

---

# 公交移动支付问题的评估

## 摘要

随着智能手机的普及和移动支付技术的提高，公交移动支付手段越来越受到重视。本文通过大量的用户出行数据挖掘和分析，建立动态的第三方支付平台的商业盈利模型，并尝试提出移动支付平台的可行性分析和提高盈利的策略。

针对问题一，首先对附件一数据预处理，剔除异常数据、还原丢失数据、提取出行特征统计量。接着使用**关联规则**分析出行方式和支付方式间的联系。得到了乘坐地铁的人多数为移动支付的结论。最后作出统计量关于时间的一维关系曲线并结合关联规则结果，对该城市人口出行与支付特征进行了总结。

针对问题二，本文结合**价值网络理论**模型，首先分析移动支付平台的主要收益来源，重点考虑有形收益，从手续费，广告费，沉淀资金利息和服务费四个方面入手，建立商业盈利的数学模型，然后依据用户数量和服务定价的制约关系，运用**动态规划**模型，定量分析公交第三方支付平台的收支盈利情况。

针对问题三，首先对四个城市的任务点分别使用**Q-型聚类分析**，给出聚类图，引入不一致系数  $a$ 。然后，引入偏好系数  $\lambda$ ，同时考虑任务优先给予信誉度较高的会员和尽可能多的完成任务，使其成为**多目标规划**。其次，通过对问题二中模型约束条件参数进行修正，得到新的改进模型。改变不一致系数  $a$  和偏好系数  $\lambda$  的值，找出最适合的打包聚类距离。最后，通过与问题二方案比较，分析对最终完成情况的影响。

针对问题四，结合上文的分析结论和当前移动支付的环境和市场发展情况，给出一份关于移动支付的商业可行性报告，并着重从**成本**和**市场**方面，提出目标客户市场细分，提供更多服务以提升产品溢价等可行性方案。

**关键词：**

## 一、问题重述

### 1.1 问题背景

随着智能手机的普及和移动支付技术的提高，越来越多的支付手段可以转移到手机端。现有的现金缴费和实体公交卡刷卡的付费方式存在缺点，如公交卡在使用过程中存在着充值不方便、容易丢失、刷卡记录个人无法查看、跨地区无法使用等问题，现金支付带来很多不便，增加人工成本等，而公交移动支付则可以很好的解决这些问题。

### 1.2 问题提出

目前公交移动支付一般是通过第三方支付平台进行，现需要对该项目进行分析和评估。

1. 根据某城市的部分公交支付的信息和数据说明（附件 1，2），分析该城市乘车人的出行支付特征。

2. 建立一个公交第三方支付平台的商业盈利数学模型，定量分析公交第三方支付平台的收支和盈利情况。

3. 在问题 1 四分之一公交地铁安装移动支付在运营期间得到的数据下，估计该城市全部公交实现第三方平台支付后的盈利情况。

4. 结合前面的计算结果和结论，给移动支付公司写一份 500 字以内的商业计划可行性报告，并给出增加公司盈利的可行性方案建议。

## 二、问题分析

### 2.1 问题的总分析

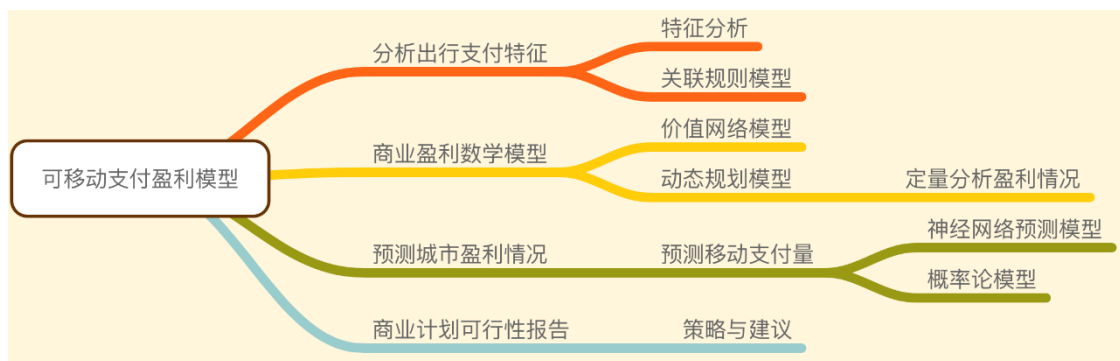


图 2.1 问题分析流程图

## 三、名词解释与符号说明

### 3.1 名词解释

**中心距离 $\rho_i$** : 该城市第  $i$  个任务点到中心点的距离

**道路交通运行指数**: 是对路网交通总体运行状况进行定量化评估的综合性指标。简称交通指数，取值范围为 0~5，分为畅通（0~1）、基本畅通（1~2）、缓行（2~3）、较拥堵（3~4）和拥堵（4~5）五个等级。交通指数越大，表明一次出行相比顺畅状况多花费的时间越长。

**不一致系数  $r$** : 聚类图中的纵坐标，可以利用不一致系数确定分类个数。

**偏好系数  $\lambda$** : 多目标规划中，各个决策变量所占的权重。

### 3.2 符号说明

表 3-1 符号说明表

符号	符号说明
$\alpha_i$	第 $i$ 个任务点附近 2Km 范围内的会员数
$\rho_i$	中心距离
$\beta_i$	信誉平均值
$\theta_i, i \in (1,2, \dots, 7)$	逻辑回归参数
$x_1$	任务点定价
$x_2$	任务点 2 公里范围内的会员数
$x_3$	任务点 2 公里范围内的信誉度平均值
$x_5$	任务点所在区域的人口密度
$x_6$	任务点所在区域的可支配收入
$x_7$	任务点所在区域的道路交通运行指数
$x_i^* \quad i \in (2,3,5,6,7)$	对 $x_i$ 的修订
$y$	(0/1) 任务是否完成
$\lambda$	偏好系数
$g$	Sigmoid 函数表示
$h_{\theta}(x)$	假设函数
$q_m$	第 $m$ 个任务点
$d_{ij}$	第 $i$ 个任务点到第 $j$ 个任务点间距离
$x_{ij}$	(0/1) 是否走过任务点 $q_i$ 与 $q_j$ 之间的路

## 四、模型假设

本文根据求解需要，作出如下合理假设：

---

## 五、模型建立与求解

### 5.1 问题一模型建立与求解

#### 5.1.1 模型的准备

##### 数据预处理

##### 1. 异常数据处理

考虑刷卡机器故障数据（0001-1-1）的两种情况：

- 刷卡机故障，但仍能正常显示刷卡方式。

UPTIME	PAYTYPE
0001-1-1	0
0001-1-1	0
0001-1-1	0
0001-1-1	0
0001-1-1	1
0001-1-1	0
0001-1-1	0
0001-1-1	1

- 刷卡机故障，且刷卡方式数据遗失。

0001-1-1	Null
0001-1-1	Null
2017/2/7 8:50	0
2017/2/7 13:36	1
0001-1-1	Null
2017/2/7 8:08	1
2017/2/7 13:17	1
2017/2/7 14:52	0
0001-1-1	Null

对于第一种情况的数据，由于不影响客观统计因此我们选择正常保留。

对于后者，产生数据丢失。初步统计数据丢失量大约为 10 万组，占总数百分之十。若直接忽略，可能对最后统计结果产生影响。因此我们使用**最近邻插值法**，将数据进行还原。

##### 2. 异常数据剔除

在统计时，发现刷卡支付方式存在 0、1 之外的方式，如 30、31。由于此类数据量较少，因此我们不予考虑，将其直接剔除。

---

2017/2/7 21:59	31
2017/2/7 18:26	31
2017/2/7 21:31	31
2017/2/7 21:38	31
2017/2/7 20:26	30
2017/2/7 21:12	30
2017/2/7 20:12	30
2017/2/7 18:21	30
2017/2/7 19:44	30

### 3. 出行特征提取

为更好的描述出行特征，我们参考论文《出行特征》，定义了如下统计量。

- 平均公交使用次数 ( $T_{av}$ )

定义为地铁公交使用总次数 ( $T$ ) 与不重复 ID 总数 ( $N$ ) 的比值，即

$$T_{av} = \frac{T}{N}$$

- 付款方式比例 ( $R_1$ )

定义为移动支付总量 ( $M$ ) 与支付总量 ( $P$ ) 的比值，即

$$R_1 = \frac{M}{P}$$

- 地铁公交使用比例 ( $R_2$ )

定义为地铁使用总次数 ( $T_d$ ) 与地铁公交使用总次数 ( $T$ ) 的比值，即

$$R_2 = \frac{T_d}{T}$$

对原始数据进行处理后，我们分别提取不同季度、不同时间段的  $T_{av}$ 、 $R_1$ 、 $R_2$  的信息。详细见附录 1。

## 5.1.2 模型的建立与求解

- 基于 Logistic 回归的最近邻插值法

下面为准备模型，主要叙述基于逻辑回归算法的最近邻插值法数学模型。

### ➤ 逻辑回归

逻辑回归 (Logistic Regression) 是机器学习中的一种分类模型。分析数据得，出行方法因变量  $y \in \{0,1\}$ ，而要使线性回归输出结果可能满足上述要求，为使得最后输出  $h_{\theta}(x) \in \{0,1\}$ ，首先需要使  $h_{\theta}(x) \in [0,1]$ 。逻辑回归中使用 Sigmoid 函数记作  $g(x)$ ，作为外层函数，内层使用线性回归。即

$$\begin{cases} h_{\theta}(x) = g(\theta^T x) \\ \theta^T x = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_n x_n \end{cases}$$

由于 Sigmoid 函数<sup>[3]</sup>的特殊性质，使得最终 $h_{\theta}(x)$ 可以满足 $h_{\theta}(x) \in \{0,1\}$ 。  
具体步骤如下：

Step1: 构造预测函数 $h_{\theta}(x) = g(\theta^T x)$ 。其中

$$\theta^T x = \sum_{i=1}^n \theta_i x_i = \theta_0 + \theta_1 x_1 + \cdots + \theta_n x_n$$

$$\text{则, } h_{\theta}(x) = \frac{1}{1 + e^{-\theta_0 + \theta_1 x_1 + \cdots + \theta_n x_n}}$$

Step2: 构造损失函数 J

$$\begin{aligned} J &= -\frac{1}{m} l(\theta) \\ &= -\frac{1}{m} \sum_{i=1}^m y^{(i)} \log(h(x^{(i)})) + (1 - y^{(i)}) \log(1 - h(x^{(i)})) \end{aligned}$$

Step3: 运用梯度下降，不断更新 $\theta_j$ ， $j \in (1, 2, \dots, n)$ ，直到样本被训练完。得出参数 $\theta$ 拟合结果

#### ➤ 最近邻插值法

**最近邻插值法**最初用于从离散分布气象站的降雨量数据中计算平均降雨量，在一个文件中，数据紧密完整，只有少数点没有取值，可用最近邻点插值法来填充无值的数据点。

最终模型步骤如下：

1. 找到每个数据丢失点 ID。由最近邻插值法概念，从该周数据中找到所有相同 ID 的出行刷卡信息。
2. 根据找到的数据使用逻辑回归模型，建立此人刷卡方式与出行方式间的回归方程，
3. 将数据丢失点代入回归方程，对结果进行预估，实现插值。

#### ● 基于 Apriori 算法的单维布尔关联规则模型 (Association Rules Mining)

关联规则用来描述数据库中数据项之间某种潜在关系的规则，它是数据挖掘 (Data Mining) 的一个研究方向。

为方便理解，解释一些术语如下：

- 
- **项目**：交易数据库中的一个字段，对本文来说是指一次出行中的一个特征，如：移动支付
  - **事务**：某个客户在一次交易中，发生的所有项目的集合：{乘坐地铁，乘坐公交，移动支付，乘车卡支付}
  - **项集**：包含若干个项目的集合如：{乘坐地铁，移动支付}、{乘坐公交，乘车卡支付}等
  - **支持度**：项集 {X, Y} 在总项集中出现的概率
  - **频繁项集**：某个项集的支持度大于设定阈值（人为设定或者根据数据分布和经验来设定），即称这个项集为频繁项集。
  - **置信度**：在先决条件 X 发生的条件下，由关联规则 {X→Y} 推出 Y 的概率（见下面的例子）
  - **提升度**：表示含有 X 的条件下同时含有 Y 的概率，与无论含不含 X 含有 Y 的概率之比。

具体步骤如下：

- 1) 数据筛选，根据数据进行项集筛选，清洗掉公共项目。比如：移动支付，乘车卡支付。
- 2) 根据支持度（support），从事务集合中找出频繁项集（使用算法：Apriori 算法，算法具体实现见附录 2）。
- 3) 根据置信度（confidence），从频繁项集中找出强关联规则（置信度阈值需要根据实验或者经验而定）。
- 4) 根据提升度（lift），从强关联规则中筛选出有效的强关联规则（提升度的设定需要经过多次试验确定）。

说明：满足最小支持度和最小置信度的关联关系叫做强关联关系

- 如果  $lift > 1$ ，叫做有效的强关联关系，
- 如果  $lift \leq 1$ ，叫做无效的强关联关系
- 特别的如果  $lift(X \rightarrow Y) = 1$ ，则称 X 与 Y 相互独立

### 5.1.3 结论分析

使用插值后得到的数据（具体数据见附录 2）进行数据挖掘，得到最终结论，并对其进行时间维度（时间段和季度）分析得到最终结论如下。

#### 1. 关联规则模型结论分析

---

分别定义 4 种关联规则，分别为 {乘坐地铁→公交卡支付}、{乘坐公交→公交卡支付}、{乘坐地铁→移动支付}、{乘坐公交→移动支付}。  
并根据数据预处理中结果，设定支持度和置信度的阈值分别为 3/7 与 5/7。

在满足最小支持度和最小置信度的条件下得到：

lift (乘坐地铁→公交卡支付) = 1

lift (乘坐公交→公交卡支付) = 1

lift (乘坐地铁→移动支付) = 1

lift (乘坐公交→移动支付) = 1

可以看出东莞市平均会员数小于总任务点平均会员数，但区域完成率却异常的高。而深圳市平均会员数高于总平均会员数，完成率却很高。因此可以得出，任务未完成的原因除任务点附近会员数外，还与具体城市的相关性质，如经济、交通、人口等有关。下面进行详细分析。

a). 本问涉及的城市深圳，广州，东莞，佛山分别是广东省内经济发展前 4 所城市，城市经济发展水平较高。

b). 总完成率在全局上，深圳的任务完成率最低，东莞的任务完成率最高。

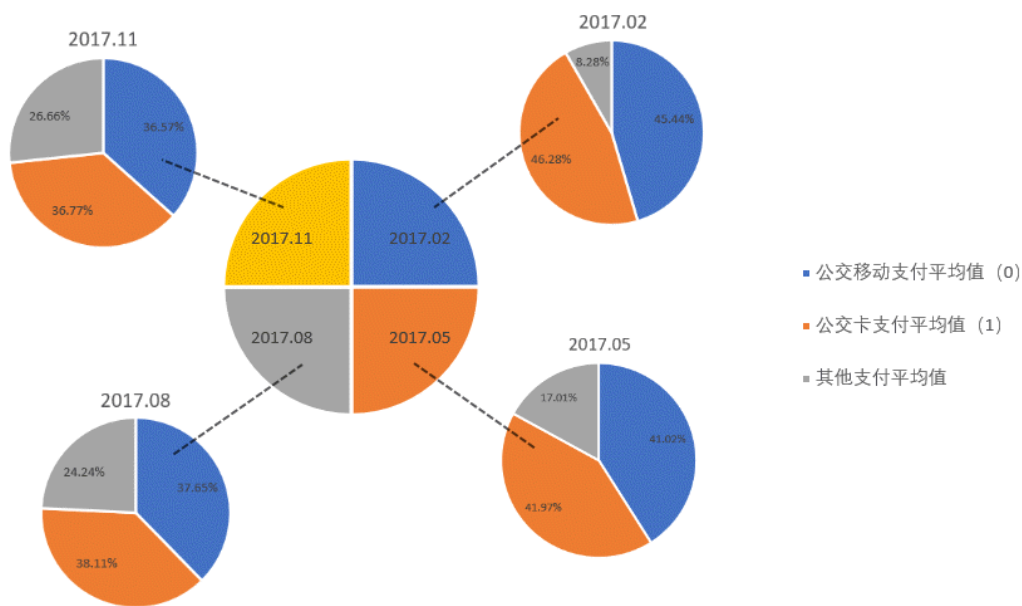
c). 深圳市人均可支配收入最高，佛山人均可支配收入最低。

d). 深圳市平均交通指数最高，佛山，东莞较低。

## 2. 出行特征比值分析

使用 MATLAB 编程（具体程序见附录 2），分别得到每个月一周内出行特征  $T_{av}$ 、 $R_1$ 、 $R_2$  的平均值如下图所示。





- a) 完成率较高的区域，普遍交通指数较低，交通情况良好
- b) 未完成区域的会员信誉均值低于完成区域。

### 3. 出行特征随时间维度变化分析

我们在地图上找到了未完成任务点聚集的区域案例，针对这些案例，我们具体调查区域地理位置以及周边环境，得出任务点未完成的个别典型原因。

- a). 对于任务 A0034 虽然远离城市中心，但是靠近地铁，交通情况良好，可能利于完成任务。
- b). 对于任务 A0181，靠近城市中心，人均密度大，可能由于会员的主观因素或者高程信息导致任务不易完成。

### 4. 总结

通过对未完成的任务点定量分析比较，结合未完成的一些典型原因及普遍原因。我们得出结论如下：

- 1. 任务价格会对任务执行成功与否产生影响。

2. 当城市人均可支配收入较高时，可能价格因素对会员执行任务的吸引力降低，导致任务完成率低。
3. 任务完成情况与其附近的交通运行状况有很大的联系。
4. 人口密度在宏观上与任务完成率成反比。

红色代表任务完成，紫色代表任务未完成。

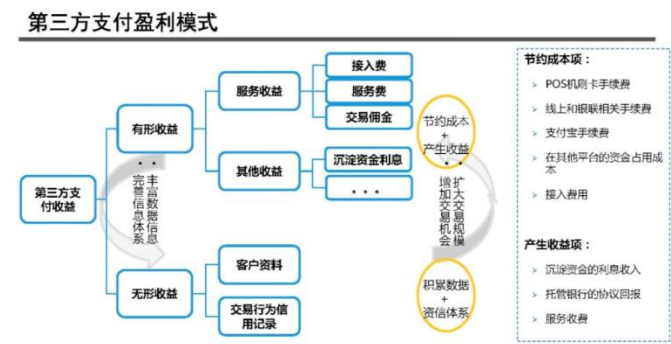
图 5.6 任务完成程度的三点图

## 5.2 问题二模型建立与求解

### 5.2.1 模型的准备

#### 1. 模型参考

参考……，盈利模型如下图所示。

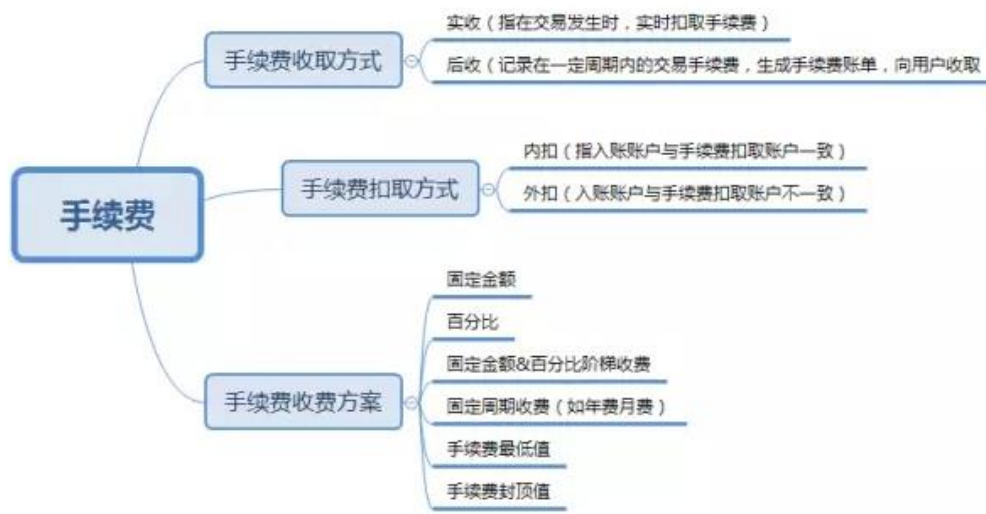


盈利模式主要从有形收益和无形收益方面进行分析和建立模型。

#### 2. 概念解释

- 手续费（）

手续费是第三方支付平台费用的最传统的盈利模式之一，即第三方支付平台向使用者收取手续费与银行支付的手续费之差。



- 广告费

第三方支付平台费用中的广告费，即有商户在第三方支付平台的电脑 PC 端和手机 APP 上发放宣传广告，第三方支付平台会收取一定的广告费用。而此笔广告费用的高低取决于第三方支付平台的知名度和使用客户的覆盖度。

- 沉淀资金利息

沉淀资金主要指支付机构为办理客户委托的支付业务而实际收到的预收货币代付资金。

- 服务费

服务费是指第三方支付平台为其客户提出支付解决方案，提供支付系统以及各种增值服务。盈利模式主要考虑该平台提供的增值服务收费和相应的用户人数。

## 5.2.2 模型的建立与求解

- 基于价值网络理论的第三方移动支付平台支付盈利模型

- 价值网络理论

价值网络是由客户、供应商、合作企业及其之间的信息流构成,由真实的客户需求所触发,并能快速可靠地对客户偏好作出反应的动态分析网络。

- 第三方移动支付平台支付盈利模型

为了简化求解过程，假设只考虑有形收益，不考虑无形收益的影响。

---

## 指标的量化

### 1. 手续费

$$y_1 = (q - p) * R$$

$$R = (\sum a_i) * b$$

其中， $y_1$  表示手续费总盈利， $q$  表示手续费比率， $p$  表示银行利率， $R$  表示第三方总共收到用户的金额， $a_i$  表示不同支付方式的人数， $b$  表单次公交，地铁的费用。

### 2. 广告费

本文投放广告价格主要考虑支付平台使用人数。

盈利模型：

$$y_2 = k_2 a + k_1 - k_0$$

其中， $y_2$  表示广告费总盈利， $k_2$  表示广告费随支付平台使用人数  $a$  变化的部分， $k_1$  表示广告费固定部分， $k_0$  表示投放广告所需要的人力物力支出。

### 3. 沉淀资金利息收入

盈利模型：

$$y_3 = (1 - D) * (m - n) * p$$

其中， $y_3$  表示沉淀资金利息收入， $D$  表示风险准备金比例， $m$  表示预收代付资金总额， $n$  表示现阶段需要实际支出， $p$  表示银行定期利率

### 4. 服务费总盈利

$$y_4 = \sum h_i * g_i$$

其中， $y_4$  表示服务费总盈利， $h_i$  表示增值服务收费， $g_i$  表示该增值服务用户人数。

### 5. 总盈利模型：

$$Y = \sum_{i=1}^4 y_i$$

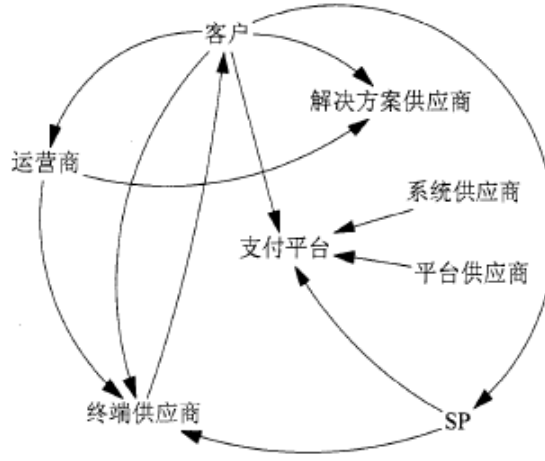


图 6-4 基于价值网络的移动支付模式仿真模型

考虑到移动支付平台用户数量和服务定价成反比的关系，不妨假设用户数量满足  $g = \frac{k}{h}$ ，

其中  $k$  为常数。建立动态规划模型：

$$\begin{aligned} \max Y &= \sum_{i=1}^4 y_i \\ \begin{cases} y_1 = (q - p) * R \\ R = (\sum a_i) * b \\ y_2 = k_2 a + k_1 - k_0 \\ y_3 = (1 - D) * (m - n) * p \\ y_4 = \sum h_i * g_i \\ g = \frac{k}{h} \end{cases} \end{aligned}$$

有上述规划模型可知，在假设成立的情况下：

当满足  $Y \geq 0$  时，第三方支付平台可实现盈利。

当满足  $\max Y$  时，第三方支付平台可实现最大盈利。

## 5.3 问题三模型建立与求解

### 5.3.1 模型的准备

刷卡设备的增加对盈利情况的影响，我们将从收入和支出两方面分析。

收入：刷卡设备增加主要增加使用移动端刷卡的人数，因此使用移动端产生的手续费会提升。所以只需估计设备增加后对应的移动支付人数增加量便可估计出收入额。

支出：刷卡设备的增加必定会增加设备成本以及安装成本，我们将这类成本统一定义为  $M$ 。本文假设增加 4 倍设备量，该成本  $M$  也线性增加 4 倍。

由前文分析可得：在知道四分之一的公交地铁安装移动支付设备的数据，估计全部公交实现第三方支付平台支付后的盈利情况。显然这是一个预测问题。同时，由于数据量巨大，因此可将频率等效为概率。在先验概率已知的情况下求后验概率，可用贝叶斯公式进行求解。

本文打算同时使用传统概率模型和神经网络模型进行求解，并最后比较两种结果。

### 1. 神经网络模型准备

BP 神经网络算法（简称“BP 算法”）采用梯度下降法的原理对神经进行训练，它是一类有导师指导学习的训练算法，其网络结构主要是由输入层、隐含层和输出层组成。

本文神经网络模型结构如下图所示：

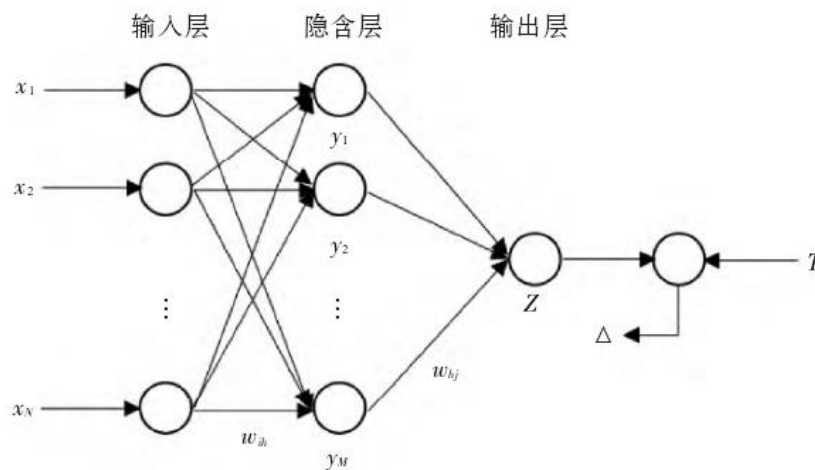


图 4.7 神经网络模型结构图

### 2. 概率模型支付特征分析

模型假设：

1. 假设乘客有移动支付习惯，且公交有相应设备，则认定必然使用移动支付。不考虑因特殊情况造成的支付方式改变的情况。
2. 假设所给数据足够反应该市出行特征。
3. 对于两种支付习惯均不满足的乘客，作为异常数据剔除。

---

4. 假设所有公交装有设备与部分公交装有设备情况中的装有设备的公交中移动与非移动支付人口比值相同。

**有移动支付习惯人口 ( $N_1$ )**

考虑到仅有 1/4 的公交装有移动支付设备，因此认为有过移动支付行为的乘客为有移动支付习惯人口。

**无移动支付习惯人口 ( $N_2$ )**

由于数据量巨大，因此假设在所有数据中出现次数超过 5 次，但未使用过移动支付的人口定为无移动支付习惯人口。

**有、无移动支付习惯人口之比 ( $P_1$ )**

遍历所有数据，剔除重复 ID，得

$$P_1 = \frac{N_1}{N_1 + N_2}$$

其中， $N_1$  为有移动支付习惯人口， $N_2$  为无移动支付习惯人口。

### 5.3.2 模型的建立与求解

- **基于 BP 神经网络的预测模型**

- **BP 神经网络**

输入层为历史样本数据，理论上样本越多，训练效果越好。然而，不同时期内支付方式和出行特征规律会有偏移。比如夏季偏多于冬季；随着时间推进，使用公共交通人数先增多后减少等，故用于训练的历史数据要考虑时效性。另外，输入样本越多，计算开销越大。本文规定以每个月数据中 4 天的数据作为神经网络的 7 个输入层。

因为支付方式与设备量之间非线性度很高，总体上无规律、较模糊，需要长时间更充分的训练，故隐含层节点数应较高。本文选取 12 个中间节点。

输出层即为待估计的移动支付量分布。当网络结构确定后，就可以将准备好的样本数据输入到网络中进行训练。

- **BP 算法的训练步骤简述如下：**

- 1) 初始化网络训练参数；
- 2) 网络的正向传播计算；
- 3) 网络误差的反向传播及其对各层连接权值的调整；
- 4) 再次进行正向训练。

可知在不同状态下（如不同月份）移动支付使用情况有较大差别，故应分类预测。本文考虑分为 4 类：工作日、工作日、周末、周末，根据待预测日期状态在对应的类中用 BP 神经网络进行预测。每过去 1d，当天使用数据即化为历史样本数据进行再学习、改进。

## ● 基于支付特征分析的概率模型

### ➤ 模型思路

由于数据量足够大，用频率代替概率，估计出 1/4 安装设备公交中的总人口数量。进而得到剩余未使用移动支付的人口数量。最后估计出所有公交中的移动支付人数。

### ➤ 具体步骤

1. 遍历所有数据，求的支付习惯人口比例  $P_1$
2. 遍历第  $j$  天数据，求的该天移动支付人口  $N_{1j}$ ，总人口  $N_j$ ，并依次求的  $N_{1j}$ ， $N_j(j = 1, 2 \dots n)$

3. 计算无设备公交总人口  $N'_j$

$$N'_j = N_j - N_{1j}/P_1$$

4. 计算全部公交安装设备后，每天移动支付增加总人口数  $T_j$

$$T_j = N' \times P_1 + N_{1j} = N_j \times P_1$$

5. 计算由于设备增加，而每天增加的移动支付人口数  $T_{zj}$

$$T_{zj} = T - N_{1j} = N_j \times P_1 - N_{1j}$$

6. 对  $N$  天求平均得到最终结果  $T$

$$T = \frac{1}{N} \sum_{j=1}^N T_{zj}$$

## 5.3.3 结论分析

考虑到，安装设备成本和手续费收入的同时增加，结合第二问盈利模型，并假设安装前收益为  $G$ ，安装后收益为  $G'$ ，则

$$G' = G + \alpha T - 4L$$

其中， $\alpha$  为手续费收益由模型二中定义， $L$  为安装设备成本。

对数据进行处理，我们得到第  $j$  天结果如下（剩余天结果见附录 1）

$P_1$	0.796
$N_{1j}$	476450
$N_j$	961760



$T_{zj}$	289111
----------	--------

所以可得最后

$$G' = G + \alpha 289111 - 4L$$

使用神经网络同样对第 j 天分析，得到最终结果如下

$T_{zj}$	330983
----------	--------

最终结果为：

$$G' = G + \alpha 330983 - 4L$$

## 5.4 问题四模型建立与求解

商业计划可行性报告

- 产品和服务内容

第三方公交可移动支付平台，为用户提供手机支付等一系列服务。

- 项目的独特性和优势

便捷性：不用找零，资金实时到账，充值方便，降低人工成本

开放性：可跨区域使用，公交地铁等通用

安全性：资金即时到账无风险，同时也无需保存大量的现金在钱包里

准确性：可避免操作中的错误，盗窃的发生

- 利润来源及持续盈利的商业模式

利润主要来源：手续费，广告费，沉淀资金利息收入和服务费。通过用户人数来对广告费等相关收入项目进行动态调整，提高平台的边际利润，保持商业上的持续盈利。前期目标主要扩大用户的数量及降低成本，后期是增加一系列增值和溢价服务，提高用户的粘性，改善技术，最终不断提高平台的盈利水平。

- 面临的风险及对策

1. 通过数据的挖掘，可以发现城市的可移动支付人数比率普及率不高，市场潜力较大，但需要扩大支付技术优点的宣传，相关基础设施的建立，即是机遇，也是挑战。

2. 现有移动支付平台较多，竞争程度比较大，必须提高平台的核心竞争力，为用户提供资金安全的保证，及其他一系列相关的增值服务。

- 增加盈利，降低风险的可行性方案

第一阶段：在公交移动支付业务的发展初期,通过对目标客户市场的细分,掌握目标市场的客户群体特征,快速锁定移动支付的目标市场,进行精准营销,以远程支付、短信小额支付方式,针对中高端市场客户、学生群体等消费多样化,实施特定群体的低成本集聚化战略。

第二阶段：已拥有一定规模的移动支付客户群体,如拥有占手机用户以上的客户,同时已

---

具备完善、稳定的移动支付网络、平台和资金结算系统后,应将战略目标转到大市场战略上。在维系客户群体关系,保持移动支付市场占有份额的同时,进行大规模的移动支付业务推广,全面铺设移动支付终端设备,提高移动支付的快捷性、方便性和安全性,增强客户粘性,进而扩大移动支付客户群体,形成规模经济效益,实现大市场战略。

#### 模型的创新点:

1. 对公交移动支付用户市场数据进行深度分析,并将关联规则数据挖掘技术用于客户群体特征分析,提出了第三方支付平台在移动支付目标市场的发展策略和建议。
2. 将价值网络植入移动支付模式中,提出了一种基于价值网络的,涵盖手续费,服务费,沉淀金利息,广告费等收入模式的商业盈利模型。
3. 通过概率知识和神经网络预测模型进行相关参数的模拟,不是简单的系数相乘,更加科学和准确。

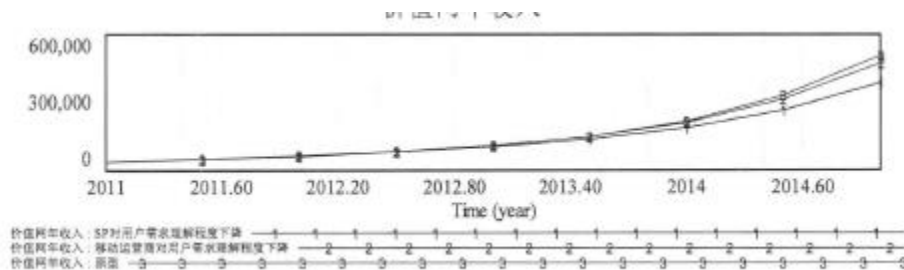
#### 模型的局限:

1. 由于模型使用的数据跨越时间较短,收集的数据有限,基于历史数据得到的结论需要在今后的实践中进一步修正和完善。
2. 商业盈利模型中,可以进一步考虑时间,空间上的影响,对模型进行进一步修正。

#### 移动支付商业盈利模式的系统模拟和有效性检验:

常用的系统动力学模型有效性检验方法有直观检验、灵敏度分析、运行检验和历史检验四种。其中,历史检验是最主要的一种方法,本模型的检验主要是基于历史检验。该检验是通过已建立的模型的模拟结果与历史数据进行比较,如果模型的仿真结果与移动支付市场业务发展数据一致或比较比较吻合,且二者经验数据化趋势一致,说明建立的模型可以正确地对实际的问题进行描述,便可以在此模型基础上对移动支付系统行为进行分析,即假定现有条件不变,改变模型中的主要指标,对模型进行仿真,找出影响系统行为的主要因素,也可以用于预测。如果模拟结果与现实数据有一定的差距,应调节系统预先设定的参数,直到模拟结果与现实结果相吻合

关键要素变动对模型的影响因素分析 一张类似的图就好了



## 模型的拓展:

### 博弈论

亦名“对策论”、“赛局理论”,属于应用数学的一个分支,主要考虑游戏中主体的预测行为和实际行为,并研究其优化策略,即被公式化的激励结构间的相互作用,以及具有斗争或竞争性质现象,在经济学、国际关系、政治学、军事战略和其他很多学科都有广泛的应用

### 博弈论在移动支付模式研究的应用

在移动支付产业链上,各个参与方都希望在这个新型业务市场上分得利益,而移动支付是一个横跨通信、金融、以及服务等多个行业的复杂程度较高的产业,其产业链上的任何一环节都不可能独霸整个市场,因此产业链的分工协作就成了如何做大“蛋糕”和如何分“蛋糕”的关键。那么移动支付应该由移动通信运营商主导,或者由金融机构主导,或者第三方支付机构主导,还是采用其他的运营模式呢在此问题上,博弈理论即可用于研究我国移动支付产业链主要参与方的决策博弈关系和博弈均衡,以分析适合我国当前移动支付发展的运营模式。其中,产业链各参与方之间不同的博弈结果将带来不同的市场战略决策。

## 八、参考文献

- [1]姜启源. 谢金星. 数学模型. [M]. 高等教育出版社. 2015-3. 15-37. 89-125
- [2]储昌木. 沈长春. 数学建模及其应用. 成都: 西南交通大学出版社, 2015.
- [3]余运伟, 李平, 陈林. 校园移动支付的用户采纳行为——基于电子科技大学喜付用户的实证分析[J/OL]. 电子科技大学学报(社科版), 2018(02): 1-10[2018-04-08]. [https://doi.org/10.14071/j.1008-8105\(2017\)-0013](https://doi.org/10.14071/j.1008-8105(2017)-0013).

- 
- [4]万紫芊. 浅谈商业银行移动支付产品的发展及营销——以中国建设银行“龙支付”为例[J/OL]. 中国战略新兴产业[2018-04-08]. <https://doi.org/10.19474/j.cnki.10-1156/f.004121>.
- [5]吕文蓉, 胡曼. 第三方移动支付用户感知与顾客品牌忠诚度相关性研究[J]. 中国集体经济, 2018(09):59-60.
- [6]林朝阳. 社交电子商务模式盈利困境及突破——以蘑菇街、美丽说为例[J/OL]. 商业经济研究, 2018(06):70-72[2018-04-08]. <http://kns.cnki.net/kcms/detail/10.1286.F.20180323.1545.042.html>.
- [7]黄水金. 基于 HCE 和 NFC 技术的交通一卡通手机虚拟卡应用[J/OL]. 交通世界, 2018(02):10-11[2018-04-08]. <https://doi.org/10.16248/j.cnki.11-3723/u.2018.02.002>.
- [8]刘新华. 网购平台盈利模式研究——以聚美优品为例[J]. 财会通讯, 2018(08):52-55.
- [9]张建梅. 移动支付下的银行业务[J/OL]. 现代营销(下旬刊), 2018(01):18[2018-04-08]. <http://kns.cnki.net/kcms/detail/22.1256.F.20180319.1053.022.html>.
- [10]袁圣兰, 封思贤. 移动支付操作风险测度问题研究[J]. 软科学, 2018(03):128-133.
- [11]仲洪丽, 邓少灵. 移动社交对用户支付行为的影响因素研究[J]. 电子商务, 2018(03):35-37.
- [12]李晖. 代表专家谈移动支付[N]. 中国经营报, 2018-03-12(B08).
- [13]王文佳. 社会化电子商务背景下企业盈利模式创新研究[J]. 商业经济研究, 2018(05):91-93.
- [14]吴存钱. 杭州公交推行一体化公交移动支付 “互联网+公交” 激发公共出行新活力[J]. 城市公共交通, 2018(02):22-24.
- [15]潘志宏, 万智萍, 谢海明. 跨平台框架下基于移动感知的智慧公交应用研究[J/OL]. 计算机工程与应用:1-7[2018-04-08]. <http://kns.cnki.net/kcms/detail/11.2127.TP.20171225.0925.002.html>.
- [16]罗在文. 基于移动智能平台的公交查询系统[D]. 电子科技大学, 2015.

## 九、附录

### 1. 源代码

问题一:

```
load data1
```

---

```

gps1=data1(:,1);gps=data1(:,2);price=data1(:,3);

shenzhen=zeros(835,2);

shenpri=zeros(835,1);

n=1;

for i=1:835

    if (gps(i)>22.45&&gps(i)<22.8667)&&(gps1(i)>113.7666&&gps1(i)<114.6166)

        shenzhen(n,1:2)=[gps1(i) gps(i)];

        shenpri(n)=price(i);

        n=n+1;

    end

end

shenzhen(n:835,:)=[];

shenpri(n:835,:)=[];

sj=[114.05 22.53;shenzhen] ;

sj=sj*pi/180; %C¶È»³É»i¶È

d=zeros(n); %dÀË%00óó³ òÊ%»

for i=1:n

    for j=i+1:n

        d(i,j)=6370*acos(cos(sj(i,1)-

sj(j,1))*cos(sj(i,2))*cos(sj(j,2))+sin(sj(i,2))*sin(sj(j,2)));

    end

end

d=d+d';

d1=d(2:n,1);

sj=[gps1 gps;VarName2 VarName1] ;

n=length(sj);

sj=sj*pi/180; %C¶È»³É»i¶È

d=zeros(n); %dÀË%00óó³ òÊ%»

for i=1:n-1

    for j=i+1:n

```

---

```

d(i,j)=6370*acos(cos(sj(i,1)-
sj(j,1))*cos(sj(i,2))*cos(sj(j,2))+sin(sj(i,2))*sin(sj(j,2)));

end

end

d=d+d' ;

```

## 问题二

```

load data

logis=zeros(835,8);

for i=1:835
    for k=1:6
        if (gps1(i)>j1(k)&&gps1(i)<j2(k))&&(gps(i)>w1(k)&&gps(i)<w2(k))
            logis(i,1:4)=[y1(i) price(i) z(i) xinpian(i)];
            sj=[VarName9(k) VarName10(k);gps1(i) gps(i)] ;
            sj=sj*pi/180;
            d=zeros(2);
            for e=1
                for j=e+1:2
                    d(e,j)=6370*acos(cos(sj(e,1)-
sj(j,1))*cos(sj(e,2))*cos(sj(j,2))+sin(sj(e,2))*sin(sj(j,2)));
                end
            end
            d=d+d' ;

            logis(i,5:8)=[d(1,2) x5(k) x6(k) x7(k)];
        end
    end
end
end

```

---

整数规划

```
function [f,g]=mengte(x)
```

```
gc=-0.0450891603153425*x+2;
```

```
gc1=1+exp(-gc);
```

```
f=sum(1./gc1);
```

```
sum(x.*(1./gc1));
```

```
g=[sum(x.*(1./gc1))-1.036450000000000e+04];
```

```
rand('state',sum(clock)); %³ õÊ¼»¯Ëæ»úËý • ¢ÉúÆ÷
```

```
p0=0;
```

```
x0=0;
```

```
tic %¼ÆÊ±¿ªÊ¼
```

```
for i=1:10^6
```

```
    x=randi([50,80],172,1); %² úÉúÒ»ÐÐÎáÁÐµÃÇø¼ä[1,99]ÉÏµÃËæ»úÕûËý
```

```
    [f,g]=mengte(x);
```

```
    if all(g<=0)
```

```
        if p0<f
```

```
            x0=x; p0=f; %¼ÇÂ¼Îµ±Ç° ½Ï°ÃµÃ½â
```

```
        end
```

```
    end
```

```
end
```

```
x0,p0
```

```
toc %¼ÆÊ±±½ûÊø
```

---

### 问题三聚类

```
load shenzhen

sj=shenzhen;

y=pdist(sj,'cityblock'); %C6aμĀĀ½Ā½ĐĐİðÁ¿¼ăμĀ¼ø¶ŎÖμ³ðĀë

yc=squareform(y) %±ă»»³É¼ăĀë•½Ŏó

z=linkage(y) %²úÉúμÈ¼¶¼ŬĀðÊ÷

[h,t]=dendrogram(z) %»-¼ŬĀðÍ¼

T=cluster(z,'maxclust',30) %° Ñ¶Ŏİó»®•Ŏ³É3Āð

for i=1:30

    tm=find(T==i); %C6μŬiĀðμĀ¶Ŏİó

    tm=reshape(tm,1,length(tm)); %±ă³ÉĐĐİðÁ¿

    fprintf('μŬ¼dĀðμĀŎĐ%s\n',i,int2str(tm)); %İŎÊ¼•ŎĀð¼đ¹û

end
```

### TSP 问题

```
function main

a=zeros(6);

a(1,2)=3.24194319461787;a(1,3)=0.761001584363399;a(1,4)=4.40580395384149;a(1,5)=2.917772

15455841;a(1,6)=2.42412892771473;

a(2,3)=3.53496336242321;a(2,4)=4.49949019735733;a(2,5)=2.25541214706232;a(2,6)=2.1430767

7321315; a(3,4)=3.81971930868623;a(3,5)=2.67331563248018;a(3,6)=2.21517857351134;

a(4,5)=2.26945169385222;a(4,6)=2.56382016305362;

a(5,6)=0.496631934154043; a=a+a'; L=size(a,1);

c=[5 1:4 6 5]; %Ñ;È;³ðÊ¼È!

[circle,long]=modifycircle(a,L,c) %μ÷ÓĀĪĀĀæĐĐ,ĀĒ! μĀ×Ó°ĒÊý

%*****

%ŎŎĪĀĪªĐĐ,ĀĒ! μĀ×Ó°ĒÊý

%*****

function [circle,long]=modifycircle(a,L,c);

for k=1:L
```



---

```

flag=0;    %İË³ ö±êô%

for m=1:L-2    %mîª Ēä • ¨ ÖÐµĀi

for n=m+2:L    %nîª Ēä • ¨ ÖÐµĀj

    if a(c(m),c(n))+a(c(m+1),c(n+1))<a(c(m),c(m+1))+a(c(n),c(n+1))

        c(m+1:n)=c(n:-1:m+1); flag=flag+1; %ÐĒ, Āò»´ î£¬±êô%401

    end

end

end

if flag==0    %Ô»İö±ßô² Ā»ÓÐÐĒ, Ā, %î • µ»Ø

    long=0;    %È|³ ƧµĀ³ ôÊ%Öµ

    for i=1:L

        long=long+a(c(i),c(i+1)); %Çó, ĀĀ%È| µĀ³ Ƨ¶È

    end

    circle=c;    % • µ»ØÐĒ, ĀÈ|

    return

end

end

```

#### 问题四

```

load data3

load data

sj=data3;

n=length(sj);

sj=sj*pi/180; %¼Ç¶È»³ Ê»i¶È

d=zeros(n); %¼äĀë¼00óó³ òÊ%»

for i=1:n-1

    for j=i+1:n

        d(i,j)=6370*acos(cos(sj(i,1)-

sj(j,1))*cos(sj(i,2))*cos(sj(j,2))+sin(sj(i,2))*sin(sj(j,2)));

    end

end

end

```

---

```
d=d+d' ;
h=zeros(2066,1877) ;
z=zeros(2066,1) ;
D=zeros(2066,1) ;
xinpin=zeros(2066,1) ;
for i=1:2066
    xinzong=0;
    for j=2067:3943
        if d(i,j)<2
            k=j-2066;
            h(i,k)=d(i,j) ;
            xinzong=xinzong+xin(k) ;
            z(i)=z(i)+1;
        end
    end
    if (z(i)~=0)
        xinpin(i)=xinzong/z(i) ;
    end
end
```

```
he=zeros(835,1) ;
hepin=zeros(835,1) ;
for i=1:835
    for j=1:1877
        if h(i,j)~=0
            he(i)=he(i)+xin(j) ;
        end
    end
    hepin(i)=he(i)/z(i) ;
end
```

```
shenzhen=zeros(5776,2) ;
```

---

```

shenpri=zeros(577,1);
shenxinpin=zeros(577,1);
shenz=zeros(577,1);
j=1;
for i=1:2066
    if (GPS(i)>22.45&&GPS(i)<22.8667)&&(GPS1(i)>113.7666&&GPS1(i)<114.6166)
        shenzhen(j,1:2)=[GPS(i) GPS1(i)];
%        shenpri(j)=price(i);
        shenxinpin(j)=xinpin(i);
        shenz(j)=z(i);
        j=j+1;
    end
end

logis=zeros(2066,8);

for i=1:2066
    for k=1:6 %Ôë¶ÔÊiÔÚ'Àí
        if (GPS1(i)>j1(k)&&GPS1(i)<j2(k))&&(GPS(i)>w1(k)&&GPS(i)<w2(k))
%            logis(i,1:4)=[y1(i) price(i) z(i) xinpin(i)];
            sj=[VarName9(k) VarName10(k);GPS1(i) GPS(i)] ;
            sj=sj*pi/180; %½Ç¶Ê»-3É»i¶Ê
            d=zeros(2); %¼ÀË¼0Ôóð³ òÊ¼»-
            for e=1
                for j=e+1:2
                    d(e,j)=6370*acos(cos(sj(e,1)-
sj(j,1))*cos(sj(e,2))*cos(sj(j,2))+sin(sj(e,2))*sin(sj(j,2)));
                end
            end
            d=d+d';
            logis(i,5:8)=[d(1,2) x5(k) x6(k) x7(k)];
        end
    end
end

```

---

end

end