

## 为什么要在游戏中推荐好友？

游戏朋友之间的互动可以提高用户活跃和付费表现。在海量用户中，为用户推荐什么样的人才能增加好友互动呢？潜在的好友可能是下面几类人，

1. 可能认识的人 真实世界的线下熟人或其他游戏的线上熟人,有互动的基础。
2. 志同道合的人 相同的玩法偏好,相似的付费能力，共同语言，有互动的动机。

所以，游戏好友推荐就是为每个用户找到上面两类人（预测空间），然后科学的对潜在好友排序，使得用户添加好友的可能性最大化。

## 推荐方案

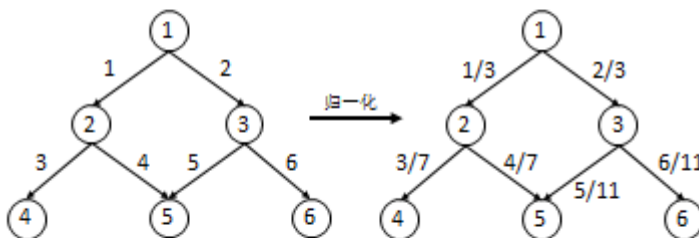
### 方案1：基于相似性的好友推荐方案

目前该方案在**smoba**好友推荐中实现，主要是根据用户人脉相似性（利用社交拓扑结构）和兴趣相似性（利用用户玩法偏好）推荐。

#### 王者荣耀好友推荐

方法：

- 1) 以好友关系做边，好友亲密度做边的权值，构建图网络
- 2) 推荐距自己距离为2, 3的用户
- 3) 推荐人数不够用相似好友填充



- 1) 高优先级推荐高亲密度好友的好友
- 2) 高优先级推荐共同好友个数多的用户

用户1与用户4的关系值为：

$$\frac{1}{3} * \frac{3}{7}$$

用户1与用户5的关系值为：

$$\frac{1}{3} * \frac{4}{7} + \frac{2}{3} * \frac{5}{11}$$

### 方案2：基于机器学习的好友推荐方案

相比于基于相似性的方案 方案，预测空间类似，都是二度好友和相似用户，但是排序时使用了机器学习而不是直接用相似性排序，这样可以利用用户其他关联数据作为特征，增加了整个系统的可扩展性和潜能。具体步骤如下：

1. 预测空间：对每个用户，计算可能认识的人（二度好友）与志同道合的人（LSH），过滤掉已成对的好友，得到潜在好友集合F。
2. 标签：在t时刻，得到F；观察t+1时刻时，F中成为好友设置为1，否则设置为0。
3. 训练：得到用户基础信息和基于网络拓扑结构抽取特征，进行训练与交叉验证，调整线下模型。
4. 预测：与预测空间的逻辑一致，得到预测数据，并预测后续可能成对情况。

5. 效果验证：通过MAP，召回等手段线上验证，并进行A/B test。

## 方案总结

上面两方案的预测空间类似，方案1成本相对较低；方案2由于引入机器学习过程，成本较高，但可塑性可能强于方案1。应用上，如果时间压力较大，可以先低成本快速上线方案1，优化时尝试使用方案2，并进行A/B test。

## 推荐效果评估

### 好友推荐模型评估

- 精度 P, AP, MAP

- $P(k)$ 是截止前k个的准确率 $P(k) = \begin{cases} 0, & \text{第}k\text{个位置没有命中} \\ \frac{h}{k}, & \text{第}k\text{个位置命中, } h\text{是前}k\text{个位置的命中数} \end{cases}$

- $AP = \sum_{i=1}^n P(k) / \min(m, n)$ , m是所有好友数，n是预测的好友数。m是新增用户数，如果m为负数，这个用户的AP=0。

- $MAP = \sum_{i=1}^N AP_i / N$ , AP是每个用户的指标，MAP是AP的均值。

- Top n准确率 预测到正确的好友/提供的用户数

- 召回：预测到正确的好友/真实添加的好友数

- 二元分类器评估指标

- 召回率/准确率

- f值

- AUC/ROC

- Mean Squared Error:  $\frac{1}{n} \sum (l - p)^2$

### 好友推荐对社交关系结构评估

- 网络拓扑结构

- 图直径

- 图的组成：孤立点，小社群和巨片参数（参考雅虎论文(2006)Structure and Evolution of Online Social Networks）

- 孤立点数量

- 小社群数量，平局成员数，星形结构数量

- 巨片点数

- 度分布：均值，方差等

- 网络弹性，计算方法参考fanky之前总结的资料。

- 小世界结构

- 平均距离：是否有缩短

- 平均聚集系数是否增加，任意点i的聚集系数可表示为 $C(i) = \frac{2|\{e_{jk}: e_{jk} \in E, v_j, v_k \in N(v_i)\}|}{k_i(k_i-1)}$

- 游戏表现
  - 活跃时长
  - 付费金额
  - 好友互动次数和频率