CLAIMS

What is claimed is:

1. A method comprising:

obtaining a plurality of candidate page display methods for displaying a ~~page to a user through reinforcement learning, comprising:~~succession of pages;

~~obtaining~~ obtaining a plurality of candidate page element arrangement methods for arranging different page elements on a same page;

obtaining a plurality of candidate attributes of the page elements;

generating, from the candidate page display methods, the candidate page element arrangement methods, and the candidate attributes of the page elements, a first plurality of candidate page layouts;

generating a display policy that defines a mapping relationship between the first plurality of candidate page layouts and a plurality of different user features ~~of a first user~~ ;

in response to a request ~~for accessing~~to access a first page by ~~the~~a first user, obtaining a plurality of first user features, wherein the plurality of first user features ~~comprise~~comprises one or more user profile features and one or more user context features;

~~mapping,~~selecting, according to the display policy and based on ~~a display policy,~~ the plurality of first user features ~~to,~~ a first page layout ~~in a~~from the first plurality of candidate page layouts ~~generated for the first page~~;

displaying the first page ~~to the first user~~ with the first page layout;

obtaining operation information of the first user ~~on~~interacting with the first page with the first page layout, and determining a reward score ~~for current page display~~ based on the operation information; and

updating the display policy ~~for displaying pages~~ by ~~using~~implementing a reinforcement learning algorithm based at least on the reward score ~~to perform page display.~~.

2. The method according to claim 1, wherein

the user profile features comprise at least one of the following: an age, a gender, registration duration, an education level, an occupation, ~~and~~or a user tag; and

the user context features comprise at least one of the following: time when the request is sent, an interface that the request is sent through, or operation track information of the first user.

3. The method according to claim 1, wherein ~~the first plurality of candidate page layouts are generated in advance in the following way:~~

~~determining candidate page display methods;~~

~~determining a candidate method of arranging page elements on the first page;~~

~~determining candidate attributes of the page elements; and~~

~~combining the page display methods, the candidate arrangement method, and the candidate attributes to generate the first plurality of candidate page layouts.~~

~~4. The method according to claim 3, wherein~~

the candidate page display methods comprise waterfall display ~~and~~or page turning display; and

the plurality of candidate attributes of the page elements comprise display sizes, display shapes, ~~and~~or resolution of the page elements.

~~5~~

4. The method according to claim 1, wherein the display policy is an initial policy, and selecting the ~~mapping the user features to a~~ first page layout ~~in a~~from the first plurality of candidate page layouts ~~that is~~ generated ~~in advance~~ for the first page comprises:

randomly assigning a first traffic bucket in a first plurality of traffic buckets to the first user, wherein the first plurality of traffic buckets correspond to the first plurality of candidate page layouts; and

using a candidate page layout ~~corresponding~~from the first plurality of candidate page layouts that corresponds to the first traffic bucket as the first page layout.

~~6~~

5. The method according to claim 1, wherein the display policy is implemented by one of a logistic regression model, a gradient boosting decision tree GBDT model, ~~and~~or a deep neural network.

~~7~~

6. The method according to claim 1, wherein the operation information comprises at least one of the following: hit count on the first page, first page view count, view depth to the first page, ~~and~~or time on the first page; and

the determining ~~a~~the reward score for current page display based on the operation information comprises: determining the reward score as a weighted sum of at least one of the operation information.

~~8~~

7. The method according to claim 1~~ or 7~~, wherein the operation information comprises a transaction amount achieved based on the first page; and

the determining ~~a~~the reward score for current page display based on the operation information comprises: determining the reward score to be directly proportional to the transaction amount.

~~9~~

8. The method according to claim 1, wherein ~~the~~ updating the display policy for displaying pages by using ~~a~~the reinforcement learning algorithm based at least on the reward score to perform page display comprises:

adding the first user features, the first page layout, and the reward score to a training sample set as one training sample; and

updating the display policy for displaying the pages ~~by using~~on the training sample set ~~through~~and by using the reinforcement learning algorithm.

~~10~~

9. The method according to claim 1~~ or 9~~, wherein the reinforcement learning algorithm comprises a SARSA algorithm, an evolution strategy algorithm, a Q-learning algorithm, a policy gradient algorithm, ~~and~~or a cross information entropy method.

10. A computer-implemented system, comprising:

one or more computers; and

one or more computer memory 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 operations comprising:

obtaining a plurality of candidate page display methods~~11. An apparatus~~ for displaying a ~~page to a user through reinforcement learning, comprising:~~succession of pages;

obtaining a ~~feature acquisition unit, configured to obtain~~plurality of candidate page element arrangement methods for arranging different page elements on a same page;

obtaining a plurality of candidate attributes of the page elements;

generating, from the candidate page display methods, the candidate page element arrangement methods, and the candidate attributes of the page elements, a first plurality of candidate page layouts;

generating a display policy that defines a mapping relationship between the first plurality of candidate page layouts and a plurality of different user features ~~of a first user~~;

in response to a request ~~for accessing~~to access a first page by ~~the~~a first user, obtaining a plurality of first user features, wherein the plurality of first user features ~~comprise~~comprises one or more user profile features and one or more user context features;

~~a layout determining unit, configured~~selecting, according to ~~map, based on a~~ the display policy~~,~~ and based on the plurality of first user features ~~to~~, a first page layout ~~in a~~from the first plurality of candidate page layouts ~~generated for the first page~~;

~~a display unit, configured to display the first page to the first user with the first page layout;~~

~~a feedback acquisition unit, configured to obtain operation information of the first user on~~displaying the first page with the first page layout~~, and determine a~~;

obtaining operation information of the first user interacting with the first page with the first page layout, and determining a reward score ~~for current page display~~ based on the operation information; and

~~an update unit, configured to update~~updating the display policy by ~~using~~implementing a reinforcement learning algorithm based at least on the reward score ~~to perform page display~~.

~~12.~~

11. The ~~apparatus~~computer-implemented system according to claim ~~11~~10, wherein

the user profile features comprise at least one of the following: an age, a gender, registration duration, an education level, an occupation, ~~and~~or a user tag; and

the user context features comprise at least one of the following: time when the request is sent, an interface that the request is sent through, or operation track information of the first user.

~~13~~

12. The ~~apparatus~~computer-implemented system according to claim ~~11~~10, wherein ~~the first plurality of candidate page layouts are generated in advance in the following way:~~

~~determining candidate page display methods;~~

~~determining a candidate method of arranging page elements on the first page;~~

~~determining candidate attributes of the page elements; and~~

~~combining the page display methods, the candidate arrangement method, and the candidate attributes to generate the first plurality of candidate page layouts.~~

~~14. The apparatus according to claim 13, wherein~~

the candidate page display methods comprise waterfall display ~~and~~or page turning display; and

the plurality of candidate attributes of the page elements comprise display sizes, display shapes, ~~and~~or resolution of the page elements.

~~15~~

13. The ~~apparatus~~computer-implemented system according to claim ~~11~~10, wherein the display policy is an initial policy, and selecting the first page layout ~~determining unit is configured to~~from the first plurality of candidate page layouts generated for the first page comprises:

randomly ~~assign~~assigning a first traffic bucket in a first plurality of traffic buckets to the first user, wherein the first plurality of traffic buckets correspond to the first plurality of candidate page layouts; and

~~use~~using a candidate page layout ~~corresponding~~from the first plurality of candidate page layouts that corresponds to the first traffic bucket as the first page layout.

~~16~~

14. The ~~apparatus~~computer-implemented system according to claim ~~11~~10, wherein the display policy is implemented by one of a logistic regression model, a gradient boosting decision tree GBDT model, ~~and~~or a deep neural network.

~~17~~

15. The ~~apparatus~~computer-implemented system according to claim ~~11~~10, wherein the operation information comprises at least one of the following: hit count on the first page, first page view count, view depth to the first page, ~~and~~or time on the first page; and

the ~~feedback acquisition unit is configured to determine~~determining the reward score for current page display based on the operation information comprises: determining the reward score as a weighted sum of at least one of the operation information.

~~18~~

16. The ~~apparatus~~computer-implemented system according to claim ~~11 or 17~~10, wherein the operation information comprises a transaction amount achieved based on the first page; and

the ~~feedback acquisition unit is configured to determine~~determining the reward score for current page display based on the operation information comprises: determining the reward score to be directly proportional to the transaction amount.

~~19~~

17. The ~~apparatus~~computer-implemented system according to claim ~~11~~10, wherein ~~the update unit is configured to~~updating the display policy for displaying pages by using the reinforcement learning algorithm based at least on the reward score to perform page display comprises:

~~add~~adding the first user features, the first page layout, and the reward score to a training sample set as one training sample; and

~~update~~updating the display policy ~~by using~~for displaying pages on the training sample set ~~through~~and by using the reinforcement learning algorithm.

~~20~~

18. The ~~apparatus~~computer-implemented system according to claim ~~11 or 19~~10, wherein the reinforcement learning algorithm comprises an evolution strategy algorithm, a Q-learning algorithm, a SARSA algorithm, a policy gradient algorithm, ~~and~~or a cross information entropy method.

~~21~~

19. A non-transitory, computer-readable ~~storage~~ medium~~, wherein a computer program is stored on the computer readable storage medium, and when the computer program is executed in a computer, the computer performs the method according to any~~ storing one ~~of claims of 1 to 10.~~

~~22. A computing device, comprising a memory and a processor, wherein the memory stores~~or more instructions executable ~~code, and when executing~~by a computer system to perform operations comprising:

obtaining a plurality of candidate page display methods for displaying a succession of pages;

obtaining a plurality of candidate page element arrangement methods for arranging different page elements on a same page;

obtaining a plurality of candidate attributes of the ~~executable code,~~page elements;

generating, from the ~~processor implements~~candidate page display methods, the ~~method~~candidate page element arrangement methods, and the candidate attributes of the page elements, a first plurality of candidate page layouts;

generating a display policy that defines a mapping relationship between the first plurality of candidate page layouts and a plurality of different user features;

in response to a request to access a first page by a first user, obtaining a plurality of first user features, wherein the plurality of first user features comprises one or more user profile features and one or more user context features;

selecting, according to ~~any one of claims 1 to 10~~the display policy and based on the plurality of first user features, a first page layout from the first plurality of candidate page layouts;

displaying the first page with the first page layout;

obtaining operation information of the first user interacting with the first page with the first page layout, and determining a reward score based on the operation information; and

updating the display policy by implementing a reinforcement learning algorithm based at least on the reward score.