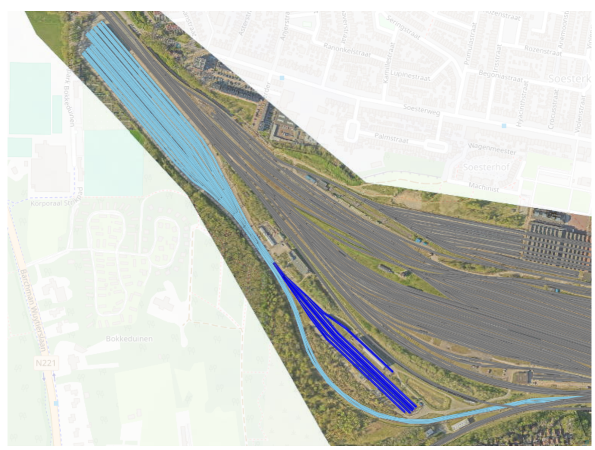

Figure 1: Amersfoort Bokkeduinen Shunting Yard (light blue) and Service Yard (dark blue)

3.1 Spatial and temporal usage analysis

Finding a statistically significant pattern in the usage of the shutting yards throughout the day is vital for building a heuristic based on past usages. Specifically, an algorithm may incorporate the usage of the yard on the same part of the day, month, or season, and that is why the investigations are split into finding patterns on such scales. For the trains within a yard, they can be further categorized into trains that are being parked, shunted, or serviced. The "ActivityType" column could be used to determine the state of a train at a given time, however, this approach was not preferred due to conclusions drawn later in this paper about its accuracy. Instead, the novel sub-path classification, described further below, is used to determine the state of a train within the yard. The investigations of the usages of shunting yards are valuable because a solution should distribute movements as much as possible to the periods with less activity in the yard.

In all experiments, the day is split into hourly bins and the number of trains that match a criteria is recorded. This approach was preferred over a continuous analysis where trains are recorded as soon as they meet criteria, such as entering the shunting yard. The reason that this approach was not chosen is that the data itself is not continuous and that the time bins can be made as small as needed, approximating a continuous function. When analyzing months, weekends, and official public holidays are excluded. The results are then averaged and displayed alongside the standard deviation of the samples per bin.

The spatial usage of a shunting yard is also analyzed to determine whether tracks are utilized evenly. If the usage of the tracks is not even, higher maintenance and bottlenecks may be introduced. The shunting yard is split into its tracks and trains are recorded into categories and time bins similarly as above. In addition to hourly usage graphs, when analyzing track usage, heatmaps are also created with the monthly averages of their maximum capacities.

3.2 Sub-paths classification

To classify the direction of movement of a train, its path needs to be firstly split into sub-paths. The "ActivityType" column classifies movements into the following categories: "Short Stop", "Long Stop", "Shunting", "Entering", "Entry", and "Exiting". From these, "Short Stop" and "Long Stop" are assumed to represent what they claim, since these are the only activities that have time duration. To split a path into sub-paths, appropriate criteria need to be selected as to when to split, as described in Section 2. In this experiment, sub-paths are split by any type of stop - both long and short. The decision to split only by long stops was considered, however, train units have short stops in the service yards, where it is assumed that they are serviced and not just shunted.

Sub-paths classification is vital in any frequency analysis of the usage of shunting yards, as briefly mentioned above. Knowing when a train is entering, exiting, being serviced, or shunted can provide insight into all of these activities. With the newly created sub-paths, they can be classified according to the following certain rules. The starts and ends of the sub-path can be used to establish whether the train is entering or exiting the shunting yard or service yard. Also, the location of a stop signifies whether the train is parked, shunted, or serviced. Thus, all components of a sub-path are labeled the same way because a train does not change its purpose of movement without a stop. In the original data "ActivityType", classification this rule has not been followed, so when comparing with the novel method above, the first label is taken to signify the entire sub-path.

3.3 Clustering moving train units

Identifying train units that belong to the same train is vital for any multi-day matching on trains based on, for example, arrival times. This is because the same train can have more individual train units. Furthermore, pattern investigations may choose to ignore the train units after the first one to avoid bias from larger trains. The Coherent Moving Cluster (CMC) algorithm is examined in this research to cluster train units that move together based on their latitude and longitude GPS location.

To perform clustering on multiple time frames, an initial preprocessing of the dataset is required. The locations of individual train units are recorded every 10 seconds, however, they are offset from each other. All trains need to be clustered at the same time in the specified algorithm. This requires interpolation of the position of a train unit when the time frame falls in between its recorded locations. The choice made in this experiment is taking the closest recorded location in time. The starting and ending times of the path were recorded to construct a trajectory, as defined by the algorithm.

After the initial processing, the parameters of the algorithm were set. The minimal number of points in a cluster was set at two as that is the minimal number of train units used to make a composition. Multiple values for the number of consecutive time frames required to form a moving cluster were examined. The value of epsilon, required by the inner DBSCAN clustering algorithm within CMC, was set at 250 after manually visualizing the movements of train units. Finally, the algorithm has been slightly modified to return the id of the paths, as described by the "GroupIdHash" column instead of the points themselves.

4 Results

4.1 Spatial and temporal usage analysis

The most promising results in the investigations were analyzing the usage of the Amersfoort Bokkeduinen yard throughout the day. For conciseness, only the results from May 2023 are displayed, however, the numbers from other months closely resemble the following analysis and are displayed for completeness in Appendix A.

Analyzing the total number of trains in the shunting yard reveals the broadest insight into its usage. Figure 2 showcases the average total number of trains in the yard throughout the day during the workdays of May 2023. Similar such graphs were constructed for each of the 10 following months, that the dataset covered. There are two spikes in the number of trains from 12:00 to 14:00 and from 1:00 to 5:00, however, the spike in the latter is, on average, around twice as large. Additionally, the sharpest influx of trains in the shunting yard occurs between 21:00 and 1:00, and the sharpest decrease is from 5:00 to 7:00.

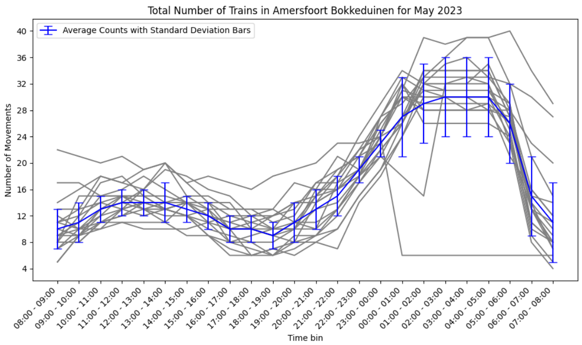

Figure 2: Average number of trains in Amersfoort Bokkeduinen in the workdays of May 2023.

Additional insight is gained by splitting the trains within the shunting yards into groups based on their classifications. For example, the vast majority of the trains in the peaks are accounted for by the parked trains. Interestingly, the vast majority of servicing of trains, as accounted for by stops in the service yard, occurs from 23:00 to 6:00 with no spike in the afternoon, as demonstrated in Figure 3. Figure 4, on the other hand, demonstrates that there is a single peak of the number of shunted trains at one time between 21:00 and 1:00.

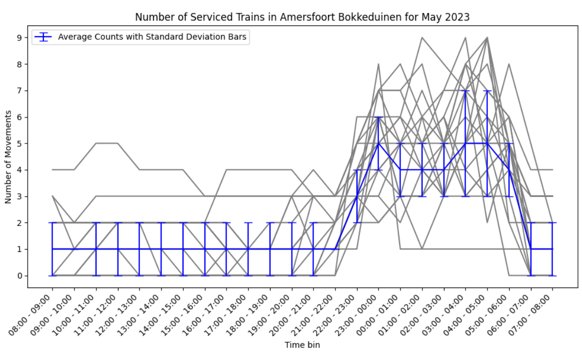

Figure 3: Average number of serviced trains in Amersfoort Bokkeduinen in the workdays of May 2023.

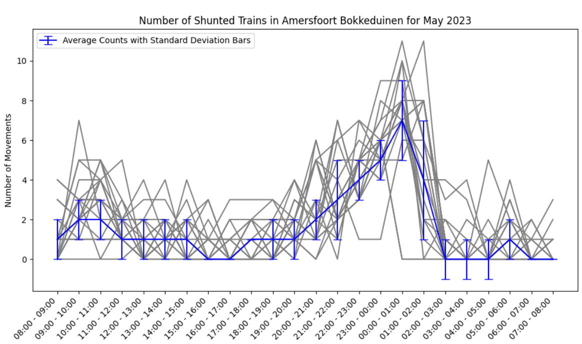

Figure 4: Average number of shunted trains in Amersfoort Bokkeduinen in the workdays of May 2023.

Some such plots of the usage throughout the day did not yield significant results. Investigations into the usages of different tracks revealed only that some tracks have a smaller maximum capacity than others on particular days, however, patterns over longer periods were not established. Finally, no statistically significant conclusions arise from comparing the average monthly usage across seasons.

4.2 Sub-Paths classification

When the train paths were split into sub-paths using their stops and classified according to the origin and destination, the produced classification did not produce results similar to the ”ActivityType” column. Firstly, as explained in Section 3, the provided classification in that column sometimes changes in the middle of a sub-path, and therefore, only the classification of the first point of a sub-path was used in the following comparison. Figure 5 shows a confusion map of the produced classifications versus the one extracted from the provided column. For this experiment, only the data from the Amersfoort Bokkeduinen yard from all paths in May 2023 was analyzed because it already reveals significant differences.

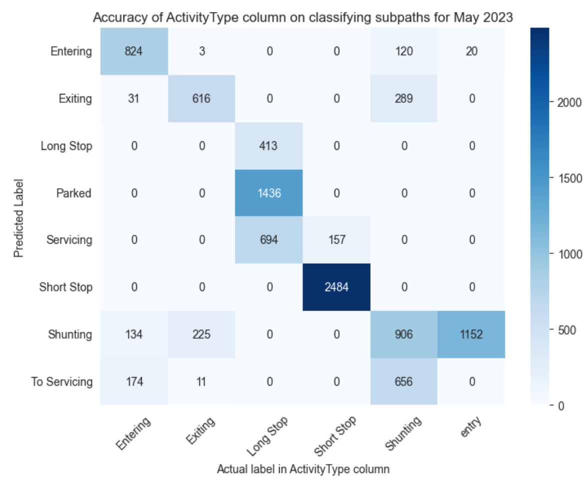

Figure 5: Confusion matrix of the identified classification of a sub-path (rows) vs the classification in the ”ActivityType” column (columns) for all sub-paths in Amersfoort Bokkeduinen over May 2023.

From the confusion matrix, multiple conclusions can be drawn about the actual labels in the ”ActivityType” column (columns in the matrix). Firstly, the ”Short Stop” and ”Long Stop” are classifying stops correctly, as expected, since they were taken as ground truth when splitting the paths. However, the ”Entering” and ”Exiting” labels are somewhat inaccurate. There are rare occasions when the predicted sub-path classification is the complete opposite of their actual classification - 31 and 3 times respectively. The most widely inaccurate classification, however, is ”Shunting” where it matched with the predicted label for only 46% of the times when it was given. Finally, a very strong correlation exists between the ”entry” actual label and the ”Shunting” predicted label.

4.3 Clustering moving train units

The Coherent Moving Clustering algorithm was for the combined data of all 7 geographic locations included. However, it is important to note that there is no ground truth, which means that the best way to assess the accuracy of the found moving clusters is by professionals examining them. The algorithm was run for each day of May 2023 separately, as otherwise, it would be computationally infeasible. The resulting clusters were then analyzed and it was established that zero clusters were formed from train units in different cities. Additionally, the maximum size of an identified moving cluster is three. An interactive frame-by-frame visualization was created to make the verifying of the correctness of this algorithm easier.

5 Responsible Research

The main ethical consideration is data security because this research analyses sensitive data about the locations of trains. If data security is breached, malicious actors may utilize location data to disrupt, vandalize, and conduct terrorist activities [6]. Withing the railway service, this would affect the "Availability" category from the CIA triad, a widely accepted model within Information Security. It lists three categories all necessary for the normal operations of a service. "Availability" is seen as a vital component of a digital system by clients and business executives [7], and the relevance of the CIA triad is growing with the increasing digitization of the railway service. Furthermore, by adhering to data security, this research addresses the Responsibility principle in the Netherlands Code of Conduct for Research Integrity towards the railway companies that shared their data with such expectations [8]. The measures taken to address data security are using Azure Blob storage and not storing any data locally. Azure Blob storage requires 2-step authentication each time the data is used in code. By using such measures, data security is ensured by a reputable service, and authorization is done regularly.

Another consideration is data reliability because the provided data has in no way been checked or processed by myself. Furthermore, the gathering and processing of the data have not been disclosed and scrutinized. The consequences of unreliable data are that inaccurate patterns would be discovered and disclosed, which may negatively impact future research.

6 Discussion

The patterns discussed in this paper can be applied in approaches aside from the main local search algorithm under consideration by the railway companies. A Deep Reinforcement Learning algorithm has been developed and tested, with more limited success than the algorithm presented in Section 2 [3]. The main advantages described are that this method generates solutions within seconds, which are also more consistent. However, officials at NS and the literature confirm that the limiting factor for further development of this approach has been that it often generates only partial solutions [3]. Partial solutions are caused because this method has a limited look-ahead capability [9] which may result in an irrecoverable state. However, the authors of this approach have demonstrated its success for small problem instances [9]. Therefore, this approach still has the potential to be tested when the shunting yards are least active, the exact times of which have been documented in this research. This method utilizes a simple reward function based on correctness, which an agent attempts to maximize [9]. The authors themselves leave creating more complex value functions as future research. Therefore, incorporating human insight and the heuristics developed for the main algorithm into the reward function may increase the size of instances that this method can solve.

6.1 Spatial and temporal usage analysis

The analysis of the Amersfoort Bokkeduinen yard's usage throughout the day reveals significant patterns that can inform operational strategies and resource allocation. These findings are consistent across the subsequent ten months, indicating stable and predictable usage patterns. The results closely align with the peak hours of the railways, as widely recognized by NS, which are from 6:30 to 9:00, and from 16:00 to 18:30. The fact that they are peak hours of usage indicates that there should be the least amount of trains in the shunting yards, which is exactly what is observed. Conversely, both the largest rise in the number of trains and the number of shunted trains coincide between 21:00 and 1:00, which indicates that this period is most difficult to plan. Given these results, any activities conducted in the lowest-intensity periods should be rewarded, while those in the high-intensity periods should be diminished.

The daily analysis of the entire dataset also revealed anomaly days, which greatly differ from the movement patterns of the rest of the workdays within that month. Weekends and holidays are excluded precisely to diminish such anomalies, therefore, the reasons for their appearance can be applicable depending on the reason. To achieve this, anomaly detection algorithms could be examined in future work. When identified and removed from consideration, the monthly usage patterns and conclusions drawn from them will be strengthened.

6.2 Clustering moving train units

The analysis of the clusters generated by the Coherent Moving Cluster algorithm showcases its potential to correctly classify train units that belong to the same train. The only piece of ground truth in the data is the different cities where the trains are positioned. If the algorithm had returned a cluster containing train units from different cities, its credibility would have been diminished as no such train composition can exist in the data. Additionally, if the size of the largest clusters was more than three, that would indicate an impossibly long train. These two main factors are the first step towards validating the accuracy and reliability of the algorithm in the data.

Many false positives could still exist within the identified clusters but could be filtered out using additional restrictions. One particular problem is that trains within shunting yards are physically very close to each other, and the examined algorithm does not factor in the topology of the rail tracks. This can be solved by representing the tracks as a graph and calculating the distances of train units based, not on the Euclidean distance of their coordinates, but on the shortest distance from their tracks. Another approach is to factor in the property of train units moving together to have the same velocity. Thus, train units can be classified on similar positions and similar velocities.

7 Conclusions and Future Work

This research highlights the significant potential of utilizing GPS data to enhance the scheduling and operational efficiency of shunting yards in the Netherlands. By analyzing the spatial and temporal usage patterns of shunting yards, particularly at Amersfoort Bokkeduinen, this study identifies critical insights into yard utilization and train movements. The minima and maxima of activity in that shunting yard are established and analyzed through the use of multiple train activity graphs. This insight can be incorporated into existing planners that create the daily planning of movements within shunting yards by rewarding the distribution of activities away from busy periods. To do so, both the historic daily and average monthly usage can be utilized. Establishing whether incorporating such a reward function results in increased performance without inducing drawbacks is left to future research.

Additionally, this research aids in conducting future pattern analysis by questioning the validity of the identified movement classification. If train sub-paths are correctly classified, more accurate patterns can be established on when and where various activities are conducted within shunting yards. The path-splitting approach based on train stops was described and compared to the existing movement activity type labels. This analysis demonstrated the inaccuracy of the labels within the dataset to specifically identify activities within shunting yards, such as when a train is entering or leaving. Future research may consider and compare different characteristics to split a train path into sub-paths and potentially compare their effectiveness with the official daily planning.

Finally, this research analyzed whether an inherent property of the data can be uncovered - train units moving together as part of a single train. If successful, trains from different days can successfully be matched despite having different lengths and patterns related to the splitting and combining of train units can be analyzed. The analyzed algorithm has promising initial results, however, analysis of experts is required to measure its precise success. In particular, the classified train compositions do not break immediately obvious physical constraints. Nevertheless, once ground truth is available, potential avenues to improve accuracy are detailed.

A The total number of trains in Amersfoort Bokkeduinen over 10 months

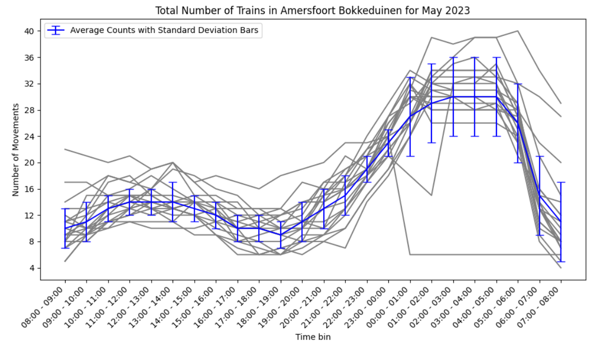

Figure 6: Average number of trains in Amersfoort Bokkeduinen in the workdays of May 2023.

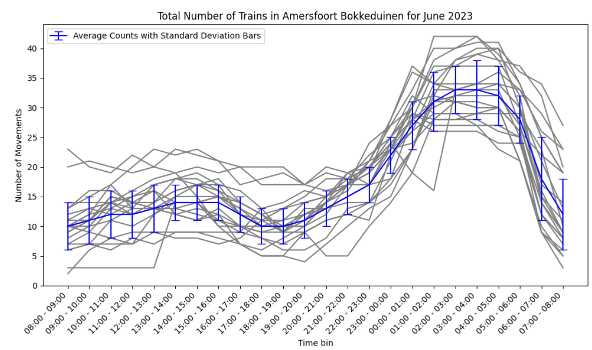

Figure 7: Average number of trains in Amersfoort Bokkeduinen in the workdays of June 2023.

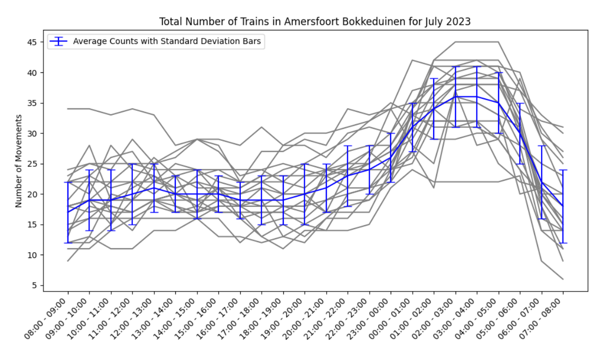

Figure 8: Average number of trains in Amersfoort Bokkeduinen in the workdays of July 2023.

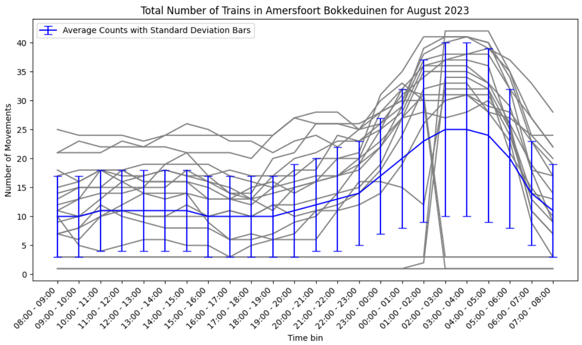

Figure 9: Average number of trains in Amersfoort Bokkeduinen in the workdays of August 2023.

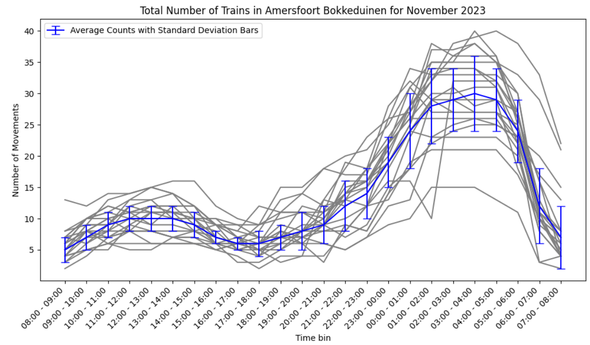

Figure 12: Average number of trains in Amersfoort Bokkeduinen in the workdays of November 2023.

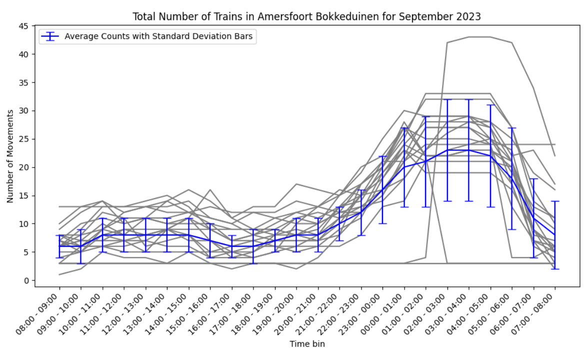

Figure 10: Average number of trains in Amersfoort Bokkeduinen in the workdays of September 2023.

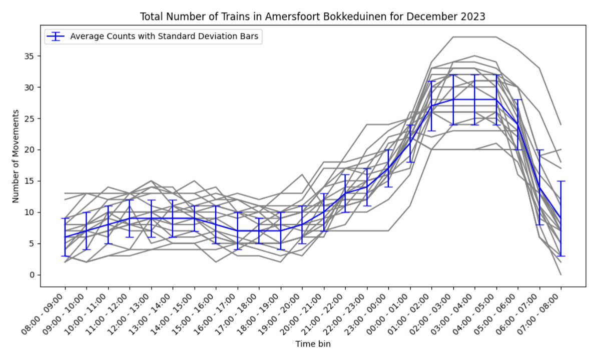

Figure 13: Average number of trains in Amersfoort Bokkeduinen in the workdays of December 2023.

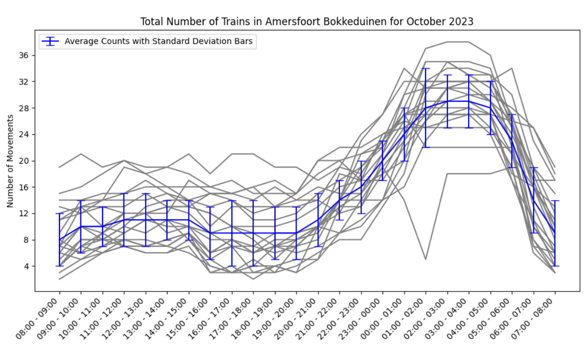

Figure 11: Average number of trains in Amersfoort Bokkeduinen in the workdays of October 2023.

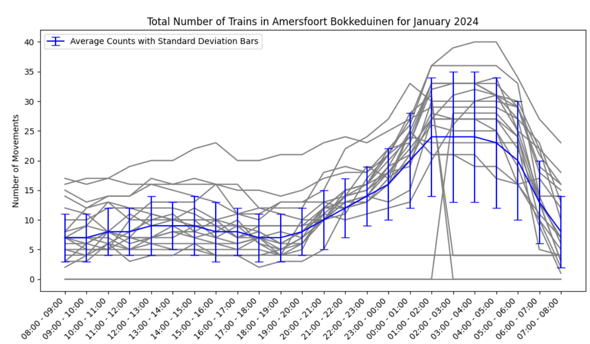

Figure 14: Average number of trains in Amersfoort Bokkeduinen in the workdays of January 2024.

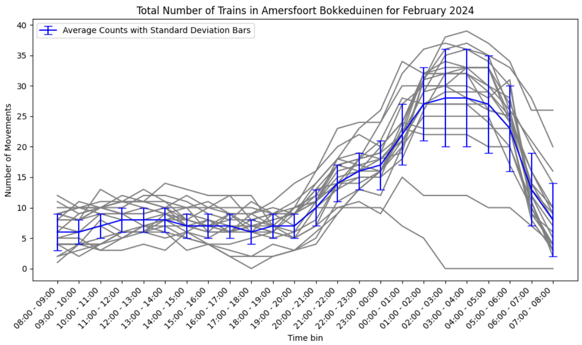

Figure 15: Average number of trains in Amersfoort Bokkeduinen in the workdays of February 2024.

References

[1] I. K. Hanou, S. Dumancic, M. de Weerdt, P. van der Voort, R. van den Broek, and M. van den Akker, "Increasing the capacity of shunting yards within the current infrastructure: A computational perspective," Ph.D. dissertation, Delft University of Technology, 2023.

[2] R. van den Broek, H. Hoogeveen, M. van den Akker, and B. Huisman, "A local search algorithm for train unit shunting with service scheduling," *Transportation Science*, vol. 56, 11 2021.

[3] L. van de Gevel, "How human knowledge can support algorithmic decision-making in the train unit shunting problem - an exemplary study," Master's thesis, Eindhoven University of Technology, 2022.

[4] H. Edelhoff, J. Signer, and N. Balkenhol, "Path segmentation for beginners: an overview of current methods for detecting changes in animal movement patterns," *Movement Ecology*, vol. 4, no. 1, p. 21, 2016. [Online]. Available: https://doi.org/10.1186/s40462-016-0086-5

[5] H. Jeung, M. L. Yiu, X. Zhou, C. S. Jensen, and H. T. Shen, "Discovery of convoys in trajectory databases," 2010.

[6] A. M.-B. Pablo López-Aguilar, Edgar Batista and A. Solanas, "Information security and privacy in railway transportation: A systematic review," *Sensors (Basel)*, vol. 22, no. 20, 2022.

[7] D. C. Spyridon Samonas, "The cia strikes back: Redefining confidentiality, integrity and availability in security," *Journal of Information System Security*, vol. 10, 2014.

[8] KNAW, N. NWO, TO2-federatie, V. Hogescholen, and VSNU, "Nederlandse gedragscode wetenschappelijke integriteit," *DANS*, 2018. [Online]. Available: https://doi.org/10.17026/dans-2cj-nvwu

[9] E. Peer, V. Menkovski, Y. Zhang, and W.-J. Lee, "Shunting trains with deep reinforcement learning," in *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2018, pp. 3063–3068.