# Chapter 3
# Industrial Applications of Blockchain to IoT Data


**Steven Pu**

## 1 IoT and Blockchain

### *1.1 Challenges Facing the IoT Space*

Blockchain has been long touted as the perfect technological complement to IoT systems. To understand why there has been such enthusiasm for the synergies between these two seemingly unrelated technology systems, we first examine some of the largest challenges facing the IoT space, divided into several broad categories: technological, commercial, and social.

#### 1.1.1 Technical Challenges

Contemporary IoT systems increasingly exist and interface in a sea of connected devices that are not only potentially adversarial, but also often operate on heterogenous infrastructure and standards. This coupled with the fact that IoT devices are being deployed at an accelerated rate (Columbus 2018) makes these hitherto rather obscure technological concerns increasingly relevant to our daily lives. Here we examine several key technical challenges to IoT systems.

From a network perspective, IoT devices predominantly exist in networks that have a hub-and-spoke topology, or a server–client paradigm. Each connected device can be considered as an endpoint that constantly needs to communicate with a central

This article is based on our final report to the study group "Blockchain and Society 5.0—The Creation of a New Marketplace based on Distributed Consensus" at the Research Institute of Economy, Trade, and Industry (RIETI).

S. Pu (✉)
Taraxa, San Jose, CA, USA
e-mail: steven.pu@taraxa.io



M. Yano et al. (eds.), *Blockchain and Crypto Currency*,
Economics, Law, and Institutions in Asia Pacific,
https://doi.org/10.1007/978-981-15-3376-1_3

server to upload data, communicate with other devices, and receive commands. In most networks, even when the IoT devices are just a few feet apart, they cannot communicate with each other directly and must rely upon this centralized server to broker such communication. This centralized server, while it may be a distributed network of computers, is still a centrally administered entity and therefore presents a single point of failure. This means that to compromise (to render inoperable, or to take outright control over) a large network of IoT devices, all the attacker needs to do is to compromise or take control of the central server these devices are reliant upon for everything from sending and receiving commands to data uploads. This presents not just a significant security risk but also an administrative nightmare to those who operate such central IoT management services.

In addition to presenting a single point of failure, centrally managed IoT networks also place the entire upfront investment, ongoing management costs, storage and computation **workload** involved with the management and maintenance on a single entity. As IoT networks become more ubiquitous, interconnected, and scale from hundreds of millions to trillions of devices, this type of centralized workload becomes rapidly untenable. This especially becomes a problem for device **maintenance** as technology advances forward and each centralized network management system needs to keep ever-increasing versions of software and firmware (many of which have become obsolete) and be able to make them available on demand to ensure the longevity of IoT devices that have been deployed in the field.

At the endpoints (often sensors) within the network, most IoT devices still rely upon plaintext passwords and worse, manufacturers' default or commonly reused passwords to establish **identity** and privileges on the network across devices, making them vulnerable to attacks by malware such as Mirai (Graff 2017). Such poor security practices are not only driven by a general lack of security awareness and understanding but also by the complexity that comes with managing such a large and disparate set of connected devices in a central system. These passwords further limit the security of these devices' communications because there is no way beyond communicating with the central server to validate the identity, origin, and, by extension, veracity of the messages (or collected data) as is commonly guaranteed by modern crypto graphical methods.

Without cryptographically guaranteed identities, signatures, and identity-based encryption, data collected by and sent from most IoT devices today cannot establish **provenance** and therefore cannot be trusted unless the data (and any fallout from bad data) is guaranteed by a trusted third party, which greatly increases the communication and more importantly, transactional friction between devices. This presents a further security risk that the unencrypted or poorly encrypted data could have been intercepted or worse, tampered with while in transmission, which further erodes the trust other entities (e.g., other people, companies, devices) have for the resultant data and could potentially damage the reputation of the IoT network's owner.

Looking at IoT as a sector, IoT networks are invariably made up of extremely long value chains comprising many disparate components and players. Using dataflow as a connecting dimension, there are sensors that collect the data at the endpoints, gateways that manage the sensors and aggregate as well as upload the data, storage

systems (e.g., cloud) that store and make the data available, and analytics engines that digest and generate actionable insights from the data. Within each step and between these steps, all the hardware and software involved must agree to a set of common **standards** by which to communicate, and those standards are just as disparate as the innumerable number of players in the IoT space. This results in the entire IoT industry being severely siloed, with completely disparate IoT systems that do not and technically cannot communicate, much less transact, with one another. The difficulty in facilitating communications between these siloed and heterogenous networks is one of the biggest technical challenges in IoT today and is holding back the massive network effect potential of the IoT space.

### 1.1.2 Business Challenges

Despite the many rosy predictions for the future of IoT (Columbus 2018), most businesses still have serious reservations when it comes to making serious investments into IoT and IoT-related systems. Besides the numerous technical challenges, there are serious business challenges such as the generally unclear (or outright lack of) business case, data sensitivity, and the potential strategic risk of sharing data.

Return on investment inevitably drives business decisions, and investment into IoT is no different. One of the biggest challenges for IoT is the lack of a viable **business case** that justifies its investments, either by generating revenue or shaving costs. Business cases are difficult to come by because it is extremely hard to figure out how to analyze and generate value from the data collected by IoT devices.

To fully capture the value of data often requires specialized expertise, an expertise that businesses that generate data generally lack. This lack of internal expertise requires businesses to seek outside help, which often raises concerns for **data sensitivity**, driving businesses to be very careful and highly selective about which partners and vendors they collaborate with to analyze the data. This cautious approach no doubt severely circumscribes the extent to which any business has access to the best possible talent to analyze and generate value from their data sets and greatly reduces the possibility of finding a viable business case. This problem is further exacerbated given that many breakthrough value-generating insights come from data that is aggregated from many businesses and often across industry verticals, but with each business closely guarding their data nest eggs such insights become nearly impossible to discover.

Even when businesses are comfortable sharing data with a specific vendor, there still exists the potentially fatal **strategic risk** that the vendor (usually a technology platform) will overtake the business with superior aggregation of and insights generated from data. As data is increasingly seen as a critical driver of performance, efficiency, and profitability, it has also become a strategic resource. Large technology platforms (e.g., Google, Amazon, Facebook) gain long-term sustainable competitive advantages through effective aggregation and analytics of data and have established de facto monopolistic power. Not only are such platforms able to dominate the technology markets they were born out of, but with their proprietary technology and

massive data aggregation and analytics, they have proven consistently capable of disrupting a variety of markets that are not even adjacent to their original core businesses (e.g., Google with automotive, Apple with gaming, Amazon with cloud) . Hence, by aggregating and effectively analyzing data, the "vendor" can then turn back on the "client" and invade its markets.

### 1.1.3 Social Challenges

With the rapid proliferation of digitized technologies, the public at large has become increasingly aware of the omnipresence of data-collecting sensors as well as concerned about how they are being used. Recent scandals involving Facebook (Granville 2018) and Google's (MacMillan and McMillan 2018) mishandling of user data sparked worldwide concerns amongst the public as well as regulators. The EU's General Data Protection Regulation (GDPR) (EU GDPR.ORG 2018) that came into effect in May of 2018 further placed privacy and data ownership at the center of civil discourse. These regulatory trends, however, are still extremely limited in scope in that they mostly require user consent upon visiting websites that only *acknowledges* the problem without fundamentally solving it. These concerns are especially thorny in the case of IoT devices, because they have increasingly become embedded directly into our environments without our knowledge, tracking everything from location and movement to voice and video. Much of this also happens with numerous third parties whose involvement and activities are difficult to track, as well as across political jurisdictions each with their uniquely different regulatory requirements, further complicating social concerns. If IoT technology is to continue to proliferate, it must address **data privacy** concerns head-on and provide socially acceptable solutions to guarantee secure data ownership and usage without triggering innovation-killing regulatory backlashes.

## *1.2 Blockchain Empowers IoT Devices*

Although the first and most widely known application of blockchains is Bitcoin, its underlying technologies provide a unique suite of functionalities that make it uniquely complementary to IoT by empowering them to become independent entities within a decentralized network. In doing so, this development directly or indirectly addresses many of the challenges currently facing IoT technologies.

### 1.2.1 Blockchain Grants Devices Independence

IoT devices in today's networks do not exist as independent entities outside of their centrally managed networks. As far as the outside world is concerned, they are dealing with a large server sitting in the cloud that has some data, without any idea

of the provenance of the data or any means to interact directly with the devices that collected the data in the first place. On a blockchain network, each node—any participant connected to the network—has a unique private and public key pair that uniquely identifies it as an independent participant on the network. Specifically, these identities are enforced largely using cryptographic signatures, or digital messages that unmistakably (and next to impossible to forge) identify the sender.

Having unique identities is the foundation for achieving independence, giving each device the ability to act on its behalf. This enables a decentralized mesh network topology rather than a centralized server–client network topology, with each node able to make its own decisions, and, more importantly, to make use of its own resources independently of the other nodes. This type of network is much more secure, because hackers can no longer gain control over millions of devices by hacking a single server (a single point of failure). Rather, the hacker has to compromise millions of devices one by one, with each compromised device likely to be rejected by the network for misbehavior, resulting in the hacker taking over a useless, disconnected device.

A decentralized network with a smart consensus algorithm is also much better at balancing workloads that were formerly handled by a single entity. This makes network deployment as well as maintenance far less costly because the workload of connectivity, storage, and even computation can now be done by many devices in the network, without the need for a costly centralized arbiter.

### 1.2.2 Blockchain Grants Devices Awareness for Ownership

Blockchain also endows devices with the concept of ownership through the very same cryptographic primitives that guaranteed unique identities. Any device can now sign for as well as encrypt any form of digital asset it has access to. Specifically, a device can now own cryptocurrencies (like Bitcoin) as well as other forms of assets that it has control over (e.g., data, bandwidth, storage). By having this concept of ownership, the IoT device is now an independent economic entity able to not only act, but act in its own best economic interests. For example, instead of remaining idle, a device might decide to put its capabilities on auction and collect customized data on-demand; to avoid obsolescence, it could network with other similar devices to contract order a firmware upgrade, etc. While they like science fiction, these examples may not be too far off in the future.

Guaranteeing ownership of digitized assets also guarantees the privacy of the asset generator. Without the explicit permission of the originator, e.g., without a decryption key, no one can access the data. Today's rampant, and, more importantly, hidden data collection and aggregation processes will be brought to the forefront and forced to seek explicit permission from the data generator and owner.

### 1.2.3 Blockchain Enables Devices to Trade

What does an independent, asset-owning economic entity do? It trades with other independent, asset-owning entities. At the core of every blockchain network is a consensus algorithm that makes sure every node on the network agrees on the network's historical set of state transitions, or, more simply, what has changed about the network. This consensus enables the defining functionality of blockchain—decentralized trading of digitized assets.

The ability to securely trade assets and resources becomes even more consequential when you consider the global ecosystem of open-source developers that are naturally part of any open-source blockchain ecosystem. Now there is a way to reward and enable better usage of the data collected by devices in a decentralized manner. Any device or a network of devices can choose to publish a segment of its collected data and put up a bounty with a specific objective (e.g., lower energy consumption, faster processing throughput) on the blockchain marketplace, locking the reward in a cryptographically guaranteed smart contract, and incentivize people (and intelligent algorithms) to discover and be rewarded for the solution. Discovering uses (business models) for data was an extremely difficult problem for a centralized entity, but with blockchain, it could potentially become a much simpler decentralized problem, tapping into a globalized talent pool from all over the world.

## *1.3 Current Limitations to the IoT + Blockchain Vision*

With blockchain technology, IoT devices are empowered to make independent decisions, work together to distribute workload and maintenance, and freely trade assets and resources with localized decision-making. Continuation down this path will see the development of an intelligent, self-evolving, self-governing network that we have seen described only in science fiction.

Although the advent of blockchain technology means that we are many steps closer to this futuristic vision, we are not quite there yet. Some of the key limitations of the system are listed below.

- There lacks a mainstream, demonstrable low-latency, high-throughput blockchain network designed specifically for IoT devices.
- Device manufacturers have yet to embed cryptographic keys into every piece of hardware or make them blockchain-compatible as a generalized standard.
- Software cryptographic methods to guaranteeing privacy-preserving computations are grossly inefficient and not practical (IBM Research Editorial Staff 2018), while hardware solutions require trust in the manufacturer and the entire manufacturing supply chain, making it difficult to protect against data piracy.
- Artificial intelligence is not sufficiently sophisticated to enable such extraordinarily autonomous decision-making behavior in devices.

- Legal recourse is still required to further de-risk trading over blockchain, but only limited jurisdictions (De 2018) have recognized smart contracts on blockchain as legally binding contracts off-chain.

In time, however, we are optimistic that all the above-mentioned limitations will be overcome.

Even with these limitations, blockchain is still well positioned to resolve many of the technological, business, and social challenges faced by IoT with wide-ranging potential for value-adding applications. We now dive deeper into the current state of blockchain technology to see what else can be done to improve upon the state of the art.

## 2 A Blockchain Network Created for IoT Devices

Given all the synergies between blockchain and IoT, what are the characteristics of a blockchain network that would be well suited for IoT needs? Although much blockchain technology is infrastructural in nature and is not obviously application specific, there are many design and optimization choices in the public ledger level that should reflect what application stacks the designers were thinking about during the development process.

### *2.1 Characteristics IoT Devices and Implications on the Design of Blockchain Networks*

When thinking about IoT, specifically in contrast with the nodes that operate on existing blockchain networks, it is useful to know that all blockchain networks today rely on the services of powerful and constantly connected servers to perform all the record-keeping and consensus duties. What is immediately apparent is that most of what we think of as "IoT" devices, or smaller, sometimes mobile, connected devices, have limited and unique characteristics that do not fit this profile.

While the term "IoT" is used to refer basically to any connected device, we could make several general statements about the characteristics of these devices.

- **Massive scale**: by some estimates (Tung 2017) the number of IoT devices has already surpassed the human population in the world, and will continue to grow at an accelerated rate.
- **Limited computing power**: IoT devices are usually not processing powerhouses often by orders of magnitude even compared to the processing power in regular laptop computers (TrueBench 2018).
- **Limited storage**: most IoT devices are not meant to store information locally and are simply meant to relay information (e.g., to a cloud), hence have very limited storage.

- **Limited bandwidth and connectivity**: many IoT devices operate out in the field without reliable connections and costly connectivity (e.g., satellite network in the middle of the woods).
- **Limited power consumption**: many IoT devices operate on batteries or via energy-harvesting mechanisms that place severe constraints on its energy consumption.

The design challenge can then be formulated thus: what are the critical metrics required to design a blockchain network that can best serve IoT devices?

1. **Network needs to be scalable**: given there could be potentially billions of devices connected to any given blockchain network, the network must be able to scale its capacity in processing transactions and requests.
2. **Network needs to support discovery and trading of generic digital assets**: IoT devices have many digital assets and resources (e.g., data) to trade, not simply currency, and they need means of discovering these assets.
3. **Network needs to support selective memory**: given all the limitations of IoT devices, they will only be able to participate in a small subset of the network and must be selective in what each device stores and processes.
4. **Network cannot solely depend on "work" to maintain security**: network security cannot be purely based on solving complex cryptographic puzzles, making blockchain transactions impractical for IoT devices.
5. **Network needs to support trustless light nodes**: IoT devices today cannot support full node operations but still need to maintain their independence on a blockchain network. The "light" nodes run on IoT devices therefore cannot be naïve (i.e., blindly trusting another full node) and must have some means of validating network state and state transitions.
6. **Network needs to support point-to-point transactions**: many transactions between IoT devices are highly localized—the devices are right next to each other—and cannot be expected to wait for the latency of network-wide validation every time.

With these design goals in mind, the Taraxa project was started to help IoT devices democratize their data and maximize the value generated by that data.

## *2.2 An Evolving Landscape*

When Taraxa was first being conceived in 2017, significant research was conducted into the existing slate of blockchain networks as well as relevant technologies to understand not just the current landscape but also how the space has evolved over time. While there are many amazing projects doing important work and making major contributions to the blockchain space, we acknowledge a few projects that not only inspired but made possible in many ways our work here at Taraxa.

### 2.2.1 Bitcoin

As of this writing, it has been exactly 10 years (Investopedia 2018) since the first publication of Satoshi Nakamoto's whitepaper (Nakamoto 2008) and the beginning of the blockchain revolution. While all the technologies underlying Bitcoin were not new and in fact similar incarnations had been proposed and even implemented before (Narayanan and Clark 2017), Bitcoin was unique in that its designs not only incorporated these technologies in an innovative way, but also built in the ideas of decentralization, trustless transactions, and a sophisticated understanding of human incentives.

Probably most consequentially, Bitcoin's arrival coincided with a global crisis of trust as the world was descending into one of the worst financial crises in recorded history (Bernanke 2018). Ordinary citizens worldwide were questioning not only the seemingly absolute authority that centralized entities such as the global banking system and large multinational corporations have over everyone's daily lives, but also the implicit trust that is placed in these institutions. Bitcoin is unique for being the very first representation of *value outside the system* (a term coined by investor and blockchain entrepreneur Jianbo Wang) of existing institutions' underwriting, approval, or participation.

Bitcoin is the technological and philosophical inspiration for the entire blockchain space.

### 2.2.2 Ethereum

By expanding beyond (or rather completely rewriting) Bitcoin's simple scripting language into a Turing complete application layer called smart contracts, Ethereum (Ethereum White Paper 2018) has enabled potentially an infinite number of applications to take advantage of blockchain's unique properties beyond simply currency.

Ethereum made possible many decentralized applications, including games, marketplaces, and even decentralized corporations. The explosion of applications drew interest and participation from far beyond just the financial sector, but also from many mainstream academic, industrial, and public institutions. Along with Ethereum also arose the initial coin offerings, a fundraising model that offers the first viable alternative to the existing and highly centralized global investment apparatus, giving many nascent decentralized projects a chance to grow.

Ethereum is what sparked our imagination that blockchain could be much more than just a currency.

### 2.2.3 IOTA

IOTA (Popov 2018) was the first widely known project (many lesser-known projects proposed similar technologies during roughly the same time period) to propose an

alternative data structure (a directed acyclic graph, or DAG) as opposed to the typical blockchain pioneered by Bitcoin. It was also the first project to educate the wider market of the synergies between IoT and blockchain. Although at times controversial (Narula 2017), IOTA nevertheless has made and continues to make important contributions to the blockchain space.

#### 2.2.4 ByteBall

ByteBall (Churyumov 2016) was the first widely known project to propose total ordering within a DAG blockchain network by identifying a main chain as a set of anchors. Via this main chain, every node would run a deterministic algorithm that eventually converges onto the same total-network ordering with minimal communication overhead. This mainchain resolves the convergent ordering issue for DAG networks while making use of every vertex (in the case of Byteball, they are transactions) on the DAG.

#### 2.2.5 Phantom

Proposed by authors of the influential papers Ghost (Sompolinsky and Zohar 2013) and Spectre (Sompolinsky et al. 2016), Phantom (Sompolinsky and Zohar 2018) is a blockchain that proposes the blockDAG, a way to organize sets of transactions like those in Bitcoin and Ethereum blocks into a DAG topology, and then converges upon a single chain via a deterministic algorithm that each node executes individually. The blockDAG combined many of the concurrent properties of a DAG while also maintaining the idea of a transaction set, enabling many of Taraxa's innovations in concurrency.

### *2.3 Taraxa's Innovations*

Every blockchain infrastructure project should seek to introduce technical innovations to the blockchain space and contribute to the cumulative pool of open-source knowledge, and Taraxa is no different. Building on the existing body of knowledge and technologies, we set out to make the following key contributions as roughly summarized below.

#### 2.3.1 Concurrent Smart Contracts

As it stands today, smart contracts are processed in sequential order by nodes on blockchain networks. Taraxa implements a way to process them concurrently (i.e., in parallel) to increase the processing throughput of smart contracts.

There are several obstacles to running smart contracts in parallel. First, because smart contracts modify shared storage (their persistent storage), it is crucial to keep track of which processes are accessing which areas of storage at any given moment to avoid conflicting access. Second, because the programming language is Turing complete, it is impossible to determine statically whether different contract calls will conflict during parallel execution.

We propose that the Taraxa nodes execute smart contract code as speculative actions. A node schedules multiple smart contract calls for parallel execution, and then keeps track of their access to persistent storage via the Taraxa runtime APIs. Should there be conflicting access (i.e., read/write, write/write), the access is rejected, the conflict is reported to the scheduler, with the scheduler terminating the process, rolling back its speculative changes to the persistent storage, and reschedules these conflicting contract calls for sequential processing.

We further propose that to minimize the number of conflicts during execution, we endow the virtual machine with partial semantic understanding for the code. In general, a computer simply executes code it is given without the need or capability to understand what it is actually doing; that is, the code has no meaning (semantics) to the machine. However, many types of executions may look like conflicts but are in fact not true conflicts if the computer understands their purpose. For example, many contracts make use of counters to enforce a specific range; hence, the order of operations (i.e., increments, decrements) on this counter is not important, because the result remains the same no matter the order they occur, as long as they do not exceed the range. Hence, what may look like conflicts with multiple calls accessing the same counter is in fact not necessarily a conflict. The virtual machine may be endowed with such semantic understanding through analysis of the byte code and automatically tagging operations that fit a specific pattern; for example, a counter.

In addition to executions, we also propose that the process of committing (writing) state transitions into persistent storage could also be parallelized.

Note that all concurrency gains are obtained without the developers needing to alter their coding behavior or their code. This is especially important because any new technology that involves more work on the part of the developers is less likely to be adopted.

With contracts now processed in parallel, it is important for other nodes to follow the same concurrent schedule, or else every node will select a different set of contract calls with different concurrent schedules and there is no convergent consensus. Hence, a concurrent schedule will be embedded along with the concurrent set to ensure that all nodes execute the concurrent set in the exact same order as agreed upon (via consensus) and arrive at the same resultant state.

### 2.3.2 Fuzzy Sharding

To take full advantage of multiple nodes working together to make progress on the network, Taraxa makes use of a blockDAG topology, pioneered by researchers of the Phantom (Sompolinsky and Zohar 2018) paper. This topology has the advantage

of enabling multiple nodes to work together to propose blocks and help the network make progress, but it then potentially suffers from nodes simultaneously performing redundant work.

Taraxa proposes a set of algorithms that elegantly resolves these issues without coordination. In most other networks, the functionality of which nodes are responsible for which separate tasks require the election of a leader, who has temporary power over a certain set of decisions, such as which node is assigned which work. The election of a leader is expensive in terms of network resources and exposes that specific leader to attacks once its identity is known. Using a set of cryptographic operations (cryptographic sortition), Taraxa allows each node to independently verify proposal eligibility and transaction jurisdiction—in other words, they are assigned non-overlapping tasks, randomly and fairly, without the need for a leader to coordinate them.

### Trustless Light Nodes

While IoT devices lack the resources necessary to host a "full node," that does not mean they cannot retain their independence or contribute to core ledger tasks. Taraxa creates a series of light node designs that can accommodate the full spectrum of IoT devices, from the most resource-starved to those that are less so. Indeed, the term "full node" is one extreme on a spectrum, referring to powerful computers with significant computation, storage, bandwidth resource, and high uptime, while the term "light node" refers to the remaining spectrum of devices that do not fit this profile. Any light node design must accommodate the spectrum, giving each device the choice to participate as much or as little as it is capable. Any network that constrains proper execution of its protocol to only those devices with a high threshold of computing power is inherently creating a centralizing force. In addition, keeping the bar low for device participation means that more powerful nodes present an attack vector, and more generally represent wasted computing effort. Therefore, given the twin desires to both maximize decentralization and performance, the protocol must enable this wide spectrum of devices to participate to their fullest.

In Taraxa's designs, light nodes will be able to be more trustless, in that they are better able to validate the information they receive from the network. In conventional designs, light nodes simply latch onto a specific full node and request an update, while only able to validate the internal consistency of what it has been told (e.g., there is no contradictory information). Taraxa allows a light node to randomly sample a subset of the network's full nodes to compare their responses to become more trustless in its validation process.

In addition to validation, light nodes could be made even more trustless by gaining the ability to propose concurrent sets. Given the blockDAG topology, we could enable an efficient merging of concurrent sets proposed by active nodes on the network, allowing for smaller sets to be proposed and still be useful. While light nodes cannot propose arbitrarily large concurrent sets, a sufficiently well-connected node with reasonable storage could propose concurrent sets pertaining to accounts it has on

store, perhaps including not just its own account states but also those entities the device regularly interacts with. By enabling light nodes to propose concurrent sets, we also move away from reliance upon powerful computers to maintain the blockchain network (as all existing blockchain projects do) and into the edge with the IoT devices themselves.

Lastly, instead of relying upon solving cryptographic puzzles (i.e., proof-of-work; PoW) to deter spamming attacks, Taraxa will rely on a system of fees, because most IoT devices are unable to complete such puzzles in timely fashion and simplifying such puzzles would render them a useless deterrent to more powerful machines. PoW is just another form of fees that requires upfront capital outlay for better hardware, an option unavailable to most IoT devices.

These are some of what we consider to be Taraxa's primary innovative contributions that could help IoT devices to transact more freely and simply with each other and the world at large.

# 3 Potential Applications

## *3.1 Blockchain Application Suitability*

Here we briefly outline a few key characteristics of blockchain technology that help to guide what blockchain should and should not be used for. For the purposes of this chapter, blockchain refers to purely public ledgers.

### 3.1.1 Blockchain Bridges Trust Gaps

Blockchain is a decentralized and distributed network. Being distributed means there are redundant copies of ownership and transactional history so it is difficult to attack the network, while being decentralized means that no participant needs to trust any other participant because agreement is reached through consensus and cryptographic algorithms. Not only is this guaranteed during transactions; blockchain's unique interlocking data structure enables trivial ex post facto auditing, making data tampering immediately obvious.

By bridging trust gaps, blockchain enables formerly impossible or grossly inefficient market-making, and simplification or entire removal of inefficient trust-building apparatus within existing markets.

### 3.1.2 Blockchain Is Digital

Although the digital nature of blockchain may sound obvious, blockchain technologies are only applicable when dealing with digital assets. The first application

of blockchain was Bitcoin, a digital currency (a form of digital asset) encoded as balances and transfers that are wholly self-contained within the blockchain. Other digital assets might represent data created off-chain and then anchored and traded on-chain (e.g., IoT data, digital content), or are digital representations of physical assets (e.g., asset-backed securities).

### 3.1.3 Blockchain's Guarantees Are on-Chain Only

While blockchain provides many security and trustless guarantees on-chain, it is important to note that all such guarantees are *only* on-chain. Should parties make off-chain arrangements that erode the integrity of on-chain transactions (e.g., collusion between Bitcoin miners), there is little that the blockchain can do about it. Robust blockchain design can seek to minimize or thwart bribery in critical consensus and validation steps, but it cannot solve the fundamental problem that potential off-chain value may be greater than the value proposition of honest on-chain behavior. In contrast, when on-chain, blockchain protocols are designed to tolerate a fair proportion of "dishonest" nodes (up to 30% or even up to 50% of the network) without fundamental loss of network integrity.

### 3.1.4 Blockchain Is Inefficient

Blockchain's trustless transactions come at the cost of efficiency. Having to constantly reach consensus and replicate transactions across the network, possibly in the presence of faulty, confused, and dishonest nodes, makes blockchain networks fundamentally inefficient. This means blockchain should be used sparingly when security and trust concerns outweigh the benefits of efficiency.

Above all, blockchain should be leveraged as a way to keep centralized systems honest. Centralized systems have clear performance advantages over decentralized systems such as blockchain and are required for any performance-sensitive applications. There is no reason to want only replace centralized systems with decentralized systems.

## *3.2 Potential IoT Applications by Archetype*

At Taraxa, we identified three distinct categories of IoT-relevant blockchain applications. Because this is a nascent and rapidly evolving space, this only represents our latest thinking at the time of writing.

### 3.2.1 IoT Data Anchoring

IoT devices generate a great deal of data, and when that data is shared across entities it needs to be trusted. One way to augment trust is to establish unique identities for each data-generating device and have the devices anchor the data they have collected onto the blockchain.

The anchoring process involves placing a hash (a function that maps data of arbitrary size into data of fixed size in a way that minimizes collisions—different pieces of data cannot map into the same hash) of a data set collected by the device along with its signature into say a smart contract as a record of the provenance of the data as well as proof that the data has not been tampered with. Only the hash of the data set should be stored on-chain, and not the full data set, because we want to minimize the load and cost of using the blockchain, and the signature guarantees that the data came from the device. Of course, this would also require that the device manufacturer publish the public keys embedded (preferably via secure hardware) into their devices.

Examples:

- **Cold chain logistics**: a supermarket that is taking delivery of milk shipped via cold chain would like to have guarantees that the milk has been properly refrigerated throughout its route. If the refrigeration units were turned off during shipping for a few hours and turned back on, the supermarket would not be able to tell the difference until the milk started to spoil much earlier than expected. Location and temperature sensors could be installed on each refrigerated truck. These would intermittently upload data as well as anchor that data onto the blockchain, ensuring that the shipping company has not tampered with the data after the event.
- **Public infrastructure monitoring**: with the advent of public–private partnerships, many local governments are increasingly outsourcing the management and maintenance of public infrastructure such as bridges, roads, and tunnels to private companies. Once outsourced, the government has a responsibility to ensure that public infrastructure is being well maintained and that the data reported are accurate (e.g., toll income, maintenance expenditure). Sensors installed on such public infrastructure (e.g., cameras, strain gauges, moisture) will also anchor the data they collect onto the blockchain to prove to local governments that the data has not been tampered with.

### 3.2.2 Machine Monetization and Eventual Tokenization

Many assets being monitored by sensors are revenue-generating machines that require significant upfront capital outlay to deploy. For many new ventures, obtaining funding or loans through traditional financing channels may be challenging.

With proper data anchoring, the earning capabilities of such machines can be tracked in real time on the blockchain, providing proof and generating expectations for future income. Businesses can then solicit potential customers to invest in shares

of these machines by issuing digital tokens on the blockchain. This has the dual benefit of not only raising funds from a much wider pool of potential investors, but also that the tokens are likely to end up in the hands of customers who have interest in purchasing the services of the machine in the future.

Examples:

- **Shared vehicles**: require the operator to not only make large upfront investments to purchase a fleet of cars but also run continuous and aggressive marketing campaigns to generate awareness. If each vehicle's location, movement, mileage, and, most importantly, income data were to be released to the public (anonymized to protect driver privacy), each vehicle could be tokenized and those who purchase the tokens would receive a heavy discount when renting the shared vehicle. Not only does this alleviate the pressure of extremely large upfront investments, it also ties customers into an ecosystem that they now have a stake in and financial incentives into drive the service's adoption, elegantly killing two birds with one stone.
- **Smart vending machines**: vending machines are becoming increasingly popular around the world as a low-cost and highly convenient alternative to manned storefronts, but their value remains underutilized. With so many machines deployed on every street corner, they could easily be outfitted with additional IoT infrastructure to enable them to collect data of their surroundings (e.g., foot and vehicle traffic, localized weather) which could be sold, become distribution points for humanitarian aid (e.g., disaster relief, charitable giving), or even serve up a far more accurate alternative geo-location service (e.g., via WiFi triangulation) in an urban environment to GPS. All of which could not only help further monetize these vending machines but also provide social services far beyond their original design goals.

Machine to Machine Economy

The ultimate application is to enable machines to trade with one another autonomously. This of course would require a very high level of intelligence and autonomy on the part of the machines, but a variety of mechanisms could be designed where machines can discover and purchase resources they require to optimally complete their stated objectives.

One of the most common mechanisms is a marketplace; here we look at two potential applications.

- **Agricultural drones** could make real-time decisions on plant protection (e.g., insecticide) deployment across a field based on its internal models and by purchasing data from the surrounding sensors. It could purchase or trade data with cameras in the surrounding fields for analytics on pest detection, from local weather sensors to predict chances of rain so that the chemicals will not be washed off minutes after deployment, from local soil moisture sensors to customize the level of dilution, etc. Once machines are automated by AI and empowered by blockchain, they can operate and interact with one another with minimal intervention.

- **Traffic routing** for autonomous vehicles may be very different than for vehicles operated by humans. Vehicles could, in real time, communicate their intended destination, urgency, and willingness to pay surrounding vehicles so that traffic could be routed and reshaped in real time according to supply and demand. Usual visual and audio cues created for human drivers would not be necessary (e.g., traffic lights, turn signals), and in their place would be a dynamic real time bidding market for road space as a commodity.

## References

Bernanke BS (2018) The real effects of the financial crisis. In: BPEA conference draft, vol Fall

Churyumov A (2016) Byteball: a decentralized system for storage and transfer of value, 1 Oct 2016 [Online]. Available: https://byteball.org/Byteball.pdf. Accessed 17 Dec 2018

Columbus L (2018) 10 charts that will challenge your perspective of IoT's growth. Forbes, 6 June 2018 [Online]. Available: https://www.forbes.com/sites/louiscolumbus/2018/06/06/10-charts-that-will-challenge-your-perspective-of-iots-growth/#47c74fd13ecc. Accessed 15 Nov 2018

De N (2018) Smart contracts now recognized under Tennessee Law. Coindesk, 23 Mar 2018 [Online]. Available: https://www.coindesk.com/blockchain-bill-becomes-law-tennessee. Accessed 15 Nov 2018

Ethereum White Paper (2018) 21 Aug 2018 [Online]. Available: https://github.com/ethereum/wiki/wiki/White-Paper. Accessed 17 Dec 2018

EU GDPR.ORG (2018) GDPR FAQs. EU GDPR.ORG [Online]. Available: https://eugdpr.org/the-regulation/gdpr-faqs/. Accessed 15 Nov 2018

Graff GM (2017) How a dorm room minecraft scam brought down the internet. WIRED, 13 Dec 2017 [Online]. Available: https://www.wired.com/story/mirai-botnet-minecraft-scam-brought-down-the-internet/. Accessed 15 Nov 2018

Granville K (2018) Facebook and Cambridge analytica: what you need to know as fallout widens. The New York Times, 19 Mar 2018 [Online]. Available: https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html. Accessed 15 Nov 2018

IBM Research Editorial Staff (2018) Elegant, disgusting cryptography. IBM, 21 Mar 2018 [Online]. Available: https://www.ibm.com/blogs/research/2018/03/elegant-disgusting-cryptography/. Accessed 15 Nov 2018

Investopedia (2018) Bitcoin. Investopedia [Online]. Available: https://www.investopedia.com/terms/b/bitcoin.asp. Accessed 31 Oct 2018

MacMillan D, McMillan R (2018) Google exposed user data, feared repercussions of disclosing to public. Wall Street Journal, 8 Oct 2018 [Online]. Available: https://www.wsj.com/articles/google-exposed-user-data-feared-repercussions-of-disclosing-to-public-1539017194. Accessed 15 Nov 2018

Nakamoto S (2008) Bitcoin: a peer-to-peer electronic cash system, 31 Oct 2008 [Online]. Available: https://bitcoin.org/bitcoin.pdf. Accessed 31 Oct 2018

Narayanan A, Clark J (2017) Bitcoin's academic pedigree. acmqueue 15(4)

Narula N (2017) Cryptographic vulnerabilities in IOTA. Medium, 7 Sept 2017 [Online]. Available: https://medium.com/@neha/cryptographic-vulnerabilities-in-iota-9a6a9ddc4367. Accessed 2 Nov 2018

Popov S (2018) The tangle, 30 Apr 2018 [Online]. Available: https://assets.ctfassets.net/r1dr6vzfxhev/2t4uxvsIqk0EUau6g2sw0g/45eae33637ca92f85dd9f4a3a218e1ec/iota1_4_3.pdf. Accessed 17 Dec 2018

Sompolinsky Y, Zohar A (2013) Secure high-rate transaction processing in bitcoin. 2013 [Online]. Available: https://eprint.iacr.org/2013/881.pdf. Accessed 17 Dec 2018

Sompolinsky Y, Lewenberg Y, Zohar A (2016) SPECTRE: serialization of proof-of-work events: confirming transactions via recursive elections, 2016 [Online]. Available: https://eprint.iacr.org/2016/1159.pdf. Accessed 17 Dec 2018

Sompolinsky Y, Zohar A (2018) PHANTOM, GHOSTDAG: two scalable blockDAG protocols. 2018 [Online]. Available: https://eprint.iacr.org/2018/104.pdf. Accessed 17 Dec 2018

Tung L (2017) IoT devices will outnumber the world's population this year for the first time. ZDNet, 7 Feb 2017 [Online]. Available: https://www.zdnet.com/article/iot-devices-will-outnumber-the-worlds-population-this-year-for-the-first-time/. Accessed 31 Oct 2018

TrueBench (2018) TrueBench, 31 Oct 2018 [Online]. Available: http://truebench.the-toffee-project.org/

World Economic Forum (2018) Blockchain beyond the hype, World Economic Forum, Geneva

**Steven Pu** is an entrepreneur and strategy consultant. Prior to founding Taraxa, he co-founded multiple venture-funded startups in wireless mesh networks, mobile health, and healthcare services. He was also a director at Monitor Deloitte's China office, spearheading the firm's digital strategy practice, advising senior executives of multinational corporations on how large organizations could effectively capture innovation and respond quickly to market disruptions.

Mr. Pu holds undergraduate and master's degrees in electrical engineering from Stanford University.