

DOI:10.13657/j.cnki.gxkxyxb.20180601.001

骆亮. 基于内容推荐算法和余弦相似度算法的领导决策辅助信息系统[J]. 广西科学院学报, 2018, 34(2): 143-150.

LUO L. A leader decision assistant information system based on content recommendation and cosine similarity algorithm[J]. Journal of Guangxi Academy of Sciences, 2018, 34(2): 143-150.

基于内容推荐算法和余弦相似度算法的领导决策辅助信息系统

A Leader Decision Assistant Information System based on Content Recommendation and Cosine Similarity Algorithm

骆 亮*

LUO Liang

(广西壮族自治区经济信息中心, 广西南宁 530022)

(Guangxi Economic Information Center, Nanning, Guangxi, 530022, China)

摘要:【目的】为研究符合党政机关辅助决策特点的决策辅助系统,帮助领导快速、准确收集查找决策辅助信息,提高党政机关领导决策效率和准确性。【方法】基于内容推荐和余弦相似度算法设计了决策信息搜索引擎,并结合决策信息库、用户决策特征库、决策信息采集子系统、决策信息推送子系统设计研发了领导决策辅助系统。【结果】系统在党政机关的决策辅助系统中进行了应用验证。系统可根据领导输入的关键字及领导的关注偏好进行决策辅助信息推送,推送结果准确率较高,符合辅助领导决策预期。【结论】基于内容推荐和余弦相似度算法的领导决策辅助信息系统根据用户偏好和特征进行决策辅助信息推荐,具有较好的准确性和针对性,开辟了党政机关决策辅助系统设计建设的新思路。

关键词:基于内容推荐 余弦相似度 决策辅助

中图分类号:TP391.7 **文献标识码:**A **文章编号:**1002-7378(2018)02-0143-08

Abstract:【Objective】In order to help the leaders to collect and find the information quickly and accurately, improving the efficiency and accuracy of the decision-making of leader of the government, the decision support system that meets the decision-making characteristics of the party and government agencies is studied. 【Methods】We design a leadership decision support system which is based on the content recommendation and cosine similarity algorithm. The system includes decision information database, user decision feature database, decision information collection subsystem and decision information push subsystem. 【Results】The system is applied in the government decision support system, which can push information according to the key words and the preference of the leader. The accuracy of the pushing results is high, which is consistent with the expectation of the leader. 【Conclusion】The system has better accuracy and pertinence according to the user preferences and characteristics, which has opened up a new idea for the design and construction of the decision-making assistant system of the party and government agencies.

Key words: content-based recommendation, cosine similarity, decision-making assistance

收稿日期:2018-01-10

作者简介:骆 亮(1984—),男,高级工程师,主要从事信息系统设计开发、数据库、计算机网络及系统集成等研究,E-mail:2015812@qq.com。

* 通信作者。

0 引言

【研究意义】党政机关领导政策制定、项目审批等决策研判往往比较复杂,无法通过传统的决策支持系统进行问题结果推导,更多地是通过查阅资讯信息、法律法规、政策文件等各类材料,结合领导所分管工作、关注重点和决策偏好进行辅助决策,研发适合党政机关的辅助决策信息系统显得尤为重要。

【前人研究进展】通过查阅文献,已有许多决策系统的研究和设计,其中,张伟^[1]、白玥^[2]、许海玲等^[3]的研究支持系统利用推理机以自动推理的方式进行问题求解或通过使用数量化方法将问题模型化,利用对数值模型的计算结果来进行决策支持。常玲慧^[4]通过建立决策知识模型辅助卫生应急决策。赵旭东等^[5]侧重对文本信息挖掘和分析辅助不良信息过滤。蒙杰等^[6]侧重通过对文本信息分析和比较辅助科研项目申报查重管理。朱青等^[7]通过对历史案件文本信息分析和辅助法院判决。这些研究都侧重于利用学习样本和历史数据进行问题求解而辅助决策,很少从用户决策偏好或特点进行辅助决策系统进行研究。**【本研究切入点】**通过研究党政机关领导决策特点,基于内容推荐和余弦相似度算法设计了由决策信息库、用户特征库、决策信息检索引擎、信息采集子系统、信息推送子系统组成的决策系统辅助领导决策,并在党政机关中进行了应用和验证。**【拟解决的关键问题】**在党政机关决策系统中应用基于内容推荐算法,构建信息特征矩阵和用户决策特征矩阵并设定权重,利用余弦相似度算法计算并推荐符合领导决策偏好的辅助决策信息。

1 基于内容的推荐算法

基于内容的推荐是指根据用户选择的对象,推荐其他类似属性的对象。这类算法源于一般的信息检索方法,不需要依据用户对对象的评价意见,常用于电子商务中商品推荐。对象使用通过特征提取方法得到的对象内容特征来表示,系统基于用户所选择对象的特征,学习用户的兴趣,从而考察用户资料与待预测对象相匹配的程度^[8-9]。原理上分为3个步骤:(1)为每个待推荐对象构建一个对象的属性资料;(2)为每个用户构建一个用户的偏好资料;(3)计算用户偏好资料与待推荐对象属性资料的相似度,相似度高意味着用户可能喜欢这个对象,相似度低往往意味着用户不喜欢这个对象。

2 余弦相似度算法

余弦相似度算法主要用于计算空间中两个向量的夹角的余弦值来度量它们之间的相似性。在空间中可把两向量(a 和 b)想象成空间中的两条线段(图1),都是从原点出发,指向不同的方向。两条向量间形成夹角,如果夹角为 0° ,意味着方向相同、线段重合;如果夹角为 90° ,意味着形成直角,方向完全不相似;如果夹角为 180° ,意味着方向正好相反。因此,我们可以通过夹角的大小,来判断向量的相似程度。夹角越小,就代表越相似^[10-11]。

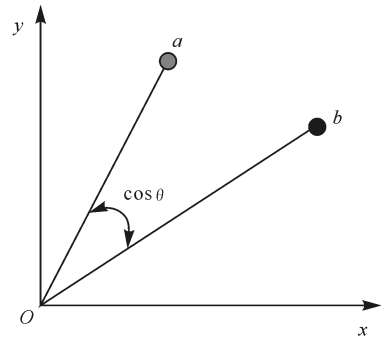


图1 余弦相似度算法向量示意图

Fig. 1 Cosine similarity algorithm vector diagram

在二维空间中,根据向量点积公式,有如下公式:

$$\cos\theta = \frac{a \cdot b}{\|a\| \|b\|}$$

若向量 a 、 b 的坐标分别为 (x_1, y_1) 、 (x_2, y_2) ,则有:

$$\cos\theta = \frac{x_1 x_2 + y_1 y_2}{\sqrt{x_1^2 + y_1^2} \times \sqrt{x_2^2 + y_2^2}}$$

进一步的,对于向量 $A = (A_1, A_2, \dots, A_n)$, $B = (B_1, B_2, \dots, B_n)$,推广到多维,有:

$$\cos\theta = \frac{\sum_1^n (A_i \times B_i)}{\sqrt{\sum_1^n A_i^2} \times \sqrt{\sum_1^n B_i^2}}$$

余弦相似度常结合 TF-IDF 算法在文本分类、计算文本相似度等通过计算两向量夹角来判断两对象相似度的场景中应用。在本研究中,基于内容推荐算法构建了待推荐对象特征向量和用户偏好向量,可通过余弦相似度算法计算两向量相似性,并根据相似程度(余弦值大小)由高到低进行信息推荐。

3 领导决策辅助系统中的应用

本研究以广西党政机关的领导辅助信息系统为例。领导辅助信息系统包括2库1引擎2子系统,

即决策信息库(信息特征矩阵)、用户特征库(用户特征矩阵)、决策信息检索引擎、决策信息采集子系统和决策信息推送子系统。

系统根据内容推荐算法构建信息特征矩阵和用户特征矩阵,利用检索引擎进行两矩阵相似度计算并推荐决策信息。系统运行过程如下(数据流图见图 2):

(1)辅助决策工作人员根据领导决策偏好,利用决策信息采集子系统进行决策信息采集(信息来源一般为工作中撰写或从国家、地方参考资料中引用)。工作人员根据信息内容手工标注关键字并将信息录入决策信息库中。

(2)领导在决策信息检索引擎中输入关键字进行辅助信息检索。

(3)检索引擎记录领导操作记录和检索信息的特征,调整领导的用户决策特征矩阵,并将用户特征信息推送至信息采集子系统帮助工作人员了解领导决策偏好。

(4)信息检索引擎利用余弦相似度算法计算每条信息特征与用户决策特征的相似度,并在信息推送子系统中按照相似度由高到低向领导推送辅助信息。

推荐偏好调整。在本研究系统案例中,关键字特征常数 1,权值设定为 1,其他特征权重根据决策偏好重要性进行设定,其中信息涉及所属地最高,主题次之,原创性最低。为此,系统设定各特征权重分别为 1,1,2,3。

关键字特征值 s_1 :用于初步筛选和索引,其特征 s_1 值为常数 1。

原创性特征值 s_2 :信息为原创时特征值为 1,非原创时特征值为 0。

信息所属主题特征值 s_3 :本研究中应用的党政机关系统中包含 15 个主题(权威信息、委内简介、规章制度、重要文件、产业政策、行政审批、业务管理、区情数据、经济运行、重大项目进展、发展规划、发改要报、桂经专报、能源发展、PPP 专栏),分别赋值 1 至 15。主题特征可以根据业务需要随时增减。为避免特征分项过多,特征分项赋值过大时影响决策特征权重,需对特征分项值进行 min-max 标准化处理。min-max 标准化处理后,对应特征分项值分别为 0,0.07,0.14,0.21,0.29,0.36,0.43,0.5,0.57,0.64,0.71,0.79,0.86,0.93,1。

信息涉及所属地 s_4 :包含 14 个地市(南宁、柳州、桂林、梧州、北海、钦州、防城港、贵港、河池、玉林、贺州、来宾、百色、崇左)及自治区层面、国家层面、其他省市等 17 个分项,分别赋值 1 至 17。min-max 标准化处理后,对应特征分项值分别为 0,0.08,0.15,0.23,0.31,0.38,0.46,0.54,0.62,0.69,0.77,0.85,0.92,1。

表 1 决策信息库表结构

Table 1 Structure of decision information database table

字段名 Field name	字段属性 Field properties
编号 ID	Int
标题 Title	Nvarchar(255)
内容 Content	Ntext
关键字 Keyword	Nvarchar(255)
原创性 Original	Nvarchar(255)
主题 Theme	Nvarchar(255)
地区 Area	Nvarchar(255)
原创性特征值 s_2 Original value s_2	Float
主题特征值 s_3 Theme value s_3	Float
地区特征值 s_4 Area value s_4	Float

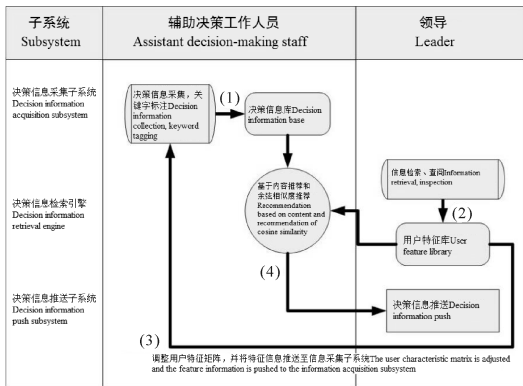


图 2 系统数据流图

Fig. 2 System data flow diagram

3.1 决策信息库(特征信息矩阵)

决策信息库包含特征信息矩阵,收集各类辅助决策信息,包括公文、项目管理信息、产业政策、经济分析报告、法律法规、业务管理、重大项目进展、专题报刊、宏观决策信息、参考资料等内容,各条信息通过人工方式标注关键字,并选取了“信息关键字、信息原创性、信息所属主题、信息涉及所属地”等作为信息特征进行量化、索引和存储,分别表示为 s_1 、 s_2 、 s_3 、 s_4 ,信息特征矩阵表示为 $I = [s_1, s_2, s_3, s_4]$,表结构见表 1。为更准确地反映各特征在检索过程中重要性的不同,为各特征值设定权值 q ,并可根据决策

如原创信息“2017 年南宁市国民经济运行分析报告”包含关键字“南宁”“经济运行”,信息涉及所属地为“南宁”,所属主题为“经济运行”,则其信息特征

矩阵可表示为 $I = [1, 1 \times 1, 0.57 \times 2, 0 \times 3] = [1, 1, 1.14, 0]$ 。

3.2 用户决策特征库(用户特征矩阵)

用户决策特征库(用户特征矩阵表)存储各用户检索偏好特征信息,表结构见表2。系统中用户特征矩阵和信息特征矩阵具有同样维度,矩阵表示 $U = [s'_1, s'_2, s'_3, s'_4]$,其中 s'_1 为用户关键字特征值, s'_2 为用户原创性特征值, s'_3 为用户主题特征值, s'_4 为用户地区特征值。

用户关键字特征值 s'_1 按照关键字字数匹配度和关键字权重的乘积计算,如关键字“南宁”与“南宁”“经济运行”匹配了1组关键字,相似程度为匹配关键字数/信息关键字总数 = $1/2 = 0.5$,此时 s'_1 为0.5。

其他特征值(s'_2, s'_3, s'_4)以对应特征中用户日常检索次数最多的分项值与对应特征权值乘积计算。公式如下:

$$\max P(s_{ji}) = \max \frac{n_{ji}}{m},$$

$$s'_j = s_{ji} \times q_j.$$

其中 j 为第 j 个特征, i 为第 j 个特征中的第 i 个分项值, m 为用户的检索次数, n_{ji} 表示第 j 个特征第 i 个分项的检索次数, q_j 为第 j 个特征的权值, s_{ji} 表示第 j 个特征值的第 i 个分项值, $P(s_{ji})$ 表示第 j 个特征的第 i 个分项值检索概率。如用户在以往检索中查阅原创性信息3次,非原创性信息1次, $s'_2 = 1$; 查阅“经济运行”主题4次, $s'_3 = 0.57$; 查阅“南宁”地区信息3次,“北海”地区信息1次, $s'_4 = 0$; 则矩阵表示为 $U = [1, 1 \times 1, 0.57 \times 2, 0 \times 3] = [1, 1, 1.14, 0]$ 。

表2 用户特征表结构

Table 2 Structure of user feature table

字段名 Field name	字段属性 Field properties
用户编号 User ID	Nvarchar(255)
用户原创性特征值 s'_2 User original value s'_2	Float
用户主题特征 s'_3 User theme value s'_3	Float
用户地区特征 s'_4 User area value s'_4	Float

3.3 决策信息检索引擎

提供领导提问接口,基于内容推荐算法和余弦相似度函数,按信息特征和决策偏好特征优先向领导推送符合其关注重点和决策习惯的辅助决策信

息,记录、统计和分析领导关注的重点和决策习惯。

3.3.1 基于余弦相似度的决策信息推荐过程

用户以关键字 k 进行决策信息检索时,先在数据库中匹配所有包含关键字 k 的信息,获得信息集合 K 。利用余弦相似度公式来计算用户 U 和集合 K 中各信息 I 之间的距离。遍历计算待推荐信息和用户特征相似性,并按照相似性由高到低进行排序向用户推荐。余弦相似度的值越大说明 U 越有可能喜欢 I ,该信息越优先推荐。

余弦相似度的具体计算方法如下:

$$\cos(U, I) = \frac{\sum_i^n (U_a \times I_a)}{\sqrt{\sum_1^n U_a^2} \times \sqrt{\sum_1^n I_a^2}},$$

U_a 表示用户第 a 项特征值, I_a 表示信息第 a 项特征值。

推荐过程算法如下:

```
Class Vector { float S1,S2,S3,S4; } //定义特征向量,
包括 S1,S2,S3,S4 向量
Class Info { //定义信息
    Vector vector; //信息特征向量
    float cosValue; //特征余弦相似度(余弦值)
}
List<Info> K = Search(k); //在数据库中匹配所有包含关键字 k 的信息,获得信息集合 K
Foreach(info in K)
    info.cosValue = cos(user.vector, info.vector); //计算集合 K 中 user 特征向量和信息 info 特征向量的余弦相似度
Sort(K, Info.cosValue); //按照余弦相似度对集合 K 排序
```

3.3.2 信息检索操作记录

用户在进行关键字检索或点击主题模块查阅决策信息时,系统自动将领导的操作记录在检索操作日志表中,包括关键字、检索时间、查阅信息特征等信息,表结构见表3。领导的操作信息在系统后台自动进行重新计算,调整用户特征矩阵,并将领导的操作特征信息推送至信息采集子系统。辅助决策工作人员可及时获知领导用户操作习惯和所关心的决策信息,并开展信息采集补充工作。

3.4 决策信息采集子系统

实现决策信息录入、编辑、更新、删除等管理功能(图3)。提供领导关注点分析控制台,自动向信息采集员推送各领导近期所关心的问题 and 操作统计信息。网络爬虫以领导的问题为导向,自动从各信息库和网络上爬取相关信息,信息采集员对收集的信息二次加工,并将信息标注关键字、划分主题和归

类补充进入决策信息库^[12]。

表 3 检索操作日志表结构

Table 3 Structure of search log table

字段名 Field name	字段属性 Field properties
用户编号 User ID	Nvarchar(255)
关键字 Keyword	Nvarchar(255)
检索时间 Search time	Date time
原创性 Original	Float
主题 Theme	Float
地区 Area	Float

3.5 决策信息推送子系统

将决策信息库内的决策信息按主题推送至领导操作查阅界面,同时提供决策信息搜索界面,供领导模糊查询和检索(图 4)。领导用户输入关键字进行检索时,推送子系统利用搜索引擎对决策信息库进行检索查询,并向领导推送决策辅助信息。领导用户也可以根据已分类的主题点击查看已整理好的决策信息(图 5)。

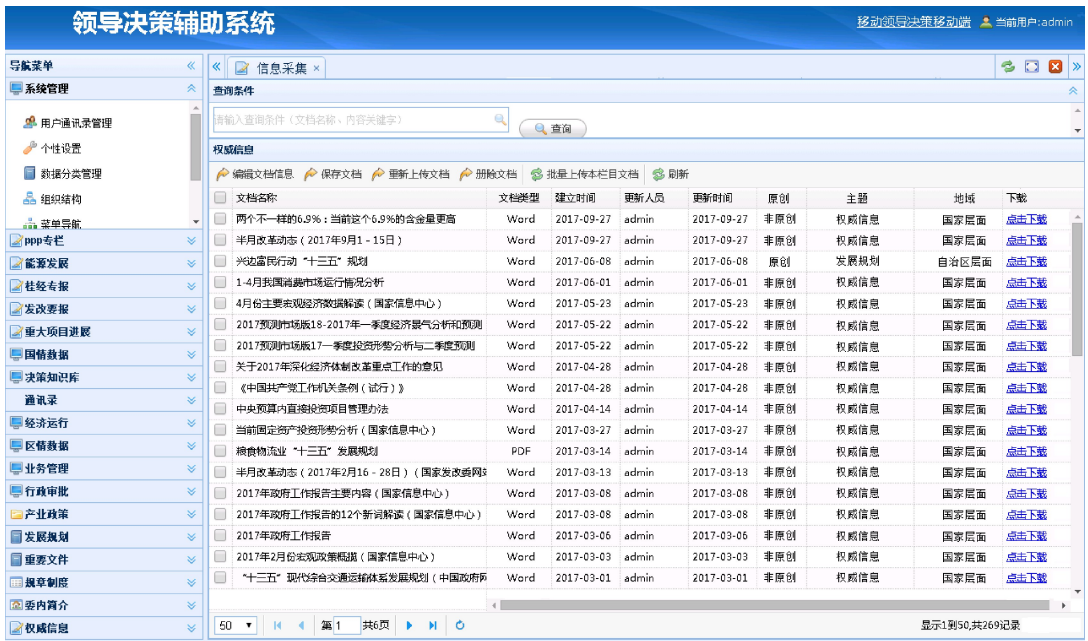


图 3 决策信息采集子系统界面

Fig. 3 Interface of decision information collection subsystem



图 4 信息搜索、推送界面

Fig. 4 Information search and push interface



图5 信息分类查询界面

Fig. 5 Interface of information classification query

表4 辅助决策信息表(部分)

Table 4 Assisting decision information table (part)

编号 ID	关键字 Keyword	原创性 Original	主题 Theme	地区 Area
01	西江经济带 Xijiang economic belt	是 Yes	发改要报 Report	自治区层面 Autonomous
02	经济数据 Economic data	是 Yes	发改要报 Report	自治区层面 Autonomous
03	重点生态功能区政策改革发展 Key ecological functional zones, policy, reform, development	是 Yes	发改要报 Report	自治区层面 Autonomous
04	经济分析 Economic analysis	是 Yes	发改要报 Report	自治区层面 Autonomous
05	动态 Trends	是 Yes	发改要报 Report	自治区层面 Autonomous
06	经济数据 Economic data	是 Yes	发改要报 Report	自治区层面 Autonomous
07	进展情况 Progress	是 Yes	发改要报 Report	自治区层面 Autonomous
08	主体功能区 Main function area	是 Yes	发改要报 Report	自治区层面 Autonomous
09	工作方案规划方案 Programme of work, plan, scheme	是 Yes	发改要报 Report	自治区层面 Autonomous
...

4.2 领导决策偏好特征构建

系统自动将领导每次检索或查阅信息的特征信息记录在检索操作日志表中,根据信息特征自动计算和调整用户特征矩阵。假设领导进行了15次检索和查阅,其中检索和查阅的信息原创性特征中原创9次、非原创性6次,主题特征中发改要报8次、经济运行4次,地区特征中自治区层面10次、南宁3次、北海2次,本研究所述方法,计算出决策偏好为“原创、发改要报、自治区层面”。系统中的浏览和检索次数记录为:用户编号1,用原创性特征值 $s'_2 = 1$,用户主题特征值 $s'_3 = 0.79 \times 2 = 1.58$,用户地区特

4 实例验证

4.1 辅助决策信息数据来源

本研究进行实例研究的数据来自于广西某党政机关辅助决策系统,除去敏感文件,信息数据共754条。这些信息由机关工作人员在工作中撰写或引用国家、地市有关工作资料整理而成,部分数据见表4。

征值 $= 0.875 \times 3 = 2.625$ 。

4.3 余弦相似度计算结果和排序推荐

假设领导以“十三五”为关键字进行检索。系统结合信息特征值和余弦相似度计算,在754条辅助信息中筛选出了30条包含“十三五”关键字的信息。系统根据关键字、信息原创性、主题、地区等属性自动标注特征值 s_1, s_2, s_3, s_4 ,并通过余弦相似度算法计算用户决策特征矩阵和各信息特征矩阵的相似度,按从大到小排列。推荐结果前6条信息结果见表5。

表 5 推荐信息表(部分)

Table 5 Recommended information table(part)

编号 ID	关键字 Keyword	原创性 Original	主题 Theme	地区 Area	余弦相似度 Cosine similarity
64	十三五 铁路 13th Five-Year, railway	是 Yes	发改要报 Report	自治区层面 Autonomous	1.000 0
65	十三五 规划 总结 13th Five-Year, railway, plan, summary	是 Yes	发改要报 Report	自治区层面 Autonomous	1.000 0
93	十三五 规划 国家 13th Five-Year, railway, plan, state	是 Yes	发改要报 Report	国家层面 Autonomous	0.999 5
111	十三五 规划 物流 13th Five-Year, railway, plan, logistics	否 No	发展规划 Development plan	自治区层面 Autonomous	0.949 6
112	十三五 规划 交通运输 13th Five-Year, railway, plan, transportation	是 Yes	发展规划 Development plan	自治区层面 Autonomous	0.999 0
113	十三五 规划 新兴产业 13th Five-Year, railway, plan, new industry	否 No	发展规划 Development plan	自治区层面 Autonomous	0.949 6
...

4.4 结果分析

通过对结果信息观察, ID 为 64, 65 的余弦相似度为 1, 原创性: 原创, 主题: 发改要报, 地区: 自治区层面, 对应 s_1, s_2, s_3, s_4 , 特征值分别为 0.33, 1, 0.79, 0.875 0, 与假设的领导用户决策偏好完全一致, 优先推荐。

ID 为 93, 112 的信息, s_1, s_2, s_3, s_4 特征值及余弦相似度结果分别为 0.33, 1, 0.79, 0.937 5, 0.999 5 和 0.33, 1, 0.71, 0.875 0, 0.999 0, 与 ID64, 65 的信息在主题和地区上有差异, 且 ID112 相较 ID93 的信息偏差更大。按照余弦相似度计算, ID112 排在 ID93 之后, 符合优先推荐与领导用户决策偏好相近的原则。

5 结论

与前人研究^[1-7]的决策辅助过程关注点不同, 本研究基于内容推荐和余弦推荐算法, 注重根据领导决策偏好和特点进行辅助信息检索和推荐, 将算法结合党政机关决策辅助信息检索应用场景, 把辅助信息进行特征量化, 以用户检索的关键字及其他信息为用户特征, 基于余弦相似度算法对辅助信息进行检索、排序和推荐, 创新地研发了符合党政机关决策特点的决策辅助系统。经过实例检验, 验证了基于用户决策偏好特点的计算辅助决策方法。辅助决策结果具有较高的准确性。同时, 系统帮助辅助决策人和信息采集员及时了解领导需求, 迅速收集整理决策辅助信息供领导查阅和参考, 增强了用户

体验, 开辟了党政机关决策辅助系统设计建设的新思路。

由于本研究研究实践应用深度有限, 系统还在许多方面可进一步完善。如在信息关键词方面可应用 TF-IDF 算法进行分词和特征词提取并辅助关键词标注^[13-14]; 在用户特征矩阵建立过程中, 用户存在第一偏好和第二偏好, 甚至是并列偏好的可能性, 本研究只考虑了第一偏好的辅助决策。同时由于提供研究数据的粗糙性, 余弦相似度算法会因信息特征值设定的不同而导致推荐结果的不同, 以上的不足可在今后的研发过程中可进一步改进。

参考文献:

[1] 张伟. 智能决策支持系统(IDSS)研究综述[J]. 现代商贸工业, 2009, 21(14): 252-253.
ZHANG W. An overview of the research on intelligent decision support system(IDSS)[J]. Modern Business Trade Industry, 2009, 21(14): 252-253.

[2] 白玥. 云计算条件下 IDSS 的发展现状及发展趋势[J]. 大科技, 2015, 18: 208.
BAI Y. Development status and development trend of IDSS under cloud computing[J]. Super Science, 2015, 18: 208.

[3] 许海玲, 吴潇, 李晓东, 等. 互联网推荐系统比较研究[J]. 软件学报, 2009, 20(2): 350-362.
XU H L, WU X, LI X D, et al. Comparison study of internet recommendation system[J]. Journal of Software, 2009, 20(2): 350-362.

[4] 常玲慧. 知识管理在突发公共卫生事件应急决策系统中的应用研究——以山西省为例[D]. 太原: 太原理工大学, 2013.

- CHANG L H. Application research on knowledge management in public health emergency decision support system taking Shanxi Province as an example[D]. Taiyuan: Taiyuan University of Technology, 2013.
- [5] 赵旭东, 亚森·艾则孜. 基于互信息和余弦相似度的维吾尔文不良文档信息过滤方案[J]. 电子设计工程, 2016, 24(16): 109-112.
- ZHAO X D, YASEN A. A uyghur bad text information filtering scheme based on mutual information and cosine similarity [J]. Electronic Design Engineering, 2016, 24(16): 109-112.
- [6] 蒙杰, 杨生举, 施韶亭. 基于文本挖掘的科研项目辅助决策系统研究与实现[J]. 计算机应用与软件, 2016, 33(9): 24-26, 55.
- MENG J, YANG S J, SHI S T. Study and implementation of text mining-based assistant decision support system for scientific research project management [J]. Computer Applications and Software, 2016, 33(9): 24-26, 55.
- [7] 朱青, 卫柯臻, 丁兰琳, 等. 基于文本挖掘和自动分类的法院裁判决策支持系统设计[J]. 中国管理科学, 2018, 26(1): 170-178.
- ZHU Q, WEI K Z, DING L L, et al. Design of court decision support system based on text mining and automatic classification[J]. Chinese Journal of Management Science, 2018, 26(1): 170-178.
- [8] 刘明昌. 基于内容的推荐技术研究[J]. 现代营销, 2016, 6: 243.
- LIU M C. Research on content based recommendation technology[J]. Marketing Management Review, 2016, 6: 243.
- [9] 杨武, 唐瑞, 卢玲. 基于内容的推荐与协同过滤融合的新闻推荐方法[J]. 计算机应用, 2016, 36(2): 414-418.
- YANG W, TANG R, LU L. News recommendation method by fusion of content-based recommendation and collaborative filtering[J]. Journal of Computer Applications, 2016, 36(2): 414-418.
- [10] 刘冰, 李文书. 基于余弦相似度的指纹匹配算法的室内定位方法[J]. 科技通报, 2017, 33(3): 198-202.
- LIU B, LI W S. Indoor positioning method based on cosine similarity of fingerprint matching algorithm [J]. Bulletin of Science and Technology, 2017, 33(3): 198-202.
- [11] 李梦洁, 邵曦. 基于文本属性的微博用户相似度研究[J]. 计算机技术与发展, 2018, 5. <http://kns.cnki.net/kcms/detail/61.1450.TP.20180207.1917.082.html>
- LI M J, SHAO X. Research of micro-blog user similarity based on text similarity[J]. Computer Technology and Development, 2018, 05. <http://kns.cnki.net/kcms/detail/61.1450.TP.20180207.1917.082.html>
- [12] 乔峰. 基于模板化网络爬虫技术的 Web 网页信息抽取[D]. 成都: 电子科技大学, 2012.
- QIAO F. Web page information extraction based on formwork web crawler technology[D]. Chengdu: University of Electronic Science and Technology of China, 2012.
- [13] 武永亮, 赵书良, 李长镜, 等. 基于 TF-IDF 和余弦相似度的文本分类方法[J]. 中文信息学报, 2017, 31(5): 138-145.
- WU Y L, ZHAO S L, LI C J, et al. Text classification method based on TF-IDF and cosine similarity[J]. Journal of Chinese Information Processing, 2017, 31(5): 138-145.
- [14] 董洋溢, 李伟华, 于会. 基于混合余弦相似度的中文文本层次关系挖掘[J]. 计算机应用研究, 2017, 34(5): 1406-1409.
- DONG Y Y, LI W H, YU H. Hierarchical relation mining of Chinese text based on mixed cosine similarity[J]. Application Research of Computers, 2017, 34(5): 1406-1409.

(责任编辑: 米慧芝 符支宏)

(上接第 142 页 Continue from page 142)

- [9] CAMBAZARD H, HADZIC T, O' SULLIVAN B. Knowledge compilation for itemset mining[C]//Conference on ECAI 2010: European Conference on Artificial Intelligence. [S. l.]: IOS Press, 2010: 1109-1110.
- [10] TSEITIN G. On the complexity of derivations in the propositional calculus [J]. Zap Nauchn Sem Lomi, 1968: 234-259.
- [11] DAVIS M, LOGEMANN G, LOVELAND D W. A machine program for theorem-proving[J]. Communications of the ACM, 1962, 5(7): 394-397.
- [12] MARQUES-SILVA J P, SAKALLAH K A. GRASP—A new search algorithm for satisfiability[C]//IEEE/ACM International Conference on Computer-Aided Design, 1996. Iccad-96. IEEE: [s. n.], 1997: 220-227.
- [13] ZHANG L, MADIGAN C F, MOSKEWICZ M H, et al. Efficient conflict driven learning in a Boolean satisfiability solver[C]//IEEE/ACM International Conference on Computer-Aided Design. IEEE Press: [s. n.], 2001: 279-285.
- [14] BRYANT R E. Graph-based algorithms for boolean function manipulation[J]. IEEE Trans Comput, 1986, C- 35(8): 677-691.
- [15] GU T L, XU Z B. Ordered binary decision diagram and its application[M]. Beijing: Science Press, 2009.
- [16] BRYANT R E. Symbolic Boolean manipulation with ordered binary decision diagrams [J]. ACM Computing Surveys, 1992, 24(3): 293-318.

(责任编辑: 米慧芝)