



EUROPEAN PATENT APPLICATION

(43) Date of publication:
02.11.2022 Bulletin 2022/44

(51) International Patent Classification (IPC):
G08G 1/01 ^(2006.01) **G08G 1/14** ^(2006.01)
G08G 1/00 ^(2006.01)

(21) Application number: **21171719.4**

(52) Cooperative Patent Classification (CPC):
G08G 1/0112; G08G 1/0116; G08G 1/0129;
G08G 1/0133; G08G 1/142; G08G 1/147;
G08G 1/148; G08G 1/207

(22) Date of filing: **30.04.2021**

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB
GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO
PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME
Designated Validation States:
KH MA MD TN

(71) Applicant: **Bayerische Motoren Werke**
Aktiengesellschaft
80809 München (DE)

(72) Inventor: **Gomari, Syrus**
80634 München (DE)

(54) **QUALITY ASSESSMENT OF MOBILITY-RELATED PREDICTION MODELS**

(57) An apparatus (100) for assessing a quality of a prediction model (110) for predicting mobility-related information is provided. The apparatus (100) comprises a processor (120) configured to provide vehicle fleet data (130) indicative of a usage of vehicles in slices (140) corresponding to different geographical areas and/or different time periods. The processor (120) is further configured to assign an importance weight (150) to the slices (140) in accordance with a density of the usage of vehicles in the respective slice of geographical area and/or time period. The processor (120) is further configured to

compare, for at least one of the slices (140), predicted mobility-related information (162) of the prediction model (110) against predefined mobility-related information (172) in said slice, wherein the processor (120) is configured to weight said slice in accordance with its assigned importance weight (150). The processor (120) is further configured to assess the quality of the prediction model (110) based on the weighted comparison of predicted and predefined mobility-related information (162, 172).

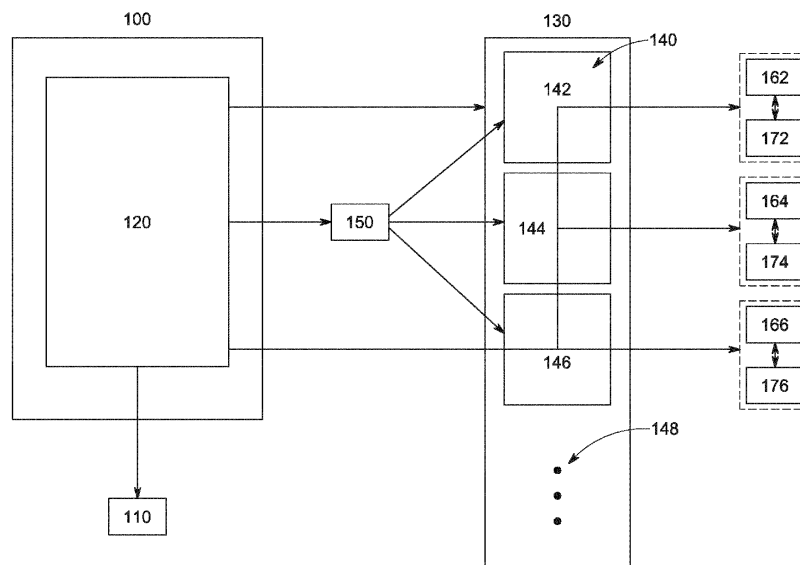


FIG. 1

Description**Field**

- 5 **[0001]** Examples relate to apparatuses and methods for assessing a quality of a prediction model for predicting mobility-related information. For instance, the prediction model may be part of an on-street parking information system of a vehicle.

Background

- 10 **[0002]** Prediction models for predicting mobility-related information, such as on-street parking information systems (OSPI), have become more popular in the original equipment manufacturer industry over the last decade. Vehicles cruising for on-street parking contribute to a significant amount of congestion within a city's inner urban area. Based on 22 studies in different cities ranging from 1927 to 2015 (R. C. Hampshire and D. Shoup, "What Share of Traffic is Cruising for Parking?," J. Transp. Econ. Policy, vol. 52, no. July, pp. 184-201, 2018.), the average cruising traffic share in a city
15 is around 34% and drivers spend around 8 minutes searching for parking. OSPI navigate drivers in search for on-street parking and reduce traffic congestion caused by cruising drivers. OSPI may also provide pre-departure information of parking availability at destination.

- [0003]** Prediction models of OSPI differ in a collection of training data and in features considered for training, validating, and testing the prediction models. Data sources for training data may be smart parking meters, real-time ground sensors, images captured by a camera mounted on a moving vehicle, crowd-sensing information from probe vehicles, e.g., taxis, with on-board sensors, cameras, or ultrasonic sensors. Differences in the training data affect the reliability and quality of the prediction model. The quality of the prediction models is conventionally validated by comparing randomly selected predefined mobility-related information, so-called ground truth data, against predicted mobility-related information.

- [0004]** Most prediction models aim to achieve a real-time prediction, some estimate parking availability for a given time interval, like 10 to 20 minutes. Further added value for users come with a capability to correctly assess the quality of the prediction model. A "true" quality of OSPI needs to be addressed. An accuracy and true quality of OSPI determine benefits gained in a transport network.

- [0005]** Conventional quality assessment of prediction models like OSPI has mainly focused on measuring an accuracy of the predictions using randomized testing, i.e., collecting ground truth randomly from any street within a specific area. This quality assessment may have a low significance as it does not consider features important to system objectives and user or management expectations.

- [0006]** On-street parking information systems (OSPI) are a special case as high efforts are required for a quality assessment in comparison to traffic prediction, for instance. The prior involves a high number of small streets where low volume of on-street parking events occurs, whereas the latter deals with low number of major roads where high volume of traffic events are measured easier. This makes quality assessment of OSPI comparably more error-prone and a high volume of so-called ground truth is needed. For instance, ground truth is a measured or observed parking availability which is used for a comparison with a corresponding prediction to assess the quality of the OSPI. An elaborate collection of ground truth increases efforts for quality assessment. Although many strategies for ground truth collection exist, there are so far no scalable method that reduces efforts and costs of ground truth collection. An approach may be acquiring local knowledge about land use and daily parking behaviour. However, this approach is not scalable. Thus, a volume of ground truth shall be kept low without affecting the quality assessment.

- [0007]** The subject matter claimed herein is not limited to embodiments that solve any disadvantages or that operate only in environments such as those described above. Rather, this background is only provided to illustrate one exemplary technology area where some embodiments described herein may be practiced.

Summary

- [0008]** The present disclosure may introduce apparatuses and methods which may lower a volume of ground truth data and increase a significance of quality assessment of prediction models for predicting mobility-related information, such as parking availability. For example, the relative importance of information about an arrival of a train with a 15-minute headway is higher than of a train with a 2-minute headway. Another instance is, information about a parking availability is more important for a driver in a busy central area compared to a parking availability in a periphery of a city with minimal traffic. This leads to a gap between an assessed quality by a service provider and a users' expected quality based on perceived utility. Ideally, the quality of a prediction model for predicting mobility-related information may be assessed based on a relative importance of the prediction to a user. A collection of ground truth data in areas or at times of low relative importance may be omitted.

- [0009]** According to a first aspect of the present disclosure, it is provided an apparatus for assessing a quality of a prediction model for predicting mobility-related information. The apparatus comprises a processor configured to provide

vehicle fleet data indicative of a usage of vehicles in slices corresponding to different geographical areas and/or different time periods. The processor is further configured to assign an importance weight to the slices in accordance with a density of the usage of vehicles in the respective slice of geographical area and/or time period. The processor is further configured to compare, for at least one of the slices, predicted mobility-related information of the prediction model against predefined mobility-related information in said slice, wherein the processor is configured to weight said slice in accordance with its assigned importance weight. The processor is further configured to assess the quality of the prediction model based on the weighted comparison of predicted and predefined mobility-related information.

[0010] According to a second aspect of the present disclosure, it is provided a computer-implemented method for assessing a quality of a prediction model for predicting mobility-related information. The method comprises providing vehicle fleet data indicative of a usage of vehicles in slices corresponding to different geographical areas and/or different time periods. The method further comprises assigning an importance weight to the slices in accordance with a density of the usage of vehicles in the respective slice of geographical area and/or time. The method further comprises comparing, for at least one of the slices, predicted mobility-related information of the prediction model against predefined mobility-related information in said slice, wherein said slice is weighted in accordance with its assigned importance weight. The method further comprises assessing the quality of the prediction model based on the weighted comparison of predicted and predefined mobility-related information.

Brief description of the Figures

[0011] Some examples of apparatuses and/or methods will be described in the following by way of example only, and with reference to the accompanying figures, in which

Fig. 1 illustrates an apparatus for assessing a quality of a prediction model for pre-dicting mobility-related information;

Fig. 2 illustrates a method for assessing the quality of the prediction model for pre-dicting mobility-related information, with optional benefits analysis;

Fig. 3 illustrates a temporal distribution of parking events of example vehicle fleet data;

Fig. 4 illustrates a spatial distribution of the parking events of the example vehicle fleet data;

Fig. 5 illustrates a distribution of the parking events over specific areas and time periods;

Fig. 6 illustrates a spatial distribution of quadkeys and their encoded labels;

Fig. 7 illustrates a spatial distribution of clusters.

Detailed Description

[0012] Some examples are now described in more detail with reference to the enclosed figures. However, other possible examples are not limited to the features of these embodiments described in detail. Other examples may include modifications of the features as well as equivalents and alternatives to the features. Furthermore, the terminology used herein to describe certain examples should not be restrictive of further possible examples.

[0013] Throughout the description of the figures same or similar reference numerals refer to same or similar elements and/or features, which may be identical or implemented in a modified form while providing the same or a similar function. The thickness of lines, layers and/or areas in the figures may also be exaggerated for clarification.

[0014] When two elements A and B are combined using an 'or', this is to be understood as disclosing all possible combinations, i.e., only A, only B as well as A and B, unless expressly defined otherwise in the individual case. As an alternative wording for the same combinations, "at least one of A and B" or "A and/or B" may be used. This applies equivalently to combinations of more than two elements.

[0015] If a singular form, such as "a", "an" and "the" is used and the use of only a single element is not defined as mandatory either explicitly or implicitly, further examples may also use several elements to implement the same function. If a function is described below as implemented using multiple elements, further examples may implement the same function using a single element or a single processing entity. It is further understood that the terms "include", "including", "comprise" and/or "comprising", when used, describe the presence of the specified features, integers, steps, operations, processes, elements, components and/or a group thereof, but do not exclude the presence or addition of one or more other features, integers, steps, operations, processes, elements, components and/or a group thereof.

[0016] This document describes systems and techniques for assessing a quality of a prediction model for predicting

mobility-related information. The assessment may therefore reflect a relative importance of the mobility-related information to a user of the prediction model. The subsequent description may mainly refer to predicting parking availability, e.g., of on-street parking information systems (OSPI). This may be due to OSPI being a special case where an assessment of predictions needs a high volume of ground truth data, thus, it may lead to high efforts in collecting comparative data for the assessment. The present disclosure may seek an automated and scalable method to reduce efforts of ground truth collection. Persons skilled in the relevant art will recognize that the present disclosure may also be applicable to other systems using prediction models in a mobility-related context, such as traffic prediction, prediction of public transport delay, or alike.

[0017] The said prediction model for predicting mobility-related information may be a computer-implemented program which may output the mobility-related information upon a request of a user of the prediction model. The user may be any entity that could benefit from the prediction model, for example, a driver or passenger of a vehicle or a provider of mobility-related services. The prediction model may be integrated into a board computer of a vehicle and may be part of an information system of the vehicle. The prediction model may alternatively be integrated into an external computing system and exchange data with the board computer or other user devices comprising processing circuitry to forward a user request and receive the mobility-related information. The mobility-related information may be information relevant to the user for a usage of the vehicle. For example, the mobility-related information may comprise probability values of a parking availability in the case of an OSPI. The prediction model may be part of a navigation system. It may predict a parking availability in a certain area, like a street near a destination defined by the user, and at a certain time, like an hour of the day as defined by the user. Alternatively, the prediction model may be used for predicting a traffic status of a vehicle infrastructure. The prediction may comprise probability values for an availability of a vacant parking spot in the area and time defined by the user or the navigation system. The prediction model may be developed using complex machine learning techniques and may have previously been trained with appropriate training data. The processing circuitry integrating the prediction model may have an interface to a navigation system which may further process the prediction and navigate the user according to the prediction via, for instance, a graphical user interface in the vehicle.

[0018] Quality assessments of said prediction models may usually be made by a service provider of the prediction models. Quality assessments may be necessary to improve a prediction model, to advertise a quality of the prediction model, or to choose an appropriate prediction model among a plurality of prediction models. Conventional quality assessment of prediction model may be based on randomly collecting ground truth data within a certain area, like a city. The ground truth data may be observed or measured vacant parking spots collected in different areas within the city, like in different streets, and at different dates, like different hours of several days. For example, the ground truth data used for a BMW study may comprise 20000 random observations throughout the city of Munich, Germany. The ground truth data may have been collected between June 2018 and October 2020. Each observation may be made for a block at a certain time. The block may be a stretch of a street measured from one intersection to the other. For each observation, the area of the block and the time of the observation may be recorded. When at least one vacant legal parking spot is observed on the block, a parking availability value for the corresponding block and time may be set to 1, otherwise it may be set to 0. Regardless of the number of open spots, the parking availability may be recorded as a binary outcome - available, 1, or not available, 0.

[0019] Conventional quality assessment may comprise requesting predictions of the parking availability from the prediction model. The parking availability may be requested for the same areas and times like those of the ground truth data. Conventional quality assessment may then compare the predictions against the ground truth data for each area and time. It may determine an average difference between the predictions and the ground truth data over all areas and times. The average difference may be considered a loss of the prediction.

[0020] However, conventional quality assessment may not consider a relative importance of the areas and times to a potential user. For instance, a certain area may be located within a center of a city where vacant parking spots are rare and frequently demanded by drivers. Likewise, a certain time, like an afternoon hour during a working day, may be critical in terms of parking availability. So, such an area or time may be seen as relatively important to users of the prediction model. Another area may be located in outskirts of a city where vacant parking spots are much easier to find and seldomly requested by a potential user. Another time may be a night hour on Sunday when vacant parking spots are usually non-critical. So, the latter mentioned area and time may be seen as relatively unimportant to users of the prediction model. As the relative importance is not considered in conventional quality assessment, a significance of the quality assessment may be diminished.

[0021] Thus, it may be seen as an objective of the present disclosure to increase the significance of the quality assessment of the prediction model.

[0022] Fig. 1 illustrates an apparatus 100 for assessing a quality of a prediction model 110 for predicting mobility-related information according to some embodiments of the present disclosure. The apparatus 100 comprises a processor 120 configured to provide vehicle fleet data 130 indicative of a usage of vehicles in slices 140, such as 142, 144, 146 and others as indicated by the dots 148, corresponding to different geographical areas and/or different time periods.

[0023] The vehicle fleet data 130 may comprise timestamps and/or geographical coordinates of parking events or

traffic events in the slices 140. Alternatively, the vehicle fleet data 130 may indicate other usage of vehicles in slices 140. In the subsequent, OSPI is used as non-limiting example. In the case of OSPI, the vehicle fleet data 130 may comprise parking events. Geographical coordinates of the parking events may lie within a certain area, like a city. For instance, vehicle fleet data 130 may be retrieved from a vehicle fleet of a plurality of vehicles equipped with GPS sensors, real-time clocks or timers, controllers with an interface to backends of the vehicles' board computers, and storage devices. The vehicles may usually move within the certain area. For instance, BMW's OSPI for Munich, Germany may be examined. The vehicle fleet data 130 may be gathered from a vehicle fleet of BMW vehicles. A data collection of the vehicle fleet data 130 may be performed at BMW's backend services which may include anonymization according to EU defined data privacy standards. A parking event of a vehicle may be generated when a vehicle's engine is switched off or on, triggering a parked-in event or parked-out event, respectively. Switching off or on the engine may be tracked by the backend of the vehicle's board computer. The backend may trigger the controller via the interface to the backend to store current geographical coordinates of the GPS sensor, a current timestamp generated by the real-time clock, and a type of the parking event (parked-in or parked-out) on the storage device. After a predefined study time, data of the storage devices of all vehicles of the vehicle fleet may be collected and read out by a computing system, such as the processor 120. For the BMW study, data from February 2020 to September 2020 may have been taken. The collected data may be considered as vehicle fleet data 130. The computing system may post process the vehicle fleet data 130 to contain only parking events within a proximity of a street. So, parking events triggered by parking in a garage or car park may be excluded.

[0024] Referring back to **fig. 1**, the processor 120 is further configured to assign an importance weight 150 to the slices 140 in accordance with a density of the usage of vehicles in the respective slice 140 of geographical area and/or time period.

[0025] As mentioned above, the vehicle fleet data 130 may be used for determining a relative importance of areas and times to the user and using the relative importance for a quality assessment of a prediction model 110. For this purpose, the slices 140 within an examined area and/or time for which the vehicle fleet data 130 is collected, like the city of Munich and the year 2020, may be defined. The slices 140 may correspond to different geographical areas and/or different time periods. For example, firstly, a map of the examined area may be portioned into map tiles which may be square shaped and equally sized spatial divisions of the examined area. The map tiles may be labelled with numbers. Secondly, an optional temporal dimension of the slices 140 may be defined as 168 week-hours being one-hour divisions of a week. Thirdly, the slices 140 may be defined as spatial-temporal divisions of the examined area and time. The slices 140 may result in an array of the map tiles over the 168 week-hours. The parking events of the vehicle fleet data 130 may be assigned to the slices 140 according to the geographical coordinates and timestamps of the parking events. The relative importance may be defined by the percent volume weight or density of the parking events that occur within the slices 140. An importance weight 150 may be assigned to each slice 142, 144, 146, 148 according to the relative importance. For instance, the importance weight 150 for the slices 140 may be calculated as follows:

$$w_s = \frac{PEVolume_s}{\sum_{s=1}^N PEVolume_s},$$

where w_s may be the importance weight 150 for a slice with a number s . The slices 140 may in this case be numbered consecutively from 1 on. With the number s , a specific slice may be selected. $PEVolume_s$ may be a volume of the parking events in the slice with the number s . N may be a total number of slices 140. The skilled person in the art will recognize that the slices 140 may be defined as any other geographical areas, e.g., clusters of map tiles or area divisions of other shapes, and/or time periods, e.g., 2-hour intervals of a year or irregular time sections, that are appropriate for the application. The spatial dimension of the slices 140 may not be combined with a temporal dimension or the other way round. More examples of slices 140 will be given below. Furthermore, the slices 140 may be of any number other than demonstrated in the examples.

[0026] A portioning of areas and times for defining slices may be referred to as prioritization-based subsampling strategy PSS. In other words, different PSS may look at vehicle fleet data from several perspectives. A PSS may be chosen by robustness of an assessment design. A first PSS may be purely based on spatial slices, referred to as neighbourhoods. The first PSS may only consider a density of vehicle fleet data in each neighbourhood within a city over an entire study period. The first PSS may be based on a quadkey concept (Microsoft, "Zoom levels and tile grid," 2020, <https://docs.microsoft.com/en-us/azure/azure-maps/zoom-levels-and-tile-grid?tabs=csharp#quadkey-indices%5C.>). The quadkey concept may be an indexing convention and unique identifier of a standard map tile at a specific zoom level. The quadkey concept may be a method for partitioning a map into map tiles. The quadkey concept may be a standard used by Microsoft's Azure Maps. The zoom level of quadkeys varies from 0 to 24, corresponding to a map tile size of 40,075,017 m x 40,075,017 m to 2.39 m x 2.39 m, respectively. The finer a map tile level the lower may be a volume of the vehicle fleet data per map tile, and thereby the higher may be a relative error. The quadkey concept may be favorable to generate reproducible and comparable results for similar applications. Each map tile may equate to a slice; the densest map tile, or the map tile with the highest number of parking events, may then be considered most

important. In this manner, the map tiles may be sorted from most to least important.

[0027] A second PSS may define slices 140 as 168 week-hours. A time interval of one hour may be due to heuristics as it is not too small, and not too large, while maintaining interval consistency. In the mentioned BMW study half-hour slices were also experimented with, but with negligible differences in overall scores, hence, omitted from further analysis. A busiest week-hour may be the densest slice and, thereby, the most important. Typically, morning and afternoon peak hours may have the highest densities and after midnight hours may be the quietest.

[0028] A third PSS may be a combination of the first and the second PSS. Each neighbourhood may be divided into 168 week-hours, resulting in a number of slices 140 equal to a number of neighborhoods times 168. The first and the second PSS may be on a higher aggregated level, while the third PSS may create lower aggregated priority, in other words, result in more slices 140. The third PSS may be a generic strategy that can be used in any city use case. It may divide a study area spatially based on the quadkey concept combined with the week-hour basis. This may allow for a precise identification of important areas. For instance, it may allow for pinpointing neighbourhoods that are more important at specific hours during a week. The slices 140 may be sorted according to a density of vehicle fleet data 130 in the slices 140, more concretely, the slices 140 may be sorted by the number of parking events having occurred in the respective slices. Since third PSS may be both across neighbourhood and time, the sequence of slices 140 ordered by importance may discontinuously step from one neighbourhood and week-hour to another. For example, a topmost important slice may be neighbourhood A on Monday 13:00-14:00, while a second most important may be from neighbourhood B on Monday 8:00-9:00.

[0029] A fourth PSS may be based on neighbourhood clusters and time. Neighbourhood clusters may be generated based on vehicle fleet data within the neighbourhoods of an examined area, like a city. Neighbourhoods that have a similar spread of vehicle fleet data 130 over time may be grouped together and be treated as one entity. The clustering may be done by defining a behaviour of each neighbourhood through an aggregation method of the vehicle fleet data 130. And then, the clustering may be performed on a corresponding behavioural pattern. For the case of OSPI, the neighbourhood clusters may be based on a temporal trend of parking dynamics (TTPD) inferred from parking events. TTPD may be a week-hour time-series of a cumulative sum of a difference of week-hour normalized average parked-in and parked-out events per 30-minute intervals at quadkey zoom level 14. For OSPI, zoom level 14 may be selected as an optimum since a more localized level may generate high relative errors given that a volume of parking events within 30-minute intervals was small. Each neighbourhood at zoom level 14 may have a particular normalized TTPD. The TTPDs may be used as a base for clustering similar neighbourhoods. Each neighbourhood cluster may consist of multiple neighbourhoods and may be spatially treated together. The logic behind this approach may be that different neighbourhoods with similar parking behaviour may be analyzed on the same level and therefore be combined in a neighbourhood cluster. As a next step, the neighbourhood clusters may then be divided into 168 week-hours to form slices 140 as previously. The slices 140 may then be sorted according to a density of the vehicle fleet data 130.

[0030] Referring back to **fig.1**, the processor 120 is further configured to compare, for at least one of the slices 140, for example, slice 142, 144, 146, predicted mobility-related information 162, 164, 166 of the prediction model 110 against predefined mobility-related information 172, 174, 176 in said slice 142, 144, or 146, respectively. The predicted mobility-related information 162, 164, 166 may be probability values of a parking availability in the respective slices 142, 144, 146; predefined mobility-related information 172, 174, 176 may be measured or observed ground truth data of a parking availability in the respective slices 142, 144, 146. The processor is configured to weight said slice 142, 144, or 146 in accordance with its assigned importance weight 150.

[0031] So, once an appropriate PSS is identified, the slices 140 for the PSS may be generated. The slices 140 may then be used for subsampling a collected ground truth data. In other words: the ground truth data may be reduced by excluding ground truth data of the least important slices 140. Likewise, the most important slices 140 may be prioritized in a quality assessment of the prediction model 110. For instance, the processor 120 may be configured to select a certain percentage of the slices 140 with the lowest importance weight 150 and set the respective importance weight 150 to zero. The weighted comparison performed by the processor 120 may be considered a key performance indicator (KPI) which may serve as a quality metric for assessing the prediction model 110.

[0032] The KPI may ensure that prioritizations are consistent over the slices 140 and may measure a "true" quality

correctly. A popular KPI may be the Brier Score, as described in the following equation:

$$KPI = \frac{1}{N} \sum_{t=1}^N (p_t - o_t)^2$$
 where p_t may be the predicted mobility-related information 162, 164, or 166 for an instance t , e.g., for a certain slice; o_t may be the predefined mobility-related information 172, 174, or 176 for the instance t (o_t may be 0 if there was no occurrence, 1 if there was an occurrence); and N may be a total number of instances. The KPI may be calculated for each slice 140 of a PSS. A total KPI for a PSS may be calculated based on an evidence-based multi-criteria decision making method called weighted sum model WSM as described in the following equation:

$KPI_{total} = \sum_{s=1}^N KPI_s \times w_s$, where KPI_s may be the KPI of the slice s , w_s may be the importance weight

assigned to the slice s according to the equation above and N may be a total number of slices 140. WSM may be favorable due to its objectivity and not being prone to score skewness. The total KPI may be dependent on which prediction model 110 and which PSS is used. The KPI may be a loss function of the predictions. Hence, a low value of KPI may indicate a good prediction, in other words, it may indicate a prediction is close to corresponding ground truth data.

[0033] Referring back to **fig. 1**, the processor 120 is further configured to assess the quality of the prediction model 110 based on the weighted comparison of predicted mobility-related information 162, 164, 166 and predefined mobility-related information 172, 174, 176. The weighted comparison may be the total KPI of the prediction model 110 for a specific PSS.

[0034] The apparatus 100 may use the assessed quality for selecting a prediction model among a plurality of prediction models for a user-selected slice. For instance, a driver of a vehicle may use an OSPI integrated into a navigation system or other information system of a board computer in the vehicle. The driver may request via a user interface of the board computer an information about a parking availability in a destination area at a destination time. The destination area and destination time may be assigned to a slice of a PSS, being designated as the user-selected slice. The apparatus 100 may compare the assessed quality of the plurality of prediction models and select the prediction model with the best assessed quality, e.g., the lowest total KPI, for the user-selected slice. The selected prediction model may be used for generating an estimated parking availability for the destination area and destination time. The processing circuitry of the board computer may answer the request of the driver by displaying the estimated parking availability on a display in the vehicle. The information system of the vehicle may use the estimated parking availability to estimate a time for searching a parking spot or to propose streets near the destination area for parking. A navigation system may navigate the driver to the proposed streets or consider the time for searching a parking spot for a proposed departure time. The apparatus 100 may increase a reliability of such a navigation system.

[0035] Additionally or alternatively, the apparatus 100 may have an interface to a computing system generating the predictions of the prediction model 110. The apparatus 100 may send via the interface the assessed quality of the prediction model 110 to the computing system. The computing system may use the assessed quality to adapt the prediction model 110, for instance, to adjust computational weights of a machine-learning model used for the prediction model 110. The apparatus 100 may increase an accuracy of the predictions of the prediction model 110.

[0036] Aside from importance weighting based on a percent volume share of vehicle fleet data in each slice, another weighting technique may be considered in this document: equal weighting for all slices, which may be computed by one divided by total number of slices 140. This may be done to validate importance weighting and to see whether importance weighting may shift a penalty or incentive of the assessment to important slices.

[0037] Randomized subsampling may be the norm to reduce any biases in subsampling. In contrast, the present disclosure may introduce a prioritization-based subsampling based on PSS as a competing method to the conventional randomized subsampling. Prioritization-based subsampling may enable an assessment of a "true" quality of prediction models as it prioritizes slices 140 based on a relative importance to a user.

[0038] **Fig. 2** illustrates a flow-chart of a method 200 for assessing a quality of a prediction model 110 for predicting mobility-related information according to some embodiments of the present disclosure. The method 200 may optionally include a validation of the above-mentioned prioritization-based subsampling. The method 200 may comprise providing 210 vehicle fleet data 130 indicative of a usage of vehicles in slices 140 corresponding to different geographical areas and/or different time periods. Providing 210 the vehicle fleet data 130 may comprise acquiring vehicle fleet data 130 from an external source, like an external server or other computing system. Providing 210 the vehicle fleet data 130 may also comprise processing the vehicle fleet data 130, e.g., sorting the vehicle fleet data 130 by the slices 140. Processing the vehicle fleet data 130 may comprise identifying relevant information in the vehicle fleet data 130 for the usage of vehicles, e.g., identifying parking events in the case of OSPI or traffic events in the case of traffic prediction. Providing 210 the vehicle data 130 may further comprise determining an appropriate PSS, e.g., the first, second, third, or fourth PSS as explained above. The PSS may define slices 140. The method 200 may further comprise assigning 220 an importance weight 150 to the slices 140 in accordance with a density of the usage of vehicles in the respective slice of geographical area and/or time. Assigning 220 may comprise generating priority slices as a subset of the slices 140 with the most important slices according to the importance weight 150. The method 200 may further comprise comparing 230, for at least one of the slices 140, predicted mobility-related information 162, 164, 164 of the prediction model 110 against predefined mobility-related information 172, 174, 176 in said slice, wherein said slice is weighted in accordance with its assigned importance weight 150. Comparing 230 may comprise acquiring the predicted mobility-related information 162, 164, 164 of the prediction model 110 and the predefined mobility-related information 172, 174, 176. Comparing 230 may comprise calculating a Brier Loss, e.g., for determining a KPI, or using another loss function for each priority slice. Comparing 230 may comprise determining a total KPI as weighted product sum of KPI over all priority slices of the PSS. The method 200 may further comprise assessing 240 the quality of the prediction model 110

based on the weighted comparison of predicted mobility-related information 162, 164, 164 and predefined mobility-related information 172, 174, 176. In other examples, there may be more or less than the presented predicted mobility-related information 162, 164, 164 and predefined mobility-related information 172, 174, 176.

[0039] Optionally, the method 200 may comprise validating 250 the assessed quality by performing a randomized comparison of predicted mobility-related information 162, 164, 164 and predefined mobility-related information. The randomized comparison may be a baseline for validating the weighted comparison of a prioritization-based subsampling, for instance. Validating 250 may also comprise comparing the total KPI of different PSS for selecting an appropriate PSS for the prediction model 110. Validating 250 may be useful for a benefits analysis of the prioritization-based subsampling against conventional randomized subsampling. Optionally, the method 200 may comprise reducing 260 a collection of predefined mobility-related information 162, 164, 164. As priority slices are determined, a collection of ground truth data in non-priority slices may be omitted for further assessments of the prediction model 110.

[0040] The method 200 may further comprise assessing (not shown) the quality of a plurality of prediction models based on a weighted comparison of respective predicted and the predefined mobility-related information. For example, the processor 120 may run through the above-mentioned steps for each of a plurality of the prediction models. The method 200 may further comprise comparing the assessed quality of the plurality of prediction models for determining a suitable prediction model for at least one of the slices. For example, the processor 120 may determine the lowest loss of the predicted mobility-related information for the at least one of the slices among respective losses of the prediction models. The method 200 may further comprise using the determined suitable prediction model to predict the mobility-related information in the slice. For example, the processor 120 may select the suitable prediction model with the lowest loss for the slice.

[0041] In the case of OSPI, the processor 120 may select the suitable prediction model with the highest chances to predict correctly a parking availability in a certain slice. A user of a car may use an application on a board computer of his or her car or an application on another user device, like a smartphone. In the application, like a navigation or mapping application, the user may select a certain destination and/or time. The board computer or user device may have an interface to the apparatus 100 and convey the selection of the user to processing circuitry of the apparatus, like the processor 120. According to the selection of the user, the processor 120 may define a slice following a PSS. The slice may correspond to an area around the selected destination and/or to a time interval around the selected time which is considered as potential parking area and/or parking time. Alternatively, the user may indicate in the application that he or she looks for a parking spot and automatically select a slice according to a location of the user (as indicated by a GPS sensor) and a time (as indicated by a clock). The selected suitable prediction model may then be used to predict the parking availability of the slice. The processor 120 may convey the predicted parking availability, for example in form of probability values, via the interface to the board computer or user device where the application uses the prediction to answer a request of the user, such as to navigate the user to a street within the slice with high probability to have vacant parking spots or to display on a graphical user interface in the car or on the user device the probability for a vacant parking spot in the slice. The application may give, based on the prediction, a recommendation to park in a different area than selected by the user or to depart to the destination to a later point of time. By using the suitable prediction model selected by the apparatus 100, the application may be able to be more precise in its indications to the user. The user may benefit from better parking predictions which may help the user to choose a time for departure that decreases the cruising time for parking. The predictions may help the user to make better decisions in his or her mobility behavior, for example, the user may prefer to do without the car and use public transport when faced with a low possibility to find a parking spot at time.

[0042] On a global scale, the suitable prediction model selected by the apparatus 100 may reduce traffic in an area where the application is used. In particular, traffic in busy times and areas may be equalized as those busy times and areas would be considered as important in the slices and would shift the incentive of prediction models to those busy times and areas, in other words: parking predictions would be very precise for those busy times and areas. Users of the application would precisely see their chances to get a parking spot in those busy times and areas. Thus, they may plan their car trip differently to avoid long cruising for parking. A navigation system may use the predictions to smartly navigate the users to quieter areas near a destination of the users.

[0043] An experimental design may be defined to test the prioritization-based subsampling against thousands of random subsamples of the randomized subsampling. The experimental design may be defined to test the chances of falsely assess the quality of the prediction models 110. After the total KPI are calculated for all considered PSS, the next step may be to check the benefits of "true" quality assessment. This may be done by comparing the total KPI against the total KPI of a baseline, which is equal weighting of slices 140 and randomized subsampling of ground truth data. The experimental design for randomized subsampling of ground truth data may be necessary to assess and ensure the robustness of the prioritization-based subsampling. One objective may be to ensure that if any of the PSSs are followed for a collection of ground truth data, the followed PSSs are likely to be representative of the "true" quality of the prediction system 110. The goal of randomized subsampling may be to generate different random slices independently from vehicle fleet data density. The ideal, however, unrealistic randomized subsampling may be adjusted to give the best results for

assessing the quality of the prediction model 110. This may be useful to provide a baseline for the benefits analysis of the PSS implemented. Validating 250 may aim to identify weakly designed prediction models 110 that only perform well in rare instances. The experimental design may ensure that the randomized subsamples cover most possible combinations of subsets of slices 140. The comparison of priority slices with random slices may be done for comparing effects of a reduction of ground truth data. This also may provide an opportunity to check the benefits of the PSS with a smaller volume of ground truth data, which may result in higher relative error. It is noted that the top importance weights may correspond to the vehicle fleet data percent share that is attributed to a slice, and therefore may not correspond to a volume of the ground truth data in said slice. For instance, within the top 50% most important slices, it may be possible to only have a sample size of 30% of ground truth observations occurring in the top 50% most important slices. In summary, exemplary steps taken for validating 250 the assessed quality are shown in the following:

1. Sorting the slices 140 of each PSS based on corresponding importance weights 150.
2. Selecting top 30% up to top 90% most important slices, at 5% interval steps and calculating the total KPI for all PSS.
3. Getting a share of ground truth data selected randomly for randomized subsampling, the share of ground truth data may be equivalent to a share of ground truth data resulting from step 2.
4. Running n-number of random trials covering different subsets of the ground truth data and calculating the KPI for all trials.
5. Calculating a variance of the KPI of an m-number of PSS.
6. Calculating a variance of a KPI of the n-number of random trials.
7. Using an interquartile range (IQR) method of outlier detection for robustness of KPI scores. The IQR may be calculated as follows: $IQR = Q3 - Q1$, where Q1 may be a third quartile value and Q1 may be a first quartile value. A lower bound and upper bound outlier may be detected with the following inequations: $Lower\ bound\ outlier > Q1 - 1.5 \times IQR$ and $Upper\ bound\ outlier > Q3 + 1.5 \times IQR$.
8. Comparing the KPI variance for random trials with the KPI variance for the PSS.
9. Assessing the robustness of the PSS.
10. Analyzing if it is feasible to reduce a collected ground truth data to the most important slices and if an assessment based on the reduced ground truth data is representative of the "true" quality.

[0044] As an example, the above-described steps may be performed for a use case of OSPI. Various parking prediction models may be utilized to generate parking availability predictions. However, a development of the prediction models may not be of essence in this document and may be considered as exemplary predictions models to generate adequate predictions for quality comparison between the prediction models. Some real feature-based prediction models and random parking prediction models may be used for the use case as described below. The code carrying out the benefits analysis may be written in the programming language Python. The main packages used may be: Pandas, GeoPandas, Folium, Numpy, OSMnx, Matplotlib, Sea-born, Statsmodel, PySal, and Scikit-learn.

[0045] The experimental design setup of the PSSs implemented for the use case are shown in the following table:

PSS setup #	Neighbourhood zoom level				TTPD Clusters level	Time interval
	14	15	16	17	14	
1	X					
2		X				
3			X			
4				X		
5	X					X
6		X				X
7			X			X
8				X		X
9						X
10					X	X

[0046] The above-mentioned BMW's vehicle fleet data may be used for the examined use case. It may be observed that the parking events from Mondays to Friday evening have a similar temporal distribution with small day to day

discrepancies, as illustrated in fig. 3. Hence, those parking events may be grouped together. During a working weekday there may be peaks in the morning and afternoon, as expected since the study area of Munich is mainly commercial. On weekends, a peak may occur at around noon during lunch hours and shopping directly before or after lunch hours.

[0047] For the examined use case, the volume of parking events of BMW vehicles in the vehicle fleet data may be an indicator for relative importance of a corresponding slice. Only parking event pairs (parked-in, parked-out) with a duration of more than 5 minutes between parking in and out may be considered to eliminate noise generated by standing by cars. Hundreds of thousands of parking events that were recorded in Munich during the indicated collection period may show a spatial-temporal importance of a slice.

[0048] For the first PSS which is based on neighbourhoods, a total volume of parking events in each quadkey may be considered as importance weight. The highest and lowest quadkey zoom levels considered as neighbourhood may be level 14 (2457.6m x 2457.6m) and 17 (250m x 250m), respectively. These quadkey zoom levels may be heuristically determined as assumption of a cruising distance range for on-street parking search. A spread of the parking events may be mainly focused on city hubs within a polygon, as illustrated in fig. 4. This may correspond to areas to focus on for the KPI calculation.

[0049] The second PSS which is based on time periods, may show that most parking events occur in early mornings during weekdays and at noon during weekends. It may be observed that on a global level, the importance weights in the second PSS are not well distinguishable since the importance weights are similar during the day hence making it difficult to prioritize. This may confirm a nature of the study area as being mainly commercial and business centered. With prioritization only based on temporal slices, a small trend shift of a collection of ground truth data may be done by selecting the following top prioritized hours: period 7:00-15:00 during weekdays, 9:00-14:00 on Saturdays. Sundays can essentially be left out, as it may not be as busy as weekdays.

[0050] The third PSS based on neighbourhoods and time periods applied to the vehicle fleet data may provide detailed prioritized subsamples of slices, as illustrated in the distribution of importance weights in fig. 5 for a zoom level of 14. The third PSS may be performed for zoom levels 14 to 17, only level 14 may be discussed hereafter as example. The study area may be divided into quadkeys which have encoded labels, as illustrated by fig. 6. The neighbourhoods at quadkeys 6 and 8 may have the highest hourly importance. The neighbourhood 8, which is located around the central station of Munich, may have the highest share of parking events. Within a duration of 6:00-18:00, most neighbourhoods may have stable hourly importance. In neighbourhood 14, a slight increase in importance may be observed on Saturday afternoon; this may be traced to neighbourhood 14 including lots of shopping and dining facilities. Neighbourhoods 0, 4, 10, and 8 may be located at a periphery of the study area, hence, considered as less important. As an example, slices within the top 50th percentile of importance weights may be examined. It is noted that the importance weights may not be normalized. The top 50th percentile may have prioritized 539 slices out of 3671. And instead of 23 neighbourhoods, such a reduction may reduce the number of neighbourhoods to 10. For top 10%, 20%, 30%, or 40%, respectively, 76, 167, 276, or 398 slices may be considered. In this manner, the third PSS may provide narrowed down slices that need to be prioritized for quick quality assessments.

[0051] The fourth PSS based on clusters of neighbourhoods and time periods may build on the third PSS by aggregating similar neighbourhoods. The logic behind neighbourhood clustering may be grouping neighbourhoods based on similar temporal trend of parking dynamics (TTPD). Applying the fourth PSS on the use case of OSPI in Munich may generate 7 neighbourhood clusters, where 2 of them occurring at peripheries may have negligible importance. So, 5 valid clusters may remain. 5 clusters in the study area may be sufficient since neighbourhoods within central Munich are quite similar according to the BMW's vehicle fleet data. The fourth PSS may be applied on zoom level 14 as this is considered an optimal size for modelling TTPD in 15-minute intervals. The clusters may have encoded labels, as illustrated by fig. 7. Cluster 1 may contain most areas in Munich city center and may be considered important in almost all week-hours between 6:00-18:00, with lesser importance on Sundays. For the same period, Cluster 2 has the same stable hourly distribution but with lesser magnitude in the importance weights. For cluster 6, the importance weights may be lower in the morning hours and higher in the late afternoon and evening hours, and then fade shortly after the evening. Clusters 0 and 4 may include neighbourhoods in the periphery, where the importance weights are lower in magnitude, but uniform during the week. The benefit of the fourth PSS may be that instead of being limited to certain neighbourhoods like in the third PSS, similar neighbourhoods can be grouped together.

[0052] The priority slices selected for the respective PSS may now be used as input for quality assessment of parking prediction models. A Brier Loss Score may be used as KPI. The focus of this use case may lie mainly on assessing the quality of various prediction models and not on modelling improvement or development. Only the predictions as output of the prediction models may be used for comparison of the prediction models. Twelve prediction models may have been used as examples for testing the quality assessment. For the calculation of the KPIs, two weighting techniques may be applied: equally weighted slices and importance weighted slices. The following table may display the algorithms and the KPI for the twelve prediction models, as well as an average of equally and importance weighted KPI scores:

	Model #	Algorithm	KPI	Average KPI	
				Equally	Importance
5	1	Xgboost	0.249	0.249	0.249
	2	Random Forest	0.303	0.306	0.307
	3	Xgboost	0.227	0.224	0.229
	4	Random Forest	0.236	0.233	0.238
	5	Xgboost	0.228	0.226	0.231
10	6	Random Forest	0.231	0.231	0.235
	7	Xgboost	0.233	0.232	0.232
	8	Random Forest	0.248	0.247	0.248
	9	Random	0.332	0.334	0.335
	10	Optimistic Random	0.273	0.267	0.273
15	11	Pessimistic Random	0.486	0.493	0.487
	12	Single Optimum Value	0.226	0.224	0.227

[0053] Models 1 to 8 may use actual on-street parking related features, while 9 to 11 are random models. Model 12 may essentially be an unrealistic random guesser that only has a single optimum prediction value determined based on an expected parking availability from the ground truth data. Nonetheless, model 12 may be useful as a baseline reference for a benefits analysis of the PSS. The best randomized subsampling models in this case may be model 3, 5, 7, and 12, whereas the worst by large may be model 11.

[0054] The KPI scores may be calculated for the PSS and their importance weightings. All feature-based models may have on average a slightly worse importance weighted KPI, but not significantly different than the equally weighted and normal KPI. It may be observed that the importance weights do not shift the KPI scores significantly compared to the equally weighted scores. The KPI scores may be on average -1.06% worse compared to importance weighting over all prediction models, while -1.07% over all PSSs. The first PSSs may have the largest negative relative difference between the equal and importance weighting. The third PSS, starting at zoom level 14 and 168 week-hours, may have incurred a positive effect but as the zoom level increase, there may be a gradual decrease in KPI scores. For the second and fourth PSS, the importance weighting improved the KPI scores compared to the equal weighting. This may imply that, temporally, the prediction models are assessed to perform better than the measured normal KPI, while considering spatial importance, the KPI may punish the prediction models' scoring. Essentially, the temporal and spatial aspects of the PSSs may create a push and pull effect in the KPI scoring, thus, a difference between equally weighting and importance weighting cannot be clearly distinguished. Since from multiple angles, it may have been confirmed that the KPI scores may not significantly show a difference between equal and importance weighting, further detailed comparison between the two weightings is not necessary and only the importance weighting may be used hereafter.

[0055] Having calculated the KPI scores for the considered PSSs and weighting techniques, the next step may be to validate the quality assessment. This may be done by proving that the quality assessment using PSSs with priority slices may give better insights about on-street parking prediction models compared to randomized subsampling of ground truth.

[0056] The benefits analysis may be done by comparing the KPI scores of the topmost important slices (prioritization-based subsampling) against the KPI scores for the baseline case of non-prioritized randomized subsampling (NPRS) of ground truth data. The NPRS may be done for the same slices as the ones from the PSSs, but the importance weight was not considered, hence they are non-prioritized. In the case of prioritization-based subsampling, the size of ground truth data may be reduced by sorting the importance weights of the PSS slices and then taking a certain top fraction percentage of the slices. For example, using the prioritization-based subsampling of ground truth data considering top 90% most important slices of PSS setup 6, the ground truth data size may be reduced to 3563 observations (30% decrease) out of 5152 observations. However, if reduction of the ground truth data is done randomly, 90% of the ground truth data may be 4637 observations. There may be two reasons for the large reduction: (1) slices may only be generated in areas and time frames that have recorded a parking event, hence, the ground truth data outside these slices may automatically be disregarded as less important, in the case of the example, only 4838 observations (6% decrease) exist for PSS setup 6; and (2) there may be a disproportionate distribution of the ground truth data throughout the city since the observations were conducted randomly. Based on the performed reduction, a substantial amount of the collected ground truth data may be outside important areas.

[0057] Further prioritization-based subsampling may be performed at fractions ranging from 30% - 90% at 10% intervals as a preliminary heuristic step. For the main analysis here the 50% top fraction may be considered as an example. The same experimental design was setup for the NPRS. For the NPRS, at each fraction, 1000 random subsampling sets may be created, resulting from 10 PSS setups and 100 unique random sampling trials each. For prioritization-based

subsampling and NPRS, this may be done to see a difference in the information retained about quality as compared to calculating the KPI score for the entire ground truth data. As a counterpart to the average sample size of the top 50% fraction based on the different PSSs, only a 30% fraction was used for NPRS. Top 50% importance was selected, as the variances of KPI scores from this fraction size onwards to 90% may be relatively small.

[0058] The robustness indicator used herein may be the IQR method, as defined above. The IQR method may be used to measure a spread of KPI variance for each prediction model and to identify KPI scores far from the median. KPI Scores that are considered as outliers may be interpreted as subsampling strategies with an unfortunate selection of slices; it may be an indication that a strongly biased quality assessment is present. Outliers are not to be considered as part of decisive factors. In the case of NPRS, 60% of KPI scores across the first 8 feature-based prediction models may be worse than the normal KPI, while this may be 69% for the prioritization-based subsampling. Moreover, it may be observed that for the pessimistic prediction model (number 11), the KPI scores may improve in a PSS-based quality assessment since the important slices may include busy areas and times, suggesting some pessimism necessary for a prediction model to perform well in such slices.

[0059] The benefits assessment may prove to detect weakly designed prediction models that only outperform other prediction models because of unfortunate subsampling selection. It may also verify that marginal benefits comparison between several competing models can be assessed. In summary, the present disclosure may enable assessing a "true" quality of a prediction model while reducing the size of ground truth data. It may also help decide between competing prediction models.

[0060] The proposed data-driven methodology in this document may show that it is possible to smartly reduce ground truth data and still assess the "true" quality of different prediction models using prioritization-based subsampling strategies (PSS). Important neighbourhoods and/or time periods, called slices, may be identified based on a volume share of parking events in the vehicle fleet data. For the use case of on-street parking information (OSPI), the methodology was applied using the vehicle fleet data of Munich, Germany. For OSPI, a particular strategy was created using neighbourhood clusters based on the concept of temporal trend of parking dynamics (TTPD).

[0061] The benefits assessment of the methodology may confirm that, the prioritization-based subsampling can identify weakly designed parking prediction models. This was evaluated based on a comparison with non-prioritized randomized subsampling (NPRS) on a 30% fraction of the ground truth data. The NPRS may be used to quantify the chances of misfortunately selecting slices that do not necessarily represent the true quality. This was accomplished by comparing the quality metric, here KPI, scores at the automatically defined slices across 10 PSS design setups. The prioritization-based subsampling considered the top 50% important slices as a subset of slices to assess the "true" quality of different OSPI models. The methodology also may allow the quality managers of OSPI service providers to gain first valuable insights of a prediction model fast at a lower cost with less ground truth data needed. Thus, the introduced methodology may be useful to companies that are increasing their resources for quality assessment of mobility-related information systems.

[0062] The aspects and features described in relation to a particular one of the previous examples may also be combined with one or more of the further examples to replace an identical or similar feature of that further example or to additionally introduce the features into the further example.

[0063] Examples may further be or relate to a (computer) program including a program code to execute one or more of the above methods when the program is executed on a computer, processor, or other programmable hardware component. Thus, steps, operations, or processes of different ones of the methods described above may also be executed by programmed computers, processors, or other programmable hardware components. Examples may also cover program storage devices, such as digital data storage media, which are machine-, processor- or computer-readable and encode and/or contain machine-executable, processor-executable or computer-executable programs and instructions. Program storage devices may include or be digital storage devices, magnetic storage media such as magnetic disks and magnetic tapes, hard disk drives, or optically readable digital data storage media, for example. Other examples may also include computers, processors, control units, (field) programmable logic arrays ((F)PLAs), (field) programmable gate arrays ((F)PGAs), graphics processor units (GPU), application-specific integrated circuits (ASICs), integrated circuits (ICs) or system-on-a-chip (SoCs) systems programmed to execute the steps of the methods described above.

[0064] It is further understood that the disclosure of several steps, processes, operations, or functions disclosed in the description or claims shall not be construed to imply that these operations are necessarily dependent on the order described, unless explicitly stated in the individual case or necessary for technical reasons. Therefore, the previous description does not limit the execution of several steps or functions to a certain order. Furthermore, in further examples, a single step, function, process, or operation may include and/or be broken up into several sub-steps, -functions, -processes or -operations.

[0065] If some aspects have been described in relation to a device or system, these aspects should also be understood as a description of the corresponding method. For example, a block, device or functional aspect of the device or system may correspond to a feature, such as a method step, of the corresponding method. Accordingly, aspects described in relation to a method shall also be understood as a description of a corresponding block, a corresponding element, a

property or a functional feature of a corresponding device or a corresponding system.

[0066] The following claims are hereby incorporated in the detailed description, wherein each claim may stand on its own as a separate example. It should also be noted that although in the claims a dependent claim refers to a particular combination with one or more other claims, other examples may also include a combination of the dependent claim with the subject matter of any other dependent or independent claim. Such combinations are hereby explicitly proposed, unless it is stated in the individual case that a particular combination is not intended. Furthermore, features of a claim should also be included for any other independent claim, even if that claim is not directly defined as dependent on that other independent claim.

Claims

1. An apparatus (100) for assessing a quality of a prediction model (110) for predicting mobility-related information, the apparatus (100) comprising a processor (120) configured to

provide vehicle fleet data (130) indicative of a usage of vehicles in slices (140) corresponding to different geographical areas and/or different time periods;
 assign an importance weight (150) to the slices (140) in accordance with a density of the usage of vehicles in the respective slice of geographical area and/or time period;
 compare, for at least one of the slices (140), predicted mobility-related information (162) of the prediction model (110) against predefined mobility-related information (172) in said slice, wherein the processor (120) is configured to weight said slice in accordance with its assigned importance weight (150); and
 assess the quality of the prediction model (110) based on the weighted comparison of predicted and predefined mobility-related information (162, 172).

2. Apparatus (100) according to claim 1, wherein the mobility-related information comprises probability values of a parking availability or a traffic status of the respective slice.

3. Apparatus (100) according to one of the previous claims, wherein the vehicle fleet data (130) comprises timestamps and/or geographical coordinates of parking events or traffic events in the slices (140).

4. Apparatus (100) according claim 3, wherein the processor (120) is further configured to determine the importance weight (150) of the slices (140) by dividing a number of parking events or traffic events in the respective slice by a total number of parking events or traffic events.

5. Apparatus (100) according to one of the previous claims, wherein the slices (140) comprise at least one of

map tiles with one or more zoom levels, the zoom level indicating a size of the map tiles;
 a time interval of week-hours, a week-hour being a specific hour within 168 hours of a week;
 the time interval of week-hours combined with the map tiles with one or more zoom levels;
 the time interval of week-hours combined with a cluster of the map tiles.

6. Apparatus (100) according to one of the previous claims, wherein the processor (120) is further configured to select a certain percentage of the slices (140) with the lowest importance weight (150) and set the respective importance weight to zero.

7. Apparatus (100) according to one of the previous claims, wherein the weighted comparison is based on a loss function.

8. Apparatus (100) according to claim 6, wherein the loss function is based on a Brier score.

9. Computer-implemented method (200) for assessing a quality of a prediction model (110) for predicting mobility-related information, comprising

providing (210) vehicle fleet data (130) indicative of a usage of vehicles in slices (140) corresponding to different geographical areas and/or different time periods;
 assigning (220) an importance weight (150) to the slices (140) in accordance with a density of the usage of vehicles in the respective slice of geographical area and/or time;
 comparing (230), for at least one of the slices (140), predicted mobility-related information (162) of the prediction

model (110) against predefined mobility-related information (172) in said slice, wherein said slice is weighted in accordance with its assigned importance weight (150); and
assessing (240) the quality of the prediction model (110) based on the weighted comparison of predicted and predefined mobility-related information (162, 172).

5

10. Method (200) according to claim 9, comprising

assessing the quality of a plurality of prediction models (110) based on a weighted comparison of respective predicted and the predefined mobility-related information (162, 172);
comparing the assessed quality of the plurality of prediction models (110) for determining a suitable prediction model for at least one of the slices (140); and
using the determined suitable prediction model to predict the mobility-related information in the slice.

10

15

20

25

30

35

40

45

50

55

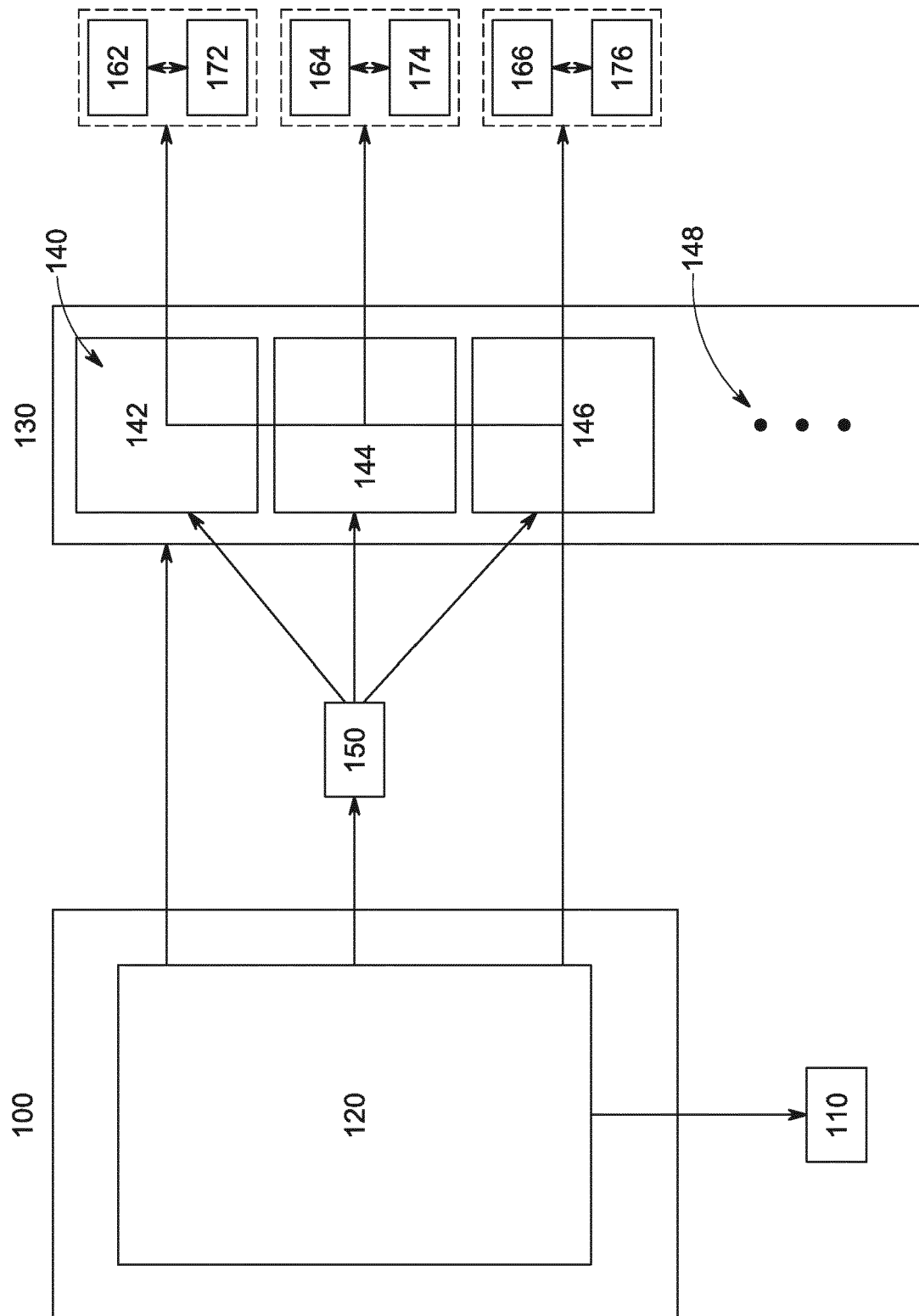


FIG. 1

200

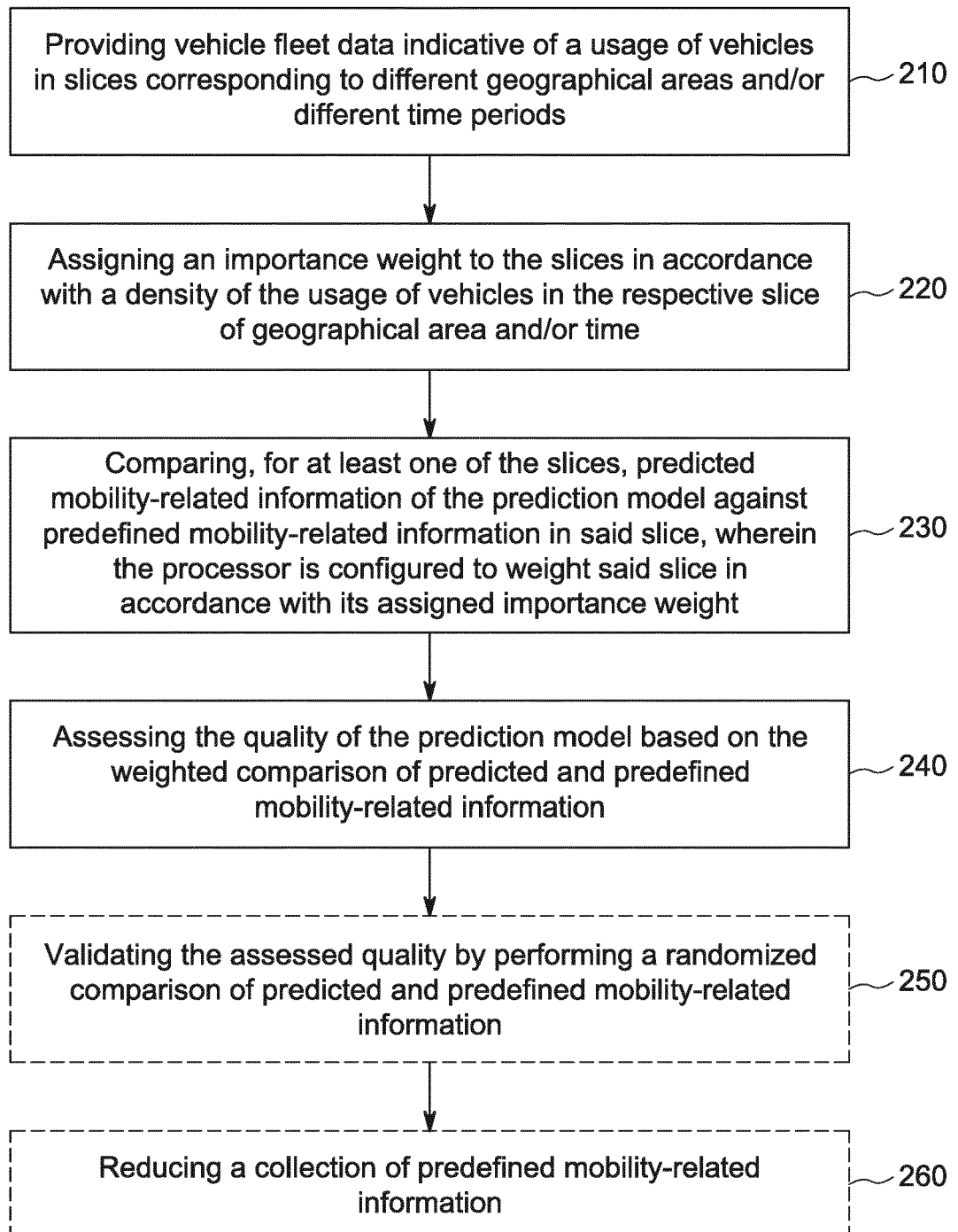


FIG. 2

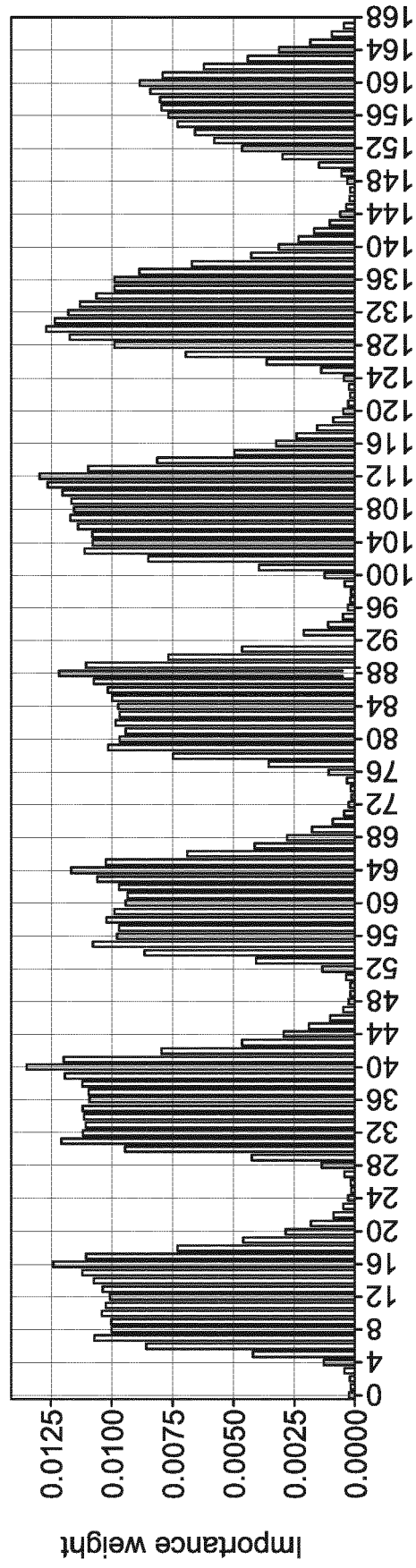


FIG. 3

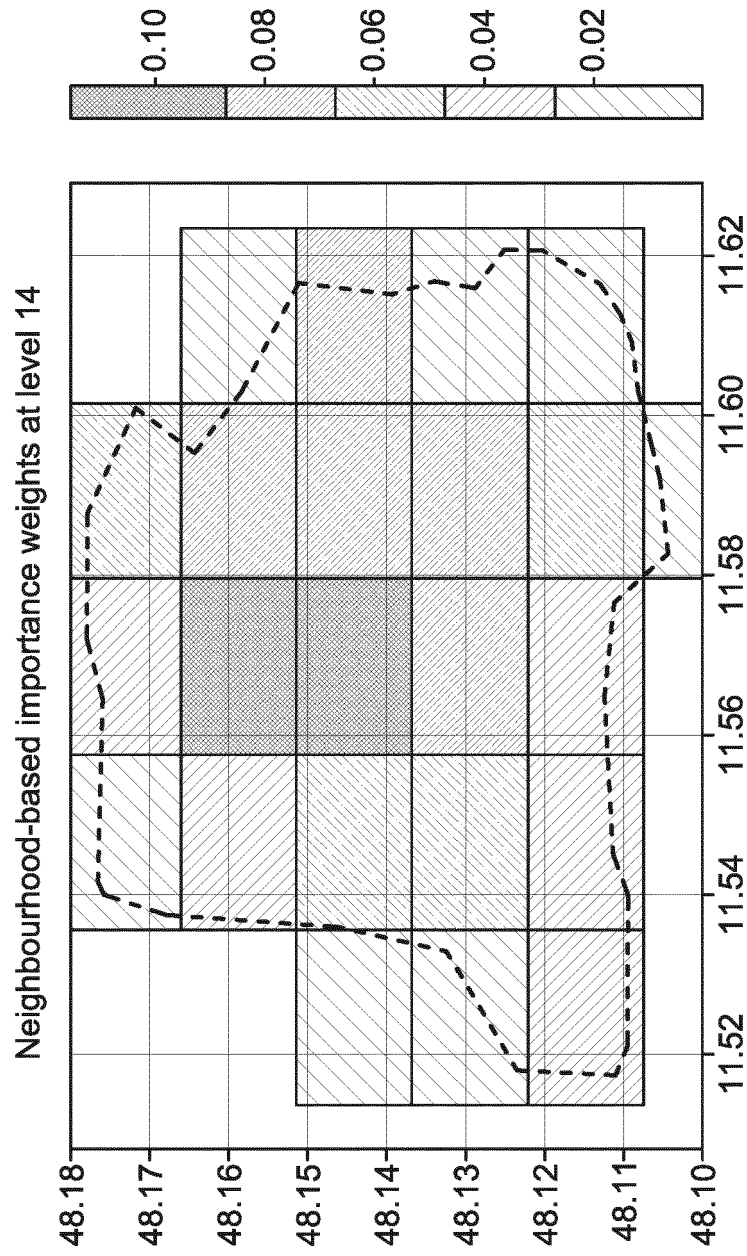


FIG. 4

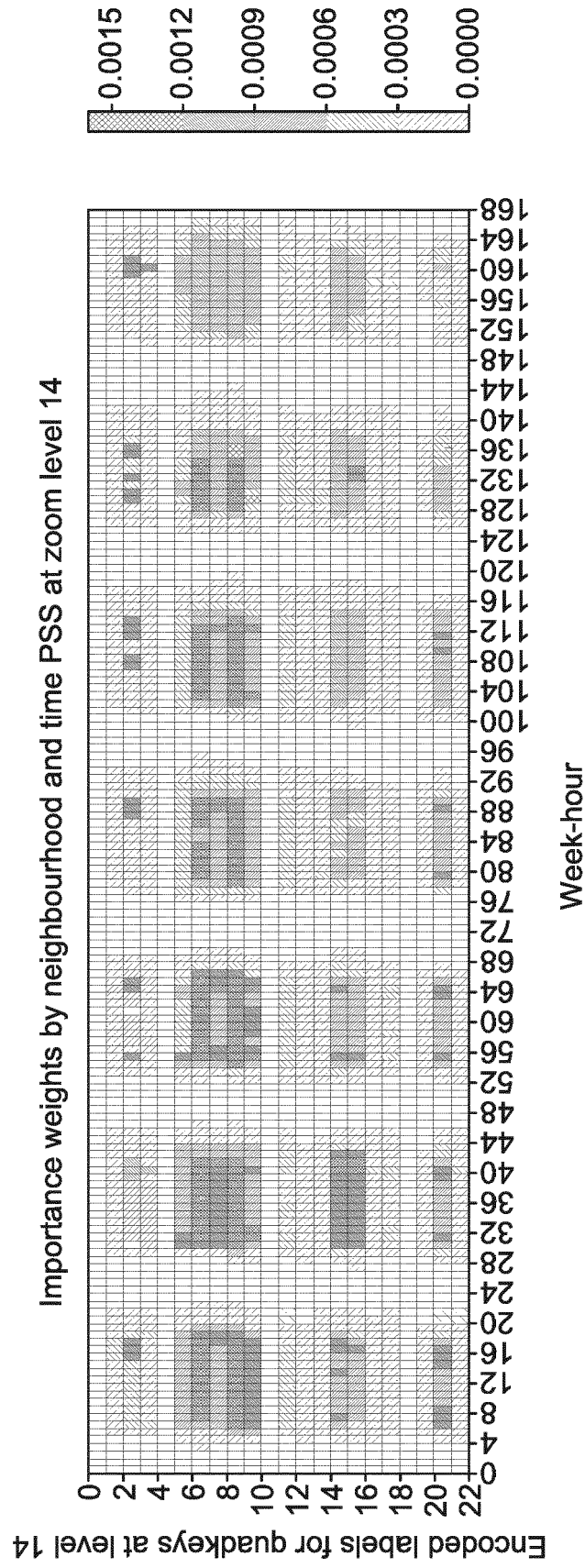


FIG. 5

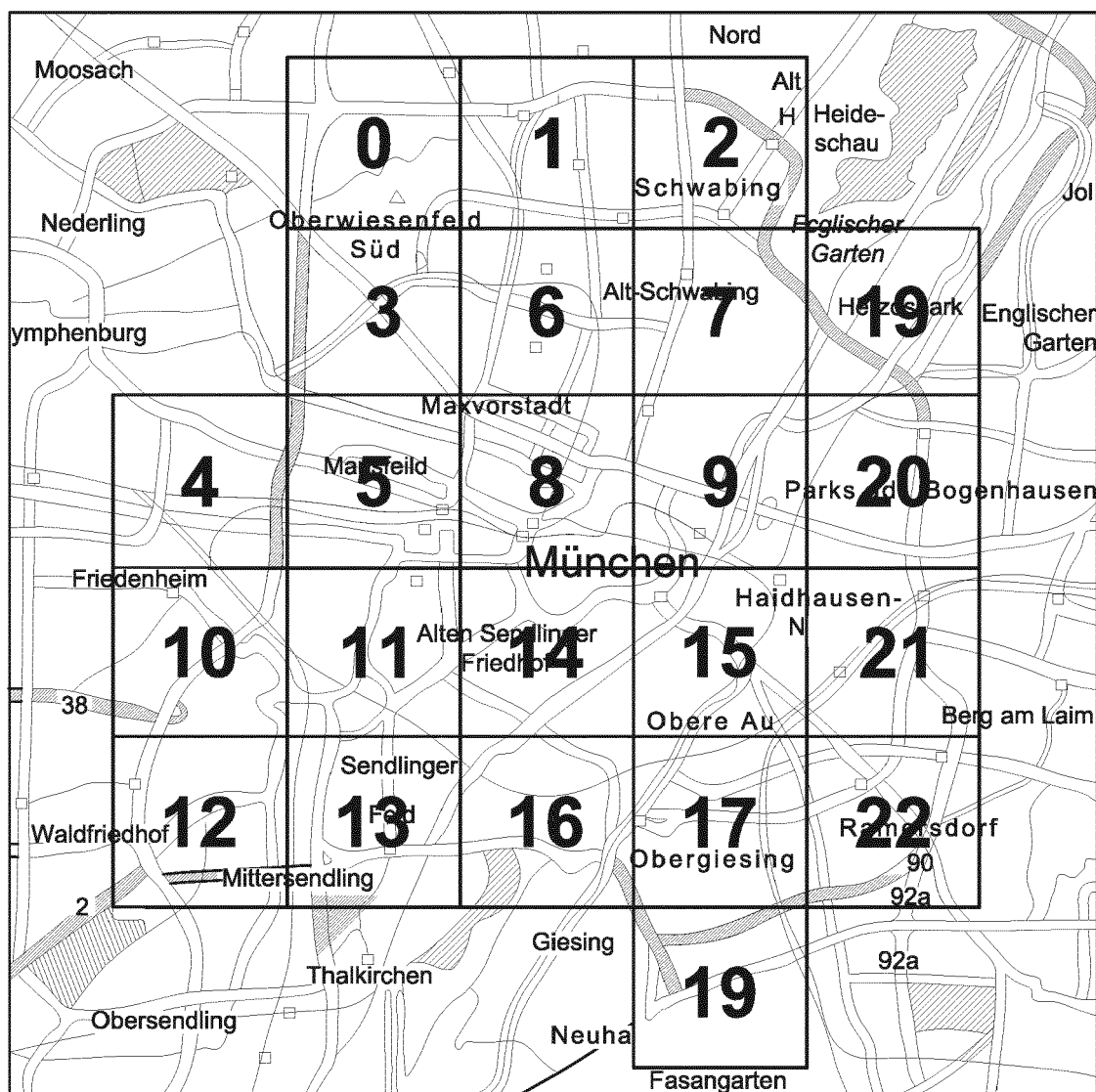


FIG. 6

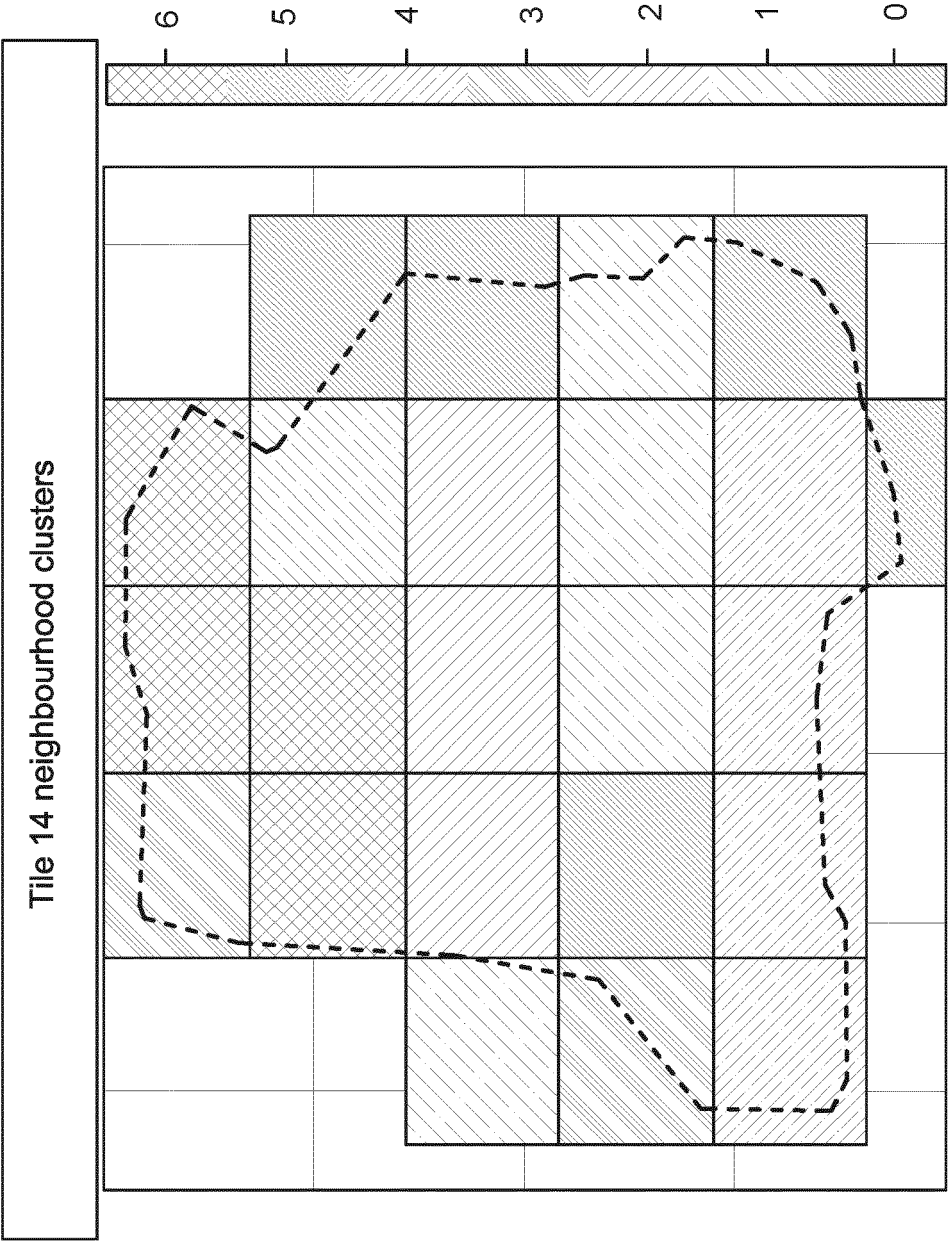


FIG. 7



EUROPEAN SEARCH REPORT

Application Number
EP 21 17 1719

5

10

15

20

25

30

35

40

45

50

55

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	TIAN ZHIYONG ET AL: "Real-Time Charging Station Recommendation System for Electric-Vehicle Taxis", IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, IEEE, PISCATAWAY, NJ, USA, vol. 17, no. 11, November 2016 (2016-11), pages 3098-3109, XP011627014, ISSN: 1524-9050, DOI: 10.1109/TITS.2016.2539201 [retrieved on 2016-10-28]	1-9	INV. G08G1/01 G08G1/14 G08G1/00
Y	* abstract * * paragraph [000I] * * paragraph [011I] - paragraph [0V.A] * * paragraph [00VI] * * figures 1-8 * * tables II-IX * * example algorithm 1 * * example algorithm 2 *	10	
Y	US 2016/300150 A1 (WATTS-FITZGERALD COLIN [US] ET AL) 13 October 2016 (2016-10-13) * abstract * * paragraph [0001] - paragraph [0010] * * paragraph [0026] - paragraph [0111] * * figures 1A-12 *	10	TECHNICAL FIELDS SEARCHED (IPC) G08G
X	US 2016/104377 A1 (FRENCH LESLIE JOHN [US] ET AL) 14 April 2016 (2016-04-14) * abstract * * paragraph [0002] * * paragraph [0058] - paragraph [0071] * * paragraph [0094] - paragraph [0276] * * paragraph [1195] - paragraph [1559] * * figures 1-50 *	1-10	
The present search report has been drawn up for all claims			
Place of search The Hague		Date of completion of the search 4 October 2021	Examiner Dujardin, Corinne
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document	

EPO FORM 1503 03.82 (P04C01)

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 21 17 1719

5 This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

04-10-2021

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2016300150 A1	13-10-2016	NONE	
US 2016104377 A1	14-04-2016	CA 2905503 A1	25-09-2014
		MX 368600 B	09-10-2019
		US 2016104377 A1	14-04-2016
		US 2021264776 A1	26-08-2021
		WO 2014153130 A1	25-09-2014

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Non-patent literature cited in the description

- **R. C. HAMPSHIRE ; D. SHOUP.** What Share of Traffic is Cruising for Parking?. *J. Transp. Econ. Policy*, July 2018, vol. 52, 184-201 **[0002]**
- Zoom levels and tile grid. Microsoft, 2020 **[0026]**