Automation in Construction 83 (2017) 237–246

Contents lists available at ScienceDirect

Automation in Construction

journal homepage: www.elsevier.com/locate/autcon

# Identification of workstations in earthwork operations from vehicle GPS data

Jiali Fu[a,*], Erik Jenelius[a], Haris N. Koutsopoulos[a,b]

[a] Department of Transport Science, KTH Royal Institute of Technology, Teknikringen 10, 10044 Stockholm, Sweden
[b] Department of Civil and Environmental Engineering, Northeastern University, 360 Huntington Avenue, Boston, USA

ARTICLE INFO

*Keywords:*
Earthwork operations
Global Positioning System (GPS)
Location detection
Probabilistic model
Kernel density estimation

ABSTRACT

The paper proposes a methodology for the identification of workstations in earthwork operations based on GPS traces from construction vehicles. The model incorporates relevant information extracted from the GPS data to infer locations of different workstations as probability distributions over the environment. Monitoring of workstation locations may support map inference for generating and continuously updating the layout and road network topology of the construction environment. A case study is conducted at a complex earthwork site in Sweden. The workstation identification methodology is used to infer the locations of loading stations based on vehicle speeds and interactions between vehicles, and the locations of dumping stations based on vehicle turning patterns. The results show that the proposed method is able to identify workstations in the earthwork environment efficiently and in sufficient detail.

## 1. Introduction

Heavy construction refers to large-scale projects such as construction of infrastructure (highways, streets and railways), flood control, and mining and quarry operations. Earthwork, in particular, is the processing and moving of large quantities of soil from the earth's surface, and is an important part of the early stages of heavy construction projects. Earthwork operations involve processes such as excavating, hauling, dumping, crushing and compacting soil. Activities are typically equipment-driven, and most frequently include excavators, loaders, compactors and hauling trucks.

Earthwork environments are highly dynamic by nature. Especially during the early stages of projects, the layout of the construction site and the road geometry change continuously. As the operation proceeds, the construction site expands and workstations move further away from each other, driving times between workstations become longer and the loading units may stand idle while waiting for hauling units. The performance of the fleet thus declines, including productivity reduction, increased cost and reduced utilization of certain types of equipment [1]. In such situations, it is crucial that management take actions and adjust the operations so that performance does not deviate much from the plan. Such actions may include increasing the number of hauling vehicles or changing to hauling vehicles with higher capacities. Even if there is no possibility to change the fleet composition, the negative effects due to site expansion can be reduced by adjusting operating methods. Such alterative operating methods may include, e.g., not loading the hauling trucks to their full capacity and thus reducing the driving time, fuel consumption of hauling units, and idle time for loading units.

The project management thus needs to regularly update the map of the environment in order to accurately plan and monitor the work process. It is hence important to be able to track the number and locations of significant areas, including loading and dumping workstations, haulage roads, road intersections etc., and the network of transport paths between them [2]. The rate of change of workstations is high at construction sites of moderate scales. In large construction sites, the workstations are more significant and the locations are most likely known in the designing phase of the projects. In any case, the locations of certain types of workstations, such as loading stations, are moving gradually further away as the projects progress. It is thus necessary to follow the development of projects and make appropriate adjustment in the operations so that the performance is not much affected.

Traditionally, mapping of the site layout requires manual data collection, which can be time and resource consuming [3]. Some large-scale heavy construction operations even engage helicopters with advanced laser scanning technology to take aerial photographs with 3-dimensional geographical information of the site at regular intervals. This method can help managers to accurately deduce the site layout and road geometry at a high accuracy, but is expensive to repeat frequently.

The recent spread of mobile GPS devices opens up the possibility for

* Corresponding author.
*E-mail addresses:* jiali.fu@abe.kth.se (J. Fu), erik.jenelius@abe.kth.se (E. Jenelius), h.koutsopoulos@neu.edu (H.N. Koutsopoulos).

http://dx.doi.org/10.1016/j.autcon.2017.08.023
Received 7 December 2016; Received in revised form 2 June 2017; Accepted 10 August 2017

easy collection of location data, and facilitates the development of a variety of location-aware applications. For example, information from vehicles' movement data is already finding applications in areas such as city planning [4], real-time traffic management [5], and fleet management [6]. GPS devices have become standard equipment for construction vehicles in recent years.

The paper proposes a methodology for inferring the locations of distinct types of workstations in earthwork operations from vehicle GPS data logs. The proposed method makes use of the "free" GPS trace data from daily operations, and does not require other signals from vehicle Controller Area Network (CAN) buses or prior knowledge of the operating environment. Compared to manually identifying workstations from aerial photographs, the proposed approach does not require input or experience in the construction field on the part of the users. Thus, the approach may identify the locations of critical workstations at a fraction of the cost of traditional approaches. The identification of workstation locations may be used as part of a wider map inference framework for generating and updating the layout and road network topology of the construction environment as it evolves over time. The problem is important in light of the use of simulation-based optimization tools for modelling the complex characteristics of earthwork operations, and for allocating the most suitable vehicle fleet and operating methods [7].

The methodology employs various characteristics in vehicle movements and interactions during earthwork operations to generate probability distributions over the site geometry for the locations of different workstations. Specific models are developed in order to infer the locations of loading stations, where excavated material is loaded onto hauling trucks, based on interactions between loading and hauling vehicles, and the locations of dumping stations based on hauling vehicles' turning movements. Furthermore, a clustering method is used to calculate the number of distinct workstations of each type.

The methodology is applied in a case study at an earthwork site in Sweden. GPS data are collected from a group of construction vehicles working together, and processed using the proposed methodology to extract locations of various workstations. The experimental results indicate that the proposed method is capable to infer the most important workstations in earthwork operations.

The remainder of the paper is structured as follows. A review of related prior work is given in Section 2. Section 3 presents the probabilistic framework for the identification of different workstations. The case study is described in Section 4, with results discussed in Section 5. Section 6 concludes the paper and identifies future research directions.

## 2. Related work

A number of studies have presented methods for inferring significant locations (i.e., locations that plays a significant role in the activities of the agent) from GPS traces, focusing primarily on road traffic and pedestrian movement. One proposed method automatically clusters places where the user spends a minimum pre-defined amount of time into significant locations, and uses a Markov model to predict the agent's movements [8]. Other approaches include spatio-temporal clustering to identify stops and movements in trajectories based on the speed profile [9], and a stay extraction algorithm using pre-defined scale parameters such as a maximum distance from or a minimum duration at a location [10]. Other studies use clustering and fixed-threshold-based criteria to identify significant locations [11–13].

A drawback of methods based on fixed thresholds is that inference is sensitive to the choice of threshold parameters. Furthermore, suitable thresholds can vary significantly depending on application context and GPS data source. In practice, there is no standard threshold that leads to a satisfactory detection of all significant places. Thus, success of the approach depends heavily on the design decisions as well as the quality and frequency of the GPS data.

Machine learning methods have also been employed for the inference of significant places in various applications, mainly for the analysis of individuals' travel and activity patterns. The proposed methods include a Gaussian mixture model for inferring agent's significant locations from GPS data [14], a dynamic Bayesian network model for inferring transport modes between significant places [15], and a conditional random fields model for inferring a person's activities and significant places [16]. Based on GPS traces, the latter model first segments an agent's day into activities of work, visits or travel, and then identifies significant places such as workplace, home, or restaurant. Other approaches include a probabilistic place extraction algorithm based on the density of data points [17], where high-density regions are ranked by importance using a density scoring metric. Similarly, kernel density estimation has been applied to GPS traces to create a continuous density surface from which local maxima are retained as stop locations [18].

In contrast to the settings of the studies above, earthwork environments are highly unstructured and dynamic, and vehicles do not always follow predictable routines. Thus, it is necessary to tackle the location identification problem in these applications in a different manner [19]. Vehicle speed, for example, may provide meaningful information in the case of off-road construction environments. If a vehicle is moving very slowly in a particular area, it may indicate that the vehicle is performing an activity (loading, unloading, queuing or yielding). In the context of safety, one approach is to identify activity locations with potential interactions between vehicles and risk of accidents uses a probabilistic model based on vehicle speed profiles [19]. Examples from an open pit mine show that the algorithm detected important locations successfully. The detected locations are forwarded to a collision avoidance system for mining operations, which provides warnings and driving assistance to vehicle operators [20]. This method detects important locations from a safety point of view, but is not designed for identifying the contextual information of a construction environment. With the same reasoning, a speed-based method is presented in [21] for automatic work zone detection for trucks, and suggested using GPS data and the concept of work zones for the extraction and analysis of cycle time information. However, the proposed zone detection method is not able to identify the type of work zones and manual input is therefore required.

In contrast to [19], an important motivation for this study is to use the inferred up-to-date layout as input to a simulation-based optimization framework for optimization of various aspects of site operations, such as vehicle fleet combination and alternative operating methods [7]. For this application, it is essential to have accurate information on the locations of specific workstations rather than general significant places. In the proposed method, more complex activity characteristics than vehicles' speed profiles, such as interactions between difference vehicle types and vehicle turning movements, are employed for the inference of different types of workstations. Information regarding the movements and interactions between construction vehicles is extracted from synchronized GPS data from multiple vehicles.

## 3. Methodology

This section presents a methodology for inferring the locations of multiple types of workstations in an earthwork environment. In earthwork operations, various units of equipment interact with each other to perform various tasks. Indicators based on vehicles' driving patterns and the interactions between different vehicles extracted from GPS measurements provide valuable insights into the type of activity undertaken at a particular location. Hence, such indicators are utilized to identify distinct workstations. A general methodology is first introduced, which is then specified to infer the locations of loading and dumping stations in particular, the two most important types of workstations in earthwork operations.