

Research on Plant-Protecting Operation Recommendation Algorithm Based on Log Data Processing

Zhichao Lv¹, Lihua Zheng^{2*}, Ze Yang², Wei Yang², Minzan Li²

¹Network Technology Center, China Agricultural University, Beijing

²Key Laboratory of Modern Precision Agriculture System Integration Research, Ministry of Education, China Agricultural University, Beijing

Email: lvzhichao@cau.edu.cn, *zhenglh@cau.edu.cn

Received: Nov. 7th, 2018; accepted: Nov. 21st, 2018; published: Nov. 28th, 2018

Abstract

The scientific and reasonable recommendation algorithm for plant-protecting operation can effectively assure the safety of UAV (Unmanned Aerial Vehicles), as well improve the efficiency of plant protecting work and its management level. This paper proposed a method of operation recommendation for plant-protecting UAV users based on log data processing, built a log data real-time analysis and calculation system using Storm framework technology, and developed an operation recommendation algorithm combining user feature and collaborative filtering and another plant-protecting operation recommendation algorithm based on content filtering. In the former algorithm, the most similar neighbor users of the target user were investigated and calculated, and then the appropriate operation recommendation list was obtained according to the calculation result of their weighted scores. The later recommendation algorithm was based on content filtering which recommended the potential plant-protecting operations for the corresponding user by analyzing his search and browser history and records. System run and analysis showed that the log data processing based algorithm could give users the best plant-protecting operation recommendations according to the setup rules, which could potentially improve the system's intellectuality, convenience and usability.

Keywords

Unmanned Aerial Vehicle, Log Data Processing, Storm, Collaborative Filtering, Recommendation Algorithm

基于日志处理的植保作业推荐算法研究

*通讯作者。

吕智超¹, 郑立华^{2*}, 杨 泽², 杨 玮², 李民赞²

¹中国农业大学网络技术中心, 北京

²中国农业大学现代精细农业系统集成研究教育部重点实验室, 北京

Email: lvzhichao@cau.edu.cn, zhenglh@cau.edu.cn

收稿日期: 2018年11月7日; 录用日期: 2018年11月21日; 发布日期: 2018年11月28日

摘 要

本文提出了一种基于日志处理的植保作业推荐方法, 利用Storm框架技术搭建了一个日志实时分析计算系统, 设计并开发了混合用户特征和协同过滤的植保作业推荐算法与基于内容过滤的植保作业推荐算法。混合用户特征和协同过滤的植保作业推荐算法首先计算与目标用户最相似的邻居用户, 按其作业加权评分结果获取适宜的植保作业推荐给无人机用户; 基于内容过滤的植保作业推荐算法基于用户的搜索与查看记录, 为用户推荐可能感兴趣的植保作业。系统测试与分析表明, 基于日志处理的植保作业推荐算法能够按设定规则为用户推荐植保作业, 提高了系统的智能性、便捷性和可用性。

关键词

植保无人机, 日志处理, Storm, 协同过滤, 推荐算法

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

保证粮食安全是我国的基本国策, 农业植保是农业生产中最重要的一环之一。而当前国内农业植保装备比较落后, 农作物保护形式仍以人工、半机械化作业为主。据相关统计, 我国目前使用的植保设备仍以手动和小型电动喷药机为主, 其中手动施药药械、背负式机动药械分别占国内植保机械的93%和5.5%以上, 拖拉机悬挂式植保机械占不到1% [1]。传统农作物保护形式农药用量大, 浪费严重, 作业成本高, 工作效率低, 无法保障植保质量与作物质量产量, 同时也带来了严重的环境污染等问题[2], 这与现代化农业发展的趋势相悖。更令人担心的是, 传统植保作业方式易造成人身药害, 据统计, 每年有数十万人遭遇农药中毒事件, 其致死率高达8.9% [3]。为缓解上述问题, 农业植保正在向更为现代化、专业化的方向发展。

在植保航空领域, 美国、日本、澳大利亚等国处于领先地位。其中日本的情况与我国相似: 人均耕地面积小, 不适合有人驾驶的大型农用飞机进行作业[4] [5]。美国近些年在农业航空方面, 应用了一系列的新技术[6] [7]。随着植保无人机应用走向市场, 无人机植保作业量将会迅猛增加, 如何快速匹配植保无人机用户租赁需求与农业生产用户植保作业需求至关重要。尽管在植保无人机监管系统中已有了多种便捷的植保作业检索方式, 但是这一过程仍然需要用户主动去搜索, 尚不够智能, 尚无无人机植保作业匹配方法或系统发表。为了能够使无人机与植保作业的匹配更为便捷、有效, 提升系统效率与用户体验, 本文考虑利用用户行为日志设计一个科学、合理的无人机植保作业推荐算法, 把可能感兴趣的植保作业推荐给用户, 供用户选择。

日志分析可通过批量离线处理或实时计算来进行。离线批处理涉及 Hadoop 大数据处理技术,如薛文娟[8]设计了基于层次聚类的日志分析技术;杨锋英、刘会超[9]基于 Hadoop 的在线网络日志分析系统。而实时日志分析,需要用到常见的实时计算框架,如 Spark、Storm、Dremel 等,如刘季函[10]基于 Spark 的海量日志数据分析平台;屈国庆[11]基于 Storm 设计并实现了实时日志分析系统;孙思源[12]基于 MongoDB 分布式数据库的高效 Web 日志分析方案;陈飞、艾中良[13]基于 Flume、ElasticSearch、Kibana 等技术研究并实现了一种分布式的日志采集分析系统。由于植保作业的喷药期有比较严格的时间限制,为了能够及时计算出用户可能感兴趣的植保作业,本系统选择使用实时计算。而在实时计算框架中,常见的计算框架有 Storm 和 SparkStreaming。Storm 平台以其纯实时的流式计算、极高的运算速度、清晰的拓扑结构和编程接口而成为了工业界最流行的实时计算框架,与 Spark 相比它拥有更低的计算延迟、更完善的事务机制、更好的健壮性,但吞吐量相较逊色。考虑到植保作业推荐系统对于高实时性、高可用性的需求,本文选择 Storm。

前人针对推荐算法的研究成果为本文设计提供了可贵的借鉴。王嫣然[14]等设计了基于内容过滤的科技文献推荐算法;庄景明[15]等将内容过滤算法应用在农业信息推荐系统上;袁先虎[16]设计了基于混合用户模型的协同过滤推荐算法;胡勋[17]等混合了内容推荐算法与基于用户的协同过滤算法;刘树栋[18]等基于位置设计了社会化网络推荐系统。推荐系统已成为最重要的个性化技术发展方向[19][20]。在植保无人机用户匹配植保作业的过程中,需要将农业作业的地理空间属性、作业时间属性以及用户属性均考虑在内,才能获得较为理想的作业推荐算法。通过研究与实践,本文基于 Storm 框架设计、采用内容过滤算法、协同过滤算法、以及多种方法混合的推荐算法最终实现了一个智能化的无人机作业快速推荐算法。

2. 材料与方法

2.1. Storm 技术

Storm 是一个分布式的流式数据处理系统[21]。Storm 内部组件处理流程图如图 1 所示。它把任务分配给不同的任务处理流,流中的每个组件负责处理一项特定的任务。Storm 中的数据接收单元为 Spout,每个 Spout 都会把数据传递给一个或多个数据处理单元 Bolt,经过处理后 Bolt 把数据保存到某存储器中,或者传递给其他的 Bolt 继续进行处理,数据就如流水一样经过这些组件,完成整体处理流程。

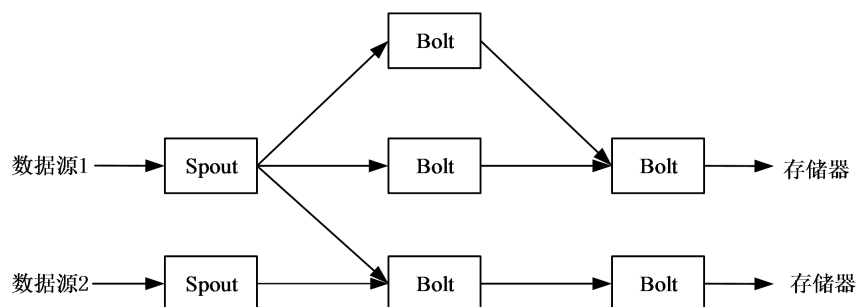


Figure 1. The processing flow of Storm's internal components
图 1. Storm 内部组件处理流程图

2.2. 日志获取

在本文开发的植保无人机飞行监管系统中,为了实现植保无人机作业推荐,设计了相应字段来记录无人机用户对植保作业的操作内容作为日志信息,如表 1。

Table 1. The log information
表 1. 日志信息

日志分类	字段	用途
查看日志	该用户的 ID、该用户所处位置、该植保作业的 ID、该作业的作业类型和植保类型、查看时间	用户检索到某植保作业，并查看了其详情
申请日志	该用户的 ID、该用户所处位置、该植保作业的 ID、该作业的作业类型和植保类型、申请时间	用户检索到某植保作业，并申请了这项作业
区划搜索日志	用户 ID、行政区划代码、搜索时间	利用行政区划搜索植保作业的日志
R 树搜索日志	用户 ID、矩形区域坐标、搜索时间	利用 R 树的区域搜索植保作业的日志

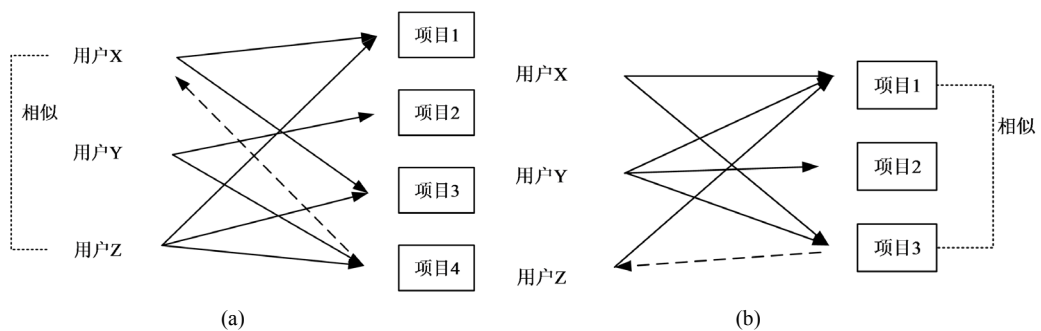
利用以上信息，一方面可以计算出用户对哪一类的植保作业比较感兴趣；另一方面可以利用基于用户的协同过滤算法计算出用户之间的相似度，并以此作为植保作业推荐的依据。

2.3. 作业推荐算法

2.3.1. 协同过滤算法

协同过滤算法包括基于记忆的协同过滤和基于模型的协同过滤[22]。对于农业植保作业用户来说，根据其无人机的不同类型、植保作业类型以及执行项目所在区域等信息可较为精确地定位植保作业用户的适宜项目，因此本研究采用基于记忆的协同过滤算法实现用户作业推荐。

基于记忆的协同过滤算法具体分为基于用户的协同过滤和基于项目的协同过滤。基于用户的协同过滤根据用户之间的相似性来互相推荐可能感兴趣的项目，其原理如图 2(a)所示。假设用户 X 喜好项目 1 和 3，用户 Y 喜好项目 2 和 4，用户 Z 喜好项目 1、3 和 4。从这些历史喜好中，可以推论用户 X 和用户 Z 具有相似偏好。由于用户 Z 还喜好项目 4，那么用户 X 也可能喜好项目 4，因此可以向用户 X 推荐项目 4。



注：虚线段表示相似的用户或项目，实线箭头表示用户喜好，虚线箭头表示向用户推荐

Figure 2. The principle of collaborative filtering based on memory
图 2. 基于记忆的协同过滤原理

图 2(b)中展示了基于项目的协同过滤。用户 X 喜好项目 1 和 3，用户 Y 喜好项目 1、2 和 3，用户 Z 喜好项目 1。从这些历史记录可发现，喜好项目 1 的用户大多同时也喜好项目 3，因此可以认为项目 1 和 3 类似，而用户 Z 喜好项目 1，因此可以把项目 3 推荐给他。

2.3.2. 内容过滤算法和人口统计学算法

内容过滤算法根据用户的喜好记录分析出喜爱哪一类的项目，最后便可以把这一类项目推荐给用户，其原理如图 3 所示。假设用户 X 喜好项目 1，通过相似度计算，发现项目 1 和 3 具有较高相似，由此推论，用户 X 也喜好项目 3。于是将项目 3 推荐给用户 X。

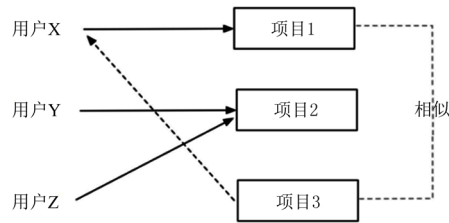


Figure 3. The Content filtering recommendation
图 3. 基于内容过滤推荐原理

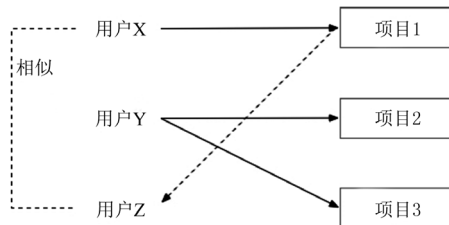


Figure 4. The Demographics recommendation
图 4. 基于人口统计学推荐原理

图 4 中展示了基于人口统计学推荐算法，根据特征计算用户间的相似度，用户 X 喜好项目 1，系统通过计算发现用户 X 和 Z 具有较高相似度，就会把用户 X 喜好的项目推荐给 Z。

3. 设计与实现

3.1. 基于 Storm 框架的日志实时计算系统搭建

为了搭建分布式日志实时分析系统，在单台服务器上部署了 4 台虚拟机，分别命名为 node1, node2, node3, node4，这些虚拟机承担着 Web 服务器和日志分析的任务。其中，Web 服务器部署在 node1 上，作为网站的主节点；Flume 为单点部署，部署在 node1，方便日志的收集；Kafka 为分布式部署，共 3 台虚拟机，分别部署在 node2, node3, node4 上；Storm 为分布式部署，共 4 台虚拟机，其中 node1 为 Master 节点，即 Nimbus；而 node2、node3、node4 则为 Worker 节点，即 Supervisor；Zookeeper 分布式集群管理组件共 3 台虚拟机，分别部署在 node2、node3、node4 上。系统拓扑结构如图 5 所示。

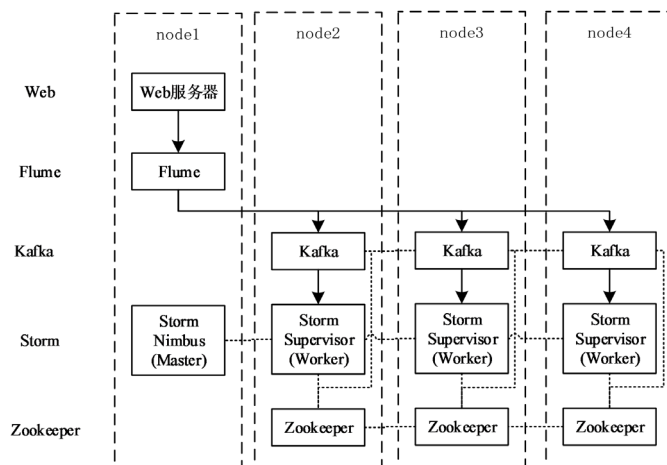


Figure 5. Topological diagram of plant protection operation recommendation system
图 5. 基于日志分析的植保作业推荐系统拓扑图

3.2. 混合用户特征与协同过滤的植保作业推荐算法

考虑到植保作业和无人机都有着较强的地理属性，而无人机仅能在有限的地理范围进行服务，这意味着一般来说无人机只为本地的植保作业提供服务。因此，无人机所属的地理因素强烈地影响着用户对于植保作业的偏好，本系统默认无人机的位置即用户的位置，因此用户地理位置可以作为用户特征，以人口统计学推荐算法进行推荐。本文将人口统计学推荐与基于用户的协同过滤推荐相结合，以取得改进的效果。本文取植保作业记录并混合用户地理属性来计算用户相似度，具体用户作业评分数据结构为：

{用户 ID: {同省: 分值, 同市: 分值, 同区: 分值, 作业 1: 分值, 作业 2: 分值, 作业 3: 分值}, ...}

获得了目标用户与其他用户的上述评分结构后，就需要计算哪个用户距离当前最近。在计算最近邻时，皮尔逊相关系数和余弦夹角相似度都可用于离散评分和连续评分。皮尔逊相关系数主要用于衡量用户公共评分的相关程度[23]，但由于植保作业推荐算法混合了基于用户特征的推荐，因此获取的某些记录可能与目标用户不存在公共评分，因此本文选用余弦夹角相似度算法，具体计算方法如式(1)所示。

$$\text{sim}(c, v) = \frac{\overline{R_c} \times \overline{R_v}}{\|R_c\| \times \|R_v\|} = \frac{\sum r_{ci} r_{vi}}{\sqrt{\sum r_{ci}^2} \sqrt{\sum r_{vi}^2}} \quad (\text{式 1})$$

式中， c 与 v 分别代表两个用户， $\text{sim}(c, v)$ 代表两个用户之间的相似度。二者的评分向量分别为 $R_c = \{r_{c1}, r_{c2}, \dots, r_{ci}, \dots, r_{cn}\}$ 和 $R_v = \{r_{v1}, r_{v2}, \dots, r_{vi}, \dots, r_{vn}\}$ 。

计算出目标用户与其他用户的相似度评分后，将结果排序，分值范围在 0~1 之间，分值越高，相关程度越好。排序后选择固定数量的最近邻用户，通过最近邻来进行评分并预测目标用户可能感兴趣的项目(植保作业)。此时采用加权求和方法，计算得到的相似度作为评分的权重，所有相似度的加权评分总和即为结果。最后将相关植保作业评分排序，取前若干条作业作为推荐结果，具体流程见图 6。

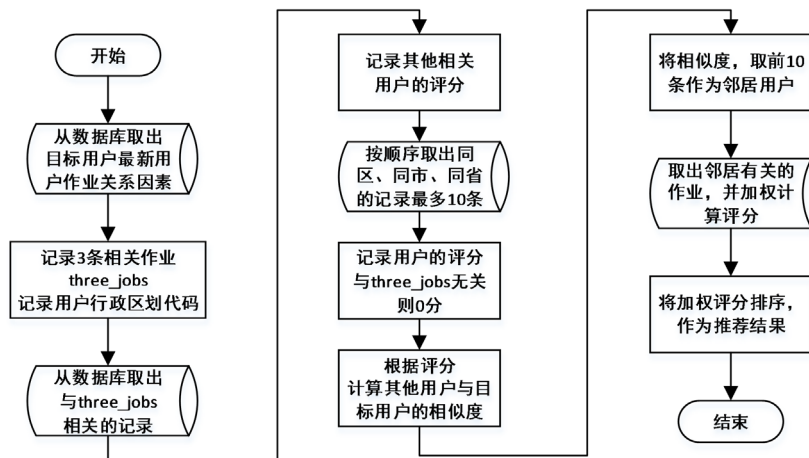


Figure 6. The plant protection operation recommendation algorithm based on mixing user characteristics and collaborative filtering

图 6. 混合用户特征与协同过滤的植保作业推荐算法流程图

3.3. 基于内容过滤的植保作业推荐算法

在推荐植保作业时，首先从 MySQL 数据库中，获取该用户的地理偏好因素的最新一条记录，利用记录的地理位置偏好坐标，在植保作业 R 树空间索引中搜索该坐标附近的若干条植保作业。而作为相对次要的植保类型偏好和作物类型偏好，将影响已经获得的若干条植保作业的排序。首先从 MySQL 数据库中，获取最新若干条的该用户的类型偏好因素记录，统计出用户最感兴趣的植保类型和最感兴趣的作

物类型。然后对已经从 R 树中获取的若干条植保作业进行排序，排序规则如下：用户最为感兴趣的植保类型的作业排在靠前，不是的排在靠后；而在前两部分中，用户感兴趣的作物类型排在靠前，不是的排在靠后。基于内容过滤的植保作业推荐算法流程图如图 7 所示：

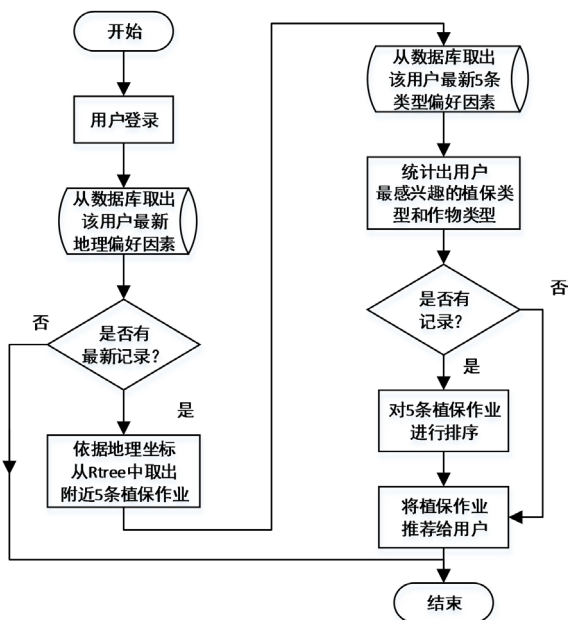


Figure 7. The plant protection operation recommendation algorithm based on content filtering
图 7. 基于内容过滤的植保作业推荐算法流程图

4. 结果与分析

4.1. 混合用户特征与协同过滤的植保作业推荐结果与分析

首先测试了目标用户是新注册用户。用户 2 与 3 有相关操作记录。按照规则，在目标用户没有记录时，仅按照用户之间的行政区划位置关系来评价相似度，用户 2 的评分应比用户 3 的评分高。经过余弦相似度计算用户 2 为 1 分，用户 3 为 0.816 分，作业 A 到作业 E 的加权分数分别为：8.16、9.08、10、4.08、10 分。

接着测试了目标用户已有对植保作业操作的情况，用户 2 与 3 有相关操作记录，目标用户则申请了作业 A 和 D，可见目标用户与 3 更为相似。此时同时考虑作业记录的相似度和行政区划位置的相似度。经过余弦相似度计算用户 2 为 0.775 分，用户 3 为 0.868 分，作业 B、作业 C、作业 E 的加权评分分别为 8.21、7.75、7.75 分。为了对比，做如下测试，测试数据描述如表 2 所示。

Table 2. Testing data
表 2. 测试情况

用户	位置	行政区划代码	作业 A	作业 B	作业 C	作业 D	作业 E
目标用户	海淀区	110108	10			10	
用户 1	海淀区	110108	5			5	10
用户 2	海淀区	110108		5	10		10
用户 3	东城区	110101	10	5		5	

目标用户申请了作业 A 和 D，目标用户与用户 1 和 3 比较相似。同时考虑作业记录的相似度和行政区划位置的相似度。经过余弦相似度计算用户 1 为 0.956 分，用户 2 为 0.775 分，用户 3 为 0.868 分，作业 B、作业 C、作业 E 的加权评分分别为 8.21、7.75、17.30 分。目标用户得到的结果如图 8 所示。

为您推荐以下植保作业								
作业流水号	所属用户	作业状态	作物类型	作业类型	地点	开始时间	价格	操作
作业E	Deluxe	未开始	果树	粉状药	北京,北京市,朝阳区	2016-08-23 21:19:49	1000	详情
作业B	Deluxe	未开始	高秆作物	液态药	北京,北京市,海淀区	2016-08-25 16:57:39	1000	详情
作业C	Deluxe	未开始	高秆作物	液态药	北京,北京市,海淀区	2016-08-25 16:54:45	1000	详情

Figure 8. Test results

图 8. 测试情况结果

从上述测试结果可以看出，混合用户特征与协同过滤的植保作业推荐算法，能够计算出与目标用户相似度最好的邻居用户，并能够预测出与目标用户评分最高的植保作业，即目标用户最可能感兴趣的植保作业。此外，由于混合了用户特征中地理位置这一属性，使得本算法在一定程度上缓解了协同过滤算法的冷启动问题：新注册用户无需对植保作业进行操作，只要附近有其他的用户有过操作记录，就可得到推荐结果。

4.2. 基于内容过滤的植保作业推荐结果与分析

本文在实验中设计了一些模拟数据，分布在北京市海淀区和东城区等地。对基于行政区划搜索日志的推荐进行实验，模拟了用户利用行政区划搜索植保作业，对北京海淀区的植保作业进行了筛选，之后返回首页。推荐了海淀区中心点附近 5 条植保作业。此时仅按照距离远近排序，最近的排在最上。接着又搜索了北京市朝阳区的植保作业，距离朝阳区中心点最近的 5 条作业。由此可以得出结论，基于行政区划搜索日志的推荐可以根据用户搜索的不同地点而更有针对性地推荐附近的植保作业。对基于 Rtree 的区域搜索日志的推荐结果进行了验证。首先利用区域搜索植保作业功能，搜索了海淀区香山一带的植保作业。又搜索东城区东直门一带的植保作业，推荐结果如图 9 所示。

根据您的搜索历史为您推荐以下植保作业								
作业流水号	所属用户	作业状态	作物类型	作业类型	地点	开始时间	价格	操作
20160824	Deluxe	未开始	高秆作物	液态药	北京,北京市,东城区	2016-08-23 21:20:38	1000	详情
20160708	Deluxe	未开始	果树	液态药	北京,北京市,东城区	2016-07-08 18:44:37	1000	详情
20160823	Deluxe	未开始	果树	粉状药	北京,北京市,朝阳区	2016-08-23 21:19:49	1000	详情
20160827	Deluxe	未开始	高秆作物	粉状药	北京,北京市,朝阳区	2016-08-26 09:48:33	1000	详情
Deluxe20150513	Deluxe	未开始	高秆作物	液态药	北京,北京市,海淀区	2015-05-20 00:00:00	5000	详情

Figure 9. Recommendation results based on Rtree region search log (Beijing, Dongcheng District, Dongzhimen)

图 9. 基于 Rtree 区域搜索日志的推荐结果(北京市东城区东直门)

由此可以得出结论，基于 Rtree 区域搜索日志的推荐可以根据用户搜索的不同区域而有针对性地推荐附近的植保作业。

验证类型偏好的推荐结果，本文通过点击查看多个植保作业详情获得了相关模拟数据。首先查看了作业类型为液态药，作物类型为高秆作物的作业。对比可见，在查看了液态药、高秆作物的植保作业后，推荐结果得到了正确的排序结果。首先液态药作业类型排在靠前的位置，粉状药靠后，其次高秆作物排在靠前，果树排在靠后。接着，又查看了粉状药、高秆作物的植保作业和粉状药、果树的植保作业。结

果如图 10 所示。首先粉状药作业类型排在靠前的位置，液态药靠后，其次高秆作物排在靠前，果树排在靠后。实验显示该算法得到了正确的结果。

根据您的搜索历史为您推荐以下植保作业								
作业流水号	所属用户	作业状态	作物类型	作业类型	地点	开始时间	价格	操作
20160827	Deluxe	未开始	高秆作物	粉状药	北京,北京市,东城区	2016-08-26 09:48:33	1000	详情
20160823	Deluxe	未开始	果树	粉状药	北京,北京市,朝阳区	2016-08-23 21:19:49	1000	详情
20160824	Deluxe	未开始	高秆作物	液态药	北京,北京市,东城区	2016-08-23 21:20:38	1000	详情
Deluxe20150513	Deluxe	未开始	高秆作物	液态药	北京,北京市,海淀区	2015-05-20 00:00:00	5000	详情
20160708	Deluxe	未开始	果树	液态药	北京,北京市,东城区	2016-07-08 18:44:37	1000	详情

Figure 10. Recommendation results based on type preference (Beijing, Dongcheng District, Dongzhimen)

图 10. 基于类型偏好的推荐结果(北京市东城区东直门)

5. 结论

本文研究并实现了基于日志处理的高效、可靠的无人机植保作业推荐算法，该推荐算法提高了植保无人机飞行监管系统的可用性、实用性、便利性和智能性，对于充分优化资源配置和保障农作物植保作业具有重要意义和实用价值。系统经过验证和实验，得出结论如下：

1) 混合用户特征和协同过滤的植保作业推荐算法，能够计算出与目标用户最相似的邻居用户，并能够基于此计算出目标用户可能评分最高的植保作业，并推荐给用户；基于内容过滤的植保作业推荐，基于用户的搜索与查看记录，能够为用户推荐可能感兴趣的植保作业。本文推荐算法能够产生正确且有意义的推荐结果，说明其具有智能性。

2) 使用 Storm 框架实现了植保无人机监管和作业推荐系统，其具有较好的实时性、便捷性以及实用性。

3) 目前仅设计了针对植保无人机用户的植保作业推荐算法，未来可以为农业生产用户提供植保无人机的推荐，让双方都能获得理想的推荐结果。

致 谢

感谢郑立华老师及所有在写作期间给予我指导的老师们！

基金项目

北京市科委重大专项——智能型农业植保无人直升机系统研发及工程应用(D1511000012150002)。

参考文献

- [1] 周海燕. 我国高效植保机械应用现状及发展展望[J]. 农业工程, 2014(11): 4-6.
- [2] 何雄奎. 改变我国植保机械和施药技术严重落后的现状[J]. 农业工程学报, 2004, 20(1): 13-15.
- [3] 徐文斌, 吕建国, 陈仕芬. 农药中毒患者死亡危险因素的 Logistic 回归分析[J]. 中国医药导报, 2014, 11(10): 143-145.
- [4] Cubber, G.D. Remotely Piloted Aircraft Systems—The Global Perspective—Yearbook 2013/2014. Citation, 2013: 133-134.
- [5] 温源, 薛新宇, 邱白晶, 等. 中国植保无人机发展技术路线及行业趋势探析[J]. 中国植保导刊, 2014(S1): 30-32.
- [6] 薛新宇, 兰玉彬. 美国农业航空技术现状和发展趋势分析[J]. 农业机械学报, 2013, 44(5): 194-201.
- [7] 郭宝录, 李朝荣, 乐洪宇. 国外无人机技术的发展动向与分析[J]. 舰船电子工程, 2008, 28(9): 46-49, 112.

- [8] 薛文娟. 基于层次聚类的日志分析技术研究[D]: [硕士学位论文]. 济南: 山东师范大学, 2013.
- [9] 杨锋英, 刘会超. 基于 Hadoop 的在线网络日志分析系统研究[J]. 计算机应用与软件, 2014, 31(8): 311-316.
- [10] 刘季函. 基于 Spark 的网络日志分析系统的设计与实现[D]: [硕士学位论文]. 南京: 南京大学, 2014.
- [11] 屈国庆. 基于 Storm 的实时日志分析系统的设计与实现[D]: [硕士学位论文]. 南京: 南京大学, 2016.
- [12] 孙思源. 基于 MongoDB 的网站日志分析系统的设计与实现[D]: [硕士学位论文]. 北京: 中国地质大学(北京), 2014.
- [13] 陈飞, 艾中良. 基于 Flume 的分布式日志采集分析系统设计与实现[J]. 软件, 2016, 37(12).
- [14] 王嫣然, 陈梅, 王翰虎, 等. 一种基于内容过滤的科技文献推荐算法[J]. 计算机技术与发展, 2011, 21(2): 66-69.
- [15] 庄景明, 王明文, 叶茂盛. 基于内容过滤的农业信息推荐系统[J]. 计算机工程, 2012, 38(11): 38-41.
- [16] 袁先虎. 基于混合用户模型的协同过滤推荐算法研究[D]: [硕士学位论文]. 重庆: 重庆大学, 2010.
- [17] 胡勋, 孟祥武, 张玉洁, 等. 一种融合项目特征和移动用户信任关系的推荐算法[J]. 软件学报, 2014(8): 1817-1830.
- [18] 刘树栋, 孟祥武. 基于位置的社会化网络推荐系统[J]. 计算机学报, 2015, 38(2): 322-336.
- [19] 任磊. 推荐系统关键技术研究[D]: [博士学位论文]. 上海: 华东师范大学, 2012.
- [20] Adomavicius, G. and Tuzhilin, A. (2005) Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Transactions on Knowledge & Data Engineering*, 17, 734-749. <https://doi.org/10.1109/TKDE.2005.99>
- [21] Leibusky, J., Eisbruch, G. and Simonassi, D. (2012) Getting Started with Storm. O'Reilly Media, Inc.
- [22] 冷亚军, 陆青, 梁昌勇. 协同过滤推荐技术综述[J]. 模式识别与人工智能, 2014, 27(8): 720-734.
- [23] 李季红. 协同过滤推荐技术的研究与改进[D]: [硕士学位论文]. 上海: 上海大学, 2014.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2164-5507, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>
期刊邮箱: hjas@hanspub.org