

改进型动态度量推荐算法

张洪军¹ 杨霞²

(1、南京国电南自电网自动化有限公司,江苏 南京 210000 2、易思奇汽车电子(上海)有限公司,江苏 南京 210000)

摘要:随着电子商务的普及,越来越多的用户通过电子商务网站来购物,电子商务网站上的商品也越来越多。电子商务系统使用最多的推荐算法是协同过滤算法,但是传统的协同过滤推荐算法中存在覆盖率低、冷启动、数据稀疏、推荐准确率低、不安全等问题。基于协同过滤的推荐系统默认可信网络内的用户都是可信的,没有对可信网络内的信任度进行测量。为了解决这种问题,本算法在基于 CF 的推荐系统中引入可信度测量的概念,将低于预定义阈值的可信网络来重构重新生成可信网络。重新生成的可信网络可以提高推荐结果的准确性和用户的覆盖率。

关键词:推荐系统;信任感知;可信网络;可信度计算

中图分类号:TP391

文献标识码:A

文章编号:2096-4390(2021)20-0093-03

1 概述

电子商务、搜索引擎中应用范围最广的推荐算法是由 Goldberg 等人提出的协同过滤推荐算法 (collaborative filtering recommendation algorithm)。传统的协同过滤算法基于两种假设:(1)用户的喜好是固定不变的;(2)对商品评分相同的用户之间具有相同的喜好。

协同过滤算法基于这两个假设来构建用户的可信网络并使用该可信网络来计算推荐条目。随着社会的发展、新事物的出现、用户习惯的变化、历史数据时间跨度的变大,用户的兴趣爱好发生了很大的变化。这就导致协同过滤算法中的两种假设误差变大,相同用户集误差的变大会导致可信网络的准确度变低。为了解决协同过滤算法中存在的这些问题,本文在协同过

滤算法的基础上引入了动态度量的概念,提出了一种动态度量可信度的推荐算法 (dynamic reliability measure recommendation algorithms, DRMAR)。DRMAR 算法通过相似度与信任的耦合值来替代相似度值,通过耦合的相似度值来构建可信网络并计算初始评分值,然后利用 DRMAR 算法来重构所有低于预定值的可信网络。DRMAR 算法在 Epinions、Flixster 数据集上进行了对比测试,测试结果显示与主流的推荐算法比较 DRMAR 算法在用户覆盖率、推荐条目准确率上面有所提高。

2 推荐系统常用算法介绍

协同过滤算法:

Goldberg 等人提出的协同过滤算法是推荐系统中使用率最高的算法,协同过滤算法两种假设:(1)用户的喜好 (转下页)

结果如图 5 所示,可以看出 AR1 选择了 AR2、AR3 的路径去往电信机器 3.3.3.3,实现了路由的冗余备份。

```
PC>tracert 2.2.2.2

tracert to 2.2.2.2, 8 hops max
(ICMP), press Ctrl+C to stop
 1  1.1.1.1    16 ms  <1 ms  15 ms
 2  60.1.1.2   31 ms  16 ms  16 ms
 3  23.1.1.1   31 ms  16 ms  15 ms
 4  2.2.2.2    31 ms  16 ms  31 ms
```

图 4 端口故障后去往联通的测试结果

```
PC>tracert 3.3.3.3

tracert to 3.3.3.3, 8 hops max
(ICMP), press Ctrl+C to stop
 1  1.1.1.1    31 ms  <1 ms  16 ms
 2  50.1.1.2   16 ms  15 ms  32 ms
 3  23.1.1.2   31 ms  31 ms  31 ms
 4  3.3.3.3    47 ms  47 ms  31 ms
```

图 5 端口故障后去往电信的测试结果

4 结论

本文基于 ENSP 仿真平台,根据公司网络出口实际需求,在路由器静态路由配置的基础上,对浮动静态路由进行了详细的分析、设计和配置,并进行了网络路径测试,实现了网络的高可靠性,为多出口网络的路由设计提供了有效的参考。

参考文献

- [1]徐鹏.在华为 eNSP 平台上实现企业级网络模拟与仿真[J].电脑知识与技术,2019,15(36):63-65.
- [2]叶涛,王思齐,杨建彪.基于 eNSP 的大规模路由综合设计与仿真实验[J].实验室研究与探索,2019,38(04):109-114.

基金项目:防灾科技学院教育研究与教学改革项目(项目编号:JY2020B30)。

作者简介:潘志安,男,硕士,研究方向为网络工程、软件定义网络。

是固定不变的;(2) 对商品评分相同的用户之间具有相同的喜好。协同过滤算法计算方法如下:(1)预定义相似值,并使用该值来计算用户之间的距离权重;(2) 利用距离权重来生成可信网络;(3)根据可信网络来生成推荐推荐条目。协同过滤算法检索历史数据中用户对不同条目的评分并根据历史评分来查找邻居,通过检索邻居的购买条目来生成当前用户的推荐条目。协同过滤算法可以分为两类:基于项目的方法(Project-Base)、基于模型的方法(Mode-base)。Project-Base 算法会首先遍历该用户的历史评分,通过历史评分记录来计算用户、条目之间的相似值,并根据该相似值来直接访问包含所有用户评分、建议的用户-商品评分矩阵并为当前活跃用户生成预测评分。Mode-Base 算法使用 AI 算法来生成统计模型,并使用生成的统计模型来为不可见商品进行评分。

3 改进型可信度动态度量推荐算法

为了解决传统的协同算法中由于用户矩阵准确率低、依赖于用户历史评分造成的推荐结果准确率低及恶意评分导致的一系列安全风险。本文引入了可靠性动态度量的概念提出了一种新的算法:可信度动态度量的推荐算法(dynamic reliability measure recommendation algorithms, DRMRA)。DRMRA 算法将可靠性度量值、信任声明进行耦合成新的动态度量值来跟提高推荐系统准确性。算法的改进主要体现在以下方面:(1)将相似性矩阵与用户信任声明矩阵耦合来生成全新的动态可信网络;(2)提出积极/消极因子动态度量可信值;(3)使用基于信任的动态可靠性度量值来重建用户的信任网络。DRMRA 工作流程由六部分组成:(1)耦合信任网络构造;(2)初始评分预测;(3)基于信任的动态可信度计算;(4)耦合可信网络重构;(5)用户最终评分预测;(6)Top-N 商品选择。

3.1 耦合信任网络构建

DRMRA 算法在传统的协同过滤算法信任网络构建的基础上进行了改进,改进如下:

(1)将用户信任值、皮尔逊系数测量值合并生成耦合相似度值;(2)使用耦合相似度值来计算用户邻居的加权有向图。每个用户是加权有向图的一个节点,用户之间的相似度值是用户节点之间的权重值。用户对之间的信任声明 $T_{a,u}$ 、信任传播距离 d_{\max} 计算如下:

$$T_{a,u} = \frac{d_{\max} - d_{a,u} + 1}{d_{\max}} \quad d_{\max} = \frac{\ln(n)}{\ln(k)}$$

用户之间的权重调整方法如下:

$$w_{a,u} = \frac{2 \times \text{sim}(a,u) \times T_{a,u}}{\text{sim}(a,u) + T_{a,u}},$$

如果 $\text{sim}(a,u) + T_{a,u} \neq 0, \text{sim}(a,u) \times T_{a,u} \neq 0$

$$w_{a,u} = T_{a,u}, \text{ 如果 } \text{sim}(a,u) = 0, T_{a,u} \neq 0$$

$$w_{a,u} = \text{sim}(a,u), \text{ 如果 } \text{sim}(a,u) \neq 0, T_{a,u} = 0$$

$$w_{a,u} = 0, \text{ 其他}$$

3.2 初始评分预测

用户未选购商品的默认评分是通过初始可信网络来计算的,用户 u 对于商品 i 的预测评分计算方法如下:

$$P_{a,i} = \bar{r}_a + \frac{\sum_{u \in K_{a,i}} w_{a,u} (r_{u,i} - \bar{r}_u)}{\sum_{u \in K_{a,i}} w_{a,u}}$$

其中 \bar{r}_a 表示用户 a 的平均评分, $K_{a,i}$ 表示商品 i 跟用户 a 具有相同评分的用户集, $w_{a,u}$ 表示用户 a 跟用户 u 之间的相似权重即可信网络中两个节点之间的边长。

3.3 基于信任的动态可信度计算

传统的协同过滤算法中是在用户评分矩阵中使用一个正因子一个负因子来计算可信度,因此这种方法不适用于基于信任的协同过滤算法。本文在传统的协同过滤算法的基础上做了如下改进:(1)在传统协同过滤算法基础上引入了信任声明;(2)将信任值引入正负因子中,提出了一种全新的基于信任的动态可信度计算方法,将基于信任的动态可信度度量值做为可信网络重构的反馈依据,基于信任的动态可信度度量值计算方法如下:

$$R_{a,i} = (n_s(S_{a,i}) \times n_v(V_{a,i})^{n_t(S_{a,i})})^{\frac{1}{1+n_t(S_{a,i})}}$$

其中 $f_s(S_{a,i})$ 表示正因子 $f_v(V_{a,i})$ 表示负因子,正负因子计算方法如下:

$$n_s(S_{u,i}) = 1 - \frac{\bar{S}}{S + S_{u,i}}, n_v(V_{u,i}) = \left(\frac{\max - \min - \partial_{u,i}}{\max - \min} \right)^\beta$$

$$\beta = \frac{\ln 0.8}{\ln \frac{\max - \min - \partial}{\max - \min}}, \partial_{u,i} = \frac{\sum_{u \in L_{a,i}} w_{a,u} \times (r_{u,i} - \bar{r}_u - P_{a,i} + \bar{r}_a)^2}{\sum_{u \in L_{a,i}} w_{a,u}}$$

3.4 耦合可信网络重构

可信度计算完成后利用可信度度量值来评估预测评分的质量,当前用户 a 对于商品 i 的可信值 $R_{a,i}$ 小于预定义的阈值 θ 那么就需要重构可信网络。我们在可信网络重构过程中引入正因子 $w_{a,u}$ 和负因子 $V(u)$ 来重构可信网络,将可信网络中的无用用户删除来提高可信网络的质量。负因子 $V(u)$ 计算方法如下:

$$V(u) = \frac{w_{a,u} \times (r_{u,i} - \bar{r}_u - P_{a,i} + \bar{r}_a)^2}{4 \times (\max - \min)^2}$$

3.5 Top-N 商品选择

当可信网络重构完成后,我们用可信网络来预测 Top-N 条商品推荐给用户。

4 实验结果与分析

4.1 实验环境设定

为了评估 DRMRA 算法的效果我们选取了 Epinions、Flixster 两个数据集进行验证,初始用户-商品评分矩阵、用户信任矩阵初始值如表 1 所示,商评分值的范围为[1,10]步进值 1。

我们将数据集中的用户分为三类:新注册用户、活跃用户、恶意用户。我们将商品分为如下两类:冷门商品、有争议商品。为了比对 DRMRA 算法在推荐准确率、用户覆盖率方面的改进,我们选择了 CF、TARS、TCF、RTW、BIBR 五种算法来做对比测试,使用留一法(LOO)在平均绝对值误差(MAE)、平均用户误差(MAUE)、评分覆盖率(RC)、用户覆盖率(UC)四个方面进行对比。实验过程中试验参数设定如下,可信网络重建阈值 $\theta = 0.7$,

表 1 用户信任、用户商品评分矩阵

	i_1	i_2	i_3	i_4	i_5		u_1	u_2	u_3	u_4	u_5
u_1	-	5	8	10	-	u_1	-	8	-	-	2
u_2	-	5	6	8	4	u_2	-	-	4	-	-
u_3	9	-	-	1	6	u_3	-	3	-	3	2
u_4	8	10	4	2	-	u_4	6	-	5	-	-
u_5	-	1	6	5	7	u_5	4	-	3	-	4

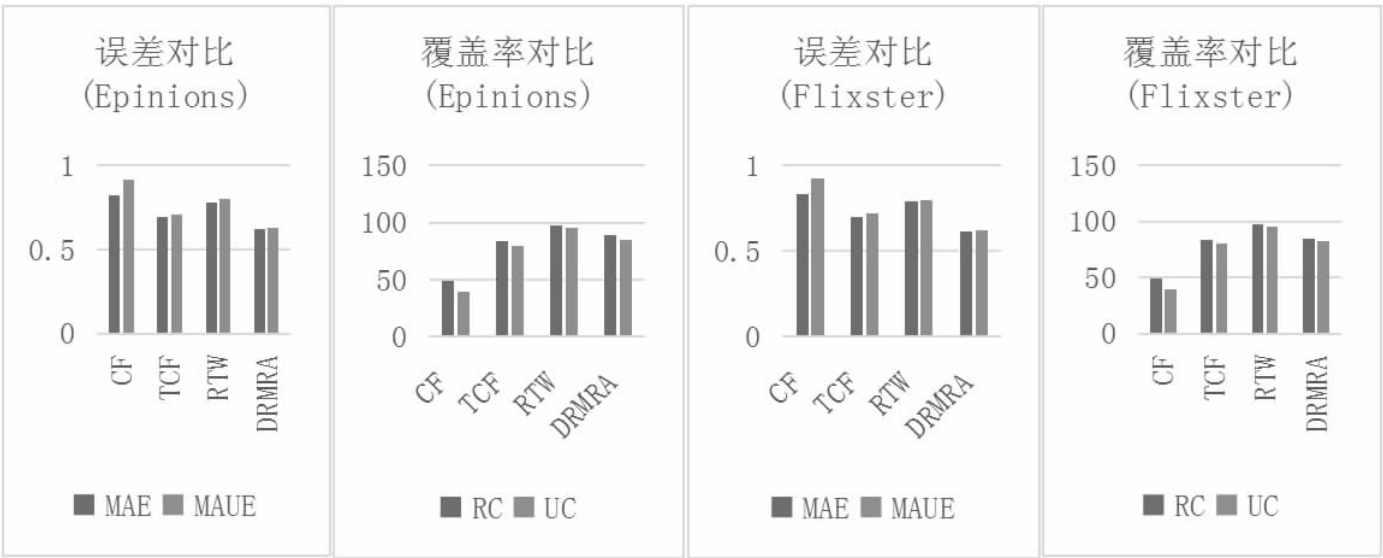


图 1 数据集 Epinions、Flixster 测试结果

可信网络重建过程中阈值 $\alpha=0.6$, $\beta=0.7$, 用户邻居选在范围 $K=500$ 。

4.2 实验结果及分析

Epinions、Flixster 数据集所有数据测试结果（图 1）所示，DRMRA 算法具有最低的 MAE 仅次于 RTW 算法的用户覆盖率。DRMRA 算法在 RC、UC 指标跟 TCF 算法很接近但是准确率最高。

参考文献

[1]王祖高,李卫华,徐艳玲.基于神经网络组合模型的软件可靠性评估研究[J].计算机工程,2016,32(33):214-215.
[2]张文博,红兵.融合极限学习机[J].电子与信息学报,2013,35(11):2728-2732
[3]KARYDI E,MARGARITIS K G. Multithreaded implementation of the Slop One algorithm for collaborative filtering [C] // 8th IFIP WG 12.5 International Conference on Artificial Intelligence Applications and Innovations. New York: Springer,2012:117-125.
[4]M.Ranjbar, P.Moradi, M.Azami, M.Jalili, An imputation - based matrix fac - torization method for improving accuracy of collaborative filtering systems, Eng. Appl,Artif.Intell.46 (2015) 58-66,http://dx.doi.org/10.1016/j.engappai.2015.08.010.
[5]C.Cechinel,M.Sicilia,S.S.Alonso,E.G. Barriocanal,Evaluating collaborative fil - tering recommendations inside large learning object repositories,Inf,Process.Manage.49(1)(2013)34-50.

[6]X.Luo,Y.xia,Q Zhu,Incremental collaborative filtering recommender based on regularized matrix factorization, Know. Based Syst. 27 (2012) 271-280,http://dx.doi.org/10.1016/j.knosys.2011.09.006

作者简介:张洪军(1983-),男,硕士,研究领域为人工智能、大数据分析。杨霞(1984-),女,硕士,研究领域为信息安全、大数据分析。