支付风控系统设计：支付风控场景分析(一)

风控是一个让人爱恨交加的话题。 对支付来说风控是必不可少的功能。这个系列的文章将试图从这两个领域简单梳理下支付风控面临的问题，以及如何从技术角度来解决这些问题。

风控是一个让人爱恨交加的话题。 对支付来说风控是必不可少的功能。只要老板不想把底裤都赔掉，那就必须上风控。可对互联网公司来说，风控是一个谜一般的话题，无论是对风控专家还是 IT 工程师而言。随着互联网和大数据技术的引入，风控变成了一个跨学科的领域，可这无疑是互联网公司里面最同床异梦的跨学科。

机器学习、深度学习、规则推理、随机森林……光这些名词就足以让人风控专家望而怯步；而风险事件、尽职调查、巴塞尔协议……这些名词，一提起来 IT 人员就头大。这个系列的文章将试图从这两个领域简单梳理下支付风控面临的问题，以及如何从技术角度来解决这些问题。

概念定义

按照教科书的说法， 风险是指在特定场景下，特定时间内某个损失发生的可能性，或者说是在某一个特定时间段里，人们所期望达到的目标与实际出现的结果之间的差距。金融领域自从诞生以来，就一直伴随着风险。风险控制是指风险管理者采取各种措施和方法，消灭或减少风险事件发生的各种可能性，或风险控制者减少风险事件发生时造成的损失。这里又引入了一个词，风险事件， 它和风险因素经常容易混淆。 风险事件指造成风险的直接原因，风险因素则是间接原因。 如下雨天路滑导致发生车祸造成人员伤亡。 则车祸是人员伤亡的直接原因，是风险事件。而下雨天是间接原因，属于风险因素。 先看一条小道消息惊悚下：

高价回收线报

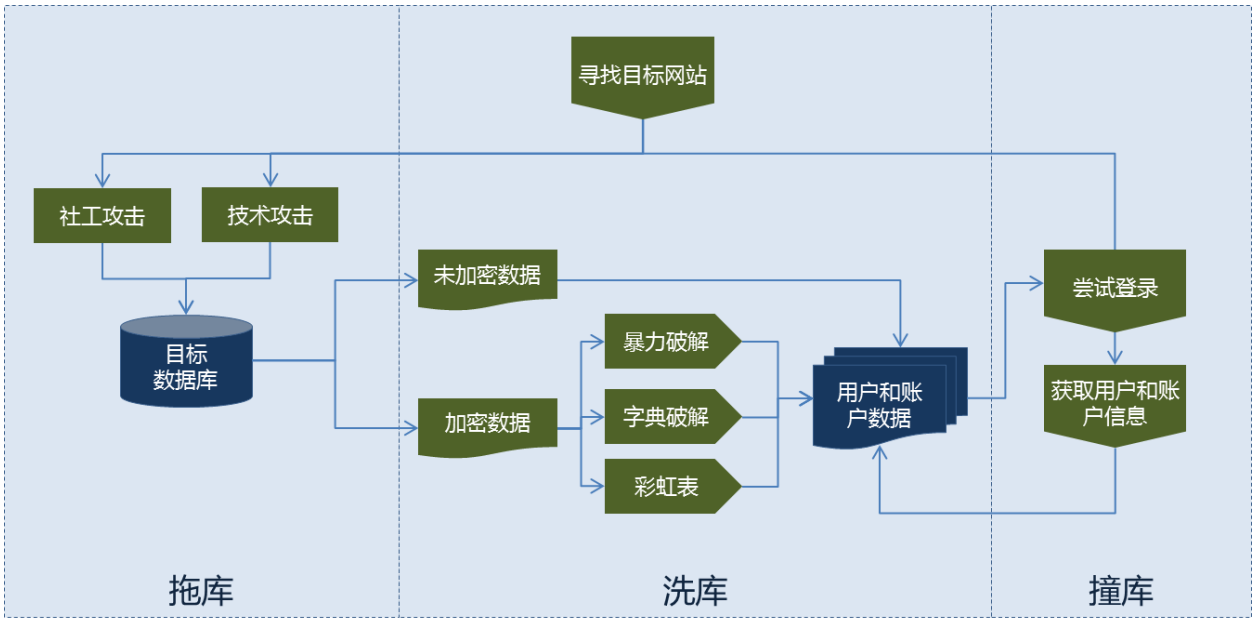
今天22:21 试探你的温柔-单色凌

4月12日消息，据业内人士微博爆料，[redacted]充值系统于昨晚14点30分左右出现重大漏洞亏损在2亿元左右，遭用户篡改数据充值流量和话费。目前，[redacted]方面回应称已修复该漏洞，并已报警。《刑法》第286条规定，“违反国家规定，对计算机信息系统中存储、处理或者传输的数据和应用程序进行删除、修改、增加的操作，后果严重的，处五年以下有期徒刑或者拘役，后果特别严重的，处五年以上有期徒刑。

风控做不好，一个晚上 2 个亿就出去了。饶是该公司财大气粗，也扛不住几次折腾。一个漏洞搞垮一个小公司也是常有的事。对支付系统来说，安全是第一考虑的问题，特别是资金安全，这需要风控系统来保驾护航。那一般来说，支付系统会面临哪些风险？ 不同文献有不同的风险分类，本文试图从账户、资金、交易、操作、信用风险角度来详细描述。

账户风险

支付系统最常见的，也是在黑产圈中最为成熟的，那就得算账户的风险，即俗话说的“盗号”。近几年来，各大型互联网网站的账户泄露事故层出不穷，携程，京东，CSDN 等都中过招，每一次都能引起轩然大波。而在黑产圈，账号窃取都形成了一套完整成熟的产业链。



这是目前在黑产圈中账户攻击的主要流程，以下分析在这个过程中每个阶段的具体操作，为风险系统设计提供依据。

拖库

拖库是实施账户攻击的第一步。考虑到大型网站一般防守比较严密，黑客一般选择从小型网站入手，入侵到一些防守薄弱或有漏洞的网站，将注册用户的资料窃取出来。常见手法包括：

1. 利用操作系统和系统组件漏洞

比如近年来杀伤力最大的漏洞之一 Heartlbleed 漏洞。这个漏洞，在 2012 年 OpenSSL 软件发布时带有这个 bug，而正式公开时间是 2014 年。Heartbleed 漏洞使得黑客有可能通过 memory dump 的手段来获取到服务器上接受的用户请求、密码、甚至是服务器的私钥。

只要持续不断的攻击，任何被加载到内存中且不幸被加载到和 OpenSSL 在同一个区块内存中的数据，都会被黑客所获取。 这两年期间，有多少黑客使用了这个漏洞来窃取网站信息，就不得而知了。 由于此类漏洞的发现和修复往往有一定的时间差，这也给黑客利用漏洞窃取信息带来了便利。

2. 利用网站所使用的第三方组件漏洞

如臭名昭著的 Apache Struts 系列漏洞。从 2010 年开始， 不断地有漏洞暴露出来， 这些漏洞直奔 struts 所使用的 OGNL 表达式，通过构建各种匪夷所思的表达式， 可以远程执行任意命令，包括访问根目录。由于 SSH（Springframework + Apache Struts + Hibernate）架构入门简单、上手容易，再加上各种 IT 培训机构不遗余力的推广，在国内电商、银行、运营商网站上被大量使用。每次 Apache Struts 漏洞的发布，都能够掀起一番血雨腥风。 而 Apache 组织对这些漏洞响应不及时，修复慢，更让这些机构雪上加霜。远离 Apache Struts 更是支付系统的基本要求。

3. SQL 注入攻击

基本上所有网站都会用到数据库。而一些新手在写代码的时候，对用户输入数据不做验证或者验证不到位，就把这些数据直接通过拼接 SQL 语句写入到数据库中，这就很容易导致 SQL 注入攻击。 比如系统在判断用户名和密码是否正确时，会使用这个 SQL 语句来查询数据库：

```
SELECT 1 FROM users WHERE username = 'admin' AND password = 'guest'
```

攻击者可以尝试修改密码为 ‘ OR ‘a’ = ‘a ， 拼接成 SQL 语句：

```
SELECT 1 FROM users WHERE username = 'admin' AND password = 'guest' OR 'a'='a'
```

由此执行成功，获取管理后台的权限。

这三个是常见的攻击方式。 当然还有其他的方式，如木马，钓鱼网站等等，不再详细描述。

洗库

在攻入服务器，获取到资料，特别是数据库的信息后， 需要对信息进行分析。 不是所有的信息都可以直接使用，部分信息，如密码，身份证等，一般都会加密存储。 通过暴力、字典或者彩虹表的方式来破解，获取到破解后的信息，就拿到用户名，密码等资料。

暴力破解：如果知道用户名或者密码的范围，可以通过枚举的方法逐个尝试。 对密码来说，会受限于密码的长度，如果长度在 8 位以上，那可枚举项就太多了，需要几天甚至几年的计算才能破解。
字典表：其实也是暴力破解的一种，区别是可以预先计算出来一些常见的组合，比如生日之类的，然后使用这些组合来进行破解。
彩虹表：这是一种破解哈希算法的技术，是黑客必备的跨平台密码破解方法，可以破解 MD5 进行哈希处理的密码。它的性能优异，在一台普通 PC 上辅以 NVidia CUDA 技术，对于 Microsoft Windows 操作系统使用的 NTLM 密码加密算法，可以达到最高超过 1 千亿次每秒的明文尝试。对于广泛使用的 MD5 也接近一千亿次。

撞库

第三步是撞库，就开始进攻真正的目标网站了。 把拿到的账户信息去尝试登陆大型网站。因为大部分用户，习惯于在多个网站使用同一套账户和密码。如果登录成功，则可以进一步窃取更多的用户信息，比如信用卡信息等。 由此可见， 撞库攻击本质上是利用用户相同的注册习惯，以大量的用户数据为基础，尝试登陆目标网站，从而窃取更多的用户资料。这也使得黑客无需进行系统攻击的情况下，即可轻易获取目标用户信息。

更进一步，黑客们会把这些资料整理后，形成社工库。这个库也日益壮大， 目前有千万规模。除了用户名密码，还有大量的个人隐私也被挖掘出来。比如如家 2000 万数据泄露，其中包含开房信息；QQ 群用户信息泄露、京东 2015 年初用户信息泄露。这都导致大量的个人隐私被窃取甚至出售。

由此可见，账户被窃取，往往是网站防护薄弱和用户安全意识薄弱两种因素导致。

交易风险

支付的交易风险主要是交易过程中的各种恶意行为，而这些行为在电商系统中表现特别突出，包括 自动刷单、人工批量下单以及异常大额订单等场景。在秒杀的时候， 由于其价格有很大的优惠力度， 黄牛会采用机器批量注册账号、机器抢购等方式来争取秒杀商品，普通消费者很难享受到秒杀的实惠，使得秒杀活动效果大打折扣。此外，在商家侧，主要的风险在于刷单。不少商家使用刷单、刷评价的方式来以非正常途径提升销量，积分，信誉等。甚至通过刷单的方式来套取补贴，帮助套现。 从阿里公司发布的《互联网信任环境调查报告》来看，大部分用户在购买的时候，会看中商家的资质和诚信，商品的销量、评论也往往会成为购买的一个参考。 在这种情况下， 刷单就成为一个提升店铺交易量的重要手段。而刷单和反刷的猫鼠游戏，也推高了刷单识别的难度。 以电商为例，一般刷单行为有如下特征：

小号刷单。谁也不会用自己的注册账号来刷单，这样被封的代价就太大了。 小号的来源，可能是商家自己组织注册的，但大部分还是从专业刷单机构手中获取的。
使用虚拟机。大部分网站都会为访问设备植入识别码。通过虚拟机，可以在一个物理机上模拟多台机器访问，随用随建。一般使用 VMWare 来建立虚拟机。而对手机设备，则会采用手机模拟器。
使用 VPN。 这样可以伪装使用全国任何一个地区的 IP，甚至可以使用国外的 VPN。
使用手机 IP：移动和联通的 IP 出口少，所以大部分手机端的出口 IP 并不多。 这些 IP 是电商的白名单，把某个 IP 封了，那会有大量的手机无法正常访问。 所以刷单人员会选择使用这些 IP。
刷虚拟物品：虚拟物品不涉及到物流环节，交易流程简单，很容易就可以把量刷上去。
低价刷单：为了降低成本，往往会将单品价格调低，或者成交金额调低来支持刷单。
交易商品少：刷单时，仅选择少量几个商品进行。
互刷： 一些商家会勾结起来，相互刷单。

这些是从刷单行为的角度来分析的结果。看来简单，可对支付系统来说，如何交易记录中识别出小号、互刷、低价等这些特征，都需要使用大量的数据进行分析才能搞定。

资金风险

2016 年 11 月份的时候， 网上突然出现了大量怀疑支付宝沉淀资金用途的帖子，这些帖子在有意无意地引导一个观点： 支付宝将沉淀资金用于恒生 HOMS 系统的场外配资，用户将资金投放到余额宝有巨大资金风险。 毫无疑问，从监管的角度来看，这是不可能的事情。但这谣传也揭示了支付系统的另一个风险：资金风险。 发展沉淀资金成为支付系统，特别是第三方支付系统的一个公开的秘密。

沉淀资金主要有两种形式：

在途资金 ：指买卖双方在确认交易后，完成结算前尚未到达卖方账户的资金。在买方没有最终确认收货之前，资金暂时交由第三方支付进行保管。这样在买卖双 方从开始交易到最终完成货款两清的这段时间差内，这些存在于第三方支付平台内部的资金，被称为在途资金。
留存资金：对采用交易担保型账户的支付机构，客户需要开立虚拟账户来完 成交易。机构也会吸引客户进行充值操作，即留存一些资金用于交易。比如微信支付和支付宝的钱包。当有交易需求时，可以直接从这里进行扣款。这些留存于虚拟账户中的资金也是沉淀资金的一部分。

沉淀资金对支付来说是必要的， 通过这个资金来帮助买卖双方解决信任的问题，有利于提升用户体验。但这个资金也带来不少风险。2013 年 央行出台了《支付机构客户备付金存管办法》，其中明确要求第三方支付机 构对于客户的备付金要进行严格的区分管理，这一定程度上限制了沉淀资金风险的发生。也就是说，沉淀资金是客户的钱，支付公司不能挪用。支付公司可以获得沉淀资金的利息收益，但是不能够用这个资金来进行投资或者公司内部的消费。对这笔资金进行合理监控避免出现风险，也是支付系统需要考虑的问题。

套现风险

我国法律明确禁止使用信用卡套现，使用信用卡套现是违法的。但是在线支付系统中，使用信用卡进行套现，几乎是不需要成本的。 信用卡套现的手段也很多，一般是通过客户和商家的勾结来完成，比如：

虚假购买，客户通过信用卡购买某商品后，商品并未实际发货，商家将购买的款项打回给客户，完成套现。
退货套现：或者通过信用卡来购买商品，然后退货，将退款返回到借记卡或者其他可提现的渠道，也能完成套现。
自买自卖：商家通过信用卡购买自己的商品，将货款打入到借记卡中，完成套现。

上述的套现手段，很难识别。套现很难完全杜绝，除了要求退款资金必须原路返回外，还可以通过数据分析手段来减少发生的频率。

操作风险

按照巴塞尔委员会《操作风险管理》的定义， 操作风险主要是指那些由于用户支付终端操作失误、工作人员违规操作、内控机制失灵等人员操作上的原因引致损失的风险，或者说是外部风险、员工风险和流程风险。

流程风险指由公司的规章制度管理、业务流程不完善而引发的风险。对一些支付公司而言，作为新兴的经济形式，不像银行那样有一套成熟、规范的流程以及完善的培训机制，这就容易触发流程风险。在以“快”为特征的互联网公司，功能创新非常重要，但往往也容易忽视了风险管理相关配套制度的建设和落实，从而为线上运行的新功能带来隐患。当新的支付方式上线后，配套的清结算、记账、对账等功能，未必能够及时地跟上，更不用说相关的内控制度建设、岗位人员配备的工作。

员工风险指的是支付机构的员工不遵守职业道德，违法违规或违章操作，单独或参与骗取、盗用机构资产和客户资金，工作疏忽等行为导致的损失。在缺乏成熟培训机制的互联网公司中，这类问题往往更加突出。

欺诈行为：员工同外部人员相勾结，通过挪用资金、职务侵占等方式非法占有公司财产或者泄露出卖公司商业秘密的行为。
越权行为：员工未经授权、或超越工作权限导致的损失，比如开发人员私自修改数据库给人送优惠券。
错误操作：员工在具体业务操作过程中的失误造成的错误操作。

合规风险

合规风险指机构因未能遵守相关的法律法规从而导致机构可能受到处罚、声誉受损的风险。从 2004 年的电子签名法开始，和支付相关的法律法规：

发布时间	发布机构	法律法规
2004.8	商务部	《中华人民共和国电子签名法》
2005.4	中国电子商务协会	《网上交易平台自律规范》
2005.6	中国人民银行	《支付清算组织管理办法》
2005.10	中国人民银行	《电子支付指引（第一号）》
2007.3	商务部	《关于网上交易的指导意见（暂行）》
2008.4	商务部	《电子商务模式规范》《网络购物服务规范》
2009.11	商务部	《关于加快流通领域电子商务发展的意见》
2010.9	中国人民银行	《非金融机构支付服务管理办法实施细则》（征求意见稿）
2011.6	中国人民银行	《关于规范商业预付卡管理的意见》
2011.10	中国人民银行	《支付机构预付卡管理办法》（征求意见稿）
2011.11	中国人民银行	《支付就够互联网支付业务管理办法》（征求意见稿）
2012.1	中国人民银行	《支付机构互联网支付业务管理办法》（征求意见稿）
2012.3	中国人民银行	《支付机构反洗钱和反恐怖融资管理办法》
2012.6	中国人民银行	《银行卡收单业务管理办法》（征求意见稿）
2013.6	中国人民银行	《支付机构客户备付金管存办法》（征求意见稿）

其中 2010 年的《非金融机构支付服务管理办法实施细则》是一个标志性的法规，标识国家开始认可第三方支付的地位并开始执行监管。 之后，央行又陆续出台一系列法规来规范支付行业的发展。可以说，支付行业的业务创新，是一个不断地由乱而治的过程。而对支付公司来说，滞后的法规建设，也给业务发展带来了巨大的风险。 2013 年支付宝推出互联网理财产品余额宝，在短期内迅速发展成为国内最大的基金。随后多家支付机构也开始开展这个业务，后续央行出台了《支付机构网络支付业务管理办法》，对支付公司的业务范围、资金转移金额进行限制，避免了该业务的过度发展。 2014 年央行也相继叫停了虚拟信用卡、二维码支付等业务。

合规风险是国内第三方公司一个无法规避的风险，在企业发展过程中，需要密切关注央行的动向，减少合规带来的负面影响。

洗钱风险

第三方支付目前成为洗钱的重灾区。 2016 年 8 月，18 家支付机构被公安部列为重点整改对象。 这些支付机构提供的服务，存在未落实实名制、风控措施不严格等问题，被犯罪分子所利用，沦为咋骗和洗钱的工具。主要手段包括：通过一些第三方支付平台发行的商户 POS 机虚构交易套现；将诈骗得手的资金转移到第三方支付平台账户，在线购买游戏点卡、比特币、手机充值卡等物品，再转卖套现；利用第三方支付平台转账功能，将赃款在银行账户和第三方支付平台之间多次切换，使得公安机关无法及时查询资金流向，逃避打击。2012 年央行发布的《支付机构反洗钱和反恐怖融资管理办法》，对支付机构如何防范洗钱风险做了明确的规范和要求，需要支付公司严格遵守。

以上是支付系统可能面临的风险分析。支付风控系统是通过采集交易、渠道、商品、账户、用户等信息，对这些数据进行实时和定时的挖掘分析，识别出各种风险，采取各种措施降低损失。这是支付风控系列的第一篇文章，这个系列将包括如下内容：

支付风控场景分析（本文）；
支付风控数据仓库建设；
支付风控模型和流程分析；
支付风控系统架构。

作者：凤凰牌老熊，程序员 & 架构师