Conquering Generals: an NP-Hard Proof of Useful Work

Angelique Faye Loe
Royal Holloway, University of London
Information Security Group
angelique.loe.2016@rhul.ac.uk

Elizabeth A. Quaglia
Royal Holloway, University of London
Information Security Group
elizabeth.quaglia@rhul.ac.uk

ABSTRACT

Proof of Work systems are used in cryptocurrencies to obtain consensus in distributed peer-to-peer systems that share no trust. Miners of cryptocurrency compete by engaging in the Proof of Work to solve a cryptographic challenge. The first to successfully provide a solution to the challenge wins by minting new currency. The process of mining also simultaneously prevents double-spending through the creation of an append-only distributed database known as the blockchain. The most widely adopted Proof of Work is the Hashcash scheme and the most widely deployed miners are ASIC-based. Despite the popularity of Hashcash, two issues are commonly identified its use. Firstly, the high energy consumption of the scheme is perceived as wasteful because the solutions found provide no useful output, and secondly, the computational complexity class of the scheme is not formally known. Based on these deficiencies, we propose a novel Proof of Work system which achieves the following goals:
- to provide a fiscally incentivized platform for algorithm research that aims to optimize an NP-Hard computational problem. This provides indirect insight into the P Versus NP Clay Institute Millennium problem, thus providing useful output.
- to construct a challenge within a known hard computational complexity class.
- to ensure the Proof of Work created is inclusive of ASIC hardware.

Our proposal is a hybrid Proof of Work system that initially uses the HashCash scheme and which subsequently constructs an instance of the NP-Hard Travelling Salesman Problem. We build on the ambitions of others to develop Proofs of Useful Work. We differentiate our paper from related work as the first to consider the current capital investment into ASIC hardware, thus including them in our proposal.

1 INTRODUCTION

The Hashcash-based [2] Proof of Work exists in Bitcoin, and several other cryptocurrencies, to serve the two main purposes of minting new currency and providing consensus in a peer-to-peer network. Peers on the network mint new currency through a process known as mining. Mining can be described as expending electric energy by iteratively running the SHA256 hashing algorithm on a candidate block header, consisting of unconfirmed transactions taken from a memory pool and concatenating a nonce until the desired output of several leading zero’s is obtained \(^1\). Mining is a competitive process in which miners race each other to be the first to find a solution to the computational challenge. The process of mining also simultaneously ensures consensus by generating the immutable distributed ledger data structure, commonly referred to as the blockchain. The blockchain is append-only and it contains the transaction history between the peers on the network. It provides cryptographic assurance that every entry is irrevocably scribed in its publicly viewable data structure. It is precisely because the blockchain provides a non-repudiable, publicly distributed database of all historical transactions that fraudulent transactions known as double-spending are broadly mitigated.

Since the Genesis Block of Bitcoin was mined, the aggregated hash rate of the participating miners has grown exponentially [5]. As a result of this growth, cryptocurrency mining has been the focus of debate in relation to the energy consumption requirements to perform the underlying Proof of Work. The “Bitcoin Energy Consumption Index” estimates the energy consumption of Bitcoin miners and compares this to various nation states and performs a comparison on a per-transaction-basis to other payment systems, such as VISA [14]. In Table 1 we present a sample of cryptocurrencies and notable academic proposals based on the classification of their underlying Proof of Work schemes, and their current market capitalization in BTC. Rows 3 - 7 outline systems which address the energy waste issue by establishing Proofs of Useful Work, whilst rows 6 - 7 additionally address the issue of heuristic hardness associated with the HashCash Proof of Work [3]. Enumerating each altcoin goes beyond the scope of this paper and [7] provides an exhaustive list of all currently traded cryptocurrencies, including those with Proofs of Useful Work.

The largest difference our proposal incorporates when compared to other Proof of Useful Work systems is the inclusion of ASIC mining hardware. If we consider one of the top end ASIC miners on the market as of March 2018, it claims a hash rate of 14 TH/s and a hardware cost of 3000 USD [6]. Based on this data we develop the following formula \( \lambda \text{TH/s} \times (3000 \text{USD}) = \lambda \text{USD} \), where \( \lambda \) is the estimated capital expenditure on ASIC hardware and \( \lambda \) is the average daily hash rate. Table 2 and Figure 1 summarize our findings \(^2\). Clearly, the capital investment into ASIC hardware runs into the billions of USD. If we consider the estimated 4.3 billion USD currently invested in ASIC hardware, this is equivalent to the Arkansas 2018 state budget for education and transportation [8].

\(^1\)For a detailed view of this process and clarification of definitions refer to [1]. History of Proof of Work algorithm development prior to the landmark paper [23] can be found most notably in [2, 15, 19]. History of digital cash and “pre-Nakamoto” cryptocurrency papers can be found in [9, 12, 16, 25].

\(^2\)We export the “Hash Rate” data from [5] for our calculations of average daily hash rate.
Table 1: Current Proof of Work Variants

<table>
<thead>
<tr>
<th>Row</th>
<th>Proof of Work Variant</th>
<th>Proof of Work Sub-Variant</th>
<th>Use Case</th>
<th>Perceived Benefits</th>
<th>Perceived Limitations</th>
<th>Market Cap (BTC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hash Based</td>
<td>SHA256 output less than target value</td>
<td>Bitcoin [23]</td>
<td>original cryptocurrency Proof of Work, widely used, matured</td>
<td>computational complexity hardness heuristic, energy consumption</td>
<td>17,067,359</td>
</tr>
<tr>
<td>2</td>
<td>Hash Based</td>
<td>Scrypt key derivation</td>
<td>Litecoin [27]</td>
<td>widely used, matured</td>
<td>computational complexity hardness heuristic only, energy consumption</td>
<td>1,144,753</td>
</tr>
<tr>
<td>3</td>
<td>Number Theoretic</td>
<td>Cunningham Chain, Bi-Twin Chain search</td>
<td>Primecoin [21]</td>
<td>academic application of solutions</td>
<td>low market cap, alienates ASIC miners</td>
<td>1,367</td>
</tr>
<tr>
<td>4</td>
<td>Number Theoretic</td>
<td>Prime gap search</td>
<td>Gapcoin [17]</td>
<td>academic application of solutions</td>
<td>very low market cap, alienates ASIC miners</td>
<td>55</td>
</tr>
<tr>
<td>5</td>
<td>Resource Efficient Mining</td>
<td>trustworthy reporting on inherently useful CPU workloads</td>
<td>academic only [20]</td>
<td>fractional percentage Hash Based energy waste</td>
<td>currently only academic, possible implications of Spectre bug, alienates ASIC miners</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>Graph Theoretic</td>
<td>Cycle and structure search in large graphs</td>
<td>academic only [26]</td>
<td>academic application, constructed problems have strongly conjectured computational complexity classes</td>
<td>currently only academic, alienates ASIC miners</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>Linear Algebra Based</td>
<td>Orthogonal Vector Problem</td>
<td>academic only [3]</td>
<td>academic application, verbose fine-grained computational complexity class conjecture provided</td>
<td>currently only academic, alienates ASIC miners</td>
<td>0</td>
</tr>
</tbody>
</table>

We thus reason that Proof of Useful Work systems that do not consider the use of ASIC hardware to be quixotic, and likely to face opposition in adoption. They additionally risk industrial scale waste of purpose-built physical hardware, that by design, cannot be repurposed. It is specifically for this reason that we do not consider Proof of Stake or Proof of Space systems in our proposal as these schemes purposely exclude ASIC miners.

We therefore propose a hybrid two-round Proof of Useful Work scheme with a deterministic time interval which continues the use of ASIC hardware in the first round and subsequently constructs an instance of the NP-Hard Travelling Salesman Problem. Our hybrid approach ensures the current use of ASIC miners that have already been purchased and deployed, thus avoiding the latter wasted energy scenario. Additionally, a hybrid approach addresses ASIC mining energy usage by limiting the Hashcash-stage to a single round, bound by a deterministic time interval. Furthermore, by including an NP-Hard computational problem we build upon the ideas presented in [3] by offering a challenge within a well-known computational complexity class whilst presenting an incentive to enhance research into algorithm design, rather than the current sole ambition of increasing brute force hashing power. Also, we argue that a hybrid scheme could shift the dependence from ASIC hardware in a gradual manner thereby making adoption of this Proof of Useful Work scheme less contentious. We finally recall that, should anyone find an algorithm to “solve” the TSP the implications would have far reaching consequences because: “NP-Complete problems have the intriguing property that if any NP-Complete problem can be solved in polynomial time, then every problem in NP has polynomial time solution, that is P = NP [11].”

2 BACKGROUND, ASSUMPTIONS AND PREREQUISITES

2.1 Background

As we aim to construct instances of the Euclidean TSP, the subsequent claims attest to the computational hardness associated with seeking optimal solutions:

**Claim 1:** There is a polynomial time reduction from the Hamiltonian Cycle Problem to the TSP, Figure 2 [11, 20].

**Claim 2:** The Travelling Salesman Decision Problem is NP-Complete in the Strong Sense [18].

**Claim 3:** It is possible to construct the NP-Hard Travelling Salesman Optimization Problem from the Travelling Salesman Decision Problem [18].

2.2 Assumptions

Our construction of the TSP is on a Euclidean plane, therefore we recall the following:

**Definition 1 The TSP (Optimization) [18].** Input: Set Ω of n cities \{ω1,...,ωn\}, distance d(ωi,ωj) ∈ \mathbb{R}^+, where 1 ≤ i, j ≤ n and i ≠ j, for each pair of cities ωi, ωj ∈ Ω. Output: The minimal tour length of a permutation \( ω_1, ω_2, ..., ω_n \) > of Ω, i.e. find a permutation that gives the following:

\[
\min \left\{ \sum_{i=1}^{n-1} d(ω_π(i), ω_π(j)) + d(ω_π(n), ω_π(1)) \right\}
\]

**Property 1** The Triangle Inequality. For Cities ωi, ωj, ωk, where 1 ≤ i, j, k ≤ n, the following property holds:

\[ d(ω_i, ω_j) ≤ d(ω_i, ω_k) + d(ω_k, ω_j). \]

**Property 2** Symmetric TSP. We also assume that the distances between any pair of cities is symmetric, that is: \( d(ω_i, ω_j) = d(ω_j, ω_i) \).

\[d(ω_i, ω_j) \]
Table 2: Estimated Capex on ASIC Hardware

<table>
<thead>
<tr>
<th>Time Interval</th>
<th>d average daily hash rate (TH/s)</th>
<th>λ est. capex on ASIC hardware (USD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jul. 1 - Dec. 31, 2017</td>
<td>8,826,086</td>
<td>1,891,304,038</td>
</tr>
<tr>
<td>Jan. 1 - Mar. 4, 2018</td>
<td>20,146,175</td>
<td>4,317,037,509</td>
</tr>
</tbody>
</table>

Table 3: Algorithm Classification for the TSP

<table>
<thead>
<tr>
<th>Algorithm Classification</th>
<th>Example Algorithm</th>
<th>Big O run time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exhaustive Search</td>
<td>Naive Algorithm</td>
<td>O(n!)</td>
</tr>
<tr>
<td>Exact</td>
<td>Held-Karp Algorithm</td>
<td>O(2^n n^2)</td>
</tr>
<tr>
<td>Approximation</td>
<td>Christofides Algorithm</td>
<td>O(n^3)</td>
</tr>
<tr>
<td>Heuristic</td>
<td>LKH Algorithm</td>
<td>O(n^{2+ε})</td>
</tr>
</tbody>
</table>

2.3 Prerequisites

Our proposal relies on the following prerequisites that miners must have:

- synchronized clocks obtained via reputable NTP Servers, (see section 3.4 for further discussion)
- the computational resource to run:
  - (i) the SHA256 cryptographic hash function
  - (ii) an algorithm of their choice to optimize the TSP
  - (iii) a sorting algorithm
  - (iv) an IEEE 754-2008 [24] floating-point arithmetic operation to find square roots using the Pythagorean theorem to calculate \( \ell_2 \) norms.

3 PROPOSED PROOF OF USEFUL WORK

A Proof of Useful Work [3] typically consists of three algorithms (Gen, Solve, Verify), where Gen(x) takes an input \( x \) and generates a challenge \( c \), Solve(c) solves the challenge \( c \), outputting a solution \( s \), and Verify(c, s) is a (possibly randomized) algorithm that verifies the solution \( s \) to \( c \). The detailed formal requirements of Proofs of Useful Work, namely, Efficiency, Completeness, Soundness, Hardness and Usefulness can be found in [3]. Our proposed Proof of Useful Work relies on the efficiency of the Gen and Verify algorithms running in polynomial time. The Solve algorithm satisfies the notion of associated hardness, with a running time of \( O(t(n)) \), where the values of \( t(n) \) are dependent on algorithm selection as outlined in Table 3. Currently the only algorithm (excluding the naive) to guarantee finding an optimal solution to the TSP is the Held-Karp, which has an intractable run time. Formal security proofs for our proposed construction are presented in [22].

3.1 Notation

\( \omega_j \) denotes a city, and \( x \) is the input of the problem. In Round 1 \( |x| = 1 \text{MB} \) (as per the bitcoin maximum candidate block size), in Round 2 \( n \) will indicate the number of cities used to construct the Euclidean TSP, and \( m \) is the number of miners online, where \( n, m \in \mathbb{Z}_+^* \). The deterministic time interval for each blocks Proof of Useful Work is \([t_s, t_c]\), where \( t_c - t_s = 10 \text{ minutes} \). \( H \) is the keyless hash function SHA256, \( \eta \) is a cryptographic nonce and \( \psi \) is a SHA256 cryptographic hash output. The subscripts are defined as 1 or 2 to identify the round, \( i \) to represent the identification of the miner, \( j \) to represent the block number in the blockchain and \( k \) to represent the counter number on the nonces. The superscripts are: \( (\ell) \), in Round 1 represents the lowest value in the set of all attempted hashes, \( (\ell) \) in Round 2 represents the permutation of \( n \) cities giving the lowest sum, and \( (\mathcal{L}) \) represents the lowest of the \( (\ell) \) solutions.

3.2 Example

For instance, \( s_{x_i(j)(k)} \) is a solution output by the miner with identification \( (33) \), and where \( x_i \) indicates the input of \( x \) relevant to Round 1. \( 512253 \) indicates the block number in the blockchain, and \( 67987 \) indicates the counter number of the attempted nonce.

3.3 Conquering Generals: an NP-Hard Proof of Useful Work

Conquering Generals is a two round Proof of Useful Work. In the first round the Hashcash scheme is used and subsequently in the second round an instance of the Travelling Salesman Problem is constructed. The set of \( \{1, \ldots, m\} \) miners online will perform the following:

### Round 1

1.1 Gen\(_1(x_j)\) each miner will generate a unique challenge \( c_{x_1(i)(j)} \) by selecting \( |x| = 1 \text{MB} \) of unconfirmed transactions from the memory pool to create a candidate block, Figure 5.

1.2 Solve\(_1(c_{x_1(i)(j)})\), will operate the double iterated SHA256 algorithm as follows:

#### Attempt 1:

\[
H(H(c_{x_1(i)(j)}||\eta_{x_1(i)(j)(1)})) = s_{x_1(i)(j)(1)}
\]

#### Attempt 2:

\[
H(H(c_{x_1(i)(j)}||\eta_{x_1(i)(j)(2)})) = s_{x_1(i)(j)(2)}
\]

...
Attempt \( t_1: \quad H(H(x_{1(i,j)})) = s_{x_1(i,j)}(s_{x_1(i,j)}(s_{x_1(i,j)}(s_{x_1(i,j)})))) \)

1.3 Propagate solution

\[ s_{x_1(i,j)}^{(t)} = \min \{ s_{x_1(i,j)}(1) s_{x_1(i,j)}(2) \ldots s_{x_1(i,j)}(t) \} \]

and corresponding \( c(s_{x_1(i,j)}(k)) \) to other miners.

1.4 Verify \( c_{x_1(i,j)}(t) \) and \( s_{x_1(i,j)}(t) \) for each solution

\[ \{ s_{x_1(i,j)}^{(t)}(1), \ldots, s_{x_1(i,j)}^{(t)}(m-1) \} \]

as follows:

If \( H(H(H(c_{x_1(i,j)})) \neq s_{x_1(i,j)}^{(t)} \) then filter from the next step. Else include.

1.5 Sort and store solutions into an ascending order sequence. Output solution

\[ s_{x_1(i,j)}^{(L)} = \min \{ s_{x_1(i,j)}^{(t)}, \ldots, s_{x_1(m-1)}^{(t)} \} \].

\[ |x| \quad |y| \]

Round 2

2.1 Gen2 \( x_2 \) will take \( x_2 = s_{x_1(i,j)}^{(L)} \) and recursively perform a double iterated hash \( n \) times as follows:

Initialize: \( H(x_2) = \psi_{x_{1(i,j)}} \),

\( H(H(x_2)) = \psi_{x_{2(i,j)}} \)\ldots \( H(H(s_{x_{1(i,j)}})) = \psi_{x_{2n}} \).

2.2 Gen2 will extract the 128 LSB’s from each hash to overlay the cities \( \{ (1, \ldots, m) \} \) on a \( 2^{64} \times 2^{64} \) Euclidean plane as follows: If \( \psi_{x_{1(i,j)}} = 000000000000000B \ldots E360568F29EF54005F2A8E8787A28759E \), Then 360568F29EF54005 F2A8E8787A28759E. Repeat for \( \psi_{x_{2(i,j)}} \), \( \psi_{x_{2n}} \).

2.3 Gen2 will calculate \( \ell_2 \) norms for each of the \( \frac{n(n-1)}{2} \) edges on the complete graph and construct the graph distance matrix as follows:

\[
\begin{bmatrix}
0 & d(1,2) & d(1,3) & \ldots & d(1,n) \\
d(2,1) & 0 & d(2,3) & \ldots & d(2,n) \\
d(3,1) & d(3,2) & 0 & \ldots & d(3,n) \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
d(n,1) & d(n,2) & d(n,3) & \ldots & 0
\end{bmatrix}
\]

2.4 Sort and store solutions in ascending order. Output solution

\[ s_{x_2(i,j)}^{(L)} = \min \{ s_{x_2(i,j)}^{(t)}, s_{x_2(i,j)}^{(t)}, \ldots, s_{x_2(i,j)}^{(t)} \} \].

If a tie for \( c_{x_2(i,j)}^{(t)} \) then the solution belonging to the miner with the lowest hash in step 1.5 will be the tie-breaking criteria.

2.5 Verify \( c_{x_2(i,j)}^{(t)} \) \( s_{x_2(i,j)}^{(L)} \) will run as follows: Look up associated distances from \( s_{x_2(i,j)}^{(t)} \). Namely:

\( d(\omega_{x_{(1)}}, \omega_{x_{(2)}}, \ldots, \omega_{x_{(1)}}) \)\ldots \( d(\omega_{x_{(n-1)}}, \omega_{x_{(n)}}, \ldots, \omega_{x_{(1)}}) \)\ldots \( d(\omega_{x_{(n)}}, \omega_{x_{(n)}}, \omega_{x_{(1)}}) \) directly from \( c_{x_2(i,j)}^{(t)} \).

2.6 Increment Block Number, Mint Currency, Commit Transactions and Loop as follows: Increase the counter of \( i \) to \( (j + 1) \). Miner \( (i) \) \( W \) will mint the new currency and the unconfirmed transactions chosen in 1.1 will be those committed in block \( (j + 1) \).

For the solutions provided by miner \( (i) \) \( W \), we compute:

\[ H(H(c_{x_2(i,j)})), s_{x_2(i,j)}^{(L)} \] or \( H(H(c_{x_2(i,j)})) \).

From Figure 2: Polynomial reductions from SAT to the TSP [20]

Figure 3: Visual of toy example of Gen2 \( x_2 \)

Note: \( \min \) indicates the absolute minimum, whereas \( \min \) indicates the minimum obtained by time \( t_i \).
step 1.3 and inject them into the \( (j + 1) \) candidate block header. We then remove the sorted values from step 1.5 and 2.4 and Loop to step 1.1. Figure 5.

3.4 Timelines
As we propose a deterministic mining interval, the time line of this scheme will be as follows:

- step 1.1 (generating the candidate block/ challenge) occurs as a background process during steps 1.2 through 2.6,
- step 1.2 (the Hashcash stage) \([t_1, t_1 = t_5 + 3]\),
- steps 1.3 through 1.5 (propagation, verification, sorting) \([t_1, t_2 = t_1 + 1]\),
- step 2.1 (generating instance of the TSP) \([t_2, t_3 = t_2 + 1]\),
- step 2.2 (solving instance of the TSP) \([t_3, t_4 = t_3 + 4]\),
- steps 2.3 through 2.6 (propagation, sorting, verification) \([t_4, t_c = t_4 + 1]\) \(^5\).

It will also be critical to ensure protection against time shifting attacks, which can be achieved using Chronos [13].

3.5 Toy Example of Step 2.1.2
On a \( 2^4 \times 2^4 \) Euclidean Plane we take the 8 LSB’s from each hash output \( \{\psi_{i1}, \ldots, \psi_{i8}\} \) as follows:

- \( \psi_{i1} = 36F63E2A0B7C96BF84 \ldots F6 \) \( A \) \( x_1 \)
- \( \psi_{i2} = BEBEF3761858AC5113 \ldots 8B \) \( D \) \( x_2 \)

\(^5\)It may also be necessary between steps 1.3 and 1.4 and between 2.3 and 2.4 to introduce another time limit for the receipt of other miners propagated solutions prior to sorting and verification.

3.6 Background to the Conquering Generals
The name “Conquering Generals” was the merger of ideas from the Travelling Salesman Problem and the Byzantine Generals Problem associated with the consensus achieved in a Proof of Work. We equate the work as the planning effort undertaken by a group of generals with a conquering mission. First, they perform reconnaissance to determine the cities they wish to overthrow (the Hashcash stage) then they establish the optimal order in which to conquer the cities (the TSP stage).

3.7 Implementation Considerations
Hashcash based schemes subject transaction clearing times to jitter as noted in Figure 4. Our proposal aspires to ease erratic block confirmation times by using a deterministic mining interval, requiring the clock synchronization from a legitimate time source, as noted in the prerequisites.

Our selection on the elapsed CPU years recorded by the Concorde code when attempting all the 110 TSPLIB instances [10]. To aid miners in determining how many propagated solutions they expect to receive in step 1.3 and 2.3, they may use a recursive algorithm such as the getaddr Python crawler to discover the other miners online [4].

In [22] we also discuss the probability and impact of the construction of geometrically congruent instances of the TSP in different blocks. In summary we argue that the combinatorial search space of all possible instances of the TSP with \( n \) cities on a \( 2^6 \times 2^6 \) plane would make the computational resources required for storage and search infeasible. Furthermore, we recall Claims 1 - 3 justify that...
We also note that our proposal presents an instance of a randomly generated computational challenge and thus lacks an explicit link to a real-life problem. Our original ambition was to design a system that would solve real-life instances of the TSP. We contemplated that a logistics company may consider outsourcing instances of the TSP to be solved by the cryptocurrency miners concerning delivery routes for their transportation fleets. However, with the latter scenario, we noted the time sensitive and confidential nature associated with commercially based real-life problems. This introduced an anticipated dependence on a Trusted Third Party to handle typical business requirements such as job queueing, prioritization and confidentiality. Reliance on a TTP would of course break the untrusted distributed consensus model inherent in cryptocurrencies. We therefore opted to focus on a framework which precluded the use of any centralized elements. However, in [28] a framework entitled “Resource-Efficient Mining” addresses the issue of solving real-life problems. In this system miners are free to choose which challenges to solve (including self-submitted problems), and to decide the order in which they wish to solve them. As noted previously, real-life problems often have time restraints associated with solution submission and may also contain confidential information. Therefore, the outstanding issue of handling these properties without a TTP remains. We leave this challenge open for further research.

5 CONCLUSION

We have presented a Proof of Useful Work which avoids the potential loss of an estimated 4.3 billion USD of currently deployed ASIC hardware whilst introducing a method to subsequently construct an instance of the NP-Hard Travelling Salesman Problem. Our proposal ensures that the hardness of the scheme belongs to a formally defined computational complexity class whilst simultaneously providing a fiscally incentivized platform to perform algorithm optimization research, thereby providing useful insight into the P versus NP problem.

REFERENCES