

CLAIMS

What is claimed is:

1. A computer-implemented method based on deep reinforcement learning ~~method for referral marketing, wherein the method comprises the following, comprising:~~

obtaining state information of an execution environment, of a deep reinforcement learning system, wherein the execution environment comprises a referral marketing environment, and wherein the state information of the execution environment comprises at least user information of a current user; and information associated with a current market of the deep reinforcement learning system, wherein the user information comprises user basic attribute information and user historical behavior information, and wherein the information associated with the current market comprises marketing services and transactions associated with the current user;

determining, by the deep reinforcement learning system, a marketing behavior corresponding marketing behavior for to the state information based on of the execution environment using a marketing strategy that is trained by a deep neural network in the deep reinforcement learning system and has classified offers, wherein the marketing behavior comprises a combination of a marketing channel, marketing content, and a marketing time period; and wherein determining the marketing behavior comprises:

obtaining the a current user historical behavior information from the user information of the current user, wherein the user historical behavior information comprises at least one of a historical channel, a historical marketing content, or a historical transaction record;

generating a next marketing behavior using the user historical behavior information of the current user based on the classified offers in the marketing strategy, wherein the next marketing behavior comprises the combination of the marketing channel, the marketing content, and the marketing time period corresponding to the at least one of a historical channel, a historical marketing content, or a historical transaction record;

selecting the marketing channel, the marketing content, and the marketing time period from the next marketing behavior applicable to the execution environment; and

determining the marketing behavior using the selected marketing channel, the marketing content, and the marketing time period;

adjusting the execution environment by updating the marketing behavior;
in response to adjusting the execution environment by updating the marketing behavior,
obtaining a current reward score of the execution environment for the marketing behavior, wherein
the current reward score is determined based on a response of the current user to the marketing
behavior; and
updating the marketing strategy based ~~at least~~ on the current reward score.

~~2. The method according to claim 1, wherein the user information comprises user basic attribute information and user historical behavior information.~~

23. The method according to claim 12, wherein the user historical behavior information comprises at least one of the following: a historical operation sequence, a list of recently browsed pages, a list of recently accepted marketing information, a consumption record, a transaction record, a loan record, a travel record, and an insurance record.

34. The computer-implemented method according to claim 1, wherein the user information comprises user profile information, ~~and the user profile information is~~ predetermined based on at least one of the user basic attribute information ~~and/or~~ the user historical behavior information.

45. The computer-implemented method according to claim 1, wherein the state information further comprises scenario information of a scenario ~~in which~~that the current user is located.

56. The computer-implemented method according to claim 35, wherein the scenario information comprises location service scenario information, payment scenario information, and wallet operation event information.

67. The computer-implemented method according to claim 1, wherein the marketing content comprises a marketing message and a credit package, and the credit package comprises one or more of ~~the following:~~ a red envelope, a coupon, ~~and/or~~ a right.

78. The computer-implemented method according to claim 1, wherein the marketing channel comprises one or more of ~~the following~~: message sending, card display, badges reminder, and notification dots.

89. The computer-implemented method according to claim 1, further comprising ~~the following~~:

determining a return value based ~~at least~~ on the current reward score, wherein the return value comprises a weighted sum of the current reward score and at least one future reward score; and

~~the updating the marketing strategy based at least on the current reward score comprises~~ determining a subsequent marketing strategy to maximize the return value.

910. ~~A deep reinforcement learning apparatus for referral marketing, wherein the apparatus comprises the following:~~

~~a state acquisition unit, configured to obtain~~ A non-transitory, computer-readable medium storing one or more instructions executable by a computer system to perform operations comprising:

obtaining state information of an execution environment; ~~of a deep reinforcement learning system, wherein the execution environment comprises a referral marketing environment, and wherein the state information~~ of the execution environment comprises ~~at least~~ user information of a current user; and information associated with a current market of the deep reinforcement learning system, wherein the user information comprises user basic attribute information and user historical behavior information, and wherein the information associated with the current market comprises marketing services and transactions associated with the current user;

~~a behavior determination unit, configured to determine~~ determining, by the deep reinforcement learning system, a marketing behavior corresponding ~~marketing behavior for to the state information of the state information based on execution environment using~~ a marketing strategy that is trained by a deep neural network in the deep reinforcement learning system and has classified offers, wherein the marketing behavior comprises a combination of a marketing channel, marketing content, and a marketing time period; and wherein determining the marketing behavior comprises:

~~a score acquisition unit, configured to obtain a current~~ obtaining the user historical behavior information of the current user from the user information of the current user, wherein the user historical behavior information comprises at least one of a historical channel, a historical marketing content, or a historical transaction record;

~~generating a next marketing behavior using the user historical behavior information of the current user based on the classified offers in the marketing strategy, wherein the next marketing behavior comprises the combination of the marketing channel, the marketing content, and the marketing time period corresponding to the at least one of a historical channel, a historical marketing content, or a historical transaction record;~~

~~selecting the marketing channel, the marketing content, and the marketing time period from the next marketing behavior applicable to the execution environment; and~~

~~determining the marketing behavior using the selected marketing channel, the marketing content, and the marketing time period;~~

~~adjusting the execution environment by updating the marketing behavior;~~

~~in response to adjusting the execution environment by updating the marketing behavior, obtaining a current~~ reward score of the execution environment for the marketing behavior, ~~wherein the current reward score is determined based on a response of the current user to the marketing behavior;~~ and

~~an update unit, configured to update~~ updating the marketing strategy based ~~at least~~ on the current reward score.

~~11. The apparatus according to claim 10, wherein the user information comprises user basic attribute information and user historical behavior information.~~

~~1012.~~ The apparatus according to claim ~~944~~, wherein the user historical behavior information comprises at least one of the following: a historical operation sequence, a list of recently browsed pages, a list of recently accepted marketing information, a consumption record, a transaction record, a loan record, a travel record, and an insurance record.

~~1113.~~ The ~~apparatus~~ non-transitory, computer-readable medium according to claim ~~940~~, wherein the user information comprises user profile information, ~~and the user profile information~~

~~is~~ predetermined based on at least one of the user basic attribute information ~~and/or~~ the user historical behavior information.

~~124. The apparatus~~The non-transitory, computer-readable medium according to claim ~~940~~, wherein the state information further comprises scenario information of a scenario ~~in which~~that the current user is located.

~~135. The apparatus~~non-transitory, computer-readable medium according to claim ~~124~~, wherein the scenario information comprises location service scenario information, payment scenario information, and wallet operation event information.

~~146. The apparatus~~non-transitory, computer-readable medium according to claim ~~940~~, wherein the marketing content comprises a marketing message and a credit package, and the credit package comprises one or more of ~~the following~~: a red envelope, a coupon, ~~and/or~~ a right.

~~157. The apparatus~~non-transitory, computer-readable medium according to claim ~~940~~, wherein the marketing channel comprises one or more of ~~the following~~: message sending, card display, badges reminder, and notification dots.

~~168. The apparatus~~non-transitory, computer-readable medium according to claim ~~940~~, the operations further comprising ~~a return value determination unit, configured to determine;~~

determining a return value based ~~at least~~ on the current reward score, wherein the return value comprises a weighted sum of the current reward score and at least one future reward score; and

~~the update unit is configured to determine~~determining a subsequent marketing strategy to maximize the return value.

~~179. A computing device~~computer-implemented system, comprising ~~a;~~
one or more computers; and

one or more computer memory and a processor devices interoperably coupled with the one or more computers and having tangible, non-transitory, machine-readable media storing one or more instructions that, when executed by the one or more computers, perform one or more operations comprising:

obtaining state information of an execution environment of a deep reinforcement learning system, wherein the memory stores executable code, and execution environment comprises a referral marketing environment, and wherein the processor executes state information of the executable code to implement execution environment comprises user information of a current user and information associated with a current market of the method according to any one of claims 1 to 7 deep reinforcement learning system, wherein the user information comprises user basic attribute information and user historical behavior information, and wherein the information associated with the current market comprises marketing services and transactions associated with the current user;

determining, by the deep reinforcement learning system, a marketing behavior corresponding to the state information of the execution environment using a marketing strategy that is trained by a deep neural network in the deep reinforcement learning system and has classified offers, wherein the marketing behavior comprises a combination of a marketing channel, marketing content, and a marketing time period, and wherein determining the marketing behavior comprises:

obtaining the user historical behavior information of the current user from the user information of the current user, wherein the user historical behavior information comprises at least one of a historical channel, a historical marketing content, or a historical transaction record;

generating a next marketing behavior using the user historical behavior information of the current user based on the classified offers in the marketing strategy, wherein the next marketing behavior comprises the combination of the marketing channel, the marketing content, and the marketing time period corresponding to the at least one of a historical channel, a historical marketing content, or a historical transaction record;

selecting the marketing channel, the marketing content, and the marketing time period from the next marketing behavior applicable to the execution environment; and

determining the marketing behavior using the selected marketing channel, the marketing content, and the marketing time period;

adjusting the execution environment by updating the marketing behavior;
in response to adjusting the execution environment by updating the marketing
behavior, obtaining a current reward score of the execution environment for the marketing
behavior, wherein the current reward score is determined based on a response of the current user
to the marketing behavior; and
updating the marketing strategy based on the current reward score.

18. The computer-implemented system according to claim 17, wherein the user information
comprises user basic attribute information.

19. The apparatus according to claim 17, wherein the user historical behavior information
comprises at least one of the following: a historical operation sequence, a list of recently browsed
pages, a list of recently accepted marketing information, a consumption record, a transaction
record, a loan record, a travel record, and an insurance record.

20. The computer-implemented system according to claim 17, wherein the state
information further comprises scenario information of a scenario that the current user is located.

21. The computer-implemented system according to claim 20, wherein the scenario
information comprises location service scenario information, payment scenario information, and
wallet operation event information.

22. The computer-implemented system according to claim 17, wherein the marketing
content comprises a marketing message and a credit package, and the credit package comprises
one or more of a red envelope, a coupon, or a right.

23. The computer-implemented system according to claim 17, wherein the marketing
channel comprises one or more of message sending, card display, badge reminder, and notification
dots.

24. The computer-implemented system according to claim 17, the operations further comprising:

determining a return value based on the current reward score, wherein the return value comprises a weighted sum of the current reward score and at least one future reward score; and
determining a subsequent marketing strategy to maximize the return value.