Modeling and Analysis of Market Blockchain-based Data Trading in NB-IoT Networks

Lam D. Nguyen, Student Member, IEEE, Israel Leyva-Mayorga, Member, IEEE, Amari N. Lewis, Student Member, IEEE, and Petar Popovski, Fellow, IEEE

Abstract—Mobile devices with embedded sensors for data collection and environmental sensing create a basis for a cost-effective approach for data trading. For example, these data can be related to pollution and gas emissions, which can be used to check the compliance with national and international regulations. The current approach for IoT data trading relies on a centralized third-party entity to negotiate between data consumers and data providers, which is inefficient and insecure on a large scale. In comparison, a decentralized approach based on distributed ledger technologies (DLT) enables data trading while ensuring trust, security, and privacy. However, due to the lack of understanding on the communication efficiency between sellers and buyers, there is still a significant gap in benchmarking the data trading protocols in IoT environments. Motivated by this knowledge gap, we introduce a model for DLT-based IoT data trading over the Narrowband Internet of Things (NB-IoT) system, intended to support massive environmental sensing. We characterize the communication efficiency of three basic DLT-based IoT data trading protocols via NB-IoT connectivity in terms of latency and energy consumption. The model and analyses of these protocols provide a benchmark for IoT data trading applications.

Index Terms—Distributed Ledger Technology, Blockchain, Data Trading, Internet of Things (IoT), NB-IoT, Smart Contract, Smart City.

I. INTRODUCTION

In 2025, the volume of sensing data generated by personal IoT devices is expected to reach 79.4 ZB globally [1]. Many attempts have been made to improve and adapt business workflows to exploit the availability of IoT data [2], [3]; among these, IoT data trading is the most popular approach. Various services for trading of IoT data are emerging, connecting various devices and distributed IoT data sources, thereby facilitating data providers to exchange their data [4].

Interesting use cases for data trading include public transport systems, for example, the bus network in Aalborg, Denmark. In these systems, the density of personal travel card swipes at specific bus stations could be useful information, not only to the administration of transport systems, but also to the local taxi companies. The taxi companies benefit from the data of anomalous passenger traffic patterns for the purposes of improving ride-sharing and private services [5]. Also, analyzed traffic data of passengers can be collected via IoT infrastructure and recommendation services to taxi companies can be sold. Besides, drivers can exchange information about the traffic status of a particular street with others to avoid traffic jams or to exchange green house gas emission information with manufactures. Hence, IoT data can be considered as a tradable digital asset.

Traditional trading systems (e.g. Paypal) feature a single point of failure, the lack of trust, transparency, and incentive for data trading, which is preventing the availability of digital information from data providers to customers. On the other hand, Distributed ledger technologies (DLTs) and Blockchain support immutable and transparent information sharing among involved untrusted parties [6]. Outside of its role in financial transactions, DLTs are seen as a key enabler for trusted and reliable distributed monitoring systems. The authentication process for DLTs relies on consensus among multiple nodes in the network [7]. In Blockchain-enabled IoT networks [8], transactions can include sensing data, or monitoring control messages, and these are recorded and synchronized in a distributed manner in all the participants of the system. These participants are called miners or peers and, in some specific DLTs, users are charged a transaction fee to deploy and execute transactions.

In addition, DLTs allow the storage of all transactions into immutable records and every record is distributed across many participants. Thus, security in DLTs comes from the decentralized operation, but also from the use of strong public-key cryptography and cryptographic hashes. The benefits of the integration of DLTs into IoT data trading systems include: i) guarantee of immutability and transparency for environmental sensing data, ii) removal of the need for third parties, iii) development of a transparent system for heterogeneous IoT data trading networks to prevent tampering and injection of fake data from the stakeholders [9].

With the spread of ubiquitous marketplaces, it became relevant to explore the use of IoT data trading in marketplace environments. For instance, in [10], Gupta et al. introduced the architecture for a dynamic decentralized marketplace for trading IoT data. The approach involves a 3-tier design: 1) provider, 2) broker and 3) consumer. The use of DLTs in their work is primarily to manage the terms of agreement between involved parties. Additionally, a reputation system is used in the design to penalize the participants and reduce their rating. Bajoudah et al. present a marketplace model and architecture

1The terms DLT and Blockchain will be used interchangeably throughout this paper. Blockchains are a type of DLT, where chains of blocks are made up of digital pieces of information called transactions and every node maintains a copy of the ledger.
for the trading of IoT streaming data in [11]. Within their work, periodic checkpoints during data exchange are introduced to limit fraudulent activity on either side. In [12], Missier et al. propose another marketplace, where streams of IoT data are the main assets traded utilizing Oracles for the off-chain queries. Xiong et al. [13] present a trading mode based on smart contracts. It incorporates machine learning to guarantee fairness of data exchange and utilizes arbitration institution to deal with the dispute over the data availability in the data trading. However, the arbitration institution in the trading mode is a trusted entity of trading parties.

The related work indicates a gap in terms of: 1) a benchmark for IoT data trading, and 2) analysis of the cost of IoT data trading in terms of communication, specifically in city-level networks. The efficiency of a Blockchain-based data trading protocol is a major concern for data traders. The futures markets are highly dynamic and low latency trading is critical to to maximize the efficiency of the the marketplace. However, currently there is a lack of general framework that provides a guideline for the use of trading protocols based on a set of neutral and commonly accepted rules. A proper benchmark helps the interested parties to understand the tradeoffs in Blockchain-based systems and the associated performance indicators.

In this paper, first, we design a DLT-based trading system for exchanging IoT data. We have chosen the NB-IoT standard [15] as the underlying connectivity solution, as it is seen by the mobile operators as a major candidate to dominate wide-range connectivity for future smart cities. One of the reasons for this is that, unlike other IoT technologies, it is able to offer symmetric uplink/downlink throughput, which is a feature essential from the viewpoint of a DLT [7], [16]. The proposed trading system includes the following IoT data trading protocols: General Trading (GT), Buying on Demand (BoD), and Selling on Demand (SoD). Here we use the term "on demand" from the viewpoint of the smart contracts that implement the transactions between buyers and sellers. Each trading protocol is customized for different scenarios. GT could be considered as the usual trading protocol in data marketplace, while the BoD and SoD are protocols used to support particular demands from either sellers or buyers.

The analysis and simulation results show that the GT protocol has outstanding performance in terms of latency and energy consumption; however, it requires mechanisms to guarantee the continuous availability of data. On the contrary, the BoD protocol can be adapted in Vehicle-to-Infrastructure (V2I) networks, where vehicles can trade their emission information with manufactures. Finally, the SoD protocol is particularly useful when customers are interested in collecting specific data as soon as it is available on the market, for example, in the for of updates. It can also be deployed in Vehicle-to-Vehicle (V2V) networks where the drivers want to buy traffic jam information of a specific street from other vehicles on the road. Clearly, SoD protocols, on their own, would face situations in which the data is no longer available for customers after the initial advertising phase. In practice, the three trading protocols present interesting synergies and can be implemented together in a single system, which will select the best one based on the actual situation.

The contributions of this paper can be stated as follows. First, we present a solution for systematic DLT-based IoT data smart trading towards city-level networks using NB-IoT connectivity. Next, we propose three IoT data trading protocols namely General Trading (GT), Buying on Demand (BoD), and Selling on Demand (SoD). The cost model of each trading protocol is derived and analyzed along with NB-IoT connectivity. Both resources consumed by executing DLT/smart contracts and NB-IoT devices are investigated. Finally, the analysis and the associated experimental results provide a benchmark for data trading protocols in wide-area IoT networks.

The remainder of this paper is organized as follows. In the next section, we outline the general architecture of DLT-based trading system and introduce three IoT data trading protocols. Then, we present the system model, including the physical deployment of the devices. In Section III and IV, we model and analyze the performance of Blockchain-enabled IoT network in terms of latency and energy. Then, we evaluate and prove the derived model and design in Section V. Finally, we conclude the paper in Section VI.

II. DLT-ENABLED IoT DATA SMART TRADING ARCHITECTURE AND PROTOCOLS

This section presents the general system model of DLT-based IoT data trading as well as the data trading system with the three protocols tailored to different scenarios. Table I summarizes the used notation.

A. DLT-enabled Data Trading via NB-IoT

The general architecture of DLT-based IoT data trading includes three main components: data providers (sellers $S$), data consumers (buyers $B$) and a distributed ledger, shown in Fig. 1. Each seller or buyer can own one or more devices in the network. Here we assume that buyers and sellers act as digital wallets in a distributed network. During a trade denoted by $T_j$, the seller $S_i \in S$ and buyer $B_j \in B$ communicate using the wide-area NB-IoT links. The trading procedure occurs to complete a deal between $S_i$ and $B_j$, exchanging data $D_i \in D$.
TABLE I: NOMENCLATURE

| Parameters | Descriptions | Values |
|------------|--------------|--------|
| **General** |              |        |
| N          | Total number of NB-IoT devices | 10000  |
| M          | Number of DLT miners | ≤ 20   |
| S          | Set of Data Providers (Sellers) | | [S] ≤ 10^4 |
| B          | Set of Data Consumers (Buyers) | | [B] ≤ 10^4 |
| D          | Data to buy or sell | –      |
| T          | Set of trades | –      |
| α^u        | Uplink request arrival rate | Eq.3   |
| α^d        | Downlink request arrival rate | Eq.3   |
| G_i, G_r   | Transmitter and receiver antenna gains | 1      |
| β          | Path loss exponential | {2.4, 2.7, 3.0, 3.3} |
| γ          | Receiver sensitivity | 3.6 x 10^{-10} |
| λ_c        | Computing speed of a miner | 0.3    |
| λ_0        | Scaling factor | 0.05   |
| P_i        | Power of miner | 6      |
| τ          | The unit length | 10 ms  |
| d          | The average time interval between two NPDCCCH | | [0.05 : 2] |
| R^u, R^d   | Uplink and Downlink transmission rate | –      |
| **Energy** |              |        |
| E_T        | Energy required to complete a trade T_i | Eq. 7 |
| E^u        | Uplink energy consumption | Eq. 7 |
| E^d        | Downlink energy consumption | Eq. 7 |
| E_{DLT}    | Blockchain energy consumption | Eq. 6 |
| E_{sync}   | Energy required for synchronization | Eq. 11 |
| E^r        | Uplink: Energy for resource reservation | Eq. 11 |
| E^r       | Downlink: Energy for resource reservation | Eq. 11 |
| E^r        | Downlink: Resource reservation energy | Eq. 11 |
| P_L, P_I   | Listening Power | 0.1 W  |
| P_I        | Idle Power | 0.2 W  |
| P_T        | Transmission Power | 0.2 W  |
| P_c        | Power consumption in electronic circuits | 0.01 W |
| **Latency** |              |        |
| L_W        | Average computation latency of a miner | Eq.19 |
| L_{DLT}    | Total DLT latency | Eq.    |
| L_{TM}     | DLT average transmission latency | Eq.18 |
| L_{sync}   | Synchronization Latency | 0.33s  |
| N_{max}    | Maximum number of attempts | 10     |
| P_r        | Probability of resource reservation | Eq.8   |

* Eq: Equation

and payment P_i. First, B_i completes the payment P_i to S_i in reference to the requested data, D_i, and S_i delivers D_i to B_i immediately. The general procedure from Fig. 1 can be described as follows:

1) **Buyer B_i:** Subscribes to the IoT data in distributed ledger generated and published by S_i, and B_i makes a data request, b_i regarding its preferred data, D_i. The b_i will be transmitted to S_i and recorded in the ledger via transaction T_{i,add} for negotiation based on factors such as amount of data, quality of data, price, discount, etc. After choosing D_i from the list, B_i generates a transaction T_{i,commit} which executes payment from B_i’s wallet. Once B_i receives the D_i via T_{i,commit}, it will generate a confirmation back to ledger.

2) **Seller S_i:** Has two main roles: to collect data from the environment (e.g., environmental sensing data, geographical data or data from surveillance systems) and to act as a hub gathering data from neighboring devices to sell on the market. S_i aims to earn the payment P_i from B_i by delivering D_i to B_i. After publishing a hashed version of its data and prices to the market via T_{i,add}, S_i waits for buying requests. Based on the predefined rules in the smart contract system, upon receiving a request from B_i and the appearance of T_{i,commit}, generated by B_i, as being recorded in the ledger, the seller S_i can receive the payment P_i. Finally, it confirms to the ledger that the trade T_i is completed.

3) **Distributed Ledger:** The DLT manages a distributed ledger to record all data trading history which is grouped into blocks and linked together chronologically. The deployed smart contracts autonomously control the order and automate payments from parties without the need of human interaction. The smart contracts guarantee trust, transparency and speed of exchanging information. These can be deployed based on the negotiation between data providers and customers via T_{i,deploy}. Any change in smart contract (e.g. change of price, amount of data, or discount) can be performed via T_{i,update}.

In order to minimize the cost of storage, the sensing data could be hashed and recorded at more powerful DLT nodes, and only the hash of data is recorded to ledger. Then, a message is sent back to confirm that the data has been added to the ledger. After both S_i and B_i are satisfied with the terms of the contract, T_{i,settle} is executed to get the payment P_i from B_i to transfer to S_i’s wallet, while the data D_i is transmitted to the storage address of B_i. We assume that the data services, (e.g., data storage, trading and task dispatching) are implemented on top of a permissionless Blockchain. The sensing data are formatted into normal transactions of fixed size. To enhance efficiency, only the digest of each transaction is stored on the chain, and the content of the transactions are stored by each consensus node off-chain or at the IPFS (InterPlanetary File System) storage.

### B. IoT Data Trading Protocols

1) **General Trading (GT):** The GT protocol procedure is shown in Fig. 2. In a trade T_i, the buyer B_i sends a buying request r_i to market via transaction T_{i,add} to express its interest in specific data D_i. The data producer and seller S_i, after collecting sensing data, start publishing its data information D_i to the market. The smart contracts receive the requests from both customers B and producers S and then map the buying requests and selling requests to satisfy both parties based on their expected data and price. The buyer B_i omits the requests with fund transfer via T_{i,commit}. After the smart contract receives the payment from B_i, it executes T_{i,settle} to transfer requested D_i to B_i and P_i to S_i. Finally, both B_i and S_i confirm to the ledger that they received P_i and D_i, respectively.

A marketplace exchange of streaming IoT data, with a massive amount of data, requests, and a large number of parties, is an appropriate use case for the GT protocol. The environmental sensing data such as accurate real-time measurement data for control and alarm systems are exchanged between interested customers. More specifically, this protocol is used for the aforementioned use case due to its wide range of data continuously being pushed to the market. The open advertisement style of the GT protocol is appealing to potential buyers, encouraging the safe buying and selling of IoT data in a decentralized IoT data marketplace.
2) Buying on Demand (BoD): BoD protocol describes a process where the producer $S_i$ publishes data information $D_i$ to the market for selling via $b_i$ request. The smart contract will broadcast information of received data by buying offer $b_o$ to other parties, e.g. $B_i$ to ask whether they are interested in buying $D_i$. If $B_i$ is interested in $D_i$, it will accept the offer by generate $T_{i,\text{addr}}$, and commit by $T_{i,\text{commit}}$ when get the payment requests from smart contract. Then, the deal is settled as GT protocol via $T_{i,\text{settle}}$. The process of BoD protocol is described in Fig. 2b.

Vehicle-to-Industry (V2I) emission trading, with a frequent exchange of data between vehicles and the vehicle industry, is an appropriate use case for the BoD protocol. In this scenario, the vehicles on the network act as the sellers of their emissions data, e.g., CO$_2$, NOx, while manufacturers (vehicle industry), GoV e.g air quality management department, and data analytic organizations act as the buyers for maintaining accurate, secure tamper-proof vehicular emissions data. In V2I, the data being exchanged are used for the purpose of creating a trusted life-cycle emissions or fuel economy monitoring.

3) Selling on Demand (SoD): The SoD protocol is described in Fig. 2c In this case, the smart contract receives the buying requests $b_i$ from a customer $B_i$, but there is no available appropriate data on the ledger to satisfy the requirements from $B_i$. Hence, the smart contract sends an ask-for-data request $aD_i$ to producers to ask whether they can provide the required data $D_i$. The providers $S_i$ after a while can gather data from the environment or from other sources then answer to the market by $s_i$ including $D_i$ information as well as price $P_i$. Then, the smart contract asks $B_i$ for fund transfer with amount of $P_i$. The $B_i$ make payment via $T_{i,\text{commit}}$. Then $T_{i,\text{settle}}$ are executed to complete the deal between $B_i$ and $S_i$. Finally, the confirmations are sent to the ledger from both parties.

This protocol is beneficial, for example, when a party needs a type of data that is not available on the market and there is a need to trade in real time. In the scope of this study, we assume that, when a provider receives the request ask-for-data from smart contract, it can provide the required data to the market. In real-life scenario, the requests from customer sometimes cannot be satisfied immediately, so these requests are queued in the systems until completed. A Vehicle-to-Vehicle (V2V) use case is appropriate for this protocol where vehicles can purchase the information on the traffic jam in specific street which drivers expect to use in the near future. The vehicles that have the requested information can be traded with the buyers on the road. Finally, similar to V2I, V2V involves the continuous wireless exchange of IoT data collected from vehicle sensors. The V2V use case contributes in generating a life-cycle emissions or fuel economy monitoring system amongst vehicles. This form of communication helps to manage the safety of the road, as well as increase vehicle awareness.

C. Communication System Model

We consider an NB-IoT cell with eNB located in its center, including $N$ devices uniformly distributed within the area. A data provider or consumer can consist of a single or multiple NB-IoT devices. For simplicity, we assume that each buyer or seller owns a single NB-IoT device to exchange assets and all devices belong to the normal coverage class. The DLT nodes are NB-IoT devices that have more computational power than seller/buyer nodes. In our model, involved sellers and buyers use NB-IoT as wireless network interfaces. In reality,
the involved parties can use various wireless interfaces or networks for trading purposes but, our general model and analysis can be applied in these cases because of its modular and versatile design satisfies a broad range of interfaces and networks.

Our propagation model takes into account shadowing, but not small-scale fading; which is a sufficient first approximation as detailed physical layer modeling is not the focus of the work. Hence, for a given transmission power $P_t$ and carrier frequency $f$, the received power at a distance $d$ between the base station $BS$ and sensor $i$ is:

$$P_r(d) = 10 \log_{10} \left( \frac{P_t G_t G_r c^2}{(4\pi f)^2 d^2} \right) + N(0, \sigma_{dB}) \text{ dB} \quad (1)$$

where $G_t$ and $G_r$ are the transmitter and receiver antenna gains, respectively, $c = 3 \cdot 10^8 \text{ m/s}$ is the speed of light, $N(0, \sigma_{dB})$ is a zero-mean Gaussian RV with standard deviation $\sigma_{dB} \text{ dB}$, and $\beta$ is the path loss exponent. From there, the outage probability at a given distance and receiver sensitivity $\gamma = 3.65 \cdot 10^{-10} \text{ W}$ is:

$$p_{out} = 1 - Q \left( \frac{1}{\sigma_{dB}} 10 \log_{10} \left( \frac{\gamma (4\pi f)^2 d^2}{P_t G_t G_r c^2} \right) \right) \quad (2)$$

Fig. 3 demonstrates the delivery probability $p_d = 1 - p_{out}$ at varying distances, for four different $\beta$ path loss exponent values, a standard deviation of $\sigma_{dB} = 6 \text{ dBs}$. In this work, we choose $\beta = 2.7$ for urban area. We are aware that the model lacks a mobility aspect, however for this initial work, we have decided to use a simple model as previously described.

The arrival rate of uplink including selling and buying requests, respectively, to the system are: $\lambda_s = |S|T_p p_s p_d$ and $\lambda_b = |B|T_p p_b p_d$, in which $T$ is number of communication sessions that an IoT device performs daily; $p_s$ and $p_b$ are probability a device request a selling service and a buying service, respectively. When an NB-IoT sensor device attempts to join the network, it first listens for the cell information, e.g. NPSS and NSSS messages to synchronize with the eNB. NB-IoT UEs have only two modes of operation, namely radio-resource control (RRC) idle and RRC connected [17]. In the former, the UEs can only receive the system information from the BS and, only in the latter, data can be transmitted. UEs are in idle mode before initial access to the network, but may also enter this mode during power saving or after an explicit disconnection request. To transition from idle to connected mode, the UEs (clients) must first acquire the basic system information and synchronization as illustrated in the upper part of Fig. 4. For this, the UE receives the master information block (MIB-NB) and the system information blocks 1 (SIB1-NB) and 2 (SIB2-NB). These are transmitted periodically through the downlink shared channel (DL-SCH) and carry the basic cell configuration, timing, and access parameters [18]. In addition, SIB1-NB carries the scheduling information for the rest of the SIBs.

After the system information has been acquired, the UEs must perform the RA procedure to transition to RRC connected mode. The RA procedure is a four-message handshake, initiated by the UEs by transmitting a single-tone frequency-hopping pattern, called preamble, through the NB Physical Random Access Channel (NPRArch). In most cases, the RA procedure is contention-based, hence, the preamble is chosen randomly from a predefined pool of up to 48 orthogonal sub-carrier frequencies. Consequently, the main reasons for an access failure are the lack of power in the transmission and simultaneous transmissions of the same preamble, which lead to collisions.

After completing the RA procedure, and if the control-plane (CP) cellular IoT (CIoT) is used, UEs may piggyback short UL data packets along with the RRC Connection Setup Complete message. Otherwise, the non-access stratum (NAS) setup must be completed before eNB allocates resources for uplink transmission through the NB Physical UL Shared Channel (NPUSCH) and data can be transmitted. The resource unit (RU) is the basic unit for resource allocation in the NPUSCH and comprises a set of sub-frames in the time domain and sub-carriers in the frequency domain. The downlink (DL) data is transmitted through the NB physical DL shared channel (PDSCH). Data is exchanged based on the three defined trading protocols, GT, BoD and SoD. Fig. 4 shows the physical operations of GT protocol as an example. BoD and SoD could be considered as extensions of the GT protocol, those protocols are especially beneficial when the data is not available in the market.

D. Performance metrics

Latency and the time required to complete a trade is one of the most important concerns of involved users. Latency
directly influences the amount of time it takes for a trader to interact with the data market, the timely reception of relevant market information and the ability to act upon its receipt. The spread of the automatized data trading amplifies the impact of latency in terms of its competitive advantage. On top of this, IoT environments should be characterized with high energy efficiency. All these factors have motivated this investigation on the total E2E latency and energy consumption to complete a trade \( \mathcal{T} \).

1) Latency: The latency to complete a trade \( \mathcal{T}_i \) between seller and buyer are formulated as:

\[
L_{\mathcal{T}_i} = L_{UD} + L_{DLT},
\]

where \( L_{UD} \) is the transmission latency between \( S_i \) and \( B_i \), which act as light nodes and full DLT nodes; While, \( L_{DLT} \) represents the DLT mining and synchronization latency. In detail, \( L_{UD} = L_u^u + L_d^d \), where \( L_u^u, L_d^d \) are NB-IoT uplink and downlink latency, respectively; \( L_{DLT} = L_o + L_{DLTsync} \), where \( L_o \) is block verification time at DLT nodes, and \( L_{DLTsync} \) is synchronization time between DLT nodes via NB-IoT connectivity.

2) Total energy consumption: Similarly, the energy consumption model of a trade includes the energy consumption for uplink \( E_u^u \), downlink \( E_d^d \) transmission between NB-IoT sensors with DLT full nodes, among DLT full nodes, and the energy consumed in verification process known as mining in DLT nodes \( E_{DLT} \).

\[
E_{\mathcal{T}_i} = E_{UD} + E_{DLT}
\]

where \( E_{UD} \) and \( E_{DLT} \) are energy consumed by communication between sellers/buyers and DLT nodes and the energy performed among full DLT nodes, respectively. The transmission power and latency depend significantly on the physical deployment, such that we analyze both analyze the resource consumed in physical communication and the application layer. In next parts, we formulate the latency and energy consumption of each process.

III. NB-IoT Transmission Latency and Energy Consumption Models

As described in previous section, the total E2E latency includes two parts, the latency of transmissions of uplink and downlink between buyers/sellers and DLT nodes, and latency occurs in DLT verification process. For the first part, we define an adapted queuing model for DLT-based NB-IoT based on queuing model of the NB-IoT access network \([19]\), where the uplink and downlink radio resources are modeled as two servers which visit and serve their traffic queues in both directions.

1) End-to-End latency: The E2E latency of NB-IoT uplink and downlink can be formulated as:

\[
L_{UeD} = L_u^u + L_d^d
= L_u^{sync} + L_{rr}^u + L_{tx}^u + L_{sync}^u + L_{rr}^d + L_{tx}^d
\]

where \( L_u^{sync}, L_u^{rr}, L_u^{tx}, L_{sync}^u, L_{rr}^d, L_{tx}^d \) are energy consumption of synchronization, resource reservation, and data transmission of uplink and downlink, respectively. \( L_{sync} \) has been defined in \([18]\) with the values of 0.33s. \( L_{rr} \) is given as:

\[
L_{rr} = \sum_{l=1}^{N_{r_{max}}} (1 - P_{rr})^{l-1} P_{rr} l (L_{rr} + L_{rar})
\]

in which \( N_{r_{max}} \) is the maximum number of attempts, \( P_{rr} \) is the probability of successful resource reservation in an attempt, \( L_{rr} = 0.5t + \tau \), is the expected latency in sending an RA control message, \( \tau \) is the unit length and equal to the NPRACH period for the coverage class 1 which is varied from 40 ms to 2.56 s \([18]\), and \( L_{rar} = 0.5d + 0.5Qf_u + u \), is the expected latency in receiving the RAR message, where \( Q \) is requests waiting to be served.

In the following, we provide a simple technique based on drift approximation \([20]\) to calculate \( P_{rr} \) recursively. Therefore, we treat the mean of the random variables involved in the process as constants. Besides, we assume that sufficient resources are available in the PDCCH so that failures only occur due to collisions in the PRACH or to link outages.

Let \( \lambda^a = \lambda^u + \lambda^d \) be the arrival rate of access requests per PRACH period and \( \lambda^a(l) \) be the mean number of devices participating in the contention with their \( l \)-th attempt. Note that in a steady state \( \lambda^a(l) \) remains constant for all PRACH periods. Next, let \( \lambda^a_{tot} = \sum_{l=1}^{N_{r_{max}}} \lambda^a(l) \) and that the collision probability in the PRACH can be calculated using the drift approximation for a given value of \( \lambda^a_{tot} \) and for a given number of available preambles \( K \) as:

\[
P_{\text{collision}}(\lambda^a_{tot}) = 1 - \left(1 - \frac{1}{K}\right)^{\lambda^a_{tot} - 1} \approx 1 - e^{-\frac{\lambda^a_{tot}}{K}}.
\]

From there, we approximate the probability of resource reservation as a function of \( \lambda^a_{tot} \) as:

\[
P_{rr}(\lambda^a_{tot}) \approx p_d e^{-\frac{\lambda^a_{tot}}{K}}.
\]

This allows us to define \( \lambda^a_{tot} \) as:

\[
\lambda^a_{tot} = \lambda^a + \left(1 - P_{rr}(\lambda^a_{tot})\right) \sum_{l=2}^{N_{r_{max}}} \lambda^a(l)
\]

since \( \lambda^a(l) = (1 - P_{rr}(\lambda^a_{tot})) \lambda^a(l-1) \) for \( l \geq 2 \) and \( \lambda^a(1) = \lambda^a \). Finally, from the initial conditions \( \lambda^a(l) = 0 \) for \( l \geq 2 \), the values of \( \lambda^a(l) \) and \( \lambda^a_{tot} \) can be calculated recursively by: 1) applying \([9]\), 2) calculating \( P_{rr}(\lambda^a_{tot}) \) for the new value of \( \lambda^a_{tot} \); and 3) updating the values of \( \lambda^a(l) \). This process is repeated until the values of the variables converge to a constant value. The final value of \( P_{rr}(\lambda^a_{tot}) \) is simply denoted as \( P_{rr} \) and used throughout the rest of the paper.

Assuming that the transmission time for the uplink transmissions follows a general distribution with the first two moments \( l_1, l_2 \), then first two moments of the distribution of the packet transmission time are \( s_1 = (f_1l_1) / (R\omega) \), and \( s_2 = (f_1l_2) / (R^2\omega^2) \). Applying the results from \([21]\), considering \( L_{tx} \) as a function of scheduling of NPUSCH, we have:

\[
L_{tx} = \frac{f \lambda^u s_1 s_2}{2s_1(1 - fGs_1)} + \frac{f \lambda^u s_1^2}{2(1 - fGs_1)} + \frac{l_1}{R\omega}
\]
task for the POW of current block, the time period required to complete this PoW can be formulated as an exponential random variable $W$ whose distribution is $f_W(w) = \lambda_c e^{-\lambda_c w}$, in which $\lambda_c = \lambda_0 P_c$ presents for the computing speed of a miner, $P_c$ is power consumption for computation of a miner, and $\lambda_0$ is a constant scaling factor. Once a miner completes its PoW, it will broadcast messages to other miners, so that other miners can stop their PoW and synchronize the new block.

$$L_{tM} = L_{newB} + L_{getB} + L_{transB}$$  \hspace{1cm} (18)

where $L_{newB}$, $L_{getB}$, and $L_{transB}$ are latencies of sending hash of new mined block, requesting new block from neighboring nodes, and new block transmission, respectively. $L_{newB}$ and $L_{transB}$ are computed using as uplink transmission, while $L_{getB}$ is computed based on downlink transmission as described in previous section.

For the PoW computation, a miner $i$s, first finds out the desired PoW hash value, $i$s = min$_{i\in M}$ $W_i$. The fastest PoW computation among miners is $W_i$s, the complementary cumulative probability distribution of $W_i$s could be computed as $Pr(W_i < x) = Pr(i\in M(W_i) > x) = \prod_{i=1}^{H} Pr(W_i > x) = (1 - Pr(W < x))^M$. Hence, the average computational latency of miner $i$s is described as:

$$L_{W,i} = \int_{0}^{\infty} (1 - Pr(W \leq x))^M dx = \int_{0}^{\infty} e^{-\lambda_c M x} dx = \frac{1}{\lambda_c M}$$  \hspace{1cm} (19)

The total latency required from DLT verification process is $L_{DLT} = L_{tm} + L_{W,i}$. The average energy consumption of DLT to finish a single PoW round is:

$$E_{DLT} = P_L L_{W,i} + P_L L_{tm}$$  \hspace{1cm} (20)

The performance of DLT system is shown in Fig. 5. The figure demonstrates that the energy consumed and latency by DLT nodes are reduced with the number of miners. As the number of miners increases, this leads to a higher probability that miners verify transactions, and the mining speeds increases.

\section*{A. Analysis of data trading protocols}

In this section, the E2E latency and energy consumption of three protocols are formulated and compared approximately. The resource consumed by each data trading protocol is separated into two parts, namely, 1) the connectivity between $S_i$ and $B_i$ acting as light nodes in DLT network with full nodes, and 2) the communication among DLT full nodes.

The E2E latency of trade $T_i$ using $GT$ protocol including the transmission latency between $B_i$, $S_i$ and DLT verification nodes is described as below:

$$L_{T_i}^{P1} = L_{UD}^{P1} + L_{DLT}^{P1} + L_{ud}^{P1} + L_{d}^{P1} + L_{DLT}^{ud}$$  \hspace{1cm} (21)

$$E_{T_i}^{P1} = E_{UD}^{P1} + E_{DLT}^{P1} + E_{ud}^{P1} + E_{d}^{P1} + E_{DLT}^{ud}$$  \hspace{1cm} (22)

Assuming that $L_{ud}^{P1}$ = $L_{ud}^{sync}$ = 0.33 s, $L_{ud}^{P1}$ and $L_{d}^{P1}$ are computed as (7), $L_{ud}^{P1}$ and $L_{d}^{P1}$ are calculated based on (8) and (9) with the defined packet length of uplink and downlink.

Then, the battery lifetime of a NB-IoT device can be computed as below: $BTL = E_0 \left[ T_p (E_u) + T_p (E_d)^{-1} \right]$, where $E_0$ is the energy storage at device battery. Similarly, the performance of $BoD$ protocol and $SoD$ protocol can be formulated as $GT$ protocol.
TABLE II: Comparison in Smart Contract Execution Cost

| Protocols | Operations | Ether | Gas Cost | Approx. USD |
|-----------|------------|-------|----------|-------------|
| GT        | Deploy SC  | $1.2 \times 10^{-4}$ | 1132443 | 0.2862      |
| BoD       | Deploy SC  | $1.3 \times 10^{-4}$ | 1268369 | 0.2879      |
| SoD       | Deploy SC  | $1.5 \times 10^{-4}$ | 1582349 | 0.3783      |

* 1 Ether = 10^9 Gwei; 1 USD = 4,182,471.9949 Gwei

V. PERFORMANCE EVALUATION

In this section, we will introduce the settings in terms of simulations and experiments. Then, we analyze the performance of proposed trading protocols in terms of latency and battery lifetime.

A. Settings

In this section, we evaluate the derived data trading model, compare and analyze the designed trading protocols. In order to evaluate the derived model and compare the three proposed protocols, we setup a network with $N = 10000$ NB-IoT devices, where devices randomly play roles as sellers or buyers. We validate the results via Monte Carlo Simulations, where we run 1000 realizations for each trading protocol and experiments. The buyer nodes and seller nodes generate requests following a Poisson distribution process with rates of $\lambda_b$ and $\lambda_s$, respectively. The number of buying and selling requests per day varied from 1 to 20, $T_i = [1, 20]$. The number of buyer and seller nodes varied and less than $N$. The transmission power $P_t = 0.2W$, $E_0 = 1000$. The number of DLT miners are up to 20 miners as maximum, $M = [1, 20]$.

B. Cost of Smart Contracts

The proof of concept for the three proposed trading protocols are deployed in Ganache Ethereum network to evaluate the complexity and the cost of execution of different trading strategies. The smart contracts are implemented and deployed using Remix IDE. In the Ethereum platform, any operation or transaction executions that changes the Blockchain or its state requires fees and the involved parties need to pay that fees. These costs are calculated by using the amount of gas executed and the unit of gas price. The gas required during any activity reflects the computational complexity or size of the smart contracts, while the gas prices are determined by the Ethereum miners in the network. In this work, we used Gwei to evaluate the cost of different operations in trading process.

Table II shows the cost of three protocols deployment and transactions to complete a deal between a seller and a buyer. We observe that the approximate cost in USD for GT is the cheapest in comparison to BoD and SoD protocols. The cost of smart contracts execution is generally expensive, therefore, it is preferred to use the GT protocol. In an environment with a massive number of involved parties and transactions (e.g., marketplace), the transactions are executed autonomously to reduce costs using the available resources. While, BoD and SoD are preferred when the users have requests with specific resources.

C. Latency to complete a deal

1) Impact of number of Miners: Fig. 6 shows the latency of three trading protocols. Both the analysis and the simulation results show that the SoD protocol has higher latency to complete a deal between $S_i$ and $B_i$ because of extra steps. Note that the comparison is evaluated approximately because the latencies are dependent on various factor such as number of DLT miners, the length of blocks, and level of difficulty.

The verification latency of DLT miners are measured based on Ganache Ethereum network. In GT protocol, the smart contracts map selling requests $r_i$ with available $D_i$ stored in the ledger and make a deal between $S_i$ and $B_i$ immediately, so that it guarantees efficient trading in the market. The average latency to complete a deal of GT protocol is around 4.5 seconds including latency of NB-IoT and DLT procedures. The BoD and SoD latencies are higher because of extra procedures necessary to gather required information between customers and producers. We observe that GT could be used in terms of applications which require low latency. The downside of GT protocol is that the data requests have to always be available to settle the trade, so that it is matched with applications (e.g. smart metering) where the type of information is fixed.

2) Seller and Buyer Ratio: The impact of ratio between the number of sellers and buyers is demonstrated in Fig. 7. The figure also shows a comparison between trading protocols under varying NB-IoT uplink transmission rate.
The number of uplink requests are varied from 1 to 20 requests per day. We observe that the number of requests daily up to 5, the battery lifetime reduced significantly to around 2 years. Because for each buying or selling request, the NB-IoT devices start running protocol with multiple operation until the trade is settled.

D. Battery lifetime of NB-IoT devices

In general, the power consumption of battery lifetime during a reporting period depends on length of data transmitting, bandwidth, MCL, latency, and RF module. Hence, the power consumption of one trading protocol will be higher or lower than the other depending on the values of these parameters. The battery lifetime capabilities of NB-IoT devices among three trading protocols are compared and demonstrated in Fig. 8. The number of uplink requests are varied from 1 to 20 requests per day. We observe that the number of requests per day significantly impact to the battery lifetime of NB-IoT devices. The fact is that the battery lifetime of around 10 years can be achieve with one report per day, however, for more frequent transmissions (e.g. 8 requests per day) the battery lifetime is reduced to around 1 year. Specifically, the GT trading protocol achieves over 11 years for 1 report per battery lifetime is reduced to around 1 year. Specifically, the BoD and SoD protocols can be interchangeably used when there are particular demands from either buyers or sellers.

To the best of our knowledge, this is the first work of its kind, providing a general benchmark framework for data trading protocols in IoT environments. In the next iteration of this work, we will first consider more elaborate utility models for the parties involved in trading. Second, we will evaluate the performance of trading schemes in diverse network interfaces.

VI. CONCLUSION

In this paper, we proposed the first benchmarking framework for evaluating data trading protocols. The framework includes a model and analysis of systematic DLT-based IoT data smart trading protocols in massive NB-IoT deployments. We have proposed and analyzed three IoT data trading protocols named General Trading, Buying on Demand, and Selling on Demand. Considered collectively, these protocols cover a wide range of interesting scenarios, such as carbon emission trading or monitoring of vehicle emissions. We have conducted a comprehensive analysis of these protocols in terms of communication and evaluated end-to-end latency, battery lifetime and resource consumption. In terms of performance, each protocol is tailored to a different scenario. We conclude that the GT protocol should be used as primary protocol in a data marketplace where massive amounts of data are available. Additionally, the BoD and SoD protocols can be interchangeably used when there are particular demands from either buyers or sellers.

ACKNOWLEDGMENT

This work has been in part supported by the European Research Council (Horizon 2020 ERC Consolidator Grant Nr. 648382 WILLOW).

REFERENCES

[1] ”The Growth in Connected IoT Devices Is Expected to Generate 79.4ZB of Data in 2025, According to a New IDC Forecast”, IDC report, 2019. [Online]. Available: https://www.idc.com/getdoc.jsp?containerId=prUS45213219. Accessed: July 9, 2020.
[2] Huang, Zhiqing, Xiongye Su, Yanxin Zhang, Changxue Shi, Hanchen Zhang, and Luyang Xie. “A decentralized solution for IoT data trusted exchange based-on Blockchain.” In 2017 3rd IEEE International Conference on Computer and Communications (ICCC), pp. 1180-1184. IEEE, 2017.
[3] Perera, Charith. “Sensing as a service (S2aaS): Buying and selling IoT data.” arXiv preprint arXiv:1702.02380 (2017).
[4] Mao, Weichao, Zhenzhe Zheng, and Fan Wu. "Pricing for revenue maximization in iot data markets: An information design perspective." In IEEE INFOCOM 2019-IEEE Conference on Computer Communications, pp. 1837-1845. IEEE, 2019.
[5] Jernigan, Stephanie, David Kiron, and Sam Ransbotham. “Data sharing and analytics are driving success with iot.” MIT Sloan Management Review 58, no. 1 (2016).
[6] Nakamoto, Satoshi. Bitcoin: A peer-to-peer electronic cash system. Manubot, 2019.
[7] Danzi, Pietro, Anders E. Kalor, Rene B. Sorensen, Alexander K. Hagelskjær, Lam D. Nguyen, Cedomir Stefanovic, and Petar Popovski. “Communication aspects of the integration of wireless iot devices with distributed ledger technology.” IEEE Network 34, no. 1 (2020): 47-53.
[8] Lam D. Nguyen, . "Witness-based Approach for Scaling Distributed Ledgers to Massive IoT Scenarios." arXiv preprint arXiv:2004.06300 (2020).
[9] L. D. Nguyen, A. E. Kalor, I. Leyva-Mayorga and P. Popovski, "Trusted Wireless Monitoring Based on Distributed Ledgers over NB-IoT Connectivity", in IEEE Communications Magazine, vol. 58, no. 6, pp. 77-83, June 2020, doi: 10.1109/MCOM.001.2000116.
[10] Gupta, P., Katherere, S., & Jurdak, R. (2019). A Decentralized IoT Data Marketplace. arXiv preprint arXiv:1906.01799.
[11] Bajouah, S., Dong, C., & Missier, P. (2019, July). Toward a Decentralized, Trust-less Marketplace for Brokered IoT Data Trading using Blockchain. In 2019 IEEE International Conference on Blockchain (Blockchain) (pp. 339-346). IEEE.
[12] Missier, F., Bajouah, S., Caposelle, A., Gaglione, A., & Nati, M. (2017, October). Mind my value: a decentralized infrastructure for fair and trusted iot data trading. In Proceedings of the Seventh International Conference on the Internet of Things (pp. 1-8).
[13] Wei, X., and Li, X., “Smart Contract Based Data Trading Mode Using Blockchain and Machine Learning,” in IEEE Access, vol. 7, pp. 102331-102344, 2019, doi: 10.1109/ACCESS.2019.2928325.
[14] Wei, D., Chankai, D., Kim, Khawng R. C., Changze, C., Deiqing, Z., and Hai, J., “SDTE: A Secure Blockchain-Based Data Trading Ecosystem,” in IEEE Transactions on Information Forensics and Security, vol. 15, pp. 725-737, 2020, doi: 10.1109/TIFS.2019.2928256.
[15] Popli, Sakshi, Rakesh Kumar Jha, and Sanjeev Jain. "A survey on energy efficient narrowband internet of things (NBIoT): architecture, application and challenges." IEEE Access 7 (2018): 16739-16776.

[16] 3GPP, “Physical layer procedures,” TS 36.213 V15.5.0, Mar. 2019.

[17] 3GPP, “Radio Resource Control (RRC); Protocol specification,” TS 36.331 V15.0.0, Sept. 2018.

[18] “Technical specification group GSM/EDGE radio access network; cellular system support for ultra-low complexity and low throughput Internet of Things (CIoT),” 3GPP, Sophia Antipolis, France, Rep. TR 45.820, 2015.

[19] Azari, A., Stefanović, Č., Popovski, P. and Cavdar, C., 2019. On the Latency-Energy Performance of NB-IoT Systems in Providing Wide-Area IoT Connectivity. IEEE Transactions on Green Communications and Networking.

[20] C.-H. Wei, G. Bianchi, and R.-G. Cheng, “Modeling and Analysis of Random Access Channels With Bursty Arrivals in OFDMA Wireless Networks,” IEEE Transactions on Wireless Communications, vol. 14, no. 4, pp. 1940–1953, 2015.

[21] H. Akimaru and K. Kawashima, Teletraffic: theory and applications. Springer Science and Business Media, 2012.