用图来分析以太坊论文解读

16350027 黄俊

文章做了什么

1. money flow graph (MFC)
2. contract creation graph (CCG)
3. smart contract invocation graph (CIG)

论文主要的贡献就是通过构建了这三个能够表现以太坊活动的特征的图，希望借助图的信息和已有的分析方法来对于以太坊当下的状况进行分析，并且提出改进的方法。可能一看上去会觉得没什么大不了的因为会觉得这篇论文其实没什么大不的不就是收集数据嘛，又没有提出什么新颖的算法来改进整个系统。但是其实这篇文章的价值其实就在于数据，就像imagenet一样，往往技术的发展其实很大程度上是数据跟不上，有的时候数据可以大大扩宽人们的想象力。

具体而言，文章所说的分析是基于cross-graph analysis的，然后针对的问题是安全方面的，包括了attack forensics也就是对于攻击发生的追溯以至于给所有攻击者以威胁 和 anomaly attack也就是匿名攻击，也就是怎么追溯的问题。

##如何得到数据

首先据论文所讲，所谓的外部交易的数据都很好获得，因为这一些数据都被记录在了区块上，但是至于内部的数据（智能合约内部）就没办法直接通过分析区块的方式来获取。所以这个团队就对于evm整个虚拟机进行了修改以便能够得到所有内部交易的数据。

通过进行图的构建，从大致的统计数据可以看出来几个有趣的insight. 分别是

1. 现在还是直接交易的记录占大多数，使用智能合约的还是少数。
2. 现在大多数,dapp都是和金融有关的应用。
3. 很多人对于智能合约的时候都只停留在部署一个玩具的智能合约
4. 发现很多智能合约都是重复的，独立的智能合约只占到6%，这就意味着由于智能合约的开放性，其实要想复制别人的合约是很容易的。与此同时还有一些对于磁盘特别不友好的合约会造成每一次同步的时间都被大大拉长。

我的看法就是其实这一些insight其实算是大家的共识了，但是不一样的是通过这一次非常全面的数据收集，真正给我们的预想做了一个证据的backup，之后再去讨论现在以太坊智能合约的趋势时我们可以说自己的观点是有数据依据的而不是想原来一样只能说这是自己觉得的，是自己的看法，根本一点说服力都没有。

##如何感知有问题的智能合约 因为当初从智能合约的角度去分析如果只是通过被复制的次数来判断而不去实际使用其实是分辨不出到底智能合约是不是有用的，因为智能合约可能会被很多种方式正常的复制比如exchange market这种类型的。 所以要想分析智能合约的合理性还是要从对于用户形象做一个刻画的角度去分析，如果一个用户创造了很多没什么人用的智能合约，那么大概率可以得到这个用户所产生的智能合约都是存在问题的印象，所以进一步可以对于所属的智能合约给出一个定义。

但是真正这篇论文所提出的检测算法远远没有这么简单，这个是结合了之前通过统计得到了三种特征图，依照统计得出的信息来对于智能合约行为的正常与否来进行判断。

结论

这篇论文最突出的贡献应该对于evm进行了定义，使其能够得到所有详尽的数据并且通过这一些数据定义了一些系统的用图的角度来分析的方法。

并且通过结合起以太坊三种不同特征的图的性质，有效地提出了能够解决或者避免检测和追溯效果的方法。

我的读后感

这一篇论文可真的算是一篇大制作了，本身对于区块上的数据能够有一个分析就已经需要很大的工作量了，但是这篇论文为了准确性竟然还通过定制evm的方式来希望能够收集到内部的交易数据，最后得到了极为完整的数据。在这个完备的数据上再进行分析势必能够发现新的规律和趋势，但是如何使用这么庞大的数据又成了一个问题所以之后使用了图的方式，希望能够站在一个更好，更加能够表现数据的角度上去分析问题。

并且在最后还针对区块链中存在两个问题给出了非常使用的解决方法。

如此大制作能被infocom收录确实是实至名归。