

基于特征的单向数字电视节目推荐系统

朱晓航, 吕锋

武汉理工大学信息工程学院, 湖北武汉 (430000)

Email: zxhandyou@163.com

摘 要: 本文针对数字电视中业已存在的“信息过载”问题, 分析该问题的产生原因、现有的对信息过载问题的各种解决办法, 并提出了专门针对单向网络应用环境下, 前端按一定规则计算并发送节目特征值, 由终端负责根据用户行为采集用户特征值, 并根据节目特征与用户特征的匹配情况, 选择出与用户特性具有某些相似性的节目集合, 推荐给用户进一步选择。此过程对用户透明, 无需用户的主动介入, 便可提供用户其感兴趣的节目、以及前端推送针对性的增值内容。本文针对数字电视单向传输网络环境下, 提出了一种经济有效、易于部署的推荐系统的实现思路。

关键词: 数字电视 推荐系统 节目特征 用户特征

中图分类号: TN949.197

1 引言

在我国广电数字化进程中, 数字电视传输最主要的承载网为有线电视网络 (cable network, 如各地区的整转项目), 卫星 (如 NDS 直播星项目)、地波 (北京、上海等试验性项目) 等。而受限于承载网络的单向下行特性, 以及我国目前保有的大量单向数字电视接收机的现实, 单向数字电视系统仍将存很长一段时间。

模拟有线电视, 一个频点传输一路电视信号, 整条有线电视传输节目量一般在 50 个节目左右。而对数字电视系统, 前端把各个节目内容、描述数据、控制信息等封装并调制到各个频点上, 典型的一个频点带宽约 8M、符号率为 6.875MSym/s、调制方式为 64QAM, 那么一个频点的传输容量大概可以传输 8 个左右的标清视频节目。如果前端对有线电视调制传输范围为 300M~800M, 那么可以传送大约 500 个左右的标清视频节目。如果加上占用带宽较少的广播节目、数据服务等, 那么用户可以选择的节目数量相对于模拟电视时代将是爆炸性的增长。数字电视的“信息过载”问题因此而生。并且将来伴随调制解调技术的发展、节目内容的丰富、运营商的大量参与, 信息过载问题必将越发突出。

解决信息过载问题, 目前主要解决方法便是通过前端发送的 psi/si 信息^[1], 提供给终端诸如节目分类信息、节目播放信息等内容, 但此种方法仅在一定程度上减轻了用户搜寻自己真正感兴趣节目的难度, 用户仍然需要不停的主动寻找其感兴趣的节目。基于有些地方双向改造进程的发展, 已出现基于上行链路获取用户的需求信息, 进而对个人或群体提供针对性的收视方案或是其他增值服务^{[2][3][4]}。但这些方案在终端的双向更新、前端开发与改造的高昂成本制约下, 仍具有很大的局限性。

2 系统设计

单向节目推荐系统由前端按一定规则计算并发送节目特征值, 由终端负责根据用户行为采集用户特征值, 并根据节目特征与用户特征的匹配情况, 选择出与用户特性具有某些相似性的节目集合, 然后推荐给用户进一步选择。整个过程对用户透明, 无需用户主动参与。

为了在数字电视系统中实现基于特征的单向节目推荐系统, 系统前端需要定义节目的特征集合, 每个节目的特征描述按照一定组织方式, 插入 TS 流适当位置传输到终端; 终端需要实现用户特征的采集、节目特征描述的接收分析、节目特征和用户特征的匹配、用户特征值的反馈修正等工作。

2.1 特征的定义

用户特征信息是指用户兴趣爱好、关注焦点等能够反映用户个性、差异的信息。典型的用户的特征信息可以包括：（1）身份相关信息，如年龄段、性别、受教育程度、职业等等；（2）节目相关信息，喜爱的节目类别（包括一、二级节目类别）、喜爱的收视时段、喜爱的演员/导演等等。

在系统设计中，应当确定一个特征集合，确定前端对节目特征的采集方式、终端对相应用户特征的采集和反馈修正方式、终端根据节目特征和用户特征的匹配方式、以及在特征匹配中各个特征的权重值等。

2.2 特征描述符

在数字电视前端，需要根据定义的特征值的种类，对每个节目提供若干的节目特征值。前端给定的节目特征越多，则终端越容易把该节目和用户需求匹配准确。

前端需要一种方法把每个节目的特征传递到终端。可以采用的一个有效的手段就是扩展 SI，新定义一种节目特征描述符，描述符结构见下图 1：

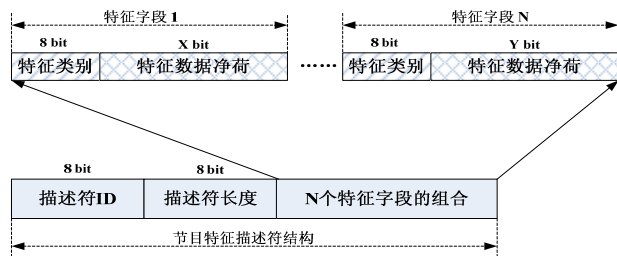


图 1 节目特征描述符

节目特征描述符是从数字电视前端传递到终端的，用来反映某个节目的若干特征，被终端用来匹配用户需求的一种描述符。

在《数字视频广播中文业务信息(SI)规范》^[1]中，规定有用户可自行使用的私有描述符 ID 范围：0x80-0xFE。可以从此 ID 范围，取一个 ID 作为节目特征描述符的 ID，图 1 中的“描述符 ID”字段应该填此约定的 ID。“描述符长度”字段指名其后跟着的若干个特征字段的总字节长度。从描述符第 3 个字节开始填写特征字段。特征字段的结构如图 1 所示，依次为 1byte 的“特征类型”，若干 byte 的“特征数据净荷”。“特征数据净荷”的长度及格式与“特征类型”相关，需要在系统设计中其进行语法和语义上的约定。

2.3 特征描述符的传输

前端不但需要在每个频点下插入频点所属节目的特征描述符，还需要在每个频点下传输其他频点节目的特征描述，这样终端可以在当前锁定频点下接收并匹配所有频点节目的特征。目前对大多数数字电视系统，EIT 表的传输符合此要求，在一个频点下一般以较短间隔传输本频点节目 EIT 事件(actual eit p/f & sch)，同是也会以稍长间隔发送其他频点所有节目的 EIT 事件(other eit p/f & sch)。因此可以简单的把特征描述符插入 EIT 表中每个事件块下的描述符集合中，随 EIT 表传送。

根据中文 SI 规范中的定义，数字电视前端发送的 EIT 表，通常包含着一个或多个事件描述块。每个事件描述块，通常包含着有短事件描述符、扩展事件描述符、内容描述符等，用来描述某一节目的事件信息。前端可以把普通节目的特征描述符插入对应的事件描述块(EIT p/f & EIT sch)中，对于 NVOD 业务可以插入参考事件描述块中。然后再把表复用到 TS 流中。

当终端从 TS 流中过滤出 EIT 表，解析并保存节目事件信息的同时，也可以解析该节目

的特征描述符,以便获取前端传来的该节目的若干特征。当终端获取到该节目的特征后,接着需要从节目特征描述符中解析出节目的各个特征,最后根据节目特征和用户特征进行匹配,按照匹配顺序把节目推荐给用户。

2.4 特征的匹配

对于事先确定的 N 个特征,终端解析出的节目特征分别表示为 $K_1 \sim K_n$, 用户特征分别为 $R_1 \sim R_n$ 。对这 N 个特征,分别给定初始权重 $W_1 \sim W_n$ 。对 N 个特征,分别定义 N 个匹配函数 $F_1(K_1, R_1) \sim F_n(K_n, R_n)$, 返回 $0 \sim 1$ 之间的值,表示对应节目和用户特征从不匹配到完全匹配的状态。

匹配指数 E 为:

$$E = \sum F_i(K_i, R_i)$$

节目计算出的匹配值 E 越高,则表示该节目和用户越匹配。

终端根据各个节目的匹配结果排序,取匹配程度最高的一定数量的节目构成推荐列表,推荐给用户选择。

2.5 反馈修正

系统终端需要能够根据用户收视行为,不断的修正内部记录的用户特征,以便动态的跟随用户喜好、收视行为的变迁。一方面终端提供用户直接干预修改特征值机会;另一方面终端可以根据用户平时收视行为以及用户对推荐系统推荐节目列表的选择情况,对特征值做出相应调整。

系统在设计时应当给出各个特征权重的动态变化范围、以及特征值取值的动态变化范围。终端在实际运行中根据用户收视行为和对节目推荐列表的选择情况,相应的对不同特征权重在规定范围内上下浮动或是终端根据用户选择节目的特征项来修正用户的特征值。

3 系统部署

系统前端需要扩展其 ψ/σ 采集与发送系统,增加节目特征值采集系统,发送时则把节目特征描述符插入对于 EIT 表中发送。系统终端可以直接把相应功能模块集成进终端软件。

对于待建的数字电视系统,可以直接在前端和终端相应集成所需功能。对于已建成的系统,前端部分可以扩展其 ψ/σ 采集与发送系统;终端功能的集成可以采用 OTA 升级等方式自动更新终端软件。

4 结 论

基于特征的单向节目推荐系统,针对单向网络,根据节目特征和用户特征匹配情况,匹配出符合用户需求的节目给用户选择,其直接目的是为了解决信息过载问题,改善用户体验。用户特征的建立和不断修正完善,也可以促进其它数字电视定向增值业务,比如广告定投等,可以有效提高 ARPU 值。本文给出了基于特征的单向节目推荐系统的一种思路。

参考文献

- [1] 国家广播电影电视总局. 数字视频广播中文业务信息规范, GY/Z174-2001, 2001
- [2] 徐江山、卢增祥等. 基于显性用户特征的数字电视节目推荐系统, 清华大学学报(自然科学版)2007 年第 47 卷第 1 期.
- [3] Zimmerman J, Kurapati K, Schaffer L. TV personalization system : Design of a TV show recommender engine and interface [C] //Personalized Digital Television. Netherlands: Kluwer Academic Publishers, 2004:27-51.
- [4] Baudisch P. Dynamic Information Filtering [D]. Sank taugustin: Forschungszentrum information stechnik GmbH, 2001.

- [5] Cotter P, Smyth B. PTV : Intelligent personalized TV Guides [C]//Seventeenth National Conference on Artificial Intelligence. Austrin, Tx, USA: AAAI Press/The MIT Press, 2000: 957-964.

Digital TV program recommendation system based on user&program profiles under one-way network

Xiaohang Zhu, Feng lv

College of Information Engineering, Wuhan University of Technology(430000)

Abstract

This paper analyze causes of "information overload" problem existing in digital television system, provide a solutions of the problem for digital television system in one-way network environment. In this solutions of information overload specifically for one-way network application, the front-end according to certain rules of shows characteristic values is calculated and sented, the terminal is responsible for collecting user behavior characteristic values, and match user's characteristics with received program characteristics, choose a user has some similarities features a collection of programs and recommend it to the user for further selection. This process transparent to the user without the user's active intervention, it could provide the user programs of interest, as well as value-added front-end push targeted content. This one-way transmission networks for digital television environment, presents a cost-effective, easy deployment system to achieve the recommended idea.

Keywords: digital TV; recommendation system; program profile; user profile

作者简介: 朱晓航, 男, 1984 年生, 硕士研究生, 主要研究方向是嵌入式、流媒体。