

基于语义关联和情景感知的个性化推荐方法研究*

李枫林 陈德鑫 梁少星

(武汉大学信息资源研究中心 武汉 430072)

摘要 当前基于内容的推荐方法存在两大缺陷:一是受项目和用户模型匹配算法的限制,使得推荐结果过于专门化;二是较少考虑情景信息,使得推荐系统缺乏情景敏感性。笔者提出了一种基于语义关联和情景感知的推荐方法,该方法通过基于语义关联的实例相似度算法改善过于专门化问题,通过情景感知推荐提升系统的情景敏感性。最后,通过实验验证了本文所提方法的有效性。

关键词 本体 个性化推荐 情景感知 语义关联 语义相似度

中图分类号 G350

文献标识码 A

文章编号 1002-1965(2015)10-0189-07

DOI 10.3969/j.issn.1002-1965.2015.10.032

Research on Personalized Recommendation Method Based on Semantic Association and Context Awareness

Li Fenglin Chen Dexin Liang Shaoxing

(Center for Studies of Information Resources of Wuhan University, Wuhan 430072)

Abstract There are two drawbacks in content-based recommendation method: one is the overspecialization problem due to the matching algorithm between items and users; the other is the lack of sensitivity for context due to limited consideration for context. This paper presents a personalized recommendation method based on semantic association and context awareness, which improves the overspecialization problem through the instances similarity algorithm based on semantic association, and upgrades the sensitivity for context through context perception recommendation. Finally, the effectiveness of the method proposed is verified by experiments.

Key words ontology personalized recommendation context awareness semantic association semantic similarity

0 引言

个性化推荐技术是缓解“信息过载”问题比较有代表性的解决方案之一,目前受到认可的基本推荐策略有两种,分别是基于内容的推荐和协同过滤推荐。基于内容的推荐方法避免了协同过滤推荐的稀疏性和冷启动问题,但是当前基于内容的推荐方法存在两个缺点:一是受余弦相似度算法的限制,只有与用户偏好有相同属性的项目才会被加入最终推荐结果集推送给用户,导致用户将仅限于获得跟以前历史偏好类似的推荐结果集,很难出现新颖性的推荐结果,即过于专门化问题(Overspecialization);二是较少考虑情景信息对推荐结果集的影响,使得推荐系统缺乏情景敏感性(Lack of sensitivity for context)。

根据本体中概念或实例的语义关联信息对推荐对象特征和用户模型进行语义拓展,以提高推荐结果的丰富性,可以改善过于专门化问题^[1]。当前学者侧重研究本体中概念相似度计算来判断文本或网页中信息资源的相似度,然而在实例作为信息资源特征表示对象的应用中,则需要计算实例相似度。为简化本体构建和相似度计算,学者较多考虑本体中的层次关系,忽略属性关系,影响了相似度计算的准确性。如文献[2]计算相似度时仅分析了层次关系,忽略了属性关系对相似度的影响,导致丢失了很多语义描述。将语义关联和情景感知结合起来,通过情景后过滤(Contextual Post-filtering)的方式,为语义关联和情景感知分配不同的权重对资源进行推荐排名^[3-5],可以改善缺乏情景敏感性的问题

收稿日期:2015-05-14

修回日期:2015-08-13

基金项目:教育部人文社会科学重点研究基地重大项目“多网络融合环境下 Context 计算模式与商务应用研究”(编号:13JJD870005)成果之一。

作者简介:李枫林(1962-),男,教授,博士生导师,研究方向:情景计算、个性化信息服务、数字策展;陈德鑫(1989-),女,博士研究生,研究方向:社会情景计算、个性化信息服务;梁少星(1990-),女,硕士研究生,研究方向:个性化信息服务。

本文提出了一种基于语义关联和情景感知的个性化推荐方法来改善基于内容推荐的两个问题。该方法首先进行基于本体的项目建模、用户建模和情景建模；其次，从语义关联实例的角度，分别分析了用户兴趣本体与推荐项目实例之间层次相交关联相似度和属性相交关联相似度对语义相似度的影响，最终给出了实例之间基于语义关联的综合语义相似度；最后，通过情景后过滤为用户兴趣和用户情景设置不同的权重，将同时符合用户兴趣和当前情景的项目推荐给用户。

1 基于本体的建模

本体是共享概念模型明确的形式化规范说明^[6]，本体由于具有良好的层次结构、支持逻辑推理以及方便知识的共享和复用，已成为主流的建模方法。本部分使用基于本体的方法分别对项目维 (Item)、用户维 (User) 和情景维 (Context) 进行建模，建立了领域本体 DomO、用户模型 UserModel 和情景模型 UserCont。

1.1 推荐项目建模 领域本体 (Domain Ontology) 是一种重要的知识组织方式，它阐释了某专业领域中的概念、实例及它们之间的关系。文献[7]和[8]利用本体技术对推荐对象建模，但在构建本体模型时只考虑了层次关系，忽略了概念之间及实例之间的属性关系，从而丢失了项目的部分语义描述。丢失语义描述将影响后续语义相似度计算的精确度。

本文全面考虑层次关系和属性关系，将领域本体形式化表示为： $\text{DomO} = \{C, R_H, R_P, I, A\}$ ，其中 C 代表概念，这些概念构成一个分类层次； R_H 指概念或实例之间的层次关系 (Hierarchy Relationship)，包括概念之间的 Subclass of 关系和概念与实例之间的 Instance of 关系； R_P 指概念之间或实例之间的属性关系 (Property Relationship)，是由用户自定义的对象属性，该属性将不同的概念和不同的实例联系起来； I 指实例，代表某个类的具体实体； A 指公理，代表永真断言。

本文使用 protégé4.3 建立电影领域本体，protégé4.3 自带的 OntoGraf 插件可以将本体中的类、实例及关系以图形的形式呈现出来。为了后面论述的需要，抽取 MOVIE 本体片段 (如图 1 所示)，该本体片段包含 14 个概念类 (1 个顶级类、4 个二级类、9 个三级类)、1 个数据属性 (MovieID) 和 3 个对象属性 (Has-

Actor、HasDirector、HasTopic)、33 个实例 (8 个电影实例、4 个电影主题实例、13 个演员实例、8 个导演实例)。

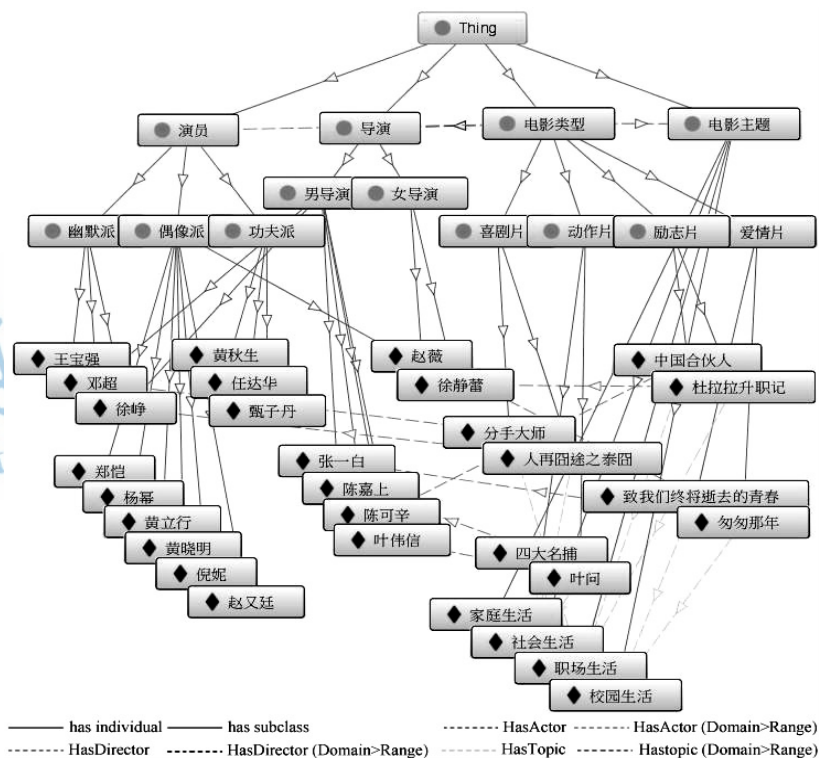


图 1 MOVIE 领域本体片段

1.2 用户建模 可以将用户模型形式化表示为一个三元组： $\text{UserModel} = (\text{UserInfo}, \text{UserDOI}, \text{UserOnto})$ 。

1.2.1 UserInfo UserInfo 代表用户基本信息，表示为 $\text{UserInfo} = \{ \text{UserID}, \text{Name}, \text{Sex}, \text{Age}, \text{Profession} \}$ ，分别表示用户 ID、姓名、性别、年龄、职业。该信息需要用户首次登录时填写。

1.2.2 UserDOI UserDOI 代表用户兴趣度，表示为 $\text{UserDOI} = \{ (I_1, D_1(t)), (I_2, D_2(t)), \dots, (I_n, D_n(t)) \}$ ，其中 $I_i (1 \leq i \leq n)$ 为用户已评分的第 i 个实例， $D_i(t)$ 为用户在 t 时刻对实例 I_i 的兴趣度 (Degree of Interest，简称 DOI)，DOI 为 $-1 \sim 1$ 的实数，正数表示用户喜欢该实例 (正偏好)，负数表示用户不喜欢该实例 (负偏好)。

其中，初始兴趣度由用户自己指定或者由推荐系统推理得出。用户的兴趣偏好会随时间发生变化，所以用户模型需要不断进行学习和更新。

1.2.3 UserOnto UserOnto 代表用户兴趣本体，表示为 $\text{UserOnto} = \{ C, R_H, R_P, I, A \}$ ，其中 C 表示用户感兴趣的概念集合， I 代表用户感兴趣的实例集合。用户兴趣本体可以通过复用领域本体建立，因此 UserOnto 一般是 DomO 的子集。

实际上，用户模型只需存储 UserInfo 和 UserDOI

即可,因为只要确定了 UserDOI,就可以根据 UserDOI 中的实例 ID 在领域本体中访问和获得这些实例的语义描述(包括实例所属概念、实例的属性及实例之间的关系等),这些语义描述构成了用户兴趣本体。

1.3 情景建模 情景维度模型是表示情景综合信息的模型。用户偏好会随着所处情景(如地点、时间、环境、用户状态等)的不同而发生变化,因此进行推荐时需要考虑用户的情景。

综合分析了 Schilit^[9]、Juan Ye^[10] 和顾君忠^[11] 对情景的不同分类方法,本文把情景模型划分为 5 个子情景维度,表示为 $UserCont = \{TempCont, SpatCont, EquiCont, EnviCont, UserStatCont\}$ 。

1.3.1 时间情景(Temporal Context) 时间情景是指用户与推荐系统发生交互的时间,可以根据具体需要按照不同的分层粒度对时间情景进行组织。

以电影推荐为例,时间维情景可以表示为 $TempCont = \{Date, DayOfWeek, TimeOfWeek, Month, Quarter, Year\}$,其中 $DayOfWeek = \{“Mon”, “Tue”, “Wed”, “Thu”, “Fri”, “Sat”, “Sun”\}$, $TimeOfWeek = \{“Weekday”, “Weekend”\}$ 。

1.3.2 空间情景(Spatial Context) 空间情景是指用户与推荐系统发生交互时所处地点的相关信息。

以电影推荐为例,用户观影地点有四种可能:电影院、家庭、公司和路上,可以表示为 $SpatCont = \{“cinema”, “home”, “company”, “way”\}$ 。

1.3.3 设备情景(Equipment Context) 设备维情景是指用户与推荐系统发生交互所使用设备的基本情况,包括硬件设施(如存储器容量)、软件设施(如操作系统)和网络互联情况(如 WiFi、4G、3G、GPRS 等)。

设备维情景可以表示为 $EquiCont = \{hardware, software, network\}$ 。

1.3.4 环境情景(Environment Context) 环境维情景是指用户周围的外在环境,如温度、湿度、亮度、噪音等,这些外在因素一般可以通过传感器获得。

环境情景可以表示为 $EnviCont = \{temperature, humidity, luminance, noise\}$ 。

1.3.5 用户状态情景(User-State Context) 用户状态情景是指与推荐系统发生交互时的用户状态,用户处于不同状态下的信息需求是不同的。

以电影推荐为例,影响用户观影决策的状态情景包括伴侣(Companion)情景和情绪(Emotion)情景。同伴类型影响观影决策,如和恋人在一起倾向于看浪漫温馨的爱情片,和朋友在一起倾向于看喜剧片等,情绪情景也会直接影响观影决策^[12]。文献[13]采集用户生理信息,并通过特征提取算法识别用户情绪,将情

绪分为 Joy、Anger、Sadness、Pleasure 四大类,本文对其进行改进,将 Joy 和 Pleasure 归为积极情绪(“positive”),将 Anger 和 Sadness 归为消极情绪(“passive”)。

因此用户状态情景可以表示为 $UserStatCont = \{Companion, Emotion\}$, $Companion = \{“alone”, “friends”, “girlfriend/boyfriend”, “family”, “co-workers”, “others”\}$, $Emotion = \{“positive”, “passive”\}$ 。

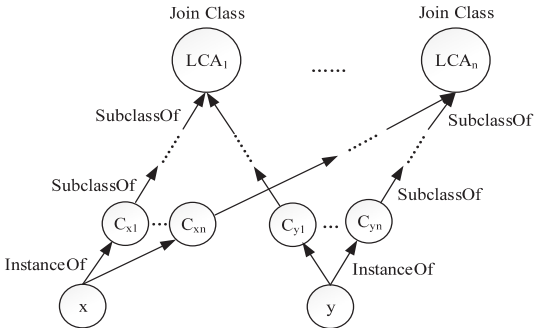
需要注意的是,不同的情景信息对推荐任务的影响程度不尽相同,因此,在推荐生成之前需要分析与推荐任务紧密相关的有效情景,在情景建模时可以视具体情况适量减少或增加一些维度。

2 基于语义关联的实例相似度算法

语义相似度计算是基于内容的推荐中项目和用户模型匹配算法的基础。在领域本体中,实例通过类或属性联系起来,因此实例之间的连通路径体现了实例之间的语义关系。本文综合层次相似度和属性相似度计算实例的语义相似度,将实例之间的连通路径分为层次相交关联和属性相交关联两类。

2.1 层次相交关联相似度 若实例 x 和 y 属于同一个类或相似类的实例,则称 x, y 之间存在层次相交关联(Hierarchy Join Association, 记为 HJA)。层次相交关联是指两个实例因拥有公共祖先而引起的关联,层次相交关联侧重于实例之间的层次关系,体现了实例之间的层次相似性。

实例存在多重继承关系,所以一个实例可能属于多个类,导致实例 x 和 y 之间存在多条层次相交关联。假设实例 x 和 y 之间存在 n 条层次相交关联且相交类为实例 x 和 y 的最近共同祖先(Lowest Common Ancestor, LCA)(如图 2 所示), $depth(x)$ 和 $depth(y)$ 表示实例 x 和 y 的深度,实例 $depth(x)$ 的深度等于实例所属类 C_x 的深度 $depth(C_x)$, $depth(LCA_i)$ 为第 i 条层次相交关联中相交节点的深度,其中 $depth(LCA) = depth(LCA_1) = depth(LCA_2) = \dots = depth(LCA_n)$ 。



第一,相交节点即最近公共祖先在本体层次树中的深度 $\text{depth}(\text{LCA})$ 。最近公共祖先的深度越深,实例 x 和 y 越具体,层次相交关联相似度也越大。

第二,层次相交关联所在分支的最大深度 $\max(\text{depth}(x), \text{depth}(y))$ 。分支的最大深度越深,节点离共同祖先越远,节点之间的层次相交关联相似度也越小。

因此,实例 x 和 y 之间的层次相交关联相似度表示为:

$$\text{SimHJA}(x, y) = \frac{\text{depth}(\text{LCA})}{\max(\text{depth}(x), \text{depth}(y))} \quad (1)$$

如果两个实例的最近公共祖先(LCA)为根节点,则它们的层次相交关联相似度为0。

如果实例 x 和 y 属于同一个类,则层次相交关联相似度为1,因为 $\text{depth}(\text{LCA}_i) = \text{depth}(x) = \text{depth}(y)$ 。

以图1中MOVIE本体的两个电影实例 I_1 (《匆匆那年》)和 I_3 (《致青春》)为例,两者共存在1条层次相交关联(如图3所示),根据公式(1)计算 I_1 和 I_3 的层次相交关联相似度为:

$$\text{SimHJA}(I_1, I_3) = \frac{\text{depth}(\text{LCA})}{\max(\text{depth}(I_1), \text{depth}(I_3))} = \frac{\text{depth}(\text{爱情片})}{\max(\text{depth}(C_n), \text{depth}(C_b))} = \frac{2}{2} = 1$$

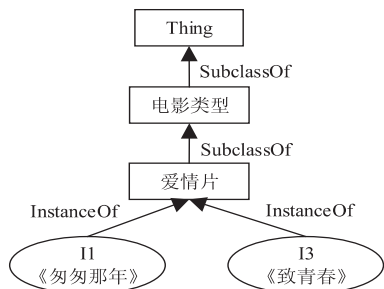


图3 电影实例 I_1 与 I_3 之间的层次相交关联

2.2 属性相交关联相似度 若实例 x 和 y 具有共同或相似的属性,即实例 x 和 y 通过属性连接起来,则称 x 、 y 之间存在属性相交关联(Property Join Association, 记为 PJA)。属性相交关联是两个实例因存在共同属性或相似属性而产生的关联,属性关联侧重用户自定义的对象属性关系,体现实例之间的属性相似性。

假设实例 x 和 y 之间存在 n 条属性相交关联(如图4所示),第 i 条路径相交关联的长度为 $\text{length}(\text{PJA}_i)$,实例 x 和 y 的属性个数为 k 。通过对图4进行分析,可知 x 和 y 之间的属性相交关联相似度与以下因素有关:

第一,属性相交关联条数 n 。属性相交关联越多,说明两个实例相同或相似的属性值越多,两个实例也越相似。

第二,语义关联长度 $\text{length}(\text{PJA}_i)$ 。语义关联的

长度越长,说明两个实例相隔的连通路程越长,相似度也越小。实例间语义关联长度等于实例间连通路程所包含的边的数量。

因此,实例 x 和 y 之间的属性相交关联相似度可以表示为:

$$\text{SimPJA}(x, y) = \frac{1}{k} \cdot \sum_{i=1}^n \frac{1}{\text{length}(\text{PJA}_i)} \quad (2)$$

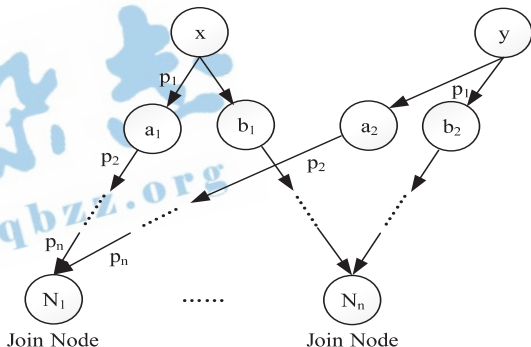


图4 实例 x 和 y 之间的属性相交关联

仍然以图1中MOVIE本体的两个电影实例 I_1 (《匆匆那年》)和 I_3 (《致青春》)为例,两者共存在6条属性相交关联(如图5所示),根据公式(2)计算 I_1 和 I_3 的属性相交关联相似度为:

$$\text{SimPJA}(I_1, I_3) = \frac{1}{6} \cdot \left(\frac{1}{\text{length}(\text{PJA}_1)} + \frac{1}{\text{length}(\text{PJA}_2)} + \frac{1}{\text{length}(\text{PJA}_3)} + \frac{1}{\text{length}(\text{PJA}_4)} + \frac{1}{\text{length}(\text{PJA}_5)} + \frac{1}{\text{length}(\text{PJA}_6)} \right) = \frac{1}{6} \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{4} + \frac{1}{6} + \frac{1}{2} \right) = \frac{23}{72} = 0.32$$

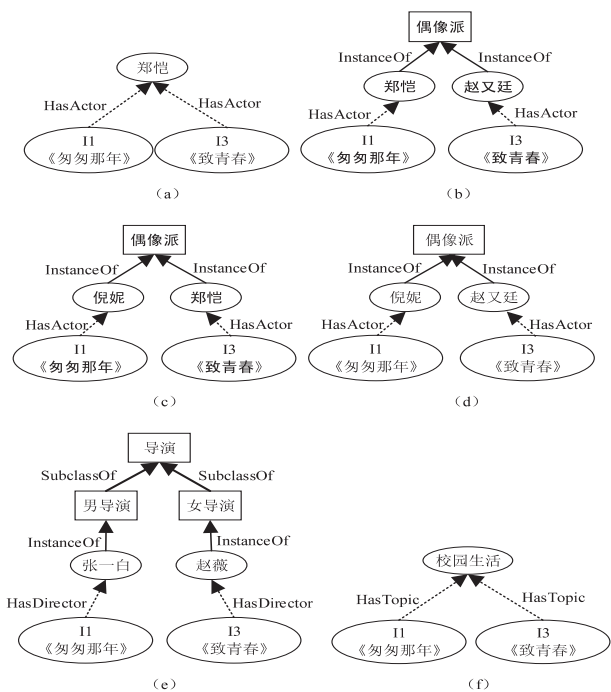


图5 电影实例 I_1 与 I_3 之间的属性相交关联

2.3 基于语义关联的综合语义相似度算法 两个

实例之间可能存在多种语义关联,设层次相交关联相似度的权重为 $\alpha (0 \leq \alpha \leq 1)$,属性相交关联相似度的权重为 $\beta (0 \leq \beta \leq 1)$,且 $\alpha + \beta = 1$,则本体中任意两个实例 x 和 y 的综合语义相似度为:

在具体应用中,可以根据需要对公式(3)中的权重进行调整。

$$\text{SemSim}_{(x,y)} = \alpha \cdot \text{SimHJA}_{(x,y)} + \beta \cdot \text{SimPJA}_{(x,y)} \quad (3)$$

3 基于语义关联和情景感知的个性化推荐方法

情景感知环境下的个性化推荐不仅需要考虑用户维与项目维的相似度匹配,而且需要考虑情景维对推荐结果的影响。基于语义关联和情景感知的推荐方法具体实现步骤如下:

第一,基于语义关联的推荐。根据兴趣本体 UserOnto 和兴趣度 UserDOI,利用本文第 2 节提出的基于语义关联的实例相似度算法进行用户维与项目维 (User * Item) 的匹配,并按照预测兴趣度的高低产生基于语义关联的推荐列表。

第二,基于情景感知的推荐。根据当前情景 UserCont 进行情景维与项目维 (Context * Item) 的匹配产生基于情景感知的推荐列表。

第三,基于语义关联和用户情景的综合推荐。根据用户兴趣和用户情景对推荐结果的影响权重,将第一步和第二步生成的推荐列表进行加权 (User * Context * Item),生成最终的综合排序推荐列表。

3.1 基于语义关联的推荐 基于语义关联的推荐通过对初始偏好实例进行语义扩展,可以发现更多语义相关的实例,从而提高了推荐结果的多样性,改善了过于专门化问题。

基于语义关联的推荐包含以下几点:

第一,用户偏好扩散。对用户兴趣本体 UserOnto 进行偏好扩散的目的是为了发现更多与用户模型中实例存在语义关联的实例,从而丰富最终的推荐结果集,其步骤如下:

第 1 步,根据兴趣模型 UserDOI 中的实例 ID 号,在领域本体 DomO 中定位所有已评分的实例。

第 2 步,根据层次相交关联和属性相交关联的特点,在领域本体 DomO 中找出与 UserOnto 中的实例有语义关联的节点。

第 3 步,随着所有的语义关联被遍历,最终发现所有与 UserOnto 中用户已评分实例存在语义关联的其他实例,记为 $I = \{I_1, I_2, \dots, I_n\}$ 。

第二,实例兴趣度预测。根据用户偏好扩散后某实例与用户兴趣本体的相似度,预测用户对该实例的

兴趣度,其计算步骤如下:

第 1 步,根据公式(1)、公式(2)和公式(3),分别计算实例 $I_i (I_i \in I)$ 与 UserOnto 中实例 $I_j (I_j \in \text{UserOnto})$ 的相似度 $\text{SemSim}(I_i, I_j)$ 。

第 2 步,根据实例 I_i 与用户兴趣本体的语义相似度来预测用户对该实例的兴趣度 $D(I_i)$,计算公式为:

$$D(I_i) = \text{SemSim}(I_i, \text{UserOnto}) = \sum_{j=1}^n \text{SemSim}_{(I_i, I_j)} \cdot D(I_j) \quad (4)$$

第 3 步,根据预测兴趣度的高低产生 Top-N 基于语义关联的推荐列表。

3.2 基于情景感知的推荐 传统推荐系统较少考虑情景信息对推荐结果的影响,使得推荐系统缺乏情景敏感性。将情景信息用于推荐系统可以改善信息推荐准确度和用户满意度^[5,14]。

对于特定的推荐领域和推荐任务,不同类型的情景信息对推荐结果的影响程度不尽相同,因此需要分析有效情景对推荐结果的影响,最终生成 Top-N 基于情景感知的推荐列表。

以电影推荐中的情绪情景为例,用户情绪分为积极和消极两类,积极情绪下 (motion = “positive”) 的观影偏好为:爱情片、喜剧片、动作片、励志片,消极情绪下 (motion = “passive”) 的观影偏好为:喜剧片、励志片、爱情片、动作片^[11]。可以根据用户的情绪状态向其推荐符合当前情景的电影。

3.3 基于语义关联和用户情景的综合推荐 在不同的推荐应用中,用户兴趣和用户情景对推荐结果的影响权重不同,如在基于位置的旅馆推荐中,用户位置情景的权重远远超过用户兴趣的权重。

假设用户兴趣对推荐结果的影响权重为 m ,用户情景对推荐结果的影响权重为 n ,其中 $m + n = 1$ 。设实例 I_i 在基于语义关联的推荐列表中排序为 $R_1(I_i)$,在基于情景感知的推荐列表中排序为 $R_2(I_i)$,则在基于语义关联和用户情景的综合推荐中,实例 I_i 加权后的排序为:

$$R_3(I_i) = m * R_1(I_i) + n * R_2(I_i) \quad (5)$$

最后,按照加权排序的高低生成最终同时符合用户兴趣和用户情景的综合排序推荐列表,并将这些项目推荐给目标用户。

4 实验及结果分析

为了验证基于语义关联和情景感知的个性化推荐方法的有效性,本节将设计实验并对实验结果进行分析和评估。

4.1 实验设计

4.1.1 推荐场景与本体模型

本实验对用户情

景进行了简化,仅考虑用户的兴趣偏好和情绪情景进行电影推荐。假设某用户小明(UserID = 1)看过 2 部电影 I_1 《匆匆那年》和 I_2 《四大名捕》,兴趣度分别为 0.9 和 0.2,系统感知小明当前情绪为“passive”,系统将根据用户兴趣和情景向小明推荐电影。

小明的用户模型可以表示为 $UserModel = (UserInfo, UserDOI, UserOnto)$, 其中 $UserInfo = \{“1”, “小明”, “男”, “20”, “大学生”\}$, $UserDOI = \{(I_1, 0.9), (I_2, 0.2)\}$, $UserOnto$ 通过复用本文 1.1 节定义的电影领域本体建立(如图 6 所示)。

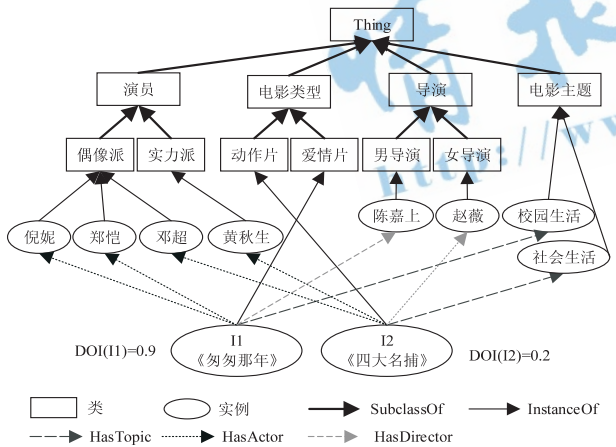


图 6 基于领域本体的用户兴趣本体 UserOnto

小明的情景模型可以表示为 $UserCont = \{Emotion\} = \{“passive”\}$ 。

4.1.2 基于语义关联和情景感知的电影推荐
根据小明的用户模型和情景模型,使用本文提出的基于语义关联和情景感知的个性化推荐方法为小明推荐电影。

a. 基于语义关联的电影推荐。首先,对图 6 的用户兴趣本体 UserOnto 进行偏好扩散,在领域本体中找出与 UserOnto 中实例 I_1 和 I_2 存在语义关联的其他实例,用户偏好扩散后的实例集为 $I = \{(I_3, 《致我们终将逝去的青春》), (I_4, 《人再囧途之泰囧》), (I_5, 《分手大师》), (I_6, 《叶问》), (I_7, 《中国合伙人》), (I_8, 《杜拉拉升职记》)\}$ 。

其次,根据公式(1)、(2)和(3)计算偏好扩散后的电影实例与 UserModel 中 2 个电影实例 I_1 和 I_2 的相似度。公式(3)中取 $\alpha = 0.5, \beta = 0.5$,则计算结果如表 1 所示。

最后,根据公式(4)预测用户对待推荐实例的兴趣度,它们分别为 $DOI(I_3) = 0.66, DOI(I_4) = 0.39, DOI(I_5) = 0.4, DOI(I_6) = 0.44, DOI(I_7) = 0.4, DOI(I_8) = 0.401$ 。因此根据预测兴趣度大小产生的基于语义关联的推荐列表 R_1 为:

$R_1 = \{(1, I_3), (2, I_6), (3, I_8), (4, I_5), (4, I_7), (5, I_4)\}$ 。

表 1 UserModel 中 2 个电影实例与其他 6 个实例的相似度

电影 ID	与 I_1 的相似度			与 I_2 的相似度		
	层次相交	属性相交	综合语义	层次相交	属性相交	综合语义
	关联相似度	关联相似度	相似度	关联相似度	关联相似度	相似度
I_3	1	0.32	0.66	0.5	0.18	0.34
I_4	0.5	0.19	0.35	0.5	0.26	0.38
I_5	0.5	0.22	0.36	0.5	0.25	0.38
I_6	0.5	0.19	0.35	1	0.26	0.63
I_7	0.5	0.22	0.36	0.5	0.25	0.38
I_8	0.5	0.24	0.37	0.5	0.18	0.34

b. 基于情绪情景的电影推荐。消极情绪下(motion = “passive”)用户的观影偏好顺序为:喜剧片、励志片、爱情片、动作片。因此,依据小明的情绪情景“passive”生成的基于情景感知的推荐列表 R_2 为:

$R_2 = \{(1, I_4), (1, I_5), (2, I_7), (2, I_8), (3, I_3), (4, I_6)\}$

c. 基于语义关联和用户情景的电影推荐。假设用户兴趣对推荐结果的影响因子 $m = 0.4$, 用户情景对推荐结果的影响因子 $n = 0.6$, 根据公式(5) 计算实例的综合排序为: $R_3(I_3) = 2.2, R_3(I_4) = 2.6, R_3(I_5) = 2.2, R_3(I_6) = 3.2, R_3(I_7) = 2.8, R_3(I_8) = 2.4$ 。

最后,根据按综合排序由低到高生成综合推荐列表 R_3 为:

$R_3 = \{(1, I_3), (1, I_5), (2, I_8), (3, I_4), (4, I_7), (5, I_6)\}$ 。

4.2 基于语义关联的推荐结果分析
传统相似度算法中,只有当两个实例有相同属性值时(如“相同导演”“相同演员”)才认为两者有相似性,而本文提出的相似度计算方法认为:不仅两个实例有相同属性值时存在相似性,而且属性值不同时(如“不同演员”但“同派系”)也存在相似性。

图 5 展示了电影 I_1 与 I_3 之间的 6 条属性相交关联,不仅考虑了相同属性对语义相似度的影响(如图 5(a)共同演员“郑恺”和图 5(f)的共同主题“校园生活”),而且考虑了相似属性对语义相似度的影响(如图 5(b)、图 5(c)和图 5(d)不同姓名但同派系演员等)。通过该相似度算法不仅可以向用户推荐与用户偏好有相同属性的项目,而且可以推荐与用户偏好存在语义关联的项目,最终提高推荐结果的多样性,改善过于专门化问题。

4.3 基于语义关联和情景感知的推荐结果分析
传统推荐方法仅考虑了兴趣偏好,忽略了情景信息的影响。本文提出的基于语义关联和情景的方法,通过赋予兴趣偏好和当前情景不同的权重因子,综合考虑了两者对推荐结果的影响。下面使用表 2 对比了三种情况下(即仅考虑用户兴趣、仅考虑用户情景和综合用户兴趣和情景) 电影推荐列表 R_1, R_2 和 R_3 中实例排

序的变化:

表 2 推荐列表 R_1 、 R_2 和 R_3 中电影实例排序的对比			
Movie ID	基于语义关联 的推荐列表 R_1	基于情景感知 的推荐列表 R_2	综合推荐 列表 R_3
I_3	1	3	1
I_4	5	1	3
I_5	4	1	1
I_6	2	4	5
I_7	4	2	4
I_8	3	2	2

由表 2 可知,如电影 I_6 ,虽然该实例与用户兴趣本体相似度较大,在基于用户兴趣的排序中排序为 2,但是由于用户当前处于“消极情绪”,观影偏好顺序为:喜剧片、励志片、爱情片、动作片,而 I_6 属于“动作片”,因此综合考虑用户兴趣和情景后,其排序由 2 降为 5。因此,根据用户情景对推荐结果进行调整,可以提高系统的情景敏感性,使推荐系统更加人性化。

5 结 语

本文提出了一种基于语义关联和情景感知的个性化推荐方法,该方法通过基于语义关联的实例相似度算法改善了基于内容推荐中的过于专门化问题,通过情景感知推荐提升系统的情景敏感性,最终将同时满足用户兴趣和当前情景的项目推荐给用户。

基于语义关联的实例相似度算法不仅可以用于基于内容的推荐中来改善专门化问题,而且可以用于协同过滤推荐和信息检索中以改善前者的稀疏性问题和后者的查全率和查准率不高的问题。我们后续工作将研究该算法在协同过滤推荐和检索领域的应用。

参 考 文 献

[1] Cantador I, Bellogín A, Castells P. Ontology-based personalised and Context-aware Recommendations of News Items [C]//Proceedings of the 2008 IEEE/WIC/ACM International

Conference on Web Intelligence and Intelligent Agent Technology-Volume 01 [A]. IEEE Computer Society, 2008:562-565.

[2] 陈沈焰, 吴军华. 基于本体的概念语义相似度计算及其应用 [J]. 微电子学与计算机, 2009, 25 (12):96-99.

[3] Aleman-Meza B, Halaschek C, Arpinar I B. Context-aware Semantic Association Ranking [C]// Cruz I F, Kashyap V, Decker S, et al. Proceeding of the 1st International Workshop on Semantic Web and Databases [A]. Berlin, Germany: Humboldt-Universität, 2003:33-50.

[4] Barnaghi P M, Abdul Kareem S. A Context-Aware Ranking Method for the Complex Relationships on the Semantic Web [C]// Advanced Language Processing and Web Information Technology, International Conference on IEEE Computer Society [A], 2007:129-134.

[5] 杨 君, 吴菊华, 艾丹祥. 一种基于情景相似度的多维信息推荐新方法研究 [J]. 情报学报, 2013, 32 (3):262-269.

[6] Studer R, Benjamins V R, Fensel D. Knowledge Engineering: Principles and Methods [J]. Data & Knowledge Engineering, 1998, 25 (97):161-197.

[7] 周 莉, 潘旭伟, 谢玉开. 情境感知的电子商务个性化商品信息服务 [J]. 图书情报工作, 2011 (10):130-134.

[8] 陈 钰, 张功亮, 阚述贤, 等. 一种基于领域本体的用户建模方法 [J]. 计算机与数字工程, 2011, 39 (2):86-89.

[9] Schilit B, Adams N, Want R. Context-Aware Computing Applications [C]// In Proceedings of the Workshop on Mobile Computing Systems and Applications [A], 1994:85-90.

[10] Ye J, Coyle L, Dobson S, et al. Ontology-based Models in Pervasive Computing Systems [J]. The Knowledge Engineering Review, 2007, 22 (4):315-347.

[11] 顾君忠. 情景感知计算 [J]. 华东师范大学学报 (自然科学版), 2009 (5):1-20.

[12] 李 晟. 基于情境感知的个性化电影推荐 [D]. 北京:北京邮电大学, 2012.

[13] 聂 聆, 王晓韡, 段若男, 等. 基于脑电的情绪识别研究综述 [J]. 中国生物医学工程学报, 2012, 31 (4):595-606.

[14] 王立才, 孟祥武, 张玉洁. 上下文感知推荐系统 [J]. 软件学报, 2012, 23 (1):1-20.

(责编:刘影梅)