# 说明书摘要

本申请实施例提供一种信息流构建方法、系统、设备及存储介质。在本申请实施例中，可确定与目标用户适配的多个候选信息，并按照多个候选信息各自关注的效果指标，分别为多个候选信息进行出价，这可实现出价阶段的独立性，在出价阶段可灵活支持不同商户各自关注的效果指标。而且，可基于为多个候选信息确定的出价，选取在当前调控指标下胜出的获胜信息，并根据获胜信息，为目标用户构建信息流，这可实现调控阶段的独立性，而且，调控指标也是灵活可变的，因此，可支持任意切换调控指标。据此，本申请实施例中，可将商户和展示平台的利益诉求进行解耦，充分释放不同参与方的诉求优化能力，从而可提高分配效率，还可灵活支持不同参与方的利益诉求。

# 摘要附图



# 权利要求书

1、一种信息流构建方法，其特征在于，包括：

响应于目标用户触发的信息流生产指令，确定与所述目标用户适配的多个候选信息；

按照所述多个候选信息各自关注的效果指标，分别为所述多个候选信息计算针对所述目标用户的出价，不同候选信息关注的效果指标不完全相同；

基于所述多个候选信息对应的出价，选取在当前调控指标下胜出的至少一个获胜信息；

根据所述至少一个获胜信息，为所述目标用户构建信息流。

2、根据权利要求1所述的方法，其特征在于，所述按照所述多个候选信息各自关注的效果指标，分别为所述多个候选信息计算针对所述目标用户的出价，包括：

评估所述目标用户在所述第一候选信息关注的效果指标下的抢占价值；

基于所述第一候选信息对应的抢占价值，为所述第一候选信息确定出价；

其中，所述第一候选信息为所述多个候选信息中的任意一个。

3、根据权利要求2所述的方法，其特征在于，所述评估所述目标用户在所述第一候选信息关注的效果指标下的抢占价值，包括：

从至少一个出价模型中，为所述第一候选信息确定与其关注的效果指标对应的出价模型，不同效果指标对应的出价模型中的出价逻辑不同；

利用所述第一候选信息对应的出价模型，为所述第一候选信息计算针对所述目标用户的出价。

4、根据权利要求2所述的方法，其特征在于，所述基于所述第一候选信息对应的抢占价值，为所述第一候选信息确定出价，包括：

获取所述第一候选信息对应的基础出价和基础价值；

基于所述抢占价值与所述基础价值之间的第一差异，为所述第一候选信息确定出价，以使所述出价与所述基础出价之间的第二差异适配于所述第一差异。

5、根据权利要求1所述的方法，其特征在于，所述基于所述多个候选信息对应的出价，选取在当前调控指标下胜出的至少一个获胜信息，包括：

基于所述多个候选信息对应的出价，计算所述多个候选信息在所述当前调控指标下的排名得分；

根据所述多个候选信息的排名得分，确定在所述当前调控指标下胜出的至少一个获胜信息。

6、根据权利要求5所述的方法，其特征在于，所述基于所述多个候选信息对应的出价，计算所述多个候选信息在所述当前调控指标下的排名得分，包括：

获取所述多个候选信息各自对应的状态特征，所述状态特征中包含所述出价；

将所述多个候选信息各自对应的状态特征输入支持所述当前调控指标的调控模型中；

在所述调控模型中，基于状态特征与排名得分之间的映射关系，输出所述多个候选信息的排名得分。

7、根据权利要求6所述的方法，其特征在于，所述调控模型的训练过程，包括：

获取样本信息经过展示后的真实反馈参数；

若根据所述真实反馈参数确定所述样本信息的展示效果达到所述当前调控指标，则向所述调控模型施加激励，以供所述调控模型强化学习所述样本信息在所述当前调控指标下的状态特征与排名得分之间的映射关系。

8、根据权利要求6所述的方法，其特征在于，还包括：

响应于调控指标切换指令，将所述调控模型支持的调控指标调整至目标调控指标。

9、根据权利要求8所述的方法，其特征在于，所述将所述调控模型支持的调控指标调整至目标调控指标，包括：

获取样本信息经过展示后的真实反馈参数；

若根据所述真实反馈参数确定所述样本信息的展示效果达到所述目标调控指标，则向所述调控模型施加激励，以供所述调控模型强化学习所述样本信息在所述目标调控指标下的状态特征与排名得分之间的映射关系。

10、根据权利要求6所述的方法，其特征在于，所述状态特征包括出价、信息特征、用户特征、商户特征和上下文信息中的一种或多种。

11、根据权利要求1所述的方法，其特征在于，所述根据所述至少一个获胜信息，为所述目标用户构建信息流，包括：

分别计算所述至少一个获胜信息各自对应的展示质量得分；

按照所述目标用户下的排序约束条件，确定所述目标用户对应的信息流中可放置获胜信息的数量；

基于所述至少一个获胜信息各自对应的展示质量得分，确定可加入所述信息流中的目标信息以及目标信息所放置的位置。

12、根据权利要求11所述的方法，其特征在于，所述分别计算所述至少一个获胜信息各自对应的展示质量得分，包括：

计算第一获胜信息对应的用户体验得分；

获取所述第一获胜信息对应的扣费金额；

对所述第一获胜信息对应的用户体验得分和扣费金额进行加权求和，以获得所述第一获胜信息对应的展示质量得分；

其中，所述第一获胜信息为所述至少一个获胜信息中的任意一个。

13、根据权利要求12所述的方法，其特征在于，所述用户体验得分包括点击率和成交额中的一种或多种。

14、根据权利要求1所述的方法，其特征在于，所述效果指标包括点击率、加购率和转化率中的一种或多种，所述当前调控指标采用收入、成交额、点击率、转化率和加购率中的一种或多种。

15、一种信息流构建系统，其特征在于，包括：

媒体智能体，用于响应于目标用户触发的信息流生产指令，确定与所述目标用户适配的多个候选信息；

出价智能体，用于按照所述多个候选信息各自关注的效果指标，分别为所述多个候选信息计算针对所述目标用户的出价，不同候选信息关注的效果指标不完全相同；

调控智能体，用于基于所述多个候选信息对应的出价，选取在当前调控指标下胜出的至少一个获胜信息；

所述媒体智能体，还用于根据所述至少一个获胜信息，为所述目标用户构建信息流。

16、一种计算设备，其特征在于，包括存储器、处理器和通信组件；

所述存储器用于存储一条或多条计算机指令；

所述处理器与所述存储器和所述通信组件耦合，用于执行所述一条或多条计算机指令，以用于：

响应于目标用户触发的信息流生产指令，确定与所述目标用户适配的多个候选信息；

按照所述多个候选信息各自关注的效果指标，分别为所述多个候选信息计算针对所述目标用户的出价，不同候选信息关注的效果指标不完全相同；

基于所述多个候选信息对应的出价，选取在当前调控指标下胜出的至少一个获胜信息；

根据所述至少一个获胜信息，为所述目标用户构建信息流。

17、一种存储计算机指令的计算机可读存储介质，其特征在于，当所述计算机指令被一个或多个处理器执行时，致使所述一个或多个处理器执行权利要求1-14任一项所述的信息流构建方法。

# 说明书

一种信息流构建方法、系统、设备及存储介质

技术领域

本申请涉及数据处理技术领域，尤其涉及一种信息流构建方法、系统、设备及存储介质。

背景技术

信息流是一种可以滚动浏览的内容流，信息流中包含的多个信息可出现在外观相似、一个挨一个的显示区块中。在电商领域中，信息流的构建方案主要涉及三个参与方：商户、展示平台和用户。基于信息流，商户可将信息触达至用户，展示平台可获得信息展示收益。因此，信息流中的信息布局将牵动多方的利益。

目前，信息流构建过程中采用的分配机制通常是以目标转化为优化方式的点击出价OCPC（Optimized Cost Per Click），但是在OCPC机制下，多方的优化目标严重耦合，导致分配过程的迭代效率很低，因此，分配效率不佳；而且，优化目标通常是固定配置好的，无法调整，这导致该机制能够支持的优化目标单一，灵活性不足。

发明内容

本申请的多个方面提供一种信息流构建方法、系统、设备及存储介质，用以提高信息流构建过程中的分配效率和/或灵活支持不同参与方的利益诉求。

本申请实施例提供一种信息流构建方法，包括：

响应于目标用户触发的信息流生产指令，确定与所述目标用户适配的多个候选信息；

按照所述多个候选信息各自关注的效果指标，分别为所述多个候选信息计算针对所述目标用户的出价，不同候选信息关注的效果指标不完全相同；

基于所述多个候选信息对应的出价，选取在当前调控指标下胜出的至少一个获胜信息；

根据所述至少一个获胜信息，为所述目标用户构建信息流。

本申请实施例还提供一种信息流构建系统，包括：

媒体模块，用于响应于目标用户触发的信息流生产指令，确定与所述目标用户适配的多个候选信息；

出价模块，用于按照所述多个候选信息各自关注的效果指标，分别为所述多个候选信息计算针对所述目标用户的出价，不同候选信息关注的效果指标不完全相同；

调控模块，用于基于所述多个候选信息对应的出价，选取在当前调控指标下胜出的至少一个获胜信息；

所述媒体模块，还用于根据所述至少一个获胜信息，为所述目标用户构建信息流。

本申请实施例还提供一种计算设备，包括存储器、处理器和通信组件；

所述存储器用于存储一条或多条计算机指令；

所述处理器与所述存储器和所述通信组件耦合，用于执行所述一条或多条计算机指令，以用于：

响应于目标用户触发的信息流生产指令，确定与所述目标用户适配的多个候选信息；

按照所述多个候选信息各自关注的效果指标，分别为所述多个候选信息计算针对所述目标用户的出价，不同候选信息关注的效果指标不完全相同；

基于所述多个候选信息对应的出价，选取在当前调控指标下胜出的至少一个获胜信息；

根据所述至少一个获胜信息，为所述目标用户构建信息流。

本申请实施例还提供一种存储计算机指令的计算机可读存储介质，当所述计算机指令被一个或多个处理器执行时，致使所述一个或多个处理器执行前述的信息流构建方法。

在本申请实施例中，可在目标用户触发信息流生产指令的情况下，确定与目标用户适配的多个候选信息，并按照多个候选信息各自关注的效果指标，分别为多个候选信息进行出价，这可实现出价阶段的独立性，在出价阶段可灵活支持不同商户各自关注的效果指标，从而为商户提供的候选信息进行更合理的出价，因此，可在出价阶段独立优化商户的利益诉求。而且，可基于为多个候选信息确定的出价，选取在当前调控指标下胜出的获胜信息，并根据获胜信息，为目标用户构建信息流，这可实现调控阶段的独立性，而且，调控指标也是灵活可变的，因此，可支持任意切换调控指标。据此，本申请实施例中，可将商户和展示平台的利益诉求进行解耦，充分释放不同参与方的诉求优化能力，从而可提高分配效率，还可灵活支持不同参与方的利益诉求。

附图说明

此处所说明的附图用来提供对本申请的进一步理解，构成本申请的一部分，本申请的示意性实施例及其说明用于解释本申请，并不构成对本申请的不当限定。在附图中：

图1为本申请一示例性实施例提供的一种信息流构建系统的结构示意图；

图2为本申请一示例性实施例提供的一种调控智能体的示例性工作原理示意图；

图3为本申请一示例性实施例提供的一种信息流构建系统的工程部署架构图；

图4为本申请一示例性实施例提供的一种信息流构建方法的流程示意图；

图5为本申请又一示例性实施例提供的一种计算设备的结构示意图。

具体实施方式

为使本申请的目的、技术方案和优点更加清楚，下面将结合本申请具体实施例及相应的附图对本申请技术方案进行清楚、完整地描述。显然，所描述的实施例仅是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

目前，在OCPC机制下，信息流构建过程中的分配效率不佳；而且，能够支持的优化目标单一，灵活性不足。为此，本申请的一些实施例中：可在目标用户触发信息流生产指令的情况下，确定与目标用户适配的多个候选信息，并按照多个候选信息各自关注的效果指标，分别为多个候选信息进行出价，这可实现出价阶段的独立性，在出价阶段可灵活支持不同商户各自关注的效果指标，从而为商户提供的候选信息进行更合理的出价，因此，可在出价阶段独立优化商户的利益诉求。而且，可基于为多个候选信息确定的出价，选取在当前调控指标下胜出的获胜信息，并根据获胜信息，为目标用户构建信息流，这可实现调控阶段的独立性，而且，调控指标也是灵活可变的，因此，可支持任意切换调控指标。据此，本申请实施例中，可将商户和展示平台的利益诉求进行解耦，充分释放不同参与方的诉求优化能力，从而可提高分配效率，还可灵活支持不同参与方的利益诉求。

以下结合附图，详细说明本申请各实施例提供的技术方案。

图1为本申请一示例性实施例提供的一种信息流构建系统的结构示意图。如图1所示，该系统包括：媒体智能体10、出价智能体20和调控智能体30，媒体智能体10、出价智能体20和调控智能体30之间可通信连接。其中，智能体可以是驻留于环境中的实体，它可以解释从环境中获得的反映环境中所发生事件的数据，并执行对环境产生影响的行动，智能体可被看作是一种在环境中“生存”的实体，它既可以是硬件（如计算设备等），也可以是软件（如神经网络模型等）。

本实施例提供的信息流的构建方案可应用于各种信息流应用场景中，例如，电商领域、媒体领域等，本实施例对应用场景不做限定。在不同的场景中，信息流中包含的信息类型可以是多种多样的，例如，在电商领域中，信息流中包含的信息可以是广告、短视频、商品等。

为此，本实施例中，媒体智能体10可用于确定信息流中信息布局，例如，在电商领域中，媒体智能体10可用于决定如何在信息流中选择坑位插入广告，以实现多方利益的同时优化。出价智能体20可用于帮助商户出价，例如，在电商领域中，出价智能体20可从广告主的视角出发，决定如何帮助广告主出价，使得广告主只需要简单的设置即可清晰表达利益诉求。调控智能体30可以用于对信息展示机会进行调控，例如，在电商领域中，调控智能体30可用于决定将广告展示机会分配给哪个广告主，以及怎么收费等问题。具体来说，调控智能体30可通过一系列明确的规则和买者出价来决定资源配置。

正如前文提及的，在信息流的构建过程中，主要涉及三个参与方：商户、展示平台和用户，通常，商户期望高效、低成本地达到精准的感兴趣用户触达，展示平台则关注收入及成交额等，而用户则期望拥有更好的用户体验，例如，高效地找到自己感兴趣的信息等。本实施例中，展示平台，作为商户和用户之间的中介，需要实现良好的双边匹配，并且为了系统的生态健康和长期发展，需要兼顾商户、展示平台和用户的三方利益诉求。为此，本实施例中，可将图1提供的信息流构建系统部署在展示平台一侧，当然，本实施例并不限于此，信息流构建系统也可部署在前述的三个参与方之外。本实施例中，在信息流的展示过程中，媒体智能体10可记录用户的各种交互行为数据，例如点击、加购、购买等，还可记录商户在信息流构建过程中的各种决策行为数据，例如，加价、降价等，以及信息点击量、商品购买量等店铺情况数据。通过上述海量的数据，使得本实施例中的信息流构建系统可充分理解用户和商户的行为。

基于此，本实施例中，信息流构建系统可基于对用户和商户的深入理解，为用户和商户提供个性化的决策服务，为信息展示机会提供个性化的高效匹配，实现灵活的多方利益诉求优化。

参考图1，为了实现商户与用户的高效匹配，同时实现商户、展示平台和用户的多方利益共同优化，本实施例中，提供一种框架，该框架可包括三层智能体：媒体智能体10、出价智能体20和调控智能体30，三层智能体独立优化自身的目标，且三层智能体之间可相互通信，以保证三层智能体各自的独立性。

本实施例中，目标用户可触发信息流生产指令，例如，当一个用户浏览电商网站时，可触发信息流生产指令。信息流生产指令可产生一条信息展示机会。值得说明的是，这里的信息展示机会是指向目标用户展示信息的机会，而非指信息流中的一个信息展示坑位。对于一个用户而言，其触发的一次信息流生产指令可向多个商户分别提供一条信息展示机会。此时，信息流中可加入的信息数量及位置等均未进行决策。

媒体智能体10可响应于目标用户触发的信息流生产指令，确定与目标用户适配的多个候选信息。其中，商户期望在展示平台中通过信息流进行展示的信息通常会有很多，本实施例中，媒体智能体10可在目标用户触发信息流生产指令的情况下，从海量的待展示信息中，筛选出与目标用户适配的多个候选信息。在一种可选的实现方案中：可获取目标用户的用户特征，包括但不限于性别、年龄、收入等；还可确定各待展示信息的触达倾向，包括但不限于性别、年龄、收入等；选择具有与目标用户的用户特征适配的触达倾向的待展示信息，作为候选信息。这样，可有效提高用户触达效果，而且，可减少大量的计算量。实际应用中，媒体智能体10可构建候选信息集合，该集合中的元素为候选信息，且可记录各候选信息的信息特征，包括但不限于信息类型、点击率预估值、转化率预估值等。

正如上文提及的，候选信息是商户期望展示的信息，而不同商户关注的效果指标可能不完全相同，其中，效果指标可包括但不限于转化率、点击率、加购率等。本实施例中，出价智能体20的工作可包括为不同类型的商户优化出价逻辑，例如，有的商户希望优化点击率、有的商户希望优化加购率，而有的商户则希望优化购买量，为此，出价智能体20可提供多个出价模型来支持多样的优化目标（效果指标），不同出价模型可独立优化自身的目标；相应地，不同出价模型内的出价逻辑也将不同，具体的出价逻辑可按需进行设定，将在后文中进行举例说明。

基于此，基于多样的出价模型，本实施例中，出价智能体20可按照多个候选信息各自关注的效果指标，分别为多个候选信息计算针对目标用户的出价。也即是，可利用与候选信息所关注的效果指标对应的出价模型，为候选信息计算针对目标用户的出价。这保证了为候选信息确定的出价忠于候选信息所属商户的效果诉求，且可灵活支持不同候选信息各自关注的不完全相同的效果指标。

另外，考虑到出价环境波动频繁，本实施例中，出价智能体20可按照指定的时间量级执行上述的出价操作，例如，可按照5分钟的时间量级执行出价操作，通过这种方式可以吧频繁波动的出价环境带来的影响最小化，达到接近实时地出价效果。

出价智能体20可将为候选信息确定的出价发送至调控智能体30。调控智能体30可基于多个候选信息对应的出价，选取在当前调控指标下胜出的至少一个获胜信息。本实施例中，调控智能体30支持的调控指标可包括但不限于收入、成交额、点击率、转化率或加购率等，其中，点击率、转化率及加购率等均是从用户利益角度出发进行的调控，这使得展示平台不仅可对平台收入负责，还可对整个信息流构建系统的长期健康稳定负责，也即兼顾商户、展示平台和用户三方的利益诉求。而且，调控智能体30支持的（优化目标）调控指标可按需进行切换，具体的实现方案将在后文中进行详述。

据此，出价智能体20和调控智能体30相互配合，可实现将商户的效果诉求和展示平台的效果诉求解耦，商户视角的抓手为智能调价/出价，而展示平台视角的抓手为信息展示机会分配/扣费机制，充分释放两部分诉求优化能力，获得更好的商户营销效果，但同时又给展示平台留有宏观调控的手段和能力，让商户的效果诉求和展示平台的效果诉求达到平衡，本实施例中采用了多目标联合优化的思路，以先商户后展示平台的顺序，在展示平台（调控智能体30）调控指标优化的过程中保证商户（出价智能体20）效果指标的优化，从而兼容展示平台和商户的效果诉求。

参考图1，调控智能体30可将确定出的至少一个获胜信息发送给媒体智能体10，媒体智能体10可根据至少一个获胜信息，为目标用户构建信息流。应当理解的是，为目标用户构建的信息流中不仅可包含获胜信息，还可包含其它信息。例如，在电商领域中，获胜信息可以是广告，在为目标用户构建信息流时，可将获胜广告加入信息流中，而目标用户的信息流中还可包含商品、短视频等其它信息。

其中，获胜广告的数量可根据实际需要进行设定，本实施例对此不作限定。在一种可选的实现方式中：媒体智能体10分别计算至少一个获胜信息各自对应的展示质量得分；按照目标用户下的排序约束条件，确定目标用户对应的信息流中可放置获胜信息的数量；基于至少一个获胜信息各自对应的展示质量得分，确定可加入信息流中的目标信息以及目标信息所放置的位置。以至少一个获胜信息中的第一获胜信息为例，实际应用中，可计算第一获胜信息对应的用户体验得分；获取第一获胜信息对应的扣费金额；对第一获胜信息对应的用户体验得分和扣费金额进行加权求和，以获得第一获胜信息对应的展示质量得分。其中，用户体验得分可包括但不限于点击率或成交额等能够体现与用户偏好的适配程度的评价指数。这样，展示质量得分可反映获胜信息在平台收入和用户体验两个维度上的效果情况，基于展示质量得分，媒体智能体10可对至少一个获胜信息进行排序。媒体智能体10还可按照目标用户对应的信息流的排序约束条件（例如，广告比例约束PVR、用户对广告的偏好等），来确定可加入信息流中的获胜信息的数量和位置，从而从获胜信息中选出前述的目标信息并确定各目标信息在信息流中所放置的位置。例如，若获胜的广告为K个，按照PVR确定目标用户的信息流中只能放置M（M<K）个广告，则可从K个广告中选择展示质量分数最高的M个广告，并将展示质量分数更高的广告放入信息流中更靠前的位置。当然，这仅是示例性的，本实施例并不限于此。这样，本实施例中，媒体智能体10再根据至少一个获胜信息为目标用户构建信息流的过程中，可兼顾展示平台的收入和用户体验。

值得说明的是，本实施例中，媒体智能体10可在调控智能体30选取出获胜信息之后再对信息流中可加入获胜信息的数量和位置进行决策，基于这种处理方式，媒体智能体10可在充分了解获胜信息的展示质量的情况下再进行信息布局的决策，可保证信息流中的信息布局更加符合用户的效果诉求，且可兼顾展示平台的利益诉求。

另外，本实施例中，上述的媒体智能体10、出价智能体20和调控智能体30均可通过工作过程中发生的真实数据进行训练而习得相应的处理能力，这些智能体可根据真实的动态环境的变化而不断调整，以实现最好的效果。实际应用中，这些智能体提供的学习模型可以通过历史数据进行离线训练得到初始化模型，然后部署到真实场景中，之后在工作过程中不断收集用户、展示平台和商户的真实反馈参数作为训练样本，并通过实时的训练或参数调整方案进行学习。

据此，本实施例中，可在目标用户触发信息流生产指令的情况下，确定与目标用户适配的多个候选信息，并按照多个候选信息各自关注的效果指标，分别为多个候选信息进行出价，这可实现出价阶段的独立性，在出价阶段可灵活支持不同商户各自关注的效果指标，从而为商户提供的候选信息进行更合理的出价，因此，可在出价阶段独立优化商户的利益诉求。而且，可基于为多个候选信息确定的出价，选取在当前调控指标下胜出的获胜信息，并根据获胜信息，为目标用户构建信息流，这可实现调控阶段的独立性，而且，调控指标也是灵活可变的，因此，可支持任意切换调控指标。据此，本申请实施例中，可将商户和展示平台的利益诉求进行解耦，充分释放不同参与方的诉求优化能力，从而可提高分配效率，还可灵活支持不同参与方的利益诉求。

在上述或下述实施例中，出价智能体20可采用多种实现方式为各个候选信息确定出价。

在一种可选的实现方式中：以多个候选信息中的第一候选信息为例，出价智能体20可评估目标用户在第一候选信息关注的效果指标下的抢占价值；基于第一候选信息对应的抢占价值，为第一候选信息确定出价。

在该实现方式中，抢占价值用于表征目标用户对于候选信息的吸引力。实际应用中，抢占价值可采用目标用户在候选信息关注的效果指标下的预估效果值。例如，在电商领域中，若第一候选信息关注的效果指标为转化率，则出价智能体20可评估目标用户在第一候选信息下的转化率，作为抢占价值，显然，目标用户在第一候选信息下的转化率越高，其对于第一候选信息来说，抢占价值越大，也即对第一候选信息所属商户的吸引力越高。

在此基础上，出价智能体20可基于目标用户在第一候选信息关注的效果指标下的抢占价值，来为第一候选信息确定出价，其中，出价通常与抢占价值成正比。可选地，出价智能体20可获取第一候选信息对应的基础出价和基础价值；基于抢占价值与基础价值之间的第一差异，为第一候选信息确定出价，以使出价与基础出价之间的第二差异适配于第一差异。其中，基础出价和基础价值可以是第一候选信息所属的商户自主设定的，当然，也可以是根据该商户的历史决策行为数据，进行取均值、取中值等计算而获得的，在此不做限定。例如，若第一候选信息所属的商户A的基础出价为5元，基础转化率为1%，若目标用户对应的转化率为1.2%，相比于基础转化率提高了20%，则第一候选信息的出价也可相对于基础出价提高20%，也即6元，且未超出商户A的出价边界。

在另一种可选的实现方式中：还是以多个候选信息中的第一候选信息为例，出价智能体20可确定第一候选信息所属目标商户的可用总金额以及出价边界，以利用目标商户的可用总金额最大化满足目标商户所关注的效果指标为目标，确定目标商户下的至少一个待展示信息的出价，以获得第一候选信息对应的出价。在该实现方式中，出价智能体20可统筹目标用户的可用总金额，涉及目标商户下的至少一个待展示信息的出价过程的统筹优化，目标商户下的不同待展示信息可能在不同的信息流构建过程中作为不同用户的候选信息。在该可选方案中，可将第一候选信息所属目标商户的可用总金额最优化利用作为目标，来将可用总金额合理分配至各个待展示信息，从而确定出第一候选信息的出价。

当然，应当理解的是，上述两种实现方式也仅是示例性的，出价智能体20还可采用其它实现方式来为多个候选信息确定出价，本实施例并不限于此。

另外，正如前述实施例中提及的，出价智能体20可提供多种出价模型，基于此，本实施例中，可从多种出价模型中，为第一候选信息确定与其关注的效果指标对应的出价模型；利用第一候选信息对应的出价模型，为第一候选信息计算针对目标用户的出价。对应于上述两种示例性的为第一候选信息确定出价的实现方式中，可利用第一候选信息对应的出价模型，实现上述两种示例性的实现方式中的出价操作。

据此，本实施例中，基于相互独立的出价模型，可独立优化不同出价模型的自身目标，而不同出价模型可对应不同的效果指标，因此，可为不同类型的商户优化出价策略，以满足不同类型商户的效果诉求。且在出价过程中，不受展示平台的利益诉求的影响，而是忠于商户的效果诉求进行出价，这可准确满足商户的效果诉求。

在上述或下述实施例中，调控智能体30可采用多种实现方式选取至少一个获胜信息。

在一种实现方式中：调控智能体30可基于多个候选信息对应的出价，计算多个候选信息在当前调控指标下的排名得分；根据多个候选信息的排名得分，确定在当前调控指标下胜出的至少一个获胜信息。

图2为本申请一示例性实施例提供的一种调控智能体30的示例性工作原理示意图。参考图2，在该实现方式中，计算排名得分的示例性方案可以是：获取多个候选信息各自对应的状态特征，状态特征中包含出价；将多个候选信息各自对应的状态特征输入支持当前调控指标的调控模型中；在调控模型中，基于状态特征与排名得分之间的映射关系Rϴ，输出多个候选信息的排名得分。

在调控模型中，可使用公式一来计算排名得分：

ri=Rϴ（bi，xi） -----公式一

其中，ri表示第i个候选信息的排名得分，bi表示第i个候选信息对应的出价，xi则表示第i个候选信息的除出价外的其它状态特征，xi可包括但不限于信息特征（类别、点击率预估值、转化率预估值等）、用户特征（性别、年龄、收入等）、商户特征（预算、触达倾向等）或上下文信息（场景、会话session内全局统计信息等）等。这样，调控模型的输入可以是状态特征，而输出则为排名得分。

在该实现方式中，调控模型可采用深度神经网络模型。其中，调控模型可满足博弈均衡约束（Game Equilibrium Contraints），以保证所有商户能够达到博弈均衡状态，对分配结果感到满意（也即是分配结果的改变不能使商户的收益变多）。实际应用中，可采用对称纳什均衡（Symmetric Nash Equilibrium）来实现博弈均衡约束，使得每个商户在最终分配到的展示位置下的收益最佳，不会嫉妒其它展示位置。

在该实现方式中，调控模型的训练过程可以是：获取样本信息经过展示后的真实反馈参数；若根据真实反馈参数确定样本信息的展示效果达到当前调控指标，则向调控模型施加激励，以供调控模型强化学习样本信息在当前调控指标下的状态特征与排名得分之间的映射关系。这样，调控模型可不断强化学习符合当前调控指标的样本信息所反映的状态特征与排名得分之间的映射关系，使得调控模型更好地支持当前调控指标。

正如前文提及的，调控模型可支持的调用指标是可切换的，基于此，在该实现方式中，可响应于调控指标切换指令，将调控模型支持的调控指标调整至目标调控指标。

对此，调控模型还可满足调控指标平滑切换约束（Smooth Transition Contraints）：由于信息流构建过程的动态性很强，不同参与方在不同时刻的利益诉求也不尽相同，当调控模型从服务于一个调控指标切换至另一个调控指标时，基于平滑切换约束可保证各参与方的利益效果平稳过渡。实际应用中，可通过真实反馈参数对调控模型进行训练来实现平滑切换约束。

相应地，在调控模型切换至目标调控指标后，调控模型的训练过程变成：获取样本信息经过展示后的真实反馈参数；若根据真实反馈参数确定样本信息的展示效果达到目标调控指标，则向调控模型施加激励，以供调控模型强化学习样本信息在目标调控指标下的状态特征与排名得分之间的映射关系。这样，调控模型切换至学习符合目标调控指标的样本信息所反映的状态特征与排名得分之间的映射关系，使得调控模型顺利切换至支持目标调控指标。可选地，激励可采用真实反馈参数的加权聚合值。

参考图2，本实施例中，调控智能体30可输出多个候选信息的排名得分，因此，可获得多个候选信息的排名顺序。基于该排名顺序，调控智能体30可确定多个候选信息的扣费。实际应用中，可采用GSP（Generallized Second Price）的方式确定扣费，也即是，排名顺序在前的候选信息按照排在其后的候选信息的出价进行扣费。当然，这种扣费方式仅是示例性的，本实施例并不限于此，在其它扣费方式下，还可按照候选信息的自身出价进行扣费，或者参考媒体智能体10按照展示质量得分进行排序后的顺序，来执行GSP方案以确定扣费。

据此，本实施例中，可在信息流构建过程所处的动态博弈环境下，解决多诉求指标优化问题。通过引入深度神经网络来提升调控智能体30的分配能力，并通过强化学习朝着提升给定的优化目标的方向直接优化调控模型的参数，在扣费上延续GSP的二价计费，而且调控模型的优化可在满足机制激励兼容/机制平滑切换的条件下进行，优化过程是基于真实反馈进行的，因此也可以支持任意切换定制目标并实现优化，从而实现基于数据驱动的方式优化调控智能体30。

在上述或下述实施例中，在工程实现上，信息流构建系统至少可包含在线推理部分和离线训练部分。

图3为本申请一示例性实施例提供的一种信息流构建系统的工程部署架构图。参考图3，在线推理部分可基于图Graph处理包括自动出价、调价、信息展示机会分配等功能，对应于前文提及的媒体智能体、出价智能体和调控智能体的功能。由于信息之间不存在依赖关系，而是存在并行的可能性，因此，本实施例中，可并行处理不同的信息。实际应用中，可基于tensorflow的数据流执行引擎以实现专门面向指定类型信息的信息流构建服务，例如，指定类型信息可以是广告，以基于tensorflow实现不同信息处理过程的并行。

另外，参考图3，在线推理部分还可包括系统指标监控组件和服务容量控制组件，其中，系统指标监控组件可用于采集指定的系统指标，还可周期性地检查集群指标是否满足正常运行条件，如果不满足，服务容量控制组件可调控系统服务级别。为此，本实施例中，可预先配置多个服务级别，在不同服务级别下可配置不同的信息流构建逻辑，例如，前述实施例中涉及到的信息流构建方案可作为某一个服务级别下的信息流构建逻辑，而在其它服务级别下可采用传统的OCPC机制确定获胜信息，以构建信息流，当然，还可采用其它方式来构建信息流。不同服务级别下的处理逻辑的复杂度以及所需的资源量可呈阶梯式变化，据此，可利用系统指标监控组件来感知不同的场景需求（例如：节假日、大促期间、购物淡季等）等，并可利用服务容量控制组件确定与感知到的场景需求适配的服务级别。以保证各种场景需求下的获得适配的处理效果。

系统指标监控组件还可用于在系统运行过程中，发现系统异常时间，并及时通知系统维护人员，以将影响控制在尽量小的范围内，并且保证系统维护人员及时获知系统运行状态，保证系统可用性。

参考图3，本实施例中，离线训练部分可包括数据收集组件，以收集信息流展示过程中产生的真实反馈数据，包括但不限于图3中的点击、购买、加购等数据，收集的真实反馈数据可作为训练样本，用于对信息流构建系统中涉及到的各种模型进行训练。其中，信息流构建系统中可提供模型中心，模型中心可管理前述实施例中涉及到的各种模型，包括但不限于调控模型、各种出价模型以及媒体智能体可采用的网络模型等。

图4为本申请一示例性实施例提供的一种信息流构建方法的流程示意图，该方法可由前述实施例中的信息流构建系统执行。参考图4，该方法包括：

步骤400、响应于目标用户触发的信息流生产指令，确定与目标用户适配的多个候选信息；

步骤401、按照多个候选信息各自关注的效果指标，分别为多个候选信息计算针对目标用户的出价，不同候选信息关注的效果指标不完全相同；

步骤402、基于多个候选信息对应的出价，选取在当前调控指标下胜出的至少一个获胜信息；

步骤403、根据至少一个获胜信息，为目标用户构建信息流。

在一可选实施例中，步骤按照多个候选信息各自关注的效果指标，分别为多个候选信息计算针对目标用户的出价，包括：

评估目标用户在第一候选信息关注的效果指标下的抢占价值；

基于第一候选信息对应的抢占价值，为第一候选信息确定出价；

其中，第一候选信息为多个候选信息中的任意一个。

在一可选实施例中，步骤评估目标用户在第一候选信息关注的效果指标下的抢占价值，包括：

从至少一个出价模型中，为第一候选信息确定与其关注的效果指标对应的出价模型，不同效果指标对应的出价模型中的出价逻辑不同；

利用第一候选信息对应的出价模型，为第一候选信息计算针对目标用户的出价。

在一可选实施例中，步骤基于第一候选信息对应的抢占价值，为第一候选信息确定出价，包括：

获取第一候选信息对应的基础出价和基础价值；

基于抢占价值与基础价值之间的第一差异，为第一候选信息确定出价，以使出价与基础出价之间的第二差异适配于第一差异。

在一可选实施例中，步骤基于多个候选信息对应的出价，选取在当前调控指标下胜出的至少一个获胜信息，包括：

基于多个候选信息对应的出价，计算多个候选信息在当前调控指标下的排名得分；

根据多个候选信息的排名得分，确定在当前调控指标下胜出的至少一个获胜信息。

在一可选实施例中，步骤基于多个候选信息对应的出价，计算多个候选信息在当前调控指标下的排名得分，包括：

获取多个候选信息各自对应的状态特征，状态特征中包含出价；

将多个候选信息各自对应的状态特征输入支持当前调控指标的调控模型中；

在调控模型中，基于状态特征与排名得分之间的映射关系，输出多个候选信息的排名得分。

在一可选实施例中，调控模型的训练过程，包括：

获取样本信息经过展示后的真实反馈参数；

若根据真实反馈参数确定样本信息的展示效果达到当前调控指标，则向调控模型施加激励，以供调控模型强化学习样本信息在当前调控指标下的状态特征与排名得分之间的映射关系。

在一可选实施例中，该方法还包括：

响应于调控指标切换指令，将调控模型支持的调控指标调整至目标调控指标。

在一可选实施例中，步骤将调控模型支持的调控指标调整至目标调控指标，包括：

获取样本信息经过展示后的真实反馈参数；

若根据真实反馈参数确定样本信息的展示效果达到目标调控指标，则向调控模型施加激励，以供调控模型强化学习样本信息在目标调控指标下的状态特征与排名得分之间的映射关系。

在一可选实施例中，状态特征包括出价、信息特征、用户特征、商户特征和上下文信息中的一种或多种。

在一可选实施例中，步骤根据至少一个获胜信息，为目标用户构建信息流，包括：

分别计算至少一个获胜信息各自对应的展示质量得分；

按照目标用户下的排序约束条件，确定目标用户对应的信息流中可放置获胜信息的数量；

基于至少一个获胜信息各自对应的展示质量得分，确定可加入信息流中的目标信息以及目标信息所放置的位置。

在一可选实施例中，步骤分别计算至少一个获胜信息各自对应的展示质量得分，包括：

计算第一获胜信息对应的用户体验得分；

获取第一获胜信息对应的扣费金额；

对第一获胜信息对应的用户体验得分和扣费金额进行加权求和，以获得第一获胜信息对应的展示质量得分；

其中，第一获胜信息为至少一个获胜信息中的任意一个。

在一可选实施例中，户体验得分包括点击率和成交额中的一种或多种。

在一可选实施例中，效果指标包括点击率、加购率和转化率中的一种或多种，当前调控指标采用收入、成交额、点击率、转化率和加购率中的一种或多种。

值得说明的是，上述关于信息流构建方法各实施例中的技术细节，可参考前述的系统实施例中的相关描述，为节省篇幅，在此不再赘述，但这不应造成本申请保护范围的损失。

上述实施例所提供方法的各步骤的执行主体均可以是同一设备，或者，该方法也由不同设备作为执行主体。比如，步骤400至步骤403的执行主体可以为设备A；又比如，步骤400和403的执行主体可以为设备A，步骤402的执行主体可以为设备B；等等。

另外，在上述实施例及附图中的描述的一些流程中，包含了按照特定顺序出现的多个操作，但是应该清楚了解，这些操作可以不按照其在本文中出现的顺序来执行或并行执行，操作的序号如401、402等，仅仅是用于区分开各个不同的操作，序号本身不代表任何的执行顺序。另外，这些流程可以包括更多或更少的操作，并且这些操作可以按顺序执行或并行执行。需要说明的是，本文中的“第一”、“第二”等描述，是用于区分不同的信息、消息、模块等，不代表先后顺序，也不限定“第一”和“第二”是不同的类型。

图5为本申请又一示例性实施例提供的一种计算设备的结构示意图。如图5所示，该计算设备包括：存储器50、处理器51以及通信组件52。

处理器51，与存储器50耦合，用于执行存储器50中的计算机程序，以用于：

响应于目标用户触发的信息流生产指令，确定与目标用户适配的多个候选信息；

按照多个候选信息各自关注的效果指标，分别为多个候选信息计算针对目标用户的出价，不同候选信息关注的效果指标不完全相同；

基于多个候选信息对应的出价，选取在当前调控指标下胜出的至少一个获胜信息；

根据至少一个获胜信息，为目标用户构建信息流。

在一可选实施例中，处理器51在按照多个候选信息各自关注的效果指标，分别为多个候选信息计算针对目标用户的出价时，用于：

评估目标用户在第一候选信息关注的效果指标下的抢占价值；

基于第一候选信息对应的抢占价值，为第一候选信息确定出价；

其中，第一候选信息为多个候选信息中的任意一个。

在一可选实施例中，处理器51在评估目标用户在第一候选信息关注的效果指标下的抢占价值时，用于：

从至少一个出价模型中，为第一候选信息确定与其关注的效果指标对应的出价模型，不同效果指标对应的出价模型中的出价逻辑不同；

利用第一候选信息对应的出价模型，为第一候选信息计算针对目标用户的出价。

在一可选实施例中，处理器51在基于第一候选信息对应的抢占价值，为第一候选信息确定出价时，用于：

获取第一候选信息对应的基础出价和基础价值；

基于抢占价值与基础价值之间的第一差异，为第一候选信息确定出价，以使出价与基础出价之间的第二差异适配于第一差异。

在一可选实施例中，处理器51在基于多个候选信息对应的出价，选取在当前调控指标下胜出的至少一个获胜信息时，用于：

基于多个候选信息对应的出价，计算多个候选信息在当前调控指标下的排名得分；

根据多个候选信息的排名得分，确定在当前调控指标下胜出的至少一个获胜信息。

在一可选实施例中，处理器51在基于多个候选信息对应的出价，计算多个候选信息在当前调控指标下的排名得分时，用于：

获取多个候选信息各自对应的状态特征，状态特征中包含出价；

将多个候选信息各自对应的状态特征输入支持当前调控指标的调控模型中；

在调控模型中，基于状态特征与排名得分之间的映射关系，输出多个候选信息的排名得分。

在一可选实施例中，调控模型的训练过程时，用于：

获取样本信息经过展示后的真实反馈参数；

若根据真实反馈参数确定样本信息的展示效果达到当前调控指标，则向调控模型施加激励，以供调控模型强化学习样本信息在当前调控指标下的状态特征与排名得分之间的映射关系。

在一可选实施例中，处理器51还用于：

响应于调控指标切换指令，将调控模型支持的调控指标调整至目标调控指标。

在一可选实施例中，处理器51在将调控模型支持的调控指标调整至目标调控指标时，用于：

获取样本信息经过展示后的真实反馈参数；

若根据真实反馈参数确定样本信息的展示效果达到目标调控指标，则向调控模型施加激励，以供调控模型强化学习样本信息在目标调控指标下的状态特征与排名得分之间的映射关系。

在一可选实施例中，状态特征包括出价、信息特征、用户特征、商户特征和上下文信息中的一种或多种。

在一可选实施例中，处理器51在根据至少一个获胜信息，为目标用户构建信息流时，用于：

分别计算至少一个获胜信息各自对应的展示质量得分；

按照目标用户下的排序约束条件，确定目标用户对应的信息流中可放置获胜信息的数量；

基于至少一个获胜信息各自对应的展示质量得分，确定可加入信息流中的目标信息以及目标信息所放置的位置。

在一可选实施例中，处理器51在分别计算至少一个获胜信息各自对应的展示质量得分时，用于：

计算第一获胜信息对应的用户体验得分；

获取第一获胜信息对应的扣费金额；

对第一获胜信息对应的用户体验得分和扣费金额进行加权求和，以获得第一获胜信息对应的展示质量得分；

其中，第一获胜信息为至少一个获胜信息中的任意一个。

在一可选实施例中，户体验得分包括点击率和成交额中的一种或多种。

在一可选实施例中，效果指标包括点击率、加购率和转化率中的一种或多种，当前调控指标采用收入、成交额、点击率、转化率和加购率中的一种或多种。

进一步，如图5所示，该计算设备还包括：电源组件53等其它组件。图5中仅示意性给出部分组件，并不意味着计算设备只包括图5所示组件。

值得说明的是，上述关于计算设备各实施例中的技术细节，可参考前述的系统实施例中的相关描述，为节省篇幅，在此不再赘述，但这不应造成本申请保护范围的损失。

相应地，本申请实施例还提供一种存储有计算机程序的计算机可读存储介质，计算机程序被执行时能够实现上述方法实施例中可由计算设备执行的各步骤。

上述图5中的存储器，用于存储计算机程序，并可被配置为存储其它各种数据以支持在计算平台上的操作。这些数据的示例包括用于在计算平台上操作的任何应用程序或方法的指令，联系人数据，电话簿数据，消息，图片，视频等。存储器可以由任何类型的易失性或非易失性存储设备或者它们的组合实现，如静态随机存取存储器（SRAM），电可擦除可编程只读存储器（EEPROM），可擦除可编程只读存储器（EPROM），可编程只读存储器（PROM），只读存储器（ROM），磁存储器，快闪存储器，磁盘或光盘。

上述图5中的通信组件，被配置为便于通信组件所在设备和其他设备之间有线或无线方式的通信。通信组件所在设备可以接入基于通信标准的无线网络，如WiFi，2G、3G、4G/LTE、5G等移动通信网络，或它们的组合。在一个示例性实施例中，通信组件经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中，所述通信组件还包括近场通信（NFC）模块，以促进短程通信。例如，在NFC模块可基于射频识别（RFID）技术，红外数据协会（IrDA）技术，超宽带（UWB）技术，蓝牙（BT）技术和其他技术来实现。

上述图5中的电源组件，为电源组件所在设备的各种组件提供电力。电源组件可以包括电源管理系统，一个或多个电源，及其他与为电源组件所在设备生成、管理和分配电力相关联的组件。

本领域内的技术人员应明白，本申请的实施例可提供为方法、系统、或计算机程序产品。因此，本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

本申请是参照根据本申请实施例的方法、设备（系统）、和计算机程序产品的流程图和／或方框图来描述的。应理解可由计算机程序指令实现流程图和／或方框图中的每一流程和／或方框、以及流程图和／或方框图中的流程和／或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和／或方框图一个方框或多个方框中指定的功能的步骤。

在一个典型的配置中，计算设备包括一个或多个处理器 (CPU)、输入/输出接口、网络接口和内存。

内存可能包括计算机可读介质中的非永久性存储器，随机存取存储器 (RAM) 和/或非易失性内存等形式，如只读存储器 (ROM) 或闪存(flash RAM)。内存是计算机可读介质的示例。

计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存 (PRAM)、静态随机存取存储器 (SRAM)、动态随机存取存储器 (DRAM)、其他类型的随机存取存储器 (RAM)、只读存储器 (ROM)、电可擦除可编程只读存储器 (EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器 (CD-ROM)、数字多功能光盘 (DVD) 或其他光学存储、磁盒式磁带，磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括暂存电脑可读媒体 (transitory media)，如调制的数据信号和载波。

还需要说明的是，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、商品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、商品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、商品或者设备中还存在另外的相同要素。

以上所述仅为本申请的实施例而已，并不用于限制本申请。对于本领域技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原理之内所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。

# 说明书附图



图1



图2



图3



图4



图5