

METHOD AND SYSTEM FOR CONTROLLING A SURGE PRICING FOR A GEOGRAPHIC REGION FOR A TRANSPORTATION SERVICE NETWORK

TECHNICAL FIELD

5 **[0001]** The present invention generally relates to a method of controlling a surge pricing for a transportation service network, and a system thereof, such as for ride-hailing services and/or delivery services.

BACKGROUND

10 **[0002]** Transport surge pricing is one of the most important levers to maintain supply and demand balance in a transportation service network and thus the overall system health, especially during peak demand and supply crunch situations. There are existing systems for controlling a surge pricing for a transportation service network based on PID (Proportional-Integral-Derivative) control. For example, such existing systems may determine the surge
15 pricing for each geohash every one minute based on real-time demand/supply signals. However, such existing systems suffer from various drawbacks or deficiencies. In particular, since such existing systems determine the surge pricing only based on real-time demand/supply signals (i.e., current demand/supply conditions), they only react to changes in current demand/supply conditions and thus may not be optimal or effective. For example, the surge pricing determined
20 may only react to a demand peak that is currently happening (based on the real-time demand/supply signals) and thus may be too slow or ineffective in attempting to maintain supply and demand balance in the transportation service network. In other words, such existing systems provide a slow or ineffective response in controlling surge pricing for managing changes, especially fast changes, in supply/demand conditions in the transportation service
25 network. Furthermore, the PID control model in such existing systems is configured with pre-defined formulas with a large number of hyper-parameters requiring complicated and time-intensive manual tuning in an attempt to optimize its performance in dynamic price surging (e.g., regionally), which thus results in significant implementation inefficiencies and costs, as well as being prone to human tuning errors or inaccuracies.

30 **[0003]** A need therefore exists to provide a method of controlling a surge pricing for a transportation service network, as well as a system thereof, that seeks to overcome, or at least ameliorate, one or more deficiencies in conventional methods of controlling a surge pricing for a transportation service network, and more particularly, with improved effectiveness in

controlling the surge pricing for managing supply and demand in the transportation service network. It is against this background that the present invention has been developed.

SUMMARY

5 **[0004]** According to a first aspect of the present invention, there is provided a method of controlling a surge pricing for a geographic region for a transportation service network using at least one processor, the method comprising, for each time step of a series of time steps:

optimizing, using a cost function for the transportation service network, a sequence of surge pricings for a sequence of time steps, including the time step as a first time step thereof,
10 over a prediction horizon for the geographic region based on a sequence of predicted performance indicators, a sequence of predicted supply indicators and a sequence of predicted demand indicators for the sequence of time steps with respect to the geographic region; and

applying a first surge pricing of the sequence of surge pricings optimized for the first time step to the geographic region for the first time step,

15 wherein the cost function is parameterized by a sequence of surge pricing variables for the sequence of time steps for optimization to obtain the sequence of surge pricings for the sequence of time steps, and

the sequence of predicted performance indicators is predicted using a system dynamics model for the transportation service network based on the sequence of surge pricing variables
20 for the sequence of time steps.

[0005] According to a second aspect of the present invention, there is provided a system for controlling a surge pricing for a geographic region for a transportation service network, the system comprising:

at least one memory; and

25 at least one processor communicatively coupled to the at least one memory and configured to, for each time step of a series of time steps:

optimize, using a cost function for the transportation service network, a sequence of surge pricings for a sequence of time steps, including the time step as a first time step thereof, over a prediction horizon for the geographic region based on a sequence of predicted
30 performance indicators, a sequence of predicted supply indicators and a sequence of predicted demand indicators for the sequence of time steps with respect to the geographic region; and

apply a first surge pricing of the sequence of surge pricings optimized for the first time step to the geographic region for the first time step,

wherein the cost function is parameterized by a sequence of surge pricing variables for the sequence of time steps for optimization to obtain the sequence of surge pricings for the sequence of time steps, and

the sequence of predicted performance indicators is predicted using a system dynamics
5 model for the transportation service network based on the sequence of surge pricing variables for the sequence of time steps.

[0006] According to a third aspect of the present invention, there is provided a computer program product, embodied in one or more non-transitory computer-readable storage mediums, comprising instructions executable by at least one processor to perform the method of
10 controlling a surge pricing for a geographic region for a transportation service network according to the above-mentioned first aspect of the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] Embodiments of the present invention will be better understood and readily apparent
15 to one of ordinary skill in the art from the following written description, by way of example only, and in conjunction with the drawings, in which:

FIG. 1 depicts a schematic flow diagram of a method of controlling a surge pricing for a geographic region for a transportation service network, according to various embodiments of the present invention;

20 FIG. 2 depicts a schematic block diagram of a system for controlling a surge pricing for a geographic region for a transportation service network, according to various embodiments of the present invention;

FIG. 3 depicts a schematic block diagram of an exemplary computer system which may be used to realize or implement the system for controlling a surge pricing for a geographic
25 region for a transportation service network, according to various embodiments of the present invention;

FIG. 4 depicts a schematic flow diagram showing an overview of an example method of controlling a surge pricing for a geographic region for a transportation service network, according to various embodiments of the present invention; and

30 FIG. 5 depicts a schematic drawing illustrating two consecutive time steps and three consecutive timestamps, according to various embodiments of the present invention.

DETAILED DESCRIPTION

[0008] Various embodiments of the present invention provide a method and a system for controlling a surge pricing for a transportation service network, such as for ride-hailing services or delivery services.

5 **[0009]** As explained in the background, there are existing systems for controlling a surge pricing for a transportation service network based on PID control. However, such existing systems suffer from various drawbacks or deficiencies, including providing a slow or ineffective response in controlling surge pricing for managing changes in supply/demand conditions in the transportation service network, as well as significant implementation
10 inefficiencies and costs. In this regard, various embodiments of the present invention provide a method of controlling a surge pricing for a transportation service network, as well as a system thereof, that seeks to overcome, or at least ameliorate, one or more deficiencies in conventional methods of controlling a surge pricing for a transportation service network, and more particularly, with improved effectiveness in controlling the surge pricing for managing supply
15 and demand in the transportation service network. The method also improves implementation efficiencies and costs.

[0010] FIG. 1 depicts a schematic flow diagram of a method 100 of controlling a surge pricing for a geographic region for a transportation service network according to various
20 embodiments of the present invention. The method 100 comprises, for each time step of a series of time steps: optimizing (at 106), using a cost function for the transportation service network, a sequence of surge pricings for a sequence of time steps, including the time step (i.e., the above-mentioned time step of the series of time steps) as a first time step thereof, over a prediction horizon for the geographic region based on a sequence of predicted performance indicators, a sequence of predicted supply indicators and a sequence of predicted demand
25 indicators for the sequence of time steps with respect to the geographic region; and applying (at 108) a first surge pricing of the sequence of surge pricings optimized for the first time step to the geographic region for the first time step. The cost function is parameterized by a sequence of surge pricing variables for the sequence of time steps for optimization to obtain the sequence of surge pricings for the sequence of time steps. Furthermore, the sequence of predicted
30 performance indicators is predicted using a system dynamics model for the transportation service network based on the sequence of surge pricing variables for the sequence of time steps.

[0011] Therefore, the method 100 of controlling a surge pricing for a geographic region according to various embodiments of the present invention advantageously has improved

effectiveness in controlling the surge pricing for managing supply and demand in the transportation service network. In particular, the surge pricing applied to a geographic region for a time period (e.g., corresponding to a time step or a time window) is optimized over a prediction horizon (a sequence of time steps) based on a sequence of predicted performance indicators, a sequence of predicted supply indicators and a sequence of predicted demand indicators over the prediction horizon with respect to the geographic region. In this regard, the predicted performance indicators are predicted using a system dynamics model for the transportation service network (for representing a performance of the transportation service network) based on a sequence of surge pricing variables over the prediction horizon. Therefore, the surge pricing applied to a geographic region for a time period is optimized taking into account performances of the transportation service network (e.g., taking into account impact to performances), predicted supply volumes (e.g., including forecasted replenishment supply volumes and taking into account impact to supply volumes, e.g., driver supply volumes) and predicted demand volumes (e.g., including forecasted new demand volumes and taking into account impact to demand volumes, e.g., driver demand volumes) over the prediction horizon as a result of the surge pricing being optimized. Therefore, the method 100 of controlling a surge pricing for a geographic region according to various embodiments of the present invention is able to control the surge pricing for managing supply and demand in the transportation service network with improved effectiveness. For example, according to the method 100, when demand conditions for certain subsequent time periods are forecasted to be high, the surge pricing for the surge pricing for a time period prior thereto may be controlled preemptively (e.g., increase the surge pricing) before the demand conditions become high so as to conserve supply (e.g., driver supply) for the upcoming high demand time periods and avoid (or mitigate) a sudden supply crush (which may otherwise result in a sharp spike in surge pricing), thus more effectively managing supply and demand in the transportation service network. These advantages or technical effects, and/or other advantages or technical effects, will become more apparent to a person skilled in the art as the method 100 of controlling a surge pricing for a transportation request, as well as the corresponding system for controlling a surge pricing for a geographic region, is described in more detail according to various embodiments and example embodiments of the present invention.

[0012] In various embodiments, transportation requests may be made to the transportation service network, each transportation request may have associated therewith a pick-up location, a drop-off location and a pick-up time for a transportation (i.e., a transportation request

comprising pick-up location information, drop-off location information and pick-up time information for the transportation). In various embodiments, the surge pricing for a geographic region for a time period (e.g., corresponding to a time step or a time window) applied by the method 100 may thus apply to all transportation requests having associated therewith a pick-up location belonging to (or within) the geographic region and a pick-up time belonging to (or within) the time period.

[0013] In various embodiments, the cost function comprises a first cost component configured to, for each of the sequence of surge pricing variables, determine a revenue indicator with respect to the geographic region for the time step corresponding to the surge pricing variable based on the surge pricing variable.

[0014] In various embodiments, the cost function further comprises: a second cost component configured to, for each of the sequence of surge pricing variables, determine a supply utilization indicator with respect to the geographic region for the time step corresponding to the surge pricing variable based on the surge pricing variable; and a third cost component configured to, for each of the sequence of surge pricing variables, determine a demand pressure indicator with respect to the geographic region for the time step corresponding to the surge pricing variable based on the surge pricing variable.

[0015] In various embodiments, the second cost component is configured to determine the supply utilization indicator with respect to the geographic region for the time step corresponding to the surge pricing variable based on an available supply volume at a start of the time step, the predicted supply indicator for the time step and a filled supply indicator for the time step. In various embodiments, the predicted supply indicator for the time step corresponds to a predicted replenishment supply volume for the time step, and the filled supply indicator for the time step corresponds to a supply volume filled by demand volume for the time step.

[0016] In various embodiments, the third cost component is configured to determine the demand pressure indicator with respect to the geographic region for the time step corresponding to the surge pricing variable based on a pending demand volume at the start of the time step, the predicted demand indicator for the time step and the filled supply indicator for the time step.

In various embodiments, the predicted demand indicator for the time step corresponds to a predicted amount of transportation requests (e.g., corresponding to predicted new demand volumes) for the time step.

[0017] In various embodiments, the filled supply indicator for the time step is determined based on the available supply volume at the start of the time step, the predicted supply indicator for the time step, the pending demand volume at the start of the time step, the predicted demand indicator for the time step and the predicted performance indicator for the time step.

5 **[0018]** In various embodiments, the above-mentioned system dynamics model comprises a transportation booking probability model configured to predict a probability which a transportation request with respect to the geographic region is converted into a transportation booking based on a surge pricing variable. In this regard, the predicted performance indicator for the time step corresponds to a predicted probability which a transportation request using the
10 surge pricing of the surge pricing variable corresponding to the time step and with respect to the geographic region is converted into a transportation booking.

[0019] In various embodiments, the above-mentioned optimizing (at 106), using the cost function for the transportation service network, the sequence of surge pricings for the sequence of time steps, comprises maximizing the cost function with respect to the sequence of surge
15 pricing variables for the sequence of time steps over the prediction horizon, subject to one or more constraints relating to the sequence of surge pricing variable.

[0020] In various embodiments, the one or more constraints comprises: a first constraint configured for bounding a value of each surge pricing variable of the sequence of surge pricing variables; and a second constraint configured for bounding a difference between the values of
20 two consecutive surge pricing variables of the sequence of surge pricing variables for the sequence of time steps.

[0021] In various embodiments, the above-mentioned optimize (at 106), using the cost function for the transportation service network, the sequence of surge pricings for the sequence of time steps is based on model predictive control.

25 **[0022]** In various embodiments, the transportation service network is configured to provide ride-hailing services and/or delivery services (e.g., food delivery services). In this regard, the above-mentioned sequence of predicted supply indicators is with respect to driver supply and the above-mentioned sequence of predicted demand indicators is with respect to driver demand. Furthermore, the geographic region is a geohash.

30 **[0023]** FIG. 2 depicts a schematic block diagram of a system 200 for controlling a surge pricing for a geographic region for a transportation service network according to various embodiments of the present invention, corresponding to the above-mentioned method 100 of for controlling a surge pricing as described hereinbefore with reference to FIG. 1 according to

various embodiments of the present invention. The system 200 comprises: at least one memory 202; and at least one processor 204 communicatively coupled to the at least one memory 202 and configured to perform the method 100 of for controlling a surge pricing as described hereinbefore according to various embodiments of the present invention. Accordingly, the at least one processor 204 is configured to, for each time step of a series of time steps: optimize, using a cost function for the transportation service network, a sequence of surge pricings for a sequence of time steps, including the time step as a first time step thereof, over a prediction horizon for the geographic region based on a sequence of predicted performance indicators, a sequence of predicted supply indicators and a sequence of predicted demand indicators for the sequence of time steps with respect to the geographic region; and apply a first surge pricing of the sequence of surge pricings optimized for the first time step to the geographic region for the first time step. The cost function is parameterized by a sequence of surge pricing variables for the sequence of time steps for optimization to obtain the sequence of surge pricings for the sequence of time steps. The sequence of predicted performance indicators is predicted using a system dynamics model for the transportation service network based on the sequence of surge pricing variables for the sequence of time steps.

[0024] It will be appreciated by a person skilled in the art that the at least one processor 204 may be configured to perform various functions or operations through set(s) of instructions (e.g., software modules) executable by the at least one processor 204 to perform various functions or operations. Accordingly, as shown in FIG. 2, the system 200 may comprise: a surge pricing optimization module (or a surge pricing optimization circuit) 206 configured to optimize, using a cost function for the transportation service network, a sequence of surge pricings for a sequence of time steps, including the time step as a first time step thereof, over a prediction horizon for the geographic region based on a sequence of predicted performance indicators, a sequence of predicted supply indicators and a sequence of predicted demand indicators for the sequence of time steps with respect to the geographic region; and a surge pricing implementation module (or a surge pricing implementation circuit) 208 configured to apply a first surge pricing of the sequence of surge pricings optimized for the first time step to the geographic region for the first time step.

[0025] It will be appreciated by a person skilled in the art that the above-mentioned modules are not necessarily separate modules, and two or more modules may be realized by or implemented as one functional module (e.g., a circuit or a software program) as desired or as appropriate without deviating from the scope of the present invention. For example, the surge

pricing optimization module 206 and surge pricing implementation module 208 may be realized (e.g., compiled together) as one executable software program (e.g., software application or simply referred to as an “app”), which for example may be stored in the at least one memory 202 and executable by the at least one processor 204 to perform the corresponding functions or operations as described herein according to various embodiments.

[0026] In various embodiments, the system 200 for controlling a surge pricing corresponds to the method 100 of controlling a surge pricing as described hereinbefore with reference to FIG. 1, therefore, various operations, functions or steps configured to be performed by the least one processor 204 may correspond to various operations, functions or steps of the method 100 of controlling a surge pricing described hereinbefore according to various embodiments, and thus need not be repeated with respect to the system 200 for clarity and conciseness. In other words, various embodiments described herein in context of methods (e.g., the method 100 of controlling a surge pricing for a geographic region) are analogously valid for the corresponding systems or devices (e.g., the system 200 for controlling a surge pricing for a geographic region), and vice versa. For example, in various embodiments, the at least one memory 202 may have stored therein the surge pricing optimization module 206 and/or surge pricing implementation module 208, which respectively correspond to various operations, functions or steps of the method 100 of controlling a surge pricing as described hereinbefore according to various embodiments, which are executable by the at least one processor 204 to perform the corresponding operations, functions or steps as described herein.

[0027] A computing system, a controller, a microcontroller or any other system providing a processing capability may be provided according to various embodiments in the present invention. Such a system may be taken to include one or more processors and one or more computer-readable storage mediums. For example, the system 200 described hereinbefore may include at least one processor (or controller) 204 and at least one computer-readable storage medium (or memory) 202 which are for example used in various processing carried out therein as described herein. A memory or computer-readable storage medium used in various embodiments may be a volatile memory, for example a DRAM (Dynamic Random Access Memory) or a non-volatile memory, for example a PROM (Programmable Read Only Memory), an EPROM (Erasable PROM), EEPROM (Electrically Erasable PROM), or a flash memory, e.g., a floating gate memory, a charge trapping memory, an MRAM (Magnetoresistive Random Access Memory) or a PCRAM (Phase Change Random Access Memory).

[0028] In various embodiments, a “circuit” may be understood as any kind of a logic implementing entity, which may be special purpose circuitry or a processor executing software stored in a memory, firmware, or any combination thereof. Thus, in an embodiment, a “circuit” may be a hard-wired logic circuit or a programmable logic circuit such as a programmable
5 processor, e.g., a microprocessor (e.g., a Complex Instruction Set Computer (CISC) processor or a Reduced Instruction Set Computer (RISC) processor). A “circuit” may also be a processor executing software, e.g., any kind of computer program, e.g., a computer program using a virtual machine code, e.g., Java. Any other kind of implementation of various functions or operations may also be understood as a “circuit” in accordance with various other
10 embodiments. Similarly, a “module” may be a portion of a system according to various embodiments in the present invention and may encompass a “circuit” as above, or may be understood to be any kind of a logic-implementing entity therefrom.

[0029] Some portions of the present disclosure are explicitly or implicitly presented in terms of algorithms and functional or symbolic representations of operations on data within a
15 computer memory. These algorithmic descriptions and functional or symbolic representations are the means used by those skilled in the data processing arts to convey most effectively the substance of their work to others skilled in the art. An algorithm is here, and generally, conceived to be a self-consistent sequence of steps leading to a desired result. The steps are those requiring physical manipulations of physical quantities, such as electrical, magnetic or
20 optical signals capable of being stored, transferred, combined, compared, and otherwise manipulated.

[0030] The present specification also discloses a system (e.g., which may also be embodied as one or more devices or apparatuses), such as the system 200, for performing various operations, functions or steps of various methods described herein. Such a system may be
25 specially constructed for the required purposes or may comprise a general purpose computer system selectively activated or reconfigured by a computer program stored in the computer system. In general, various algorithms that may be presented herein are not limited to being implemented or executed by any particular computer system. Alternatively, the construction of more specialized computer system to perform various operations, functions or steps of various
30 methods described herein may be provided as desired or as appropriate without going beyond the scope of the present invention.

[0031] In addition, the present specification also at least implicitly discloses computer program(s) or software/functional module(s), in that it would be apparent to a person skilled in

the art that various operations, functions or steps of various methods described herein may be put into effect by computer code. The computer program(s) is not intended to be limited to any particular programming language and implementation thereof, and it will be appreciated by a person skilled in the art that a variety of programming languages and coding thereof may be used to implement the computer program(s). Moreover, the computer program(s) is not intended to be limited to any particular control flow as there are a variety of programming languages which can use different control flows. It will be appreciated by a person skilled in the art that a computer program may be stored on any computer-readable storage medium (non-transitory computer-readable storage medium), such as but not limited to, a magnetic disk, an optical disk or a memory chip. For example, a computer program stored on a computer-readable storage medium may be loaded and executed on a computer system to implement various operations, functions or steps of various methods described herein according to various embodiments of the present invention.

[0032] Accordingly, in various embodiments, there is provided a computer program product, embodied in one or more computer-readable storage mediums (non-transitory computer-readable storage medium), comprising instructions (e.g., the surge pricing optimization module 206 and/or surge pricing implementation module 208) executable by one or more computer processors to perform a method 100 of controlling a surge pricing as described hereinbefore with reference to FIG. 1 according to various embodiments of the present invention. Accordingly, various computer programs or software modules described herein may be stored in a computer program product receivable by a system therein, such as the system 200 as shown in FIG. 2, for execution by at least one processor 204 of the system 200 to perform various operations, functions or steps of various methods described herein according to various embodiments of the present invention.

[0033] It will be appreciated by a person skilled in the art that various modules described herein (e.g., the surge pricing optimization module 206 and/or surge pricing implementation module 208) may be software module(s) realized by computer program(s) or set(s) of instructions executable by a computer processor to perform various functions or operations. Various modules described herein (e.g., the surge pricing optimization module 206 and/or surge pricing implementation module 208) may also be implemented as hardware module(s) being functional hardware unit(s) designed to perform various functions or operations. More particularly, in the hardware sense, a module is a functional hardware unit designed for use with other components or modules. For example, a module may be implemented using discrete

electronic components, or it can form a portion of an entire electronic circuit such as an Application Specific Integrated Circuit (ASIC). Numerous other possibilities exist. It will also be appreciated by a person skilled in the art that a combination of hardware and software modules may be implemented. Furthermore, various operations, functions or steps of various methods described herein may be performed in parallel rather than sequentially as desired or as appropriate (e.g., as long as it does not render the method(s) inoperable or unsatisfactory for its intended purpose).

[0034] In various embodiments, the system 200 for controlling a surge pricing may be realized by any computer system (e.g., desktop or portable computer system) including at least one processor and at least one memory, such as an example computer system 300 as schematically shown in FIG. 3 as an example only and without limitation. Various methods/steps or functional modules may be implemented as software, such as a computer program being executed within the computer system 300, and instructing the computer system 300 (in particular, one or more processors therein) to conduct various functions or operations as described herein according to various embodiments. The computer system 300 may comprise a system unit 302, one or more input devices 304 such as a keyboard, a touchscreen and/or a mouse, and a plurality of output devices such as a display 308. The system unit 32 may be connected to a computer network 312 via a suitable transceiver device 314, to enable access to e.g., the Internet or other network systems such as Local Area Network (LAN) or Wide Area Network (WAN). The system unit 302 may include a processor 318 for executing various instructions, a Random Access Memory (RAM) 320 and a Read Only Memory (ROM) 322. The system unit 302 may further include a number of Input/Output (I/O) interfaces, for example I/O interface 324 to the display device 308 and I/O interface 326 to the one or more input devices 304. The components of the system unit 302 typically communicate via an interconnected bus 328 and in a manner known to a person skilled in the art.

[0035] For example, the system 200 may receive a transportation request made to the transportation service network over a wireless communication network (e.g., cellular network (e.g., 4G, 5G or a future generation cellular network)) from a computing device or system (e.g., a portable computing device, such as a mobile communication device) of a user requesting a transportation, such as for a ride-hailing service. Each transportation request may have associated therewith a pick-up location, a drop-off location and a pick-up time for a transportation. In various embodiments, the surge pricing for a geographic region for a time period applied by the method 100 as described herein according to various embodiments of the

present invention may thus apply to all transportation requests having associated therewith a pick-up location belonging to (or within) the geographic region and a pick-up time belonging to (or within) the time period. In this regard, an overall price or fare for the transportation request (e.g., based on a base price and the above-mentioned surge pricing determined/applied) may then be transmitted from the system 200 over the wireless communication network to the device of the user whom made the transportation request for displayed thereat for consideration or confirmation by the user on whether to make the transportation booking.

[0036] It will be appreciated by a person skilled in the art that the terminology used herein is for the purpose of describing various embodiments only and is not intended to be limiting of the present invention. As used herein, the singular forms “a”, “an” and “the” are intended to include the plural forms as well, unless the context clearly indicates otherwise. It will be further understood that the terms “comprises” and/or “comprising,” when used in this specification, specify the presence of stated features, integers, steps, operations, elements, and/or components, but do not preclude the presence or addition of one or more other features, integers, steps, operations, elements, components, and/or groups thereof.

[0037] Any reference to an element or a feature herein using a designation such as “first”, “second” and so forth does not limit the quantity or order of such elements or features, unless stated or the context requires otherwise. For example, such designations may be used herein as a convenient way of distinguishing between two or more elements or instances of an element. Thus, a reference to first and second elements does not necessarily mean that only two elements can be employed, or that the first element must precede the second element, unless stated or the context requires otherwise. In addition, a phrase referring to “at least one of” a list of items refers to any single item therein or any combination of two or more items therein.

[0038] In order that the present invention may be readily understood and put into practical effect, various example embodiments of the present invention will be described hereinafter by way of examples only and not limitations. It will be appreciated by a person skilled in the art that the present invention may, however, be embodied in various different forms or configurations and should not be construed as limited to the example embodiments set forth hereinafter. Rather, these example embodiments are provided so that this disclosure will be thorough and complete, and will fully convey the scope of the present invention to those skilled in the art.

[0039] In particular, for better understanding of the present invention and without limitation or loss of generality, and unless stated or the context requires otherwise, various example

embodiments of the present invention will now be described with respect to an example practical application of the transportation service network being configured for ride-hailing services for illustration purposes only. It will be understood by a person skilled in the art that the present invention is not limited to such an example practical application and may be employed in a variety of other practical applications as desired or as appropriate without going beyond the scope of the present invention, as long as a surge pricing for a transportation service network is desired to be controlled (or dynamically controlled), such as but not limited to, delivery services (e.g., food delivery services, where a surge pricing for a food delivery request to a food delivery service network may be controlled).

[0040] As explained in the background, there are existing systems for controlling a surge pricing for a transportation service network based on PID control. However, such existing systems suffer from various drawbacks or deficiencies, including providing a slow or ineffective response in controlling surge pricing for managing changes in supply/demand conditions in the transportation service network and significant implementation inefficiencies and costs. In this regard, various embodiments of the present invention provide a method of controlling a surge pricing for a transportation service network, as well as a system thereof, that seeks to overcome, or at least ameliorate, one or more deficiencies in conventional methods of controlling a surge pricing for a transportation service network, and more particularly, with improved effectiveness in controlling the surge pricing for managing supply and demand in the transportation service network. The method of controlling a surge pricing also improves implementation efficiencies and costs.

[0041] In particular, various example embodiments provide a method (or a system) for controlling a surge pricing for a geographic region for a transportation service network based on model predictive control (MPC) (which may be referred to herein as the MPC-based surge pricing method (or the MPC-based surge pricing system)). In this regard, the MPC-based surge pricing method is configured to determine an optimal surge pricing (spatial and temporal surge pricing) for a geographic region and for a time period (e.g., corresponding to a time step or a time window) based on predicted performance indicators predicted using a system dynamics model for the transportation service network (for representing a performance of the transportation service network) and predicted supply and demand indicators (e.g., forecasted replenishment supply and new demand volumes or conditions) over a prediction horizon. The MPC-based surge pricing method advantageously overcomes or addresses various disadvantages or deficiencies of existing systems for controlling a surge pricing based on PID

control (which may be referred to herein as the PID-based surge pricing systems), resulting in improved effectiveness in controlling the surge pricing for managing supply and demand in the transportation service network. For example, the MPC-based surge pricing method is advantageously able to control the surge pricing for a geographic region which optimizes the transportation request conversion rate (i.e., the probability which transportation requests are converted into transportation bookings, e.g., the ride conversion rate, which may also be referred to herein as the book through rate (BTR)) and thus, the revenue generated in the transportation service network (e.g., the gross merchandise value (GMV)) with respect to the geographic region, while maximizing supply utilization and minimizing demand pressure with respect to the geographic region.

[0042] As discussed in the background, in existing PID-based surge pricing systems, the surge pricing is determined only based on real-time demand and supply signals (i.e., current demand/supply conditions). Therefore, they are only able to react to changes in the current demand/supply conditions, which may not be optimal or effective. For example, the surge pricing determined may only react when a demand peak is currently happening (based on the real-time demand/supply signals) and thus may be too slow or ineffective in attempting to manage supply and demand balance in the transportation service network. In contrast, when determining the surge pricing, the MPC-based surge pricing method takes into account predicted supply and demand indicators (e.g., forecasted replenishment supply and new demand volumes or conditions) over a prediction horizon, and thus is advantageously able to control or adjust the surge pricing in advance (e.g., preemptively) to certain demand/supply conditions happening to more effectively control the surge pricing for better managing supply and demand in the transportation service network. For example, when the demand conditions for certain subsequent time periods are forecasted to be high (e.g., expected peak hours or major events), the surge pricing for certain time periods prior thereto may be controlled preemptively (e.g., optimized to progressively increased preemptively) before the demand conditions become high so as to conserve supply (driver supply) for the upcoming high demand time periods and avoid (or mitigate) a sudden supply crunch (which may otherwise result in a sharp spike in surge pricing), thus effectively managing supply and demand in the transportation service network for the benefit of both drivers and riders. Accordingly, enabling the surge pricing to preemptively increase based on forecasted high demand time periods helps to avoid (or mitigate) hard kick-in and sudden price surge increase issues when such high demand time periods occur. In various example embodiments, the MPC-based surge pricing method may

further take into account forecasted weather conditions in controlling the surge pricing. In this regard, weather conditions influence the supply and demand balance of the transportation service network. Therefore, by taking into account forecasted weather conditions, the MPC-based surge pricing method is advantageously able to control or adjust the surge pricing in advance (preemptively) to the demand/supply conditions changing due to changes in weather conditions to more effectively control the surge pricing for managing supply and demand in the transportation service network. In addition, the MPC-based surge pricing method also takes into account predicted performance indicators predicted using a system dynamics model for the transportation service network (for representing a performance of the transportation service network) based on surge pricing variables over a prediction horizon.

[0043] Therefore, the MPC-based surge pricing method is advantageously able to take into account the impact to future supply/demand conditions due to the surge pricing being optimized/determined. In other words, the MPC-based surge pricing method optimizes the surge pricing taking into account its impact to future supply/demand conditions. In contrast, existing PID-based surge pricing systems do not take this into account and simply determine the surge pricing based on real-time/current demand and supply signals, regardless of the surge price's impact to future supply/demand conditions. As a result, for example, for some demand and supply density imbalanced areas, with low allocation rate (AR) (e.g., high amount of transportation requests not being fulfilled) or high occupancy rate (OR) (e.g., high amount of drivers are occupied), the surge pricing may be controlled to be high based on current demand/supply conditions but doing so may dampened too much demand if subsequent supply conditions are high or subsequent demand conditions are low.

[0044] Furthermore, the PID control model in such existing systems is configured with pre-defined formulas with a large number of hyper-parameters requiring complicated and time-intensive manual tuning in an attempt to optimize its performance in dynamic price surging (e.g., regionally), which thus results in significant implementation inefficiencies and costs, as well as being prone to human tuning errors or inaccuracies. In contrast, by taking into account predicted performance indicators, predicted supply indicators and predicted demand indicators, the control parameter (e.g., surge pricing variable) of the MPC-based surge pricing method for managing supply/demand conditions in the transportation service network is advantageously optimized based on such indicators and thus is auto-adaptive based on such indicators for optimally managing supply/demand conditions in the transportation service network.

[0045] Accordingly, various example embodiments seek to overcome or address various disadvantages or deficiencies of existing PID-based surge pricing systems by providing a method (and corresponding system) for controlling the surge pricing (spatial temporal surge pricing) based on model predictive control, resulting in improved effectiveness in controlling the surge pricing for better managing supply and demand in the transportation service network. In this regard, the method is based on MPC which uses a system dynamics model for the transportation service network (a model configured to describe, or represent, a system dynamics of the transportation service network) and forecasts (or predicts) changes in dependent variables (e.g., replenishment supply volume (e.g., number of newly available drivers) in subsequent time steps and new demand volumes (e.g., number of new transportation requests) in subsequent time steps) of the modeled system given by changes in independent variables (e.g., historical demand and supply volumes) for determining an optimal solution for a control input (e.g., surge pricing) for each time step based on optimization objective and constraints. The optimization is performed based on a cost function, including one or more cost components (which may also be referred to as an objective function, including one or more objective components). As a result, the method for controlling the surge pricing according to various example embodiments possesses a number of technical advantages.

[0046] For example, the surge pricing for each time period (e.g., each time step) is determined (optimized) based on both real-time/current supply/demand signals and forecasted (or predicted) supply/demand signals (e.g., replenishment supply volume and new demand volumes in subsequent time steps). As an example, by utilizing the forecasted supply/demand signals, the method (and corresponding system) is able to take into account (e.g., foresee) upcoming peak time periods (e.g., over-demand time periods) when determining (optimizing) the surge pricing for time periods prior thereto, and thus is able to preemptively increase surge pricing in advance before the forecasted upcoming peak time periods. As a result, for example, an amount of supply (e.g., driver supply) can be conserved from certain time periods prior thereto for the upcoming peak demand time periods, which not only helps to avoid or minimize hard kick-in and sudden significant price surge when the peak demand periods arrive but also enable the amount of supply be conserved for higher value rides during the peak demand time periods. On the other hand, for example, when transitioning from peak to non-peak time periods, the method is also able to reduce surge pricing in advance, which can help to boost demand for the upcoming over-supply time periods. Therefore, the method is able to improve

or optimize the ride conversion rate (which may be referred to as the book through rate (BTR)) and the revenue generated in the transportation service network (e.g., GMV).

[0047] Furthermore, in various example embodiments, since the method (and the corresponding system) uses a system dynamics model (e.g., a transportation booking probability model, such as a book through rate (BTR) model) to represent a system dynamics of the transportation service network (e.g., complex system dynamics mechanism) as a performance indicator thereof (which may be configured as a machine learning model), the method is advantageously able to take into account the impact to future supply/demand conditions due to the surge pricing being optimized/determined (i.e., the impact to future supply/demand conditions due to decisions in control parameter (surge pricing variable) being made by the method), and thus, is able to optimize the surge pricing taking into account its impact to future supply/demand conditions. In contrast, existing PID-based surge pricing systems do not take this into account and simply determine the surge pricing based on real-time demand and supply signals, regardless of its impact to future supply/demand conditions.

[0048] In addition, since the method determines an optimal solution (optimal surge pricing) in real-time based on the predicted performance indicators and predicted supply and demand indicators, the method is able to avoid or significantly reduce the amount of manual parameter tuning as required in existing PID-based surge pricing systems. Therefore, the risk of suboptimal surge pricing due to changes, especially sudden changes, in supply/demand conditions is significantly reduced or mitigated.

[0049] To facilitate better understanding, a method of controlling a surge pricing for a geographic region (e.g., a geohash or a group of geohashes) for a transportation service network will now be described in further details according to various example embodiments of the present invention. FIG. 4 depicts a schematic drawing showing an overview of the method 400 of controlling the surge pricing for a geographic region. In various example embodiments, the method 400 may be performed for each geographic region (e.g., in parallel) to determine (or optimize), for each time step of a series of time steps in turn (e.g., each time step successively for as long as the method 400 is being executed to control the surge pricing), the surge pricing for the time step for the geographic region. Accordingly, for clarity and conciseness, the method 400 of controlling the surge pricing may simply be described with respect to one geographic region since the method 400 may be applied to each geographic region to control the surge pricing for each geographic region in the same or similar manner. Furthermore, for simplicity

and without limitation or loss of generality, the geographic region may correspond to one geohash and each time step may be a time period or time window of 1 minute.

[0050] For illustration purpose, FIG. 5 depicts a schematic drawing illustrating two consecutive time steps (which may also be referred to as time windows) T-1 and T and three consecutive timestamps t-1, t and t+1, where timestamp t-1 corresponds to a start of time step T-1, timestamp t corresponds to an end of time step T-1 or a start of time step T and timestamp t+1 corresponds to an end of time step T.

[0051] The method 400 comprises, for each time step of a series of time steps: optimizing, using a cost function for the transportation service network, a sequence of surge pricings for a sequence of time steps, including the time step (i.e., the above-mentioned time step of the series of time steps) as a first time step thereof, over a prediction horizon for the geographic region based on a sequence of predicted performance indicators, a sequence of predicted supply indicators and a sequence of predicted demand indicators for the sequence of time steps with respect to the geographic region; and applying (or implementing) a first surge pricing of the sequence of surge pricings optimized for the first time step to the geographic region for the first time step. In this regard, the cost function is parameterized by a sequence of surge pricing variables for the sequence of time steps for optimization to obtain the sequence of surge pricings for the sequence of time steps. Furthermore, the sequence of predicted performance indicators is predicted using a system dynamics model for the transportation service network based on the sequence of surge pricing variables for the sequence of time steps.

[0052] In various example embodiments, the method 400 of controlling the surge pricing is performed with respect to each geographic region respectively, and thus, the sequence of predicted performance indicators, the sequence of predicted supply indicators and the sequence of predicted demand indicators for the sequence of time steps are with respect to (e.g., dedicated to) the respective geographic region (e.g., one geohash), and not shared with other geographic regions (e.g., other geohashes). In various example embodiments, the cost function may comprise a number of cost components, including a first cost component relating to revenue generated, a second cost component relating to supply utilization and a third cost component relating to demand pressure with respect to the geographic region. As an illustrative example, the cost function (or objective function) may be configured to maximize the revenue generated (e.g., gross merchandise value (GMV)) and supply utilization (e.g., driver supply utilization) with respect to the geographic region and minimize demand pressure (e.g., driver over-demand) with respect to the geographic region.

[0053] Therefore, in various example embodiments, the cost function comprises a first cost component configured to, for each of the sequence of surge pricing variables, determine a revenue indicator with respect to the geographic region for the time step corresponding to the surge pricing variable based on the surge pricing variable. In various example embodiments, the cost function further comprises: a second cost component configured to, for each of the sequence of surge pricing variables, determine a supply utilization indicator with respect to the geographic region for the time step corresponding to the surge pricing variable based on the surge pricing variable, and a third cost component configured to, for each of the sequence of surge pricing variables, determine a demand pressure indicator with respect to the geographic region for the time step corresponding to the surge pricing variable based on the surge pricing variable.

[0054] As an illustrative example, the revenue indicator with respect to a geographic region for a time step is the gross merchandise value (GMV) which may be defined as follows:

$$GMV = AvgNSNSFare * Surge * UCP * BTR(surge)$$

(Equation 1)

where *AvgNSNSFare* denotes an average Non-Surge-Non-Surcharge (NSNS) fare (which corresponds to the base fare without any surcharge or surge applied; the *Surge* denotes the surge pricing variable to be optimized; the *UCP* denotes the amount of unique check price instances (e.g., corresponding to the number of unique transportation requests received) and the *BTR(surge)* denotes a book through rate (BTR) prediction model configured to predict or estimate the rate (or probability) which a transportation request with respect to the geographic region is converted into a transportation booking (which may also be referred to the transportation request conversion rate, such as the ride conversion rate) based on the *Surge* variable (which in the illustrative example corresponds to the system dynamics model for predicting the performance indicator (the BTR)) for the time step with respect to the geographic region. For example, transportation requests may each have associated therewith a pick-up location, a drop-off location and a pick-up time for a transportation. The BTR prediction model may thus be configured to predict or estimate the probability which a transportation request having associated therewith a pick-up location belonging to (within) the geographic region and the pick-up time belonging to (within) the time step is converted into a transportation booking based on the *Surge* variable. In various example embodiments, the number of transportation requests for a time step with respect to a geographic region may refer to the number of unique transportation requests (i.e., from unique users (or riders)) for the time step with respect to the

geographic region to avoid repeated transportation requests from same users which does not actually correspond to, or result in, additional demand. As a result, a more accurate indication or measure of the level of demand can be obtained. For example, the *AvgNSNSFare* may be determined as desired or as appropriate such as the average NSNS fare per city/day or per geohash/hour. For example, the UCP (corresponding to the predicted demand indicator $\hat{z}_N$) for a time step N may be determined or predicted based on a demand prediction model, such as but not limited to, a time-series forecasting model configured to predict the UCP for a time step (e.g., one minute) using a certain period of historical data point (e.g., the past 60 minutes of data points). As an illustrative example and without loss of generality, the time-series forecasting model may be an LSTM (Long Short-Term Memory) model.

[0055] In various example embodiments, the second cost component is configured to determine the supply utilization indicator with respect to the geographic region for the time step corresponding to the surge pricing variable based on an available supply volume (e.g., driver supply volume) at a start of the time step, the predicted supply indicator for the time step and a filled supply indicator for the time step. In this regard, the predicted supply indicator for the time step corresponds to a predicted replenishment supply volume (e.g., replenishment driver supply volume) for the time step, and the filled supply indicator for the time step corresponds to a supply volume filled by demand volume for the time step.

[0056] In various example embodiments, the third cost component is configured to determine the demand pressure indicator with respect to the geographic region for the time step corresponding to the surge pricing variable based on a pending demand volume at the start of the time step, the predicted demand indicator for the time step and the filled supply indicator for the time step. In this regard, the predicted demand indicator for the time step corresponds to a predicted amount of transportation requests for the time step.

[0057] In various example embodiments, the filled supply indicator for the time step is determined based on the available supply volume at the start of the time step, the predicted supply indicator for the time step, the pending demand volume at the start of the time step, the predicted demand indicator for the time step and the predicted performance indicator for the time step.

[0058] In various example embodiments, the system dynamics model comprises a transportation booking probability model (e.g., the above-mentioned BTR prediction model) configured to predict a probability which a transportation request with respect to the geographic region is converted into a transportation booking based on a surge pricing variable, and the

predicted performance indicator for the time step corresponds to a predicted probability which a transportation request using the surge pricing of the surge pricing variable corresponding to the time step and with respect to the geographic region is converted into a transportation booking.

5 **[0059]** In various example embodiments, the above-mentioned optimizing, using the cost function for the transportation service network, the sequence of surge pricings for the sequence of time steps, comprises maximizing the cost function with respect to the sequence of surge pricing variables for the sequence of time steps over the prediction horizon, subject to one or more constraints relating to the sequence of surge pricing variables.

10 **[0060]** To facilitate a better understanding of the present invention and without limitation or loss of generality, an illustrative example implementation for determining the surge pricing for a geographic region for a transportation service network will now be described according to various example embodiments of the present invention. It will be appreciated by a person skilled in the art that the present invention is not limited to such an illustrative example
15 implementation and various changes or modifications may be made as desired or as appropriate without going beyond the scope of the present invention.

[0061] In the illustrative example implementation, the cost function for the transportation service network for optimization to obtain a sequence of surge pricings for a sequence of time steps over a prediction horizon for a geographic region may be expressed as follows:

$$\begin{aligned}
 & \text{20} \quad \text{Max} \sum_{T=1}^N \gamma^T (F * a_t * M_T - \lambda * S_{t+1} - \mu * D_{t+1}^p) \\
 & \text{s. t.} \quad S_{t+1} = \eta (S_t + \hat{S}_T^R - M_T), \forall t = 1, 2, \dots, N \\
 & \quad D_{t+1}^p = D_t^p + \hat{z}_T * h(a_t) - M_T, \forall t = 1, 2, \dots, N \\
 & \quad M_T = l(a_t) * \min (S_t + \hat{S}_T^R, D_t^p + \hat{z}_T * h(a_t)), \forall t = 1, 2, \dots, N \\
 & \quad u_t = a_t - a_{t-1}, \forall t = 1, 2, \dots, N \\
 & \text{25} \quad u_{lb} \leq u_t \leq u_{ub}, \forall t = 1, 2, \dots, N \\
 & \quad a_{lb} \leq a_t \leq a_{ub}, \forall t = 1, 2, \dots, N
 \end{aligned}
 \tag{Equation 2}$$

where t denotes a time step of a sequence of time steps ($t = 1, 2, \dots, N$) over a prediction horizon (N time steps), T denotes a length or an interval of the time step and a_t denotes a surge
30 pricing variable of a sequence of surge pricing variables (a_1 to a_N) for the sequence of time steps.

[0062] Accordingly, in the illustrative example implementation, it can be seen that, for each time step t of a series of time steps (e.g., each time step successively), a sequence of surge pricings (values of a_t to a_N) for a sequence of time steps ($t = 1, 2, \dots, N$) are optimized over

the prediction horizon N using the cost function based on a sequence of predicted performance indicators (values of $h(a_1)$ to $h(a_N)$), a sequence of predicted supply indicators ($\hat{S}_1^R$ to $\hat{S}_N^R$) and a sequence of predicted demand indicators ($\hat{Z}_1$ to $\hat{Z}_N$) for the sequence of time steps with respect to the geographic region. In this regard, the cost function is parameterized by the sequence of surge pricing variables (a_t to a_N) for the sequence of time steps for optimization to obtain the sequence of surge pricings for the sequence of time steps, respectively. Furthermore, the sequence of predicted performance indicators (values of $h(a_1)$ to $h(a_N)$) is predicted using a system dynamics model $h(a_t)$ for the transportation service network based on the sequence of surge pricing variables (a_1 to a_N) for the sequence of time steps.

10 **[0063]** In the illustrative example implementation, the first cost component is defined by $F * a_t * M_T$, where F denotes the average base fare (e.g., the average non-surcharge non-surge (NSNS) fare) and M_T denotes the filled supply indicator for the time step (which corresponds to a supply volume which has been filled (fulfilled or taken up) by demand volume in the time step). In the illustrative example, the filled supply indicator M_T for a time step t is determined
 15 based on the available supply volume S_t at the start of the time step t , the predicted supply indicator $\hat{S}_T^R$ for the time step t , the pending demand volume D_t^p at the start of the time step t , the predicted demand indicator $\hat{Z}_T$ for the time step t and the predicted performance indicator $h(a_t)$ for the time step t . For example, as shown in Equation (2) above, the filled supply indicator M_T for the time step t may be determined or predicted based on the minimum of a
 20 predicted fulfillment capacity (i.e., predicted overall supply volume, which may be defined by $S_t + \hat{S}_T^R$) and a predicted overall demand volume (e.g., defined by $D_t^p + \hat{Z}_T * h(a_t)$) for the time step t . In this regard, the overall demand volume for the time step t may be predicted based on the predicted demand indicator $\hat{Z}_T$ and the predicted performance indicator $h(a_t)$ (i.e., $(\hat{Z}_T * h(a_t))$, which corresponds to the predicted amount of transportation requests converted into
 25 transportation bookings, or simply referred to as predicted converted transportation bookings (e.g., the predicted number of rides allocated)) as well as the pending demand volume D_t^p at the time step t . Accordingly, the filled supply indicator M_T is the predicted overall demand volume if it is lower than the predicted fulfillment capacity since the predicted overall demand volume can all be taken up by the predicted fulfillment capacity. On the other hand, if the predicted
 30 overall demand volume is more than the predicted fulfillment capacity, the filled supply indicator M_T equals the predicted fulfillment capacity since it cannot exceed the predicted overall demand volume. In various example embodiments, as shown in Equation (2) above, the

value of the filled supply indicator M_T may be weighted by an estimated acceptance rate (EAR) (e.g., driver EAR) determined by an EAR function $l(a_t)$ based on the surge pricing variable a_t . As an illustrative example and without loss of generality, the EAR function may be a simple linear function such as $l(a_t) = m \times a_t + c$ where the parameters m and c may be obtained from regression based on the historical EAR data plotted against surge pricing.

[0064] In the illustrative example implementation, the second cost component is defined by $\lambda * S_{t+1}$, where S_{t+1} denotes the available supply volume at the start of the next time step $t+1$ (or equivalently at the end of the time step t). As described above, the second cost component determines the supply utilization indicator with respect to the geographic region for the time step t . In this regard, determining the available supply volume S_{t+1} at the start of the next time step $t+1$ corresponds to determining the supply utilization for the time step t , for example, since the lesser the available supply volume S_{t+1} at the start of the next time step $t+1$, the more the supply has been utilized at the time step t . In the illustrative example implementation, the supply utilization indicator for the time step t corresponding to the surge pricing variable a_t is determined based on an available supply volume S_t at the start of the time step t , the predicted supply indicator $\hat{S}_T^R$ for the time step t and the filled supply indicator M_T for the time step t . As described above, the filled supply indicator M_T for the time step t is determined based on the predicted performance indicator $h(a_t)$ for the time step t . For example, as shown in Equation (2) above, the supply utilization indicator for the time step t corresponding to the surge pricing variable a_t is based on the addition of the available supply volume S_t at the start of the time step t and the predicted supply indicator $\hat{S}_T^R$ for the time step t , minus the filled supply indicator M_T for the time step t . Since the supply utilization indicator is determined for each surge pricing variable of the sequence of surge pricing variables for the sequence of time steps ($t = 1, 2, \dots, N$) over the prediction horizon N , the second cost component thus describes or represents the dynamics of supply changes across a sequence of time steps ($t = 1, 2, \dots, N$). In various example embodiments, the supply utilization indicator may be weighted (or modified) by an available supply penalty parameter (or coefficient or factor) λ . For example, the available supply penalty parameter λ enables the setting of a penalty weight on supply retention (or the ‘leftover supply’) as appropriate. For example, if it is desired for the method to control the surge pricing to utilise as many drivers’ supply as possible, then the available supply penalty parameter λ may be set to a relatively higher value which may then lead to a lower average surge pricing. For example, the value of the available supply penalty parameter λ (together with the other configurable parameters, e.g., gravity coefficient on surge) may be determined

through simulations and choosing a set of values of the parameters that maximises the objective, for instance, the simulated number of transportation bookings. In addition, the supply utilization indicator may be further weighted by a supply retention parameter (or coefficient or factor) η . For example, the supply retention parameter η may be set as appropriate to address the competition that a portion of supply may switch to a competitor with certain probability. For example, the supply retention parameter η enables the discount of the supply as appropriate from one minute to the next. For example, some of the initially available drivers in an area may not be available in the area the next minute if they pick up jobs from a competitor platform or they simply drove away from the area. For example, the supply retention parameter λ may be derived from historical data when deriving the percentage of drivers who continue to be available in the same area from one minute to the next. For example, this is different from the above-mentioned available supply penalty parameter λ as the available supply penalty parameter λ is determined by the platform and it is subjective how much it is desired to penalise any leftover drivers that were not allocated a booking. Meanwhile, the supply retention parameter λ is derived from historical data based on drivers' behaviour, e.g., whether they tend to stay in the area, drive away from the area or just balk from the platform.

[0065] In the illustrative example implementation, the third cost component is defined by $\mu * D_{t+1}^p$, where D_{t+1}^p denotes the pending demand volume at the start of the next time step $t+1$ (or equivalently at the end of the time step t). As described above, the third cost component determines the demand pressure indicator with respect to the geographic region for the time step t . In this regard, determining the pending demand volume at the start of the next time step $t+1$ corresponds to determining the demand pressure for the time step t , for example, since the higher the pending demand volume D_{t+1}^p at the start of the next time step $t+1$, the more the demand pressure at the time step t . In the illustrative example implementation, the demand pressure indicator for the time step t corresponding to the surge pricing variable a_t is determined based on a pending demand volume D_t^p at the start of the time step t , the predicted demand indicator $\hat{z}_T$ for the time step t and the filled supply indicator M_T for the time step t . For example, as shown in Equation (2) above, the demand pressure indicator for the time step t is based on the addition of the pending demand volume D_t^p at the start of the time step t and the above-mentioned predicted new demand volume ($\hat{z}_T * h(a_t)$) for the time step t minus the filled supply indicator M_T for the time step t . Since the demand pressure indicator is determined for each surge pricing variable of the sequence of surge pricing variables for the sequence of time steps ($t = 1, 2, \dots, N$) over the prediction horizon N , the third cost component thus

describes or represents the dynamics of demand changes across a sequence of time steps t ($t = 1, 2, \dots, N$), and more particularly, the update of pending demands that is carried over to the next time step. In various example embodiments, the demand pressure indicator may be weighted (or modified) by a pending demand penalty parameter (or coefficient or factor) μ . For example, the pending demand penalty parameter μ enables the setting of a penalty weight on the pending demand, which is the demand that has not been fulfilled (or taken up). For example, if it is desired for the method to control the surge pricing to fulfill as many passengers' demands as possible, then the pending demand penalty parameter μ may be set to a relatively higher value which may then lead to a higher average surge pricing. For example, the value of the pending demand penalty parameter μ (together with the other configurable parameters, e.g., available supply penalty parameter λ and gravity coefficient on surge) may be determined through simulations and choosing a set of values that maximises the objective, for instance, the simulated number of transportation bookings.

[0066] In the illustrative example implementation, as shown in Equation (2) above, the available supply volume S_{t+1} at the start of the next time step $t+1$, the pending demand volume D_{t+1}^p at the start of the next time step $t+1$ (e.g., carried over from the previous time step t) and the filled supply indicator M_T may be defined as constraints to the cost function. Additional surge pricing constraints for bounding the value of each surge pricing variable of the sequence of surge pricing variables may be provided. In this regard, as shown in Equation (2) above, a first surge pricing constraint may be configured for bounding the value of each surge pricing variable of the sequence of surge pricing variables and a second surge pricing constraint may be configured for bounding a change or difference in the values of two consecutive surge pricing variables of the sequence of surge pricing variables for the sequence of time steps ($t = 1, 2, \dots, N$). For example, the first surge pricing constraint may be defined as $a_{lb} \leq a_t \leq a_{ub}$, $\forall t = 1, 2, \dots, N$ and the second surge pricing constraint may be defined as $u_{lb} \leq u_t \leq u_{ub}$, $\forall t = 1, 2, \dots, N$, where $u_t = a_t - a_{t-1}$, a_{lb} and a_{ub} denote a lower bound and an upper bound, respectively, for the value of the surge pricing variable, and u_{lb} and u_{ub} denote a lower bound and an upper bound, respectively, for the change in the values of two consecutive surge pricing variables.

[0067] Furthermore, as shown in Equation (2), a time discount parameter γ^T for the sequence of time steps ($t = 1, 2, \dots, N$) may be applied to the cost function for optimization. For example, the time discount parameter γ^T enables the discount of the future revenue or costs as appropriate (e.g., as they rely on future predictions which may be less accurate). For example,

the time discount parameter γ^T may be set to 1 if no discounting is implemented or may be set to a value less than 1 (e.g., 0.95) to implement discounting as appropriate.

[0068] Therefore, in the illustrative example implementation, inputs for the optimization (e.g., to the surge pricing optimization module 406 shown in FIG. 4) to obtain the sequence of surge pricings for the sequence of time steps may include:

- average NSNS fare: F
- value produced by the BTR function 424: $h(\cdot)$
- initial surge pricing at time 0: a_0
- available supply at the start of time step 1 (first time step): S_1 (e.g., included in the fulfilment capacity real-time signal 432 shown in FIG. 4)
- pending demand at the start of time step 1: D_1^p (e.g., included in the online UCP real-time signal 434 shown in FIG. 4)
- a sequence of forecasted/predicted number of transportation requests (e.g., UCP) for the sequence of time steps: $\hat{z}_1, \hat{z}_2, \dots, \hat{z}_N$ (e.g., included in the UCP session forecasting signal 446 shown in FIG. 4)
- a sequence of forecasted/predicted replenishment supply volume for the sequence of time steps: $\hat{S}_1^R, \hat{S}_2^R, \dots, \hat{S}_N^R$ (e.g., included in the fulfilment capacity forecasting signal 444)
- lower bound and upper bound for surge pricing change between two consecutive surge pricing variables: u_{lb}, u_{ub}

For example, the sequence of replenishment supply volumes $\hat{S}_1^R, \hat{S}_2^R, \dots, \hat{S}_N^R$ for the sequence of time steps may be predicted based on a time-series forecasting model, for example, using a certain period of historical data point (e.g., the past 60 minutes of data points). As an illustrative example and without loss of generality, the time-series forecasting model may be an LSTM (Long Short-Term Memory) model. Similarly, the number of transportation requests (e.g., the UCPs denoted as $\hat{z}_1, \hat{z}_2, \dots, \hat{z}_N$) for the sequence of time steps may be predicted based on a time-series forecasting model, for example, also using a certain period of historical data point (e.g., the past 60 min data points), which for example may be an LSTM model. In various example embodiments, inputs for the optimization may further include a unique booking real-time signal 436 providing a count on the number of transportation requests that has been converted into transportation bookings for time step 1 (first time step).

[0069] In addition, configurable parameters for the optimization (e.g., to the surge pricing optimization module 406 shown in FIG. 4) to obtain the sequence of surge pricings for the sequence of time steps may include:

- time discount factor γ
- 5 • available supply penalty coefficients λ
- pending demand penalty coefficients μ
- supply retention coefficient η

[0070] Furthermore, outputs from the optimization may include:

- the first surge pricing optimized for the first time step to apply to the geographic region
10 for the first time step: a_1
- the predicted new demand volume for the first time step: $\hat{z}_1 * h(a_1)$

[0071] As described hereinbefore, in various example embodiments, the system dynamics model may comprise a BTR model 424 to represent a system dynamics of the transportation service network as a performance indicator thereof. In various example embodiments, as shown
15 in FIG. 4, the system dynamics model may include the BTR model 424 and/or a driver EAR model 426. In other words, for the optimization, the BTR model 424 and/or the driver EAR model 426 may represent the system dynamics model for providing one or more performance indicators/measures of the transportation service network. In this regard, the EAR model 426 is configured to determine or predict the probability of drivers accepting transportation requests
20 given a certain surge pricing (or surge value). In various example embodiments, as shown in FIG. 4, the city level surge stream block 452 may be configured to map every city-geohash6-minute to a surge pricing (or surge value), which may simply be a JSON or dictionary format. The surge RT storage block 454 is a database (e.g., the real-time data platform Redis) configured to store the above mapping. The surge API block 456 is a surge interface having an
25 external API configured to receive a query for a surge pricing and obtain the optimized surge pricing.

[0072] It will be appreciated by a person skilled in the art that the optimization to obtain the sequence of surge pricings may involve additional inputs and/or outputs as desired or as appropriate without going beyond the scope of the present invention.

30 [0073] As shown in FIG. 4, in the illustrative example implementation, real-time supply and demand signals 432, 434, 436 may be utilized as inputs to the optimization, for example, with granularity of geohash6 (corresponding to the geographic region) and one-minute level (corresponding to the length or interval of each time instance t). In this regard, real-time signals

utilized may include a supply counting signal (included in the fulfilment capacity signal 432), a demand counting signal (including the unique booking signal 436) and an online UCP signal 434. The supply counting signal indicates the available supply S_1 at the start of time step 1 (first time step). For example, this available supply S_t (e.g., available drivers) at time stamp t , together with the new replenished supply (e.g., new driver supply) during the time step (having a time period T), that is, $S_t + \hat{S}_T^R$, can be considered as the fulfilment capacity for the time step t , which corresponds to the total amount of supply available for use in the time period T . In various example embodiments, the new replenished supply (e.g., new driver supply) for the time step t may be computed from the fulfilment capacity and the available supply. In addition, the demand counting signal includes the unique booking signal 436 for providing a count on the number of transportation requests that has been converted into transportation bookings for a time step. In addition, the online UCP signal 434 is provided for indicating the pending demand D_1^p at the start of time step 1, which may then be used to determine the overall demand volume for time step 1. Accordingly, the online UCP signal 434 indicates a count of the number of unique transportation requests (which may or may not result in a transportation booking) while the unique booking signal 436 indicates a count of the number of transportation bookings.

[0074] As shown in FIG. 4, in the illustrative example implementation, forecasted or predicted supply and demand signals are also utilized as inputs to the optimization, which may also be with granularity of geohash6 and one-minute level (corresponding to the length of each time instance). For the forecasted demand signals, there may be two types of demand, namely, intended demand (corresponding to the number of unique transportation requests received, which may be referred to herein as the UCP) and converted demand (e.g., corresponding to the predicted new demand volume (e.g., $\hat{z}_T * h(a_t)$)). In this regard, since the number of transportation bookings is highly influenced by the surge pricing, various example embodiments employ a UCP forecasting model to forecast the amount of transportation requests (e.g., included in the UCP session forecasting signal 446 shown in FIG. 4), which may then be used with the BTR model $h(a_t)$ to determine (or predict) the number of transportation requests converted into transportation bookings based on a certain surge pricing (which may also be referred to as converted transportation bookings (e.g., the number of rides allocated)).

For example, the UCP forecasting model may be employed for each of 2W and 4W wheel types in the transportation supply pool in each city and update every predetermined time period (e.g., 1 minute). The granularity may be geohash6-min level and the forecasting horizon may be a

predefined number of minutes. As described hereinbefore, for example, the UCP forecasting model may be implemented based on a time-series forecasting model.

[0075] For the supply forecasting signals, various example embodiments seek to predict the new supply (e.g., drivers) coming from various resources (which may be referred to as replenishment supply), such as drivers switching from offline to online and drivers with current jobs dropping off at the geohash and will become available soon, and so on. In various example embodiments, the real time replenishment supply signal may be calculated from the fulfillment capacity and available supply signals as described above, which may then be fed into the forecasting model to generate the replenishment supply forecasting signal. In particular, various example embodiments predict the fulfillment capacity signal 444 and convert the predicted fulfillment capacity signal 444 to predicted replenished supply signal. For example, this may be performed by subtracting the latest available supply from the predicted fulfillment capacity signal 444 to obtain predicted replenishment supply signal. Then, the predicted replenishment supply signal may be provided to the optimisation module 406. For the available supply (e.g., driver) signal, since it may be influenced by the number of transportation bookings converted, it may be recalculated instead of time series forecasting during the MPC control process. In this regard, the recalculated available supply for the future timestamps may be considered as intermediate variables that are calculated based on the forecasted demand and supply signals, together with how much to surge which is the decision variable according to various example embodiments of the present invention. Similarly, the granularity may be geohash6-min level, and the forecasting horizon may be a predefined number of minutes. For example, the forecasting model may be employed for each of 2W and 4W wheel types in the transportation supply pool and update every predetermined time period (e.g., 1 minute).

[0076] In the illustrative example implementation, the BTR prediction model 424 is performed on a geohash-min level to provide the performance indicator (i.e., BTR) for each 1-minute time period with respect to the geohash. For example, various example embodiments provide two approaches for building the BTR model 424. A first approach is a simple regression function between the surge pricing variable and the BTR, focusing on causal effect and computation speed. A second approach is a two-step BTR model, a first step focusing on accuracy of BTR prediction with deep neural network as the baseline of BTR for geohash-minute, and a second step being performed by a delta BTR model focusing on causal effect given by different surge levels. By doing this, a BTR prediction model with acceptable accuracy can be achieved, while being causally correct at the same time. It will be appreciated by a person

skilled in the art that the present invention is not limited to any particular approach in forming the BTR prediction model 424 as long as a BTR is estimated or predicted based on the surge pricing variable.

[0077] Accordingly, a MPC-based surge pricing method (and corresponding MPC-based surge pricing system) is provided for solving an optimization problem for each geographic region (e.g., each geohash6) to determine an optimal surge pricing for each upcoming time step (e.g., each 1-minute). Advantageously, the MPC-based surge pricing method is able to preemptively adjust surge pricing by considering both real time supply and demand signals and their forecasted values in the future. As the optimization problem can be solved online, this approach has a very high level of auto-adaptive capability, which is able to adapt to fast changing market conditions.

[0078] For illustration purpose only and without limitation or loss of generality, a simple example of a 2 period MPC optimization problem ($N = 2$) based on Equation (2) above is now provided. For example, for the two time periods, the forecasted/predicted UCP is (30, 40), the forecasted/predicted replenishment supply is 10, 8, time discount factor $\gamma = 1$, and the average NSNS Fare $F = 10$. For simplicity, assume the BTR function is:

$$h(x) = -0.1 * x + 0.5, 1 \leq x \leq 3 \quad (\text{Equation 3})$$

[0079] Furthermore, the initial surge $a_0 = 1.5$, the available supply $S_1 = 2$, and the pending demand $D_1^p = 0$. The surge change lower bound and upper bound are $u_{lb} = -0.5, u_{ub} = 0.5$, and the surge lower bound and upper bound are $a_{lb} = 1, a_{ub} = 3$. For configurable parameters, $\lambda = 20, \mu = 30, \eta = 1$. Based on Equation (2) above, the optimization problem may thus be formulated as:

$$\begin{aligned} & \max 10 * a_1 * M_1 + 10 * a_2 * M_2 - 20 * (S_2 + S_3) - 30 * (D_2^p + D_3^p) \\ & \text{s.t. } S_2 = 2 + 10 - M_1 \\ & \quad S_3 = S_2 + 8 - M_2 \\ & \quad D_2^p = 0 + 30 * h(a_1) - M_1 \\ & \quad D_3^p = D_2^p + 40 * h(a_2) - M_2 \\ & \quad M_1 = \min(12, 30 * h(a_1)) \\ & \quad M_2 = \min(S_2 + 8, D_2^p + 40 * h(a_2)) \\ & \quad 1 \leq a_1 \leq 2 \\ & \quad a_1 - 0.5 \leq a_2 \leq a_1 + 0.5 \\ & \quad a_2 \geq 1 \end{aligned} \quad (\text{Equation 4})$$

[0080] The optimization problem may be solved according to any techniques known in the art as desired or as appropriate. In particular, it will be appreciated by a person skilled in the art

that a variety of techniques are known in the art for solving optimization problems (e.g., SLSQP (Sequential Least Squares Programming optimizer) algorithm, BFGS (Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm and so on), and thus, it is not necessary to describe such techniques herein for clarity and conciseness.

- 5 **[0081]** While embodiments of the invention have been particularly shown and described with reference to specific embodiments, it should be understood by those skilled in the art that various changes in form and detail may be made therein without departing from the scope of the invention as defined by the appended claims. The scope of the invention is thus indicated by the appended claims and all changes which come within the meaning and range of
- 10 equivalency of the claims are therefore intended to be embraced.

CLAIMS

1. A method of controlling a surge pricing for a geographic region for a transportation service network using at least one processor, the method comprising, for each time step of a series of time steps:
 - optimizing, using a cost function for the transportation service network, a sequence of surge pricings for a sequence of time steps, including the time step as a first time step thereof, over a prediction horizon for the geographic region based on a sequence of predicted performance indicators, a sequence of predicted supply indicators and a sequence of predicted demand indicators for the sequence of time steps with respect to the geographic region; and
 - applying a first surge pricing of the sequence of surge pricings optimized for the first time step to the geographic region for the first time step, wherein
 - the cost function is parameterized by a sequence of surge pricing variables for the sequence of time steps for optimization to obtain the sequence of surge pricings for the sequence of time steps, and
 - the sequence of predicted performance indicators is predicted using a system dynamics model for the transportation service network based on the sequence of surge pricing variables for the sequence of time steps.
2. The method according to claim 1, wherein the cost function comprises a first cost component configured to, for each of the sequence of surge pricing variables, determine a revenue indicator with respect to the geographic region for the time step corresponding to the surge pricing variable based on the surge pricing variable.
3. The method according to claim 2, wherein the cost function further comprises:
 - a second cost component configured to, for each of the sequence of surge pricing variables, determine a supply utilization indicator with respect to the geographic region for the time step corresponding to the surge pricing variable based on the surge pricing variable, and
 - a third cost component configured to, for each of the sequence of surge pricing variables, determine a demand pressure indicator with respect to the geographic region for the time step corresponding to the surge pricing variable based on the surge pricing variable.
4. The method according to claim 3, wherein

the second cost component is configured to determine the supply utilization indicator with respect to the geographic region for the time step corresponding to the surge pricing variable based on an available supply volume at a start of the time step, the predicted supply indicator for the time step and a filled supply indicator for the time step,

5 the predicted supply indicator for the time step corresponds to a predicted replenishment supply volume for the time step, and

 the filled supply indicator for the time step corresponds to a supply volume filled by demand volume for the time step.

10 5. The method according to claim 4, wherein

 the third cost component is configured to determine the demand pressure indicator with respect to the geographic region for the time step corresponding to the surge pricing variable based on a pending demand volume at the start of the time step, the predicted demand indicator for the time step and the filled supply indicator for the time step, and

15 the predicted demand indicator for the time step corresponds to a predicted amount of transportation requests for the time step.

6. The method according to claim 5, wherein

 the filled supply indicator for the time step is determined based on the available supply
20 volume at the start of the time step, the predicted supply indicator for the time step, the pending demand volume at the start of the time step, the predicted demand indicator for the time step and the predicted performance indicator for the time step.

7. The method according to claim 1, wherein

25 the system dynamics model comprises a transportation booking probability model configured to predict a probability which a transportation request with respect to the geographic region is converted into a transportation booking based on a surge pricing variable, and

 the predicted performance indicator for the time step corresponds to a predicted probability which a transportation request using the surge pricing of the surge pricing variable
30 corresponding to the time step and with respect to the geographic region is converted into a transportation booking.

8. The method according to claim 1, wherein said optimizing, using the cost function for the transportation service network, the sequence of surge pricings for the sequence of time steps, comprises maximizing the cost function with respect to the sequence of surge pricing variables for the sequence of time steps over the prediction horizon, subject to one or more constraints relating to the sequence of surge pricing variables.
9. The method according to claim 8, wherein the one or more constraints comprises:
 - a first constraint configured for bounding a value of each surge pricing variable of the sequence of surge pricing variables; and
 - a second constraint configured for bounding a difference between the values of two consecutive surge pricing variables of the sequence of surge pricing variables for the sequence of time steps.
10. The method according to claim 1, wherein said optimizing, using the cost function for the transportation service network, the sequence of surge pricings for the sequence of time steps is based on model predictive control.
11. The method according to claim 1, wherein
 - the transportation service network is configured to provide ride-hailing services and/or delivery services,
 - the sequence of predicted supply indicators is with respect to driver supply,
 - the sequence of predicted demand indicators is with respect to driver demand, and
 - the geographic region is a geohash.
12. A system for controlling a surge pricing for a geographic region for a transportation service network, the system comprising:
 - at least one memory; and
 - at least one processor communicatively coupled to the at least one memory and configured to, for each time step of a series of time steps:
 - optimize, using a cost function for the transportation service network, a sequence of surge pricings for a sequence of time steps, including the time step as a first time step thereof, over a prediction horizon for the geographic region based on a sequence of predicted

performance indicators, a sequence of predicted supply indicators and a sequence of predicted demand indicators for the sequence of time steps with respect to the geographic region; and

apply a first surge pricing of the sequence of surge pricings optimized for the first time step to the geographic region for the first time step,

5 wherein the cost function is parameterized by a sequence of surge pricing variables for the sequence of time steps for optimization to obtain the sequence of surge pricings for the sequence of time steps, and

 the sequence of predicted performance indicators is predicted using a system dynamics model for the transportation service network based on the sequence of surge pricing variables
10 for the sequence of time steps.

13. The system according to claim 12, wherein the cost function comprises a first cost component configured to, for each of the sequence of surge pricing variables, determine a revenue indicator with respect to the geographic region for the time step corresponding to the
15 surge pricing variable based on the surge pricing variable.

14. The system according to claim 13, wherein the cost function further comprises:
 a second cost component configured to, for each of the sequence of surge pricing variables, determine a supply utilization indicator with respect to the geographic region for the
20 time step corresponding to the surge pricing variable based on the surge pricing variable, and
 a third cost component configured to, for each of the sequence of surge pricing variables, determine a demand pressure indicator with respect to the geographic region for the time step corresponding to the surge pricing variable based on the surge pricing variable.

25 15. The system according to claim 14, wherein
 the second cost component is configured to determine the supply utilization indicator with respect to the geographic region for the time step corresponding to the surge pricing variable based on an available supply volume at a start of the time step, the predicted supply indicator for the time step and a filled supply indicator for the time step,

30 the predicted supply indicator for the time step corresponds to a predicted replenishment supply volume for the time step, and

 the filled supply indicator for the time step corresponds to a supply volume filled by demand volume for the time step.

16. The system according to claim 15, wherein
the third cost component is configured to determine the demand pressure indicator with
respect to the geographic region for the time step corresponding to the surge pricing variable
5 based on a pending demand volume at the start of the time step, the predicted demand indicator
for the time step and the filled supply indicator for the time step, and
the predicted demand indicator for the time step corresponds to a predicted amount of
transportation requests for the time step.
- 10 17. The system according to claim 16, wherein
the filled supply indicator for the time step is determined based on the available supply
volume at the start of the time step, the predicted supply indicator for the time step, the pending
demand volume at the start of the time step, the predicted demand indicator for the time step
and the predicted performance indicator for the time step.
- 15 18. The system according to claim 12, wherein
the system dynamics model comprises a transportation booking probability model
configured to predict a probability which a transportation request with respect to the geographic
region is converted into a transportation booking based on a surge pricing variable, and
20 the predicted performance indicator for the time step corresponds to a predicted
probability which a transportation request using the surge pricing of the surge pricing variable
corresponding to the time step and with respect to the geographic region is converted into a
transportation booking.
- 25 19. The system according to claim 12, wherein said optimize, using the cost function for
the transportation service network, the sequence of surge pricings for the sequence of time
steps, comprises maximizing the cost function with respect to the sequence of surge pricing
variables for the sequence of time steps over the prediction horizon, subject to one or more
constraints relating to the sequence of surge pricing variable.
- 30 20. The system according to claim 19, wherein the one or more constraints comprises:
a first constraint configured for bounding a value of each surge pricing variable of the
sequence of surge pricing variables; and

a second constraint configured for bounding a difference between the values of two consecutive surge pricing variables of the sequence of surge pricing variables for the sequence of time steps.

- 5 21. The system according to claim 12, wherein said optimize, using the cost function for the transportation service network, the sequence of surge pricings for the sequence of time steps is based on model predictive control.
22. The system according to claim 12, wherein
10 the transportation service network is configured to provide ride-hailing services and/or delivery services,
the sequence of predicted supply indicators is with respect to driver supply,
the sequence of predicted demand indicators is with respect to driver demand, and
the geographic region is a geohash.
15
23. A computer program product, embodied in one or more non-transitory computer-readable storage mediums, comprising instructions executable by at least one processor to perform a method of controlling a surge pricing for a geographic region for a transportation service network, the method comprising, for each time step of a series of time steps:
20 optimizing, using a cost function for the transportation service network, a sequence of surge pricings for a sequence of time steps, including the time step as a first time step thereof, over a prediction horizon for the geographic region based on a sequence of predicted performance indicators, a sequence of predicted supply indicators and a sequence of predicted demand indicators for the sequence of time steps with respect to the geographic region; and
25 applying a first surge pricing of the sequence of surge pricings optimized for the first time step to the geographic region for the first time step, wherein
the cost function is parameterized by a sequence of surge pricing variables for the sequence of time steps for optimization to obtain the sequence of surge pricings for the sequence of time steps, and
30 the sequence of predicted performance indicators is predicted using a system dynamics model for the transportation service network based on the sequence of surge pricing variables for the sequence of time steps.