Double spend races

Ricardo Pérez-Marco (CNRS, IMJ-PRG, Paris 7)

(joint work with Cyril Grunspan (ESILV, Paris))

Finance Seminar (Stony Brook)

October 20, 2017

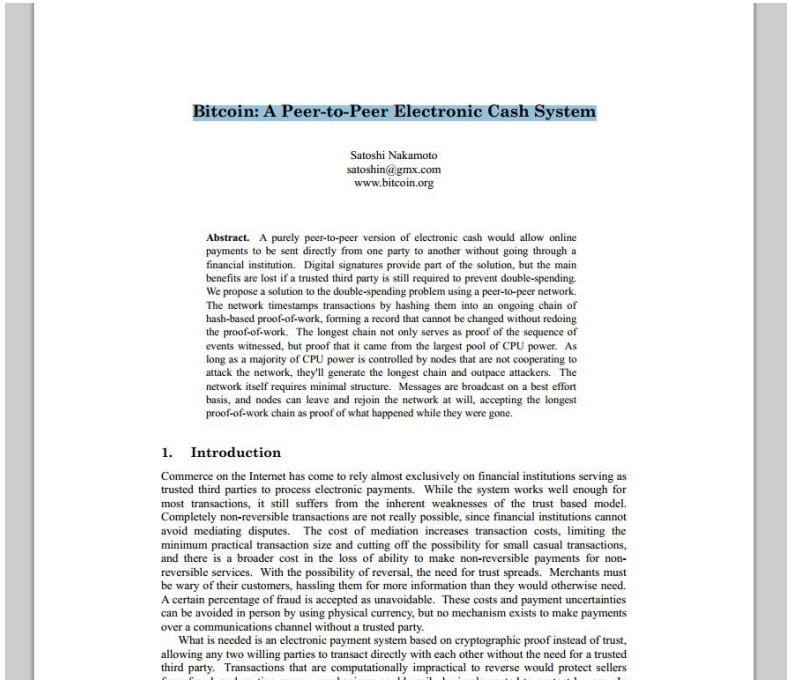
"Double spend races" (with C. Grunspan), ArXiv 1702.02867, 2017.

Double spend races

Bitcoin paper

S. Nakamoto, November 1st 2008,

“Bitcoin: A peer-to-peer electronic cash system”



Bitcoin: A Peer-to-Peer Electronic Cash System

Satoshi Nakamoto
satoshin@gmx.com
www.bitcoin.org

Abstract. A purely peer-to-peer version of electronic cash would allow online payments to be sent directly from one party to another without going through a financial institution. Digital signatures provide part of the solution, but the main benefits are lost if a trusted third party is still required to prevent double-spending. We propose a solution to the double-spending problem using a peer-to-peer network. The network timestamps transactions by hashing them into an ongoing chain of hash-based proof-of-work, forming a record that cannot be changed without redoing the proof-of-work. The longest chain not only serves as proof of the sequence of events witnessed, but proof that it came from the largest pool of CPU power. As long as a majority of CPU power is controlled by nodes that are not cooperating to attack the network, they'll generate the longest chain and outpace attackers. The network itself requires minimal structure. Messages are broadcast on a best effort basis, and nodes can leave and rejoin the network at will, accepting the longest proof-of-work chain as proof of what happened while they were gone.

1. Introduction

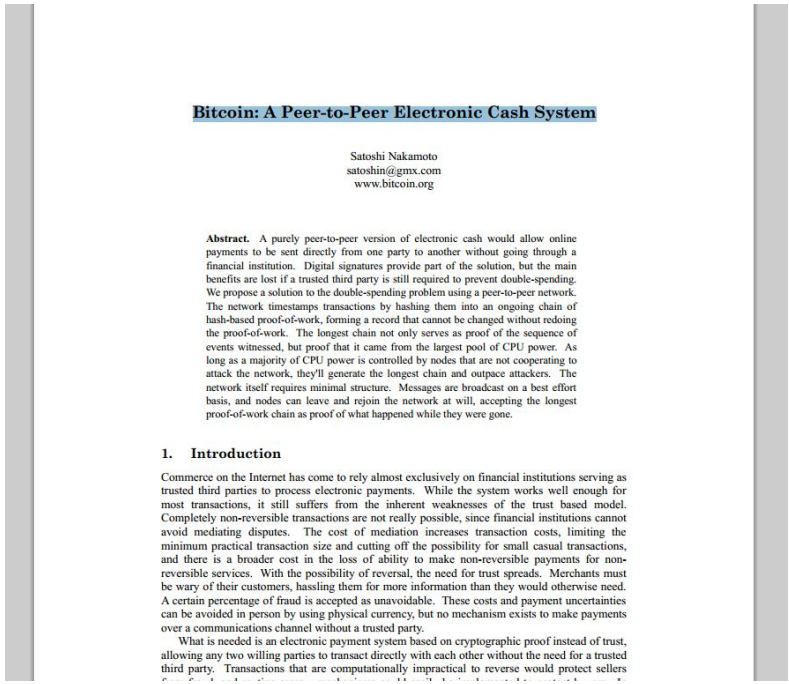
Commerce on the Internet has come to rely almost exclusively on financial institutions serving as trusted third parties to process electronic payments. While the system works well enough for most transactions, it still suffers from the inherent weaknesses of the trust based model. Completely non-reversible transactions are not really possible, since financial institutions cannot avoid mediating disputes. The cost of mediation increases transaction costs, limiting the minimum practical transaction size and cutting off the possibility for small casual transactions, and there is a broader cost in the loss of ability to make non-reversible payments for non-reversible services. With the possibility of reversal, the need for trust spreads. Merchants must be wary of their customers, hassling them for more information than they would otherwise need. A certain percentage of fraud is accepted as unavoidable. These costs and payment uncertainties can be avoided in person by using physical currency, but no mechanism exists to make payments over a communications channel without a trusted party.

What is needed is an electronic payment system based on cryptographic proof instead of trust, allowing any two willing parties to transact directly with each other without the need for a trusted third party. Transactions that are computationally impractical to reverse would protect sellers

Bitcoin paper

S. Nakamoto, November 1st 2008,

“Bitcoin: A peer-to-peer electronic cash system”



Bitcoin: A Peer-to-Peer Electronic Cash System

Satoshi Nakamoto
satoshin@gmx.com
www.bitcoin.org

Abstract. A purely peer-to-peer version of electronic cash would allow online payments to be sent directly from one party to another without going through a financial institution. Digital signatures provide part of the solution, but the main benefits are lost if a trusted third party is still required to prevent double-spending. We propose a solution to the double-spending problem using a peer-to-peer network. The network timestamps transactions by hashing them into an ongoing chain of hash-based proof-of-work, forming a record that cannot be changed without redoing the proof-of-work. The longest chain not only serves as proof of the sequence of events witnessed, but proof that it came from the largest pool of CPU power. As long as a majority of CPU power is controlled by nodes that are not cooperating to attack the network, they'll generate the longest chain and outpace attackers. The network itself requires minimal structure. Messages are broadcast on a best effort basis, and nodes can leave and rejoin the network at will, accepting the longest proof-of-work chain as proof of what happened while they were gone.

1. Introduction

Commerce on the Internet has come to rely almost exclusively on financial institutions serving as trusted third parties to process electronic payments. While the system works well enough for most transactions, it still suffers from the inherent weaknesses of the trust based model. Completely non-reversible transactions are not really possible, since financial institutions cannot avoid mediating disputes. The cost of mediation increases transaction costs, limiting the minimum practical transaction size and cutting off the possibility for small casual transactions, and there is a broader cost in the loss of ability to make non-reversible payments for non-reversible services. With the possibility of reversal, the need for trust spreads. Merchants must be wary of their customers, hassling them for more information than they would otherwise need. A certain percentage of fraud is accepted as unavoidable. These costs and payment uncertainties can be avoided in person by using physical currency, but no mechanism exists to make payments over a communications channel without a trusted party.

What is needed is an electronic payment system based on cryptographic proof instead of trust, allowing any two willing parties to transact directly with each other without the need for a trusted third party. Transactions that are computationally impractical to reverse would protect sellers

“Bitcoin and Decentralized Trust Protocols”, Newsletter European Math. Soc., 100, June 2016. ArXiv 1601.05254.

Transparency Theorem

- Bitcoin is an electronic currency modelled by gold.

Transparency Theorem

- Bitcoin is an electronic currency modelled by gold.
- Bitcoin is independent of any central authority.

Transparency Theorem

- Bitcoin is an electronic currency modelled by gold.
- Bitcoin is independent of any central authority.
- Bitcoin protocol runs on open software.

Transparency Theorem

- Bitcoin is an electronic currency modelled by gold.
- Bitcoin is independent of any central authority.
- Bitcoin protocol runs on open software.
- To avoid the "**double spend problem**" Bitcoin relies on a public ledger.

Transparency Theorem

- Bitcoin is an electronic currency modelled by gold.
- Bitcoin is independent of any central authority.
- Bitcoin protocol runs on open software.
- To avoid the "**double spend problem**" Bitcoin relies on a public ledger. This is general and necessary:

Theorem

Transparency Theorem: *An electronic decentralized currency must rely on a public ledger.*

The blockchain

- The public ledger is an incorruptible public database of all transactions called “**the blockchain**”.

The blockchain

- The public ledger is an incorruptible public database of all transactions called “**the blockchain**”.
- Anyone can write in the blockchain.

The blockchain

- The public ledger is an incorruptible public database of all transactions called “**the blockchain**”.
- Anyone can write in the blockchain.
- Anyone can have a copy of the blockchain.

The blockchain

- The public ledger is an incorruptible public database of all transactions called “**the blockchain**”.
- Anyone can write in the blockchain.
- Anyone can have a copy of the blockchain.
- The blockchain is composed by a chronological sequence of cryptological chained blocks.

The blockchain

- The public ledger is an incorruptible public database of all transactions called “**the blockchain**”.
- Anyone can write in the blockchain.
- Anyone can have a copy of the blockchain.
- The blockchain is composed by a chronological sequence of cryptological chained blocks.
- Each block contains a set of transactions.

The blockchain

- The public ledger is an incorruptible public database of all transactions called “**the blockchain**”.
- Anyone can write in the blockchain.
- Anyone can have a copy of the blockchain.
- The blockchain is composed by a chronological sequence of cryptological chained blocks.
- Each block contains a set of transactions.
- Each new block is generated in about 10 minutes.

The blockchain

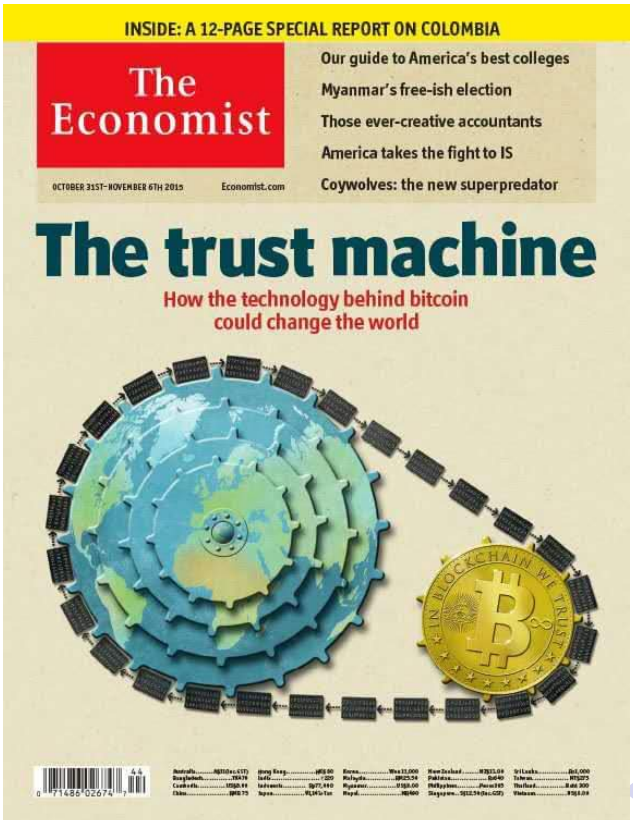
- The public ledger is an incorruptible public database of all transactions called “**the blockchain**”.
- Anyone can write in the blockchain.
- Anyone can have a copy of the blockchain.
- The blockchain is composed by a chronological sequence of cryptological chained blocks.
- Each block contains a set of transactions.
- Each new block is generated in about 10 minutes.
- The blocks are generated by “miners” that validate current transactions.

The Trust Machine

- The core of the Bitcoin protocol is the algorithm to ensure that this database cannot be forged.

The Trust Machine

- The core of the Bitcoin protocol is the algorithm to ensure that this database cannot be forged.
- The mechanism of consensus: “The trust machine”.



Nodes

- The Bitcoin Network is composed by nodes that communicate with each other.

Nodes

- The Bitcoin Network is composed by nodes that communicate with each other.
- Nodes check and broadcast transactions.

Nodes

- The Bitcoin Network is composed by nodes that communicate with each other.
- Nodes check and broadcast transactions.
- Some nodes are miners that validate transactions.

Nodes

- The Bitcoin Network is composed by nodes that communicate with each other.
- Nodes check and broadcast transactions.
- Some nodes are miners that validate transactions.
- Anyone can join and participate in the network.

Nodes

- The Bitcoin Network is composed by nodes that communicate with each other.
- Nodes check and broadcast transactions.
- Some nodes are miners that validate transactions.
- Anyone can join and participate in the network.
- To avoid Sybil attacks a “Proof of Work” (PoW) for miners is required.

Nodes

- The Bitcoin Network is composed by nodes that communicate with each other.
- Nodes check and broadcast transactions.
- Some nodes are miners that validate transactions.
- Anyone can join and participate in the network.
- To avoid Sybil attacks a “Proof of Work” (PoW) for miners is required.

Reaching Consensus

- The idea is to select randomly who validates the next block of transactions.

Reaching Consensus

- The idea is to select randomly who validates the next block of transactions.
- A decentralized “lottery” is set by the PoW.

Reaching Consensus

- The idea is to select randomly who validates the next block of transactions.
- A decentralized “lottery” is set by the PoW.
- A computationally intensive problem is set to validate a block.

Reaching Consensus

- The idea is to select randomly who validates the next block of transactions.
- A decentralized “lottery” is set by the PoW.
- A computationally intensive problem is set to validate a block.
- The problem is difficult to solve, but the solution is easy to check.

Reaching Consensus

- The idea is to select randomly who validates the next block of transactions.
- A decentralized “lottery” is set by the PoW.
- A computationally intensive problem is set to validate a block.
- The problem is difficult to solve, but the solution is easy to check.
- The difficulty is adjusted to find a solution in about 10 minutes.

Reaching Consensus

- The idea is to select randomly who validates the next block of transactions.
- A decentralized “lottery” is set by the PoW.
- A computationally intensive problem is set to validate a block.
- The problem is difficult to solve, but the solution is easy to check.
- The difficulty is adjusted to find a solution in about 10 minutes.
- The miner that solves it receives an award in newly created bitcoins.

Reversing a transaction

Reversing a transaction

- The attacker issues a second Tx with the same bitcoins to pay himself and invalidate a first Tx.

Reversing a transaction

- The attacker issues a second Tx with the same bitcoins to pay himself and invalidate a first Tx.
- Easy if the recipient does not wait for the Tx to be included in the blockchain.

Reversing a transaction

- The attacker issues a second Tx with the same bitcoins to pay himself and invalidate a first Tx.
- Easy if the recipient does not wait for the Tx to be included in the blockchain.
- Once included in the blockchhain, the attacker can only mine a longer blockchain from an earlier block where the Tx was included.

Reversing a transaction

- The attacker issues a second Tx with the same bitcoins to pay himself and invalidate a first Tx.
- Easy if the recipient does not wait for the Tx to be included in the blockchain.
- Once included in the blockchhain, the attacker can only mine a longer blockchain from an earlier block where the Tx was included.
- We have a mining race with the rest of the network.

Reversing a transaction

- The attacker issues a second Tx with the same bitcoins to pay himself and invalidate a first Tx.
- Easy if the recipient does not wait for the Tx to be included in the blockchain.
- Once included in the blockchhain, the attacker can only mine a longer blockchain from an earlier block where the Tx was included.
- We have a mining race with the rest of the network.
- The attacker must control a substantial fraction of the mining power of the network.

- This “double spend” attack is discussed in the last section of Nakamoto’s paper.

- This “double spend” attack is discussed in the last section of Nakamoto’s paper.

The recipient waits until the transaction has been added to a block and z blocks have been linked after it. He doesn't know the exact amount of progress the attacker has made, but assuming the honest blocks took the average expected time per block, the attacker's potential progress will be a Poisson distribution with expected value:

$$\lambda = z\frac{q}{p}$$

To get the probability the attacker could still catch up now, we multiply the Poisson density for each amount of progress he could have made by the probability he could catch up from that point:

$$\sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} \cdot \begin{cases} (q/p)^{(z-k)} & \text{if } k \le z \\ 1 & \text{if } k > z \end{cases}$$

Rearranging to avoid summing the infinite tail of the distribution...

$$1 - \sum_{k=0}^{z} \frac{\lambda^k e^{-\lambda}}{k!} \left(1 - (q/p)^{(z-k)}\right)$$

- Nakamoto makes the abusive simplification that the network mines blocks at average speed.

- Nakamoto makes the abusive simplification that the network mines blocks at average speed.
- Nakamoto computes numerically the probability of success of the attack.

- Nakamoto makes the abusive simplification that the network mines blocks at average speed.
- Nakamoto computes numerically the probability of success of the attack.
- Nakamoto empirically observes that it decreases exponentially with the number of validations.

- Nakamoto makes the abusive simplification that the network mines blocks at average speed.
- Nakamoto computes numerically the probability of success of the attack.
- Nakamoto empirically observes that it decreases exponentially with the number of validations.
- This is the crucial point giving security to the Bitcoin protocol!

Exact computation

Exact computation

- $0 < q < 1/2$ fraction of the mining power of the attacker.

Exact computation

- $0 < q < 1/2$ fraction of the mining power of the attacker.
- $p = 1 - q$ fraction of the mining power of the rest.

Exact computation

- $0 < q < 1/2$ fraction of the mining power of the attacker.
- $p = 1 - q$ fraction of the mining power of the rest.
- $z \geq 1$ number of validations of the first Tx.

Exact computation

- $0 < q < 1/2$ fraction of the mining power of the attacker.
- $p = 1 - q$ fraction of the mining power of the rest.
- $z \geq 1$ number of validations of the first Tx.

Theorem (C. Grunspan, R. Pérez-Marco, 2017)

The probability of success of the attacker is

$$P(z) = I_{4pq}(z, 1/2) \ ,$$

where $I_x(a,b)$ is the Regularized Incomplete Beta Function

$$I_x(a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^x t^{a-1}(1-t)^{b-1}\, dt \ .$$

An important corollary

An important corollary

Main security result for the Bitcoin Protocol.

An important corollary

Main security result for the Bitcoin Protocol.

Theorem

When $z \to +\infty$ *we have , with* $s = 4pq < 1$,

$$P(z) \sim \frac{s^z}{\sqrt{\pi(1-s)z}},$$

thus the probability $P(z)$ *decreases exponentially to* 0 *with* z

Practical implications

Practical implications

• Nakamoto discusses in function of q how many confirmations do we need to wait so that the probability of success of the attacker is less than 0.1%.

Practical implications

- Nakamoto discusses in function of q how many confirmations do we need to wait so that the probability of success of the attacker is less than 0.1%.
- The exact result proves that one needs to be more conservative.

Practical implications

- Nakamoto discusses in function of q how many confirmations do we need to wait so that the probability of success of the attacker is less than 0.1%.
- The exact result proves that one needs to be more conservative.

q	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
z	6	9	13	20	32	58	133	539
z_{SN}	5	8	11	15	24	41	81	340

A finer analysis

A finer analysis

- The probability of success of the attacker increases with the time τ it takes to validate the z blocks by the honest network

A finer analysis

- The probability of success of the attacker increases with the time τ it takes to validate the z blocks by the honest network
- τ is known to the recipient of the first Tx, so he can refine the estimate of the probability of success of the attacker.

A finer analysis

- The probability of success of the attacker increases with the time τ it takes to validate the z blocks by the honest network
- τ is known to the recipient of the first Tx, so he can refine the estimate of the probability of success of the attacker.
- We consider the dimensionless parameter that measures the deviation from average validation time

$$\kappa = \frac{\tau}{z t_0} \ ,$$

where t_0 is the average time of validation of blocks (10 minutes), and $P(z, \kappa)$ is the probability we study.

- $\kappa = 1$ corresponds to the situation studied by Nakamoto

$$P_{SN}(z) = P(z, 1) \ .$$

- $\kappa = 1$ corresponds to the situation studied by Nakamoto

$$P_{SN}(z) = P(z, 1) \ .$$

- We can also recover the exact probability $P(z)$ as a weighted average,

$$P(z) = \int_0^{+\infty} P(z, \kappa)\, d\rho_z(\kappa) \ ,$$

with the density function

$$d\rho_z(\kappa) = \frac{z^z}{(z-1)!} \kappa^{z-1} e^{-z\kappa}\, d\kappa \ .$$

A refined theorem

A refined theorem

Theorem

We have

$$P(z,\kappa) = 1 - Q(z,\kappa z q/p) + \left(\frac{q}{p}\right)^z e^{\kappa z \frac{p-q}{p}} Q(z,\kappa z)\ .$$

Where Q is the Incomplete Gamma Function,

$$Q(s,x) = \frac{\Gamma(s,x)}{\Gamma(x)}\ ,$$

where

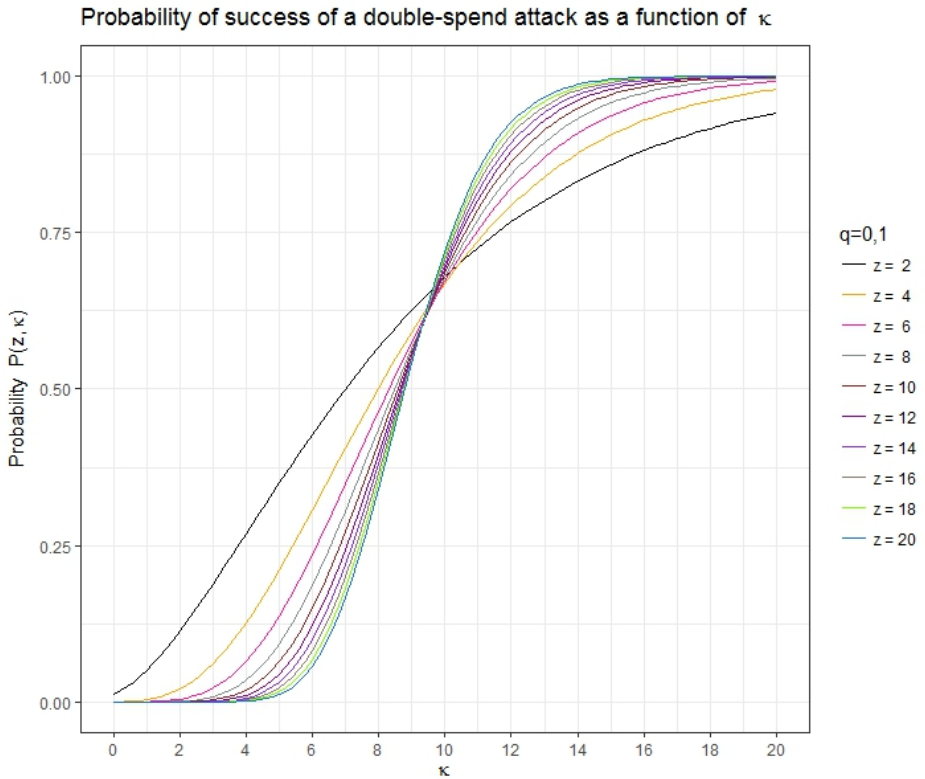
$$\Gamma(s,x) = \int_x^{+\infty} t^{s-1} e^{-t}\, dt$$

Practical implications

We can plot this family of functions for diferent values of $z \geq 1$

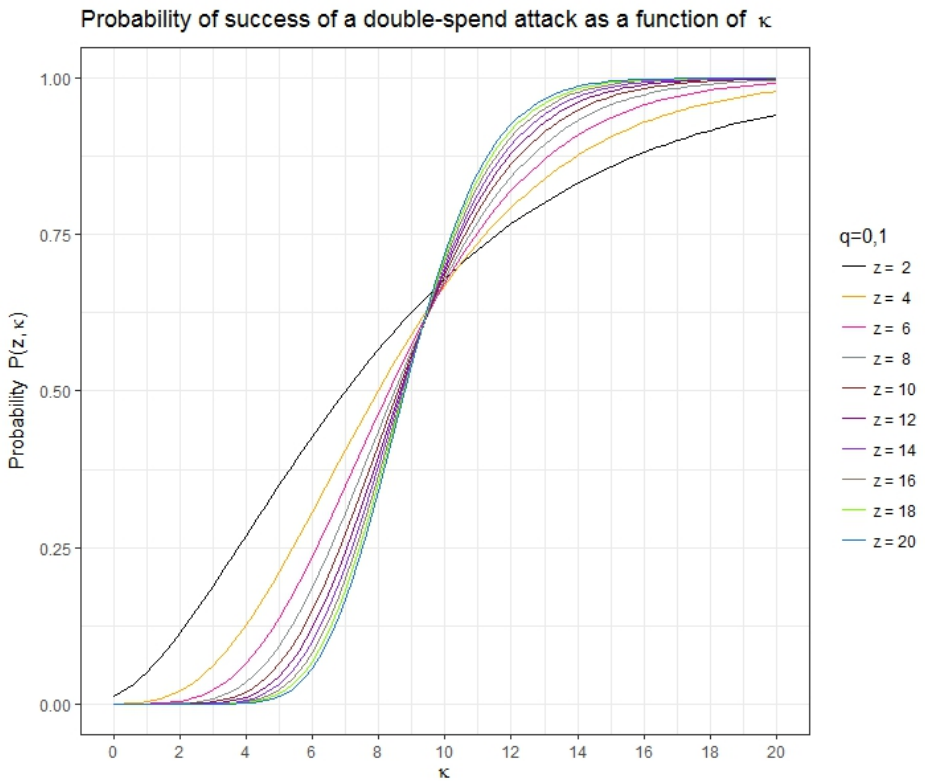
Practical implications

We can plot this family of functions for diferent values of $z \geq 1$



Practical implications

We can plot this family of functions for diferent values of $z \geq 1$



- Double entry "Satoshi risk tables" with parameter (q, κ), ArXiv 1702.04421, 2017.

Mathematics of mining

Mathematics of mining

• Mining consists in performing independent hashes changing a nounce, thus the random variable $\boldsymbol{T}$ it takes to mine a block follows an exponential distribution,

$$f_{\boldsymbol{T}}(t) = \alpha e^{-\alpha t}$$

with $t_0 = 1/\alpha = \mathbb{E}[\boldsymbol{T}]$.

Mathematics of mining

• Mining consists in performing independent hashes changing a nounce, thus the random variable $\boldsymbol{T}$ it takes to mine a block follows an exponential distribution,

$$f_{\boldsymbol{T}}(t) = \alpha e^{-\alpha t}$$

with $t_0 = 1/\alpha = \mathbb{E}[\boldsymbol{T}]$.

• If $(\boldsymbol{T}_1, \ldots, \boldsymbol{T}_n)$ are independent identically distributed exponential random variables of parameter α, then, by convolution, $\boldsymbol{S}_n = \boldsymbol{T}_1 + \ldots + \boldsymbol{T}_n$ is a random variable with a gamma density with parameter (n, α).

Mathematics of mining

• Mining consists in performing independent hashes changing a nounce, thus the random variable $\boldsymbol{T}$ it takes to mine a block follows an exponential distribution,

$$f_{\boldsymbol{T}}(t) = \alpha e^{-\alpha t}$$

with $t_0 = 1/\alpha = \mathbb{E}[\boldsymbol{T}]$.

• If $(\boldsymbol{T}_1, \ldots, \boldsymbol{T}_n)$ are independent identically distributed exponential random variables of parameter α, then, by convolution, $\boldsymbol{S}_n = \boldsymbol{T}_1 + \ldots + \boldsymbol{T}_n$ is a random variable with a gamma density with parameter (n, α).

• The random process $\boldsymbol{N}(t)$ of the number of mined blocks at time t is a Poisson process and $\boldsymbol{N}(t)$ has a Poisson distribution with expectation αt.

Elementary lemma

Elementary lemma

Similar to Gambler's Ruin problem:

Elementary lemma

Similar to Gambler's Ruin problem:

Lemma (S. Nakamoto, 2009)

Let q_n be the probability of the event E_n, "catching up from n blocks behind". We have

$$q_n = (q/p)^n \, .$$

- Let $\boldsymbol{T}$ and $\boldsymbol{S}_n$, resp. $\boldsymbol{T}'$ and $\boldsymbol{S}'_n$, be the random variables associated to the group of honest, resp. attacker, miners.

- Let $\boldsymbol{T}$ and $\boldsymbol{S}_n$, resp. $\boldsymbol{T}'$ and $\boldsymbol{S}'_n$, be the random variables associated to the group of honest, resp. attacker, miners.
- Respective random Poisson process $\boldsymbol{N}(t)$, resp. $\boldsymbol{N}'(t)$.

- Let $\boldsymbol{T}$ and $\boldsymbol{S}_n$, resp. $\boldsymbol{T}'$ and $\boldsymbol{S}'_n$, be the random variables associated to the group of honest, resp. attacker, miners.
- Respective random Poisson process $\boldsymbol{N}(t)$, resp. $\boldsymbol{N}'(t)$.
- The random variables $\boldsymbol{T}$ and $\boldsymbol{T}'$ are independent with exponential distributions with parameters α and α'

- Let $\boldsymbol{T}$ and $\boldsymbol{S}_n$, resp. $\boldsymbol{T}'$ and $\boldsymbol{S}'_n$, be the random variables associated to the group of honest, resp. attacker, miners.
- Respective random Poisson process $\boldsymbol{N}(t)$, resp. $\boldsymbol{N}'(t)$.
- The random variables $\boldsymbol{T}$ and $\boldsymbol{T}'$ are independent with exponential distributions with parameters α and α'
- We have

$$p = \frac{\alpha}{\alpha + \alpha'} \ ,$$
$$q = \frac{\alpha'}{\alpha + \alpha'} \ .$$

Exact computation

Exact computation

• Let $\boldsymbol{X}_n = \boldsymbol{N}'(\boldsymbol{S}_n)$ be the number of blocks mined by the attackers when the honest miners have just mined the n-th block.

Exact computation

• Let $\boldsymbol{X}_n = \boldsymbol{N}'(\boldsymbol{S}_n)$ be the number of blocks mined by the attackers when the honest miners have just mined the n-th block.

Proposition (C. Grunspan, R. Pérez-Marco, 2017)

The random variable $\boldsymbol{X}_n$ has a negative binomial distribution with parameters (n, p), i.e. for $k \geq 0$,

$$\mathbb{P}[\boldsymbol{X}_n = k] = p^n q^k \binom{k+n-1}{k}.$$

Probability of success of attacker

Probability of success of attacker

Proposition (Probability of success of the attacker)

The probability of success by the attackers after z blocks have been mined by the honest miners is

$$P(z) = 1 - \sum_{k=0}^{z-1} \left(p^z q^k - q^z p^k \right) \binom{k+z-1}{k} .$$

Probability of success of attacker

Proposition (Probability of success of the attacker)

The probability of success by the attackers after z blocks have been mined by the honest miners is

$$P(z) = 1 - \sum_{k=0}^{z-1} \left(p^z q^k - q^z p^k \right) \binom{k+z-1}{k} .$$

Probability of success of attacker

Proposition (Probability of success of the attacker)

The probability of success by the attackers after z blocks have been mined by the honest miners is

$$P(z) = 1 - \sum_{k=0}^{z-1} \left(p^z q^k - q^z p^k \right) \binom{k+z-1}{k} .$$

• A manipulation with special functions gives the closed form formula of the main Theorem.

Thank you for your attention!