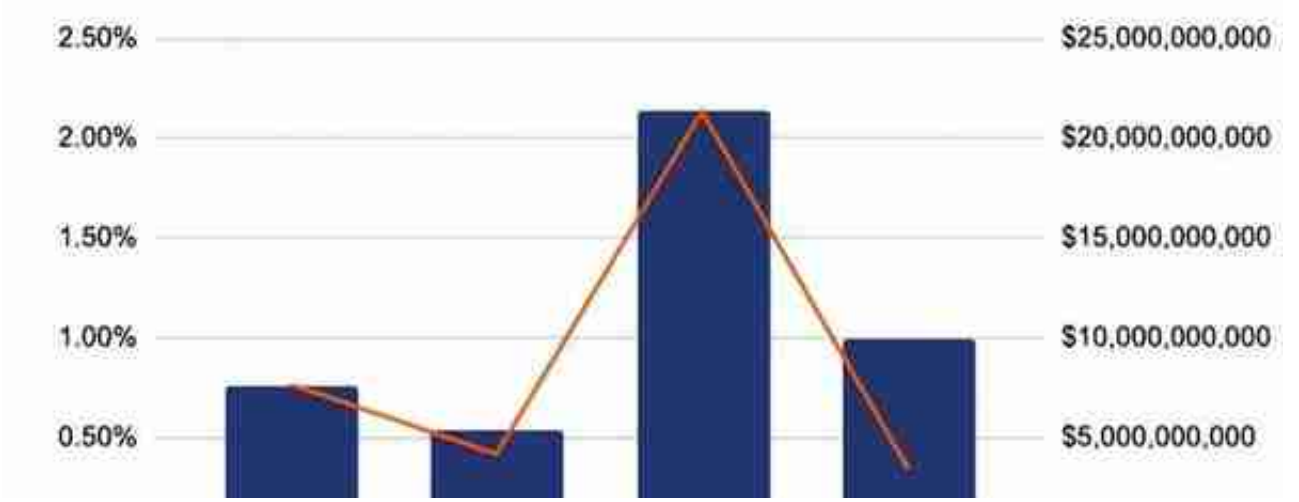

图1 2017～2020年虚拟数字货币交易额变化

另外，我国正面临着数字货币的威胁与挑战，无法规避我国激进反对分子和恐怖势力与境外组织勾结并通过数字货币筹资的风险，我国新疆、西藏甚至香港等地区都直接面临打击这种非法集资的挑战。香港修例风波中的极端反对势力已经通过比特币和发行所谓“抗争币”为暴徒募集活动资金。在虚拟数字货币时代，我国正面临打击通过虚拟数字货币进行非法融资的艰巨任务。

基于以上现实背景，在总体国家安全的需求牵引下，我们认为对非法的加密数字货币融资构建识别与追踪体系是十分迫切和需要的。

二、区块链技术基础理论

应用区块链技术可以使交易在互相不信任的情况下正常进行。区块链技术与传统数据库技术不同，它的显著优势是不可篡改性，不能伪造且具有智能合约特性，因此区块链技术的应用可能引发社会变革，促进社会科技发展。在十三五规划中，国务院对区块链发展进行规划，将其纳入战略性前沿技术发展前列。早在2017年，国外一些公司就对区块链技术发展做出战略目标，Gartner公司将区块链技术列入十大战略科技。

区块链技术最显著的五大特征：

（1）开放性:交易过程中数据时公开共享的，除了包含交易双方的个人隐私外，其余所有数据都是高度透明的，且区块链技术是开源的。

（2）去中心化：该特性是区块链技术最本质特征，由于区块链技术在应用过程中不涉及中心管制和其他依赖，在各时间段信息传输都是自我管理。

（3）安全性：想要修改或操控网络数据时，需要掌握全部数据信息，否则无法进行修改，保障了网络数据安全，避免人为修改修改数据。

（4）独立性：区块链系统在运营过程中独立存在，不依赖其他第三方，遵从协议规范在系统内不需人为干预下自主地进行数据交换。

（5）匿名性：除特定要求外，仅从技术层面来看，信息传输过程中不用公开验证节点信息，传输过程是匿名进行的。

区块链技术架构分为六层：一是数据层、二是合约层、三是共识层、四是应用层、五是激励层及六是网络层。随着科学技术的发展，区块链技术不断更新，其外延得到发展演变。本文站在数据分析的角度，从数据的类型和环境出发，认为区块链可以分为三层。最底层是交易层，对应的是以比特币为代表的区块链1.0阶段。交易是改变区块链数据的手段，同时也为区块链提供数据依据。合约层处于中间位置，通过把合约条款电子化的方式，使条件成立时，交易能自动执行，交易过程中不断有新数据产生，且各流程都在进行交易。

通过相关的区块链技术，我们可以实现用户隐私信息的隐藏，由于网络信息共享原理，想要在分布式环境中对信息进行保密，对技术要求十分高。对不同许可机制和应用场景，区块链可划分为三个层级：一是联盟链、二是公有链、三是私有链。与机构与企业隶属的传统数据库不同，区块链属于公开共享数据库，所有人都有自由获取区块链数据的权利，区块链数据并非内部人员专属。区块链数据的公开性为数据分析人员提供数据依据，通过分析公有数据，为科技进步提供前进动力，激发技术发展潜力。

三、研究内容与方法

3.1总体思路

针对加密数字货币的交易业务模式，我们团队把加密数字货币GEC全周期定义为加密数字货币从产生到流通的整个阶段过程，主要包括产生、兑换、流通等三个主要过程．通过对整个GEC周期中的不同事件建立模型，我们就可以进行全方位的监督、管理和调控，提高加密数字货币交易的可追溯性，以便所有的交易有据可循、有账可查，为后续科学监管方法的提出奠定基础。

3.2研究问题与方法

我们深入研究现有与研究问题相关性较高的文献，结合本文研究方向，文献综述部分主要从以下五个角度入手了解现有加密数字货币的相关研究情况：一是实体识别；二是网络画像；三是非法行为检测与分析；四是网络可视化；五是交易模式识别。

（1）实体识别

以比特币为例，比特币交易的一个突出特点就是交易双方均为匿名，但由于比特币交易较为复杂往往会涉及到多个输入和输出，这就导致如何在相关交易信息中识别用户成为了一个需要解决的问题，即如何识别一个用户拥有多少地址。但由于没有有效的方法可以验证识别其是否属于特定的用户，按照已有论文的研究思路，一个实体可能是一个组织或者一个用户等；反之，同一组织或者同一用户实际上可能同时操控数个不同的实体。在相关文献中可以看出，为更好地识别出潜在实体，其大多采用启发式的方法，主要包括以下两种识别方式（找零地址法和共同输入法）。

尽管如此，启发式规则也会容易产生一些错误，因此目前也一些学者也提出了不同的优化方法。随着地址标签的广泛使用，有监督的机器学习算法也被应用于捕捉带有交易特征的标记样本之间的差异。比特币地址分类任务可以分为二进制分类和多分类的类型。二元分类任务通常能达到很高的准确性，表明非法活动在其交易模式中是高度可分的。相比之下，多个分类任务的准确率明显较低。就特征而言，交易量特征在区分实体方面起着至关重要的作用。在学习算法方面，基于决策树的方法，尤其是随机森林，在大多数任务中取得了最高的性能。基于GNN的方法，尽管它们在最近的文献中很新颖和流行，但并没有显示出优于成熟方法的预测能力。

（2）用户画像

比特币主链在不断的发展过程中以为挖掘出50万块以上的区块，其中包含了大量的交易数据，相关信息超过了150GB，此时我们就需要分析相关数据中具体包含了多少用户，以及进行交易的用户具有何种特征。从某种程度来说比特币属于一种虚拟资产，其交易过程是否符合经济学规律，比特币是如何在其使用对象中分配的，都是值得思考的问题，这类研究被称为用户画像。

3. 3网络可视化

随着区块链的深入发展，相关技术不断创新，区块链交易不断增多，交易网络日趋庞大，可视化研究工具的运用对于研究的开展来说较为重要的推进意义。在现有研究中，很多学者从这一角度入手展开研究。目前很多相关研究中将区块链交易的特殊结构图作为研究重点，重点关注于相关交易中交易双方的行为，其可以较为快捷的识别出潜在的异常交易行为和交易模式等。

3 .4交易模式识别

比特币与传统银行支付有很大的差异，其实也个匿名系统，且其没有运营中心。较强的匿名性使比特币交易中滋生了一些非法行为，诸如赌钱、诈骗、洗钱等，如何从比特币交易数据中识别异常交易，发掘潜在非法交易、非法行为是一个十分具有研究价值的课题。而在这一过程中，识别并分析比特币相关交易行为、交易模式是较为重要的研究基础。

3. 5非法行为检测与分析

区块链的一个突出特征就是匿名性，这就使得我们很难获取交易双方的相关信息。目前已有学者将比特币的使用者具体划分为以下四个类型：一是计算机技术狂热爱好者；二是犯罪分子；三是投机客；四是自由主义者。事实上，目前的区块链交易中，的确存在许多潜在的违法交易行为：如诈骗、恐怖主义融资等。因此，提高区块链数据识别技术可以促进区块链网络的可以持续发展，打击犯罪，又有利于相关

监管部门提高监管能力。

最后，在线情报平台是提供深度区块链信息的网站。一些平台还允许用户将众包知识边缘发布到他们的数据库中。这些智能平台包括Blockchain.info（现在称为Blockchain.com）、Etherscan、WalletExplorer和BlockCypher。Chainalysis和Elliptic等技术公司也在加密货币数据分析和恶意活动监控方面提供全面服务。

参考文献：

[1] Chainalysis-Crypto-Crime-2021[EB/OL]. https://go.chainalysis.com/2021-Crypto-Crime-Report.html

[2] Ron D, Shamir A. Quantitative analysis of the full Bitcoin transaction graph [C] //Proc of the 17th Int Conf on Financial Cryptography and Data Security. Berlin: Springer, 2013: 6-24

[3] Androulaki E, Karame G O, Roeschlin M, et al. Evaluating user privacy in Bitcoin [C] //Proc of the 17th Int Conf on Financial Cryptography and Data Security. Berlin: Springer, 2013: 34-51

[4] Badev A, Chen M. Bitcoin: Technical background and data analysis [EB/OL]. [2018-02-08]. https://srn.com/abstract=2544331

王佳鑫[1],[2]

（[1] 南京大学信息管理学院，[2]南京大学众享科技区块链联合实验室）