2018年11月16日星期五

上午9时4分

经过一周的努力，终于抛出了原始的程序，该程序中虽然试着对每月迟到早退的用户数据进行了统计（按月统计），然而并未用于更新高危用户预测的Risk\_Ratio。

今天的主要任务是查看、分析单纯使用RLF的JSR建模方法。

分析一：基于RLV计算的JSR预测结果直接分析

不考虑用户标记的Risk\_Ratio的自动更新变化

我们先来验证检查场景二跳槽用户的情况，

首先是2010年8月离职的三个用户，需要查看2010年7月的预测结果：

VCF1602,2010-08-20,

预测：VCF1602,14.7263465016,

前100最低：AMC0854,16.561130049

CKP0630,2010-08-26,

预测：CKP0630,2.72629007281,

ZIE0741,2010-08-27,

预测：ZIE0741,12.253115567,

我们继续查看2010-9月份的Insiders2：

SIS0042,2010-09-02,

预测：SIS0042,16.6124371579,

前100最低：KCG1615,22.9780816538

本月：SIS0042,2.90294234691

本月：KCG1615,10.8445969053

TNB1616,2010-09-10,

预测：TNB1616,18.3628686663,

TRC1838,2010-09-15,

预测：TRC1838,17.1635610489,

MDS0680,2010-09-17,

**预测：MDS0680,24.2026684626, 命中第58位**

WDT1634,2010-09-20,

预测：WDT1634,3.87513518858,

OSS1463,2010-09-21,

预测：OSS1463,8.06596705673,

CIF1430,2010-09-23,

预测：CIF1430,5.28141842377,

初步查看了2010年8月与9月两个月的Insiders-2离职情况，发现，其中只有一个用户正确识别到，而其他用户却都排序较靠后。

为了进一步分析，我们查看某几个用户在2010年7月的当月RLV特征

对于CIF1480用户而言，在2010年7月时，没有离职的邮件联系人

我们试着查看下2010年12月份离职的Insiders\_2：

为此，需要分析2010年11月的预测文件

RRS0056,2010-12-10,

预测：RRS0056,32.7934833797,

ICB1354,2010-12-15,

预测：ICB1354,19.254197899,

BYO1846,2010-12-15,

**预测：BYO1846,46.0780045872**

**前100：MGS1944,45.9394562009**

HXP0976,2010-12-20,

HXP0976,19.4707901928,

HMS1658,2010-12-30,

HMS1658,24.9066235757,

HIS1394,2010-12-30,

**预测HIS1394,49.240394831,**

**我们分析下**

**RRS0056与HIS1394的不同**

**再分析RRS0056与MGS1944的不同**

**从中提取Insiders\_2的相似性，提取insiders与一般用户的差异性**

2018年11月16日星期五

下午1时56分

下午准备重新跑一遍CERT5.2所有用户的Logon迟到早退分析程序；

另一方面，自己在高危用户预测模块中设立了每个月分析完毕后，建立预测器后从当前全体用户集合中删除当月离职的用户，避免下个月再次分析这些离职的用户。其实自己读取的是邮件数据，因此如果该月用户离职，下月自然不会有邮件数据，应该同最后的结果影响不大。

2018年11月17日星期六

下午3时58分

周五下午一直在跑程序，终于跑完了一个没有阈值更新的JSR预测版本，然后写了个小程序简单分析三类场景中被预测中的比例，由于场景二是我们最为关心的主客观融合的攻击，因此我们重点来分析下30个Insiders\_2的情况。

在验证数据的时候，突然发现：自己的High\_Risk没有按照从大到小的顺序输出，导致输出错误，赶紧修正！

程序验证：

对于AAB用户而言；

2010-02的Accumulate\_JSR是：

AAB1302,1.14218527305,

2010-03的Current\_JSR是：

AAB1302,1.23029298274

则2010-03的Accumulate\_JSR是：

AAB1302,2.37247825579,

验证通过！

自己的程序应该没有问题，问题仅仅出现在2010-07，很有可能是自己人工修改文件名时重复了，因而，出现了错误。

后续需要重新运行一遍所有的程序，然后再分析结果。

VCF1602,2010-08-20,

* Current\_JSR: VCF1602,3.99814860414, 429
* Accumulate\_JSR: VCF1602,14.7263465016, 428
* The Last High-Risk: KAM0169,12.7762503211

CKP0630,2010-08-26,

ZIE0741,2010-08-27,

SIS0042,2010-09-02,

TNB1616,2010-09-10,

TRC1838,2010-09-15,

MDS0680,2010-09-17,

WDT1634,2010-09-20,

OSS1463,2010-09-21,

CIF1430,2010-09-23,

MCP0611,2010-10-06,

CHP1711,2010-10-13,

GWG0497,2010-10-15,

KSS1005,2010-10-16,

NAH1366,2010-11-17,

RRS0056,2010-12-10,

ICB1354,2010-12-15,

BYO1846,2010-12-15,

HXP0976,2010-12-20,

HMS1658,2010-12-30,

HIS1394,2010-12-30,

LVF1626,2011-01-14,

MGB1235,2011-01-21,

DCC1119,2011-01-26,

SNK1280,2011-02-11,

ITA0159,2011-02-17,

JAL0811,2011-02-25,

OKM1092,2011-04-29,

HSN0675,2011-04-29,

TMT0851,2011-05-11,