

DVD 在线租赁方案

摘要

本论文通过对 DVD 租赁问题的抽象简化，建立了一个明确的、完整的数学模型，分别用线性规划模型与递归算法对 DVD 分配进行优化，设计出一个使得会员满意度较高的分配方案。

针对问题一，我们利用正态分布等概率论知识建立了一个较为完整而又简单的数学模型

$$1.6d_j \geq Q_j \times \omega_j$$

在问题四中，关于问题一我们利用货币流通模型和信息源的最大熵原理，建立起关于需求量的另一种模型：

$$d_j = \frac{\omega_j m_j}{n}$$

针对问题二，考虑到当天会员的偏爱度加和可以用来衡量满意度，我们提出了在两种不同网站运行模式下的分配方案，方案一运用线性规划很好的解决了分配问题，使得满意度最大。方案二运用递归算法较好的解决了此问题，并且跟方案一结果相当吻合。

针对问题三，我们利用问题一和问题二建立的数学模型组合起来解决了当前 DVD 的购买和分发问题，并阐述了动态规划的方法。

针对问题四，我们从 DVD 需求预测角度出发，利用本模型特点合理的引入了传染病模型，有效的解决了该预测问题。

最后，我们对模型的科学性和现实性进行了阐述，并得到了对模型的整体评价，及急需改进之处。

关键字： 正态分布 线性规划 递归算法 货币流通 最大熵原理 SIS 模型
SIR 模型

问题重述

随着信息时代的到来，网络成为人们生活中越来越不可或缺的元素之一。许多网站利用其强大的资源和知名度，面向其会员群提供日益专业化和便捷化的服务。DVD 的在线租赁就是一种可行的服务。顾客缴纳一定数量的月费成为会员，订购 DVD 租赁服务。会员对哪些 DVD 有兴趣，只要在线提交订单，网站就会通过快递的方式尽可能满足要求。会员提交的订单包括多张 DVD，这些 DVD 是基于其偏爱程度排序的。网站会根据手头现有的 DVD 数量和会员的订单进行分发。每个会员每个月租赁次数不得超过 2 次，每次获得 3 张 DVD。会员看完 3 张 DVD 之后，只需要将 DVD 放进网站提供的信封里寄回（邮费由网站承担），就可以继续下次租赁。

- 1) 网站正准备购买一些新的 DVD，通过问卷调查 1000 个会员，得到了愿意观看这些 DVD 的人数（表 1 给出了其中 5 种 DVD 的数据）。此外，历史数据显示，60%的会员每月租赁 DVD 两次，而另外的 40%只租一次。假设网站现有 10 万个会员，对表 1 中的每种 DVD 来说，应该至少准备多少张，才能保证希望看到该 DVD 的会员中至少 50%在一个月之内能够看到该 DVD？如果要求保证在三个月内至少 95%的会员能够看到该 DVD 呢？
- 2) 表 2 中列出了网站手上 100 种 DVD 的现有张数和当前需要处理的 1000 位会员的在线订单，如何对这些 DVD 进行分配，才能使会员获得最大的满意度？请具体列出前 30 位会员（即 C0001~C0030）分别获得哪些 DVD。
- 3) 继续考虑表 2，并假设表 2 中 DVD 的现有数量全部为 0。如何决定每种 DVD 的购买量，以及如何对这些 DVD 进行分配，才能使一个月之内 95%的会员得到他想看的 DVD，并且满意度最大？
- 4) 在 DVD 的需求预测、购买和分配中还有哪些重要问题值得研究？明确提出问题，并尝试建立相应的数学模型。

表 1 对 1000 个会员调查的部分结果

DVD 名称	DVD1	DVD2	DVD3	DVD4	DVD5
愿意观看的人数	200	100	50	25	10

表 2 现有 DVD 张数和当前需要处理的会员的在线订单（表格格式示例）

DVD 编号		D001	D002	D003	D004	...
DVD 现有数量		10	40	15	20	...
会员 在线 订单	C0001	6	0	0	0	...
	C0002	0	0	0	0	...
	C0003	0	0	0	3	...
	C0004	0	0	0	0	...

注：D001~D100 表示 100 种 DVD，C0001~C1000 表示 1000 个会员，会员的在线订单用数字 1, 2, ... 表示，数字越小表示会员的偏爱程度越高，数字 0 表示对应的 DVD 当前不在会员的在线订单中。

问题分析

1) 对网站运行模式的理解

模式一：网站的会员可随时提交订单，每天网站将订单信息及现有 DVD 信息汇总，然后进行 DVD 分配，不在会员订单中的 DVD 也有可能分配给会员。不会将任何客户订单推迟到第二天处理，即不存在会员等待时间。

模式二：网站的会员可随时提交订单，每天网站按当天顾客提交订单的时间先后对其排序，并将其订单信息及现有 DVD 信息汇总，然后进行 DVD 分配，使排在前面的会员尽量被满足。公司只能向会员提供其订单中的 DVD，即会员不喜欢的 DVD 公司不能给会员（这是模式二与模式一的根本区别）。当天没有被满足的会员及没有被分派的 DVD 汇入第二天，并在第二天排序时将这些会员排在最前面。

2) 对 DVD 需求问题的分析

由于各种因素的随机性，DVD 的流通问题是很复杂的，我们考虑需求最大情况，即 DVD 需求上限，以及 DVD 月流通次数最少。

3) 对满意度的理解

我们认为满意度受会员等待时间和偏爱度影响，等待时间为从提交订单到订单中 DVD 被邮寄的时间间隔。对于方案一无等待时间，会员的满意度被会员的偏爱度唯一确定。对于方案二，等待时间的影响很小，所以我们也以当天被满足的会员的偏爱度加和来衡量满意度。

问题假设

- 1) 对于调查问卷中 1000 人和所有会员对 DVD 的选择情况是同分布的，也就是说这 1000 人对 DVD 的选择情况代表了所有会员的选择倾向。
- 2) 对于所有的会员提交订单的时间是随机的。
- 3) 无论租赁一次或两次，会员必须在 30 天之内归还所借的 DVD。
- 4) 一个月内借两次的会员这两次借的 DVD 种类不会重复。
- 5) 公司收到订单时不知道此会员在一个月内会借一次或两次。

模型分析与求解

问题 (1):

变量假设:

m : 网站现有会员人数.

p_j : 第 j 种 DVD 被选中的概率

q_j : 第 j 种 DVD 没被选中的概率

α ：每月租赁 DVD 一次的会员的比例

β ：每月租赁 DVD 两次的会员的比例

d_j ：第 j 种 DVD 应准备的数量

ω_j ：一个月内对第 j 种 DVD 有需求的会员得到满足的比例.

n ：DVD 每月可用次数的数学期望

Q_j ：某月内对第 j 种 DVD 需求的人数上限.

考虑到会员租碟的实际情况，表 1 中给出的选择某种 DVD 的人数可以认为是某月选择该 DVD 人数的数学期望，每月实际选择该 DVD 的人数会在其周围波动，我们认为对第 j 种碟片的总需求可以用正态分布 $N(mp_j, mp_j q_j)$ 近似（此处 $m=100000$ ），可以算出第 j 种 DVD 的需求人数上限 Q_j （在一定置信区间下，这里我们选取 0.95），只要在租借过程中满足上限 Q_j 的一定人数比例 ω_j （50%）即可，假设第 j 种 DVD 购买 d_j 张。

我们考虑需要 DVD 最多的情况：借一次的会员在一个月的最后一天归还，借两次的会员在一个月的最后一天第二次归还，那么对于一张碟来说借一次的会员使得它流通了一次，而借两次的会员使得它流通了两次，这相当于该 DVD 的每月可用次数 n 为 $\alpha \times 1 \times d_j + \beta \times 2 \times d_j$ ，对于本题目来说， $\alpha=0.4$ ， $\beta=0.6$ ，即 $1.6 d_j$ ，要求一个月至少 50% 有需求的会员能得到满足，即

$$1.6 d_j \geq Q_j \times \omega_j \quad (1)$$

求出 d_j 的最小值.

用 Matlab 求得置信度为 0.95 下的上限值分别为：

$$Q_1 = 20209$$

$$Q_2 = 10157$$

$$Q_3 = 5114$$

$$Q_4 = 2582$$

$$Q_5 = 1052$$

带入公式(1)解得：

$$d_1 = 6316$$

$$d_2 = 3174$$

$$d_3 = 1598$$

$$d_4 = 807$$

$$d_5 = 329$$

对于三个月的情况，想当于一个月情况的三次累积，三个月的 DVD 流通次数是一个月的 3

倍，上限 Q_j 值不便，所得公式为：

$$4.8d_j \geq Q_j \times 0.95 \quad (2)$$

带入数据计算得：

$$d_1 = 4000$$

$$d_2 = 2011$$

$$d_3 = 1013$$

$$d_4 = 512$$

$$d_5 = 209$$

问题 (2)：

模式一规划模型

对于问题中的分配问题，我们可以建立一个 0-1 分配矩阵 q_{ij} ， $q_{ij}=1$ 表示将 j 号 DVD 分配给 i 号会员， $q_{ij}=0$ 表示不分配。而根据模型分析中对满意度的理解，满意度由会员的偏爱度加和来衡量。显然该分配问题可以转化成为使得已分配的会员的偏爱度之和最小的规划问题。由于附件表格中的 0 项表示用户对该 DVD 无任何偏爱度，而题目里有表格中数据越小，偏爱程度越大，显然所有的 0 项将把我们规划模型导向错误的方向，为了方便模型的建立，我们对附录表进行了一定的技术处理—将所有的 0 项改成 11。

变量假设：

x_{ij} ：编号为 i 的会员对第 j 号 DVD 的偏爱度

y_j ：编号为 j 的 DVD 的当天库存

根据上述分析，考虑到会员一次有且仅得到 3 张 DVD，DVD 有数量限制的约束条件，我们可以建立如下的线性规划模型。

$$\begin{aligned} & \min \sum_{i=1}^{1000} \sum_{j=1}^{100} q_{ij} \times x_{ij} \\ & s.t. \begin{cases} q_{ij} = 0, 1 \quad (i = 1, 2 \dots 1000, j = 1, 2 \dots 100) \\ \sum_{i=1}^{1000} q_{ij} \leq y_j \quad (j = 1, 2 \dots 100) \\ \sum_{j=1}^{100} q_{ij} = 3 \quad (i = 1, 2 \dots 1000) \end{cases} \end{aligned}$$

用 Lingo 求解得 偏爱度之和最小为 8254。

对于 30 位会员（C0001~C0030）获得的 DVD 情况如下表所示：

会员编号	DVD 编号		
C0001	DVD8	DVD41	DVD98
C0002	DVD6	DVD44	DVD62
C0003	DVD32	DVD50	DVD80
C0004	DVD7	DVD18	DVD41
C0005	DVD66	DVD68	DVD11
C0006	DVD19	DVD53	DVD66
C0007	DVD26	DVD66	DVD81
C0008	DVD31	DVD35	DVD71
C0009	DVD53	DVD78	DVD100
C0010	DVD41	DVD55	DVD85
C0011	DVD59	DVD63	DVD66
C0012	DVD31	DVD2	DVD41
C0013	DVD21	DVD78	DVD96
C0014	DVD52	DVD23	DVD89
C0015	DVD13	DVD85	DVD52
C0016	DVD84	DVD97	DVD10
C0017	DVD67	DVD47	DVD51
C0018	DVD41	DVD60	DVD78
C0019	DVD84	DVD86	DVD66
C0020	DVD45	DVD89	DVD61
C0021	DVD53	DVD45	DVD50
C0022	DVD57	DVD55	DVD38
C0023	DVD95	DVD29	DVD81
C0024	DVD76	DVD41	DVD37
C0025	DVD9	DVD69	DVD94
C0026	DVD22	DVD68	DVD95
C0027	DVD58	DVD78	DVD50
C0028	DVD8	DVD34	DVD37
C0029	DVD55	DVD30	DVD26
C0030	DVD62	DVD37	DVD98

对该模型的评价

显然, 该模型尽可能的使得所有分配会员偏爱度数值之和最小, 根据我们对满意度的理解, 该模型得到的解肯定是满足满意度最大的最优解。

模式二递归算法

由于模式一下的规划模型使得不在会员订单中的 DVD 也有可能分配给会员, 而模式二严格规定公司只能向会员提供其订单中的 DVD, 而这个限制条件用规划模型是无法实

现的, 所以我们考虑用算法实现.

根据我们对满意度的理解及该模式下的网站的经营方案, 如果要使得会员的满意度达到最大, 应该尽量满足会员订单之中偏爱度较小的碟片请求. 根据我们的假设, 会员提交订单的时间是随机的, 当网站在处理会员的订单请求时, 必然存在一个处理次序问题, 我们假设处理次序仅与用户发出订单的次序有关.

假设当前的订单已经按会员的请求顺序排列出来, 会员被编号为 1-1000, 为了使得总体的满意度最大, 我们所要做的是尽量满足会员订单中偏爱度较小的请求, 同时尽量满足编号靠前的会员的请求. 根据以上分析我们引入下面的算法解决分配方案:

1), 提取表 2 中的数据, 经过一定的加工处理生成需求矩阵 $Metr_{ij}$, 偏爱矩阵 $UserMetr_{mn}$. 矩阵 $Metr_{ij}$ 反映了会员 i 对 DVD j 的偏爱程度, 0 表示会员对该 DVD 无需求, 1, 2, ..., n 分别表示会员对 DVD 的偏爱程度, 数值越小偏爱程度越高. 矩阵 $UserMetr_{mn}$ 反映了会员 m 偏爱度为 n 的对应的 DVD 编号.

2), 根据用户的偏爱度从低到高进行排序, 偏爱度相同的按照会员编号进行排序, 构成序列 A_i, B_i . 其中 A_i 记录了序列第 i 项对应的会员编号, B_i 记录了序列第 i 项对应会员满足一定偏爱的 DVD 编号.

3), 构建 DVD 库存序列 $Store_i, (i=1, 2, \dots, 100)$, 表示 DVD i 当前的库存数. 构建会员定购盘数序列 $Check_i$, 表示会员 j 当前还需满足的 DVD 盘数, 显然分配前有 $Count_j = 3, (j=1, 2, \dots, 1000)$.

4), 构建与 A_i 序列等长的序列 $Check_i$, 然后对 A_i, B_i 中的数据逐个检索, 假设 A_i 中第 i 项中会员的编号为 k , B_i 中第 i 项中 DVD 编号为 l . 如果 $Count(k) > 0$ 并且 $Store(l) > 0$, 则说明会员 k 希望得到 l 号 DVD, 已分配给会员 k 的光碟数不足三张, 而 l 号 DVD 有库存, 所以可以把 l 号 DVD 分配给会员 k ,

令 $Check(i) = 1, Count(k) = Count(k) - 1, Store(l) = Store(l) - 1$.

否则, 如果 $Count(k) = 0$, 说明当前的会员 k 已经分配的三张光碟, 如果 $Store(l) = 0$, 说明会员 k 希望分配的 l 号 DVD 已经分完, 显然这两种情况都不满足分配条件, 故令 $Check(i) = 0$. 显然 $Check_i$ 中 0, 1 分布反映了 DVD 的分配情况.

5), 统计 $Check_i$ 中的所有的 1 对应的会员编号及对应的 DVD 编号, 由于分配时仅考虑的时会员对某一张 DVD 的定购请求是否满足而没有考虑用户请求盘数为三的要求, 故上述分配方案可能导致某些用户请求订单上的一张, 或两张盘得到满足, 显然对于这些用户来说分配并未完成, 故统计过程中仅可以记录已经分配三张碟片的会员编号, 及其对应的 DVD 编号.

6), 上述分配完成后剩余部分未分配的碟片, 同时有一些会员未分配到 DVD, 或者仅分配到部分 DVD. 显然这些仅用这些剩余碟片无法满足剩余用户的分配需求. 但是如果把分配给会员的部分碟片回收起来, 然后对这些会员重新分配仍然可以满足部分会员的分配需求. 故我们把剩余的碟片以及分配给用户的部分碟片统一回收, 然后根据已经处理的会员, DVD 情况生成新的需求矩阵 $Metr_{ij}$, 偏爱矩阵 $UserMetr_{mm}$, 然后重复上述步骤就可以对这些剩余的碟片重新分配.

7), 重复步骤 6) 可以不断的对剩余碟片进行重新分配. 当一次分配过程中得到分配的会员个数 $M < 1$ 时, 说明当前统一回收的 DVD 已经无法满足任何剩余会员的分配请求. 这时, 上述循环步骤终止.

8), 对于已完成分配的会员编号 i , 及其对应的 DVD 编号 j , 将其在需求矩阵 $Metr_{ij}$ 中对应的偏爱度 $Metr(i, j)$ 清零.

根据上述算法可以得出每次分配过程中的会员编号以及对应的 DVD 编号, 以及整个分配过程中得到分配的会员编号及对应的 DVD 编号, 未得到满足的会员编号及其对应的 DVD 编号.

本分配方案从尽量满足会员订单中偏爱度较小的请求, 及编号靠前的会员的请求出发, 由于算法开始时创建的 A_i, B_i 序列是按照偏好数值从高到低排列的, 对于偏好数值相同的情况, 我们保证会员编号靠前的会员排在编号靠后的会员之前, 从而尽可能的保证了编号靠前的会员的最大的满意度. 本算法对剩余碟片及分配给用户的部分碟片统一回收, 然后重新分配, 从而保证了尽可能多的会员的订单请求得到满足, 如果不采用递归算法, 即仅分配一次得到分配的用户数为 792(个), 但是采用递归之后得到完全分配的用户数为 915(个), 显然绝大多数的会员得到了满足. 分配结束后会有部分会员的订单请求没有得到满足, 对于本模型来说, 有 85 名会员没有在当天得到需要的碟片, 这样显然会影响这些会员的满意度, 但是这些会员的编号一般比较靠后. 对整体的满意度影响很小. 同时, 对于这部分会员, 可以直接将他们放到第二天订单序列的最前部. 从而保证他们的订单请求在第二天先被满足. 这样, 对于所有会员来说, 绝大多数会在订单发出当天得到满足, 而剩余的会员也基本在两天之内得到满足. 综上所述, 本模型可以使得会员得到较大满意度.

根据上述分配方案可以得到前 30 位会员 (C0001~C0030) 获得的 DVD 情况如下表所示:

会员编号	DVD 编号		
C0001	DVD8	DVD82	DVD98
C0002	当天订单请求无法满足		
C0003	DVD80	DVD50	DVD4
C0004	DVD7	DVD18	DVD41
C0005	DVD66	DVD68	DVD11
C0006	DVD19	DVD53	DVD16
C0007	DVD81	DVD8	DVD26
C0008	DVD71	DVD99	DVD31
C0009	DVD53	DVD100	DVD78
C0010	DVD60	DVD55	DVD85

C0011	DVD59	DVD63	DVD19
C0012	DVD31	DVD2	DVD7
C0013	DVD96	DVD78	DVD21
C0014	DVD52	DVD23	DVD89
C0015	DVD13	DVD85	DVD66
C0016	DVD84	DVD97	DVD55
C0017	DVD67	DVD47	DVD51
C0018	DVD41	DVD60	DVD78
C0019	DVD84	DVD86	DVD66
C0020	DVD45	DVD89	DVD61
C0021	DVD53	DVD45	DVD2
C0022	DVD57	DVD55	DVD38
C0023	DVD95	DVD29	DVD81
C0024	DVD76	DVD41	DVD37
C0025	DVD9	DVD69	DVD81
C0026	DVD22	DVD68	DVD95
C0027	DVD58	DVD22	DVD50
C0028	当天订单请求无法满足		
C0029	DVD55	DVD30	DVD44
C0030	DVD62	DVD37	DVD98

对该算法的评价

这种算法的特点是每次递归的过程保证最优,但以我们现有的数学知识无法论证它最终得到的结果是全局近似最优还是局部最优.对于模拟退火算法和遗传算法,模拟退火中以一定的概率接受恶化解以及遗传算法中的交换和变异使他们避免了陷入局部最优的境地,但这个算法中似乎没有这样的因子使其避免陷入局部最优.但这种算法的计算速度是相当快的,并且同步给出一个明确的分配方案,这使其优点所在.我们考虑可由遗传算法去验证这种算法的解,但这个体系的规模过于庞大,遗传算法运算需要花费相当长的时间,最终被我们放弃.

问题（3）：

我们考虑利用问题一与问题二建立的模型来处理问题三。

1. 我们首先考虑当前分配。先将表二统计成类似表一的表格,我们的原则是,不论偏爱度,只要该会员选择了这张DVD,我们就说他愿意观看这张DVD. 下面的表格就是由表二统计出来的这100张DVD每张愿意观看的人数(按DVD序号排列):

84	92	87	99	78	87	87	100	93	90
95	97	85	102	84	94	102	91	100	116
96	101	109	93	89	101	87	83	97	97
100	87	91	82	109	97	91	94	87	87
119	104	93	90	106	94	94	88	91	94
107	91	98	92	97	99	108	77	85	103
94	103	105	108	98	105	90	96	105	101

95	106	85	82	90	86	88	99	82	98
99	77	72	84	90	78	95	73	94	98
107	94	93	90	102	78	95	101	80	86

我们仍认为需求量满足正态分布，从而由第一问的方法求出各种 DVD 需求人数的上限（置信度仍为 0.95）：

97	107	102	115	92	102	101	115	108	105
109	112	100	118	98	109	118	106	116	133
111	117	125	108	104	117	102	97	112	112
116	102	106	96	125	112	106	109	102	102
135	120	108	105	122	109	109	103	105	109
123	106	114	107	112	115	124	91	100	119
109	119	121	124	114	121	105	111	121	117
110	122	100	96	105	101	103	115	96	114
115	90	85	98	104	92	110	87	109	114
123	109	108	105	118	92	110	116	94	101

我们仍假设 $\alpha = 0.4$ ， $\beta = 0.6$ 。这样利用公式（1）求出当前各种 DVD 的需求量（这里

$\omega_j = 95\%$ ）：

58	64	61	68	55	61	60	68	64	62	65	66	59	70	58	65	70	63	69	79	66
69	74	64	62	69	61	58	66	66	69	61	63	57	74	66	63	65	61	61	80	71
64	62	72	65	65	61	62	65	73	63	68	64	66	68	74	54	59	71	65	71	72
74	68	72	62	66	72	69	65	72	59	57	62	60	61	68	57	68	68	53	50	58
62	55	65	52	65	68	73	65	64	62	70	55	65	69	56	60					

利用 Lingo 求得模式一的当前分配方案，这里给出前 30 位会员的分配：

会员编号	DVD 编号		
C0001	DVD8	DVD82	DVD98
C0002	DVD6	DVD42	DVD44
C0003	DVD4	DVD50	DVD80
C0004	DVD7	DVD18	DVD41
C0005	DVD11	DVD66	DVD68
C0006	DVD16	DVD19	DVD53
C0007	DVD8	DVD26	DVD81
C0008	DVD15	DVD71	DVD99
C0009	DVD53	DVD78	DVD100
C0010	DVD55	DVD60	DVD85
C0011	DVD59	DVD63	DVD19
C0012	DVD31	DVD2	DVD7
C0013	DVD21	DVD78	DVD96
C0014	DVD52	DVD23	DVD43
C0015	DVD13	DVD85	DVD88
C0016	DVD84	DVD97	DVD6
C0017	DVD67	DVD47	DVD51
C0018	DVD41	DVD60	DVD78

C0019	DVD84	DVD86	DVD67
C0020	DVD45	DVD89	DVD61
C0021	DVD53	DVD45	DVD65
C0022	DVD57	DVD55	DVD38
C0023	DVD95	DVD29	DVD81
C0024	DVD76	DVD41	DVD79
C0025	DVD9	DVD69	DVD94
C0026	DVD22	DVD68	DVD95
C0027	DVD58	DVD42	DVD22
C0028	DVD8	DVD34	DVD82
C0029	DVD55	DVD30	DVD44
C0030	DVD62	DVD37	DVD1

偏爱度之和最小为 6000（即每个人都满意度最高的三张 DVD）。
利用算法求得模式二的当前分配方案，给出前 30 位会员的分配：

会员编号	DVD 编号		
C0001	DVD8	DVD82	DVD98
C0002	DVD6	DVD42	DVD44
C0003	DVD4	DVD50	DVD80
C0004	DVD7	DVD18	DVD41
C0005	DVD11	DVD66	DVD68
C0006	DVD16	DVD19	DVD53
C0007	DVD8	DVD26	DVD81
C0008	DVD15	DVD71	DVD99
C0009	DVD53	DVD78	DVD100
C0010	DVD55	DVD60	DVD85
C0011	DVD59	DVD63	DVD19
C0012	DVD31	DVD2	DVD7
C0013	DVD21	DVD78	DVD96
C0014	DVD52	DVD23	DVD43
C0015	DVD13	DVD85	DVD88
C0016	DVD84	DVD97	DVD6
C0017	DVD67	DVD47	DVD51
C0018	DVD41	DVD60	DVD78
C0019	DVD84	DVD86	DVD67
C0020	DVD45	DVD89	DVD61
C0021	DVD53	DVD45	DVD65
C0022	DVD57	DVD55	DVD38
C0023	DVD95	DVD29	DVD81
C0024	DVD76	DVD41	DVD79
C0025	DVD9	DVD69	DVD94
C0026	DVD22	DVD68	DVD95
C0027	DVD58	DVD42	DVD22
C0028	DVD8	DVD34	DVD82

C0029	DVD55	DVD30	DVD44
C0030	DVD62	DVD37	DVD1

得到分配方案与模式一相同。

2. 再考虑多阶段问题。其实很简单，网站每天只需按同样的方法处理当前分配即可，只不过要多考虑两个问题。

第一，考虑归还。在求得当天各种 DVD 的需求量后，当天购买量为需求量减去归还量。

第二，考虑第二次租赁。只需在订单中将该会员第一次已租赁过的 DVD 对应的偏爱度清零即可（这一步模式二已经自动解决）。

问题(4)：

针对问题一，我们还有另外一种思路。对于这一问，如果从会员的角度出发，则要考虑很多的因素，如归还时间 租借次数 不同会员的租期等问题，而这些因素的随机性都是非常强的，这样建出的模型会很复杂，因此，我们巧妙的转换了思路，从某种碟片出发，碟片总共被出租的次数就是它能够满足的会员人数，因为我们假定借两次的会员两次借的碟片种类不同，这显然是符合实际情况的，不论是第一次借到还是第二次借到都满足了该会员的要求，换句话说，就是某一种碟片集合与会员是一一对应的关系，所以我们把人数问题转换成了碟片出租次数的问题，碟片的出租次数等于实际的碟片数与该类碟片的每月平均出租次数(即出租频率)之积。

就此，我们考虑到碟片的出租过程随机性非常强，这恰恰类似货币的流通，随即性强，货币的流通频率越高其代表的价值就越高，因此，我们用货币流通模型对其进行分析。

问题假设：

在原问题假设的基础上，我们增加一条假设：考虑到 DVD 邮寄和观看时间，假定 DVD 最短 t 天归还，

变量假设：

m_j ：100000 会员中愿意观看第 j 种 DVD 的人数

α ：每月租赁 DVD 一次的会员的比例

β ：每月租赁 DVD 两次的会员的比例

d_j ：第 j 种 DVD 应准备的数量

ω_j ：一个月内能够看到第 j 种 DVD 的会员的比例

n_1 ：DVD 在每月租赁一次的会员手中的流通频率的数学期望。

n_2 ：DVD 在每月租赁两次的会员手中的流通频率的数学期望。

n ：DVD 每月流通次数的数学期望。

t ：DVD 的最短归还周期 ($t=5$)。

得到第 j 种 DVD 应准备量的表达式：

$$d_j = \frac{\omega_j m_j}{n} \quad (3)$$

依据信息源的最大熵原理：若信源输出的幅度被限定在 $[a, b]$ 区域内，则当输出信号的概率密度是均匀分布时，信源具有最大熵。

由于影响DVD流通周期的因素很多(即信息量很大)，依据最大熵原理，我们认为此模型是均匀分布是合理的。

因为我们假设 DVD 流通周期满足均匀分布. 这样对于在月租赁一次的会员手中流通的 DVD, 它的流通周期满足 t 天至 30 天的均匀分布, 数学期望为 $(t+30)/2=(5+30)/2=17.5$ 天, 这样 $n_1 = 30/17.5 = 1.7143$. 对于在一个月租赁两次的会员手中流通的 DVD, 两次被租的周期是对等的, 那么这两个阶段必然是同分布的, 并且都是均匀分布的, 整个租借周期分成两个等份, 每个分周期 15 天, 第一个分周期满足 t 天至 15 天的均匀分布, 同样, 第二个分周期也满足这个分布, 那么, 这个阶段的数学期望是 $(t+15)/2=(5+15)/2=10$ 天, 这样 $n_2 = 30/10 = 3$. 再根据每月租一次和租两次的人数比例求得

$$n = \alpha \cdot n_1 + \beta \cdot n_2 = 1.7143 \times 0.4 + 3 \times 0.6 = 2.4857.$$

要满足 50%的会员的需求, $m_1 = 20000$ $m_2 = 10000$ $m_3 = 5000$ $m_4 = 2500$ $m_5 = 1000$

根据公式(3)求得:

$$d_1 = 0.5 \times 20000 / 2.4857 \approx 4023,$$

$$d_2 = 0.5 \times 10000 / 2.4857 \approx 2012,$$

$$d_3 = 0.5 \times 5000 / 2.4857 \approx 1006$$

$$d_4 = 0.5 \times 2500 / 2.4857 \approx 503$$

$$d_5 = 0.5 \times 1000 / 2.4857 \approx 201$$

对于三个月内满足 95%的会员看到 DVD 的情况, 很显然三个月的情况等同于一个月的情况的三次累积过程, 即对应的碟片出租频率的数学期望是上述模型的数学期望的 3 倍, 同理得:

$$d_1 = 0.95 \times 20000 / (2.4857 \times 3) \approx 2548$$

$$d_2 = 0.95 \times 10000 / (2.4857 \times 3) \approx 1274,$$

$$d_3 = 0.95 \times 5000 / (2.4857 \times 3) \approx 637$$

$$d_4 = 0.95 \times 2500 / (2.4857 \times 3) \approx 319$$

$$d_5 = 0.95 \times 1000 / (2.4857 \times 3) \approx 127$$

2. 原文中并没有考虑关于网站运行中的预测问题。实际上对于上市后的 DVD，愿意观看它的人数随着时间的推移是符合某种曲线的，网站完全可以对其提前进行预测，而不必通过当前订单才能知道这一阶段的大体需求。我们考虑某种 DVD 在这个大系统中的租赁问题其实类似于各类传染病模型，观看即是被“感染”，成为病人，然后成为治愈者，未观看者属于“易感人群”。观看者观看后会对此 DVD 进行评价，评价将作为信息反馈给未观看者，从而很大程度上决定观看者是否租赁此 DVD，这是“感染”的过程。受欢迎程度不同的 DVD 满足不同的传染病模型。

(1) 受欢迎程度高的 DVD 比较符合 SIS 模型，SIS 模型的特点是病人被治愈后变成健康者，健康者还可以再被感染成为病人，也就是看过该 DVD 的会员因为认为此 DVD 比较好看而可能再租赁此 DVD。

查阅有关传染病模型的书籍，设在时刻 t 易感染者与已感染者在总人数中所占比例分别记作 $s(t)$ 和 $i(t)$ 。每个病人每天平均感染人数是常数 λ 。每天被治愈的病人占病人总数的比例为常数 μ ，在 DVD 租赁问题中这个治愈率为 100%，病人治愈后成为仍可被感染的易感染者。定义 $\sigma = \lambda/\mu$ ，在这里 $\sigma = \lambda/100\% = \lambda$ 。书籍中有关于 i 的微分方程模型：

$$\frac{di}{dt} = -\lambda i \left[i - \left(1 - \frac{1}{\sigma} \right) \right] = -\lambda i \left[i - \left(1 - \frac{1}{\lambda} \right) \right]$$

这也是没看过此 DVD 与看过并准备再看的人所占比例的数学模型。网站可以据此模型对这类会员随时间的变化作出预测，从而对阶段购买量作出相应控制，因此这个模型有很强的可行性。（具体公式推导及参数设定参见各种有关传染病模型的书籍）

(2) 有些 DVD 很不好看，看完后没有人会再看它了，这种情况类似于 SIR 模型，该模型特点是病人治愈后不会再被感染，既非易感染者，也非易感染者，他们已经退出传染系统。

设人群分为易感染者，已感染者，病愈免疫的移出者，在总人数中比例分别为 $s(t)$ ， $i(t)$ 和 $r(t)$ 。 λ ， μ 与 σ 定义同上。可建立 SIR 模型：

$$\begin{cases} \frac{di}{dt} = \lambda si - \mu i, i(0) = i_0 \\ \frac{ds}{dt} = -\lambda si, s(0) = s_0 \end{cases}$$

但该方程无法求出解析解，我们做数值计算，设 $\lambda = 1$ ， $\mu = 100\%$ ， $i(0) = 0.02$ ，

$s(0) = 0.98$ ，利用 MATLAB 软件编程：

```
function y=ill(t,x)
a=1;b=1;
y=[a*x(1)*x(2)-b*x(1),-a*x(1)*x(2)]';
```

```
ts=0:50;
x0=[0.02,0.98];
[t,x]=ode45('ill',ts,x0);[t,x]
plot(t,x(:,1),t,x(:,2)),grid,pause
plot(x(:,2),x(:,1),grid,
```

可得到 $i(t)$, $s(t)$ 图像，网站根据此图像便可进行预测。

模型评价与改进

- 1) 在解决第一个问题上，我们在需求量最大的情况下考虑了 DVD 购买问题，虽然简单，但是得到的值偏大，使得网站的经营成本过高。
- 2) 对于问题二，针对第二种模式，我们编制了一个分配算法，该算法递归的对 DVD 资源进行分配，基本上实现了在该网站模式，满意度以偏爱度之和为衡量标准的前提下达到最优，但是该算法无法保证求解的结果是全局最优解，这是该模型最大的局限。
- 3) 对于问题三，我们是在确定了购买量之后再进行分配，使得满意度在这种购买方式中最大。其实购碟的数量和顾客满意度的大小是网站需要权衡的两个因素，网站一方面想要购买金尽可能少的 DVD，一方面又想使会员满意度尽可能大，这样构成一个多目标规划要比当前解法更有说服力。
- 4) 在 DVD 购买和预测方面分别建立了模型，这些模型具有一定的实用价值。

参考文献

- [1] 《信息论理论基础》 贾世楼编著 哈尔滨工业大学出版社
- [2] 《数学模型》(第三版) 姜启源编著 高等教育出版社
- [3] 《概率论与数理统计》 李贤平编著 复旦大学出版社
- [4] 《Handbook Of Monetary Economics》 Benjamin M.Friedman Elsevier Science Press
- [5] 《Mathematical Modeling》 Mark M.Meerschaert Elsevier Science Press

参考网站

- [1] <http://www.es88.com/Article/news/200503/447.html>

附录

matlab 源代码

main.m

```
[usermetr]=findit(2005B) %2005B 是将 excel 中的数据导入 matlab 得到的矩阵
w=0;
```

bound=1,2,3,... %bound 取不同的值将得到不同代得到的分配序列.

```
[userresult,dvdresult,metr,usermetr,w]=getR(2005,usermetr,w,bound); %result 记载了
某一代满足的分配序列
```

```

getR.m
function [userresult,dvdresult,metr,usermetr,w]=getR(metr,usermetr,w,bound)
row=size(metr,1);%1001
col=size(metr,2);%100
manyi=10;
for i=1:row-1
    usercheck(i)=3;%算法中的 Count(i);
end
for i=1:col
    arr(i)=metr(1,i);%算法中的 Store(i);
end
k=0;a1=0;a2=0;
for i=1:manyi
    for j=1:row-1
        k=k+1;

        user(k)=j;
        dvd(k)=usermetr(j,i);

        if dvd(k)>0
            if arr(dvd(k))>0
                a1=1;
            end
        end
        if usercheck(j)>0
            a2=1
        end
        if a1*a2>0
            arr(dvd(k))=arr(dvd(k))-1;
            usercheck(j)=usercheck(j)-1;
            check(k)=1;
            a1=0;
            a2=0;
        else
            check(k)=0;
            a1=0;
            a2=0;
        end
    end
end
end
num=size(check,2);
m=0;
num
for i=1:num

```



```

        if check(i)==1
            m=m+1;
            user1(m)=user(i);
            dvd1(m)=dvd(i);
        end
    end

    for i=1:row-1
        tongji(i)=0;
        usertongjiresult(i)=0;
    end
    for i=1:size(user1,2)
        tongji(user1(i))=tongji(user1(i))+1;
    end
    index=0;
    for i=1:size(user1,2)
        if tongji(user1(i))==3
            usertongjiresult(i)=3;
            index=index+1;
            x=user1(i)
            userresult(index)=user1(i);
            dvdresult(index)=dvd1(i);
        end
    end
    end
    xi=0;
    for i=1:row

        if i==1
            for u=1:row-1
                %tongji
                if tongji(u)==3
                    for v=1:size(user1,2)

                        if user1(v)==u
                            xv=user1(v)
                            xi=xi+1
                            xd=dvd1(v)
                            xm=metr(1,dvd1(v))
                            metr(1,dvd1(v))= metr(i,dvd1(v))-1;
                        end
                    end
                end
            end
        end
        else

```

```

        for j=1:col
            if usertongjiresult(i-1)==3
                metr(i,j)=0;
            end
        end
    end
    end
    end
    %metr(3,:)
    index=0;
    [a,usermetr]=findit(metr);

    if w<bound
        [userresult,dvdresult,metr,usermetr]=getR(metr,usermetr,w,bound);
        userresult(1:size(userresult,2));
    else
        return;
    end
end
end

```

线性规划lingo源代码

```

model:
    sets:
        user/1..1000/;
        dvd/1..100/:t;
        link(user,dvd):p,x;
    endsets
    data:
        x=@file('dvddata.txt');
        t=@file('dvd.txt');
    enddata
    min=@sum(link(i,j):p(i,j)*x(i,j));
    @for(link(i,j):@bin(p));
    @for(user(i):@sum(dvd(j):p(i,j))=3);
    @for(dvd(j):@sum(user(i):p(i,j))<=t(j));

end

```