

**本 科 生 毕 业 设 计**

**开题报告**



**学生姓名: \_ 夏立伟 \_\_\_\_\_**

**学生学号: 3120101964**

**指导教师: 陈为、巫英才**

**年级与专业: 大四 计算机科学与技术**

**所在学院: 计算机学院**

**一、题目：** 对MMORPG中多种行为的可视化分析系统

**二、指导教师对开题报告、外文翻译和中期报告的具体要求：**

**指导教师（签名）**

**年 月 日**

**毕业设计开题报告、外文翻译和中期报告考核**

**导师对开题报告、外文翻译和中期报告评语及成绩评定：**

|  |  |  |  |
| --- | --- | --- | --- |
| **成绩比例** | **开题报告**  **占（20%）** | **外文翻译**  **占（10%）** | **中期报告**  **占（10%）** |
| **分 值** |  |  |  |

**导师签名**

**年 月 日**

**答辩小组对开题报告、外文翻译和中期报告评语及成绩评定：**

|  |  |  |  |
| --- | --- | --- | --- |
| **成绩比例** | **开题报告**  **占（20%）** | **外文翻译**  **占（10%）** | **中期报告**  **占（10%）** |
| **分 值** |  |  |  |

**开题报告答辩小组负责人（签名）**

**年 月 日**

目录

[本科毕业设计开题报告 1](#_Toc446882148)

[1. 游戏数据可视化背景与简介 1](#_Toc446882149)

[2. 主要工作和目标 2](#_Toc446882150)

[3. 可行性分析 2](#_Toc446882151)

[4. 初步技术方案和关键技术考虑 5](#_Toc446882152)

[5. 预期工作结果 6](#_Toc446882153)

[6. 进度计划 6](#_Toc446882154)

[本科毕业设计外文翻译 7](#_Toc446882155)

# 本科毕业设计开题报告

## 游戏数据可视化背景与简介

大型多人在线角色扮演游戏(MMORPG)建立了一个虚拟社会，许多的玩家可以通过扮演不同的角色并且加入游戏中的不同的门派和工会进行玩家与玩家之间的行为交互。由于MMORPG中交互性和用户自参与的特点，成千上万的玩家们在不同的MMORPG中投入了金钱精力以及社会资源，而这些投入都成为了当今社会中电子商务中的重要的组成部分。据估计，2015年全世界范围内MMORPG的收入达到了1100万美元，并且在2017年有望达到1300万美元。

所以如果能够理解玩家在游戏中的行为模式和行为动态将会带来巨大的经济价值。除此之外，玩家在MMORPG中的个人行为的时间戳节点信息能够很好的帮助社会科学家来估计玩家的多种行为是如何被其他玩家影响的，并且探究这些影响是如何限制或者促进MMORPG游戏所创造的虚拟世界的。

之前对于MMORPG中玩家的行为模式的研究收到了一些大的限制。首先，大多数的研究知识对MMORPG游戏中个体的成长行为的一些片断进行了研究，这些研究在很大程度上限制了我们对玩家行为的连续动态行为的研究。其次，之前的研究将玩家脱离于MMORPG这个虚拟社会的上下文的情景中，并且将玩家的行为当作是个人决定的结果。这样的假设严重的限制了我们对个体，核心玩家以及其他玩家动态行为的社会结构所扮演的角色的理解。除此之外，大多数的研究知识将关注点放在MMORPG中的单种行为并且忽视了玩家的行为的多样性，这样的做法让社会学家很难找到多种行为之间的动态联系。

为了研究MMORPG玩家多种行为之间的动态关联并且调查不同社会结构所扮演的角色，我们必须面对两个主要的挑战。一个挑战就是如何如何测量玩家不同行为之间的相互作用并且对基于这些相互作用的不同的社会结构进行建模。在MMORPG游戏中每一个玩家都是从属于这个虚拟社会的多种社会结构中的，从最简单最基础的二元关系结构到三角关系结构最后到全世界级的结构。除此之外，这些社会结构都是互相渗透的。如何选取理论上重要的社会结构和从经验上获取他们所扮演的社会作用是我们所要面对的挑战。

第二个挑战是如何去可视化多种行为之间的动态关联并且可视化社会结构对动态行为的作用。动态相互作用包括了不同的商品以及不同组的核心玩家。尤其是这种相互影响会随着时间动态的变化类型，包括不同群体的玩家之间的影响以及不同组的玩家对商品消费的影响。因此，很难去建立一个简洁的，内容充足的并且具有高辨识度的可视化系统对这种影响力进行一个时间上的总结。除了对这些关系提供可视化之外，对不同组玩家之间对不同商品购买的影响力的异同点的解释也是一个挑战。同时由于玩家在游戏中的行为活动和他们自身的多种动态属性是在不同层面进行可视化的，所以如何可视化来进行这些探索和分析仍然是个难题。到现在为止，没有什么可视化方面的研究可以帮助分析动态行为和他们随时间变化的多种关系联系。

为了解决第一个挑战，我们系统的探索了那些会影响MMORPG游戏里玩家行为的社会结构并且找出了两个比较基础并且重要的结构化的机制:社会影响力和三元闭包。我们在这基础上提出了一种对影响力敏感的模型用来调查研究上述社会结构对于玩家行为互相之间作用的影响。为了解决第二个挑战，需要开发一个可视化分析系统BeXplorer，这个系统是基于一个比较流行的MMORPG游戏的开发商的需求进行开发的。

## 主要工作和目标

我们主要的工作分为两个部分，一个部分是对所获得的网易游戏的数据进行整理，分析，然后进行数学建模来构建一个对影响力比较敏感的模型来测量和分辨不同社会结构对玩家不同行为的影响。另外一个部分是构建一个可视化系统，整个系统由两个主要的可视部分组成的，一个是基于流的可视化，用来提供一个不同核心玩家之间以及玩家对不同商品消费影响的总体视图，另一个是基于像素的可视化，用来展示和可视化玩家的细节数据，比如玩家在游戏中的活动行为和玩家的属性值。这样就可以有利于我们去对基于流的可视化的一些特征进行更加深入的解释。我们所需要解决的问题如下：

* 在时间维度上，玩家之间的交流是如何影响他们的消费行为的？不同社会结构之间的影响力有什么不同？
* 哪一种玩家对其他人的影响力更大?那一类玩家更会受其他玩家的影响?
* 不同商品的消费受到的影响有什么不同？
* 在时间维度上，消费行为和交流行为是如何相互影响的？
* 当玩家影响的特定模式出现的时候玩家的交流网络结构是怎么样的？我们如何在那个时间点观察每一个玩家的社会行为?
* 当某种特定模式出现的时候，可视化系统是如何对假说在信息层面以及可靠性层面提供支持的？

## 可行性分析

1. 数据检查及用户需求

我们从一个拥有百万活跃用户的大型MMORPG中获取了数据。整个时间的区间是49周。整个游戏在许多不同的服务器上同时自发的运行。在不同时间戳上的不同玩家的不同属性以及活动行为都被存储在了CSV文件中。

游戏中的每一个服务器在研究中都被当作一个虚拟的世界。每个玩家在游戏中创建的角色被标记为*PlayerID*，每一个角色的属性都被分类成为静态属性（随着时间并不会变动）和动态属性（随着时间改变）。前者包括角色的性别（*avatar\_gender*）和职业（*job*），后者包括：*online\_time, grade, chat\_frequency, count\_pvp, consumption\_records. VIP\_level, Practice*和*Mastery*。这些属性的细节如下:

静态属性：

* *avadar\_gender*: 角色的性别是由玩家在创建时候决定的。
* *class*: 角色的门派是由玩家在创建角色时选择的。在各个游戏中一共有8个职业，并且每个职业都隶属于一个门派，现在一共有四个门派，分布是神机营，昆仑山，逍遥宫和天工阁。为了简便，我们将他们命名为门派1，2，3，4。由于同一门派在功能和行为特征上有相似之处，我们将在门派的层面对玩家的行为进行研究。

动态属性：

* *online\_time*: 玩家在游戏上所花的累计时间。
* *grade*: 玩家的等级，从1级到150级。
* *chat\_frequency*: 玩家和其他玩家之间的交流频次。玩家通过是以一对一的形式和他人交流的。每一次交流都是有时间戳以及交流上方的玩家id的记录的。对于他们所谈论的内容并不知道。
* *count\_pvp*: 游戏允许玩家在pvp的模式下与其他玩家进行战斗。而*count\_pvp*记录了玩家发生这种行为的频次。
* *consumption\_records*: 玩家可以使用他们的虚拟财富来购买商品。虚拟财富 包括元宝和银两。元宝的获得是通过充值现金或者用银两和其他玩家交换得到的。而银两则是玩家通过完成日常任务，击败其他玩家或者兑换元宝来积累的。为了让流通的货币统一，我们将银两通过每日的银两-元宝转换率转换成了元宝。所有的商品被分成了五类。商品的ID，购买时间戳，购买数量以及花费都被记录了下来
* *VIP\_level*: VIP的等级是从1到10级，游戏中充值更多，那么玩家的VIP等级也就越高。
* *Practice*: 学习和提高技能或者完成不同的任务能够提升修为。
* *Mastery*: 一旦玩家加入了一个帮会，他会被赋予修炼属性，修炼会随着玩家获得的技能以及对帮会做的贡献而提升。

这些数据基本上满足了数据分析的需要。

1. 社交网络的行为分析

在MMORPG游戏中的众多可能的行为里，有两种行为被认为是尤其重要的：消费和交流。玩家消费行为的重要性非常好理解，因为它直接关系到开发者的收益和他们对这个游戏持续的投资。然而，玩家的消费是十分容易收到其他玩家的影响的，尤其是那些有影响力的玩家。而这种影响力可以通过三种机制来进行计算，分别是：直接交流，社会影响力以及三元闭包。

在MMORPG游戏中，玩家可以通过直接的交流对其他玩家施加影响。玩家之间的交流纽带将会成为MMORPG游戏的虚拟社会系统中商品以及意见的首要传输途径。关于商品的重要的功能质量信息都能通过直接交流进行传播。

社会影响力机制表明社会个体会由于同伴的行为而改变自己的行为已使得行为接近一致。通过观察MMORPG游戏里面其他玩家的行为和技能，一个玩家可能会去消费某些商品来提高自己。三元闭包机制提出一个玩家和另一个玩家在一个网络里拥有更多的共同交流伙伴，那么这个玩家将会更倾向于去消费更多商品来提升自己从而维持与其他玩家的交流。

直接交流，社会影响机制以及三元闭包机制并不相互独立。相对的，它们将会一起影响用户的消费行为。

1. 社交网络可视化

社交网络可视化在近几年已经受到了相当大的关注。而这些研究让我们的系统提供了很好的理论支持。

早期的社交网络可视化的工作主要集中在网络的拓扑结构，比如Vizster和NodeTrix。Vizster主要使用节点连接的可视化方法来可视化个人为中心的图，并且能够探索图的连接和网络结构。NodeTrix是一个混合的展示方式，通过用矩阵的方式展示密集的连接节点。有些技术强调展示节点和边的属性，比如Semantic substrates在平面上赋予了节点不同的属性值。SaNDVis提供多种视图和交互让用户来探索个体与实体之间的关系。

也有很多的方法被用来支持动态网络的可视化分析。Steffen使用了聚类的超图来展示动态的聚类和聚合属性，使得用户能够找到并分析在时间上比较相类似的属性和子结构。Van den Elzen将高维空间的网络分布图投影到了二维平面来帮助可视化。两种方法都在宏观层面描述了网络，但是如果将这个直接应用到社交网络也许会丢失一些信息。

玩家与玩家，玩家与产品之间的复杂联系以及两种联系之间的动态关联是研究的中心。在某种程度上，我们的研究也被视作为社交网络的分析，用来辨认网络中节点和边的多种联系和结构。然而，当我们设计系统时候，我们需要展现更多的信息，比如交流，消费行为，参与的活动以及多种属性。因此，在这样的工作中需要一个更加统一的可视化设计来编码多种信息。

1. 暂存数据可视化

现在已经有许多可视化的技术用来寻找暂存数据的模式。主题流和故事线是非常经典的用来展现暂存数据的可视化技术。许许多多的研究都是基于这些可视化技术的。Tanahashi优化了故事线的布局并且提出了将故事线的可视化应用于流型数据的框架。StoryFlow将一种故事线的布局作为一种高效混合的优化方法，这样让实时的交互成为了可能。

最近几年比较流行的用来展现不断变化的文本和事件的方法是基于流的设计。LifeFlow将连续的时间聚合到一棵树上，并且将他们通过一系列不同的彩色的方块进行呈现。同这个研究相同，Outflow将不同的事件聚合到了一个图中，用边来代表同一时间内的转移。

我们所使用的游戏数据包括了玩家的动态属性，游戏的活动以及不同形式的行为。这些特征却让我们不能直接应用上面所说的方法。所以我们需要拓展现有的方法使得系统能够展现所种类型的随时间变化的的关系。

## 初步技术方案和关键技术考虑

整个项目将会分成两大块来进行，一个是对现有的数据进行整理分析并进行建模计算结果，另一个是构建可视化的系统，能够将整理分析以及计算出的数据进行一个合理的交互展示，并从中发现有用的特征模式。

数据建模方面我们首先需要计算出MMORPG中玩家的核心度，然后通过一个影响与受影响的模型来计算不同类型之间玩家的影响力，不同玩家对不同商品消费的影响力。这些影响力都是通过特定的模型进行计算的。在这一个过程中需要我们运用数据库的知识，也需要对python和R的编程有一定了解，能够运用python和R里面有关于社交网络等方面的库。

构建系统方面，我们打算整个系统是一套B/S的架构，也就是基于浏览器的系统，在框架方面我们选用了mean.js的架构，mean.js是由AngularJS, Express, bootstrap, node.js和mongodb组成的框架，而我们所用的可视化的图形库是javascript上面对svg应用比较成熟也比较有名的d3.js。

## 预期工作结果

我们的系统希望能达到以下一些效果：

1. 对玩家动态的行为和影响力进行一个总结。

我们的系统希望能够提供一个有效的，简洁的对动态数据的总视图。需要用不同颜色的彩带来代表不同玩家在不同时间的不断变换的关于交流和消费的影响力。不同的彩带是平行放置的，这样就可以比较容易的比较不同类型的玩家或者不同类型消费之间的异同点。同样彩带需要自带明暗，用来编码关于消费的信息。这样就可将玩家的交流行为，消费行为以及消费和交流的影响展现出来了。

1. 能够对一些有趣的特征进行深度探索。

像素块以及平行坐标会被放在中间视图中来展现每一个玩家个体的交流行为以及在游戏中的一些属性。不同视图之间的关联允许我们找到一些反常的交流行为的关键特征。总视图和细节之间的组合能够对假说进行更加深入的论证。

1. 提供视觉隐喻。

为了让用户有一个更加简洁和直觉的体验，我们的设计中有很多不同的隐喻。商品的消费信息被编码成流，交流信息的影响被编码成了流中的彩带，这些彩带根据流进行起起落落。这些设计构成了总视图。而交流的影响通过不同的方面展现出来：直接交流，社会影响机制以及三元闭包机制。

1. 独立的比较不同类别的玩家。

在不同缩放下的比较是必须的。在对总视图进行探究之后，我们可以将不感兴趣的消费流通过点击节点去除，在留下的流之间进行对比。

## 进度计划

1. 完成数据建模
2. 进行可视化设计
3. 搭建系统框架
4. 完成主视图
5. 完成中间视图
6. 完善细节视图
7. 测试完善

# 本科毕业设计外文翻译

文献题目：Different Aspects of Social Network Analysis

作者：Mohsen Jamali and Hassan Abolhassani, Web Intelligence Research Laboratory, Computer Engineering Department, Sharif University of Technology, Tehran, Iran

摘要：

社交网络是一群人（或者组织或者其他社会实体）的集合，通过朋友关系，工作关系或是信息交流所连接的。社交网络分析着重于对人际关系、组织和组织模式。社交网络分析同时提供了对人际关系的视觉和数学的分析。网络同样可以被当做社交网络。社交网络在网络上由超链接和网页构建而成。在这篇文章中，我们将对社交网络分析中的一些常用方法进行讨论

正文：

1. **介绍**

社交网络是一种在个体和组织之间的社会结构。它表明了实体连接的紧密性，包括从初识到紧密联系。电子邮件，交通，疾病传染和重要活动都能通过社交网 络所建模。

社交网络分析则是对人、组织、动物、电脑以及其他信息体之间的联系进行测量和映射的一种方法。网络中的节点是人和组，同时节点之间的连接则是表明 了节点之间的关系。社交网络框架中比较有趣的一点是它的子结构是由群组和圈 组成的。而这些子结构的数量，大小和连接情况可以告诉我们整个社交网络的行 为状况。

从定义可以知道，社交网络数据可以被看做是由实体和联系组成的社会关系系统。社会关系系统也包含了实体属性和多重关系等额外的信息。

1. **社交网络模型**
2. 用形式化的方法来展现社交网络在社交网络分析中用数学或者图形化的方法的原因之一是可以简洁系统地呈现对网络的描述。还有一个原因就是形式化的方法，尤其是数学，可以允许我 们使用计算机去对社交网络的数据进行计算分析。第三个原因是这种方法对数据 的呈现方式正式我们比较所希望看到的。

在整个社交网络分析中，大概可以分成以下三种方法：

* 描述方法，也可以通过图形的方法。

 分析过程，经常是基于对邻接矩阵的分解。

* 基于概率分布的数据模型

1. 使用图来展示社会关系

网络分析使用由点和线组成的图来展现实体和联系。当社会学家从数学家中得到这种图后，他们将这种图称为社会关系网图。

社会关系网图有很多的变种，但是它们都有相同的特点：所有的实体都是用一个带标注的圈来表示的，而连线则表示两者之间有一定的联系。用社会关系网图这种视觉化的方法能够提供对社交网络数据的第一手描述。不过虽然对于小的关系网是可以满足需求的，对于那些比较复杂的关系网络就比较困难了。

1. 用矩阵的方法来表现社会关系

在社交网络分析中最普遍的矩阵是由实体组成的行和列以及关系所组成的元素构成的。最简单并且最常用的是二值矩阵。也就是说，如果两者之间有联系，那么就在对应的矩阵栏里填上 1，反之则填 0。这类的矩阵可以说是所有社交网络分析的初始矩阵，并且由于它能够表现在我们的社会空间谁与谁比较接近，这种矩阵被称为是邻接矩阵。一般来说，在一个有向图中，关系纽带的发出自是列，纽带的目标是行。

1. 社交网络分析的数据模型

对社交网络的数据化分析已经了持续了 60 多年。在1970 年后，一个主要的方向是对实体之间进行可能性分析，虽然刚刚开始只是对小的群体进行分析。社交网络模型中的统计文献假设有n个实体和他们之间的联系信息。二元关系将会表现成为n×n的矩阵Y，其中如果i对j有某一种联系，那么*Yij*就为1，反之则为0。比如说*Yij* 为1如果i认为j是他的朋友。实体通常表现为节点而关系则是节点之间的箭头。如果矩阵是*Y*自反的，那么关系箭头就可以当作是无向的。

更一般的*Yij* 可以是有值的而不仅仅是二值的，代表了i和j之间的关系强度。除此之外，每一个实体都可以有特征集作为他们的统计信息。那么n维的向量

**浙江大学毕业设计**

**本科学生中期报告**

**学生姓名**

**实习单位/实验室**

**项目名称**

**导师姓名**

**合作导师姓名**

**项目开始日期**

**项目结束日期**

填表日期： 年 月 日

撰写提纲：（请以此结构，逐一论述）

1. **项目概况**

（项目简介、理论意义与价值，学生本人工作任务及内容等）

1. **工作成果及水平**

（完成工作内容、成果和水平、成果形式等）

1. **项目收获**

（工作中采用的新技术、新技能，及解决的关键问题；对团队合作，学习方法、工作方法诸方面的收获等）

1. **对工作建议**

（从个人发展，专业方向、项目，单位/实验室、学院建设等方面的建议）

1. **其它**