

团体标准

T/CCSA 551—2024

T/CAAAD 004—2024

互联网广告 群体标识技术要求

Internet advertising—cohorts identifier requirements

2024 - 07 - 03 发布

2024 - 10 - 01 实施

中国广告协会

中国通信标准化协会

发布

目 次

前言	III
引言	IV
1 范围	1
2 规范性引用文件	1
3 术语、定义和缩略语	1
3.1 术语和定义	1
3.2 缩略语	2
4 概述	2
5 一般原则	2
5.1 匿名化原则	2
5.2 使用有效性原则	2
5.3 隐私保护原则	2
6 总体要求	2
6.1 访问可控要求	2
6.2 可重置要求	2
6.3 不可重标识要求	2
6.4 阶段化时效要求	2
7 系统组成	2
8 工作原理	3
8.1 工作过程	3
8.2 标识生成	3
8.3 标识评估	3
8.4 发布与更新要求	4
8.5 保障机制要求	4
9 群体标识的安全要求	4
9.1 数据安全	4
9.2 业务安全	4
附录 A（资料性） 互联网广告群体标识的常见生成技术	5
A.1 群体生成技术	5
A.2 群体标识符的生成规则	6
附录 B（资料性） 互联网广告群体标识的常见评估模型	7
B.1 群体有效性评估模型	7
B.2 数据安全评估模型	7
附录 C（资料性） 互联网广告群体标识的基础实现方案	8
C.1 基于流量平台的广告群体标识实现方案	8

附录 D（资料性） 互联网广告群体标识的典型应用场景 9

 D.1 在个性化广告推荐场景中的应用举例 9

 D.2 在跨平台广告跟踪场景中的应用举例 10

 D.3 群体标识下的广告效果归因 10

附录 E（资料性） 互联网广告群体标识编码格式与互通接口 11

 E.1 编码格式 11

 E.2 互通接口 11

参考文献 12

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国广告协会和中国通信标准化协会共同提出，并分别归口。

本文件起草单位：北京京东尚科信息技术有限公司、中国信息通信研究院、北京快手科技有限公司、郑州信大捷安信息技术股份有限公司、北京三快在线科技有限公司、阿里巴巴(中国)有限公司、秒针信息技术有限公司、利欧集团数字科技有限公司、北京国双科技有限公司、北京勾正数据科技有限公司、北京数牍科技有限公司、国家广告研究院。

本文件主要起草人：林战刚、何杰、冯娜、张泽华、李然、杨正军、彭皓、叶朗朗、落红卫、王昕、杨阳、陈婉莹、段佳、吕瑞、谭玲斌、刘猷伦、刘为华、黄坤、李世奇、刘力泉、周崧弢、黄腾、马磊、金银玉、顾明毅、范翔、郭廓。

引 言

为适应信息通信发展对标准文件的需求,由中国通信标准化协会和中国广告协会共同组织制定本文件,推荐有关方面采用。有关对本文件的建议和意见,向中国通信标准化协会和中国广告协会反映。

伴随互联网普及程度的不断提升,互联网广告在高速发展的同时也面临着用户隐私保护方面的问题和挑战:一方面,个人设备信息采集和使用的不规范行为会导致用户个人信息存在较高泄漏风险;另一方面,较严格的信息保护制度使得数据孤岛现象进一步加深,导致跨平台用户识别、定向能力及投放效果的提升受限,长期将影响互联网广告行业的健康发展。为规范个人信息使用,同时促进数据的交换与共享,特制定广告场景下的群体识别技术标准,以同时保障个体用户信息的安全性以及广告投放收益的有效性。

本标准旨在借鉴国内外互联网广告产业对群体标识技术的研究和应用,通过对技术原理以及最佳实践方案的提炼,研究群体标识的目标、定义、原则、特性、机制以及应用举措,形成具有广泛参考意义的技术参考。

互联网广告 群体标识技术要求

1 范围

本文件规定了互联网广告的群体标识的一般原则、总体要求、系统组成、编码格式、工作原理、接口要求与安全要求。

本文件适用于互联网广告领域跨平台流量识别、广告定向、广告效果归因等场景中的群体标识生成、使用、互通等相关活动。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 35273—2020 信息安全技术 个人信息安全规范

GB/T 37964—2019 信息安全技术 个人信息去标识化指南

T/CAAAD 001—2021 互联网广告数据应用和安全技术要求

3 术语、定义和缩略语

3.1 术语和定义

GB/T 35273—2020、GB/T 37964—2019界定的以及下列术语和定义适用于本文件。

3.1.1

标识符 identifier

用来标识某个实体的一个符号，在不同的应用环境下有不同的含义，通常由字母和数字以及其它字符构成。

[来源：T/CAAAD 003—2020，2.1.1]

3.1.2

广告群体标识符 advertising cohort identifier

用于描述群体中具有共同特性用户的匿名化标识符，用户的个人信息经过处理无法识别特定自然人且不能复原，并且该匿名化标识符可用于广告业务。

注：该标识是以保证个体用户无法被唯一标识为前提，通过安全的参数和算法体系生成的一种可用于互联网广告业务的用户标识符。以下简称“群体标识”。

3.1.3

重标识 re-identification

将去标识化的数据集重新关联到原始数据主体，这个过程称为重标识或重标识攻击。

[来源：GB/T 37964—2019，3.9]

3.1.4

个人信息 personal information

以电子或其他方式记录的与已识别或可识别的自然人有关的各种信息，不包括匿名化处理后的信息。

[来源：GB/T 35273—2020，3.1]

3.1.5

个人信息主体 personal information subject

个人信息所标识或关联的自然人。

[来源：GB/T 35273—2020，3.3]

3.1.6

聚合数据 aggregate data

表征一组个人信息主体的数据。

[来源：GB/T 37964—2019，3.5]

3.2 缩略语

下列缩略语适用于本文件。

API：应用程序接口（Application Programming Interface）

SDK：软件开发工具包（Software Development Kit）

4 概述

互联网广告群体标识是一种由数字、字母和特殊符号构成的，通过将一组具有特定共性（如兴趣偏好、广告偏好等）的用户个体用同一符号来表示的标识符。与传统网络浏览器缓存标识（Cookie ID）、移动终端设备标识等常见的唯一标识符不同，互联网广告群体标识是以保证个体用户无法被唯一标识为前提，通过安全的参数和算法体系生成的一种可用于互联网广告投放、广告效果归因等业务的用户标识符，实现互联网广告业务开展过程中对用户个人信息的保护。

5 一般原则

5.1 匿名化原则

互联网广告群体标识需满足无法定位到唯一个体的原则，能一定程度上实现匿名化。这里的个体可是个人或终端等。

5.2 使用有效性原则

互联网广告群体标识需要满足广告产业对标识所需基础能力的要求，能一定程度上有效标记，识别互联网广告服务对象，提供基础的广告数据和特征服务。

5.3 隐私保护原则

互联网广告群体标识在工作过程中需要满足我国法律、法规和标准规范对个人信息安全保护的有关规定，保障个人信息安全。

6 总体要求

6.1 访问可控要求

群体标识符应被授权的指定应用访问。

6.2 可重置要求

对于去标识化的互联网广告群体标识，宜为用户提供标识重置行为的入口，重置后的互联网广告群体标识不可关联到重置前的群体标识。

6.3 不可重标识要求

在有限的数据和计算资源条件下，互联网广告群体标识应保证无法定位到唯一个体、无法重新关联原始数据主体以及其他可标识唯一个体的个人信息。

6.4 阶段化时效要求

针对同一个体，基于确定的参数与算法生成的互联网广告群体标识应保持一致。由于算法参数的时效性以及标识的可重置性，群体标识仅在一段时间内保持一致。

7 系统组成

互联网广告群体标识系统主要由用户、标识生成方和标识使用方三部分组成：

- 对于去标识化的群体标识，标识生成方在遵守个人信息保护、数据安全相关的国家法律法规和相应标准的情况下，可通过隐私政策获得用户对群体标识生成使用的授权，基于群体标识生成算法生成群体标识并存储在服务平台中，并定期更新版本；
- 标识使用方通过标识服务平台申请目标用户的群体标识，并开展互联网广告投放、归因等相关业务。使用方可定期获取最新版本的群体标识，提升标识信息的准确性；
- 用户可通过授权管理平台实现对去标识化的群体标识的生效状态管理。对于匿名化的群体标识，则不需要用户授权。

系统组成参见图1所示。

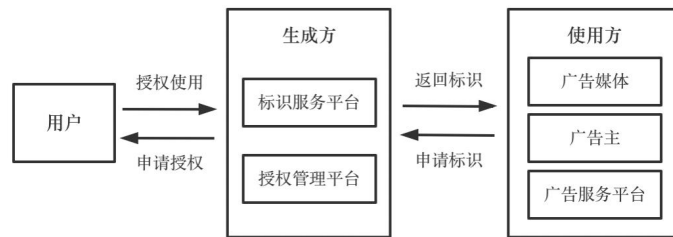


图1 互联网广告群体标识系统组成

用户：互联网广告用户包括网站、应用的注册用户，广告媒体平台用户等。

标识生成方：具备群体标识生成能力的机构或平台，提供面向使用方的标识服务平台和面向用户的授权管理平台，包括互联网流量平台、广告媒体等广告产业链主体、终端系统提供方、硬件设备提供方等。

标识使用方：基于群体标识开展互联网广告业务的参与方，包括广告媒体、广告主或广告服务平台等。

8 工作原理

8.1 工作过程

互联网广告领域群体标识工作过程见图2所示。

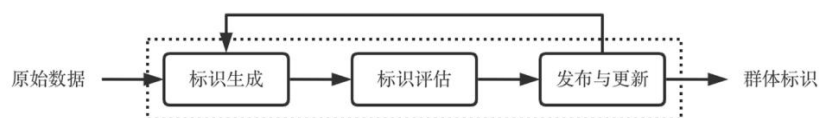


图2 互联网广告群体标识工作过程图

8.2 标识生成

互联网广告群体标识的生成过程包括群体生成和群体标识符生成两部分：

- 群体生成：使用个体的原始数据或基于推理/观察得到的行为数据，基于一定分群算法生成群体。所选算法和算法参数应在一段时间内保持稳定，保证群体内的个体作为广告受众在某些属性分布上具有一定相似性；
- 群体标识符生成：基于一定算法或编码方式为每一个群体生成一个唯一标识符。群体标识符可由数字、字母和特殊符号构成，特殊符号应避免使用互联网地址中保留字符。群体标识符可包含一定的群体属性信息，但需符合评估模型要求。

8.3 标识评估

互联网广告群体标识应具有合理的评估机制，包括群体有效性评估和数据安全性评估：

- 群体有效性评估：从信息可用性和互联网广告产业的角度，评估群体标识的使用有效性。群体标识应保证群体中的个体从分布上符合特定规律（参见附录B），能一定程度上有效标记、识别互联网广告服务对象，以满足业务使用的各项要求；

- b) 数据安全性评估：是指从个人信息保护的角度，评估群体标识的数据安全性。群体标识应保证个体的匿名性和可控的重标识风险。使用方无法基于群体标识和群体属性反推和判别群体中个体的个人数据和敏感信息。

8.4 发布与更新要求

互联网广告群体标识在发布与更新方面，包括如下的要求：

- a) 互联网广告群体标识生成方应具有合理的发布规则和定期更新能力；
- b) 互联网广告群体标识宜设定有效期限，以避免一次生成可永久使用。

8.5 保障机制要求

互联网广告群体标识在保障机制方面，包括如下要求：

- a) 互联网广告群体标识生成方应提供个人信息删除、投诉反馈渠道等方式，如果识别到风险，通过这些方式来消减风险；
- b) 互联网广告群体标识生成方应建立和提供内部的群体标识的管理机制和制度流程，确保群体标识机制能有效运行。

9 群体标识的安全要求

9.1 数据安全

互联网广告群体标识在数据安全方面，包括如下要求：

- a) 互联网广告群体标识生成时使用的数据，应遵循最小必要原则。
- b) 对于敏感信息，应保证在本地脱敏（匿名化、去标识化）处理之后进行使用。
- c) 生成群体标识时，未经用户同意不得将用户个人信息与唯一设备标识符关联使用。
- d) 应保证群体标识在传输过程中的数据安全。
- e) 数据安全要求应遵循 T/CAAAD 001-2021 第五章要求。

9.2 业务安全

互联网广告群体标识在业务安全方面，包括如下要求：

- a) 生成方和使用方在基于群体标识开展相关业务时，应保障相关程序、代码、接口、SDK、API、模型、算法等模块不引入额外的安全风险。
- b) 应通过身份认证、安全加密等机制保障未经授权的实体无法访问、篡改或恶意攻击互联网广告群体标识有关的系统。

附录 A

(资料性)

互联网广告群体标识的常见生成技术

A.1 群体生成技术

A.1.1 概述

群体生成技术是通过某种共性将个体用户进行分组的方法，主要包括属性划分、局部敏感哈希算法和聚类算法。

A.1.2 属性划分

基于属性划分的群体生成技术选取对个体有一定区分效果的单一属性或多个属性组合而成的多元组，并根据该属性或属性多元组的枚举值，对个体进行划分得到群体。

基于属性划分的群体生成技术的使用应注意以下几个方面：

- 为保证群体标识阶段一致性的特性，选择用于划分的属性也需要在一定的时间阶段内对个体保持一致性；
- 不应将违法和不良信息关键词记入用户兴趣点或者作为用户标签并据以推送信息；
- 为保证群体表示有限唯一性的特性，划分时需要确保属性或属性多元组相同的个体划分到同一群体，不应使用带有属性元组随机性的划分方式。

表 A.1 基于属性划分的群体生成示例

用户 (个体)	属性一	属性二	属性多元组 (二元组)	群组序号
用户一	A	B	(A, B)	2
用户二	A	A	(A, A)	1
用户三	A	B	(A, B)	2
用户四	B	B	(B, B)	3
用户五	A	A	(A, A)	1
用户六	B	A	(B, A)	4
...	...	...		...

A.1.3 局部敏感哈希算法

局部敏感哈希算法不同于传统的哈希算法，在对输入数据进行映射的过程中保持数据的相似性，可将相似的输入数据映射到相同的哈希分桶内，数据越相似，被分到同一个哈希桶内的概率越大，该特性使得局部敏感哈希算法被广泛应用于信息检索领域。基于局部敏感哈希算法的群体生成技术利用了局部敏感哈希的特性，首先获得可用于代表个体用户的数据，然后通过局部敏感哈希方法进行数据映射，得到该数据对应的哈希分桶。每一个哈希分桶即可作为一个群体。

基于局部敏感哈希算法的群体生成技术的使用应注意以下几个方面：

- 哈希算法参数的确定可包含随机性，但哈希计算的过程应为确定性过程，即当算法参数确定之后，对于相同的数据，其哈希计算的结果应保持一致；
- 对于个体用户的表示形式，应根据所选择的哈希方法要求的输入数据形式而确定，通常可使用向量化坐标等形式表示。

A.1.4 聚类算法

聚类算法是研究未知类别分类问题的一种统计分析方法，传统的聚类方法通常使用欧氏距离、空间密度等度量指标，通过层次化迭代、迭代重定位等方式，将分析的对象划分到多个簇中，使得这些对象与处于同一个簇的对象相似，而与处于不同簇的对象相异。除了传统的聚类方法之外，常用的聚类方法还包括基于图结构的谱聚类、基于深度学习的聚类等方法。

通常情况下,聚类算法参数的确定和个体的表达形式和基于局部敏感哈希算法的群体生成技术相似,但是在部分聚类算法中,要求的数据形式不是个体的数据表达,而是个体之间的联系,如谱聚类等方法,后者也可通过前者转换得到。

A.2 群体标识符的生成规则

A.2.1 分类

互联网广告群体标识符的生成,主要基于群体生成的结果,生成群体编码和关联属性两种标识符。

A.2.2 群体编码

主要指对群体进行直接编码,编码时通常使用数字、字母或特殊符号。群体编码仅作为群体的编号,不应包含任何群体信息及群体对应的个体信息。例如格式可定为:4-4-4的分组形式,其中,第一段为生成方的编码,第二段为版本编码,第三段为群组编码。

A.2.3 关联属性

主要指对群体属性的编码,提供对群体的描述信息。

附录 B

(资料性)

互联网广告群体标识的常见评估模型

B.1 群体有效性评估模型

B.1.1 K-匿名模型

K-匿名模型是在发布数据时保护个人信息安全的一种模型。K-匿名模型要求发布的数据中，制定标识符属性值相同的每一等价类至少包含K个记录，使攻击者不能判别出个人信息所属的具体个体，从而保护了个人信息安全。在使用K-匿名模型整合得到的数据集中，各记录之间的关联性是有限的（ $1/K$ ）。

互联网广告群体标识满足K-匿名要求，应保证每一个群体标识对应的个体数量不少于K个，即为K个以上的不同个体共同分配同一个群体标识。如表B.1所示的群体标识满足特征偏好明显的要求。

表 B.1 群体标识数据集示例

用户（个体）	群体标识符	群体属性
Alice	XXXX-YYYY-0001	性别分布：男-33.3%，女-66.7%； 年龄段：80后、90后； 购买偏好：图书、家居用品；
Bob		
Cindy		
David	XXXX-YYYY-0002	性别分布：男-66.7%，女-33.3%； 年龄段：90后、00后； 购买偏好：服装、电子产品、宠物用品；
Emma		
Frank		
...	...	...

B.2 数据安全评估模型

B.2.1 L-多样性

L-多样性是针对属性值差异性不大的数据集提出的一种增强概念。为防止确定性推导，L-多样性要求在K-匿名的基础上，实现每一等价类在每一敏感属性上存在至少L个不同值。在数据分布很不均衡时，防止推导性攻击的能力受到限制。

互联网广告群体标识满足L-多样性要求，应保证每一个群体标识对应的群体属性存在至少L个不同值。如表B.1所示的群体标识满足2-多样性要求。

B.2.2 T-接近性

T-接近性是L-多样性的增强概念，适用于发布数据集的敏感属性分布要尽可能贴近整个数据集的敏感属性分布。针对属性值分布不规则、属性值范围很小或已被分类的数据集，为防止概率性推导，要求任何等价类中敏感属性的分布与整个数据集中相应属性的分布之间的距离小于阈值T。

互联网广告群体标识满足T-接近性要求，应保证每一个群体标识对应的群体属性分布与整个数据集中的属性分布之间的距离小于阈值T。

B.2.3 群体标识的安全风险

对于无法满足L-多样性、T-接近性要求的群体对应的个体，不应为其生成广告群体标识，或设置为空值或默认值。

附录 C

(资料性)

互联网广告群体标识的基础实施方案

C.1 基于流量平台的广告群体标识实施方案

C.1.1 方案概述

基于流量平台的群体标识实施方案如图C.1所示。实现框架中主要包括三个参与方：平台用户、流量平台以及广告主。流量平台获取用户授权后，基于用户在平台使用产生的数据生成广告群体标识，面向广告主提供以群体标识为准标识符、群体特征为辅助信息的标识服务，用于广告主在本流量平台上开展互联网广告业务。

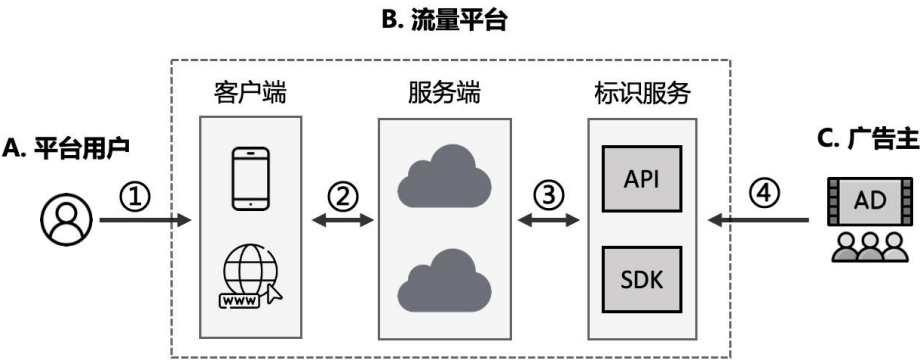


图 C.1 广告群体标识实现框架图

C.1.2 交互流程

基于流量平台的群体标识的生成与服务，主要的交互流程包括四个部分：

- a) 平台用户与流量平台的交互：流量平台在其提供的移动 APP 和网站上，提供群体标识授权管理功能，流量平台中客户端与服务端的交互：针对群体标识生成方式的不同，此处的交互分为两种情况：
 - 基于服务端生成群体标识：应取得用户充分授权后，将客户端收集的用于生成群体标识的用户信息在保证数据安全的前提下传输到服务端，或是基于横向联邦建模，在保证原始数据不出用户设备域的情况下在服务端计算得到群体标识；
 - 基于客户端生成群体标识：在客户端生成群体标识后，将群体标识在保证数据安全的前提下传输到服务端进行管理；
- b) 流量平台中服务端与标识服务的交互：流量平台通过标识服务发布和定期更新群体标识。标识服务通常以 API 或 SDK 的方式提供服务；
- c) 流量平台与广告主的交互：广告主通过调用标识服务 API、或是嵌入标识服务 SDK 获取流量平台用户对应的群体标识及群体特征信息。广告主基于群体标识及群体特征信息进行广告投放，由流量平台提供广告效果归因服务。

C.1.3 数据安全

数据安全性由以下几个方面的技术来保障：

- a) 设备域数据管理：用户个人敏感信息由客户端统一管理和存储，原始敏感数据不上传服务端，保障本地敏感数据不出用户设备域；
- b) 联邦学习技术：基于横向联邦学习技术，实现客户端与服务端间以隐私保护前提下的数据模型构建与计算；
- c) 差分隐私技术：对于必须上传服务端进行处理的数据，可使用差分隐私技术，在原始数据上引入随机噪声，在保证在数据统计可用性符合要求的同时，提供一定的隐私保护程度。

附录 D

(资料性)

互联网广告群体标识的典型应用场景

D.1 在个性化广告推荐场景中的应用举例

D.1.1 在广告投放系统中的应用举例

基于群体标识及群体属性在广告投放系统中的应用,可分为基于群体的目标受众圈选和基于群体的人群检索过滤。

在广告投放系统中,群体直接标识符及其间接标识符(如群体属性)组成群体投放端可选的属性维度,一个平台的间接标识符代表该系统维度的属性特征,这些特征即构成广告投放时的选择依赖。同时,这些特征间可进行属性特征的交并差集合的运算,完成目标受众圈选,以提升定向的丰富度。同时,不应将违法和不良信息关键词记入用户兴趣点或者作为用户标签并据以推送信息。

在人群检索过滤过程中,通过用户所带的群体标识,进行相关投放物料的过滤,选择符合投放系统设定规则的单元,再进行后续的广告环节。

对于广告投放系统中的一方数据上传,提供工具,帮助广告主将上传的数据转换为群体标识,导入投放系统进行投放。

D.1.2 在广告推荐系统中的应用举例

基于群体标识及群体属性在广告播放系统中的应用,可分为基于群体的广告召回和基于群体的广告打分排序。

以电商场景为例,群体召回使用群体统计数据以类目、品牌等属性为粒度,构建多粒度触发器,当系统离线生成群体标识及群体属性后,以群体为粒度对用户行为进行统计,得到群体近期类目、群体高频品牌等统计信息,同时通过离线算法挖掘,可得到不同类目下的高质量商品广告、各个品牌下的热门商品广告等。当广告召回系统接收到实时推荐请求时,首先根据当前请求中的群体标识,构建该群体对应的多粒度触发器,并利用这些触发器触发不同的召回分支策略,得到召回候选集。

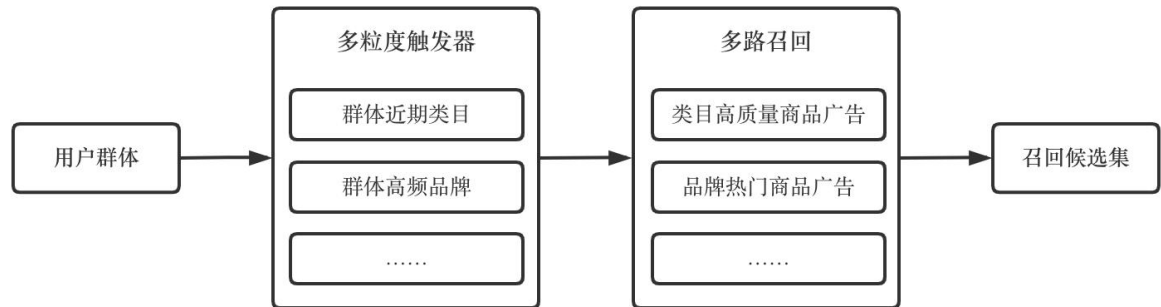


图 D.1 基于群体标识及群体属性的电商广告召回流程

群体化排序是个性化排序的一种特化,以个性化广告点击率预估算法为核心的排序算法是个性化排序的一种典型实现。以电商场景为例,个性化广告点击率预估可归纳为在给定用户属性、用户历史行为、商品(广告)信息、上下文信息的前提下,估计用户点击推荐结果的概率:

$$pCTR_{user} = p(click | user_profile, user_behavior, item, context)$$

当对用户个体进行群体标识匿名化后:

$$pCTR_{cohort} = p(click | cohort_profile, cohort_behavior, item, context), user \in cohort$$

上式中的cohort_profile指群体属性信息,可包括群体标识ID、群体统计特征(群体平均年龄、群体性别比例等),cohort_behavior指群体历史行为信息,主要为以群体粒度统计的用户历史行为数据,可包括群体高频类目等。

利用群体召回和群体排序技术,即可在用户个体经过群体标识匿名化后,在不感知具体用户及其敏感隐私信息的前提下,为用户提供一定个性化程度的广告推荐服务。

D.2 在跨平台广告跟踪场景中的应用举例

广告跟踪及归因技术是互联网广告中重要的基础技术,尤其在跨平台场景上一个在一定周期内可持续追踪的方法尤为重要。在进行追踪时,通常考虑用于追踪和归因的标识的几个重要特性:在一定周期内的区别性、一致性、一贯性。

使用群体标识进行跨平台广告追踪和归因时,可通过以下几种方式进行跨平台匹配:

- a) 群体画像匹配:平台间可将群体的描述性画像互通,依据画像的特性进行群体的匹配;
- b) 群体生成方法匹配:平台间可将生成群体的方法如算法、规则进行互通,平台间可借此识别其他平台的群体标识及所指代人群。

在群体标识和个体标识并行存在的阶段,在个体标识是合法取得用户授权的前提下,可以个体标识作为第一追踪和归因优先级,群体标识作为补充。即,群体内包含的若干个体处于已授权状态时,宜优先使用个体标识提升追踪和归因的精准性;此外宜使用群体画像匹配以及群体生成算法、规则匹配的方式将跨平台的群体标识做“对齐”处理。

D.3 群体标识下的广告效果归因

对于只能获取到群体标识的广告受众,使用群体标识进行广告效果归因。对于群体中授权了个体标识的广告受众,仍维持现有使用个体标识进行广告归因的现状。



个体数据已授权跨平台流通情况下,可以直接使用其已授权标识



隶属某群体 A_k

隶属某群体 B_s

群体 A_k 的画像 方式1: 画像匹配 群体 B_s 的画像
相似性匹配

群体 A_k 方式2: 生成方法同步 群体 B_s
算法/规则

个体数据未授权跨平台流通情况下,使用其群体标识

图 D.2 基于群体标识和个体标识的跨平台广告追踪和归因流程

附录 E

(资料性)

互联网广告群体标识编码格式与互通接口

E.1 编码格式

直接标识符主要指对群体进行直接编码，编码时通常使用数字、字母或特殊符号。直接标识符仅作为群体的编号，不应包含任何群体信息及群体对应的个体信息。例如格式可定为：2-10-6的分组形式，其中，第一段为生成域编码，第二段为版本编码，第三段为群组序号。

间接标识符主要指对群体属性的编码，提供对群体的描述信息。间接标识符可采用群体属性字段及统计值、或明文的群体标签（如兴趣偏好标签等）等形式表述，具体编码格式不作限制。

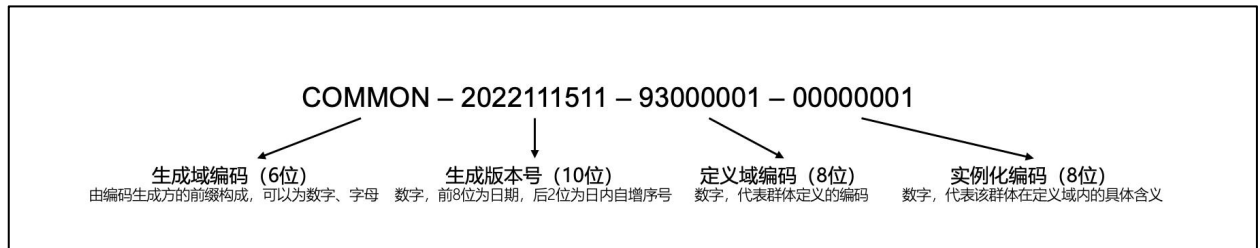


图 E.1 互联网广告群体标识直接标识符编码格式说明

E.2 互通接口

E.2.1 支持状态获取接口

表 E.1 支持状态获取接口函数

函数名 public static boolean isSupported()		
参数	返回	说明
无	boolean: 是否支持群体标识获取	true为支持，false为不支持

E.2.2 生成域版本号查询接口

表 E.2 生成域版本号查询接口函数

函数名 public static StringList getVersions(String targetDomain)		
参数	返回	说明
targetDomain	StringList: 目标生成域现存版本列表	无

E.2.3 跨生成域映射查询接口

表 E.3 跨生成域映射查询接口函数

函数名 public static String mappingCohortId(String cohortId, String targetDomain, String targetVersion, int topK, boolean returnFraction)		
参数	返回	说明
cohortId: 需要进行跨生成域映射查询的群体标识 targetDomain: 需要查询的目标生成域 targetVersion: 查询的目标生成域群体标识版本 topK: 返回的目标生成域下占比前K的群体的群体标识 returnFraction: 是否返回群体占比	String: 查询的cohortId在目标生成域给定版本下占比前K大的targetCohortId或targetCohortId及对应占比	targetVersion默认值为null，topK默认值为1，returnFraction默认值为false。在默认情况下，该接口返回cohortId对应的目标生成域最新版本的占比最高的targetCohortId

参 考 文 献

- [1]GB/T 25069—2010 信息安全技术 术语
 - [2]T/CAAAD 003-2020 移动互联网广告标识技术规范
-