隐私计算：动态的加密技术

——区块链技术引卷之八

通证通研究院
区块链研究报告

专题报告　　宏观研究

2019.05.30

通证通 x FENBUSHI DIGITAL

分析师：宋双杰，CFA

Email:master117@bitall.cc

分析师：孙含儒

Email: sunhanru@bqbase.org

分析师：吴振宇

Email: wuzhenyu@bqbase.org

特别顾问

沈波

Rin

JX

更多研究请关注通证通公众号获取

通证通研究院　　FENBUSHI DIGITAL

请务必阅读最后特别声明与免责条款

导读：隐私计算技术是密码学的一个前沿发展方向，填补了数据在计算环节隐私性问题的空白，将基于密码学的信息安全体系打造成完整的闭环，为云计算、分布式计算网络和区块链等技术的应用提供隐私性基础。本专题将简述隐私计算技术，并分析其起源、技术方向与应用前景。

摘要：

随着信息技术的不断发展，数据逐渐成为政府、企业与个人的重要资产，其发掘、存储、处理与使用变得愈发重要，逐渐产生了隐私性需求。**隐私计算，是一类数据或计算方法保持加密状态，不泄露给其他合作方的前提下，进行计算合作的技术，**其出现填补了密码学出现以来在信息的处理和使用环节的空白。目前阶段，密码学层面的隐私计算主要有全同态加密、多方安全计算、零知识证明等技术方向。

满足同态性的加密函数能够实现在不解密原始数据的前提下对加密数据进行某一运算，提供了对加密数据的计算能力。全同态加密算法则是指给定任意一种运算规则，可以通过算法构造出对加密数据的相应运算规则，并满足同态性。全同态加密是相对基础性的隐私计算技术，应用范围较广，但其目前计算效率较低，并存在一定局限性。

安全多方计算解决如何在参与计算的各方不泄露自身输入、且没有可信第三方的情况下安全地计算约定的函数并得到可验证结果的问题。主要解决的是互不信任的参与方在保护隐私的前提下协同计算的问题。其自身同样存在局限性，不能保证参与者的诚实度，也无法阻止参与者恶意输入。

零知识证明是证明者在不透露隐私数据的情况下，向任意第三方证明自己确实拥有特定数据的算法。多用于匿名区块链隐藏交易细节，实现匿名性。

隐私计算在云计算、分布式计算网络和区块链三个方向有广阔发展前景。隐私计算可以让数据在云计算过程中保持加密状态，提高了计算过程中的数据安全。隐私计算也使隐私数据上链成为可能，并且同样通过区块链技术确保其可验证性。

风险提示：技术存在瓶颈，落地应用不及预期

目录

图表目录

隐私计算技术是密码学的一个前沿发展方向，填补了数据在计算环节隐私性问题的空白，将基于密码学的信息安全体系打造成完整的闭环，为云计算、分布式计算网络和区块链等技术的应用提供隐私性基础。本专题将简述隐私计算技术，并分析其起源、技术方向与应用前景。

1 隐私计算：加密技术的另一维度

随着信息技术的不断发展，数据逐渐成为政府、企业与个人的重要资产，其发掘、存储、处理与使用变得愈发重要，逐渐产生了隐私性需求。数据科学的发展使数据的应用场景不断扩展，相应的合作也开始涌现，隐私性问题也随之而来。例如，企业可能需要使用合作方的数据以形成某种判断或结果，而合作方并不愿意将自己的数据完全交给他人，企业同样不希望自己的查询条件或分析方法被合作方得知；使用云计算资源时，使用者也希望自己的数据和运算方法能够保密，然而现实中却不得不将内容全部上传，从而面临泄露的风险。随着云计算和区块链的发展，隐私计算的需求愈发涌现，这一结合了密码学和计算科学的前沿领域再次受到了大家的关注。

隐私计算，是一类在保证数据提供方不泄露敏感数据的前提下，对数据进行计算并能验证计算结果的技术。

当代密码学起源于 1977 年，Ron Rivest，Adi Shamir 和 Leonard Adleman 发明了非对称式加密（又称公开密钥加密）算法 RSA。RSA 利用了目前计算机分解素因子运算难度的不对称性，设计了公钥用于加密、私钥用于解密的公钥加密体系，私钥不会出现在数据传输环节，极大提高了加密数据传输的安全性。RSA 算法发表于 1977 年 4 月 3 日，犹太民族的逾越节，如摩西出埃及一般，人类的加密技术突破了长期以来的瓶颈，到达了新的阶段。

密码学通过数学理论将数据转化为密文状态，无私钥不能读取其内容，解决了不安全环境下隐私存储与通信的问题，但在使用环节存在空白。到了信息的使用环节，在通讯和存储过程中处于加密状态的数据就不得不进行解密以用于查询和计算。所以，基于密码学的信息加密体系在使用环节是存在空白的，目前尚不能构成闭环的加密系统。当信息拥有者不得不提交数据使用第三方服务时，他就面临着信息泄露的风险，其他环节的加密状态也就失去了意义。针对这种情况，学术界开展了加密状态下进行数据计算的研究，也就是我们所说的隐私计算。

1978 年 Ron Rivest、Leonard Adleman 和 Michael L. Dertouzos 提出了同态加密问题，并在同年提出了满足乘法同态的算法 RSA。在此之前，密码学研究关注的都是数据在存储和传输过程中的静态安全，而同态加密问题的提出将加密技术的研究从静态引向动态，是理论上的巨大革新，也开创了隐私计算的先河。

1982 年，华人图灵奖得主姚期智开创性的提出的百万富翁问题，引入了多方安全计算概念。姚期智在他的论文《Protocols for Secure Computations》中提出了百万富翁问题，即两个百万富翁在没有可信第三方、不透露自己的财产状况的情况下，如何比较谁更富有。

20 世纪 80 年代，MIT 研究员 Shafi Goldwasser、Silvio Micali 和 Charles Rackoff 提出了零知识证明的概念。零知识证明涉及两个参与方：证明者和验证者。它的目的是解决如下问题：证明者如

何向验证者证明自己拥有某一特定的数据，但证明过程不能透露任何有关该数据的信息。

经过学界的不断研究和发展，以全同态计算、安全多方计算和零知识证明为代表的隐私计算取得了一定的成果，也成为了密码学研究的热点问题之一。

2 主要技术方向：全同态加密、安全多方计算与零知识证明

目前阶段，密码学层面的隐私计算主要有全同态加密（Full Homomorphic Encryption，FHE）、多方安全计算（Secure Multi-Party Computation，sMPC）、零知识证明（Zero-knowledge Proof）三种主要的技术方向。此外，还有可信执行环境、不可区分混淆等方向。本篇专题将解析全同态加密、安全多方计算与零知识证明，并分析它们的优势与不足。

图表1：隐私计算主要技术方向

技术路线	描述	主要问题
全同态加密	密文数据计算得到的密文结果，解密后等于明文计算结果	加、乘法同态应用场景较少；全同态加密计算资源消耗高
安全多方计算	无可信第三方前提下，共同计算某一函数得到正确结果，并且不泄露输入值	无法保证参与方的诚实性；效率较低
可信执行环境	通过划分内存区域等硬件手段构造沙盒，如 Intel 的 SGX 和 AMD 的 PSP	依赖硬件厂商，可被攻击
不可区分混淆	将代码混淆成无法理解的形式，但保留原有功能	尚未有实用性研究
零知识证明	不了解某一秘密的前提下验证另一方确实拥有秘密	验证速度慢，可扩展性差

资料来源：通证通研究院

2.1 全同态加密

在 RSA 算法问世前使用对称式加密算法的时代，加密和解密数据遵循同样的规则。对称式加密的关键要素包括加密算法和密钥，数据发送方使用特定的密钥加密数据，并且将加密数据和密钥发送给接收方。在传输的过程中，加密数据或密钥存在被拦截的风险。一旦加密算法被破解，如果不及时更换密钥，那么双方此后的通信过程就不再安全。

非对称式加密保证了数据存储与传输的安全。非对称式加密算法通过产生成对的公钥与私钥，公钥用于加密数据，私钥用于解密对应公钥所加密的数据。在上面的例子中，数据接收方只需对发送方公开公钥，发送方使用接收方的公钥加密数据，接收方收到后使用私钥解密。整个加解密过程中双方不需要交换有关私钥的信息，因此安全性非常高。

同态加密保证了数据计算过程的安全，提供了对加密数据进行处理的功能。如果要对原始数据进行处理，一般需要先解密数据。如果执行计算过程的一方是不可信任的，那之前的加解密过程等于白费心思。

同态原本是抽象代数中的概念。用加密函数 f 表示从原始数据 M 经加密得到密文 S 的过程，f^{-1}表示其逆运算，即解密过程：

$$S = f(M)$$

如果原始数据 S1、S2 之间可以进行某种运算“+”，并且运算结果仍然是可以被加密函数处理的，此外还满足

$$f(S1 + S2) = f(S1) + f(S2)$$

就称该加密函数 f 满足对运算“+”的同态性。这里的“+”表示一种抽象的运算规则，它可以是我们熟知的加法、乘法等等。

满足同态性的加密函数能够实现在不解密原始数据的前提下对加密数据进行某一运算。对加密后的数据进行计算得出结果：

$$R = f(S1) + f(S2)$$

计算服务请求者在收到计算结果后进行解密，获得的结果：

$$f^{-1}(R) = f^{-1}(f(S1) + f(S2)) = f^{-1}(f(S1 + S2)) = S1 + S2$$

与原始数据直接运算结果相等。常见的非对称加密算法 RSA、ECC 都是加法同态的。

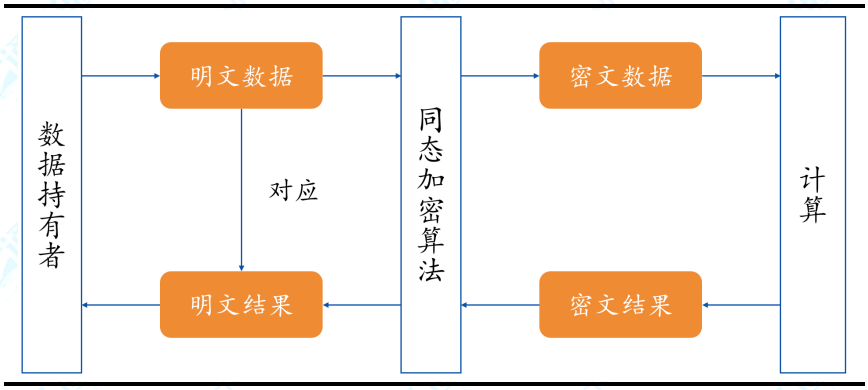
全同态加密算法则是指给定任意一种运算规则，可以通过算法构造出对加密数据的相应运算规则，并满足同态性要求。

全同态加密算法按照发展阶段可以划分为三类。2009 年，Craig Gentry 在论文《Fully Homomorphic Encryption Using Ideal Lattices》（基于理想格的全同态加密）中给出了全同态加密（FHE）的首个算法实现，也是首个同时满足加法同态和乘法同态的加密算法。但 Gentry 的算法效率极低。据测试，该算法执行单次位操作运算需要 30 分钟，远远无法满足大规模应用的需求。

Brakerski 和 Vinod Vaikuntanathan 在初代全同态加密算法的基础上作出改进，于 2011 年提出了 BGV 算法。它的数学基础是 RLWE（Ring Learning-with-errors）问题的求解难度的不对称性。

Craig Gendry、Amit Sahai 和 Brent Waters 于 2013 年发表 GSW13 算法，进一步优化了同态乘法算法。目前主流的全同态加密计算系统均使用后两类加密算法。

图表2：同态加密技术



资料来源：通证通研究院

全同态加密是相对基础性的隐私计算技术，潜在应用范围较广。全同态加密技术实现了密文的直接运算，解决了计算合作中的隐私问题。通过全同态加密，计算资源的需求方可以将涉及隐私的数据以密文状态发送给云服务器或计算公司等非可信的第三方并完成计算，不需要担心任何的信息安全问题。

现有的解决方案效率损失较大，尚不能大规模商业应用。目前，已经有一些算法相对 Gentry 最初的算法有了一定程度的改进，如 Smart 等人提出的 BGV 算法放弃了 Bootstrapping 而采用换模技术与私钥交换技术，一定意义上取得了突破，是目前效率最好的全同态加密算法。但是，从绝对性能来看，当前全同态加密算法的效率还是非常低的，同样的运算所消耗的计算资源大约要比明文高 6 个数量级。因此，全同态加密仍是一种处于研究和实验阶段的技术，尚不足以商业应用。未来，全同态加密技术的落地一方面有赖于更高效率算法的出现，另一方面也将受惠于设备计算能力的提高，可能将从关键性数据开始逐步应用，最终扩展到更多的计算场景。

同态加密自身也存在一定局限性。同态加密的所有中间结果都是加密的，因此服务器不能对中间值做出任何决策，也就是说需要计算所有的条件分支，因此只能将所有操作都包含到函数中，增加了函数复杂性，并会造成性能损失。

M.Van Dijk 和 A.Juels 已经证明同态加密必须在同一个密钥下进行，在多方合作中有共谋的可能，因此全同态加密通常适合协作计算，目前可以防止串谋攻击的同态加密安全计算效率低于其他通用协议。

2.2 安全多方计算

安全多方计算解决了如何在参与计算的各方不泄露自身输入、且没有可信第三方的情况下安全地计算约定的函数并得到可验证结果的问题。例如，在百万富翁问题中，两名富翁可以各自加密自己的财产状况 X 和 Y 作为输入，通过特定的算法，双方可以得到可信的 X 和 Y 的比较结果，但无法获知对方的财产情况。

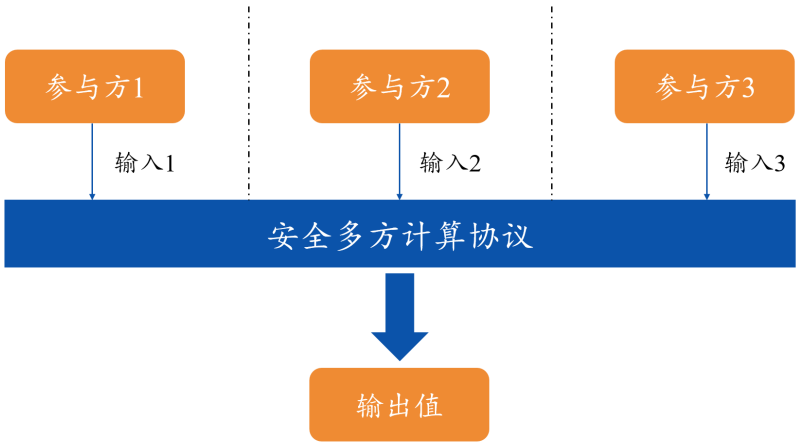
姚期智提出的混淆电路（garbled circuits）是针对两方计算的一种解决方案，Goldreich、Micali 和 Wigderson 等人对其进行了研究和发展，目前已经成为密码学领域的热点问题之一。

安全多方计算还有如下应用场景：将一个需要加密的密钥 K 拆分为 n 个不同的部分，保管在不同参与方，各方都不知道其他人保管的那部分。利用这 n 份中的任意至少 t 个部分（2<=t<=n）就能够

完整地恢复密钥 K 的内容，这样的密钥保管方式又称为（t, n）门限签名。

安全多方计算主要目的是解决互不信任的参与方在保护隐私的前提下协同计算的难题。现实世界中，会经常出现一些协作情景下，计算参与方希望运用他们的数据得到某些结果，但不愿意把自己的数据透露给其他人的情况。安全多方计算的设计理念满足了现实经济合作中参与方希望得到合作利益却不希望被泄露自己数据的客观需求，具有较大的潜在市场空间，其主要应用领域包括电子选举、电子拍卖、资产托管等。

图表3：安全多方计算技术



资料来源：通证通研究院

大部分安全多方计算方案对参与方的诚实性有所要求。在安全多方计算研究中，通常将参与者分为诚实者、半诚实者和攻击者。诚实者提供正确数据，且不尝试窃取他人输入数据；半诚实者提供正确数据，在没有不利后果的前提下愿意获取他人输入；攻击者提供虚假数据破坏合作，且试图窃取他人输入。大部分相对高效的安全多方计算协议都只有应对大部分参与方都是诚实者或半诚实者的情况具有安全性，仅 SPDZ 线的研究成果能够在大部分参与者都是攻击者的情况下保证安全性和正确性。安全多方计算协议至少应该能够在在所有参与者都是半诚实者的情况下运行，才能够适应市场经济的现实情况，保护参与者隐私。恶意的攻击者不期望得到正确的结果，所以更多地要依靠激励机制的设计来加以限制。

除了结果的正确性和输入的隐私性外，安全多方计算协议还要保证参与方之间的公平性。即计算结果只能所有人都得到或都得不到，以应对攻击者可能的提前终止行为。

安全多方计算的局限性有：效率较低；无法保证参与方输入的真实性。无法阻止参与方恶意构造输入，并从结果推测其他人的输入。

2.3 零知识证明

零知识证明，是证明者在不透露隐私数据的情况下，向任意第三方证明自己确实拥有特定数据的算法，分为交互式和非交互式两种。零知识证明算法具有以下特性：

完备性。如果证明方和验证方都是诚实的，证明方确实拥有特定的数据，那么验证一定通过。

稳定性。如果证明方没有拥有特定的数据，那么诚实的证明方几乎不可能通过验证。

零知识性。在验证过程结束后，验证方不会得到任何有关关于数据的信息。

在非对称加密中，已经出现类似零知识证明的概念。如果某人需要证明他确实是某一特定私钥的持有者，他可以让验证方随机生成一个数字，自己用私钥对数字进行签名。如果该签名可以被对应的公钥验证，那么证明通过。在整个过程中，持有者的私钥不会被泄露。但由于公钥和私钥是一一对应的，只要完成了这个验证过程，第三方很容易将私钥持有者和公钥对应起来。

zk-SNARK（Zero-knowledge succint non-interactive arguments of knowledge）是经典的非交互式零知识证明算法，被应用在匿名数字通证 ZCash 的匿名交易环节。zk-SNARK 的核心技术与同态加密、多方安全计算有一定的联系。ZCash 匿名交易的发送方可以在不泄露交易的金额、地址等细节的前提下向验证者证明交易的合法性。

零知识证明是密码学的高级应用，但并不是所有问题都可以通过零知识证明算法进行验证。Goldreich, Micali 和 Wigderson 证明了，多项式时间内可以验证解正确性的问题（NP 问题）一定存在零知识证明算法。

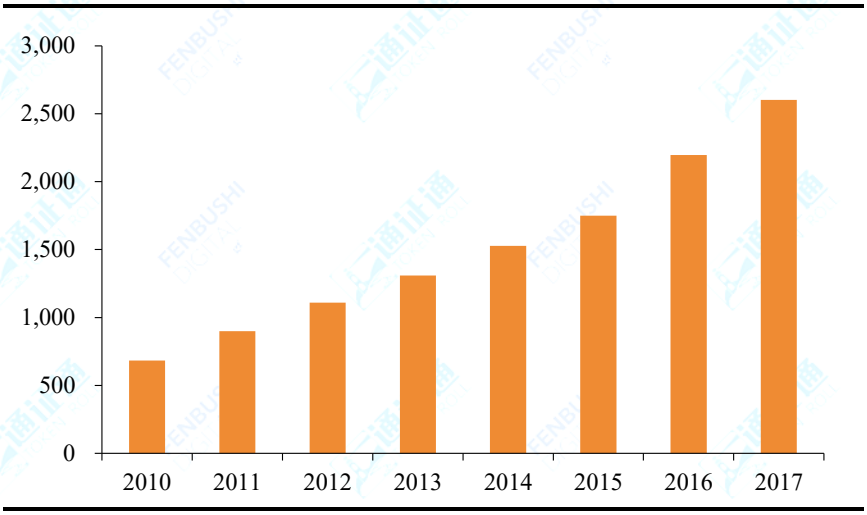
3 应用前景

目前，隐私计算技术效率较低，实际的商业应用较少。随着技术的不断发展，云计算、区块链乃至分布式计算网络逐渐走入了人们的视野，人们对数据安全的重视为隐私计算的应用提供了迫切需求。隐私计算在安全云计算、分布式计算网络和加密区块链三个方向将有较广的应用前景。

3.1 安全云计算

云计算极大的提高了算力资源的利用效率，市场规模增长迅速，具有广阔的发展前景。随着信息技术的发展，网络通信速度和服务器计算能力逐步提高，云计算产业发展迅速。一方面，云计算服务通过出租计算资源的形式，避免了个人和企业重复购买硬件设备的资源浪费，人们可以不再拥有很多使用频率较低的硬件，转而租用随时待机的线上计算资源，服务提供商的硬件也可以保持极高的使用率，提高了资源配置的效率；另一方面，云计算服务提供商拥有大量的计算硬件，可以为这些设备提供适宜的运行环境和及时的维护，数据中心的建设将极大降低单位计算资源的成本，发挥规模效应。云计算的这些优势使其近年来高速发展，目前已经形成约 2602 亿美元的市场规模。云计算技术受到了各行业的普遍认同，具有非常丰富的应用场景和广阔的发展空间，将会极大地改变计算资源的利用方式。

图表4：云计算行业规模发展迅速（亿美元）



资料来源：Gartner，通证通研究院

但是，使用云计算服务不可避免地牺牲隐私性。目前在使用云计算服务的过程中，用户数据以明文的形式存储在云服务提供商处。一方面，数据会不可避免地被云计算提供商获取，虽然有相关法规保护，但普通用户在与寡头企业的博弈中始终处于不利地位；另一方面，虽然有安全措施，但是云端服务器上的数据仍然有可能被内鬼和黑客获取。云计算虽然提高了计算资源的使用效率，但数据安全性无法得到保障。

隐私计算可以让数据在云计算过程中保持加密状态，避免被服务提供商和其他第三方获取，极大地提高了云计算过程中的数据安全。传统密码学无法在数据计算环节保证安全性，而隐私计算技术填补了这一空白，使得包括云计算在内的各种数据合作变得更加安全和隐私。机密数据可以通过全同态加密算法交由云计算服务商计算，而无需直接传输明文数据。通过安全多方计算，数据拥有者之间可以放心地进行合作计算，而无需担心自己的数据会被合作者或第三方获取。隐私计算将极大的扩大云计算的应用场景。

3.2 分布式计算网络

分布式计算网络是云计算的高级形态，利用个体广泛的计算资源，形成分布式的运算网络。随着个人计算设备性能的提高和互联网节点接入数的增加，网络计算能力的冗余越来越多。人们闲置的手机、计算机等设备能够通过分布式计算网络重新被有效利用。随着通信技术的不断发展，网络的带宽逐渐提高让分布式计算网络的建立从理论逐渐走向现实。

分布式计算网络能防止云计算寡头垄断计算资源。云计算的规模效应非常明显，发展到后期必然形成寡头垄断格局，人们的使用习惯形成后，这些企业将提高价格，获取高额垄断利润，公众难以充分享受到效率提升所带来的利益。分布式计算网络不同于云计算，是对存量计算资源的利用，个体不放弃对计算资源的所有权，只是在网络上进行互助与资源共享。企业可能持有较高的计算能力，但与分布式计算网络所整合的人类全体计算设备的力量相比，很难形成垄断性的优势。任何企图获取垄断利润的行为，都是与全世界的算力作对，不具备经济合理性。

隐私计算为分布式计算网络提供了信息安全的保证。分布式计算网络很可能会以一种点对点的形式存在，节点间直接进行计算资

源的分配，其中的安全问题尤为重要。这样的形式下，数据将直接从计算资源的需求节点发送到不受信任的提供节点，数据安全相比云计算会面临更大的挑战。如果可以使用隐私计算，就可以从根本上解决分布式计算网络的信息安全问题，用户一方面可以安全地使用他人的计算资源，另一方面也可以在不公开数据的前提下进行计算合作，将极大的丰富分布式计算网络的应用场景。

3.3 加密链上数据和隐藏交易信息

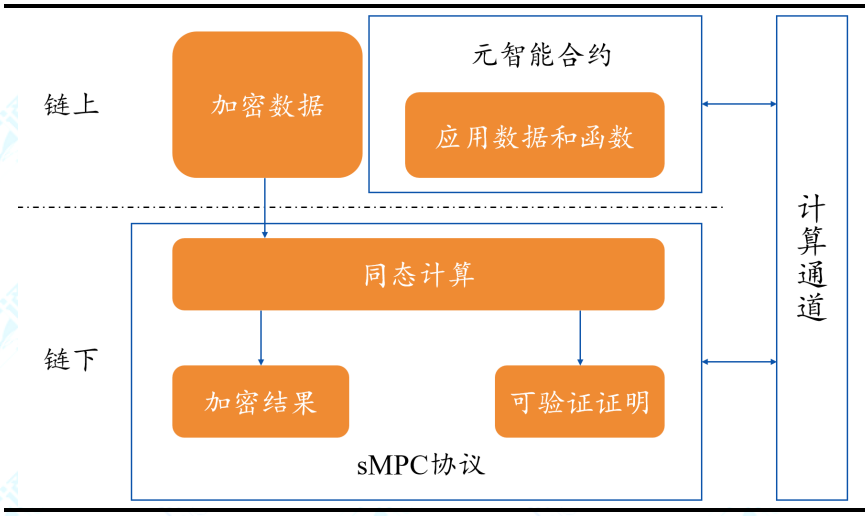
区块链可以提供可信的分布式记账，极大的降低了信任成本，但面临链上数据可验证和隐私性的两难。区块链通过共识机制实现了不可篡改的分布式记账，为价值互联网提供了一种可靠的全新信任形式，将极大的促进线上资产的发展，并以信任为切入点改变互联网乃至经济社会的生态。但是，链上的信息对所有人来说都是公开可查的，隐私性有很大牺牲。

对数据隐私的担忧限制了智能合约的应用范围。运行在区块链上的智能合约将区块链带入了 2.0 时代，实现了支付结算以外的诸多功能。但是，智能合约链上运行带来的可信性同样要以牺牲隐私性为代价，所有的处理记录都可以在区块链上被查找，其应用范围也会受数据敏感性的限制。

隐私计算使隐私数据上链成为可能，并且同样通过区块链技术确保其可验证性。通过使用隐私计算技术，数据可以在保持加密状态的进行确认、处理和计算，信息上链的可信性和隐私性之间的两难将会得到解决。人们通过全同态加密、安全多方计算和零知识证明等技术，在不泄露原始敏感数据的前提下使用区块链进行信息验证。

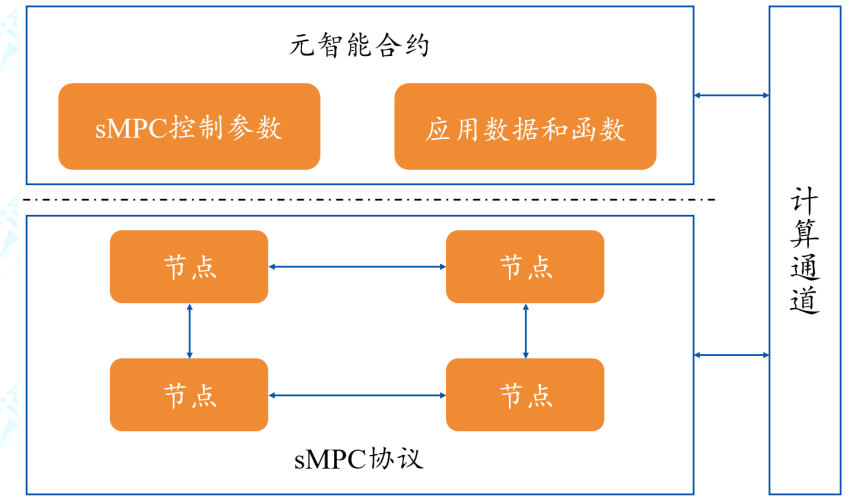
隐私计算的效率相比常规计算更低。隐私计算与区块链的结合，将解决数据可验证性与隐私性的问题，极大地拓展区块链的使用场景。但走向商业落地仍需要有硬件计算能力或技术突破。

图表5：全同态加密在区块链中的应用



资料来源：通证通研究院

图表6：安全多方计算在智能合约中的应用



资料来源：通证通研究院

附注：

因一些原因，本文中的一些名词标注并不是十分精准，主要如：通证、数字通证、数字 currency、货币、token、Crowdsale 等，读者如有疑问，可来电来函共同探讨。

免责声明

本报告由通证通研究院和FENBUSHI DIGITAL提供，仅供通证通研究院和FENBUSHI DIGITAL客户使用。本报告仅在相关法律许可的情况下发放，所提供信息均来自公开渠道。通证通研究院和FENBUSHI DIGITAL尽可能保证信息的准确、完整，但不对其准确性或完整性做出保证。

本报告的完整观点应以通证通研究院和FENBUSHI DIGITAL发布的完整报告为准，任何微信订阅号、媒体、社交网站等发布的观点和信息仅供参考，通证通研究院和FENBUSHI DIGITAL不会因为关注、收到或阅读到报告相关内容而视相关人员为客户。

本报告所载的资料、意见及推测仅反映通证通研究院和FENBUSHI DIGITAL于发布本报告当日的判断，相关的分析意见及推测可能会根据后续发布的研究报告在不发出通知的情形下做出更改，投资者应当自行关注相应的更新或修改。

市场有风险，投资需谨慎。本报告中的信息或所表述的意见仅供参考，不构成对任何人的投资建议。投资者不应将本报告为作出投资决策的唯一参考因素，亦不应认为本报告可以取代自己的判断，通证通研究院或FENBUSHI DIGITAL、通证通研究院或FENBUSHI DIGITAL员工或者关联机构不承诺投资者一定获利，不与投资者分享投资收益，也不对任何人因使用本报告中的任何内容所引致的损失负责。

通证通研究院和FENBUSHI DIGITAL对本免责声明条款具有修改和最终解释权。