



Blockene: A High-throughput Blockchain Over Mobile Devices

Sambhav Satija and Apurv Mehra, *Microsoft Research India*; Sudheesh Singanamalla, *University of Washington*; Karan Grover, Muthian Sivathanu, Nishanth Chandran, Divya Gupta, and Satya Lokam, *Microsoft Research India*

<https://www.usenix.org/conference/osdi20/presentation/satija>

This paper is included in the Proceedings of the
14th USENIX Symposium on Operating Systems
Design and Implementation

November 4–6, 2020

978-1-939133-19-9

Open access to the Proceedings of the
14th USENIX Symposium on Operating
Systems Design and Implementation
is sponsored by USENIX

Blockene: A High-throughput Blockchain Over Mobile Devices

Sambhav Satija*, Apurv Mehra*, Sudheesh Singanamalla^{†*}, Karan Grover*,
Muthian Sivathanu*, Nishanth Chandran*, Divya Gupta*, Satya Lokam*
*Microsoft Research India [†]University of Washington

Abstract

We introduce *Blockene*, a blockchain that reduces resource usage at member nodes by orders of magnitude, requiring only a smartphone to participate in block validation and consensus. Despite being lightweight, *Blockene* provides a high throughput of transactions and scales to a large number of participants. *Blockene* consumes negligible battery and data in smartphones, enabling millions of users to participate in the blockchain without incentives, to secure transactions with their collective honesty. *Blockene* achieves these properties with a novel split-trust design based on delegating storage and gossip to untrusted nodes.

We show, with a prototype implementation, that *Blockene* provides throughput of 1045 transactions/sec, and runs with very low resource usage on smartphones, pointing to a new paradigm for building secure, decentralized applications.

1 Introduction

Blockchains provide a powerful systems abstraction: they allow mutually untrusted entities (*members*) to collectively manage a *ledger* of transactions in a decentralized manner.

All blockchains today require member nodes to run powerful servers with significant network, storage, and compute resources. Blockchains based on *proof-of-work* [5, 30] push resource usage to an extreme, requiring significant compute for puzzle-solving, but even consortium blockchains [13] and blockchains based on *proof-of-stake* [21] incur significant network and storage costs to keep the blockchain up to date at a high transaction throughput. Blockchains today are therefore limited to use-cases where members have a strong incentive to participate, and can hence afford the high resource cost. For example, in consortium blockchains [13], business efficiency improves, while in cryptocurrencies [21, 30], members earn currency.

Interestingly, the high resource requirement of blockchains also weakens *reliability* for several real-world applications. Blockchains require that majority (typically two-thirds) of members are honest, a property that is easier to guarantee when a large number of members participate. However, wide-scale adoption of a blockchain is hard given the high resource requirement, especially in scenarios where members do not have a direct incentive to participate. Not surprisingly, public

blockchains with high membership today target cryptocurrencies [5, 30].

In this paper, we present *Blockene*¹, an ultra-lightweight, large scale blockchain that provides high throughput for real-world transactions. By being lightweight and scalable, it enables wide-scale adoption by millions of users. By enabling large scale of participation, *Blockene* makes it plausible to assume honest-majority. By being high-throughput, *Blockene* supports real-world transaction rates.

The key breakthrough in *Blockene* is that instead of requiring members to run powerful servers, *Blockene* is the first blockchain that enables members to participate as first-class citizens in consensus even while running on devices as lightweight as smartphones, lowering cost by orders of magnitude.

Network: Blockchains rely on peer-to-peer gossip between members; at a high transaction rate, gossip would require tens of GBs of data transfer per day; *Blockene* requires only about 60MB of data transfer per day on a smartphone.

Storage: Member nodes in blockchains keep a copy of the entire blockchain (terabytes at high-throughput); in *Blockene*, members incur only a few hundred MBs of storage.

Compute: Even the gossip cost of typical blockchains would drain battery on mobile nodes; *Blockene* ensures that battery drain is less than 3% per day.

Thus, a user incurs no perceptible cost while running *Blockene*. As the low resource usage in *Blockene* makes it feasible even in a smartphone, *Blockene* can also run on desktops, with much lighter resource usage than state-of-the-art.

Blockene achieves three conflicting properties: large scale of participation, high throughput, and lightweight resource usage, catering to even scenarios where there is no direct incentive (e.g., altruistic participation), and handling transactions across variety of use-cases including those on public funds. A comparison of *Blockene* with other blockchain architectures is depicted in Table 1.

Example application: Audited Philanthropy. Charitable donations to non-profits are in excess of USD 500 billion annually worldwide [7, 8, 10]. However, from a donor's perspective, the lack of transparency on the end-use of funds makes donations vulnerable to sub-optimal use or mismanagement by non-profits, especially in regions where regulatory enforcement is ineffective or crippled by corruption. A sys-

*Sudheesh was with Microsoft Research India while doing this work.

¹Named after *Graphene*, one of the lightest and strongest materials.

tem that provides a public, end-to-end trail of funds from the donor to the end beneficiary, will exert market pressure on non-profits, besides motivating donors. A blockchain can provide such tracking, but given the scale of funds involved, a small consortium of members cannot be trusted with operation of the blockchain. Ideally, such a blockchain should be jointly controlled by millions of citizens altruistically. Similar requirements arise in government/public spending.

Key techniques in Blockene: *Blockene* adopts a novel system design based on a split-trust architecture with a new security model. There are two types of nodes in *Blockene*: *Citizens* and *Politicians*. *Citizens* run on smartphones and are the real members of the blockchain, i.e., they have voting power in consensus protocol; hence we assume that two-thirds of the *Citizens* are honest (a reasonable assumption with millions of *Citizens*). On the other hand, *Politicians* run on servers and are untrusted, i.e., do not participate in consensus. *Politicians* are fewer in number (few hundreds), and we require *only* 20% of them to be honest. Although *Politicians* do the heavy work such as storing the blockchain, our protocols ensure that *Citizens* can detect and handle malicious behavior even if 80% of the *Politicians* collude with the one-third of malicious *Citizens*. *Citizens* deal with high dishonesty of *Politicians* by using a new primitive called *replicated verifiable reads*: the *Citizen* reads the same data from multiple *Politicians* and can get the correct value even if one (out of, say, 25) is honest.

Citizens perform transaction validation, and decide on the block and resulting global state to commit, by running Byzantine consensus. To make consensus feasible with millions of *Citizens*, *Blockene* borrows an idea from Algorand [21] (modified to make it battery-friendly), where a different random committee of (~2000) *Citizens* is cryptographically chosen to run consensus for each block. Unlike Algorand, *Blockene* exposes the set of committee members a few minutes before their participation: this enables *Blockene* to reduce data and battery cost at *Citizens*. While this may appear to increase the window for a targeted attack on the committee, we discuss in § 4.2 why this is not a serious concern.

To keep storage/communication costs at the *Citizens* low, only *Politicians* store the blockchain and the global state (i.e., key-value pairs), freeing *Citizens* from gossiping all blocks (~50GB/day). *Citizens* only read a small subset of data from *Politicians* (e.g., key-values for transactions for the current block), and write out the new block. Further, because *Politicians* are untrusted, *Citizens* cannot rely on the correct latest values returned by them for, say, a given key. *Blockene* uses a novel technique of *sampling-based Merkle tree read/write* that reduces communication cost while ensuring tolerance to 80% malicious *Politicians*.

When in the committee, *Citizens* reduce their communication cost by not gossiping directly, but through *Politicians*; data written by a *Citizen* gets gossiped among *Politicians*, and interested *Citizens* read from *Politicians*.

As participation in *Blockene* is lightweight, the system

needs to protect against Sybil attacks [17]; preventing an adversary from spinning up lots of virtual nodes to get disproportionate voting share. To thwart such attacks, *Blockene* requires the participant identity to be certified by the trusted hardware (TEE) available in most smartphones [6, 11], and enforces that each TEE can have at most one active identity on the blockchain, thus raising the economic cost of participation to the cost of a unique smartphone.

To limit damage that 80% malicious *Politicians* can cause to performance, *Blockene* employs several techniques to restrict their ability to lie. First, we use a technique called *pre-declared commitments* to make some malicious behaviors detectable. Second, to perform gossip among *Politicians* reliably and efficiently despite 80% dishonesty, we introduce a novel technique called *prioritized gossip*. These techniques reduce cost at *Citizens*, enabling *Blockene* to achieve high throughput despite running on smartphones.

We have built a prototype of *Blockene*; the *Citizen* node is implemented as an Android application, and *Politician* node is implemented as a cloud server. We evaluate *Blockene* along various dimensions, and show that it achieves good transaction throughput of 1045 transactions/sec (6.8 MB/min) while ensuring a commit latency of 270s in the 99th percentile. We also demonstrate very little data use (61 MB/day) and battery use (3%/day) at *Citizens*.

The key contributions of this paper are as follows:

- We present the first blockchain system where nodes can participate as first-class members in consensus while running on devices as lightweight as smartphones, supporting high scale of members and high throughput.
- We present a novel split-trust design with a new security model comprised of resource constrained *Citizens* (honest majority) and resource heavy *Politicians* (dishonest majority), and *Citizens* performing validation and consensus by offloading heavy work to untrusted *Politicians* in a verifiable way.
- We make several novel optimizations (e.g., pre-declared commitments, sampling-based Merkle tree read/write, prioritized gossip) that achieve good performance despite 80% malicious *Politicians*.
- With a thorough theoretical analysis, we prove that *Blockene* satisfies *safety*, *liveness*, and *fairness*.
- We perform a thorough empirical evaluation of this architecture, demonstrating its feasibility as a shared scalable blockchain service.

The rest of the paper is structured as follows: In § 2, we provide a background on blockchains, and discuss existing blockchain architectures in § 3. § 4 provides an overview of *Blockene*, and its threat model, and § 5 presents its design. We discuss optimizations for resource-heavy steps in § 6, present an overview of safety and liveness proofs in § 7, and describe the implementation in § 8. We evaluate *Blockene* in § 9, and conclude (§ 10).

2 Background

In this section, we discuss the key principles and abstractions in a blockchain, and its applications.

2.1 Basic properties

A blockchain is a *distributed ledger* of transactions. Without a trusted authority (*e.g.*, a bank) managing the ledger, a group of mutually untrusted parties collectively validate transactions, and maintain a consistent ledger, provided at least a threshold of participants (*e.g.*, two-thirds) are honest. A blockchain must provide *safety*, *liveness*, and *fairness*. Safety ensures that honest participants have a consistent view of the ledger. Liveness ensures that malicious participants cannot indefinitely stall the blockchain by preventing new block additions. Fairness ensures that all valid transactions submitted to the blockchain get eventually committed.

2.2 Building blocks

A blockchain is a replicated, peer-to-peer distributed system built on the following basic primitives:

Merkle tree for Global State: A key part of a blockchain is the *global state* database that tracks keys and their current values. This global state is managed in a tamper-proof manner, typically using a Merkle tree where the leaf nodes contain the key-value pairs, while each intermediate node contains a hash of the concatenated contents of child nodes. The root is a single hash value that represents the entire state. An update of a key requires recomputation of hashes only along the *path* from that leaf to the root. Given the root, the value of any key can be proved by a path of valid hashes to the root.

Signed transactions: The basic unit of work in a blockchain is a *transaction*. A transaction reads and updates a few keys in the global state (*e.g.*, transfer \$1000 from Alice to Bob). To be valid, (a) the transaction must be signed (b) the user signing the transaction must have access to the keys (c) “semantic” integrity must pass (*e.g.*, cannot overspend).

Cryptographic linkage: A blockchain is a list of *blocks*. A block is a list of transactions. The ordering of blocks is ensured by a *cryptographic linkage*; every block embeds the cryptographic hash of the previous block’s contents.

Gossip: Participants in a blockchain exchange state with each other in a peer-to-peer fashion. For example, when a new block gets committed to the ledger, it must be sent to other members. This communication happens through multi-hop *gossip*, with eventual consistency.

Consensus Protocol: The key primitive in blockchains is a *distributed consensus* protocol that handles *Byzantine failures* (*e.g.*, PBFT [15], Nakamoto [30], or BBA [21]), as minority of participants could be malicious. Byzantine consensus requires at least 2/3rd participants to be honest, and requires several *rounds* of communication.

3 Comparison with Existing Blockchains

In this section, we present a brief survey of related work on existing blockchain architectures. *Blockene* provides three properties: lightweight resource usage, large scale of participation, and high transaction throughput. We use the same three dimensions to compare *Blockene* with related work.

3.1 Resource usage by member nodes

Existing blockchains span a wide spectrum in resource usage by participating member nodes, depending on the mechanism used for consensus. We first discuss compute cost incurred by members, and then the network and storage cost.

Compute Cost. In terms of compute cost, the most expensive are blockchains based on Nakamoto consensus [30], also referred to as *proof-of-work*; examples are Bitcoin [30] and Ethereum [5]. In Nakamoto consensus, the first member node to solve a compute-intensive cryptographic puzzle is chosen as the winner in committing a new block. Such blockchains therefore require heavy compute resources at member nodes.

In order to address the high compute (and energy) costs of proof-of-work blockchains, two popular alternative architectures have emerged. The first is *consortium blockchains* (*e.g.*, HyperLedger [13]), which, by limiting the blockchain membership to a small number of nodes, can run traditional Byzantine consensus algorithms, instead of the compute-intensive proof-of-work based consensus. The second architecture is *proof-of-stake* blockchains, which tie the voting power of a member node with the amount of money the member node has on the blockchain. Examples of these blockchains are Algorand [21], Ouroboros [14, 22], PeerCoin [23], *etc.*. Inherently, proof-of-stake blockchains target cryptocurrency applications where such a “stake” is meaningful.

Network and Storage cost. While the above two architectures, *i.e.*, consortium blockchains and proof-of-stake blockchains, address the raw compute cost of member nodes, they are still too expensive for smartphones. In particular, they are heavy on network and storage resources, as they require the member nodes *to be always up-to-date* with the “current” state of the blockchain. Given the high transaction rate (1000s of transactions per second) that such blockchains enable, replication of the entire state across member nodes is expensive: at 1000 transactions/sec, the blockchain would commit roughly 9GB per day, which needs to be gossiped across member nodes, resulting in a network cost of roughly 45 GB/day (assuming a gossip fanout of 5 neighbors) that every member node has to incur. Further, such a blockchain would consume terabytes of storage on member nodes, as every member node stores a local copy of the blockchain.

Even blockchains that target smartphones [35] adopt the same philosophy of member nodes staying up to date, and thus incur the network and storage overheads.

Some blockchains address storage cost by *sharding*. OmniLedger [25] is a recent blockchain that allows participants to only store a *shard* of the blockchain. It uses a variant of Byzcoin [24] for fast consensus. RapidChain [37] also uses sharding to reduce storage cost. Both these works scale only to a few thousand participants and also require participants to store a large fraction ($\frac{1}{3}$ or $\frac{1}{16}$) of the entire blockchain.

Lightweight but Incapable Nodes. A class of “lightweight” blockchains adopt an approach of “unequal members”: only the first-tier, resource-heavy members participate in consensus and have voting power, while the second-tier members simply serve as read-only query frontends, and do not participate in consensus. In such a model, the “majority-honest” property must be met purely by the heavy nodes, as light nodes do not contribute to security. Not surprisingly, given the limited responsibility, the “light” nodes don’t consume much resources. An example of this architecture is the separation between light and heavy nodes in Ethereum [32].

Blockene. In contrast, *Blockene*, achieves lightweight resource usage for *first-class members* that participate in consensus and block validation. Further, unlike Ethereum which depends on honest majority among heavy nodes (only heavy nodes can vote), *Blockene* tolerates up to 80% of the “heavy” nodes (*i.e.*, *Politicians*) being corrupt. Members in *Blockene* require only a smartphone and negligible² data transfer (< 60 MB/day, *i.e.*, three orders of magnitude lower) and negligible compute (battery use of <3% per day). It achieves this by enabling member nodes to operate with minimal state needed for committing a particular block, and performing work only a few times a day, *i.e.*, not striving to stay up-to-date always.

3.2 Scale of participation

As the security of a blockchain fundamentally relies on a majority of the participating members being honest, blockchains need to protect against collusion of a large number of participants. Consortium blockchains [13] carefully structure the blockchain for a particular business process, such that members have a shared incentive in the success of the blockchain. It is sometimes infeasible/hard to structure a consortium with the above guarantee; in the philanthropy example, if a small number of members are in control of the blockchain, they may collude to, say, facilitate siphoning of donations meant for the poor. Moreover, a consortium blockchain is intricately tied to a specific business process among a set of entities, resulting in high setup and operational overhead, besides limiting inter-operability.

Another approach to guard against collusion among majority, is to enable large scale participation; by onboarding a large number of participants (say millions), majority-collusion can be made hard and unlikely. Most “public” blockchains such as Bitcoin [30], Ethereum [5], and Algorand [21] enable

²Cellular data costs in several countries are much cheaper than in the US [18]; in US/Europe, users are on WiFi/broadband most of the day.

large-scale participation. *Blockene* also supports a large number of participants, but unlike most public blockchains today that target cryptocurrencies, *Blockene* is not tied to cryptocurrency (*e.g.*, no proof-of-stake), but enables generic business transactions. Unlike consortium blockchains, *Blockene* can additionally enable real-world scenarios where there is potential for collusion among a small number of members.

3.3 Transaction throughput

Public blockchains based on proof-of-work are low in throughput (~4-10 transactions/sec). Proof-of-stake based Algorand [21] is the first public blockchain with ~1000 transactions/sec³. Consortium blockchains, due to low scale of participants and traditional consensus (*e.g.*, PBFT), provide 1000s of transactions/sec. Similar to Algorand, *Blockene* also provides a high transaction throughput. By not being tied to cryptocurrency applications, *Blockene* can serve traditional business applications similar to consortium blockchains.

<i>Blockchain</i>	<i>Scale of members</i>	<i>Trans. rate</i>	<i>Cost</i>	<i>Incentive needed?</i>
Public (<i>e.g.</i> , Bitcoin)	Millions	4-10 /sec.	Huge (PoW)	Yes
Consortium (<i>e.g.</i> , [13])	Tens	1000s /sec.	High	Yes
Algorand [21]	Millions	1000-2000/sec.	High	Yes
Blockene	Millions	1045 /sec.	Tiny	No

Table 1: Comparison of blockchain architectures.

3.4 Incentives to Participants

Because of high resource cost (compute, network, or storage), existing blockchains need an incentive for participants (*e.g.*, mining coins in cryptocurrencies, or business efficiency in consortiums). Blockchains that depend on such incentives cannot work for applications such as philanthropy (§ 1). To scale without incentives and to enable *altruistic* participation, the cost of participation has to be negligible.

Table 1 compares blockchain architectures along these dimensions. *Blockene* is the first blockchain to achieve all of the above: scale, throughput, and low cost. With low cost, *Blockene* supports real-world use-cases even where participants do not have a direct incentive, but are altruistic to run a background app with negligible battery and data usage.

3.5 Other related work

The committee-based consensus in *Blockene* is heavily inspired by Algorand [21]; Like *Blockene*, Algorand also

³Assuming 100-byte transactions and 2.2 MB in 20s, 10MB blocks @750MB/hr.

does not allow forks to occur, and one consistent view of the blockchain is always maintained. There is a tradeoff between Algorand and *Blockene* on the resilience to two kinds of targeted attacks (described in § 4.2 para 1). HoneyBadger [28] is a recent system designed for consortium blockchains with $O(100)$ participants. IOTA [19,20] is another distributed ledger system, but currently relies on a centralized co-ordinator for consensus.

Among proof-of-work-based blockchains, the most closely related work to *Blockene* is Hybrid consensus [31]. Similar to Algorand (and *Blockene*), Hybrid consensus periodically selects a group of participants and does not allow the adversary to corrupt nodes during the “participant selection interval”. However it has a long selection interval (of about 1 day) and is also open to the possibility of forks.

4 Architecture Overview

In this section, we first introduce our two-tier architecture that achieves the three conflicting properties of lightweight resource usage, large scale of participation, and high transaction throughput. We then discuss the threat model of *Blockene*.

4.1 Two-tier Architecture

Blockene employs a novel two-tier architecture with asymmetric trust. This architecture is depicted in Figure 1.

There are two kinds of nodes in *Blockene*: *Citizens* and *Politicians*. *Citizens* are resource-constrained (*i.e.*, run on smartphones), are large in number (millions), and are the only entities having voting power in the system (*i.e.*, participate in consensus). *Politicians* are powerful and run servers (similar to existing blockchains like Algorand), and are lot fewer in number (low hundreds), but they do not have voting power. *Politicians* only execute decisions taken by *Citizens*, and cannot take any decisions on their own.

The low resource usage enables a large number of *Citizens* to participate without incentives, while *Politicians* being few in number, will be run by large entities that have interest in the particular use-case (*e.g.*, in the audited philanthropy case, large donors and foundations).

As *Citizens* participate in consensus, at least two-thirds of *Citizens* are required to be honest, while others can be malicious and collude. This is reasonable as *Blockene* allows millions of *Citizens*, making large-scale corruption hard. However, *Politicians* enjoy much lower trust. *Blockene* only requires 20% of politicians to be honest; the remaining 80% of the politicians can be malicious and collude among themselves, and with one-third malicious *Citizens*.

4.1.1 Offloading work to Politicians

Intuitively, given the two-tier architecture, *Citizens* can offload expensive responsibilities such as storage and commu-

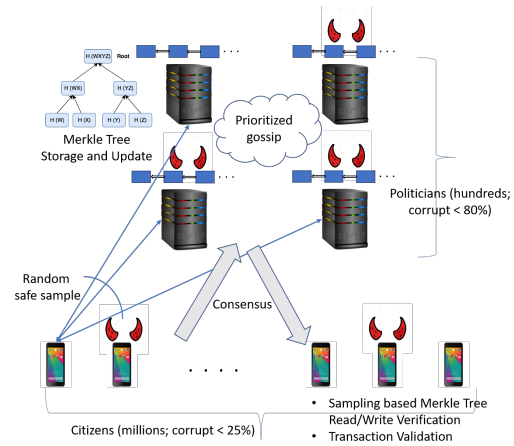


Figure 1: *Blockene*'s architecture

nication to *Politicians*. However, as 80% of *Politicians* are corrupt, a write made by a *Citizen* could just be dropped by a *Politician* or, a read could return incorrect value. To get useful work done out of *Politicians* *Blockene* uses a novel mechanism of replicated reads and writes. Reads and writes by *Citizens* to *Politicians* happen with a random *safe sample* of *Politicians*. The size of this sample is fixed such that with high probability, at least one *Politician* in the sample is honest (*e.g.*, for a sample size of 25, this probability is $1 - (0.8)^{25} = 99.6\%$). *Blockene* is resilient to a small number of *Citizens* (0.4%) picking all dishonest *Politicians*.

4.1.2 Division of responsibilities

We now describe how the *Citizens* and *Politicians* collaborate to perform the various standard blockchain tasks:

Storage: In a traditional blockchain, every participant keeps a replica of the entire blockchain, but *Citizens* in *Blockene* cannot afford to store TBs of data. In *Blockene*, only *Politicians* store the ledger and the global state (*i.e.*, database of key-values § 2). *Citizens* read subsets of this data from *Politicians* as needed. The only state *Citizens* store (and periodically update) is a list of valid *Citizen* identities (§ 5.3).

Transaction Validation: As *Citizens* are the actual participants in consensus, they validate transactions, ensuring that transactions are signed, and have semantic integrity (*e.g.*, no double-spending). To perform validation, *Citizens* read transactions from *Politicians*, and lookup latest values of the keys referenced in them, from the global state with *Politicians*. *Citizens* then propose a block with valid transactions.

Gossip: To ensure that all honest participants agree on the state of the blockchain, participants need to gossip among themselves. However, as discussed in § 3, direct gossip among *Citizens* is expensive. *Blockene* solves this problem by having *Citizens* gossip through *Politicians*. When a *Citizen* needs to broadcast information to other *Citizens*, it sends a message to a *safe sample* of *Politicians*. *Politicians* then gossip data

among themselves; they can afford to do so because they have good network connectivity. Other *Citizens* then perform a replicated read from the *Politicians* when they need to, e.g., when they are in the committee⁴. For gossip through *Politicians*, we need the guarantee that a message that reaches one honest *Politician* always reaches all other honest *Politicians* via gossip, a challenging property when 80% of the *Politicians* are malicious; our custom gossip protocol is described in § 6.1). Thus, we achieve the same semantics as direct gossip among *Citizens*, but with minimal network load on *Citizens*.

Consensus: *Citizens* participate in consensus by performing gossip through *Politicians*. Given the large scale of *Citizens*, all *Citizens* cannot participate in consensus. Instead, we cryptographically select a random *committee of citizens* (roughly 2000 members) for each block (§ 5.2).

4.2 Threat Model

While our threat model is similar to Algorand [21], there is a tradeoff between Algorand and *Blockene* on the resilience to targeted attacks. On one hand, Algorand is based on proof-of-stake, which allows an adversary infinite time to target nodes with higher stake (who will appear in the committee more frequently); *Blockene* avoids this attack, as all *Citizens* have equal votes. On the other hand, Algorand protects the secrecy of the committee members until they perform their role, but *Blockene* exposes their identities a few minutes (1-2 blocks) before they participate. To conserve battery, *Citizens* normally poll *Politicians* for current state of the blockchain roughly every 10 blocks (§ 5.2), but when they are going to be in the committee, will poll again shortly (e.g., 1 block) before their expected turn, thus exposing their identity to malicious *Politicians*. This potentially provides a window for a targeted attack (e.g., by bribing the committee: § 4.2.1).

4.2.1 Attack vector of *Citizens*

Bribing attack on *Citizens*: As *Blockene* implicitly exposes the public keys of the committee a few (e.g., 2) minutes in advance, an adversary could in theory perform a targeted attack by bribing a sufficient number of committee members. However, we believe this is not a concern for the following reasons. First, with just the IP address, it is non-trivial for an adversary to “send a message” offering bribe to a *Citizen*, because of carrier-grade NAT [4] and the architecture for push notifications in smartphones; the existing channel from a malicious *Politician* to the *Citizen* cannot be misused for this, as an untampered *Blockene* app on the *Citizen* will ignore any spurious traffic on that channel. Second, as the committee is randomly chosen every block, pull-based bribing where the *Citizens* (who know of their selection up to 10 blocks in advance - § 5.2) pro-actively reach out to the adversary

⁴Direct gossip among *Citizens* would require all *Citizens* (i.e., including those outside the committee) to participate in gossip of all data.

cannot happen, as that would imply violation of the honesty assumption on *Citizens*, i.e., greater than 70% being honest.

Sybil Attack by *Citizens*: Given the lightweight cost of participation, *Blockene* needs to ensure that an adversary cannot get disproportionate share of voting by spinning up several virtual nodes (i.e., Sybil attacks [17]). A common way of addressing Sybil attacks is *Proof-of-work* which is resource-intensive and does not fit the goals of *Blockene*; another alternative is *Proof-of-stake* [21] where a participant’s voting power is proportional to the amount of “stake” (money) on the blockchain, but it is specific to cryptocurrencies.

In *Blockene*, we protect against Sybil attack by exploiting the trusted hardware (TEE) available in smartphones [6, 11], and ensuring that a smartphone can have at most one identity on the blockchain. Thus, *Blockene* imposes an economic cost to participation, i.e., the cost of a smartphone; this is sunk cost *already incurred* in owning the smartphone, but protects against Sybil as each identity is a unique smartphone.

In particular, each TEE has a unique public key that is certified by the platform (Android/iOS) vendor. The TEE can certify an EdDSA public-private keypair generated by an app; this generated public key serves as the identity on *Blockene*. The global state of *Blockene* tracks the set of valid public keys, along with the public key/certificate of the TEE that authorized it. When a transaction for adding a new member is proposed, *Blockene* looks up the TEE public key to see if that TEE (i.e., the same smartphone) already has an identity in *Blockene*; if yes, it rejects the transaction⁵. Thus, every *Citizen* on *Blockene* is tied to a unique smartphone, making it economically infeasible/unattractive for a single entity to get large participation on *Blockene*.

Note that *Blockene* only assumes that every certificate signed by Google/Apple for a TEE public-key corresponds to a unique smartphone. It does not depend on the security of an individual TEE (unlike running the blockchain consensus inside TEE, e.g. SGX [33], that opens up side-channel attacks compromising integrity and security). As a result, the TEE identity can be replaced/combined with other unique identities. In India, one-way-hash of Aadhaar-ID [1, 12] (digitally verifiable, biometric-deduped, 1.2 billion-reach) can be used. Other de-duped IDs (e.g. SSN) augmented with digital verifiability can also be used.

4.2.2 Attack vector of *Politicians*

Dealing with 80% dishonesty among the politicians is one of the main technical challenges in the design of *Blockene*. Malicious behavior by *Politicians* falls under two kinds: detectable and covert. Detectable maliciousness where there is a succinct proof of lying, can be used to improve performance by blacklisting. For example, if a *Politician* is supposed to only send one group of transactions in a round, but there are

⁵We can also support replacing the old identity with the new one for the same TEE with appropriate bookkeeping.

two versions signed by the same *Politician*, it is detectable with proof. Covert maliciousness is harder to handle, and is the focus of our techniques. We list broad (non-exhaustive) classes of covert attacks a *Politician* can employ.

Staleness Attack: When a *Citizen* node asks the *Politician* for some state (e.g., the latest committed block), the *Politician* could return a stale block. Such a response would appear to be valid because the old block would also have been signed by a quorum of *Citizens* (§ 5.3).

Split-View Attack: A *Politician* can respond selectively to some *Citizens* and not to others, causing a split in the world-view seen by honest *Citizens*. Worse, a *Politician* can respond with two different values to different subsets of *Citizens*. In a coordinated split-view attack, the malicious *Politicians* could only gossip among themselves, so that no honest *Politician* has a certain data. Malicious *Politicians* can then selectively relay this data only to some *Citizens* (e.g., § 5.5.2).

Drop Attack: A malicious *Politician* may drop data written by a *Citizen* without committing it or gossiping it to other *Politicians*. Similarly on a read, the *Politician* may choose to not respond, even though the *Politician* has the data (§ 4.1.1).

Denial-of-Service Attack: As *Politicians* are powerful servers typically hosted in the cloud, we assume that honest *Politicians* employ standard DoS protection that public clouds offer [2, 3]. For *Citizens*, most ISPs employ carrier-grade NAT to handle the explosion of IP addresses on mobile phones [4], which also provides DoS protection. Malicious *Politicians* can make our gossip protocol more expensive by asking for more data than they need (§ 6.1).

Sybil Attack: An adversary could try pushing the dishonesty fraction of *Politicians* beyond 80% by spinning up several nodes. However, given the small number (say 200), we envision that *Politician* nodes would have an out-of-band registration mechanism (e.g., mapping them to real entities, say one per Fortune-500 company) - robust because only 20% of them need to be honest (unlike *Citizens*).

Blockene protects against both detectable and covert maliciousness of the *Politicians* including the attacks listed above.

5 Design

In this section, we present in more detail how *Citizens* and *Politicians* coordinate on the key steps in *Blockene*.

5.1 System Configuration

We first outline the system configuration for *Blockene*. *Citizens* in *Blockene* run on a smartphone, so we assume that their network bandwidth is low, i.e., 1 MB/s. We choose a block size of 9MB (to amortize fixed cost per block), containing about 90k transactions (~100 bytes each including a 64-byte signature). We assume a network bandwidth of 40 MB/s between *Politicians* (representative of bandwidth in the cloud, e.g., between an Azure and a Google Cloud VM across east-US and west-US). We choose the number of *Politicians*

as 200. The work done per block only depends on committee size, so the system scales to millions of *Citizens*.

Transaction originators submit signed transactions to a safe sample or to all *Politicians*, continuously in the background. Transactions can modify keys that the originator has access to. Transactions from the same originator can depend on each other; we preserve their order by tracking a per-originator *nonce* in the global state. In this paper, (without loss of generality) each transaction accesses three keys (debits one key and credits another, third key is nonce). *Politicians* gossip transactions among each other.

5.2 Selecting Committee of *Citizens*

The committee of citizens for validating and signing each block is chosen on the basis of a *VRF* (*Verifiable Random Function*) [27], inspired by Algorand [21] but with one key modification. Algorand requires each participant to check in each round whether it is chosen in the committee. A *Citizen* on a mobile phone cannot afford to do such frequent checks because waking up the phone every round and communicating would cause significant battery drain. Therefore, instead of computing the VRF on the hash of the previous block ($N - 1$), *Blockene* uses the hash of block $N - 10$, thus allowing a *Citizen* to wake up once every 10 blocks. Note that this modification still preserves the security guarantee required from VRFs in our threat model. Specifically, for a citizen, the VRF for block N is calculated as $\text{Hash}(\text{Sign}_{\text{sk}}(\text{Hash}(\text{Block}_{N-10})||N))$ where sk is private key known to the citizen.⁶ A *Citizen* is in the committee if the VRF has 0's in the last k bits (hence a *Citizen* is part of a committee with probability 2^{-k} ; k can be set appropriately). Only the concerned *Citizen* can generate the VRF as it requires its private key, but anyone can verify its validity based on the public key given the signature.

Committee size: The size of the committee needs to balance performance and security. A small committee is good for performance, but for security of consensus protocol, we require that in any committee, at least $2/3$ *Citizens* are honest. As our committee selection is probabilistic, by the Chernoff bound [29], this security requirement cannot be met for very small committee size even if we have $2/3$ honest *Citizens* overall. Committee size increases with the fraction of dishonest *Citizens*. We calibrate this tradeoff to obtain an expected committee size of 2000 with a citizen dishonesty threshold of 25%. The details of these computations appear in the full version [34]. We give an overview below.

Proof overview: We prove several properties about the committee for a block. We call a *Citizen* that participates in a committee as *good* if the *Citizen* is honest and speaks to at least 1 honest *Politician* through m fan-out read/write. Otherwise, we say that the *Citizen* is *bad*. For a configuration with 25% corrupt *Citizens*, 80% corrupt *Politicians*, and $m = 25$,

⁶We use EdDSA signatures. ECDSA uses random number which the adversary can exploit to brute-force itself into the committee.

we show that our committee satisfies the following properties⁷: size of all committees lies in the range [1700..2300] (Lemma 1), every committee has at least 1137 good citizens (Lemma 2), every committee has at least a 2/3 fraction of good citizens (Lemma 3), and no committee has more than 772 bad citizens (Lemma 4).

5.3 Fork-proof Structural Validation

Blockene is designed to prevent forks from occurring. To enable this, each *Citizen* periodically verifies the *structural integrity* of the blockchain to enforce that the chain of hashes and VRFs are consistent and to prevent forks.

Track local state: Each *Citizen* locally remembers the block number N until which the *Citizen* validated the structural integrity of the blockchain, and the hashes of blocks N to $N - 9$. In addition, a *Citizen* stores an up to date list of public keys of other valid *Citizens*. The total storage size is <100MB for 1 million *Citizens*.

Chained ID sub-blocks: To enable *Citizens* to efficiently update local state, the public keys of new users added as part of each block B , are tracked in an *ID sub-block (SB)* within B . SBs are chained together by embedding $\text{Hash}(SB_{i-1})$ within SB_i . To aid cheap verification, committee members sign $\text{Hash}(\text{Hash}(B_i), \text{Hash}(SB_i), \text{GlobalStateRoot}(B_i))$.

Incremental Validation: Roughly every 10 blocks (12-15 mins), each *Citizen* performs a `getLedger` call to validate the incremental structural integrity (*i.e.*, from last validation point to the latest state), and to check if it will be in the committee soon (committee for a block N is a function of the hash of block $N - 10$). To find the latest block, a *Citizen* queries a safe sample of *Politicians* for the latest block number. It picks the highest number reported by *any Politician*, and asks for proof, *i.e.*, signatures of committee of that block and the corresponding VRFs. Thus, if at least one *Politician* in the safe sample is honest, the *Citizen* will know the latest block hash. If the latest block is greater than $N + 10$, it first verifies block $N + 10$. Further, it refreshes its set of valid public keys by downloading the chained sub-blocks $SB_{N+1} \dots SB_{N+10}$ that contain new *Citizens* added in each block, verifying the integrity of SB_i based on the chained hashes.

Cool-off period for new nodes: To prevent a (low-probability) attack where an adversary can manufacture public-private keypairs⁸ to increase chances of getting higher malicious fraction for a particular block N , we allow a *Citizen* to be in the committee only $k(= 40)$ blocks after the block in which the *Citizen* was added. To verify this as part of VRF checks, a *Citizen*'s local state tracks the block number of "recently" added *Citizens*. This is similar to the "look back parameter" in Algorand [21].

⁷All references to Lemma/Theorem/Algorithm numbers below and rest of this paper refer to those in the full version.

⁸Android TEE API does not allow directly signing with the private key of TEE; instead a keypair is certified by TEE.

Proof overview: Our `getLedger` protocol [34] is used for verifying ledger height $i + 10$, given the *Citizen* v has last verified height i , *without an explicit brute-force verification of signatures of all 10 blocks*. The algorithm generalizes to verifying any height $i + j$ for $1 \leq j \leq 10$. We show (Lemma 5) that if a *good Citizen* with a verified state for height i invokes the `getLedger` protocol at round $(i + 11)$ and accepts, then the *Citizen*'s updated *structural state* is consistent with the blockchain up to height $(i + 10)$. Using this, we can show that honest *Citizens* can obtain the consistent structural state of the blockchain, along with all registered public keys, for every round of the protocol (Corollary 2).

5.4 Transaction Validation

Citizens perform the task of verifying signatures of transactions, checking the transaction nonce to detect replay attacks, and verifying semantic correctness of the transaction (*e.g.*, double spending). However, only *Politicians* store the Merkle tree (§ 2.2) of the global state; keeping a large and up to date global state in *Citizens* is unaffordable. To validate a transaction, a *Citizen* must lookup the correct value of keys referenced therein. On commit, the *Citizen* must *update* the Merkle tree with new values from the transaction, and sign the new Merkle root. The challenge lies in doing so correctly given untrusted *Politicians*.

The Merkle root (along with block number) is signed by the committee of the previous block, so the *Politician* cannot lie about the Merkle root. Once the *Citizen* learns the latest block number (§ 5.3), it learns the correct Merkle root as well. To verify a value returned for a key, *Citizen* asks the *Politician* to send the *challenge path* for this key, *i.e.*, all the sibling nodes (hashes) along the path from the leaf to the root. This enables the *Citizen* to reconstruct the Merkle path and match the root hash with the signed Merkle root. By security of hashes, the *Politician* cannot present spurious challenge paths that verify. In a tree with 1 billion key-value pairs, the challenge path would contain 30 hashes.

Update of keys in the Merkle tree follows a similar protocol. The *Citizen* could build a partial Merkle tree with the new values at the leaves, and compute the new Merkle root. Both the read and update paths mentioned above are expensive, and we optimize them in § 6.

5.5 Block Proposal

Like in any blockchain, committee members can propose a new block for committing to the blockchain.

5.5.1 Pick winning proposer

For efficiency, we allow only a subset of committee members called *proposers* to actually propose a block, based on the *VRF* of the *Citizen*. For this selection, we use an additional

VRF that is based on the hash of the previous block $N - 1$ (instead of $N - 10$); only committee members who have the last k' bits of the additional VRF set to zero can propose a block, and the winner is the one with the least VRF. Using the previous block hash in this VRF ensures that the adversary does not know about the proposers until the last minute (similar to Algorand) thus preventing a targeted attack on the proposers. Any committee member can consistently determine the winning VRF among the proposers. All proposers upload their block to *Politicians* and other committee members download the block of the winning proposer.

5.5.2 Pre-declared commitments

The upload of the proposed block by a proposer needs to be done to a safe sample of 25 *Politicians*. In *Blockene*, as the blocks are ~ 9 MB in size, assuming 1MB/s bandwidth at mobile nodes, this would take 225 sec. To optimize this step, we make the transaction selection process deterministic, so that any *Citizen* can reconstruct what the original proposer would have done, without the proposer explicitly uploading the full block. Determinism is challenging, however, because the 80% malicious *Politicians* can send different transactions to different *Citizens*. Our technique of *pre-declared commitments* to transactions addresses this.

1. Freeze Transactions. At the start of block N , each *Politician* freezes the exact set of transactions it will send to *Citizens* reading from it. It does so by creating a *tx_pool*, which includes a set of (about 2000) transactions, and then generates a *commitment* which is a signed hash of the *tx_pool* along with the block number⁹. Malicious *Politicians* are forced to issue only one commitment for a given block N , because two signed commitments from a *Politician* is a proof of malicious behavior, and can be used for efficient blacklisting; *Citizens* then drop all commitments from that *Politician* in the same round. Intuitively, with frozen commitments, a *Citizen* proposing a block, need not upload the full block, but only a *digest* with the commitments that went into the block, and other *Citizens* can reconstruct that block by downloading the *tx_pools* for those commitments from *Politicians*.

2. Ensure that enough honest citizens have commitments. A malicious *Politician* can respond with its *tx_pool* only to a subset of *Citizens*, and refuse to respond to others; thus, a *tx_pool* committed in the proposed block may not be readable by all honest *Citizens*, thus thwarting consensus. To address this, we perform three steps. First, we limit the exact set of *Politicians* from whom to pull transactions for a given block to a randomly chosen set of 45 politicians based on the hash of the block number and hash of previous block. Instead of reading *tx_pools* from a random safe sample, a

⁹To reduce overlap of transactions across *tx_pools* from multiple *Politicians* (which would reduce the unique transactions in the final block), transactions are deterministically partitioned across *Politicians* using a hash on transaction identifier and round number. Given a *tx_pool* and commitment, it is easy to detect/ blacklist a *Politician* that doesn't follow this.

Citizen reads from these 45 designated *Politicians* for a block. Second, the *Citizen* uploads a *witness list* to a safe sample of *Politicians*; the witness list contains the list of *tx_pools* the *Citizen* was able to successfully download. The witness list of all *Citizens* gets gossiped between *Politicians*. Third, the proposer reads the witness list of all other *Citizens*, and picks only commitments whose *tx_pools* were successfully downloaded by at least a threshold number of *Citizens*. This threshold is fixed to be $\tilde{n}_b + \Delta$, where \tilde{n}_b is the maximum number of malicious nodes in any committee (computed to be 772, from Lemma 4), and Δ is chosen to be 350. Intuitively, all commitments (and *tx_pools*) that pass this condition are available with at least Δ honest *Citizens*. As 20% of *Politicians* are honest, in expectation, at least 9 out of the 45 commitments will pass this test.

3. Ensure that all honest citizens get commitments. The commitments available with at least Δ honest *Citizens* now need to be propagated to all honest *Citizens*. Each *Citizen* in Step 4, re-uploads 5 random *tx_pools* it has, to 1 random *Politician*. This ensures that (with high probability) each *tx_pool* (including those from malicious *Politicians*) that belongs to at least Δ honest *Citizens* reaches at least one honest *Politician* (who then gossips it to other honest politicians). Thus, other honest *Citizens* can successfully download that *tx_pool* (by querying a safe sample of politicians), preventing a *split-view* attack by malicious *Politicians*.

4. Handle malicious proposer. When the winner of block proposal is a malicious *Citizen*, it need not respect the witness list criteria, and can pick a commitment whose *tx_pool* is known to very few *Citizens*. This attack is possible only when consensus outputs the block proposed by this malicious proposer, so we can argue that at least 1/3 honest *Citizens* had all *tx_pools* at the beginning of the consensus. To ensure that all honest *Citizens* are able to download all required *tx_pools*, a second re-upload of randomly chosen *tx_pools* happens (step 9), now including the downloaded *tx_pools* from previous step. Formal proofs capturing the guarantees provided by these re-uploads needed to prove security of our system are presented in Lemmas 10 and 11.

5.6 Block Commit Protocol

The main operation in a blockchain is adding a new block to the blockchain. We list below the key steps in the process of committing block N . The protocol for block N starts once the previous block $N - 1$ gathers a threshold number of signatures (set to 850 in our case, §E.1 [34]) from the committee members for block $N - 1$.

1. A new committee of *Citizens* is chosen for block N (using Hash of block $N - 10$), denoted by C^N . The *Citizens* in C^N keep polling for the latest committed block number, and start the protocol once that number is $N - 1$.
2. Each *Citizen* C_i^N in C^N downloads *tx_pools* & commitments from $\rho = 45$ designated *Politicians* for the block.

3. Each C_i^N uploads a signed witness list with the commitments it downloaded, to a safe sample of *Politicians*.
4. Each *Citizen* C_i^N picks 5 random *tx_pool* it has, and re-uploads them to 1 random *Politician*.
5. Each *proposer* in C^N downloads all witness lists of C^N from a safe sample of *Politicians*, and picks commitments with at least a threshold (1122) of votes (§ 5.5.2). Then, it makes a *block proposal* with those commitments, along with its VRF to prove proposer eligibility.
6. *Politicians* gossip on block proposals/VRFs and on the *tx_pools* that were re-uploaded by *Citizens*.
7. Each *Citizen* C_i^N tries to download missing *tx_pools* in step 2 from safe sample of *Politicians*, relying on the re-upload (Step 4) by other *Citizens*.
8. Each C_i^N reads the VRFs of all proposers in C^N from a safe sample of *Politicians*, and picks the lowest correct VRF as the *local winner*. If C_i^N already has all *tx_pools* in the winning proposal, it enters consensus with that set of commitments, otherwise, NULL.
9. Each *Citizen* C_i^N performs a second re-upload of 10 random *tx_pools* it has to 1 random *Politician*.
10. *Citizens* in C^N run a consensus protocol (§ 5.6.1) with gossip through *Politicians*, where each C_i^N 's vote is decided in Step 8. At the end, all honest *Citizens* either agree on same set of commitments or an empty block. C_i^N downloads the *tx_pools* missing w.r.t. the output of consensus from safe sample of *Politicians*.
11. Each *Citizen* C_i^N performs transaction validation by downloading challenge paths for all keys from *Politicians* (§ 5.4) and drops transactions that fail validation.
12. Based on valid transactions (Step 11), each C_i^N creates a block, computes the new Merkle root of the global state using updated values of keys and signs the block hash and new Merkle root, along with block number N . It uploads the block hash, new Merkle root, and this signature to a safe sample of *Politicians*.
13. When more than a threshold number of signatures have accumulated for block N , block $N + 1$ starts.

Our complete protocol description can be found in Algorithm 4. We give an overview of various properties of *Blockene*, i.e., safety, liveness and fairness, in § 7.

5.6.1 Consensus Protocol

For consensus (Step 10), we use the Byzantine Agreement (BA) algorithm for string consensus (that is based on [36]) which calls upon the bit consensus algorithm BBA [26] in a black-box manner. These are the same consensus algorithms used by Algorand. *Citizens* enter the consensus protocol with list of commitments in local winning block, as input. Two scenarios are relevant here. If the winning proposer (i.e., the one with the lowest VRF) was honest, which would happen

at least two-thirds of the time, all honest *Citizens* in the committee would enter consensus with this proposal except with small probability (Lemma 10), and the protocol will terminate in 5 rounds. However, if the winning proposer was malicious, it can collude with malicious *Politicians* to partition the view of honest *Citizens*. In general, the consensus protocol would take an expected 11 rounds [21].

6 Optimizations

In this section, we present two key optimizations crucial to achieving high transaction throughput in *Blockene*.

6.1 Prioritized Gossip

Problem. The guarantee we require in *Blockene* is that if one honest *Politician* has a message, all honest *Politicians* receive the message. Because of the high fraction of dishonesty among *Politicians*, standard multi-hop gossip with a small number of neighbors (e.g., 10) cannot provide this guarantee, because there is a non-trivial probability that all of them were dishonest, and drop the message. Hence the safe thing to do is a full broadcast to all other *Politicians*, which is expensive; when *Politicians* need to gossip *tx_pools* that were re-uploaded by *Citizens* in the committee, each *Politician* may have up to 45 *tx_pools* to gossip; with full broadcast, it would send $0.2MB * 45 * 200 = 1.8GB$ which would take 45 seconds in the critical path (@40MB/s).

Key idea. We leverage the fact that messages being gossiped by the different *Politicians* have a high overlap; each *Politician* has a subset of the same 45 *tx_pools* as *Citizens* pick a random *Politician* to re-upload a subset of *tx_pools*. Moreover, given the nature of re-upload, in expectation, any *Politician* would be missing only a few *tx_pools*, and honest *Politicians* wouldn't lie about state.

1. Handshake. Each *Politician* asks recipients B_i which *tx_pools* they already have, and send only the missing ones. While this works with honest *Politicians*, the 80% malicious ones could always lie that don't have any, to cause a higher load/latency on the system.

2. Selfish gossip. As malicious *Politicians* can lie that they have no *tx_pools*, we assign a soft-penalty to *Politicians* that miss a lot of *tx_pools*. Each sender *Politician* A favors the peer B that has the maximum number of *tx_pools* that A needs. In each round, A sends a *tx_pool* to B , and receives one in return. Given the random re-uploads by *Citizens*, each honest *Politician* would be missing only a small number of *tx_pools*, and hence would get prioritized. The list of what B has to offer keeps getting updated as B gets *tx_pools* from other peers; note that this list can only grow, not shrink.

3. Incentivize frugal nodes. Selfish gossip loses its ability to discriminate between honest and malicious recipients, once the sender receives all *tx_pools*. To address this, after getting all *tx_pools*, the sender changes its priority function for destinations B_i to be *the number of tx_pools that B_i claims to have*;

thus honest nodes which will have large fraction of *tx_pools* are favored. Again, the list of *tx_pools* that *B* advertises can only grow, not shrink, as shrinking would mean that *B* lied. Further, each honest *B_i* requests its missing chunk from at most $k = 5$ peers simultaneously; $k = 1$ will be data-frugal, but incur high latency if the peer dishonestly delays response.

6.2 Sampling-based Merkle Tree Read/Write

Problem. The Merkle tree validation in Step 11 is expensive. In a 1-billion node Merkle-tree (30-levels deep), a challenge path is 300 bytes (10-byte hashes); downloading 270K challenge paths is 81 MB (81 sec latency) ignoring compression. The compute at *Citizens* is also high (total 16.2 million hash computations for challenge path verification during read and for computing new root post update).

Key idea. We offload most of this work to *Politicians*, in a verifiable manner. Since the Merkle tree validation is done after the conclusion of the consensus run using gossip through the *Politicians*, *Politicians* know the *tx_pools* that are considered for constructing the block. Hence, all *Citizens* in committee and *Politicians* know the keys whose values need to be read and updated. We first discuss the optimization for reading values correctly from the Merkle tree.

1. Get Values. Each *Citizen* gets just the values for all 270K keys (no challenge path, 1 MB instead of 81 MB) from one *Politician*, and then asks a safe sample of *Politicians* whether those values were correct. As at least one of these *Politicians* is honest, it alerts the *Citizen* to incorrect values through an *exception list*. The *Politician* can “prove” an incorrect value by providing a challenge path from the signed Merkle root that indicates a different value for the key.

2. Spot-checks. If many values were wrong, the exception list would be quite large and eat into the savings. To avoid this, *Citizen* picks a small random subset of $k' = 4500$ keys to initially spot-check using the challenge paths. If the spot-checks pass for a sufficiently large k' , a *Politician* could have lied only for a small number (200) of keys (except with small probability). Thus, the extra spot-checks bound the size of the exception list (Lemma 6).

3. Exception list protocol. To *cross-verify* the values with a safe sample of *Politicians*, the *Citizen* deterministically puts these values into buckets (2000) and uploads the hashes of these buckets. When a *Politician* notices a mismatch for a bucket, it sends the bucket index and the correct values for all keys in that bucket. *Citizen* gets challenge paths only for keys that disagree (from first *Politician*). Our spot-checks ensure that only a small number of buckets can mismatch.

Corner case. Even after doing the above, there is a small probability ($< 2^{-10}$) that a *Citizen* may obtain an incorrect value; we count such *Citizen* nodes as malicious and account appropriately (Lemma 7). The full protocol and all proofs are provided in Algorithm 2.

Writes: Updating the Merkle tree is a trickier problem. Due to lack of old challenge paths for the all keys being updated, the *Citizen* cannot construct the root of the updated Merkle tree T' . We solve this problem by making the *Politicians* compute T' , but now the *Citizen* must verify that the *Politicians* performed the computation correctly, i.e., T' is consistent with the new values of updated keys and old tree T for unmodified keys. We achieve this by breaking T' at a level called the *frontier level* (the nodes at this level are frontier nodes). *Citizens* obtain the values of the frontier nodes of T' from a safe sample of the *Politicians*. The *Citizens* then run a spot checking algorithm - they pick a random subset of frontier nodes and ask a *Politician* to prove the correctness of that frontier node. Next, *Citizens* create exception lists with the help of the rest of the selected *Politicians*. This list denotes which frontier nodes are incorrect with the *Citizen*. The *Citizen* then proceeds to sequentially correct the incorrect frontier nodes and then finally compute the correct root of T' from the frontier nodes.

Proof Overview: In the full paper [34], we prove (in Lemma 6) that for a *good Citizen*, after successfully spot-checking only μ fraction of key-values, only (a small number of) τ values are incorrect with probability $1 - \epsilon_1$ (here, μ , τ and ϵ_1 are appropriately chosen parameters). Moreover, these values will get corrected by processing exception lists of size at most τ . Hence, a *good Citizen* gets correct values with probability $1 - \epsilon_1$ (Corollary 3). We pick our parameters (Lemma 7) such that at most 18 *good Citizens* will obtain incorrect values during read, and account for these 18, by counting them as *bad Citizens* in the committee. In the write protocol, we can show that the sizes of exception lists can be bounded (Lemma 8) and that no more than 18 *Citizens* accept an incorrectly updated Merkle tree T' (Lemma 9), which we once again factor in to the set of *bad Citizens*. We additionally also show that our algorithms are between $3 - 18\times$ more communication efficient and between $10 - 66\times$ computationally faster than the naive algorithm for global state read/write.

7 Proofs of Safety, Liveness, and Fairness

In this section, we provide a brief overview of the proofs in the full paper [34] for the safety, liveness, and fairness guarantees of *Blockene*.

A committee round N ends when a new block gets signed and committed by a threshold number (T^*), of committee members for N . T^* will be set to be 850 (done taking into account maximum number of bad citizens in any committee as well as the 36 good citizens who might have read/written an incorrect global state).

First, we show (in Lemma 10) that for a block, if a *good Citizen* is the winning proposer, then (except with bounded constant probability) all *good Citizens* will output the proposal of this *Citizen* as the output of the consensus protocol. In Lemma 11, we show that, on the contrary, if a malicious *Citizen* is the winning proposer and the consensus results

in a non-null value, then all good *Citizens* will be able to download the transactions committed in the proposal. Using Lemmas 7 and 9 (see Proof Overview of § 6), we then show (Lemma 12) that at the end of the block commit protocol all, except 36, good citizens will sign the same block hash and new global state root and that the new block is consistent with the entire blockchain and global state. Now, using Lemma 12, safety (i.e. all honest *Citizens* agree upon all committed blocks and all blocks are consistent with a correct sequence of transactions) follows via an inductive argument. Next, to argue liveness (that adversarial entities cannot indefinitely stall the system and that the empty-block probability is bounded by a small constant), we use Lemmas 12 and 10.

Additionally, we also prove bounds on throughput in Lemma 13 (in expectation, committed blocks have a threshold number of transactions in them) and fairness in Lemma 14 (all valid transactions will eventually be committed).

8 Implementation

We have built a prototype of *Blockene*, that is spread across two components, *Citizen* nodes and *Politician* nodes.

8.1 Citizen nodes

The *Citizen* node is implemented as an Android app on SDK v23 and has 10,200 lines of code. It is built to optimize battery use and runs as a background app, without user involvement after initial setup. The application caters to two main phases of the protocol that a Citizen participates in: passive and active. In the passive phase, a service using Job-Scheduler [9] periodically polls *Politicians* for `getLedger` calls. In the active phase, when the *Citizen* is part of a committee, the application runs the steps of the protocol, handling failures, timeouts and retries to deal with corrupt *Politicians*. The implementation for the active phase uses a multi-threaded event-driven model and is built on top of EventBus to parallelize and pipeline network and compute intensive crypto tasks such as signature validation.

8.2 Politician nodes

The *Politician* node is implemented in C++ (11K lines of code). The implementation scales to load from thousands of *Citizens*, and handles bursty load during gossip. Given the state-machine nature of the protocol, we have built it on top of the convenient C-Actor-Framework [16], which is based on “actors” that transition the state of the *Politician* through the steps of the protocol. For instance, the BBA actor, apart from storing and serving the votes that *Citizens* submit, also reads the votes to determine the result of consensus. Based on this, it emits an event to build the updated Merkle tree.

For the global state, we have built a SparseMerkleTree (SMT), where the leaf index is deterministically computed using the SHA256 of the key. Since the tree is of bounded depth, we allow for (a small number of) collisions in the leaf

node. The challenge path of any key includes all the collisions co-located with this key, so the leaf hash can be computed. To prevent targeted flooding of a single leaf node, we reject key additions that take a leaf node beyond a threshold, forcing the transaction originator to use a different key. We also implement a DeltaMerkleTree, which allows us to efficiently create an updated version of the SMT using memory proportional only to the touched keys.

Our gossip implementation does simple broadcast for regular messages, and runs a stateful protocol for `tx_pool` gossip. We segregate these messages into different ports/queues so the bursty gossip messages are isolated from small messages (e.g., BBA votes) that are broadcast. To prevent malicious *Citizens* from flooding an honest *Politician* with the responsibility of gossiping their writes, we limit the set of *Politicians* for a *Citizen* to be deterministic based on its VRF. *Politicians* do not gossip messages from non-conforming *Citizens*.

9 Evaluation

We evaluate our *Blockene* prototype under several dimensions. The main questions we answer in our evaluation are:

- What throughput and latency does *Blockene* provide?
- How well does *Blockene* handle malicious behaviors?
- Are the optimizations on Merkle tree & gossip useful?
- What is the load on *Citizen* nodes (battery/data usage)?

9.1 Experimental setup

In our experiments, we use a setup with 2000 *Citizen* nodes and 200 *Politician* nodes. *Citizen* nodes are 1-core VMs on Azure with a Xeon E5-2673, 2GB of RAM, and are spread across three geographic regions across WAN: 700 VMs in SouthCentralUS, 600 VMs in WestUS, and 700 VMs in EastUS. Each *Citizen* runs an Android 7.1 image, and is rate-limited to 1MB/s network upload and download. *Politician* nodes run on 8-core Azure VMs with a Xeon E5-2673, 32 GB of RAM, and are spread as 100 VMs each in EastUS and WestUS. They are rate limited to 40MB/s network bandwidth. Given the random safe sampling, the *Citizen-Politician* communication spans across WAN regions. Similarly, the gossip between *Politicians* happens across WAN regions. As our committee size is 2000, every *Citizen* is in the committee for every block. With a higher number of *Citizens*, say 1 million, a particular *Citizen* will be in the committee only once every 500 blocks. Except the per-*Citizen* load, the system performance is independent of the total number of *Citizens* and is just a function of committee size, so the numbers are representative of a large setup.

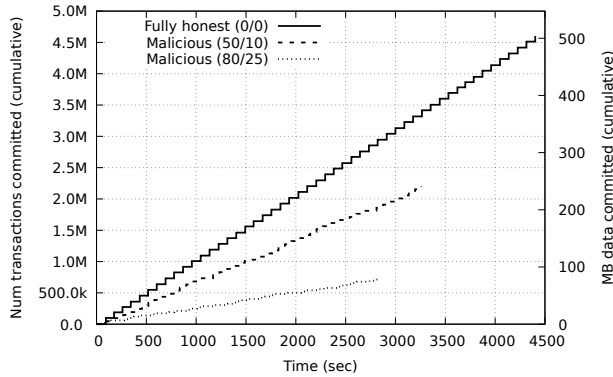


Figure 2: Throughput of *Blockene* under various configs. In 50/10, 50% *Politicians* & 10% *Citizens* are malicious.

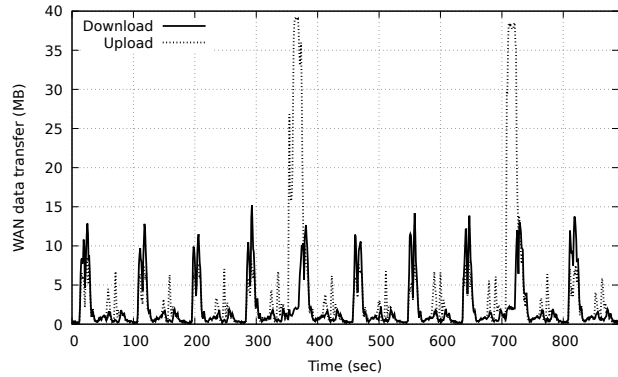


Figure 4: Network usage at a *Politician* node.

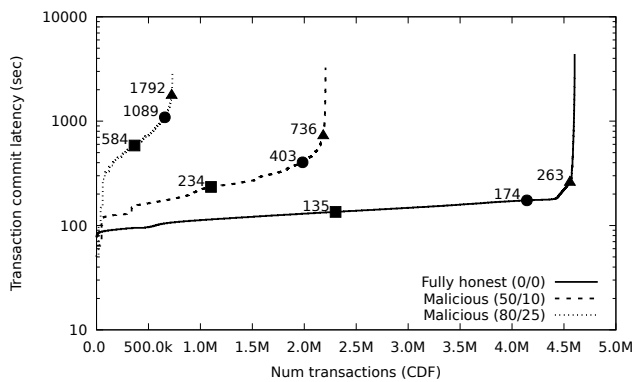


Figure 3: Transaction Latency under different malicious configs. Dots show 50th, 90th, 99th percentiles.

9.2 Transaction Throughput and Latency

Figure 2 shows the timeline of block commits in *Blockene* under fully honest and malicious configurations, for 50 consecutive blocks. In the fully honest (0/0) case, 4.6 million transactions get committed in 4403 seconds, corresponding to a throughput of 1045 transactions per second, or 114 KB/s.

Citizen dishonesty	Politician dishonesty		
	0%	50%	80%
0%	1045	757	390
10%	969	675	339
25%	813	553	257

Table 2: Transaction throughput under malicious configs.

We also evaluate *Blockene* under malicious behaviors of both *Citizens* and *Politicians*. We denote our malicious configurations in the format P/C, where P is the fraction of malicious *Politicians*, and C is the fraction of malicious *Citizens*. With our choice of parameters (e.g., committee size), *Blockene* is guaranteed to ensure safety in the presence of up to 80% malicious *Politicians* and 25% malicious *Citizens*. However,

performance can be affected because of adversarial behavior. A malicious *Citizen* in these experiments attacks in two ways (a) force an empty block by colluding with malicious *Politicians* and proposes a block with *tx_pools* that only malicious *Politicians* have. Honest *Citizens* therefore cannot download that commitment and will vote for an empty block; (b) forces additional rounds in the BBA consensus protocol by manipulating its votes. A malicious *Politician* attacks in two ways: (a) fails to give out transaction commitments, making a subset of the 45 *tx_pools* empty, potentially causing a smaller block to be committed (b) manipulates gossip by acting as sink holes and asking for same chunks from multiple peers. As Figure 2 shows, *Blockene* is quite robust to a range of malicious behaviors, and gracefully degrades in performance. With 80% dishonest *Politicians*, the effective *tx_pools* reduce to 9 out of 45, resulting in the block having only 18K transactions instead of 90K. Malicious *Citizens* cause a performance hit (empty blocks + BBA rounds) when they get chosen as the proposer (i.e., highest VRF); Table 2 shows the throughput under more configurations of malicious behaviors.

Figure 3 shows the CDF of transaction latencies of the system under different configurations, demonstrating fairness across transactions. In the fully honest case (0/0), *Blockene* ensures a median latency of 135s and a 99th%-ile latency of 263s. Under the two malicious configurations: 50/10 and 80/20, latencies are higher as expected.

9.3 Timeline of *Citizens* and *Politicians*

Figure 4 shows the network load at a typical *Politician* node during 10 blocks (each of the repetitive patterns is a block). The two large spikes in uploaded data correspond to rounds where this *Politician* was one of the 45 chosen to provide *tx_pools*. For each block, there are two small spikes of transmitted data; the first spike corresponds to gossip of *tx_pools* through prioritized gossip, and the second spike is due to gossip of votes from *Citizens* in the BBA consensus.

We also show the breakup of the 89-sec block latency by

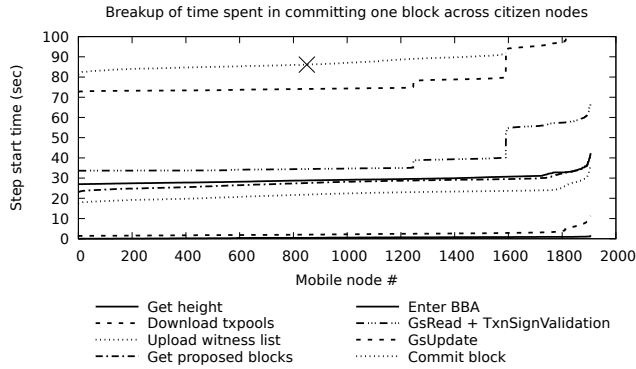


Figure 5: Breakup of time spent at *Citizen* nodes for a single block commit. Cross indicates block commit.

plotting the time taken in *Citizen* nodes during a typical block. Figure 5 shows the progress of the 2000 *Citizen* nodes during one of the blocks, separating out the key phases of the protocol; the bulk of the time goes in the transaction validation phase, and in fetching *tx_pools* from *Politicians*.

9.4 Impact of Optimizations

We now evaluate the prioritized gossip and the sampling-based Merkle tree optimizations. For gossip, we consider how much upload/download each *Politician* incurs before all other honest *Politicians* get all the *tx_pools*. For example, in the 0/0 case, we have 10K data points (across 50 blocks and 200 *Politicians* each). Across these samples, we plot the 50th, 90th, and 99th percentiles. The malicious strategy we model in the 80/25 case is where only the bare minimum number of honest *Citizens* have *tx_pools* of malicious *Politicians* (Δ from § 5.5.2) and all malicious *Politicians* ask for the full set of *tx_pools* from all honest nodes. As Table 3 shows, the network load of prioritized gossip is robust to dishonest behavior. Even in the malicious setting, the data transmitted is quite small before all honest *Politicians* get all *tx_pools*.

Table 4 compares the performance of our sampling based Merkle-tree reads and updates, with the simple solution of downloading challenge paths for all keys referenced in the block. The simple solution incurs much higher network cost (the numbers are after gRPC compression), and a significant compute cost at the *Citizen*. With our optimization, the network cost drops by $10.8\times$ while the CPU cost drops by nearly $31\times$, thus significantly improving transaction throughput.

9.5 Load on Citizens

Finally, we evaluate the load at *Citizen* nodes due to running *Blockene*. The two metrics of interest are battery usage and data usage. To get these metrics, we run an actual Android phone (a OnePlus 5) with the *Citizen* app, as part of the committee along with the 2000 committee members on VMs,

Config	Percentile	Upload (MB)	Download (MB)	Time (sec)
0/0	50	23.1	22.4	3.6
0/0	90	30.5	27.5	4.8
0/0	99	36.7	30.1	5.2
80/25	50	35.4	23.8	3.5
80/25	90	47.6	27.6	4.1
80/25	99	53.4	28.9	4.5

Table 3: Cost of gossip per honest *Politician* before all honest *Politicians* receive all *tx_pools*.

Config	Upload (MB)	Download (MB)	Compute (s)
Naive: GS Read	0	56.16	93.5
Naive: GS Update	0	0	93.5
Optimized: GS Read	0.55	1.6	1.0
Optimized: GS Update	0.01	3	5.88

Table 4: Performance of Global State Read & Write.

and measure battery use. After being in the committee for 5 blocks, the battery drain was $\sim 3\%$. The total network traffic incurred by a *Citizen* for a single block was 19.5 MB.

Now, we can extrapolate the daily cost based on the per-block cost and the number of times a single *Citizen* is expected to be in the committee. With 1 million *Citizens*, a *Citizen* will participate roughly every 500 blocks, which at our block latency of ~ 90 s, translates to 2 times per day. Thus, the expected battery use is $< 2\%$ per day, and the data use is ~ 40 MB/day. In addition, we also measured on the same OnePlus5 that waking up the phone every 10 minutes and performing `getLedger` costs about 0.9% battery and 21MB data download. Waking up every 5 minutes costs 1.7% battery and 42MB data download. With a total of 3% battery usage and 61MB data/day, a user running the *Blockene* app will hardly notice it running.

10 Conclusion

By enabling, for the first time, a high-throughput blockchain where members perform block validation and consensus on smartphones at negligible resource usage, *Blockene* opens up a much larger class of real-world applications to benefit from the security and decentralized nature of blockchains. With a novel architecture, and several new techniques coupled with a careful security reasoning, *Blockene* is able to simultaneously provide three conflicting properties: large scale of participation, high transaction throughput, and low resource usage at member nodes.

Acknowledgements

We thank our shepherd Nikolai Zeldovich and the anonymous reviewers for their valuable suggestions and feedback. We also thank Ankush Jain, Sriram Rajamani, Bill Thies, Jacki O’Neill, and Rashmi.K.Y for their support.

References

- [1] Aadhaar identity ecosystem. In <https://uidai.gov.in/aadhaar-eco-system/authentication-ecosystem.html>.
- [2] AWS Shield: Managed DDoS Protection. In <https://aws.amazon.com/shield/>.
- [3] Azure DDoS Protection. In <https://azure.microsoft.com/en-in/services/ddos-protection/>.
- [4] Carrier-grade NAT: Wikipedia. In https://en.wikipedia.org/wiki/Carrier-grade_NAT.
- [5] Ethereum blockchain. In <https://www.ethereum.org/>.
- [6] Apple Platform Security: Secure Enclaves Overview. In <https://support.apple.com/en-in/guide/security/sec59b0b31ff/web>, 2017.
- [7] Giving in Europe country reports available. In <https://ernop.eu/giving-in-europe-launched-at-spring-of-philanthropy/>, 2017.
- [8] India Philanthropy Report 2017. In <https://www.bain.com/insights/india-philanthropy-report-2017/>, 2017.
- [9] Android Docs: JobScheduler. In <https://developer.android.com/reference/android/app/job/JobScheduler>, 2018.
- [10] Charitable Giving Statistics. Americans gave \$410 billion to charities in 2017. In <https://nonprofitssource.com/online-giving-statistics>, 2018.
- [11] Android Keystore System: Hardware Security module. In <https://developer.android.com/training/articles/keystore#HardwareSecurityModule>, 2019.
- [12] Ronald Abraham, Elizabeth S Bennett, Noopur Sen, and Neil Buddy Shah. State of aadhaar report 2016-17. *IDinsight*. Available at: <http://stateofaadhaar.in>, 2017.
- [13] Elli Androulaki, Artem Barger, Vita Bortnikov, Christian Cachin, Konstantinos Christidis, Angelo De Caro, David Enyeart, Christopher Ferris, Gennady Laventman, Yacov Manevich, et al. Hyperledger fabric: a distributed operating system for permissioned blockchains. In *Proceedings of the Thirteenth EuroSys Conference*, page 30. ACM, 2018.
- [14] Christian Badertscher, Peter Gazi, Aggelos Kiayias, Alexander Russell, and Vassilis Zikas. Ouroboros Genesis: Composable Proof-of-Stake Blockchains with Dynamic Availability. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, CCS 2018, Toronto, ON, Canada, October 15-19, 2018*, pages 913–930, 2018.
- [15] Miguel Castro, Barbara Liskov, et al. Practical byzantine fault tolerance. In *OSDI*, volume 99, pages 173–186, 1999.
- [16] Dominik Charousset, Thomas C. Schmidt, Raphael Hiesgen, and Matthias Wählisch. Native Actors – A Scalable Software Platform for Distributed, Heterogeneous Environments. In *Proc. of the 4rd ACM SIGPLAN Conference on Systems, Programming, and Applications (SPLASH '13), Workshop AGERE!*, pages 87–96. ACM, Oct. 2013.
- [17] John R Douceur. The sybil attack. In *International workshop on peer-to-peer systems*, pages 251–260. Springer, 2002.
- [18] Forbes. The cost of mobile internet around the world. In https://blogs-images.forbes.com/niallmccarthy/files/2019/03/20190305_Data_Cost.jpg, 2019.
- [19] IOTA Foundation. Differences between the tangle and blockchain. In <https://docs.iota.org/docs/getting-started/1.1/the-tangle/tangle-vs-blockchain>.
- [20] IOTA Foundation. Tangle: The coordinator. In <https://docs.iota.org/docs/getting-started/1.1/the-tangle/the-coordinator>.
- [21] Yossi Gilad, Rotem Hemo, Silvio Micali, Georgios Vlachos, and Nickolai Zeldovich. Algorand: Scaling byzantine agreements for cryptocurrencies. In *Proceedings of the 26th Symposium on Operating Systems Principles, SOSP '17*, pages 51–68, New York, NY, USA, 2017. ACM.
- [22] Aggelos Kiayias, Alexander Russell, Bernardo David, and Roman Oliynykov. Ouroboros: A Provably Secure Proof-of-Stake Blockchain Protocol. In *Advances in Cryptology - CRYPTO 2017 - 37th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 20-24, 2017, Proceedings, Part I*, pages 357–388, 2017.
- [23] S King and S Nadal. Peercoin—secure & sustainable cryptocoin. <https://www.peercoin.net/whitepapers/peercoin-paper.pdf>, 2012.

- [24] Eleftherios Kokoris-Kogias, Philipp Jovanovic, Nicolas Gailly, Ismail Khoffi, Linus Gasser, and Bryan Ford. Enhancing Bitcoin Security and Performance with Strong Consistency via Collective Signing. In *25th USENIX Security Symposium, USENIX Security 16, Austin, TX, USA, August 10-12, 2016.*, pages 279–296, 2016.
- [25] Eleftherios Kokoris-Kogias, Philipp Jovanovic, Linus Gasser, Nicolas Gailly, Ewa Syta, and Bryan Ford. OmniLedger: A Secure, Scale-Out, Decentralized Ledger via Sharding. In *2018 IEEE Symposium on Security and Privacy, SP 2018, Proceedings, 21-23 May 2018, San Francisco, California, USA*, pages 583–598, 2018.
- [26] Silvio Micali. Byzantine agreement, made trivial. In <https://people.csail.mit.edu/silvio/Selected%20Scientific%20Papers/Distributed%20Computation/BYZANTYNE%20AGREEMENT%20MADE%20TRIVIAL.pdf>, 2018.
- [27] Silvio Micali, Michael O. Rabin, and Salil P. Vadhan. Verifiable random functions. In *40th Annual Symposium on Foundations of Computer Science, FOCS '99, 17-18 October, 1999, New York, NY, USA*, pages 120–130, 1999.
- [28] Andrew Miller, Yu Xia, Kyle Croman, Elaine Shi, and Dawn Song. The Honey Badger of BFT Protocols. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, Vienna, Austria, October 24-28, 2016*, pages 31–42, 2016.
- [29] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, New York, NY, USA, 2005.
- [30] Satoshi Nakamoto. A peer-to-peer electronic cash system. In <https://bitcoin.org/bitcoin.pdf>, 2008.
- [31] Rafael Pass and Elaine Shi. Hybrid Consensus: Efficient Consensus in the Permissionless Model. In *31st International Symposium on Distributed Computing, DISC 2017, October 16-20, 2017, Vienna, Austria*, pages 39:1–39:16, 2017.
- [32] Gary Rong and Felix Lange. Light ethereum sub-protocol (les). In <https://github.com/ethereum/devp2p/blob/master/caps/les.md>, 2019.
- [33] Mark Russinovich, Edward Ashton, Christine Avanesians, Miguel Castro, Amaury Chamayou, Sylvan Clebsch, Manuel Costa, Cédric Fournet, Matthew Kerner, Sid Krishna, et al. Ccf: A framework for building confidential verifiable replicated services. Technical report, Technical Report MSR-TR-2019-16, Microsoft, 2019.
- [34] Sambhav Satija, Apurv Mehra, Sudheesh Singanamalla, Karan Grover, Muthian Sivathanu, Nishanth Chandran, Divya Gupta, and Satya Lokam. Blockene: A high-throughput blockchain over mobile devices. *arXiv preprint arXiv:2010.07277*, 2020.
- [35] Kongrath Suankaewmanee, Dinh Thai Hoang, Dusit Niyato, Suttinee Sawadsitang, Ping Wang, and Zhu Han. Performance analysis and application of mobile blockchain. In *2018 international conference on computing, networking and communications (ICNC)*, pages 642–646. IEEE, 2018.
- [36] Russell Turpin and Brian A. Coan. Extending binary byzantine agreement to multivalued byzantine agreement. *Inf. Process. Lett.*, 18(2):73–76, 1984.
- [37] Mahdi Zamani, Mahnush Movahedi, and Mariana Raykova. RapidChain: Scaling blockchain via full sharding. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, CCS 2018, Toronto, ON, Canada, October 15-19, 2018*, pages 931–948, 2018.