在介绍比特币的运行原理之前，必须首先厘清以下六个重要的基本概念：散列、工作量证明、公开密钥密码体系、交易、区块与挖矿。

（1）散列（Hash）

在计算机科学中，Hash 通常被翻译为“散列”。散列函数的功能是将任意长度的不同信息（例如数字、文本或其他信息）转化为长度相等但内容不同的二进制数列（由 0 和 1 组成）。以比特币采用的 SHA256 为例，任意长度的信息输入通过这个函数都可以转换成一组长度为 256 个的二进制数字，以便统一的存储和识别。256 个 0 或 1 最多可以组合成 2256 个不同的数，这个庞大的集合能够满足与比特币相关的任何标记需要。此外，任意两个不同的信息输入，想要通过 SHA256 产生相同数字输出的概率，可以说微乎其微。因为输入信息的微小变动将会导致输出数字的巨大变化。这就保证了输入信息与输出数字的一一对应。最后，散列还有一个重要特征，即想要通过输出数字来反推出输入信息，这是极其困难的。因此，如果想要生成一个特殊的输出数字，就只能通过随机尝试的办法逐个进行正向运算，而不能由输出结果逆向推出输入信息。这个特征是比特币能够顺利运行的重要基石。

（2）工作量证明（Proof-of-Work）

倾注了更多更复杂劳动的事物具有更高的价值，这是比特币运行的哲学基础。让我们先以防范垃圾邮件为例来说明什么是工作量证明。不妨做出如下假定，即如果一个人愿意花 10 分钟写一封邮件，他就不会在意再多花一分钟对其进行处理，以证明自己写邮件付出的努力是真实的。而对垃圾邮件的传播者而言，每封邮件都要多花一分钟才能发送，这是完全不能接受的。因此我们可以设立以下规则，即在每次发送邮件之前都要算出一个随机数，以至于将这个随机数和邮件内容一起输入 SHA256 散列函数时，得到的 256 位二进制数的前 10 位均为 0。如前所述，我们无法预先选择一个前十位为 0 的数，并利用 SHA256 算法反推出这个随机数是什么。唯一可行的办法只能是随机抽取一个数，将其和邮件内容放入 SHA256 中进行计算，看结果是否满足要求。如果不满足，就换一个随机数继续进行尝试，直到要求满足为止。只要我们设定的要求足够简单（要求全为 0 的个数不太多），那么寻找这个随机数的过程也就比较简单，只不过要花去一定的时间（例如几秒或几分钟）。对于真实的邮件而言，为了证明自身价值，付出少量时间进行计算是值得的。但对于垃圾邮件而言，这将导致邮件发送者的时间成本急剧上升。因此，上述机制的引入将会显著减少垃圾邮件的产生。对比特币而言，

挖矿（Mining）也是使用随机数进行工作量证明的过程。这种过程虽然从表面上来看没有产生任何价值，但却是解决互联网中信任问题的有效办法，是在不可靠的网络环境中一种较为可靠的信用证明。

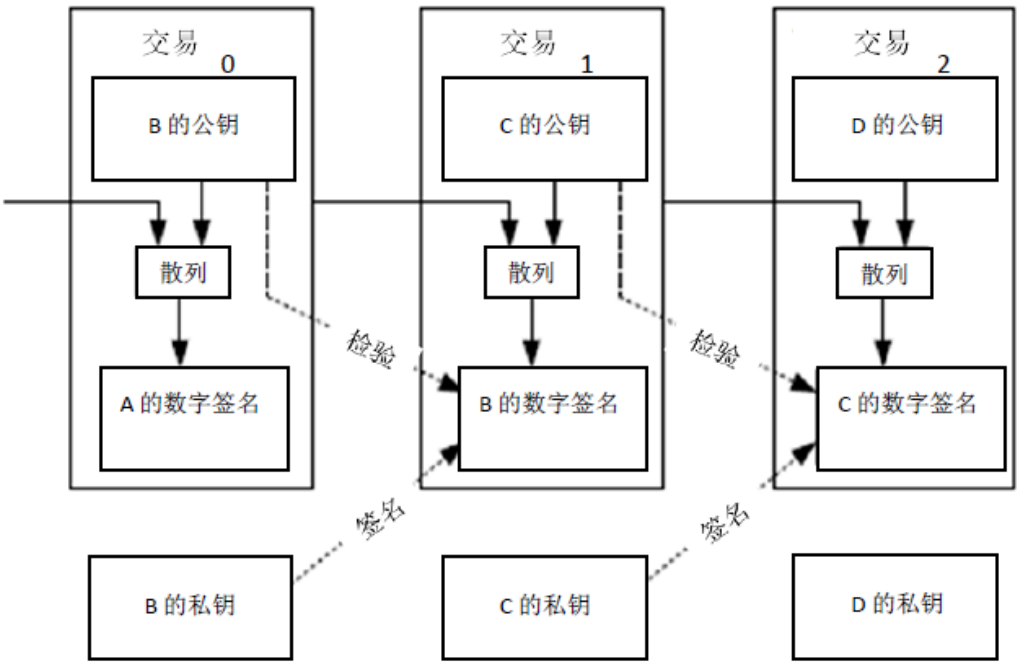
（3）公开密钥密码体系

该体系简称公钥体系。在信息传递过程中，发送方通过一把密钥将信息加密，接收方在收到信息后，再通过配对的另一把密钥对信息进行解密，这就保证了信息传递过程的私密性与安全性。而密钥无非是一组数字，通过将原始信息与这组数字放在一起进行特定运算，就能够把信息转换为另外一种格式，从而实现加密。解密过程则刚好相反。在大多数情况下，一组密钥由公钥和私钥组成。私钥由自己保存，公钥则需要向其他人公开。在信息传递过程中，公钥和私钥相互配合，既能够对持有私钥的发信人进行身份验证，也能够确保发信人对自己发出的信息不能抵赖，还能够保证收发信息的完整性、防止中间环节被截获篡改。如果公钥丢失，还可以通过私钥进行恢复。但试图通过公钥反推出私钥的努力，从理论上来讲是基本不可行的，这就保证了私钥的私密性。

（4）交易（Transactions）

交易是指一个用户用比特币向另一个用户进行支付的过程。不过，比特币的交易并非简单的支付货币本身。以图 1 中的交易 1 为例，如果 B 想支付 100 个比特币（100BTC）给 C，那么 B 不仅需要在交易单上注明金额，而且需要注明这 100 个比特币的来源。如图 1 所示，B 的 100BTC 其实来自 A，是 B 通过交易 0 得到的（交易 0 已经通过了全网用户的认证，保存在所有用户的电脑中）。为完成交易 1，B 需要在交易单上填写的信息包括：一是 100BTC 的来源，此处为交易单 0 的 ID；二是 C 的公钥，也即 C 的比特币收款地址；三是将交易单 0 的内容和 C 的公钥输入散列函数，得到一串数字。B 用自己的私钥加密这串数字，作为数字签名放在交易单 1 中。C 在收到交易单 1 之后，可以通过其中存放的 ID 找到交易单 0，并获取 B 的公钥。C 可以使用该公钥对交易单 1 中的数字签名进行解密。与此同时，C 可以把自己的公钥和交易单 0 的内容，按照同样的方式输入散列函数，并将得到的数字与数字签名解密的结果进行比对。如果比对成功，就可以确定如下两个事实：其一，100BTC 的来源属实。因为交易单 0 中包含了 A 的签名，且交易单 0 是经过全网认证过的，即 A 确实将 100BTC 给了 B；其二，交易 1 的确是经由 B 签署的。由于 B 的私钥是唯一的，他无法抵赖这单交易。

图 1 比特币交易过程



资料来源：Nakamoto（2008）

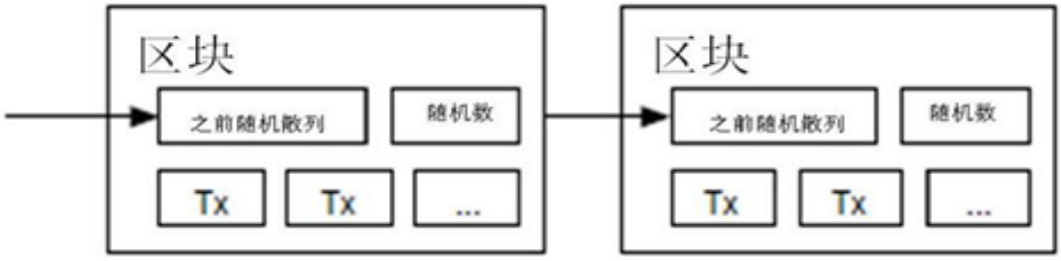
上述过程略显复杂。我们可以换一种不太精确但更容易理解的解释。依然以交易 1 为例，交易单 1 中其实包含以下六种信息：一是交易单 1 的 ID；二是资金的来源，即交易单 0 的 ID；三是 A 对资金的签名，以证明是他把 100BTC 给 B 的；四是资金的去向，即 C 的账号（公钥）；五是资金的数额，即 100BTC；六是 B 的签名（即 B 用自己私钥进行的数字签名），以证明是他自己签发的交易。由于每笔交易单都记录了该笔资金的前一个拥有者、当前拥有者以及后一个拥有者，我们就可以依据交易单实现对资金的全程追溯。这也是比特币的典型特征之一。最后，当每一笔交易完成时，系统都会向全网进行广播，告诉所有用户这笔交易的实施。

（5）区块（Block）

交易和区块的关系，就如同水和瓶子，属于内容和容器的关系。由于每笔交易是相对分散的，为了更好地统计交易，比特币系统创造了区块这一概念。每个区块均包含以下三种要素：一是本区块的 ID（散列）；二是若干交易单；三是前一个区块的 ID（散列）。比特币系统大约每十分钟创建一个区块，其中包含了这

段时间里全球范围内发生的所有交易。每个区块中也包含了前一个区块的 ID，这种设计使得每个区块都能找到其前一个节点，如此可一直倒推至起始节点，从而形成了一条完整的交易链条（图 2）。因此，从比特币的诞生之日起，全网就形成一条唯一的主区块链（BlockChain），其中记录了从比特币诞生以来的所有交易记录，并以每十分钟新增一个节点的速度无限扩展。这条主区块链在每添加一个节点后，都会向全网广播，从而使得每台参与比特币交易的电脑上都有一份拷贝。在现实世界里，每笔非现金交易都由银行系统进行记录，一旦银行计算机网络崩溃，所有数据都会遗失。而在互联网世界里，比特币的所有交易记录都保存在全球无数台计算机中，只要全球有一台装有比特币程序的计算机还能工作，这条主区块链就可以被完整地读取。如此高度分散化的交易信息存储，使得比特币主区块链完全遗失的可能性变得微乎其微。

图 2 区块链的局部结构



资料来源：Nakamoto（2008）。

（6）挖矿（Mining）

如前所述，比特币的所有交易记录都保存在主区块链中。每十分钟就会有一个新区块生成并加入进主区块链，这个新区块中记录了十分钟内全网的所有交易。由于比特币使用的是 P2P 模式，这意味着网络上的每个节点都是平等的，没有一个中心节点可以用来承担交易记录工作。因此，如此重要的交易记录任务交给谁来完成，就变成一个现实问题。而比特币创始人中本聪给出的答案居然是任何人来完成都可以。由于每笔交易完成后都会被广播给全网，因此每个人在对交易的有效性进行验证后，都可以根据这些交易数据生成新区块。但这又引发了一个新问题，即如何让所有人都信任由一个陌生人生成的新区块？这个新区块中是否记录了虚假交易或重复交易？

要解决这个问题，就要用到前文提到的工作量证明概念。基本思路是，寻找一个随机数，使得将这个数字与新区块的交易信息一起输入 SHA256 后产生的数字，前面 n 位（比如 n=100）都是 0。此项工作的意义在于，由于将会耗费很多

时间，如果一个人进行了这项计算且获得成功，那么他提供的区块很可能是真实可信的，因为花费如此大力气作假得到的好处，远远不计花费同样努力从事真实工作得到的好处。此外，其他所有节点在接收到新区块时，也会对其中包含交易的有效性进行校验，这意味着虚假交易或重复交易很难骗过其他所有用户，这就形成了节点之间的信用保障机制。

挖矿（Mining）就是指产生新区块并计算随机数的过程。具体过程可分为以下六步：第一步，由于网络上的每台计算机都保存有之前的主区块链，某台计算机以其中最后一个区块的内容为输入，计算一个散列值；第二步，该计算机在接收广播来的交易单并逐笔校验交易的准确性之后，把没有被列入之前区块的那些交易进行组合，并纳入一个新区块；第三步，该计算机任意猜一个随机数，其大小和长度没有限制；第四步，该计算机将第一步至第三步产生的数据作为输入，一起放到 SHA256 散列函数中，计算得到一个长度为 256 的二进制数；第五步，检查这个二进制数的前 n 位是否符合要求；第六步，如果该二进制数符合要求，则本轮游戏结束，该计算机会把新区块连同这个幸运随机数一起广播给网络上的其他计算机。其他人在收到这个新区块后，会以同样的方式进行校验。如果结果无误，全网就接受这个新区块，将它连同之前的主区块链一起保存。如果产生的随机数不合要求，则第二步至第六步就会重复进行，直到自己成功或者收到别人发来的新区块。

从上述流程中可以看出，挖矿就是指搜集交易数据并建立新区块的过程。这个过程虽然重要，却耗时费力，为什么所有参与者都趋之若鹜呢？最重要的原因在于，比特币系统规定，每个成功建立新区块的人都将获得 50 个新比特币的奖励，且该奖励将被记录在对应的新区块里。这 50 个新比特币是系统自动产生的，且得到全网的认同。有趣的是，这种奖励的数额每四年减半，即 2009 年至 2012 年年为每区块 50 个比特币、2013 年至 2016 年为每区块 25 比特币、2017 年至 2021 年为每区块 12.5 比特币，如此不一而足。最终，全系统的比特币容量将达到 2100 万个的上限，至此不再增加。从那时起，为保证主区块链能继续不断增长以确保比特币交易能继续正常进行，每个创建新区块的人，都将从新区块包含的交易单中抽取一定的“交易税”作为奖励。这种新的激励机制将保证比特币交易得以延续。