A Review of Prospective Applications of Blockchain Technology in the Railway Industry

J. D. Preece[a,*], J. M. Easton[a]

[a] *Birmingham Centre for Railway Research and Education, University of Birmingham, United Kingdom*

Abstract

This paper presents three potential applications for blockchain technology within the railway industry: digital ticketing; logistics and supply chain processes; and distribution of data. Problems with existing systems are identified and alternatives based on blockchain technologies are presented as solutions. A mechanism for digital ticketing is introduced along with a chronological representation of the steps involved and the potential commercial considerations. Two public blockchain platforms (Ethereum and NEO) and one bespoke platform (Hyperledger Fabric) are proposed as commercial solutions for this particular use case. Furthermore, the domain of logistics and supply chain processes is identified as another area that will benefit from blockchain technology. The public platform VeChain is introduced as a commercial solution for this use case. Finally, issues with centralised data distribution are discussed, and blockchain technology is presented as a secure and distributed alternative.

Keywords: blockchain, ticketing, logistics, data, Ethereum, NEO, VeChain, Hyperledger

1. Introduction

In the year 2009, Satoshi Nakamoto initiated the genesis block on his groundbreaking Bitcoin platform. The Bitcoin platform allows users to anonymously send other users a desired amount of Bitcoin. The platform guarantees transaction integrity by applying numerous cryptographic techniques to ensure security, privacy, and efficiency. The platform has inspired numerous other platforms which aim to improve upon the already ageing protocols of the Bitcoin platform.

As the blockchain ecosystem expands into new domains, public interest in the technology also increases. From December 2016 to December 2017, searching for the word 'blockchain' on Google increased by 670%.[1] Though the sentiment is that the majority of interest is generated by investors intending to make a considerable gain of equity, or those seeking

[*] Corresponding author

Email addresses: jdp225@student.bham.ac.uk (J. D. Preece), j.m.easton@bham.ac.uk (J. M. Easton)

URL: http://postgrad.eee.bham.ac.uk/preecej/ (J. D. Preece)

[1] Value calculated from data provided by Google Trends.

transparency of information in a world of centralisation and concealed data, the focus is shifting to how blockchain technology can support enterprise solutions in a variety of scenarios across a plethora of industries. The railway industry is no exception to these emerging technologies, offering a variety of applicable use cases.

In 2013, Network Rail released the Network Rail Technical Strategy (NRTS) in response to the Rail Technical Strategy (RTS) published by the Rail Safety and Standards Board (RSSB), outlining areas requiring innovation to transform the railway industry in the United Kingdom (UK) into a railway suitible for the 21st century. [1] Within the theme of customer experience, the document proposes "smart ticketing systems that allow access through virtual barriers and [integration] with other transport modes" as a requirement for the future. Though mobile ticketing has emerged in the UK since the document was released, innovation is still required to improve the system and eliminate paper ticketing altogether. Furthermore, the document proposes "increased data and information flow" by "securing and protecting information and systems" and "creating common information architectures and protocols". Transparency of data is crucial for stakeholders and passengers using the railway. Blockchain technology offers solutions to the aforementioned concepts by providing a distributed platform that ensures data is both transparent and secure.

2. Background

In order to properly discuss how blockchain technology will be used in the railway industry, it is crucial to understand the mechanics behind a blockchain platform. Blockchains are secure, private, and immutable data structures. This functionality is made possible by a number of governing principles, summarised in the following subsections.

2.1. Cryptographic Primitives

Cryptographic techniques are entrenched into the blockchain. Narayanan et al. discuss two primitive facets of cryptography that are crucial to cryptocurrencies; hash functions and digital signatures. Understanding the primitives will open the gate to understanding how the mechanics of a blockchain achieve the goals mentioned above. [2]

2.1.1. Hash Functions

A hash function is a defined mathematical function with the goal of mapping an input of any size to a fixed-sized output. An input can be data of any sort: a message; a document; a previously hashed value etc. The hash function taked this input, and efficiently computes the fixed size output, known as the hash digest. Though numerous hash functions exist, a has function must hold three properties to be compatible with a blockchain platform: collision resistance; hiding; and puzzle-friendliness. The Secure Hash Algorithm (SHA) is the hash algorithm used on the Bitcoin platform. [2, 3]

Collision Resistance. A collision occurs when a hash function returns an identical output value for two unique input values. It is mathematically impossible to avoid collisions, even for robust hash functions. The domain of inputs is theoretically infinite (as one can provide an input of any length) whilst the range of outputs is limited, as the output length is of fixed length. For a 256-bit output, the range of distinct outputs is $2^{256}+1$, notably smaller than the domain of inputs. Nonetheless, hash functions are designed to be collision resistant. A hash function H is said to be collision resistant if it is infeasible to find two values x, y such that $x \neq y$ and $H(x) = H(y)$. [2] This means it is unlikely for a collision to occur despite being possible, and it would take an unreasonable amount of time to discover such a collision.

Hiding. It must be infeasible to discover the original input value to a hash function once the output value is known. This keeps the original information passed to the hash function secure. In certain scenarios, it is easier to come to a conclusion about the value of the input. For example, in the scenario of hashing the result 'heads' or 'tails' from a coin toss, there will only be two possible hash outputs. The input is easily deducible by running the hash function for both possible values, and then comparing any results to these computed outputs. To avoid this technique of working back to discover the input, the input is concatenated with a random value r (chosen from a high minimum entropy probably distribution) to 'hide' the input. A hash function H is said to be hiding when it is infeasible to find x given $H(r\|x)$.

Puzzle Friendliness. This property is not essential to every hash function, but plays a crucial role for platforms with a native currency and a Proof-of-Work (PoW) consensus mechanism (explained in Section 2.2.3). As an example, take a scenario where data x is inputted to a hash function H, and the output value y must fulfil a criterion. The set of acceptable values for y is Y; if y is not in Y, then the result is rejected. Assuming $H(x)$ doesn't instantly meet this criterion, x must be modified to return a different y until a suitable result is discovered. To do this, x is concatenated with differing values of a nonce[2] i until y is in Y. Mathematically, it is necessary that $H(i\|x) \in Y$. The hash function H is said to be puzzle friendly if for every possible n-bit output value y, if a nonce i is chosen from a high minimum entropy probability distribution, then it is infeasible to find x (such that $H(i\|x) = y$) in time significantly less than $O(2^n)$. Simply put, this means that it must take an elongated period of time to find a suitable nonce.

2.1.2. Digital Signatures

A digital signature is analogous to a physical signature; the signature is unique to the signer and verifiable by anybody. This is performed using a key pair: a secret key k_{s}, which only the signer knows and uses to sign the data d; and a public key k_{p}, which anybody may know and is used to validate any signatures made. A signature σ is determined by using a signature scheme algorithm S so that $S(k_{\mathrm{s}}, d) \to \sigma$. The signature x is then verified using a verification algorithm V (provided by the same signature scheme as S) such that $V(k_{\mathrm{p}}, \sigma) \to \{\top, \bot\}$ i.e. the verification algorithm confirms signature ownership as true or false. Bitcoin uses the Elliptic Curve Digital Signature Algorithm (ECDSA) with secp256k1 parameters [3].

[2] A nonce is a number used only once.

2.2. Blockchain Mechanics

With a basic cryptograpghic toolkit, one can begin to conceptualise a blockchain and the associated technologies embedded within a blockchain platform.

2.2.1. Structure

Consider a regular linked list where each 'block' of the list contains data and a link that refers to its succeeding block. This link is known as a pointer, and is the memory location of the succeeding block. A hashed linked list is analogous to a regular linked list. However, instead of a pointer to a memory location, the hashed linked list uses a hash pointer. A hash pointer is a hash digest of the previous block. This is significant because it describes where to retrieve the information and allows instantaneous verification of the data to ensure it hasn't been tampered with. This makes the entire blockchain a tamper-evident log. Should an adversary attempt to change any data in a block, the block's hash digest will change, resulting in an incorrect hash pointer in the succeeding block, thus breaking the chain.

Within each block is: a hash pointer pointing to the previous block; a list of transactions in the form of a Merkle tree; a timestamp; and a number that fits the criterion discussed in Section 2.1.1. A Merkle tree is a "tree structure in which each leaf node is a hash of a block of data, and each non-leaf node is a hash of its children." A tree is used in order to make use of the average search complexity of $O(\log_2(n))$ in order to efficiently verify the transactions stored within the tree. [4]

A unique property of a blockchain is how and where the data is stored. Instead of a centralised database maintaining the latest version of the blockchain, it is stored by nodes on a peer-to-peer network to share the computational processes and data across the network. Further to this is the concept of decentralisation, where the nodes work on localised copies of the blockchain instead of a centralised copy. This is illustrated in Figure 1. Decentralised networks have a number of properties that provide benefits over centralised networks: they are fault tolerant as they rely on a number of nodes which are unlikely to fail simultaneously as opposed to relying on one node prone to failure; they are attack resistant, as there is no

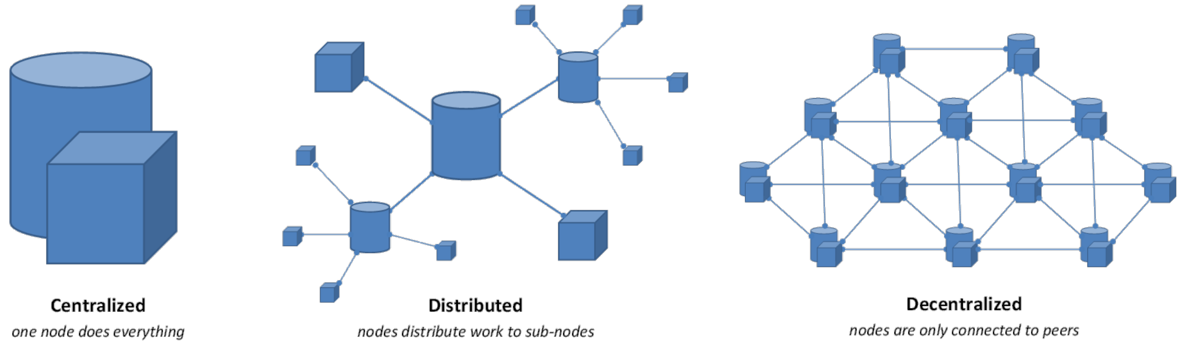
Centralized
one node does everything
Distributed
nodes distribute work to sub-nodes
Decentralized
nodes are only connected to peers

Figure 1: Visual representation of centralised, decentralised, and distributed networks. [Original image: `https://i.stack.imgur.com/hDDzg.png` Accessed: 2018-04-18]

sensitive central node; and the are collusion resistant, as it is incredibly hard for adversaries to combine efforts to manipulate the network. [5]

Each node stores a copy of the blockchain, and performs as a passive node or an active node. The purpose of a passive node is to utilise ad-hoc messaging with neighbouring nodes to maintain the latest version of the blockchain. A passive node does not perform the mining of new blocks to add to the blockchain. Instead, they wait for messages from neighbouring nodes instructing them of new blocks in order to validate them and update their locally stored blockchain. An active node participates in the mining of new blocks.

2.2.2. Transactions

A transaction takes place when a user of one wallet sends some of a platform's native currency (henceforth denoted as χ) to another wallet. For a simplified example, assume User A wants to send 0.2χ to User B. User A takes note of User B's public key. Once they know this information, they use specialised software to input the destination wallet address, the amount of 0.2χ, and confirm that they wish to proceed with the transaction. Upon confirmation the transaction is digitally signed by User A's private key (using the signature scheme signing algorithm as discussed in Section 2.1.2) and then propagated throughout the network. Active nodes begin mining blocks full of new transactions in order for them to be added to the blockchain. Once an active node contributes and transmits a new block, passive nodes verify the the transactions within the block by using the signature scheme verification algorithm with the transaction data and the public key. Anybody can validate a transaction, as the transaction data and public key are always accessible on the blockchain.

2.2.3. Mining

To expand the chain, new blocks need to be mined. Mining is designed so that any active node can contribute a block to the chain, but must perform some work to achieve this; otherwise, the system would bottleneck with an overload of potential block contributions. There are two widely used consensus mechanisms that decide how blocks are mined; PoW and Proof-of-Stake (PoS).

Proof-of-Work. With a PoW mechanism, miners compete to find a valid block first. The incentive for miners to carry out mining work is the block reward (a quantity of the currency) and transaction fees they receive for successfully mining a block. Mining can be particularly lucrative depending on the success of a platform; the reward for mining a Bitcoin block as of December 2017 was 25 BTC, worth approximately \$500,000.

A block is mined by discovering the nonce n that concatenates with the block's hashed data x to provide a new hashed value y below a certain difficulty threshold D, such that

$$H(n\|x) \in \{y | y \in \mathbb{R}_{>0}, y < D\} \tag{1}$$

holds true. This difficultly value is dynamic, and will usually be calculated to keep the time interval between blocks being mined the same length. An approximate difficulty equation is

$$D_n = \frac{D_{n-1}NT}{t_{n-1}} \tag{2}$$

where: D_n is the new difficulty; D_{n-1} is the difficulty of the previous block; N is the number of blocks this new difficulty will last for; T is approximate time between blocks being mined; and t_{n-1} is the time it took to mine the last N blocks. The values of N and T are built into the protocol and never change. For Bitcoin, these values are 2016 and 10 minutes respectively. This allows the difficulty to be adjusted approximately every two weeks in order to keep the network mining blocks every ten minutes, no matter what the hash rate[3] of the network is.

Once a suitable nonce has been discovered, the mining node transmits this block to the rest of the network in order for it to become accepted to the consensus chain. As a peer-to-peer network, this process is not instantaneous; it takes a certain amount of time for neighbouring nodes to validate the block and update their local copies of the blockchain, and then transmit this to their neighbours. Because of this latency, it is probable that on some occasions multiple blocks are discovered simultaneously. In these situations, passive nodes are susceptible to adding the first block it is informed of; this is nearly always the block mined by the closest active node. The chain is then built on whichever block the next successful miner chooses to build upon, or whichever block the network as a whole decide to accept. There is no central authority to decide this; individual nodes get to choose how to build their local blockchain, though it is beneficial to stay consistent with the accepted chain. Due to the competition, some blocks may become discontinued as nodes choose other blocks to mine on and become the consensus chain. These blocks are known as orphan blocks, and are illustrated in Figure 2.

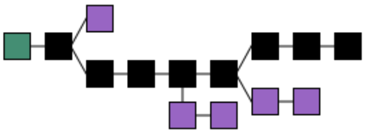

Figure 2: The blockchain illustrated: the green block is the genesis block; the black blocks are those on the consensus chain; and the purple blocks are the rejected 'orphan' blocks. [Original image: https://en.bitcoin.it/wiki/File:Blockchain.png Accessed: 2018-04-18]

In some cases, two separate groups of nodes on the network may choose to continue two chains, disagreeing on a particular consensus chain. This is known as a fork. There are two variations; a soft fork and a hard fork. A hard fork occurs when two chains become incompatible due to changes of the protocol. Some nodes may continue to use legacy software, whilst some may employ the new rules; as they are incompatible, two chains will diverge from one another. A soft fork is completely backwards compatible with the existing chain.

PoW is extremely energy consuming as all miners simultaneously use computational power to find a valid block. Moreover, it is susceptible to a 51% attack, where an individual or collaborative group control over 50% of the network hash rate. This would allow the attackers to have control over which transactions are added to the blocks in the consensus chain.

[3] The hash rate of the network is the number of hashes being calculated per second in order to mine a block.

Proof-of-Stake. Conversely, PoS randomly selects an active node to mine the next block based upon their stake of the blockchain being mined. If one particular node has a large amount of the associated currency, they are far more likely to be selected. This mechanism reduces energy expenditure as computational power is only used upon selection. Furthermore, the network is more robust against a 51% attack, as it is infeasible for a single party to hold 51% of the currency.

2.2.4. Scalability

As platforms become widely adopted, the stress on the network increases. Despite the greater number of transactions that require validating, the platform has an upper limit of how many transactions it can process per second. This is due to the choice of consensus mechanism, and the protocol. There are two technologies that exist to resolve scalability issues associated with blockchain technology; sharding and swarming. Sharding logically partitions data into shards, which are distributed across the nodes on the network. This prevents nodes from becoming cumbersome. Swarming is a collection of shard-nodes to store and manage the data. Multiple swarms make up the whole network. These technologies keep the data fragmented as to not overload the platform, yet connected to provide efficiency when accessing the data. [6]

2.3. Blockchain Use Cases

Though the initial blockchain platforms are used purely as the mechanism of a digital currency, there are been a number of alternative platforms that are emerging that focus on new applications. As of December 2017, there were over 1000 platforms listed on CoinMarketCap.

2.3.1. Currency

The intention of the Bitcoin platform is to provide a mechanism for a digital currency. Anybody can obtain a wallet address in order to send and receive Bitcoin to and from other wallets. Whilst there have been attempts to alter the usage of the Bitcoin platform in order to provide higher functionality such as multi-party lotteries and proof-of-property, the original protocol is rather restrictive in any attempts to implement changes. As such, the core purpose is still digital currency. In addition, there have been numerous alt coins released whose purposes are purely financial, which seek to improve upon Bitcoin's protocols. Examples include Monero, Litecoin, and Stellar.

2.3.2. Smart Contracts

Smart contracts were first introduced by the Ethereum platform. A smart contract is a Turing-complete script that exists in the blockchain itself. These smart contracts are be uploaded to the blockchain via a transaction. These scripts are then be executed in the Ethereum Virtual Machine (EVM) which run on each node of the Ethereum network. Because of the Turing-complete nature of the language the scripts use, they can perform many more tasks than that of Bitcoin's script language.

A particular application of smart contracts is betting. Say a user wishes to place a bet on the result on who will win the Grand National[4]. The user sends some Ether (the currency of Ethereum) to a smart contract address (written by a betting agency), which then holds this currency and runs the contract. The contract will wait for confirmation of the result and only then will send the winnings back to the user or keep the original stake. This means that both parties must follow the exact specification of the contract. The user can trust the contract as it is public on the blockchain, meaning they can read and understand the what the smart contract does before sending any Ether to it.

2.3.3. Securities

A security is defined as "a thing deposited or pledged as a guarantee of the fulfilment of an undertaking or the repayment of a loan, to be forfeited in case of default."[5] An example is holding a stock of a company which will return a dividend in time. Utilising the blockchain would allow verification of ownership and automatic dividend payment to any verified stock holders through a blockchain platform.

2.3.4. Record Keeping

A blockchain is an immutable data structure stored across numerous nodes on a network, and the data on it is secure and tamper-proof. This property is useful when proof-of-ownership is of concern. A user may own a token in their wallet which entitles them to whatever the token is representative of (a ticket, a property, or another type of asset) but has to prove that they own it in order to utilise it. By utilising digital signatures, anybody can validate the transactions of particular token in order to check who rightfully owns the them.

2.4. Existing Blockchain Platforms

There are numerous blockchain platforms in existence that aim to provide functionality for a particular set of requirements. Listed below are some highly regarded platforms which provide resources that align with the scenarios presented in Section 3.

2.4.1. Ethereum

Ethereum is a distributed computing platform that utilises smart contracts stored on the blockchain. [7] These smart contracts are written in Solidity, a Turing-complete scripting language. The native currency of the platform is called Ether, which is the underlying asset transferred between the Ethereum wallets. Ethereum uses a PoW consensus mechanism, minting new Ether upon the mining of new blocks. There are plans to move to a PoS mechanism to reduce environmental impact and to increase robustness of the network.

To execute a smart contract, Ether is sent to the wallet address the contract is stored at. The contract requires a minimum amount of Ether to execute, which is specified by the contract itself. This minimum amount is known as the amount of 'Gas'. Gas determines

[4] The Grand National is a horse race held annually at Aintree, UK

[5] Definition provided by the Oxford Dictionary

how much processing power the contract will receive on the EVM. This mitigates spam on the network, as contracts are costly to run and will time out if the Gas runs out. The block time of Ethereum is approximately 15 seconds, and the network has a practical and theoretical transaction speeds of 15 tps and 30 tps respectively.

Ethereum is designed in order for a variety of Decentralised Applications (DApps) to be designed and run upon the Ethereum platform. Hosting a DApp on such a platform means the smart contracts can run "without any possibility of downtime, censorship, fraud or third-party interference" [8], as the computation is distributed across the nodes running the EVM on the Ethereum network. Other tokens can be built on top of the Ethereum network to be transacted alongside Ether. These tokens must meet Ethereum Request for Comments Proposal 20 (ERC20) protocol standards in order to work seamlessly with the network. ERC20 is a set of rules that a token must follow in order to provide consistency between all other tokens operating on the Ethereum platform.

Ethereum has already had one major fork, after a large quantity of Ether was stolen after users exploited a vulnerability in a smart contract provided by The DAO. Whilst some users decided to continue mining blocks on top of these dishonest transactions, others decided to refund the victims with freshly minted Ether and proceed mining from before the dishonest transactions. A fork occurred, and ultimately split into two blockchains; Ethereum Classic and Ethereum respectively.

2.4.2. NEO

NEO has a similar philosophy to Ethereum; that is to say, NEO is a platform for building DApps. However, there are rudimentary differences between the platforms, both in the protocols and the objectives. NEO will focus on becoming the preferred platform for a smart economy, meaning access to and ownership of tokens on a distributed and decentralised ledger which is still under the regulation of government. This is important for the trust and safety of law-backed assets, whilst maintaining the distributed nature of the blockchain to avoid a centralised monopoly. [9]

In terms of technical specifications, NEO uses an advanced PoS consensus mechanism known as Delegated Byzantine Fault Tolerance (dBFT). Simply put, NEO token holders vote when a block is mined. Should less than 66% of the voters agree that the block is valid, this block will be ignored and another block will be proposed and voted upon. The nodes proposing these blocks are know as bookkeeper nodes, and they maintain the network. This allows NEO to scale with a greater number of transaction requests, as only the bookkeeper nodes add the blocks to the chain. This also means that no forking occurs on the NEO blockchain.

One further aspect of NEO is the decoupling of the NEO and GAS platforms. As explained above, Ether is the native currency of Ethereum, and the gas used for computation is just a small amount of Ether. NEO is different in that NEO and GAS are separated onto separate blockchains. NEO tokens are used as a share in the platform, in which holders can vote for bookkeepers and receive GAS dividends depending on their NEO stake. GAS tokens are used for all operations on the NEO platform e.g. to pay for computation for smart contracts. GAS is minted by mining NEO blocks or as dividends to NEO token holders.

2.4.3. VeChain

VeChain is a platform that exists to enhance the supply chain management process between stakeholders. Every aspect of movement within the supply chain is recorded onto the distributed ledger in order to provide transparency between all parties involved. This is achieved by the means of asset digitisation, where products are assigned unique identifiers on the platform. A technology known as VeChain Identity (VID) is associated with a particular product, using SHA to generate such a VID. The VID is stored as one of the following forms: a Near Field Communication (NFC) tag; a Quick Response (QR) code; or a Radio Frequency Identification (RFID) tag. VeChain is currently built on the Ethereum platform as a ERC20 standard token. However, the VeChain development team have laid out a plan to migrate the existing blockchain on the Ethereum network to its own platform. [10]

2.4.4. Hyperledger

Hyperledger is an umbrella term for a set of open-source blockchain platforms. The project was initiated in 2015 by the Linux Foundation in order to provide businesses with the tools to develop distributed ledger technologies. There are four main blockchain platforms: Burrow, a client including an EVM; Fabric, a permissioned blockchain infrastructure with modular architecture and delineation of roles across nodes, smart contracts, plus configurable consensus mechanisms; Iroha, which focuses on mobile applications; and Sawtooth, which uses a alternative consensus mechanism know as Proof-of-Elapsed-Time (PoET). There are also a number of developer tools to help with enterprise blockchain management.

The true purpose of the Hyperledger platforms are to provide interoperability between any companies utilising them. There is no currency affiliated with any of the projects, as no mining is required; the transactions are validated by suitably chosen consensus mechanisms that are internal to the business using them. This means that the blockchain is private to the companies using them, differentiating the Hyperledger platforms from any of the aforementioned public platforms.

3. Use Cases in the Railway Industry

This section presents various sectors of the railway industry that could utilise blockchain technology to supplement the existing Information Technology (IT) systems and regulations. There are three distinct sectors: digitised ticketing; logistics; and data distribution.

3.1. Digital Ticketing

According to the NRTS, the uptake of digital technologies has the potential to lead to significant improvements in customer experience on the UK railway network. Part of this will be down to the way customers purchase, store, and use tickets. There are currently two methods passengers may follow to obtain a valid rail ticket in the United Kingdom; buying in person at the station, or buying from an online vendor.

3.1.1. Current Process

At present, passengers purchasing in person use a ticket machine or a manned ticket kiosk. Once the user has selected their destination and provided any other information (such as railcards[6], intended time of travel etc.), they are presented with the price of the journey and can pay with cash or a debit/credit card. One advantage of paying at the station is that the passenger does not require any form of technology; mobile phones, NFC devices, and debit/credit cards are not necessary. Furthermore, there are no booking fees the passenger has to pay.

One disadvantage is that the passenger only receives one form of proof that they hold a valid ticket for the journey; the paper ticket itself. Should this ticket be misplaced or incorrect through an error whilst ordering the ticket, it is be deemed invalid and is subject to a penalty charge. Another disadvantage is that a passenger may not always have the correct form of payment, or the adequate amount of payment. Furthermore, passengers may fraudulently enter information to obtain discounts they are not entitled to. Such instances could be purchasing a child ticket as an adult, lying about the existence of a railcard, or buying a ticket to an incorrect destination to avoid paying the full price of a ticket. Purchasing tickets through an online vendor has advantages over buying tickets at the station. It is easier to compare prices between journeys, to reserve seats, and to ensure no mistakes are made during the process. Though a number of vendors exist, Trainline are the most widely used. Whereas most vendors instruct passengers to collect tickets at a station or to print them, Trainline offers mobile tickets in addition to these options.

Trainline provides a number of reasons why the mobile tickets they offer are advantageous: one can spend less time at the station and in queues and make orders wherever and whenever; there is no need to carry a paper ticket (or more than one paper ticket, should the journey demand this); it is better value for money to buy online; one can send tickets to somebody else; and, it is environmentally friendly. [11]

The disadvantages of buying online are that if the passenger does buy from a vendor that requests pick-up of the ticket from a station, then the passenger is still subject to losing the physical copy of the ticket; the receipt is not valid evidence to avoid a fine. passengers must also pay administration fees which can become costly for multiple short journeys. In addition, if a passenger opts to have a mobile ticket and runs out of battery or loses their phone, there is no way of proving ownership of a valid ticket; the passenger cannot use a receipt, and the operating franchise on that particular route does not have access to Trainline's data. Furthermore, mobile tickets are not available across the entire United Kingdom as of yet, as shown in Figure 3.

3.1.2. Blockchain Technology as a Solution

A blockchain based digital ticketing platform will resolve the issues associated with the two aforementioned methods of procuring tickets.

Prerequisites. Any passenger wishing to purchase tickets will require a digital wallet for the blockchain platform. Wallet addresses are automatically generated by generating a private

[6] A railcard is a card which indicates the holder is subject to a discount

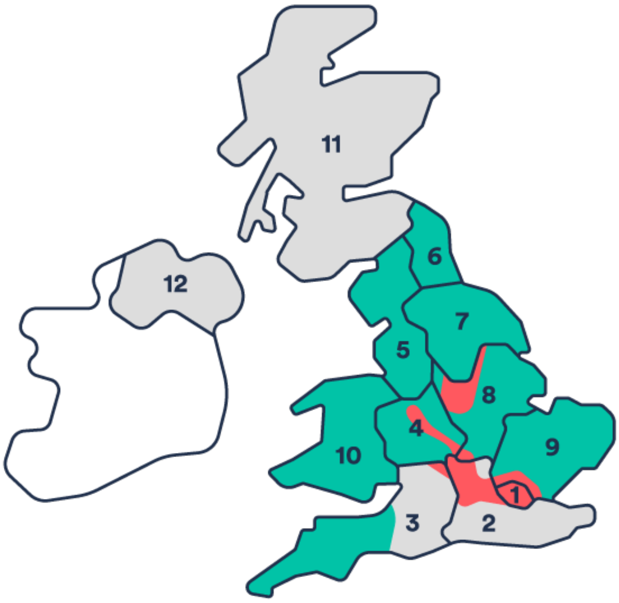

Figure 3: The Trainline provide this map of the United Kingdom, highlighting locations where their mobile tickets are (and are not) available. The colours indicate different stages of availability: green represents mobile ticket support for all routes in that area; red represents limited availability; and grey represents unavailable areas. [Source: `https://www.thetrainline.com/information/mobile-tickets`, Accessed: 2018-03-17]

key, determining a public key from the private key, and then generating a wallet address. Wallets will be accessible through a number of technologies: a mobile application with a QR code; a mobile application with access to the phone's NFC; a physical smart card with a QR code; a physical smart card with NFC capabilities; or the raw public and private keys. Because the blockchain is distributed, and the passenger will have access to their public, private, and wallet keys, they may host their wallet on multiple platforms. The transactions that occur will ultimately be made to and from the same address on the blockchain. This provides passengers with a variety of choices and reserve options.

The passenger can then apply for any discounts that they are entitled to. This may be an age-restricted railcard or an employee discount for certain routes. A digital token will be generated by a smart contract encoding all of the information regarding how much discount the token owner is entitled to, which routes the discount is to be applied, and how long the discount is valid for. The smart contract will only issue the token if all prerequisites for a discount are met, meaning there is no opportunity for the passenger to commit fraud. The token is then sent to the wallet of the applicant. The transparency of the blockchain means that vendors can check this wallet address for any discounts to apply, whilst the security of the blockchain means that nothing about the passenger is revealed.

Ticket Acquisition. The process of acquiring a ticket with a blockchain entirely depends on the choice of platform. As has been mentioned, there are a vast number of blockchain platforms that exist or are in development that provide functionality for use cases of this type.

There are three in particular which must be considered: Hyperledger Fabric, Ethereum, and NEO. The generic process flow for purchasing and holding a ticket is illustrated in Figure 4. Figure 4 summarises the interactions between the passenger, the ticket vendor, and the blockchain. Below is further clarification of the actions taking place within the flow:

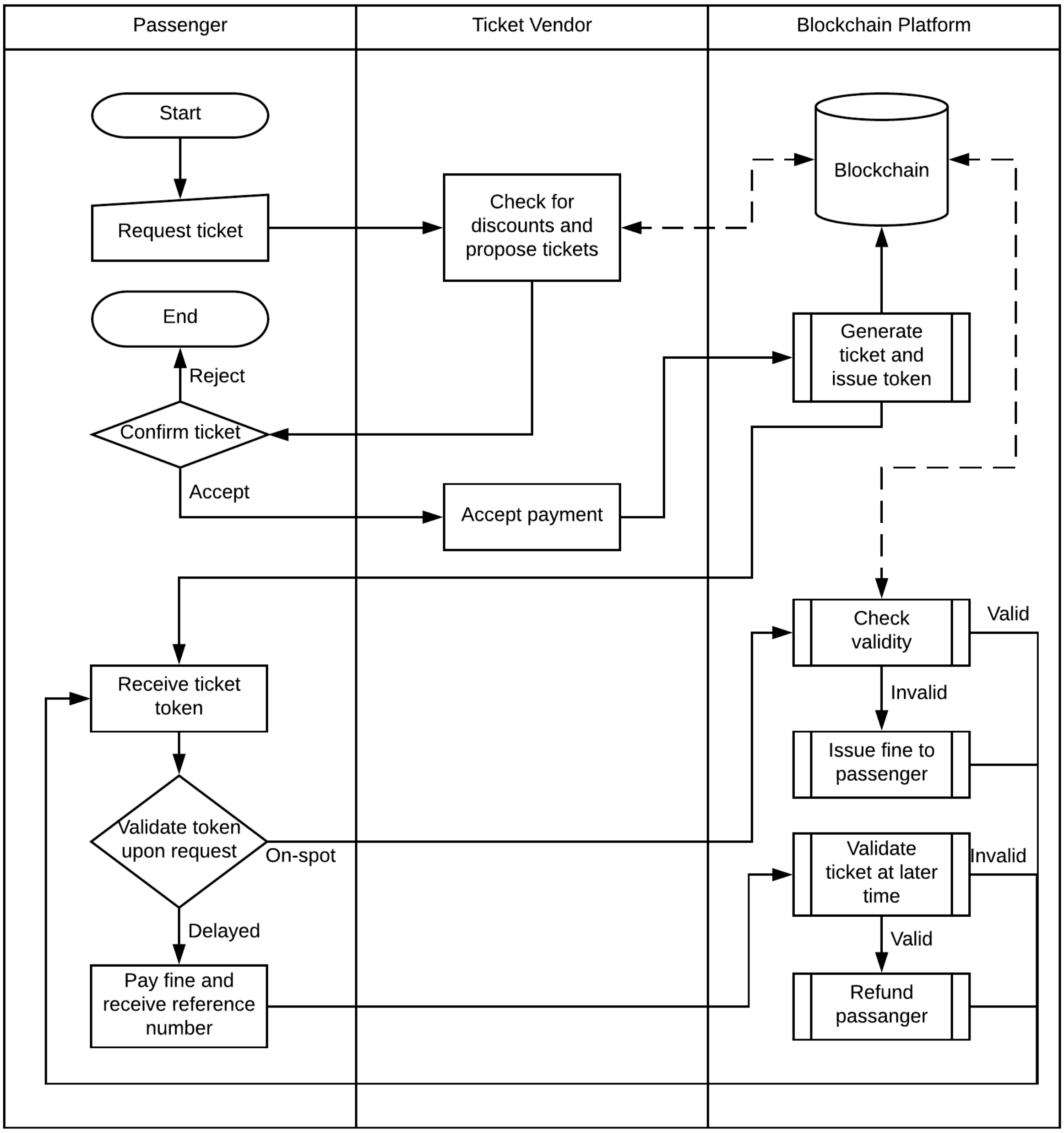

Figure 4: A summary of interactions between the passenger, the ticket vendor, and the blockchain platform.

Request The passenger supplies a vendor with their wallet credentials and the information of the journey they wish to make.

Propose The vendor cross-checks the wallet address provided to apply any discounts, should the passenger be entitled to any. The vendor compiles a suitable ticket and proposes this ticket to the passenger.

Confirm Once confirmed, the transaction is issued. The passenger is able to pay in fiat currency which the vendor receives. A smart contract is initiated which issues a ticket token to the passenger's wallet containing all information of the journey.

Validate When the passenger attempts to pass through station barriers or is requested to show their ticket by a conductor, they must validate their ticket. They send their token to a smart contract (by scanning a QR code or NFC device) which checks the validity of the ticket. The action of sending the ticket confirms ownership as the passenger must digitally sign the transaction. Once the smart contract has validated the ticket, the ticket is returned to passengers wallet.

Delayed Validation If the passenger is unable to validate their ticket on the spot, they may do so at a later time. A fine is issued as normal (in fiat currency) and the passenger receives a reference number which encodes the time of issue. Within a time period of being issued the reference number, the passenger can send their ticket token and reference number to a smart contract which will check that the ticket was valid at the time the passenger was unable to validate their ticket.

Further Considerations. There are additional aspects that must be considered when implementing a blockchain platform for railway ticketing. In some cases, only a particular number of tickets are available to purchase, or a seat reservation may be requested by the passenger. The vendor or the blockchain platform can query existing reservation systems such as LENNON to check that tickets are available to be issued to passengers as tokens. Moreover, there are particular railcards (such as the two-together railcard) that require more than one person to be used. Associating a discount token with two wallets and validating that both railcards are valid for the single journey is an important aspect to consider. Additionally, it should be considered that ticket tokens for other modes of transport may be stored in the same wallet as the train ticket tokens to increase the flexibility and ease-of-use of transport in the UK. A blockchain would provide a scope to achieve this due to the distributed nature of the network. Furthermore, should the passenger have their account information stolen, it is vital that the use of information and tokens stored on their wallet is blocked before the thief can make use of it.

3.1.3. Commercial Considerations

Existing Public Platform. The Ethereum and NEO platforms both offer the functionality to implement the process in Figure 4. Both platforms offer the capability to develop DApps with brand new tokens on the blockchain that can run alongside the native currencies. These tokens can be stored in any native wallet. This means that if a ticketing token were to be created, these tokens could be stored in existing Ethereum and NEO wallets. Should either platform become the default for hosting DApps, railway tickets could be stored alongside other useful tokens in the same wallet.

One of the biggest issues with utilising DApp platforms is the requirement for users to pay with fiat currency. Not all users of a system are going to be willing to use pure Ether or GAS to pay for transactions; it is far more convenient for the majority of people in most situations to use the fiat currency. As such, the ticket vendors would be required to store a supply of the native platform currency. When a passenger confirms and accepts a ticket, they pay in fiat currency. Once the ticket supplier has received this currency, they can use their own supply of the native currency to initiate the smart contract on the blockchain to generate and send a ticket token to the passenger's wallet. When the passenger is asked to validate their ticket, they can be sent the amount of native currency required to initiate the validation contract. These actions prevent the users from having to store their own currency. It will be the role of the ticket suppliers and franchises who validate the tickets to store the cryptocurrency in their own wallets. Though the transaction fees will cost the vendors and operating franchises money, they will have the benefit of not paying for equipment to validate the transactions between wallets, as this is performed across the distributed network by miners.

Another important aspect to consider with DApp platforms is data storage. Though data can be stored within smart contracts on the blockchain, this can become quite expensive, as storing even a byte in a smart contract will cost more gas. Instead, data within the blockchain should remain as metadata. For the tokens, it is crucial to have the relevant data associated with the token as part of the blockchain. However, for a larger data store (such as a list of all generated ticket tokens), it may be beneficial to store it external to the blockchain. One potential solution to maintain distribution of data (in order to avoid silos and centralisation of control) would be to utilise a platform such as InterPlanetary File System (IPFS).

Bespoke Platform. Hyperledger Fabric provides a framework to develop bespoke blockchain solutions. This allows flexibility in the design of the platform, validation of transactions and blocks, and access permissions to specific parties. Furthermore, it adds the option of centralised control, as opposed to the completely decentralised nature of public platforms such as Ethereum and NEO. This approach is likely to appeal to large companies that are familiar with developing bespoke computational platforms, as it removes the element of trust associated with utilising public platforms with intrinsic currencies.

Potential benefits include easier internal data storage, platform flexibility, and improved transaction speeds over public platforms. Nonetheless, there are potential flaws to not using a public platform. One such flaw is the centralisation; other parties wishing to use this blockchain would need explicit permission to access and manipulate the blockchain's data, reducing the transparency that the railway industry requires. Moreover, as the consensus mechanism is internal, this may cost the company more money than leaving the transaction validation to the peer-to-peer networks of the public platforms.

Instead, it is suggested to make a bespoke blockchain platform public to ensure transparency. Wallets can store tokens alongside a native currency that can be used for ticket purchases and refunds. The currency can be obtained by buying some with fiat currency, or as a reward for mining new blocks. Mining can use a modified PoS mechanism where rail passengers (who have opted to host a node) are selected based upon the number of tickets they have in their wallet. Passengers who use the railway more frequently will be more likely to be chosen to mine a block, and thus are more likely to receive the currency reward. This incentivises passengers to help mine transactions on the network and for passengers to use the railway network more frequently.

3.2. Logistics

Another possible area where blockchain technology could be implemented is logistics and supply chain processes. The supply chain within the railway industry concerns the logistics of shipping parts from stakeholder to stakeholder and any repairs that may happen. The particular issue that blockchain technology may address is the distribution of information between a number of stakeholders, especially within the fragmented privatised system of the UK.

The supply chain in the UK can be simplified into a three-tiered hierarchy [12]. This is illustrated in Table 1, where it is demonstrated that all stakeholders from tiers one and two depend on the raw materials manufacturers in tier three. From this table alone, it is possible to envisage the vast quantity of stakeholders involved within the railway supply chain, as multiple companies compete within each stakeholder category. Furthermore, many of the integral assets within the industry, be it infrastructure or rolling stock, will come into contact with various companies from across the hierarchy.

Table 1: The supply chain hierarchy in the UK.

Tier	Stakeholder Category	Responsibility
1	Infrastructure Builders and Contractors Maintenance and Upgrade Companies Vehicle Manufacturers	Builds the railway infrastructure Maintains/upgrades infrastructure and rolling stock Builds the rolling stock
2	Systems Manufacturers Components Manufacturers	Builds systems for the railways Takes raw materials and constructs specific components
3	Raw Materials Manufacturers	Extracts/constructs raw materials for multi-purpose usage

As an example, take the supply chain process journey of a locomotive. The vehicle manufacturer will design and request procurement of components from the components manufacturers. This is not restricted to just one company, especially if different companies specialise in different components. The companies themselves have to procure raw materials from a number of raw materials manufacturers. The request of raw materials is no concern of the vehicle manufacturers, but will still be of great interest to them; they will want to ensure that the best materials are being used for their vehicle. After the vehicle is constructed, it will be sold to a leasing company, who will proceed to lease the locomotive to an operating franchise. During the life of the locomotive, it will be switched between franchises and possibly even leasing companies. It will also go though a number of maintenance checks and repairs. A franchise looking to lease or buy this locomotive will want to know the detailed history of repairs, ownership, and manufacturing. Conventionally, this would involve tracing the history through the relevant stakeholders. Alternatively, blockchain technology could be used as a distributed ledger of all history of the locomotive (and any other assets) to provide transparency, security, and trust between all parties.

One possible blockchain platform that could be used is VeChain. As discussed in Section 2.4.3, VeChain aims to deliver a platform that can store assets on the blockchain as a native token. The asset is then linked by a NFC tag, a RFID tag, or a QR code. When scanned, a smart contract is used to access all of the information about that particular asset. This could come in useful when an operating franchise is looking to buy or lease an older locomotive and wishes to know more about the precise history of parts; where and when they were manufactured, whether they have been repaired/replaced, and how much strain they have undergone since first use. VeChain would simplify this process as the operating franchise could easily check this information instead of soliciting with multiple parties to retrieve and piece together the fragmented information.

VeChain is not the only example of supply chain blockchain platforms; Waltonchain and Ambrosus are two other platforms that follow a similar philosophy. Nonetheless, VeChain is the current market leader for this use case, and as such have the furthest developed platform. Alternatively, a bespoke platform following the same principles of VeChain could be developed in Hyperledger Fabric to provide the flexibilty that bespoke systems offer.

3.3. Data Distribution

In the UK, the railway industry is fragmented into a number of smaller stakeholders, all with varying purposes and philosophies. Naturally, these stakeholders choose to design bespoke IT systems and data storage solutions. Data exists in a vast number of digital formats, which often leads to issues with interoperability between data sets. The railway industry is no exception; the efficacy of information systems in the United Kingdom is inhibited by the use of multiple bespoke systems, incapable of interacting efficiently with one another. These self-contained entities are known as silos, and provide the majority of the challenges that must be overcome regarding data transfer and communication [13, 14]. The RSSB predicted that there are over 100 systems in the United Kingdom alone [15].

Any time that an external party wishes to access data of a particular stakeholder, they must follow the painstaking process of obtaining permission and digital access. This is time-consuming; the stakeholder may hesitate to provide full access to their data should there be any sensitive or classified data, and the process of accessing the data itself may be complex if the IT systems are not interoperable.

Research has been undertaken to address the problem of interoperability between systems. Tutcher et al. introduce the concept of digital ontologies for the railway industry, and describe the Railway Core Ontology (RaCoOn) [13]. Classically, ontology is a branch of metaphysics that concerns entities[7] and the relationships that may exist between them and

[7] According to the Oxford dictionary, an entity is "a thing with distinct and independent existence".

other entities. Though a highly theoretical concept in philosophy, the principles inspired the field of ontology within the subject of information science. In this sense, an ontology can be described as a formal model of nomenclature and definition of types, properties, and relationships of entities within a domain. Gruber describes ontology as “the specification of conceptualisations, used to help programs and humans share knowledge” [16].

Centralised databases remain an open problem. Whilst an ontology resolves the issue of what the data is and how it can be accessed, stakeholders may still be wary about distributing this data to other parties. Furthermore, centralised databases have flaws; they are a sole beacon for adversaries to target, and are prone to power outages. It is here where a blockchain could be used to solve this issue. Decentralised storage utilises the mechanics of blockchain technology whilst providing the foundations to store large quantities of data. Storage in particular is a key aspect, as the blockchain platform must be scalable in order to store the vast data.

In terms of existing platforms, there are currently very few approaches at tackling this use-case, due to the infancy of scaling technologies such as sharding and swarming which require technical development until they are suitable for deployment on any platforms. Bespoke platforms are more likely to exist in this situation, allowing multiple technologies to be integrated into one blockchain; sharding, swarming, and ontologies could all be linked through one blockchain to provide a distributed, semantic, and secure data network that could be shared between the multiple stakeholders within the rail industry.

4. Discussion

Despite the interest that blockchain technology received in 2017, there are numerous aspects of this emerging phenomenon that require scrutinising before deployment into the real world. One significant facet of blockchain platforms (those with an associated currency, such as Ethereum or Bitcoin) is the turbulent nature of the price pairing between fiat currency and the cryptocurrency. The market is notoriously volatile, and has experienced a great number of rapid surges and crashes. This behaviour is difficult to predict as the market is only young, and there are a vast number of variables which seem to affect the price. Economic models and technical analysis utilised for the stock markets are not refined enough to predict the sentiment of a cryptocurrency market. This inevitably leads to issues with adoption in the real world. Fiat currency is relatively stable against other fiat currencies, allowing holders, vendors, and banks alike to trust that a portion of the currency will be of roughly the same value at any point in the foreseeable future. Cryptocurrencies cannot provide this security as of yet; the price may vastly increase or decrease in the same timespan. This is not so much a problem if the cryptocurrency is the only currency used in a system, as all parties can agree on particular prices. However, when fiat currencies are paired with cryptocurrencies within a system, users adopt a more tentative approach, as there is

That is to say, an entity may take the form of a subject or object, either physically or non-physically, and either abstract or concrete.

increased uncertainty. Trust in cryptocurrencies and integrating such a volatile market with day-to-day systems are areas that must be considered in greater depth.

Moreover, blockchain technology is still in its infancy. The genesis block for Bitcoin was mined in 2009, under ten years ago from the time of writing this paper. As such, the market has seen new ideas blossom and fade all within its short lifespan. The genesis block of Ethereum - the first platform to introduce the concept of DApps - was mined in 2015. VeChain - one of the first platforms to approach the supply chain use case - is younger still, with its genesis block mined in 2017. Whilst these platforms have working ideas at their foundation, it seems as though there is still a long way to go in terms of making the platforms scalable and robust enough for proper real-world uses. This is noticeable for two reasons. Firstly, there are currently very few applications on these platforms. The Ethereum DApp market is saturated with games, with over $10,000,000 circulating on the Ethereum network just to run these games. Bitcoin and other purely currency focused platforms have been adopted by some small businesses and online retailers, but remains vacant from larger corporations and banks. Secondly, the road-maps of the platforms indicate that the development teams are working towards integrating new technologies to the existing platforms. Ethereum aims to migrate from a PoW mechanism to a PoS mechanism, as well as adopting sharding to scale the network.

Many compare the initial surge of interest to that of the World Wide Web (WWW) hype in the 1990s. If true, it will take many years of trust before the platforms are fully adopted. In this period, it may be the case that some platforms cease existing, and the stronger platforms survive and prosper through any difficult times. Should this be the case, it will prove to be very difficult to predict which of these platforms will survive and dominate the particular use-case they work for.

Despite these factors, it is important to discuss the potential of blockchain technology within the railway industry. Though adoption may seem a long way off with public platforms, the technologies can be integrated to bespoke systems tailored to the requirements of various areas of the railway industry. This paper has focused on three potential scenarios in particular.

Digital ticketing aligns with blockchain philosophy of shared, secure information and digitisation of particular assets. A simplified process is presented to illustrate how the mechanism may take shape. Beyond the scope of this paper, an investigation will occur into the finer details of this particular case: a decision on which platform is most suitable, or if a new platform must be developed; how to integrate fiat currency and cryptocurrency and the resulting economic effects; how tokens can be stored, validated, and hold encoded information about the valid journey and the passenger; where passengers can store their tokens in order to provide a seamless experience; and any other aspects that are mentioned.

The use of blockchain technology within the supply chain and for distributed data sharing requires investigating further to identify whether a blockchain could improve upon the current systems. However, more time is required for the specific platforms to integrate new technologies to make them at all appropriate.

5. Conclusion

This paper has discussed the idea of integrating blockchain technologies within three components of the railway industry: digital ticketing; logistics; and data interoperability and distribution. A variety of potential approaches for each component has been suggested to further the discussion on whether adopting blockchain technology is a worthwhile decision. Many technical specifications are omitted, as they are beyond the scope of this paper and will be assessed individually in future research.

Nonetheless, it is possible to see the use for a digitised ticketing infrastructure with a blockchain operating at the core. This reflects the philosophy of a number of existing blockchain platforms that use tokens to digitise assets, and utilise cryptographic techniques like digital signatures and asymmetric encryption along with smart contracts to verify ownership. Digital ticketing will be developed either on an existing platform that aligns with the requirements of ownership and validation discussed in Section 3.1 or be designed on an entirely new platform that fulfils all industry standards. Both a logistics platform and data platform are concepts that can operate successfully using blockchain technology, though may not improve on the current standards that railway industry uses. As such, a discussion is required to identify the railway-specific technical requirements for each platform to identify the areas that this technology can expand upon.

It is an exciting time for blockchain technology. The year of 2017 saw a sharp increase in interest, which has ignited interest across the commercial spectrum. The railway industry is no exception, and should seize this opportunity to develop and deliver new technologies to remain at the forefront of commercial innovation.

References

[1] Technical Strategy, Tech. Rep., Network Rail, 2013.

[2] A. Narayanan, J. Bonneau, E. Felten, A. Miller, S. Goldfeder, Bitcoin and Cryptocurrency Technologies, Princeton University Press, 2016.

[3] S. Nakamoto, Bitcoin: A Peer-to-Peer Electronic Cash System, https://bitcoin.org/bitcoin.pdf, Bitcoin Foundation, [Online; accessed 2018-04-13], 2008.

[4] A. Chunbley, K. Moore, J. Khim, Merkle Tree, https://brilliant.org/wiki/merkle-tree/, [Online; accessed 2018-04-12], 2018.

[5] V. Buterin, The Meaning of Decentralization, https://medium.com/@VitalikButerin/the-meaning-of-decentralization-a0c92b76a274, [Online; accessed 2018-04-23], 2018.

[6] J. Ray, On Sharding Blockchains, https://github.com/explore, [Online; accessed 2018-04-16], 2018.

[7] V. Buterin, A Next Generation Smart Contract and Decentralized Application Platform, https://github.com/ethereum/wiki/wiki/white-paper, Ethereum Foundation, [Online; accessed 2018-04-13], 2013.

[8] Ethereum Foundation, Ethereum Project, https://www.ethereum.org/, [Online; accessed 2018-03-01], 2018.

[9] F. Canesin, Y. Xiang, J. Lim, E. Fast, J. Lowenthal, A. Fong, J. Hepkema, NEO White Paper, http://docs.neo.org/en-us/index.html, Neo Foundation, [Online; accessed 2018-04-13], 2013.

[10] VeChain Brief, https://www.vechain.org/#brief, [Online; accessed 2018-04-16], 2018.

[11] Tickets on you mobile, https://www.thetrainline.com/information/mobile-tickets, [Online; accessed 2018-04-16], 2018.

[12] J. Lupton, A. Burrows, N. Takwale, R. Morris, Fast Track to the Future, Tech. Rep., Rail Supply Group, 2017.

[13] J. Tutcher, J. M. Easton, C. Roberts, Enabling Data Integration in the Rail Industry Using RDF and OWL: The RaCoOn Ontology, American Society of Civil Engineers .

[14] J. Tutcher, Development of a Semantic Data Model to Support Data Interoperability in the Rail Industry, Ph.D. thesis, University of Birmingham, 2015.

[15] J. Tutcher, J. Easton, C. Roberts, R. Myall, M. Hargreaves, C. Tiller, Ontology-Based Data Management for the GB Rail Industry, Tech. Rep., Rail Safety and Standards Board, 2013.

[16] T. R. Gruber, A Translation Approach to Portable Ontology Specifications, Knowledge Acquisition .