

基于BP神经网络的虚拟物品个性化设计推荐

陈雪峰, 李树刚

(上海交通大学工业工程与管理系, 上海 200240)

摘要: 通过提取玩家在网络游戏数据库中的数据特征, 用基于BP神经网络的数据挖掘方法挖掘出玩家对虚拟物品的各种属性的偏好, 为设计开发个性化的虚拟物品提供决策支持。针对传统神经网络中很难获取有广泛代表性的训练样本、常常导致普通神经网络对陌生样本推荐时精度不高的问题, 提出改进的BP神经网络, 依据专家知识对神经网络的权重进行初始化, 并根据训练样本对权重加以微调。仿真案例验证了该方法的有效性。

关键词: 数据挖掘; 网络游戏; 虚拟物品; BP神经网络

Personalized Design Recommendation of Virtual Item Based on BP Neural Network

CHEN Xue-feng, LI Shu-gang

(Industry Engineering and Management Department, Shanghai Jiaotong University, Shanghai 200240)

【Abstract】 This paper extracts gamers' personal information from the database of online game and mines to find their potential demands of virtual items by BP Neural Network (NN). To solve the problem of inaccuracy caused by lacking of supervising sample using the traditional NN, an advanced BP model with initial weights in BP NN set by expertise and adjusted by the trained samples is put forward, whose result is more accurate with fewer supervising sample. Simulation results show that the model is efficient and useful.

【Key words】 data mining; online game; virtual item; BP Neural Network(NN)

1 概述

网络游戏市场的竞争日趋激烈, 使得游戏公司无法以计时收费方式获取利润, 为了吸引玩家, 许多公司推出免费网络游戏^[1], 因此, 出售虚拟物品就成了各个游戏公司赚取利润的最重要的手段, 因此, 能否推出符合玩家需求以及个性偏好的虚拟物品来引导玩家购买成为网络游戏能否成功运营的关键。

传统的顾客消费行为影响因素量化分析主要基于综合评价法、主成分分析法、趋势分析法、相关回归分析法等^[2-3], 这些方法虽然能定量分析出整体顾客的相关需求, 却无法对顾客进行个性化需求与偏好的分析。因此, 上述方法无法实现个体玩家的个性化分析, 更无法挖掘出玩家对虚拟物品属性的具体需求。为此, 本文设计了基于神经网络的个性化虚拟物品设计推荐模型, 通过提取玩家在网络游戏数据库中的数据特征以及虚拟物品属性, 运用基于BP神经网络的数据挖掘方法挖掘出玩家对虚拟物品各种具体属性的需求和偏好量化值, 从而以清晰、明确的虚拟物品属性值指导个性化虚拟物品的设计。

此外, 在传统的BP神经网络中, 由于很难收集到精确、广泛、有代表性的训练样本, 常常导致神经网络在对陌生样本计算时, 输出精度不高^[4]。为此本文提出了改进的神经网络, 该方法的具体思想是: 依据专家知识对网络的权值进行初始化, 进而选用有代表性的训练样本进一步训练神经网络, 以便对网络的权值进行微调。最后, 结合当前比较流行的一款网络游戏——《热血江湖》进行了仿真, 仿真结果验证本文方法的有效性。

2 BP网络的输入因子-玩家特征

本文以《热血江湖》为例, 通过网络信息及专家的深入访谈总结出反映玩家潜在需求的20项特征。作为BP神经网络的输入。这些特征主要分成2大类: (1)静态特征, 包含玩家的主要属性, 如级别、性别、职业; (2)动态特征, 主要是指玩家在游戏中的行为特性。为了便于神经网络计算, 每项特征均采用5级列举比量表测量, 以1~5表示。玩家特征、特征说明及量化标准如表1、表2所示。理论上, 将来这些玩家特征数据均可从玩家数据库中抽取, 受当前条件限制, 本文的仿真数据通过问卷调查获得。

表1 玩家在游戏中的静态特征定义

静态特征	说明	量化标准
级别	玩家所扮演角色的级别	1: 低于20; 2: 20~40; 3: 40~60; 4: 60~80; 5: 80以上
职业	玩家所扮演角色的职业	1: 刀客; 2: 剑士; 3: 弓手; 4: 医生; 5: 枪客
派别	玩家所扮演角色的派别	1: 正派; 5: 邪派
性别	玩家所扮演角色的性别	1: 男; 5: 女
生命力	玩家所扮演角色的生命力	1: 非常低; 2: 低; 3: 一般; 4: 高; 5: 非常高
攻击力	玩家扮演角色配备的攻击力	1: 非常低; 2: 低; 3: 一般; 4: 高; 5: 非常高

基金项目: 国家自然科学基金资助项目(70501021)

作者简介: 陈雪峰(1983-), 女, 硕士研究生, 主研方向: 数据挖掘, 客户关系管理; 李树刚, 讲师、博士

收稿日期: 2007-07-20 **E-mail:** xfchen2004@sjtu.edu.cn

表 2 玩家在游戏中的动态特征定义

动态特征	说明	量化标准
游戏中好友的数目	玩家在游戏中的好友数目	1: 低于 10; 2: 10~20; 3: 30~40; 4: 40~50; 5: 50 以上
进行交易的次数	平均每周与其他玩家交易的次数	1: 从不; 2: 1 次以下; 3: 3 次以下; 4: 次以下; 5: 5 次及以上
做好事获搭桥费的次数	平均每周做好事获搭桥费的次数	1: 从不; 2: 1 次以下; 3: 3 次以下; 4: 次以下; 5: 5 次及以上
开店敛财的次数	平均每周开店敛财的次数	1: 从不; 2: 1 次以下; 3: 3 次以下; 4: 次以下; 5: 5 次及以上
打怪升级的偏好	玩家打怪时选择怪物级别的偏好	1: 非常低; 2: 低; 3: 一般; 4: 高; 5: 非常高
组队的偏好	玩家组队选择队员的倾向	1: 非常低; 2: 低; 3: 一般; 4: 高; 5: 非常高
丢弃无用物品的偏好	玩家丢弃无用物品的偏好	1: 从不; 2: 捡到就丢; 3: 本次下线前丢弃; 4: 下次上线的时候再丢弃; 5: 有空的时候再丢弃
加点偏好	增加点的倾向	1: 内功; 2: 生命; 3: 经验; 4: 攻击力; 5: 财富
强化/合成材料的偏好	玩家具备强化/合成材料时的行为	1: 从不; 2: 有空的时候再强化/学习; 3: 下次上线的时候再强化/学习; 4: 本次下线前强化/学习; 5: 立即强化/学习
学习武功/技能时的偏好	具备学习武功/技能条件时的行为	1: 从不; 2: 有空的时候再强化/学习; 3: 下次上线的时候再强化/学习; 4: 本次下线前强化/学习; 5: 立即强化/学习
探索地图的时间比例	平均每周探索地图的时间所占比例	1: 1/4 以下; 2: 1/4 以上; 3: 1/2 以上; 4: 3/4 以上; 5: 全部
聊天的时间比例	平均每周与其他玩家聊天的时间所占比例	1: 1/4 以下; 2: 1/4 以上; 3: 1/2 以上; 4: 3/4 以上; 5: 全部
打怪的时间比例	平均每周组队打怪的时间所占比例	1: 1/4 以下; 2: 1/4 以上; 3: 1/2 以上; 4: 3/4 以上; 5: 全部
PK 的时间比例	平均每周与对手 PK 的时间所占比例	1: 1/4 以下; 2: 1/4 以上; 3: 1/2 以上; 4: 3/4 以上; 5: 全部

3 BP 网络的输出因子——虚拟物品的属性

通过对《热血江湖》中虚拟物品属性的调研、分析、总结，把虚拟物品的属性归为如下 4 类作为 BP 网络的输出：

(1)功能属性，即虚拟物品所具备的可以提高玩家能力或改变玩家外表的功能，如在玩家打怪或斗斗时增加其攻击力或生命值。功能属性包括 4 个方面：改变外观，增加内功、生命、历练或经验，增加财富和增加职业攻击力。

(2)外观属性，每样物品都有其特有的外观，在虚拟世界中亦是如此。外观属性可以分成尊贵型、简洁型、可爱型、霸气型。

(3)社交属性，这里的社交属性是指该物品可以增强玩家的知名度或与别的玩家的情感。

(4)价格属性，虚拟社交的经济系统和现实社交非常相似，很多虚拟物品价值不菲，无疑价格也是玩家购买虚拟物品时考虑的重要因素。本文参考游戏中当前虚拟物品的交易价格划分出 5 个级别。

为了便于神经网络计算，每项属性同样采用 5 级列举比量表测量，以数值 1~5 表示。表 3 给出了各种属性的量化标准，其中，虚拟物品的价格单位为热血币，热血币与人民币的比率约为 1:10。

表 3 虚拟物品特征属性

功能属性	外观属性	社交属性	价格属性/热血币
1: 改变外观	1: 可爱型	1: 无需求	1: 低于 200
2: 增加生命	2: 浪漫型	2: 服饰类	2: 200~400
3: 增加经验	3: 简洁型	3: 情景类	3: 400~600
4: 增加攻击力	4: 尊贵型	4: 情侣类	4: 600~800
5: 增加财富	5: 霸气型	5: 广播类	5: 800 及以上

4 基于BP神经网络的个性化虚拟物品推荐设计

4.1 BP 神经网络结构设计

神经网络的功能为：采用高度非线性算法最大限度地提取不同类别样本的有效信息，再将不同类别的样本转换为不同类别的连续值量化输出。这一特点恰恰可以为匹配玩家和虚拟物品所利用，匹配的最终目的是为每个玩家(输入矢量)制定一个虚拟物品(输出矢量)。为了便于设置初始权值，本文采用的BP网络为 1 层输入层，1 层隐含层、1 层输出层的 3 层神经网络，且增加隐含层数目和隐层的节点数目并不能为分类效果带来质的变化^[4]。隐含层节点数为 20，输出层节点数为 4。神经网络的模型如图 1 所示。

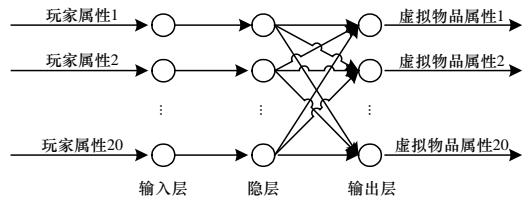


图 1 匹配玩家和虚拟物品的神经网络模型

图 1 中 BP 神经网络模型结构参数解释如下：

(1)输入层，主要输入玩家的动态特征数据，如表 1 所示。

(2)隐层，为对输入的玩家数据进行标准化处理后的数值，具体的量化标准见表 1。已知本文中输入神经网络玩家特征共有 20 项，因此，隐层含有 20 个节点。

(3)输出层，输出虚拟物品属性值，具体量化标准见表 2。共有 4 项虚拟物品属性值，因此，输出层含有 4 个节点。

4.2 BP 神经网络初始权值的设置

在传统的BP神经网络中很难获取有广泛代表性的训练样本，常常导致普通神经网络对陌生样本的输出精度不高。因此，本文基于层次分析法(AHP)^[5]，根据专家知识设置神经网络的初始权值，根据训练样本对权重加以微调，进而提出改进的BP神经网络。

层次分析法是一种实用的多准则决策方法。它把一个复杂问题表示为有序的递阶层次结构，通过人们的判断对决策方案进行排序。它采用相对权重，比传统的专家打分等经验法直接赋予权重更加客观，把对权重的判断量化。

将本文中的 4 种虚拟物品特征属性(功能属性、外观属性、社交属性和价格属性)一一通过 AHP 得出玩家行为特征对虚拟物品属性的需求重视度。层次结构如图 2 所示。

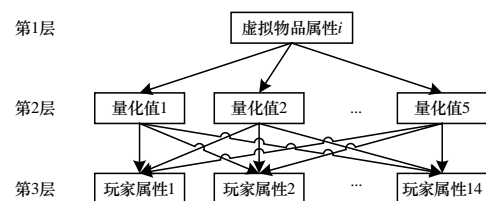


图 2 对应于应玩家需求的虚拟物品属性评价体系

在图 2 中，第 1 层表示虚拟物品特征属性；第 2 层表示属性 i 的量化值，为 1~5；第 3 层表示 14 项动态玩家行为特

征。以外观属性为例，第2层依次表示为：1: 可爱型；2: 浪漫型；3: 简洁型；4: 尊贵型；5: 霸气型。第1层与第2层之间采用最大值来决定权重，例如，玩家对于某项虚拟物品属性的5个量化值的权重为 $[w_1, w_2, \dots, w_5]$ ，如果 w_1 的值最大，则该虚拟物品属性的值为1，如果 w_2 的值最大，则为2，依此类推。第2层与第3层之间采用判断矩阵来决定其间权重。

已知BP网络结构为20-4，因此，连接权值 W_0 为 20×4 的矩阵，通过层次分析得到各项玩家行为特征对各项虚拟物品属性的需求权重。玩家的6项静态特征(级别、职业、派别、性别、生命力和攻击力)虽然会影响玩家对虚拟物品的需求，但无法反映对部分虚拟物品属性的需求倾向，因此直接设置为0，由BP网络进行优化。初始权值如表4所示。作为BP神经网络的初始权值进行训练时，将 $w_{i,j}$ 标准化到区间 $[0, 1]$ 。

表4 基于AHP得到的BP神经网络初始权重

玩家属性	虚拟物品属性			
	功能属性	外观属性	社交属性	价格属性
级别	0	0	0	0
职业	0	0	0	0
派别	0	0	0	0
性别	0	0	0	0
生命力	0	0	0	0
攻击力	0	0	0	0
游戏中好友的数目	3	5	5	1
进行交易的次数	4	3	4	5
做好事获搭桥费的次数	1	3	4	1
开店敛财的次数	4	2	3	5
打怪升级的偏好	5	1	3	1
丢弃无用物品的偏好	3	2	3	1
加点偏好	5	1	3	3
组队的偏好	5	1	4	2
强化/合成材料的偏好	5	4	2	2
学习武功/技能时的偏好	4	2	2	1
探索地图的时间比例	4	4	4	1
聊天的时间比例	3	5	5	1
组队打怪的时间比例	1	3	4	3
与对手PK的时间比例	3	4	1	1

4.3 BP神经网络的训练步骤

改进BP算法的样本训练步骤如下：

- (1) 设定输入层、输出层节点的个数，并初始化。
- (2) 设置初始权值矩阵 $W_{16 \times 4}$ ，初始权值 $w_{i,j}$ 标准化到 $[0,1]$ 之间，调整幅度为 $\pm 20\%$ 。
- (3) 输入学习样本，范围为 $[1,5]$ 。
- (4) 按BP算法对网络进行学习。
- (5) 判断权值是否超出允许范围：是，转(2)；否，转(6)。
- (6) 判断迭代步数是否超过规定的步数：是，转(7)；否，返回(4)。
- (7) 判断学习精度是否达到要求：是，转(8)；否，返回(4)。
- (8) 算法终止。

5 实验结果与分析

从《热血江湖》中随机选择玩家，并从其虚拟商城所出售的虚拟物品中选取10个不同功用和类型的物品，进行问卷调查，主要考察不同的玩家喜欢哪些虚拟物品。最终获得30份有效问卷，其中，20份作为训练样本；10份作为测试样本。改进算法的神经网络结构为20-4，特性函数 $f(\cdot)$ 为sigmoid函数，结合Matlab神经网络工具箱进行Matlab程序编写，按照表4设置初始权值进行网络训练。

推荐输出质量的评价也是推荐算法实验设计的一个重要问题。本文采用平均绝对偏差MAE作为度量标准。MAE通过计算预测的虚拟物品特征值与实际的虚拟物品特征值之间的偏差度量预测准确性，MAE越小，推荐质量越高。设预测的推荐结果矢量为 $[X_1, X_2, \dots, X_n]$ ，对应的目标矢量为 $[Y_1, Y_2, \dots, Y_n]$ ，则

$$MAE = \frac{\sum_{i=1}^n |X_i - Y_i|}{N}$$

为了验证本文方法的有效性，将改进BP神经网络算法、传统的BP神经网络算法和协同过滤推荐算法得到的推荐结果进行了比较，对比结果如图3所示。

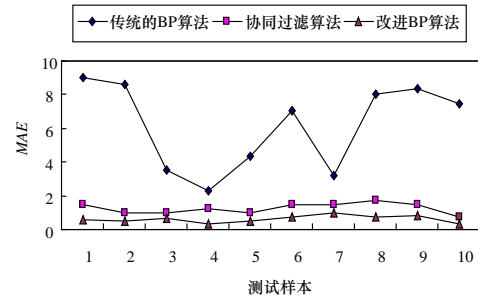


图3 样本仿真结果的MAE比较

可以看出，与其他2种算法相比，改进BP算法得到的MAE较小，能较精确地把玩家需求反映到虚拟物品的属性，从而较好地指导虚拟物品的个性化设计和改善。此外，改进BP算法保留了BP网络推荐响应时间短、推荐效率高等优点，并能有效避免协同过滤方法中通过相似顾客预测顾客需求带来的传递误差，真正实现顾客和产品的一一对应，为网络游戏中虚拟物品的设计开发提供科学、有效的决策指导依据。

6 结束语

本文针对当前网络游戏产业出售虚拟物品的盈利模式，在深入研究分析其客户特征和产品属性的基础上，以清晰、直观、便于科学分析的特征值形式表示玩家行为特征和虚拟物品产品属性，并运用BP神经网络把玩家行为量化输出为虚拟物品属性。通过引入专家知识优化BP网络的初始权值，显著提高了网络的鲁棒性和精度，并减少了样本数量。最后，通过实验与传统的顾客需求预测技术——协同过滤技术的计算结果作比较。结果表明，改进BP网络具有较高的输出精度，能较好地把玩家需求反映到产品属性，进一步指导个性化设计推荐。

参考文献

- [1] 王卓铭. 网络游戏中虚拟物品交易浅析[J]. 新闻界, 2006, (3): 59-60.
- [2] Jain C R K. Comprehensive Methodology for Evaluating the Commercial Potential of Technologies[J]. Theory Applications and Practice, 1997, 4(4): 220-235.
- [3] 李平. 主成分分析法在企业综合经济效益评价中的应用[J]. 南京邮电大学学报: 自然科学版, 1992, 12(4): 127-130.
- [4] Sultan M A, Eid S Z, Hegazi N H. Back Propagation Neural Networks for Direct Adaptive Control of Dynamic Systems[J]. Advances in Modeling & Analysis, 1994, 43(3): 13-23.
- [5] Omkarprasad V S, Sushil K. Analytic Hierarchy Process: An Overview of Applications[J]. European Journal of Operational Research, 2006, 169(1): 1-29.