2018年10月30日星期二

下午11时31分

在进行了三个基本的实验后，发现自己的假设与CERT数据集并不符合，因此需要重新考虑问题与假设。

自己重新仔细阅读了数据构造的相关论文，发现了其逻辑关系中用户的人格特征与组织结构共同结合影响了用户的人际关系（Relationship），而与被解雇用户的人际关系进一步影响了用户的工作满意度（Job Satisfactory），从而导致用户高倾向的不守时以及跳槽。

我们的目的就是尽量全面、准确地刻画出用户的JS变量，从而事先选定可疑的内部用户。

我们的研究方法是从现实场景假设出发，提出假设模型，然后在数据实验中进行修补。

从本实验开始，关于同一个主题的多个实验以数字进行标识版本，如之前的不同方法与模型分别被标识为【0.1~0.3】，本版本0.4中的多个变体则使用0.4X的形式进一步区分。

实验研究除了验证模型假设，解决问题之外，还要训练自己分析问题、提出假设、验证假设、修改假设的研究能力，这才是“金钥匙”。

1. 我们的假设

依据自己的工作经验分析，既然JS假设受影响于解雇用户的人际关系，那么刻画一个内部用户的人际关系便是首要的问题，摆在我们面前的是三个具体的问题：

1. 给定用户的LDAP、Psychometric以及Email通信统计信息，如何初步建模用户的人际关系？
2. 初步建模出用户的人际关系（即认识、接触的用户范围后），如何进一步明确不同人际联系的重要性？
3. 依据上述人际关系范围与联结重要性的研究分析，如何进一步刻画解雇用户对个体的JS影响？

为了回答上述问题，我们试着进行做出假设，并通过实验验证。

* 假设一
* 与解雇用户的人际关系影响着个体的JS；
* 人际关系（Relationship, RS）可以从两个方面刻画：

1. 反映在组织结构（Organizational Structure, OS），决定了用户日常工作所自然接触的人际范围；我们自然地分析上下结构（领导上下级关系）以及横向关系（同一个最小组织单位内以及第一个父单位下的兄弟组织人员范围），其中横向关系对于CERT5.2而言就是Team以及同一个Department下的不同Teams；
2. 反映在通讯关系（Email Relationship, ER），决定了用户日常交流的人际范围与频度，其中可以使用0.2实验中的邮件分析方法，从信息量的7元特征来进行刻画频度，并将邮件联系人补充到人际关系图中；

* 最终得到了用户的人际关系图（Relationship Graph, RG），该图中节点标识用户ID以及对应的Big5人格特质；边上存在使用四位编码标识OS位置距离的无向边，以及一条标识邮件通讯信息量的有向边；
* 一般而言，计算RG中单独一条联结的重要性，可以依据下述建议的公式：



* 计算用户的JS\_Risk则可以将解雇用户所在边的重要性相加，作为该用户的JS\_Risk

——————————————————————————————————————————————————————————————————————————————

2018年10月31日星期三

上午9时48分

实验4.1

为了进一步认识分析攻击者邮件通讯联系人与离职用户以及OS的关联，进行本实验

本实验选取分别选取三个用户进行分析：

1. BYO1846，一般意义的场景二攻击用户；
2. AAG1136，一个未跳槽的CERT5.2用户；
3. CKP0630，一个通讯Leave特征不明显的场景二攻击用户；

重点分析上述三个用户的邮件联系人的OS分布以及与Leave用户的交叉匹配结果

1. 首先我们考虑了邮件发送与接收时存在多个联系人的情况，比如以用户A的邮件分析为例，如果send邮件，则重新多次统计其发送给的多个邮件联系人；如果receive邮件，则重新多次统计其邮件来自的多个邮件联系人；

根据实验显示，就简单分析BYO1846而言，如果不考虑群发，有联系人426，而如果考虑群发邮件，则有联系人951~

接下来，进行两个统计分析，一个是两种情况下邮件联系人与离职人员匹配的个数；一个则是邮件联系人的OS集合以及相关的通信联系人个数；

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 3 | 1 | 0 | 3 | 4 |
| 3 | 4 | 1 | 7 | 23 |

BYO1846 Leave Contacts 匹配结果。。。 V2

**0 3**

TRC1838

FDS1841

RMB1821

**1 1**

KBC1390

**2 0**

**3 3**

BRM0126

OJC0930

HXP0976

**4 4**

ZVW1475

TCP0380

RPJ1159

DTB0722

BYO1846 Leave Contacts 匹配结果。。。 V4

**0 3**

TRC1838

FDS1841

RMB1821

**1 4**

VCF1602

KBC1390

CTT0639

HIS1394

**2 1**

ZHB1104

**3 7**

LLW0179

BRM0126

OJC0930

XMG1579

HBH0111

HXP0976

EAL1813

**4 23**

WHG1669

ZVW1475

TCP0380

EPG1196

SMS0432

KVF1143

LMW0837

ICB1890

SAF1942

RPJ1159

FKS1696

DTB0722

ILG0879

STH0353

KJG1121

UKM0845

ESP1198

JSB0860

BRG0728

ZJN1492

CHP1711

ACA1126

SCI0778

2018年10月31日星期三

下午3时31分

由于CERT5.2中JS变量决定于用户人际关系中被解雇用户的情况，因此，需要明确出用户的人际关系范畴，并于离职/解雇用户交叉匹配。

在进行我们的实验前，先进行一个小实验以明确统计分析单纯使用离职数据与邮件联系人的交叉匹配数据，并且对于数据进行五个层次的匹配后，采用scale正态分布化，以发现到底是否存在超出阈值的异常。

实验系列一：CERT5.2用户邮件联系人与离职关系比较

实验前提条件：

1. 假设仅分析单个事业部Business中的人际关系，因此将原始用户分析结果分为Part1与Part2两部分；
2. 分别统计每个用户五个层次上离职用户与邮件联系人匹配的数目，然后分列对于各自的事业部用户全体进行标准化，识别出其中数量超出2或3的用户；

经过实验运行，我们得到了事业部1中Insiders2攻击者的匹配数据：

VCF1602 : [ 0.3039974 -0.66991872 -0.85559626 -1.58694021]

CKP0630 : [-1.23752485 -0.66991872 -0.85559626 -1.58694021]

ITA0159 : [ 0.3039974 0.35902829 -0.85559626 -0.23320863]

ICB1354 : [-0.7236841 -0.66991872 -0.85559626 -0.91007442]

RRS0056 : [-1.23752485 -0.66991872 -0.10722306 -0.23320863]

SIS0042 : [-0.7236841 -0.66991872 -0.85559626 -1.24850732]

CIF1430 : [-0.7236841 -0.66991872 -0.85559626 -1.24850732]

HMS1658 : [ 0.3039974 -0.66991872 -0.10722306 -1.24850732]

BYO1846 : [ 0.3039974 0.35902829 -0.85559626 -0.57164153]

KSS1005 : [-1.23752485 0.35902829 -0.10722306 -1.24850732]

GWG0497 : [-0.7236841 0.35902829 0.64115014 -0.23320863]

TRC1838 : [-0.20984335 -0.66991872 -0.85559626 -1.24850732]

TNB1616 : [ 0.81783816 1.3879753 -0.10722306 -1.58694021]

HIS1394 : [ 0.81783816 -0.66991872 -0.10722306 -0.91007442]

MCP0611 : [-0.7236841 -0.66991872 -0.85559626 -1.58694021]

NAH1366 : [-1.23752485 -0.66991872 -0.85559626 -0.23320863]

LVF1626 : [ 1.84551966 4.47481633 -0.10722306 0.10522426]

WDT1634 : [-0.7236841 -0.66991872 -0.85559626 -1.58694021]

HXP0976 : [-1.23752485 -0.66991872 -0.85559626 -0.57164153]

OKM1092 : [ 0.81783816 -0.66991872 -0.10722306 -0.91007442]

上述实验考虑的是用户之间的单向通信，如果也考虑群发/群收邮件的话那么结果可能有所不同：

输出场景二的匹配特征...

VCF1602 : [ 0.25099721 -0.85792285 -0.39626153 -2.0335947 ]

CKP0630 : [-1.26980872 -0.85792285 -1.14293614 -1.87271351]

ITA0159 : [ 0.25099721 0.18685916 -1.14293614 -0.74654517]

ICB1354 : [-0.2559381 -0.85792285 -1.14293614 -0.58566398]

RRS0056 : [-1.26980872 -0.85792285 0.35041309 -0.90742636]

SIS0042 : [-0.76287341 -0.85792285 -1.14293614 -1.39006994]

CIF1430 : [-0.76287341 -0.85792285 -0.76959883 -1.71183232]

HMS1658 : [ 0.25099721 0.18685916 -0.02292422 -1.22918875]

BYO1846 : [ 0.25099721 1.23164116 -0.76959883 -0.90742636]

KSS1005 : [-1.26980872 0.70925016 -0.39626153 -0.74654517]

GWG0497 : [-0.76287341 -0.33553184 -0.02292422 -0.90742636]

TRC1838 : [-0.2559381 0.18685916 -0.76959883 -1.55095113]

TNB1616 : [ 0.75793252 0.70925016 -0.02292422 -1.71183232]

HIS1394 : [ 0.75793252 0.18685916 -0.02292422 -0.90742636]

MCP0611 : [-0.76287341 -0.85792285 -1.14293614 -1.55095113]

NAH1366 : [-0.76287341 -0.85792285 -1.14293614 -1.39006994]

LVF1626 : [ 1.77180314 3.84359618 -0.76959883 0.05786078]

WDT1634 : [-0.76287341 -0.33553184 -0.76959883 -2.0335947 ]

HXP0976 : [-1.26980872 -0.85792285 -1.14293614 -1.39006994]

OKM1092 : [ 1.26486783 -0.85792285 0.72375039 -0.90742636]

———————————————————————————————————————

2018年11月1日星期四

上午10时19分

昨日实验，统计验证邮件联系人（单向+多向）与用户周围离职关系的交叉匹配，发现邮件上看CERT5.2的30个Insiders不具备显著区分性，因而不适合从邮件角度建模分析。

在继续进行下一步分析前，我们先来看看Insiders\_2中的个别用户BYO1846的邮件通信范围，即查看其邮件联系人跨了几个OS单位；

突然想到：

方案一：从论文说明的数据角度，建模用户的relationship，然后关联到laidoff employees

方案二：直接从laidoff employees入手，然后分析目标用户与这些离职用户的关系，从而判断其受影响程度

对于单向通信而言

分析的用户为BYO1846

* Buffy Yolanda Ortiz,BYO1846,Buffy.Yolanda.Ortiz@dtaa.com,Salesman,,1 - Executive,5 - SalesAndMarketing,2 - Sales,5 - RegionalSales,Donna Erin Black
* **1 - Executive,5 - SalesAndMarketing,2 - Sales,5 - RegionalSales**

匹配的其邮件通讯的OS结构为：

BYO1846 单向通信分析完毕...

**1 - Executive,5 - SalesAndMarketing,2 - Sales,5 - RegionalSales**,30, （同团队）

1 - Executive,5 - SalesAndMarketing,**3 - FieldService,3 - RegionalFieldService**,7, （不同部门的不同团队）

2 - Executive,5 - SalesAndMarketing,2 - Sales,3 - RegionalSales,6, （不同事业部，相同职能部，不同部门，不同团队）

2 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,5 - DesktopSoftware,5,

1 - Executive,5 - SalesAndMarketing,2 - Sales,3 - RegionalSales,5,

2 - Executive,5 - SalesAndMarketing,3 - FieldService,3 - RegionalFieldService,5,

1 - Executive,5 - SalesAndMarketing,2 - Sales,1 - RegionalSales,5,

1 - Executive,5 - SalesAndMarketing,2 - Sales,2 - RegionalSales,5,

1 - Executive,2 - ResearchAndEngineering,2 - Research,2 - Lab,4,

2 - Executive,3 - Manufacturing,3 - Assembly,3 - AssemblyDept,4,

2 - Executive,5 - SalesAndMarketing,3 - FieldService,4 - RegionalFieldService,4,

1 - Executive,3 - Manufacturing,3 - Assembly,6 - AssemblyDept,4,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,5 - DesktopSoftware,4,

2 - Executive,5 - SalesAndMarketing,2 - Sales,1 - RegionalSales,4,

2 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,3 - EmbeddedSoftware,4,

1 - Executive,2 - ResearchAndEngineering,2 - Research,4 - Lab,4,

2 - Executive,5 - SalesAndMarketing,2 - Sales,2 - RegionalSales,4,

2 - Executive,1 - Adminstration,6 - Security,3 - ElectronicSecurity,4,

1 - Executive,2 - ResearchAndEngineering,4 - Engineering,1 - Publications,4,

2 - Executive,3 - Manufacturing,3 - Assembly,1 - AssemblyDept,3,

1 - Executive,2 - ResearchAndEngineering,4 - Engineering,4 - MaterialsEngineering,3,

2 - Executive,5 - SalesAndMarketing,2 - Sales,5 - RegionalSales,3,

1 - Executive,2 - ResearchAndEngineering,4 - Engineering,7 - SystemsEngineering,3,

2 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,2 - DesktopSoftware,3,

1 - Executive,5 - SalesAndMarketing,3 - FieldService,2 - RegionalFieldService,3,

1 - Executive,1 - Adminstration,6 - Security,1 - BuildingSecurity,3,

2 - Executive,5 - SalesAndMarketing,3 - FieldService,1 - RegionalFieldService,3,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,3 - EmbeddedSoftware,3,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,4 - WebSoftware,3,

1 - Executive,5 - SalesAndMarketing,3 - FieldService,5 - RegionalFieldService,3,

1 - Executive,2 - ResearchAndEngineering,4 - Engineering,2 - MechanicalEngineering,3,

1 - Executive,3 - Manufacturing,3 - Assembly,2 - AssemblyDept,2,

1 - Executive,2 - ResearchAndEngineering,2 - Research,1 - Lab,2,

2 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,6 - EmbeddedSoftware,2,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,1 - Publications,2,

2 - Executive,3 - Manufacturing,3 - Assembly,5 - AssemblyDept,2,

1 - Executive,5 - SalesAndMarketing,2 - Sales,4 - RegionalSales,2,

1 - Executive,5 - SalesAndMarketing,3 - FieldService,4 - RegionalFieldService,2,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,5 - SystemsEngineering,2,

1 - Executive,3 - Manufacturing,2 - ManufacturingEngineering,,2,

2 - Executive,5 - SalesAndMarketing,2 - Sales,4 - RegionalSales,2,

1 - Executive,3 - Manufacturing,,,2,

1 - Executive,3 - Manufacturing,3 - Assembly,7 - AssemblyDept,2,

2 - Executive,1 - Adminstration,6 - Security,2 - ElectronicSecurity,2,

1 - Executive,1 - Adminstration,6 - Security,2 - ElectronicSecurity,2,

1 - Executive,3 - Manufacturing,1 - IndustrialEngineering,,2,

2 - Executive,3 - Manufacturing,3 - Assembly,4 - AssemblyDept,2,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,7 - WebSoftware,2,

1 - Executive,2 - ResearchAndEngineering,2 - Research,3 - Lab,2,

1 - Executive,5 - SalesAndMarketing,3 - FieldService,1 - RegionalFieldService,2,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,2 - DesktopSoftware,2,

2 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,7 - WebSoftware,2,

2 - Executive,1 - Adminstration,6 - Security,1 - BuildingSecurity,2,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,6 - EmbeddedSoftware,2,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,4 - MaterialsEngineering,2,

1 - Executive,2 - ResearchAndEngineering,2 - Research,5 - Lab,2,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,,1,

2 - Executive,3 - Manufacturing,3 - Assembly,2 - AssemblyDept,1,

1 - Executive,2 - ResearchAndEngineering,4 - Engineering,6 - TestAndEvalualtion,1,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,6 - TestAndEvalualtion,1,

1 - Executive,2 - ResearchAndEngineering,4 - Engineering,8 - TestAndEvalualtion,1,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,1 - SQA,1,

1 - Executive,2 - ResearchAndEngineering,4 - Engineering,5 - SystemsEngineering,1,

2 - Executive,5 - SalesAndMarketing,3 - FieldService,5 - RegionalFieldService,1,

2 - Executive,2 - ResearchAndEngineering,2 - Research,1 - Lab,1,

1 - Executive,2 - ResearchAndEngineering,2 - Research,6 - Lab,1,

2 - Executive,6 - PurchasingAndContracts,2 - Pricing,,1,

2 - Executive,2 - ResearchAndEngineering,2 - Research,5 - Lab,1,

2 - Executive,3 - Manufacturing,1 - IndustrialEngineering,,1,

1 - Executive,2 - ResearchAndEngineering,2 - Research,,1,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,8 - TestAndEvalualtion,1,

2 - Executive,5 - SalesAndMarketing,3 - FieldService,2 - RegionalFieldService,1,

1 - Executive,4 - Finance,,,1,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,7 - SystemsEngineering,1,

2 - Executive,6 - PurchasingAndContracts,,,1,

2 - Executive,5 - SalesAndMarketing,1 - MarketAnalysis,,1,

2 - Executive,1 - Adminstration,1 - Personnel,,1,

1 - Executive,1 - Adminstration,1 - Personnel,,1,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,,1,

2 - Executive,2 - ResearchAndEngineering,2 - Research,,1,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,2 - MechanicalEngineering,1,

1 - Executive,1 - Adminstration,6 - Security,,1,

2 - Executive,2 - ResearchAndEngineering,2 - Research,2 - Lab,1,

2 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,1 - SQA,1,

2 - Executive,6 - PurchasingAndContracts,3 - Contracts,,1,

1 - Executive,1 - Adminstration,6 - Security,3 - ElectronicSecurity,1,

2 - Executive,4 - Finance,2 - Payroll,,1,

如果对于BYO1846而言，不再考虑单纯收到邮件触发的联系人，则有：（68个）

BYO1846,1 - Executive,5 - SalesAndMarketing,2 - Sales,5 - RegionalSales,

**1 - Executive,5 - SalesAndMarketing,2 - Sales,5 - RegionalSales**,30,1037.0, （同一个团队中通信人30个，总共发送1vs1邮件1037封）

1 - Executive,5 - SalesAndMarketing,3 - FieldService,3 - RegionalFieldService,6,6.0,

2 - Executive,3 - Manufacturing,3 - Assembly,3 - AssemblyDept,4,4.0,

2 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,5 - DesktopSoftware,4,4.0,

2 - Executive,5 - SalesAndMarketing,3 - FieldService,4 - RegionalFieldService,4,4.0,

1 - Executive,3 - Manufacturing,3 - Assembly,6 - AssemblyDept,4,4.0,

2 - Executive,5 - SalesAndMarketing,3 - FieldService,3 - RegionalFieldService,4,4.0,

2 - Executive,3 - Manufacturing,3 - Assembly,1 - AssemblyDept,3,3.0,

2 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,2 - DesktopSoftware,3,3.0,

2 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,3 - EmbeddedSoftware,3,3.0,

1 - Executive,3 - Manufacturing,3 - Assembly,2 - AssemblyDept,2,2.0,

1 - Executive,2 - ResearchAndEngineering,2 - Research,2 - Lab,2,2.0,

2 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,6 - EmbeddedSoftware,2,2.0,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,1 - Publications,2,2.0,

2 - Executive,3 - Manufacturing,3 - Assembly,5 - AssemblyDept,2,2.0,

1 - Executive,5 - SalesAndMarketing,2 - Sales,3 - RegionalSales,2,2.0,

1 - Executive,2 - ResearchAndEngineering,4 - Engineering,4 - MaterialsEngineering,2,2.0,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,5 - DesktopSoftware,2,2.0,

1 - Executive,5 - SalesAndMarketing,3 - FieldService,4 - RegionalFieldService,2,2.0,

1 - Executive,5 - SalesAndMarketing,2 - Sales,1 - RegionalSales,2,2.0,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,5 - SystemsEngineering,2,2.0,

1 - Executive,5 - SalesAndMarketing,3 - FieldService,2 - RegionalFieldService,2,2.0,

1 - Executive,3 - Manufacturing,3 - Assembly,7 - AssemblyDept,2,2.0,

2 - Executive,3 - Manufacturing,3 - Assembly,4 - AssemblyDept,2,2.0,

2 - Executive,5 - SalesAndMarketing,3 - FieldService,1 - RegionalFieldService,2,2.0,

1 - Executive,5 - SalesAndMarketing,3 - FieldService,5 - RegionalFieldService,2,2.0,

1 - Executive,5 - SalesAndMarketing,3 - FieldService,1 - RegionalFieldService,2,2.0,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,2 - DesktopSoftware,2,2.0,

2 - Executive,1 - Adminstration,6 - Security,3 - ElectronicSecurity,2,2.0,

1 - Executive,2 - ResearchAndEngineering,2 - Research,1 - Lab,1,1.0,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,,1,1.0,

2 - Executive,3 - Manufacturing,3 - Assembly,2 - AssemblyDept,1,1.0,

1 - Executive,2 - ResearchAndEngineering,4 - Engineering,6 - TestAndEvalualtion,1,1.0,

1 - Executive,5 - SalesAndMarketing,2 - Sales,4 - RegionalSales,1,1.0,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,6 - TestAndEvalualtion,1,1.0,

2 - Executive,5 - SalesAndMarketing,2 - Sales,5 - RegionalSales,1,1.0,

1 - Executive,2 - ResearchAndEngineering,4 - Engineering,8 - TestAndEvalualtion,1,1.0,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,1 - SQA,1,1.0,

1 - Executive,2 - ResearchAndEngineering,4 - Engineering,5 - SystemsEngineering,1,1.0,

1 - Executive,2 - ResearchAndEngineering,4 - Engineering,7 - SystemsEngineering,1,1.0,

1 - Executive,5 - SalesAndMarketing,2 - Sales,2 - RegionalSales,1,1.0,

1 - Executive,3 - Manufacturing,2 - ManufacturingEngineering,,1,1.0,

2 - Executive,5 - SalesAndMarketing,2 - Sales,1 - RegionalSales,1,1.0,

2 - Executive,5 - SalesAndMarketing,3 - FieldService,5 - RegionalFieldService,1,1.0,

2 - Executive,5 - SalesAndMarketing,2 - Sales,4 - RegionalSales,1,1.0,

2 - Executive,2 - ResearchAndEngineering,2 - Research,1 - Lab,1,1.0,

1 - Executive,3 - Manufacturing,,,1,1.0,

2 - Executive,1 - Adminstration,6 - Security,2 - ElectronicSecurity,1,1.0,

2 - Executive,5 - SalesAndMarketing,2 - Sales,3 - RegionalSales,1,1.0,

1 - Executive,1 - Adminstration,6 - Security,1 - BuildingSecurity,1,1.0,

1 - Executive,2 - ResearchAndEngineering,2 - Research,6 - Lab,1,1.0,

1 - Executive,1 - Adminstration,6 - Security,2 - ElectronicSecurity,1,1.0,

1 - Executive,3 - Manufacturing,1 - IndustrialEngineering,,1,1.0,

2 - Executive,6 - PurchasingAndContracts,2 - Pricing,,1,1.0,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,7 - WebSoftware,1,1.0,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,3 - EmbeddedSoftware,1,1.0,

1 - Executive,2 - ResearchAndEngineering,2 - Research,3 - Lab,1,1.0,

2 - Executive,2 - ResearchAndEngineering,2 - Research,5 - Lab,1,1.0,

1 - Executive,2 - ResearchAndEngineering,2 - Research,4 - Lab,1,1.0,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,4 - WebSoftware,1,1.0,

2 - Executive,5 - SalesAndMarketing,2 - Sales,2 - RegionalSales,1,1.0,

2 - Executive,3 - Manufacturing,1 - IndustrialEngineering,,1,1.0,

2 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,7 - WebSoftware,1,1.0,

1 - Executive,2 - ResearchAndEngineering,2 - Research,,1,1.0,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,8 - TestAndEvalualtion,1,1.0,

2 - Executive,5 - SalesAndMarketing,3 - FieldService,2 - RegionalFieldService,1,1.0,

1 - Executive,4 - Finance,,,1,1.0,

针对随机选择的一个普通用户而言，其邮件发送1vs1行为有以下特征

ACH1831,1 - Executive,5 - SalesAndMarketing,2 - Sales,5 - RegionalSales,

1 - Executive,5 - SalesAndMarketing,2 - Sales,5 - RegionalSales,30,1058.0,

1 - Executive,5 - SalesAndMarketing,3 - FieldService,3 - RegionalFieldService,7,7.0,

1 - Executive,5 - SalesAndMarketing,2 - Sales,3 - RegionalSales,5,6.0,

2 - Executive,3 - Manufacturing,3 - Assembly,1 - AssemblyDept,4,4.0,

1 - Executive,5 - SalesAndMarketing,3 - FieldService,4 - RegionalFieldService,4,4.0,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,6 - TestAndEvalualtion,3,4.0,

2 - Executive,3 - Manufacturing,3 - Assembly,4 - AssemblyDept,3,3.0,

1 - Executive,3 - Manufacturing,3 - Assembly,4 - AssemblyDept,2,3.0,

1 - Executive,5 - SalesAndMarketing,2 - Sales,4 - RegionalSales,3,3.0,

2 - Executive,3 - Manufacturing,3 - Assembly,2 - AssemblyDept,3,3.0,

1 - Executive,2 - ResearchAndEngineering,4 - Engineering,7 - SystemsEngineering,3,3.0,

2 - Executive,5 - SalesAndMarketing,2 - Sales,3 - RegionalSales,3,3.0,

1 - Executive,1 - Adminstration,6 - Security,2 - ElectronicSecurity,3,3.0,

2 - Executive,3 - Manufacturing,3 - Assembly,3 - AssemblyDept,3,3.0,

2 - Executive,5 - SalesAndMarketing,2 - Sales,1 - RegionalSales,2,3.0,

1 - Executive,5 - SalesAndMarketing,2 - Sales,2 - RegionalSales,3,3.0,

1 - Executive,2 - ResearchAndEngineering,4 - Engineering,2 - MechanicalEngineering,3,3.0,

1 - Executive,5 - SalesAndMarketing,3 - FieldService,1 - RegionalFieldService,3,3.0,

2 - Executive,2 - ResearchAndEngineering,2 - Research,5 - Lab,3,3.0,

2 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,3 - EmbeddedSoftware,2,3.0,

2 - Executive,2 - ResearchAndEngineering,2 - Research,2 - Lab,3,3.0,

2 - Executive,2 - ResearchAndEngineering,2 - Research,3 - Lab,2,2.0,

1 - Executive,5 - SalesAndMarketing,2 - Sales,1 - RegionalSales,2,2.0,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,1 - SQA,2,2.0,

1 - Executive,3 - Manufacturing,3 - Assembly,2 - AssemblyDept,2,2.0,

1 - Executive,2 - ResearchAndEngineering,2 - Research,5 - Lab,2,2.0,

2 - Executive,5 - SalesAndMarketing,2 - Sales,5 - RegionalSales,2,2.0,

1 - Executive,1 - Adminstration,6 - Security,1 - BuildingSecurity,2,2.0,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,7 - WebSoftware,1,2.0,

1 - Executive,2 - ResearchAndEngineering,2 - Research,4 - Lab,2,2.0,

2 - Executive,5 - SalesAndMarketing,2 - Sales,2 - RegionalSales,2,2.0,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,5 - DesktopSoftware,2,2.0,

1 - Executive,2 - ResearchAndEngineering,2 - Research,3 - Lab,2,2.0,

2 - Executive,2 - ResearchAndEngineering,2 - Research,1 - Lab,2,2.0,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,6 - EmbeddedSoftware,2,2.0,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,2 - DesktopSoftware,2,2.0,

2 - Executive,5 - SalesAndMarketing,3 - FieldService,2 - RegionalFieldService,2,2.0,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,4 - WebSoftware,2,2.0,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,3 - EmbeddedSoftware,2,2.0,

1 - Executive,3 - Manufacturing,3 - Assembly,5 - AssemblyDept,2,2.0,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,1 - Publications,2,2.0,

2 - Executive,5 - SalesAndMarketing,3 - FieldService,3 - RegionalFieldService,2,2.0,

2 - Executive,3 - Manufacturing,3 - Assembly,5 - AssemblyDept,2,2.0,

1 - Executive,2 - ResearchAndEngineering,4 - Engineering,8 - TestAndEvalualtion,2,2.0,

2 - Executive,5 - SalesAndMarketing,1 - MarketAnalysis,,1,1.0,

1 - Executive,3 - Manufacturing,3 - Assembly,7 - AssemblyDept,1,1.0,

1 - Executive,2 - ResearchAndEngineering,4 - Engineering,4 - MaterialsEngineering,1,1.0,

1 - Executive,2 - ResearchAndEngineering,1 - ProjectManagement,,1,1.0,

2 - Executive,6 - PurchasingAndContracts,3 - Contracts,,1,1.0,

2 - Executive,1 - Adminstration,6 - Security,2 - ElectronicSecurity,1,1.0,

1 - Executive,5 - SalesAndMarketing,3 - FieldService,2 - RegionalFieldService,1,1.0,

2 - Executive,5 - SalesAndMarketing,3 - FieldService,4 - RegionalFieldService,1,1.0,

1 - Executive,5 - SalesAndMarketing,,,1,1.0,

1 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,,1,1.0,

1 - Executive,3 - Manufacturing,3 - Assembly,6 - AssemblyDept,1,1.0,

1 - Executive,2 - ResearchAndEngineering,4 - Engineering,5 - SystemsEngineering,1,1.0,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,2 - MechanicalEngineering,1,1.0,

1 - Executive,6 - PurchasingAndContracts,,,1,1.0,

2 - Executive,4 - Finance,3 - FinancialPlanning,,1,1.0,

2 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,5 - DesktopSoftware,1,1.0,

1 - Executive,6 - PurchasingAndContracts,1 - Purchasing,,1,1.0,

2 - Executive,1 - Adminstration,1 - Personnel,,1,1.0,

2 - Executive,6 - PurchasingAndContracts,1 - Purchasing,,1,1.0,

2 - Executive,2 - ResearchAndEngineering,2 - Research,4 - Lab,1,1.0,

2 - Executive,5 - SalesAndMarketing,3 - FieldService,1 - RegionalFieldService,1,1.0,

2 - Executive,4 - Finance,1 - Accounting,,1,1.0,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,5 - SystemsEngineering,1,1.0,

2 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,6 - EmbeddedSoftware,1,1.0,

1 - Executive,3 - Manufacturing,3 - Assembly,1 - Stockroom,1,1.0,

2 - Executive,2 - ResearchAndEngineering,3 - SoftwareManagement,1 - SQA,1,1.0,

1 - Executive,2 - ResearchAndEngineering,2 - Research,1 - Lab,1,1.0,

1 - Executive,1 - Adminstration,2 - EmployeeRelations,,1,1.0,

1 - Executive,2 - ResearchAndEngineering,2 - Research,6 - Lab,1,1.0,

2 - Executive,2 - ResearchAndEngineering,,,1,1.0,

1 - Executive,2 - ResearchAndEngineering,2 - Research,2 - Lab,1,1.0,

2 - Executive,5 - SalesAndMarketing,2 - Sales,4 - RegionalSales,1,1.0,

1 - Executive,5 - SalesAndMarketing,3 - FieldService,5 - RegionalFieldService,1,1.0,

2 - Executive,2 - ResearchAndEngineering,4 - Engineering,3 - ElectricalEngineering,1,1.0,

2018年11月1日星期四

下午3时12分

下午。

初步分析了BYO1846和同Team的ACH1831，结果发现从邮件1v1发送联系人行为中，比较分析联系的OS部门，结果显示：

1. 主要发送的邮件都在Team内部，占据了所有发送邮件的绝大多数；
2. 除去本Team之外，在OS全体部门内存在广泛的却稀疏的联系，即涉及OS部门比较多，但是每个部门的邮件基本按照一人一封的形式存在；

因此，试图人为划定OS子集作为relationship的做法并不合适，我们接下来回到自然朴素的假设。

我们的假设：

1. 周围同事离职是重要的刺激工作满意度JS的外部因素；
2. 该因素基于用户自身的OCEAN（决定了其CPB因素，如C决定了迟到行为）和关系的紧密性决定对用户JS影响的程度
3. 因此我们设想影响用户JS的特征应包括：用户自身的Big5（静态特征）+组织编号特征（四位编码代表）+周围各个层次离职用户所占比例
4. 当然，也可以反其道行之，直接用目标用户与已有的离职用户进行聚类，分期其所在群簇中用户的个数，自然特征也包括基本的Big5 + OS编码 + Leave比例

2018年11月1日星期四

下午6时31分

实验后发现：

如果直接分析、比较OS五个层次中离职用户的比例，发现Insiders\_2用户特征并不明显

以最难搞定的ZIE为例：

ZIE0741,2010-08:

Insider\_LaidOff\_0,

Insider\_LaidOff\_1,KSW0708,CLL0306,PKS1187,USM0703,AWW0718,ESP1198,DTB0722,EPG1196,

Insider\_LaidOff\_2,JAT1218,GWO1660,CDG0770,QSG1150,GER0350,OCD1985,

Insider\_LaidOff\_3,ADL1898,WMH1300,LAS0256,MIB1265,TCP0380,CBC1504,JSB0860,CDO0684,KDP1706,FKH0864,JRC1963,QAP0266,BAR1328,OCW1127,BSS0847,MPF0690,RKW1936,RFP1918,RDP1751,FKS1696,BMR0865,CRD0272,EJO0236,DAS1320,HKK0881,MMR1458,JIB1258,SCO1719,ZJN1492,SAF1942,CEW1960,

Insider\_LaidOff\_4,RMB1821,FAM0495,WSW1091,SDL0541,CTH1812,JBG1375,LRF0549,JHP1654,NWP1609,JDM0208,HSF1115,FDS1841,NTV1777,GMM1037,MAF0467,ZAD1621,XMG1579,HBH0111,KBC0818,TAG1610,EAL1813,WBP0828,NWH0960,BRM0126,MAR1075,GWH0961,JXH1061,HFF0560,MMB0556,UAM1108,KEW0198,BNS0484,LSM1382,GFM1815,CIM1095,VCF1602,CKP0630,JKM1790,PBC0077,PTV0067,KBC1390,ZHB1104,PTM1432,

纯碎从心理模型角度去研究，关键是不知道该如何进行改进；