

流失预警模型在移动行业客户保留中的应用

华院分析技术（上海）有限公司 咨询顾问 张高峰

摘要：本文研究了数据挖掘的流失预警模型在移动通信行业客户关怀与挽留中的构建和应用。通过分析预警模型的“预测效果”、“应用效果”，作者指出，流失预警模型不是用来大幅度降低流失率的，而是用来提高客户关怀与挽留工作有效性的，而且这一工作的投资回报是相当可观的。

关键词：客户流失、预警、模型

一、客户流失严重影响移动公司的收入和利润

客户保持对公司的利润有着惊人的影响，远远超过公司规模、市场份额、单位成本和其它许多通常被认为与竞争优势有关的因素。客户保持率小小的提高能导致利润可观的改善。有关方面对美国 9 个行业进行调查的数据表明，客户保持率增加 5%，行业平均利润增加幅度在 25% — 85%之间。在各行业，客户保持已成为公司成功最至关重要的目标。

因客户流动性较强，移动通信行业的客户保持对利润的影响尤为突出。由于各移动运营商相继取消了入网费，并采取了一些入网激励政策，使得客户的入网门槛大大降低，亦即转网门槛变得极低，客户流失已经成为移动运营商非常头痛的问题。举例来讲，A 地市的移动用户有 100 万，月平均流失率为 5%，假定用户的月均 ARPU 为 80 元，则每月由于用户流失千万的收入减少将是 $100(\text{万}) * 5\% * 80 = 400$ 万元。这还不考虑由于用户流失而造成的用于新用户开拓的各种费用，如广告、渠道等方面所发生的成本。

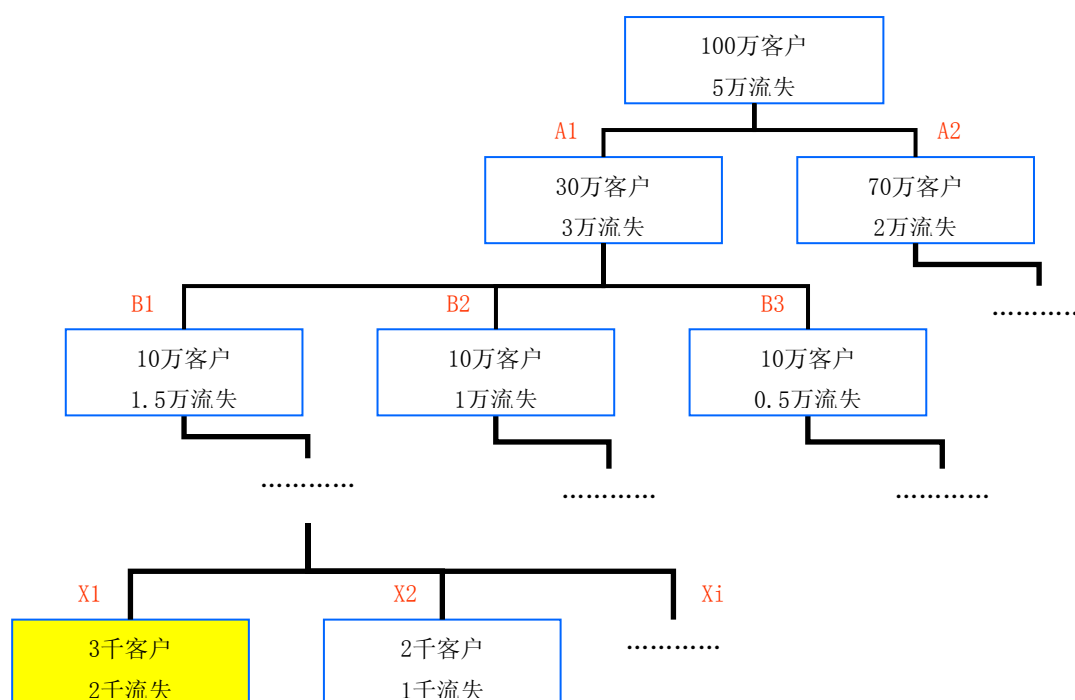
移动运营商的大量实践经验已经证明：客户保持工作的最佳时机是在其未流失时，其原因在于已流失客户的回流阻力大，策反工作成本高且效果差。但对移动公司而言，面对网上规模庞大的客户群，也不可能对每个客户都去做特殊的保持工作。那么，如何预测出将要流失的高价值客户？并了解他/她的需求，及时投入一定的资源，并通过针对性的挽留工作避免其流失，就成了移动运营商非常关注的问题。因为这样不仅可以提高效率，而且可以大大节省成本，获得可观的收益。

二、流失预警模型原理

预测哪些客户将要流失，并不是一件容易的事情。但可喜的是移动运营商已经积累了大量的有关客户的有用信息，如 BOSS 系统的通话行为数据，客户服务系统的客户服务、查询与投诉数据，帐务系统的缴费数据等等。通过对大量的客户数据信息进行分析处理，归纳出具有流失倾向客户的行为特点和规律，建立起数据挖掘模型，并不断地进行实证校验，完全可以在一定程度上预测出客户的流失倾向。

流失预警模型构建的第一步是采用数据挖掘中的聚类、神经网络、决策树等分析技术，通过对数据的探索和分析概括归纳出具有高度流失倾向的用户普遍特征。然后是从当前在网的用户中找出具有类似特征的用户群，通过大量的对比分析，最终利用用户的历史行为数据将不同用户归入不同的群组，并对不同用户群的流失倾向给予评分，形成对客户流失可能性的预测。

以决策树算法为例，图一简单地说明了流失预警模型的工作原理。即通过 A—X 等变量对用户进行分类，最终发现不同方格里的分组用户的流失倾向也不同。如下图黄色方块所示，同时满足 A1、B1、…、X1 等条件的用户的流失概率是 66.7%。



图一、流失预警模型的决策树算法举例

这样，通过构建流失预警模型，就非常准确地发现了具有高流失倾向的用户

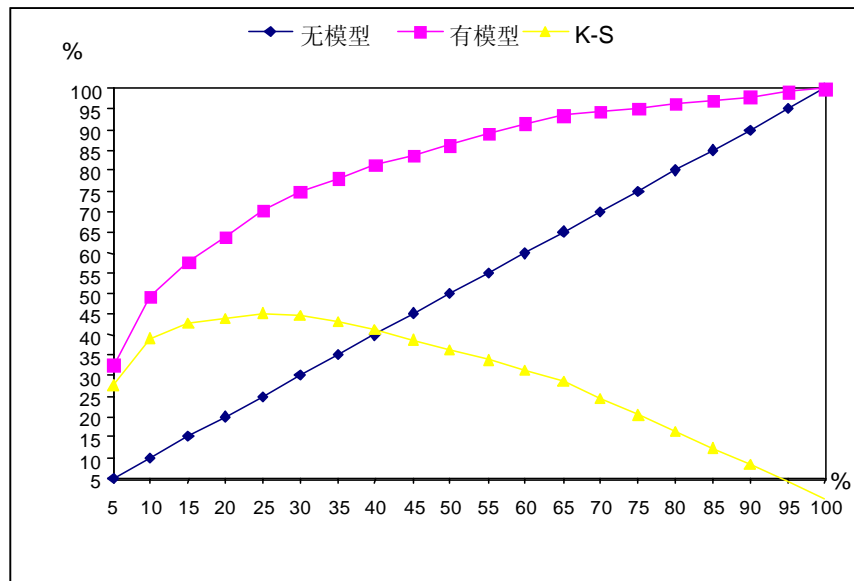
群(如上图中涂黄色的分组用户)。如果将其他用户的行为数据作为模型的输入，对这群用户的流失倾向进行评分，就可以发现具有较高流失可能的用户群体。

三、流失预警模型的“预测效果”

流失预警模型的预测效果可以用图简单表示，如图二所示：

下图的横坐标表示开展客户流失前挽留活动的目标活动用户数占总用户数的百分比，纵坐标表示选出的目标活动用户中，所包含的实际流失用户数占全部流失用户数的百分比。

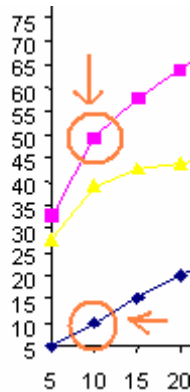
图中蓝色的曲线表示在没有模型指导的情况下随机抽取用户进行挽留活动可以获得的效果。这条曲线非常容易理解：目标活动用户数越多，能够关怀到的流失的用户量也越大。即，随机抽取一定比例的用户为我们挽留活动的目标用户，那么所能关怀到的实际流失用户也占全部实际流失用户相同的百分比。



图二、有模型、无模型目标用户查全率的比较

类似的，红色的曲线表示有流失预警模型指导的情况下，选取一定量目标用户可以达到的效果。可以发现，只要不是对全部用户做活动，有模型指导比无模型指导的目标活动用户中总能包括更多的实际流失用户。例如，我们选取流失倾向位于前 10% 用户作为目标活动用户，可以包括所有实际流失人数的 50%。这是没有模型指导情况下的 5 倍，这个“5 倍”就是使用预警模型在选取前 10% 的高危用户时可以达到的预测效果。其实，选取的目标用户量不同，使用模型可

以达到的效果也是变化的：通过模型选取的目标用户数越小，模型可以达到的效果越高(如较三，仅选取 5%的用户，模型可以达到的效果超过 6)；相反，通过模型选取的目标用户数越多，模型可以达到的效果越低；当针对全部用户都进行流失前的关怀活动时，有模型与没有模型的效果是完全一样的，因为此时的活动对象都包括了全部可能流失的用户。



图三、模型预测效果说明

四、流失预警模型所表现的“应用效果”

那么，有了这样一个流失预警模型，或者说在这个模型的指导下对高流失可能的用户开展活动，下个月的用户流失率是不是就会“显然地降低”？这样的看法是不正确的，因为流失预警模型只是揭示了“什么样的用户更可能会流失”这样一个客观规律。

下面再通过一个例子来说明在使用流失预警模型后，为什么客户流失率往往得不到大幅度降低。

依然假设某地市移动公司 A，它当前总共有 100 万在网用户，用户的月流失率平均为 5%。其中，高价值用户 20 万，月平均流失 0.6 万，流失率为 3%；一般价值用户 80 万，月平均流失 4.4 万，流失率为 5.5%。为了有效地控制用户流失率，A 公司构建了分别针对高价值用户和一般价值用户流失预警模型，它们的效果可以用上面的图示来表示，即如果选取最具流失倾向的前 10% 用户作为目标活动用户，可以包括所有实际流失人数的 50%。

由于 A 公司的大客户经理数量有限，负责客户服务的员工中可以用来进行客

户流失挽留的也不过 4、5 个人。A 公司的决定是：根据流失预警模型的高危的客户名单，对高价值用户先开展一对一的针对性挽留，而对一般价值用户，则是希望通过普遍的营销政策进行挽留。

A 公司按照流失预警模型给出的流失倾向评分从高到低，依次选择这次活动的目标用户。即从 20 万的高价值用户中选取了前 10%的预警高危用户作为目标用户，即 2 万人。然后，这 2 万人被依次安排到了这个月的工作日程中。紧接着，A 公司的大客户经理和客户服务人员在这个“挽留月”中非常尽职地根据工作安排展开一对一的挽留工作。希望在月末的流失率统计中会有一个令人满意的结果。

但是，当“挽留月”结束后，第二个月对流失率再次进行统计的时候，A 公司惊讶的发现，“挽留月”的流失率和没有开展活动的月份相差无几。症结到底在什么地方呢？

我们首先列举出流失预警模型在应用中会发生耗散的地方：

1. 在全部客户中，仅对高价值用户进行针对性挽留，假设比例 **a**
2. 目标活动用户选取时，仅对预警的高危用户进行挽留，假设选取比例为 **b**
3. 客户挽留过程中，存在目标活动用户的的接触成功率，假设 **c**
4. 客户接触成功的用户中又存在挽留成功率问题，假设 **d**

根据各地市移动公司的实际情况及其在开展客户挽留方面的经验，a、b、c、d 都是可以估算的。基本上，高价值用户的比例为全部用户比例的 20%；选取的目标活动用户数又往往是高价值用户的 10%；实际挽留过程中，成功接触率一般为 50%；对于接触成功的用户中有流失倾向的人的挽留成功率一般为 30%。

进一步地假设预警流失用户在接触到和接触不到的用户中均匀分布，我们可以计算出 A 公司在“挽留月”的活动实际成功挽留下来的用户数量¹为 450 人。

表一、有模型指导针对目标用户关怀与挽留时可以挽留到的用户

总用户数	高价值用户比例	预警高危用户比例	高价值用户平均流失率	模型提升效果	接触成功率	接触到用户的成功挽留率	可以挽留的高价值用户数
1,000,000	20%	10%	3%	5	50%	30%	450

¹ 可以成功挽留的用户数=总用户 * 高价值用户比例 * 预警高危用户比例 * 高价值用户的平均流失率 * 模型提升效果 * 接触成功率 * 接触到的用户的成功挽留率

在这种情况下，总体流失率 = $(50000 - 450) / 1000000 = 4.955\%$ ，和不做活动的 5% 确实相差无几！这时候就会有人问了，如果真是这样的结果，到底还要不要做流失预警模型了？这是一个非常现实的问题，因为如果要考虑构建流失预警模型的投资回报，这个问题就一定要清楚地测算。

五、如何正确认识流失预警模型的“实际效果”及其意义

实际上，回答流失预警模型要不要做，首先要回答的问题是客户关怀要不要做？也就是说，做流失预警模型的目的是为了降低流失率呢？还是为了提高关怀与挽留工作的有效性？

这是一个不能回避的问题，但它的答案也是早就存在的。那就是：如果是单纯是为了大幅度降低客户流失率，流失预警模型所起到的效果是相对较小的。为什么呢？原因也很简单，数据挖掘的流失模型其实是一种方法论，它并不能直接带来客户流失率的降低。打个比方，就如同病人看病，再先进的医疗设备也只能帮助病人查毛病，而不能帮助病人养好病。数据挖掘的流失预警在移动行业的客户挽留中所起的作用也只能是帮助移动公司找到流失可能性比较高的客户群，而不能直接导致流失率的下降。

再考虑到实际造成移动客户流失的原因也是多方面的。如一部分是移动用户在网内互转造成的，原因是移动自己推出了更便宜的套餐，但原来的用户又不能带号进行变换；还有一部分是由于移动本身的入网费用设置及信用政策等造成的，客户一旦发生欠费，还不如弃卡而重新入网经济省钱，在信用体系不健全的地方(全国的绝大部分地区)，这种现象也导致了客户流失的发生等等。而这些，都不通过关怀与挽留所能解决的。因此，一定程度上，客户流失率的降低是管理的问题，而不是技术的问题。

归根结底，流失预警模型的构建目的应该是为了提高关怀与挽留工作的有效性，而不是所谓的大幅度降低用户流失率。

前文也提到过，许多移动运营商的大量实践经验已经证明：客户保持工作的最佳时机是在其未流失时。事实上，所有的移动公司也都在努力实践这件事，即通过不断地投入来做好客户关怀与挽留工作。如在客户服务中心中再专门增加人手，专业通过外呼进行客户关怀与挽留；建立卓越的大客户经理队伍，对高价值

用户和集团用户定期开展回访等。但毕竟资源是有限的，客户服务的专业外呼人员和大客户经理的数量不可能无限制地发展下去。以当前的普遍情况来看，一个有着 100 万移动用户的地市级移动公司，其专业的外呼人员也不过是 4、5 个人，大客户经理也不过 100 来人。那么，如何在人手有限的情况下提高客户关怀与挽留工作的效率，如何能在较少的客户接触成本上关怀到更多实际将会流失的用户，就必须借助于数据挖掘的流失预警模型了。

接着前文的举例，A 公司每月可以达到的客户接触为 2 万人次，他们也把要接触的对象集中在了高价值用户上。如果有数据挖掘的流失预警模型指导，每月可以成功挽留的用户数为 450 人，具体计算过程见前文表（一）。但如果没有模型，每月可以成功挽留的用户数为 90 人，具体计算过程见表（二）。

表二、无模型指导针对目标用户关怀与挽留时可以挽留到的用户

总用户数	高价值用户比例	预警高危用户比例	高价值用户平均离网率	无模型	接触成功率	接触到用户的成功挽留率	可以挽留的高价值用户数
1,000,000	20%	10%	3%	1	50%	30%	90

简单比较就可以发现，基于完全相同的人员投入、完全相同的接触成功率、完全相同的挽留成功率，有模型指导的挽留比没有模型指导的挽留在每月的工作中成功地多挽留了 $450-90=360$ 个用户。假设这些挽留成功的用户可以延长在网时间半年(相当保守的估计)，其 ARPU 为 120 元(活动对象为高价值用户)，则每月由于挽留效率的提高可以获得的额外收益将为： $360*120*6=259,200$ 元。一年延续下来，年总收益的增加为 $259,200*12=3,110,400$ 。

再进行更为保守的估计，如果在模型的指导下选取前 10% 的高危用户作为目标用户时，模型的“预测效果”只能达到 3(绝大部分模型的效果都可以达得到)，那么上面估算的多出的收入再减少一半的话，每年依然可以多增加一百五十多万元。这相对于构建模型而进行的投资来算，回报依然很大！

由此推算，对于某些拥有近千万移动用户的大省来讲，通达数据挖掘流失预警模型的构建及应用，每年通过模型可以增加的纯收入也将超过一千五百万元，这是多么可观的收入啊！

六、结论及建议

通过前文的分析及论证，结论可以概括如下：第一、数据挖掘流失预警模型的构建和应用在帮助移动用户大幅度降低客户流失率上帮助不大；第二、通过流失预警模型的构建和应用，可以大幅度地提高客户关怀与挽留工作的有效性，在资源(人员、时间等)投入相同的情况下，可以获得非常大的收益回报。

最后，鉴于流失预警模型的有效应用效果离不开移动公司非常强的客户关怀与挽留方面的执行能力，对于尚没有把流失前的客户关怀作为一项常规工作抓的公司来讲，在构建客户流失预警模型上不能着急。但对于已经构建了流失预警模型的公司，也决不能在模型做出以后，只是看它预测的结果准还是不准，而不去做实际的关怀工作，这样就等同于看病时查出了毛病所在，却不去医治养病，到头来却怪罪说数据挖掘模型并没有带来实际的收益。将是非常可悲可笑的！

作者简介：

张高峰，男，1977年生，复旦大学管理学院研究生，现任华院分析技术有限公司应用咨询顾问，主要从事数据挖掘在移动通信行业的应用辅导及理论研究。