Aalborg Universitet

Stateful to Stateless

Modelling Stateless Ethereum

Johnson, Sandra; Hyland-Wood, David; Madsen, Anders L.; Mengersen, Kerrie

Link to publication from Aalborg University

Stateful to Stateless
Modelling Stateless Ethereum

Sandra Johnson

ConsenSys Software Inc, Australia[*]

ARC Centre of Excellence for
Mathematical & Statistical Frontiers (ACEMS)
QUT, Brisbane, Australia

`sandra.johnson@consensys.net`

David Hyland-Wood

ConsenSys Software Inc, Australia

University of Queensland
Brisbane, Australia

`d.hylandwood@uq.edu.au`

Anders L Madsen

HUGIN EXPERT A/S, Aalborg University
Denmark

`anders@hugin.com`

Kerrie Mengersen

School of Mathematical Sciences and
Centre for Data Science
Queensland University of Technology (QUT)
Brisbane, Australia

`k.mengersen@qut.edu.au`

The concept of 'Stateless Ethereum' was conceived with the primary aim of mitigating Ethereum's unbounded state growth. The key facilitator of Stateless Ethereum is through the introduction of 'witnesses' into the ecosystem. The changes and potential consequences that these additional data packets pose on the network need to be identified and analysed to ensure that the Ethereum ecosystem can continue operating securely and efficiently. In this paper we propose a Bayesian Network model, a probabilistic graphical modelling approach, to capture the key factors and their interactions in Ethereum mainnet, the public Ethereum blockchain, focussing on the changes being introduced by Stateless Ethereum to estimate the health of the resulting Ethereum ecosystem. We use a mixture of empirical data and expert knowledge, where data are unavailable, to quantify the model. Based on the data and expert knowledge available to use at the time of modelling, the Ethereum ecosystem is expected to remain healthy following the introduction of Stateless Ethereum.

1   Introduction

Ethereum [21] is the largest blockchain after Bitcoin [13]. When a new node wishes to join Ethereum mainnet, the public Ethereum blockchain, it has to acquire all of the blockchain history, starting from the first block up to the most recently added block. The storage space required depends largely on whether the node is operating as a 'full' node, 'archive' node, 'miner', or 'light client', and to some extent on the chosen Ethereum client software that is used to run the node.

The Ethereum world state contains all Ethereum accounts, their balances, deployed program code known as "smart contracts", and associated storage[1]. New accounts are continually being added and new smart contracts are being deployed. Therefore, by design, the state size of Ethereum keeps growing ad infinitum, leading to increased cost and time to sync to the network, which in turn leads to a less diverse ecosystem of Ethereum nodes, and slower transaction processing and block verification times.

Vitalik Buterin recognised these issues back in 2017 when he first introduced the concept of Stateless Ethereum suggesting a protocol transformation based on state transition functions (STFs) [1]. The

---

[*]Corresponding author

[1]Ethereum Development Documentation

key mechanism of Stateless Ethereum is the creation of block "witnesses" so that when clients receive validated blocks from miners, they will also receive their corresponding witness. Each witness contains a set of Merkle branches that are authenticated against the state root and contains all the data required to execute the transactions contained in that particular block, i.e. all the information required to prove the validity of the block and to perform the state transition.

With witnesses comes increased traffic on the network, because more data packets will be passed around the network. We therefore need to assess the impact that this may have on the network to ensure that the Ethereum ecosystem continues to operate securely and efficiently in this new, altered ecosystem. Since Ethereum mainnet is a public blockchain, it provides a rich data source for modelling and analysis.

2 Modelling Approach

Ethereum is a complex environment, and modelling such an ecosystem is very challenging. From the myriad of processes and interactions, we need to extract those that are most likely to be affected in some way by introducing statelessness into this equilibrium. Once we have a clearer idea of the key processes that may be impacted, we can start building a model of Ethereum as it is today, and include those processes that are specific to Stateless Ethereum.

When building a model we aim to capture what we know, and facilitate the exploration of that which we do not know. Therefore, we consult with experts to identify 'key factors' that best capture the behaviour of the system, or parts thereof. Representing expert knowledge in the model is crucial, not only to identify those processes that may be affected by the changes, but to help achieve the necessary balance between system detail and model compactness, to avoid unneccesary model clutter.

From the plethora of modelling techniques that are available, we decided on a Bayesian network (BN), a probabilistic graphical modelling approach, for several reasons.

First, the visual nature of a BN facilitates ease of interaction by all stakeholders with the model. Secondly, in a BN uncertainty is captured in the model using joint probability distributions across all the variables and their interactions with other variables. Thirdly, this modelling approach is well suited to modelling complex systems like Ethereum, and lastly, diverse data sources can be used to quantify the model, including expert knowledge, empirical data, model output, published and grey literature.

A BN is a probabilistic representation of variables and their directed relationships [14], modelled as a directed acyclic graph (DAG) comprising nodes representing variables, directed links (edges) between nodes representing relationships between the variables, and conditional probability distributions for each variable representing the nature of these relationships. The two key characteristics of BNs are the Markov property and *d*-separation [14, 11, 4]. The Markov property allows that there are no other direct dependencies in the network over and above those already represented by directed links in the network. D-separation, or 'directional' separation is a graphical criteria for reading statements of conditional dependence and independence from the structure of the BN.

Since BNs were first introduced by Pearl in 1988 [14], many papers and textbooks have been written on this modelling technique, including theoretical and applied aspects of BNs and BN learning [14, 2, 4, 10]. BNs are now common in a very wide range of applications such as medicine, computing, natural sciences and engineering [16, 17], as well as in more general fields such as decision analysis [19] and risk assessment [3].

Moreover, several variations have emerged to better suit requirements of systems being modelled. For example, a Dynamic Bayesian network (DBN) to better facilitate systems that change over time, and an Object-oriented Bayesian network (OOBN) which is useful to model large, complex hierarchical

systems [8].

OOBNs provide a more modular approach to modelling consisting of subnetworks and interfaces. Each subnetwork can be a complete BN model in its own right, or a network fragment for repeating patterns in the model. The interface(s) between these subnetworks define the information flow between them. The subnetworks can also be nested to represent a hierarchical system. A key advantage of this approach is the ability to develop the objects in parallel, using the interface to protect against internal changes in these BN sub-models [6]. Building a BN is an iterative process [5, 12].

Sensitivity analysis can be used to assess the sensitivity of specific variables to variations in the evidence entered into the network (evidence sensitivity) and to variations in the values of the parameters (parameter sensitivity) [16, 17]. Popular measures of evidence sensitivity are entropy, which can be interpreted as the average additional information necessary to specify an alternative, and mutual information, which represents the extent to which the joint probability of two variables differs from what it would have been if they were independent [14, 9]. Parameter sensitivity metrics include an evaluation of the first derivative of the sensitivity ratio of the parameter of interest and the target probability [15].

An object-oriented BN model was chosen for the Stateless Ethereum model, to best deal with the complexity of the ecosystem and take advantage of sub-models that can be run independently.

3   Stateless Ethereum Model Structure

To the best of our knowledge, a probabilistic modelling approach such as BNs has not previously been used to model Ethereum nor Stateless Ethereum.

Stateless Ethereum OOBN   Figure 1 on page 30 shows a high-level view of the complete model, comprising of four sub-models: *Block creation*, *Witness creation*, *Ethereum network*, and *Block propagation*. Each rounded rectangle represents an OOBN sub-model. In each of the sub-models, we captured the key variables and their interactions in Ethereum mainnet and focussed on the changes introduced by Stateless Ethereum. The flow of information through the sub-models are via input (ellipse with broken line) and output nodes (ellipse with solid line), the former acts as a placeholder and the latter means the node is visible to other sub-models [8, 6]. In addition to input and output nodes, an OOBN can have private nodes. These key factors may be added, changed or deleted without affecting the other sub-models, other than through the predefined interface nodes (output and input nodes) [6].

Each sub-model is explained in more detail below.

Ethereum network OOBN   The Ethereum network model (Figure 2 on page 31) has five key factors: *Block producer?*, *Node bandwidth* and *Network latency*, that are modelled as output 'nodes' and *Peer location* and *Node location*, that are modelled as private nodes.

For each of the factors in the model, we need a probability distribution across the various states that the factor can be in and this table of probabilities, known as a conditional probability table (CPT), is attached to each factor in the model.

In our study we considered only two types of Ethereum nodes: those that produced validated blocks, i.e. 'block producers', aka miners, and those that did not. We refer to the latter as 'semi-stateless' nodes. For example, in Stateless Ethereum, a full node is treated as a 'semi-stateless' node having varying degrees of 'statelessness' from no state at the time of joining the network, up to 'full state' if the node decides to keep up to date with the Ethereum world state. To quantify *Block producer?* we calculated the proportion of all Ethereum nodes that are miners using data from Ethereum mainnet. This gave
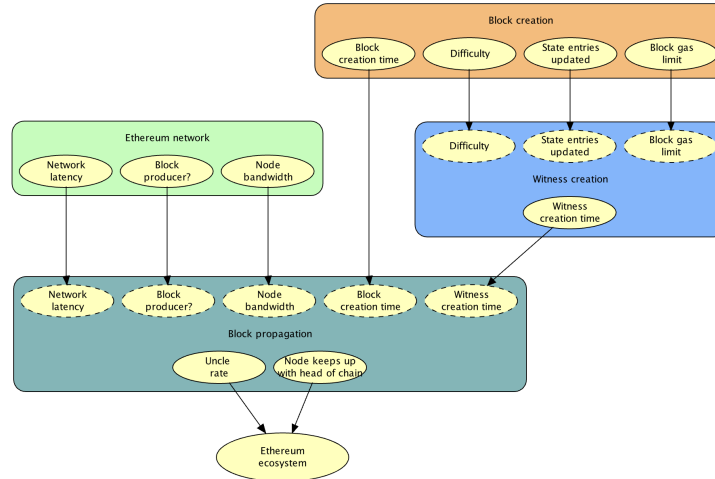
Figure 1: High level model of Stateless Ethereum showing sub-models and interfaces

us an upper limit that was then adjusted using combined expert judgement. *Node bandwidth* affects block propagation through the Ethereum network. Data from ethernodes.org [2] were used to quantify this variable and a group of experts assisted in adjusting these proportions.

Experts agreed on the threshold for 'low', 'medium', and 'high' and the description and ranges for these levels were documented with the elicitation sheets to ensure consistency and transparency. For example, for 'node bandwidth', businesses and hosting nodes were assigned as high bandwidth, colleges and residential nodes as medium bandwidth, and all other categories as low bandwidth.

Ethereum nodes are spread around the world and to make *node location* tractable, we grouped countries into five regions: *Europe*, *North America*, *China*, *Rest of Asia*, and *Rest of the world*.

Five experts participated in the elicitation exercise for node location, using a range of strategies to estimate the probability distributions for mining and semi-stateless nodes, including ranking the locations and using the data from ethernodes.org as a baseline to vary for miners. Peer location was more challenging to quantify and we had to rely entirely on experts with domain knowledge to quantify this variable.

To quantify Network latency we used GNU ping data from Wondernetwork [20] grouped into the five regions used for node and peer locations. Running the Ethereum network OOBN model shows the marginal probabilities for each factor (Figure 3 page 31).

Block creation OOBN   Block creation in Ethereum is a complicated process, however, we found that many of the complexities of block creation did not directly affect the ability of stateless nodes to keep up with the network's activities. We built the initial version of the block creation model structure based on our understanding of the current block creation process in Ethereum, extracting data from 26,595 blocks that were mined during 26-30 November 2019. In Ethereum, the term gas is used as measure of the expected work required to process a transaction. Transactions are included in a block, but the total gas that they are expected to use is limited by the *Block gas limit*, which can be adjusted by miners.

As can be seen in Figure 4, we read data such as block creation time, number of transactions per block, difficulty and more. We chose these blocks because we created witnesses for them when we

---

[2]`https://ethernodes.org/` accessed 30 June 2020
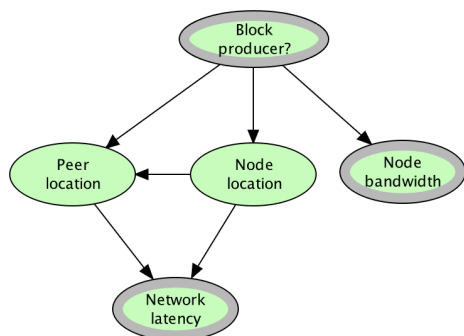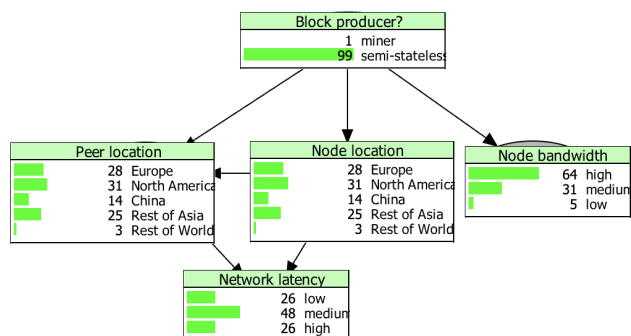
Figure 2: Ethereum network BN

Figure 3: Running Ethereum network BN

implemented the Stateless witness specification.

Using the block information we extracted, we ran multiple structure learning algorithms using general and structure restricted methods. These algorithms are pre-coded in Hugin Expert[3], the software modelling tool we used to develop the model. The output from the learning exercises resulted in minor modifications to the initial structure. For example, although there was no direct link between *Difficulty* and *Block gas limit* in the way difficulty was calculated, three of the six structure learning algorithms picked up an association between them. The learning algorithms we used were: greedy-search-and-score, tree-augmented naïve bayes, NPC (necessary path condition), Rebane-Pearl, a variant of the PC algorithm, and Chow-Liu. [4]. We therefore supplemented the model with a directed link between these two key factors. However, weak links between key factors may be deleted in subsequent iterations of the model. Figure 4 on page 32 shows the resulting block creation model structure.

Running the block creation model, produces the marginal probability distributions as illustrated in Figure 5 on page 32.

Witness creation OOBN   We implemented the Stateless Ethereum witness specification in a fork of Hyperledger Besu[5], creating witnesses for 26,595 blocks, recording the size and creation time for each witness. The dependencies between witnesses and blocks were established using a combination of expert judgement and structure learning (Figure 6 on page 32).

Once the structure was ratified, we used the witness and block information to learn the conditional probabilities.

Difficulty and *State entries updated* variables are input nodes, obtained from the Block creation OOBN. *Witness size* is private, and *Witness creation time* is an output node, and thus visible to the other OOBN sub-models.

Block Propagation OOBN   The main model endpoint of Block Propagation is the variable *Node keeps up with head of the chain*, which is modelled as an output node. Five of the factors in the block propagation BN model are part of other Stateless Ethereum sub-models. These are shown in Figure 8 on page 33 as ellipses with a broken line: *Block creation time*, *Witness creation time*, *Node bandwidth*, *Network latency* and *Block producer?*. We therefore only need to create CPTs for the remaining five nodes: *Uncle*

---

[3]https://www.hugin.com/

[4]Algorithms available in Hugin

[5]Implementation of Stateless Ethereum witness specification in fork of Hyperledger Besu
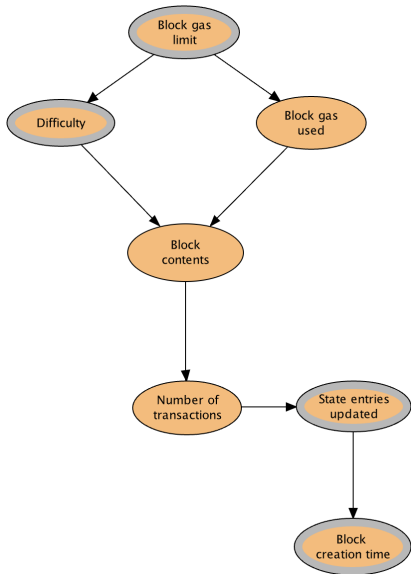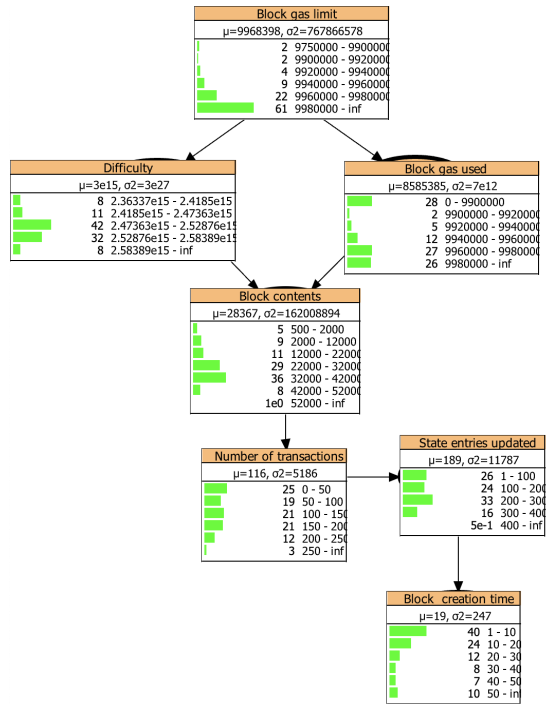
Figure 4: Block creation BN

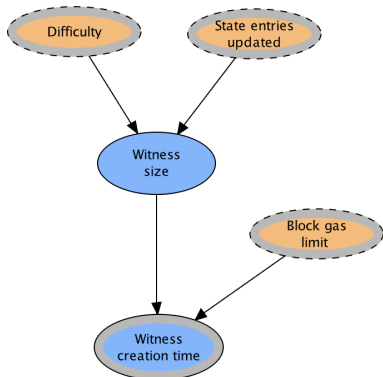Figure 5: Running Block creation BN

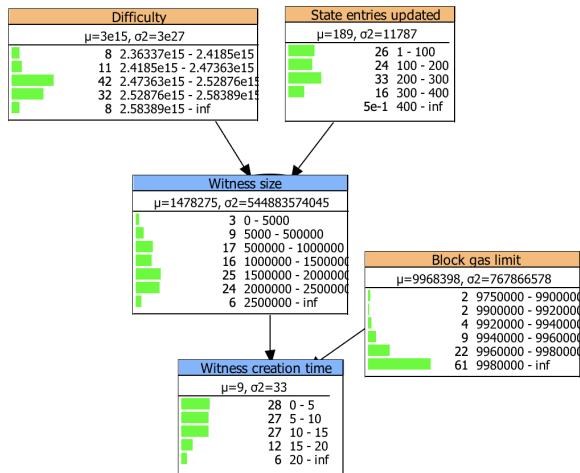Figure 6: Witness creation BN

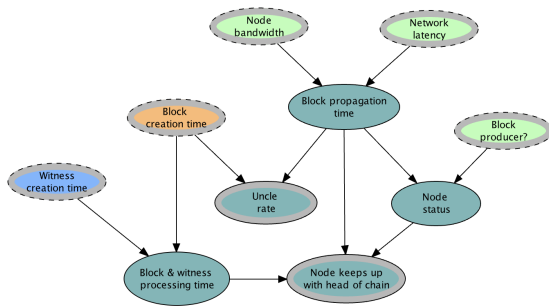Figure 7: Running Witness creation BN
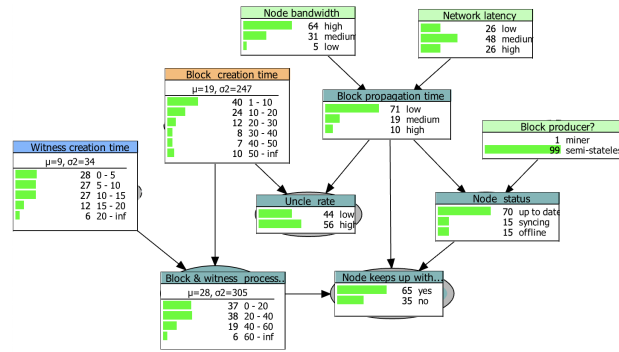
Figure 8: Block propagation OOBN sub-model

Figure 9: Running Block propagation BN

rate, *Block propagation time*, *Block and witness processing time*, *Node status* and *Node keeps up with head of the chain*.

The *Node status* can be 'up to date' (i.e. tracking the head of the chain), busy syncing to the head of the chain, or offline. When a node is offline it is difficult to know whether it is having difficulty keeping up to date with the head of the chain, has been decommissioned, having technical issues, or due to some other reason. We consulted with experts to obtain the conditional probabilities for *Node keeps up with head of the chain* and the resulting CPT entries reflect their combined input.

When two or more miners produce a valid block at the same time and attempt to add it to the chain, only one will be added to the canonical chain. The other valid blocks are known as 'uncles', or 'ommers'. We calculated daily uncle rates recorded by etherchain.org[6] and Alethio[7].

We collected block propagation times from ethstats.io[8] and ethstats.net[9]. Contribution of block propagation times to these websites are optional and do not represent all of the Ethereum network. Silva et al (2020) [18] conducted a month-long experiment from 1 April 2019 to 2 May 2019, collecting Ethereum network data, including block propagation times between geographically diverse nodes. We combined information from these three data sources for the prior probabilities, taking into account that changes in node latency and bandwidth will have an effect on block propagation time, which in turn has an effect on uncle rates. The extent to which the various factors influence block propagation time were elicited from experts.

Block and witness processing time is calculated as the sum of *Witness creation time* and *Block creation time*. The assumption therefore is that these times are additive.

To quantify the effect that block propagation times have on *Node status* for mining and non-mining nodes in the network we have to estimate the probability that a node is offline, syncing, or updated. Experts used our primary empirical data source, ethernodes.org, as a baseline to estimate the CPT entries.

Running the fully quantified block propagation model results in the probabilities shown in Figure 9.

Ethereum ecosystem Whether *Uncle rate* and *Node keeps up with head of the chain* are good indicators of the health of the Ethereum ecosystem is open to debate, but the consensus among experts was that they are valuable indicators. The reasoning is that a high uncle rate is undesirable and indicative

---

[6] https://etherchain.org/correlations
[7] https://reports.aleth.io/
[8] https://ethstats.io/
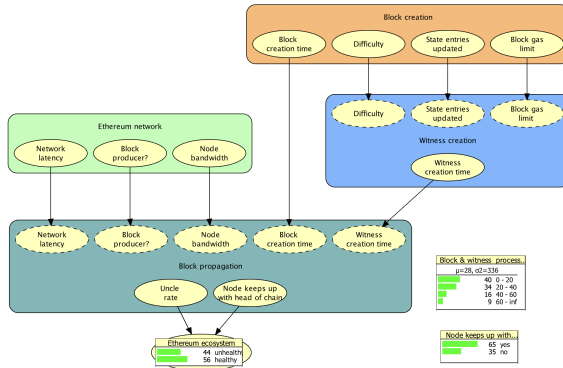[9] https://ethstats.net/
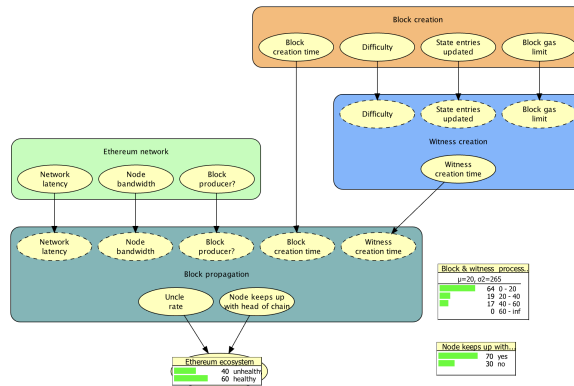
Figure 10: Running Stateless Ethereum BN

Figure 11: Running Stateless Ethereum BN - no witness generation

of a suboptimal network, adversely affecting transactional throughput[10]. Moreover, many nodes having difficulty synchronising to the network is indicative of network-related issues. (Figure 10).

With the quantification of all the sub-models and the main model endpoint, *Ethereum ecosystem*, complete, we were able to run the combined model, which ran all the sub-models (Figure 10, page 34). It showed that the higher probability was the 'healthy' state at 56%. In other words, the health of the Ethereum ecosystem with a basic Stateless implementation, i.e. no compression of witnesses, is expected to be healthy.

As a comparison, we ran the model again, excluding witness generation (Figure 11, page 34). For this scenario there was a relative percentage gain of 7% for the Ethereum ecosystem being healthy (60%).

The caveat for this result is three-fold: Firstly, the model was quantified using a subset of Ethereum mainnet data (26,595 blocks), which predates Ethereum improvement proposal EIP1559[11]. EIP1559 redesigned Ethereum's fee market. Of particular relevance to the Stateless Ethereum BN model is that post-EIP1559 block sizes are more variable and can be up to twice as large as they were at the time of building the Stateless Ethereum model. Secondly, implementing a witness size reduction technique such as Verkle tries[12] is bound to alter the model. Lastly, expert knowledge can introduce bias into the model.

For the reasons mentioned above, the probability of the final outcome, Ethereum ecosystem, being healthy, should not be interpreted as an exact result, but rather as a reference point for assessing changes in probability when exploring the effect that various scenarios have on the predicted health of the ecosystem.

We can use the quantified model to ask questions such as: "For a non-mining node and a very large witness, how does that affect the ability of a node to keep up with the head of the chain?" This will give some insight into expected changes to the Ethereum system status quo.

We selected two variations on this question and the graphical outcomes are shown in Figure 12. We observed that the probability of having this combination of evidence is 23.7% for the second largest range (Figure 12(b)), and 5.9% for the largest (Figure 12(c)). In other words, these scenarios are not expected to happen very often.

The less severe scenario (Figure 12(b)) showed the probability of keeping up with the head of the

---

[10]How do uncle blocks affect blockchain performance

[11]`https://github.com/ethereum/EIPs/blob/master/EIPS/eip-1559.md`

[12]`https://dankradfeist.de/ethereum/2021/06/18/verkle-trie-for-eth1.html`
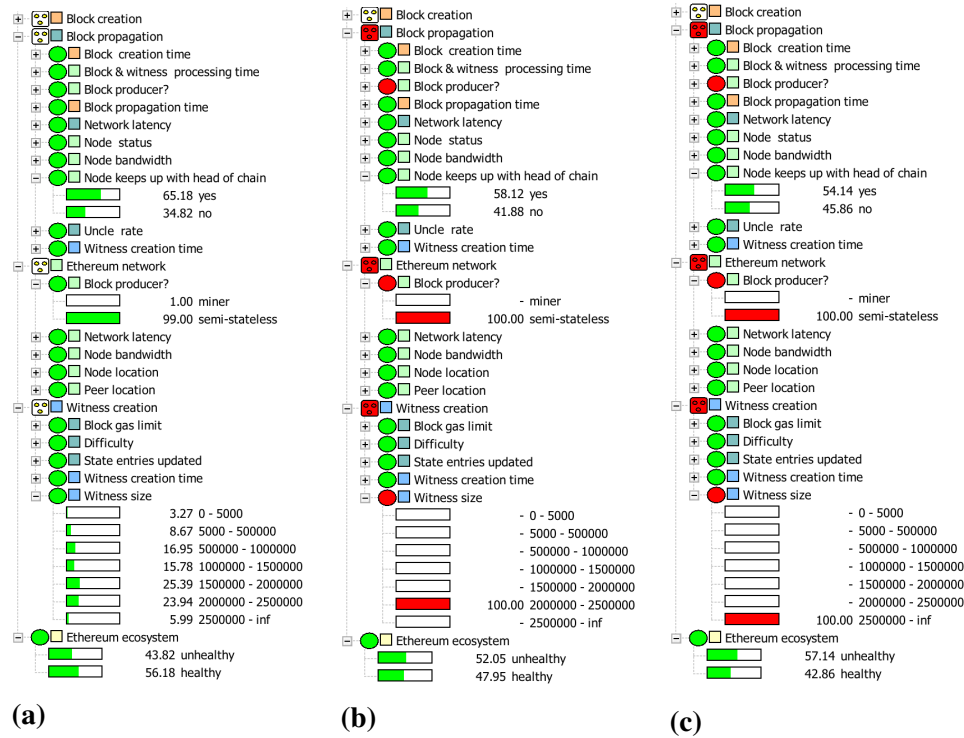
(a)   **(b)**   **(c)**

Figure 12: Running 'what if' scenarios in Stateless Ethereum OOBN

chain dropping from 65% to 58%, a percentage change of 7% which is a relative change of 11%. For the more severe scenario (Figure 12(c)), i.e. the largest witnesses, the probability of keeping up with the head of the chain dropped even further to 54%, which translates to a relative change of 17%.

4 Evidence and Parameter Sensitivity Analysis

Two additional important reasons for selecting BNs as the modelling technique are the explicit representation of assumptions in the graphical structure of the BN, which is helpful in the expert knowledge elicitation process (relevant for both structure and probability parameter elicitation), and the ability to perform different types of sensitivity analysis on the model to assess the robustness of the model and the results of the model. In this section, we will consider both types of sensitivity analysis.

The robustness of results of any model is sensitive to the underlying assumptions of the model. In a BN, one of the important assumptions is the quantification of the model, i.e., the entries of the CPTs also known as the parameters of the BN. To assess the robustness of the BN, a one-way parameter sensitivity analysis was performed under different evidence scenarios, see e.g., [7]. Here, we report on the parameter sensitivity analysis performed with respect to three different scenarios (1) no evidence, (2) base scenario, and (3) most severe scenario. The latter two scenarios are shown in Figure 13 and Figure 14, respectively.

In each scenario, we investigated the impact of probability parameter variations on the probability of the hypothesis that the Ethereum ecosystem is healthy. We denote this as $P(EthereumEcosystem = healthy \mid \varepsilon)$ where ε is the evidence scenario. We performed a one-way parameter sensitivity evaluating the impact of variations in a single probability parameter t on $P(EthereumEcosystem = healthy \mid \varepsilon)$ determining the sensitivity function denoted $P(EthereumEcosystem = healthy \mid \varepsilon)(t)$.
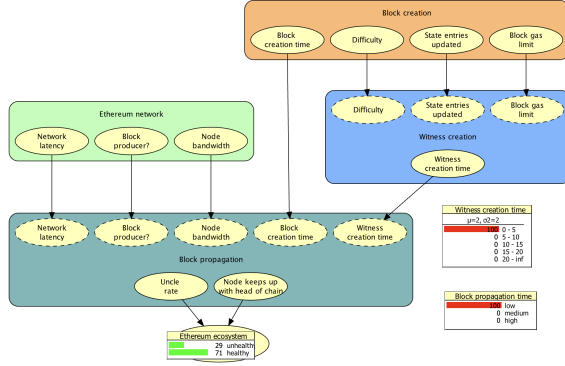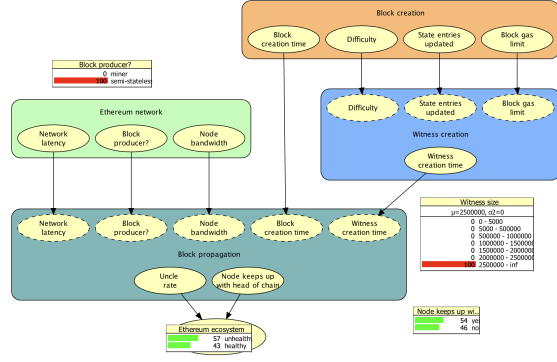
Figure 13: Base scenario

Figure 14: Most severe scenario

In general, the sensitivity function is a ratio of two linear functions in the parameter, i.e., $P(EthereumEcosystem = healthy \mid \varepsilon)(t) = \frac{\alpha \cdot t + \beta}{\gamma \cdot t + \delta}$, where α, β, γ, and δ are real numbers. The composition of the evidence makes the denominator equal to one.

For the base scenario, the parameter with highest impact was
$t = P(EthereumEcosystem = healthy \mid NodeKeepsUpWithHeadOfChain = yes, UncleRate = high)$,
i.e., this is the parameter with the highest value of the sensitivity function at the initial parameter assessment.

The sensitivity function evaluated to $P(EthereumEcosystem = healthy \mid \varepsilon)(t) = 0.3285 \cdot t + 0.3318$ with a sensitivity value less than one. Hence, the probability of the hypothesis is relatively insensitive to variations in the parameter value.
The sensitivity function for the base case was:
$P(EthereumEcosystem = healthy \mid \varepsilon)(t) = -0.7301 \cdot t + 0.8222$, where
$t = P(NodeStatus = stateOffline \mid BlockPropagationTime = low, EthereumNodeType = semiStateless)$
The sensitivity function for the most severe case was:
$P(EthereumEcosystem = healthy \mid \varepsilon)(t) = -0.4724 \cdot t + 0.6123$, where
$t = P(NodeStatus = stateOffline \mid BlockPropagationTime = low, EthereumNodeType = semiStateless)$.

Again, the sensitivity value was less than one and in all three scenarios the sensitivity function was linear due to the structure of the evidence. This suggests that the posterior probability of the hypothesis is relatively insensitive to variations in the probability parameter assessment when considered in a one-way parameter sensitivity analysis.

To further assess the performance of the model an evidence sensitivity analysis was performed. In evidence sensitivity analysis, the robustness of the results of the model was assessed with respect to variations in the evidence, see, e.g., [7]. For the no evidence scenario, the two variables that may produce the highest impact on the probability of the hypothesis, are *NodeKeepsUpWithHeadOfChain* and *NodeStatus* that may produce $0.0337 \leq P(EthereumEcosystem = healthy \mid \varepsilon) \leq 0.8439$ and $0.0536 \leq P(EthereumEcosystem = healthy \mid \varepsilon) \leq 0.6764$, respectively.

For the base scenario, the same two variables may produce $0.06 \leq P(EthereumEcosystem = healthy \mid \varepsilon) \leq 0.91$ and $0.08 \leq P(EthereumEcosystem = healthy \mid \varepsilon) \leq 0.83$, respectively. For the most severe scenario, two variables may produce $0.0176 \leq P(EthereumEcosystem = healthy \mid \varepsilon) \leq 0.7768$ and $0.0286 \leq P(EthereumEcosystem = healthy \mid \varepsilon) \leq 0.5271$, respectively. The results of the evidence sensitivity analysis showed that the two variables that have the highest impact on the hypothesis across the three different evidence scenarios were *NodeKeepsUpWithHeadOfChain* and *NodeStatus*. All other

variables can produce a much smaller variation in the probability of the hypothesis.

5 Discussion

Introducing Stateless Ethereum into a functioning, stable ecosystem presents us with an interesting study.

Ethereum has an active open source community which, by its very nature, encourages collaboration among developers and researchers. Moreover, Ethereum mainnet is a public blockchain, giving us ready access to empirical data for the BN model we built. However, contribution to websites reporting on Ethereum nodes and the distributed network is voluntary, and therefore, any insights from the data are based on a partial view of the overall network.

Modelling the known processes and quantifying them using empirical data, supplemented with the new processes and representing our knowledge about these processes, enables us to gain a more complete picture of the potential repercussions to other parts of the network. The ability of Bayesian networks to include diverse data sources such as empirical data, model output and expert knowledge, provides us with a view of the proposed system that takes into account all available information at the present time. However, the presence of expert knowledge can introduce bias into the model. Conducting sensitivity analysis and model verification are helpful in identifying and mitigating bias. Although not always the case, in this instance, the sensitivity analysis performed to assess the robustness of the model did not reveal any unexpected model behaviour requiring a revision of the model structure. However, since variables *NodeKeepsUpWithHeadOfChain* and *NodeStatus* have the highest impact on the hypothesis that the Ethereum ecosystem is healthy, additional effort and validation in estimating those parameters are recommended [6].

Bayesian network modelling is typically an iterative process, and the model presented here is the first iteration of the model. Additional versions may be created to include new and emerging research and current data, such as post-EIP1559. A key challenge in the development of this probabilistic model and its quantification has been the varying degree of data coverage of the datasets used to populate the CPTs of the BN nodes. This inconsistency introduces measurement errors into the model, compounded by elicited probabilities and interpolation, which introduce bias into the model. We advise that the reader take this into consideration when assessing the predicted probabilities from running the various scenarios. BNs are particularly useful in providing a statistically sound approach to representing the initial state of a system and identifying those areas that would benefit most from additional data collection and further in-depth statistical and mathematical modelling.

Although the main objective and aim of Stateless Ethereum remain firm, the ways in which this can be achieved has evolved over time, with some changes standing the test of time, and others being assessed and discarded in favour of more pragmatic proposals. Even at this late stage of development, a major new initiative has been introduced - *Regenesis*, which is rapidly gaining traction in the Stateless Ethereum research community. These changing conditions make modelling more challenging, but these changes can be largely mitigated or minimised by the modular approach taken in building this model, as a new sub-model can be added in the next iteration of the model to represent the changes introduced by regenesis. Moreover, as we focus our attention of data collection to those areas that have a greater impact on the probability of a node keeping up with the network, we will refine this model over time, learning from new data and information as they become available.

An online, interactive version of this BN model is hosted by Hugin on their Demo website[13]. Fur-

---

[13]`https://demo.hugin.com/example/StatelessEthereumModel`

thermore, three blog posts describe the model, providing additional detail, not included in this paper.[14] [15] [16]

References

[1] Vitalik Buterin (2017): *The Stateless Client Concept*. Available at `https://ethresear.ch/t/the-stateless-client-concept/172`.

[2] Robert G. Cowell, A. Philip Dawid, Steffen L. Lauritzen & David J. Spiegelhalter (2007): *Probabilistic Networks and Expert Systems - Exact Computational Methods for Bayesian Networks*. Information Science and Statistics, Springer.

[3] Norman Fenton & M Neil (2019): *Risk Assessment and Decision Analysis with Bayesian Networks*, second edition. Chapman & Hall.

[4] Finn V Jenson & Thomas Dyhre Nielsen (2007): *Bayesian Networks and Decision Graphs*, second edition. Information Science and Statistics, Springer, New York, doi:10.1007/978-0-387-68282-2.

[5] S. Johnson & K. Mengersen (2012): *Integrated Bayesian network framework for modeling complex ecological issues*. Integrated Environmental Assessment and Management 8(3), doi:10.1002/ieam.274.

[6] Sandra Johnson, Laurie Marker, Kerrie Mengersen, Chris H. Gordon, Jörg Melzheimer, Anne Schmidt-Küntzel, Matti Nghikembua, Ezequiel Fabiano, Josephine Henghali & Bettina Wachter (2013): *Modeling the viability of the free-ranging cheetah population in Namibia: an object-oriented Bayesian network approach*. Ecosphere 4(7), p. art90, doi:10.1890/ES12-00357.1. Available at `http://doi.wiley.com/10.1890/ES12-00357.1`.

[7] Uffe B. Kjærulff & Anders L. Madsen (2013): *Bayesian Networks and Influence Diagrams — A Guide to Construction and Analysis*, 2nd edition. Springer, doi:10.1007/978-1-4614-5104-4.

[8] Daphne Koller & Avi Pfeffer (1997): *Object-oriented Bayesian Networks*. In Dan Geiger & Prakash P. Shenoy, editors: *UAI '97: Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence, Brown University, Providence, Rhode Island, USA, August 1-3, 1997*, Morgan Kaufmann, Providence, Rhode Island, pp. 302–313.

[9] Kevin Korb & Ann Nicholson (2011): *Bayesian Artificial Intelligence*, 2nd edition. CRC Press, Chapman and Hall.

[10] T Koski & JM Noble (2009): *Bayesian Networks: An Introduction*. Wiley Series in Probability and Statistics, Wiley & Sons, Ltd, doi:10.1002/9780470684023.

[11] SL Lauritzen, AP Dawid, BN Larsen & HG Leimer (1990): *Independence properties of directed Markov fields*. Networks 20, pp. 491–505, doi:10.1002/net.3230200503.

[12] Bruce G Marcot, J Douglas Steventon, Glenn D Sutherland & Robert K McCann (2007): *Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation*. Canadian Journal of Forest Research, doi:10.1139/x06-135.

[13] Satoshi Nakamoto (2008): *Bitcoin: A Peer-to-Peer Electronic Cash System*. Available at `https://bitcoin.org/bitcoin.pdf`.

[14] Judea Pearl (1988): *Probabilistic reasoning in intelligent systems : networks of plausible inference*, rev. 2nd p edition. Morgan Kaufmann Series in Representation and Reasoning, Morgan Kaufmann Publishers, Inc.

[15] CA Pollino, AK White & BT Hart (2007): *Examination of conflicts and improved strategies for the management of an endangered Eucalypt species using Bayesian networks*. Ecological Modelling 201, pp. 37–59, doi:10.1016/j.ecolmodel.2006.07.032.

---

[14] Defining Stateless Ethereum: A Journey Into The Unknown

[15] Building A Stateless Ethereum Model

[16] Measuring The Health Of A Stateless Ethereum Ecosystem

[16] O Pourret, P Naim & B Marcot (2008): *Bayesian Networks: A Practical Guide to Applications*. Wiley & Sons, doi:10.1002/9780470994559.

[17] M Scutari & J-B Denis (2015): *Bayesian Networks: with Examples in R*. Texts in Statistical Science, CRC Press, Chapman & Hall/CRC.

[18] Paulo Silva, David Vavřička, João Barreto & Miguel Matos (2020): *Impact of Geo-distribution and Mining Pools on Blockchains: A Study of Ethereum*. arXiv:2005.06356.

[19] J Q Smith (2010): *Bayesian Decision Analysis : Principles and Practice*. Cambridge University Press, Cambridge, doi:10.1017/CBO9780511779237.

[20] Wonder Network (2020): *Global ping statistics*. Available at `https://wondernetwork.com/`.

[21] Gavin Wood (2016): *Ethereum: a secure decentralized generalised transaction ledger*. Available at `https://github.com/ethereum/yellowpaper`.