

CLAIMS

1. A network device, comprising:
 - a transceiver to send and receive data over a network; and
 - one or more processors that are operative to perform actions, including:
 - creating and training, until it is complete, a tree for a first time that has multiple branches and multiple nodes and that represents user responses affecting a feature measure, wherein the creating and training includes sending a plurality of training messages to a plurality of telecom subscriber users in a target user group, and includes creating the branches to each maximize an information gain for a message attribute or user attribute with respect to the feature measure based on responses of the plurality of telecom subscriber users to the plurality of training messages, wherein the plurality of training messages have attributes corresponding to being sent at different times and with different types of messages, and wherein each node within the tree is associated with a subset of the plurality of telecom subscriber users and includes a target distribution for the feature measure and for the telecom subscriber users in the associated subset and includes a control distribution for the feature measure and for other telecom subscriber users in a control user group;
 - using the tree for the first time to send, to multiple additional telecom subscriber users that are distinct from the plurality of telecom subscriber users in the target user group and that each has attributes based at least in part in prior activities in using telecom services, a plurality of further messages by, for each of the multiple additional telecom subscriber users, traversing the tree based at least in part on the attributes of the additional telecom subscriber user, generating an ordered ranking for the additional telecom subscriber user of the plurality of further messages based on determining a feature measure lift by selecting values from the target and control distributions in the tree, and using the ordered ranking to select one or more of the plurality of further messages to send to the additional telecom subscriber user;

repeatedly adapting the tree to changes over time by, for each of multiple additional times after the first time and after the using of the tree to send the plurality of further messages, retraining the tree for the additional time to correspond to further user responses to additional interactions with respect to the feature measure, wherein the adapting includes changing the branches and the nodes of the tree; and

after each of one or more of the additional times, using the retrained tree for the additional time to select and send additional messages to the plurality of telecom subscriber users.

2. The network device of claim 1, wherein the feature measure includes at least one of an Average Revenue Per User (ARPU), Active Base Percentage (ABP), Average Revenue Per Paying User (ARPPU), or an average margin per user (AMPU).

3. The network device of claim 1 wherein the adapting of the tree to the changes over time is performed using a sliding time window having a duration that does not include all data used for the creating and the training of the tree for the first time and that further includes additional data generated after the first time.

4. The network device of claim 1 wherein the information gain for each of the branches is further determined for each node within the tree based on a difference between an overall entropy at the node and an entropy conditioned on a candidate attribute at the node.

5. The network device of claim 1 wherein the one or more processors are further operative to perform pre-processing on at least one message attribute or user attribute to enable binary testing to be performed using the attribute during the creating and training of the tree.

6. The network device of claim 1 wherein, for each of the nodes, at least one of the target distribution or the control distribution for the node is modeled based on a gamma distribution or a Bernoulli distribution.

7. The network device of claim 1 wherein the creating and training of the tree includes creating a NULL category and performing testing based on the NULL category using one or more of the sent plurality of further messages for which at least one message attribute or user attribute is missing.

8. The network device of claim 1 wherein the creating and training of the tree includes using at least one user attribute for each of the plurality of telecom subscriber users that represents at least one of a recharge time series cluster or a usage histogram cluster.

9. A non-transitory computer readable storage device, having computer-executable instructions stored thereon, that in response to execution by a processor unit, cause the processor unit to perform operations, comprising:

creating and training, until it is complete, a tree for a first time that has multiple branches and multiple nodes and that represents user responses affecting a feature measure, wherein the creating and training includes sending a plurality of training messages having a plurality of attributes, to a plurality of users in a target user group and includes creating the branches to each maximize an information gain for a message attribute or user attribute with respect to the feature measure based on responses of the plurality of users to the plurality of training messages, wherein each node within the tree is associated with a subset of the plurality of users and includes a target distribution for the feature measure and for the users in the associated subset and includes a control distribution for the feature measure and for other users in a control user group;

using the tree created and trained for the first time to send a plurality of further messages to multiple additional users distinct from the plurality of users in

the target user group, the sending including, for each of the multiple additional users, traversing the tree based at least in part on attributes of the additional user, generating an ordered ranking for the additional user of the plurality of further messages based on determining a feature measure lift by performing a comparison between randomly selected values from the target and control distributions in the tree, and using the ordered ranking to select one or more of the plurality of further messages to send to the additional user;

repeatedly adapting the tree to changes over time by, for each of multiple additional times after the first time and after the using of the tree to send the plurality of further messages, retraining the tree for the additional time to correspond to further user responses to additional interactions with respect to the feature measure, wherein the adapting includes changing the branches and the nodes of the tree; and

after each of one or more of the multiple additional times, using the retrained tree for the additional time to select and send additional messages.

10. The non-transitory computer-readable storage device of claim 9 wherein the feature measure includes at least one of an Average Revenue Per User (ARPU), Active Base Percentage (ABP), Average Revenue Per Paying User (ARPPU), or an average margin per user (AMPU).

11. The non-transitory computer-readable storage device of claim 9 wherein the adapting of the tree to the changes over time is performed using a sliding time window that includes at least some data distinct from data used for the creating and the training of the tree for the first time.

12. The non-transitory computer-readable storage device of claim 11 wherein a duration of the sliding time window is adaptive based on a user behavior.

13. The non-transitory computer-readable storage device of claim 9 wherein the information gain for each of the branches is further determined for each node within the tree based on a difference between an overall entropy at the node and an entropy conditioned on a candidate attribute at the node.
14. The non-transitory computer-readable storage device of claim 9 wherein the computer-executable instructions cause the processor unit to further perform operations including performing pre-processing on at least one message attribute or user attribute to enable binary testing to be performed using the attribute during the creating and training of the tree.
15. The non-transitory computer-readable storage device of claim 9 wherein, for each of the nodes, at least one of the target distribution or the control distribution for the node is modeled based on a gamma distribution or a Bernoulli distribution.
16. The non-transitory computer-readable storage device of claim 9 wherein the creating and training of the tree includes creating a NULL category and performing testing based on the NULL category for at least one message attribute or user attribute that is missing.
17. A network device, comprising:
a transceiver to send and receive data over a network; and
one or more processors that are operative to perform actions, including:
creating and training, until it is complete, a model for a first time that has multiple groups of users each having common user and message attributes, wherein the creating and training includes sending a plurality of training messages to a plurality of users in a target user group and includes separating the plurality of users into the multiple groups of users to maximize an information gain for a message attribute or user attribute with respect to a feature measure, wherein

each group of users has an associated target distribution for the feature measure and for the users in the group and includes a control distribution for the feature measure and for other users in a control user group;

using the model created and trained for the first time to send a plurality of further messages to multiple additional users distinct from the plurality of users in the target user group, the sending including, for each of the multiple additional users, employing the model to generate an ordered ranking for the additional user of the plurality of further messages based on determining a feature measure lift, and using the ordered ranking to select one or more of the plurality of further messages to send to the additional user;

repeatedly adapting the model to changes over time by, for each of multiple additional times after the first time and after the using of the model to send the plurality of further messages, retraining the model for the additional time to correspond to further user responses to additional interactions with respect to the feature measure, wherein the adapting includes modifying at least one of the target distribution or the control distribution for each of one or more of the groups of users of the model; and

after each of one or more of the multiple additional times, using the retrained model for the additional time to select and send additional messages.

18. The network device of claim 17 wherein the adapting of the model for at least one of the additional times further includes adding at least one new group of users to the model.

19. The network device of claim 17 wherein at least one of the multiple groups of users includes users having a common user attribute that represents a user propensity.

20. The network device of claim 17 wherein the separating of the plurality of users into the multiple groups of users to maximize the information gain is

performed based on maximizing a difference between an overall entropy at a first decision point and an entropy conditioned on a candidate attribute at the first decision point.