An Analysis of Bitcoin and the Blockchain

Zachary Waterman

December 11, 2015

Understanding Bitcoin and the blockchain technology underlying it is a forbidding task for most laymen. Thousands of articles have been written purporting to explain the inner workings of the Bitcoin protocol. Most gloss over major details and fail to recognize the brilliance and potential of Bitcoin?s architecture. This paper attempts illuminate the fuzzier details of the Bitcoin architecture using language that is at once comprehensive and approachable, discuss the public furor elicited by Bitcoin, and illustrate the potential of Bitcoin and the Blockchain to alter human life and institutions.

To begin, what is Bitcoin? Bitcoin is a form of digital currency, created and held electronically. No one controls it. Bitcoins aren?t printed, like dollars or euros ? they?re produced by people, and increasingly businesses, running computers all around the world, using software that solves mathematical problems. It?s the first example of a growing category of money known as cryptocurrency.A software developer called Satoshi Nakamoto proposed bitcoin as an electronic payment system based on mathematical proof. The idea was to produce a currency independent of any central authority, transferable electronically, more or less instantly, with very low transaction fees.

At the heart of bitcoin is a fundamental innovation: a distributed public ledger. A ledger in accounting is a book that you cannot edit once you have written in it. Instead, if you have made a mistake, the only way to fix it is to add another transaction to the ledger that undoes the error. As we know from accounting fraud, problems arise when people figure out ways to transact without recording it in the ledger or making ex post changes to the ledger. The bitcoin ledger is the so-called blockchain which uses the fact that there are many copies of it that are broadly distributed combined with a fair bit of math to ensure that once a transaction has been recorded in the blockchain that transaction can not be changed after the fact. There is no other widely used protocol in the world today that accomplishes this: with bitcoin anyone can make a statement (a transaction) and have this be recorded in a globally visible and fixed ledger.

In order to comprehend the complexity of the technology, one must understand a few terms and processes that laid the foundation for the blockchain. First, what is cryptography?the field that underlies the blockchain? Cryptography is the practice and study of techniques for secure communication. Vital components of information security such as data confidentiality, data integrity, and data authentication are products of modern

cryptography. It exists at the intersection of mathematics, computer science, and electrical engineering. Present applications of cryptography include ATM cards, computer passwords, and nearly all electronic commerce. One of the most relevant components of cryptography to Bitcoin and the blockchain are hash functions.

Hash functions allow for the bulk of modern cryptography and are, therefore, often referred to as ?the workhorses of modern cryptography?. A hash function is any function that can be used to convert data of arbitrary size to data of a fixed size. The values returned by hash functions are called hash values or simply hashes. Cryptographic hash functions, in particular, are hash function, which are considered practically impossible to invert?to recreate the input data from its hash value alone.

There are four vital components to any valid cryptographic hash function. First, it is trivial to compute the hash value for any given message. Second, it is infeasible to generate the message from its hash. Third, it is infeasible to modify a message without changing its hash. Fourth, it is infeasible to find two different messages with the same hash.

SHA-256 hashes, in particular, are used to run the blockchain. They are a set of cryptographic hash functions designed by the NSA. SHA stands for Secure Hash Algorithm. Cryptographic hash functions are mathematical operations run on digital data; by comparing the computed "hash" (the output from execution of the algorithm) to a known and expected hash value, a person can determine the data's integrity.There are myriad applications of cryptographic hash functions, but public key cryptography is the most relevant to this paper.

Asymmetric cryptography or public-key cryptography is a branch of the cryptography field, that fuels the blockchain and all of its present and future applications, including Bitcoin. Public-key cryptography is a process by which a pair of keys is used to encrypt?convert information or data into a cipher or code, especially to prevent unauthorized access?and decrypt a message so that it arrives securely. A user of public-key cryptography will receive a public and a private key from the program orchestrating the process. These public keys are available to any user via a public directory. Any user, who wants to send an encrypted message to the owner of a public key, uses that public key to encrypt their message and then send it to the owner. When the recipient receives the encrypted message, it is easily decrypted with their private key, which no one else has access to.This process is often referred to as Pretty Good Privacy or PGP.

A digital signature, a technology birthed from public key-cryptography, is a mathematical technique used to validate the authenticity and integrity of a message or other digital document. It is the digital equivalent of a handwritten signature, but offers far more inherent security. Digital signatures were created with the intention of eliminating the problems of tampering and impersonation in digital communication. The process is relatively simple. A public key algorithm is used to generate two keys that are mathematically linked: one private and one public. Then, a digital signature is created using signing software, which creates a hash of the electronic data to be signed. The user?s private key is then used to encrypt the has. The encrypted hash and hashing algorithm compose the digital signature.

Critically, the value of the hash is unique to the hashed data. Any change in the data, even changing or deleting a single character, results in a different hash value. This attribute enables others to validate the integrity of the data by using the signer?s public key to decrypt the hash. If the decrypted has matches a second computed hash of the same data, it proves that the data hasn?t changed since it was signed. If the two hashes don?t match, the data has either been tampered with in some way or the signature was created with a private key that doesn?t correspond to the public key presented by the signer.

POW The second problem solved by Satoshi was to create a Proof of Work (POW) System. A Proof of Work System includes measures to deter attacks such as network spam by requiring some work from the service requester, usually meaning processing time by a computer. The concept was first presented by Cynthia Dwork and Moni Naor in a 1993 journal. A critical feature of a POW system is asymmetry: the work must be moderately hard (but feasible) on the requester side but easy to check for the service provider. His implementation of a POW system renders Satoshi?s invention far more secure.

A proof of work is a piece of data which is difficult (costly, time-consuming) to produce but easy for others to verify and which satisfies certain requirements. Producing a proof of work can be a random process with low probability so that a lot of trial and error is required on average before a valid proof of work is generated.

Bitcoin uses the Hashcash proof of work. Hashcash proofs of work are used in Bitcoin for block generation. In order for a block to be accepted by network participants, miners must complete a proof of work which covers all of the data in the block. The difficulty of this work is adjusted so as to limit the rate at which new blocks can be generated by the network to one every 10 minutes. Due to the very low probability of successful generation, this makes it unpredictable which worker computer in the network will be able to generate the next block.

For a block to be valid it must hash to a value less than the current target; this means that each block indicates that work has been done generating it. Each block contains the hash of the preceding block, thus each block has a chain of blocks that together contain a large amount of work. Changing a block (which can only be done by making a new block containing the same predecessor) requires regenerating all successors and redoing the work they contain. This protects the block chain from tampering.

Using these technologies, Nakamoto fashioned a solution to a number of thorny problems to create the blockchain network. The first is the Byzantine General?s Problem. In the simplest terms, the Byzantine General?s Problem is an agreement protocol involving an imaginary General who makes a decision to attack or retreat, and who must communicate his decision to his lieutenants. An unknown number of these actors?the lieutenants and possibly the General?are traitors, who cannot be relied upon to properly communicate orders. In fact, these traitors may purposefully alter messages in an effort to subvert the process.

A solution to the General?s Problem must pass three tests. First, a solution must guarantee that all Generals eventually reach a decision regarding the value of the order

they have been given. This test is known as termination. Second, all Generals must decide on the same value of the order they have been given. This test is referred to as agreement. Third, all Generals must decide on the value that was sent by the original general. This test is known as validity. Thus, even if the original General is subversive, all lieutenants still have to come to a common, unanimous decision. The problem itself is, therefore, extremely difficult to solve.

The second dilemma solved Nakamoto is double-spending, the risk that a digital currency can be spent twice. Double-spending is a problem unique to digital currencies because digital information can be reproduced relatively easily. Physical currencies do not have this issue because they cannot be easily replicated, and the parties involved in a transaction can immediately verify the bona fides of the physical currency. With digital currency, there is a risk that the holder could make a copy of the digital token and send it to a merchant or another party while retaining the original.