Drugledger: A Practical Blockchain System for Drug Traceability and Regulation

Yan Huang, Jing Wu, Chengnian Long
Department of Automation, School of Electronic Information and Electrical Engineering,
Shanghai Jiao Tong University and Key Laboratory of System Control and Information Processing,
Ministry of Education of China, Shanghai 200240, China
longcn@sjtu.edu.cn

***Abstract*—Drug traceability system is essentially important for public drug security and business of pharmaceutical companies, which aims to track or trace where the drug has been and where it has gone along the drug supply chain. Traditional centralized server-client technical solutions have been far from satisfying for their bad performances in data authenticity, privacy, system resilience and flexibility.**

In this paper, we propose a scenario-oriented blockchain system for drug traceability and regulation called Drugledger, which reconstructs the whole service architecture by separating service provider into three independent service components and ensures the authenticity and privacy of traceability data. Drugledger is more resilient than traditional solutions with its p2p architecture. Furthermore, Drugledger could efficiently prune its storage, achieving a finally stable and acceptable blockchain storage. Besides, algorithms reflecting the real drug supply chain logic (e.g. package, repackage, unpackage, etc.) are designed based on the expanded UTXO workflow in Drugledger. To our knowledge, it is the first systematic work from both a technical and practical perspective on how blockchain system could be designed for drug traceability and regulation.

I. Introduction

Drug traceability is essentially important for patients' health, business operation, and government regulation. With a reliable drug traceability system, patients and stakeholders of the drug supply chain (for description convenience, in this paper we refer stakeholders to only enterprises along the supply chain (e.g. manufacturers, packagers, wholesalers, pharmacies and hospitals), not including patients) could conveniently know where the drug has gone or has been. Indeed, the drug traceability has also become more and more a compulsory mandate by governments around the world [11, 29, 39]. For example, the U.S. Drug Supply Chain Security Act (DSCSA) [39], which was signed into law on November 27, 2013, requires that an electronic, interoperable system should be built to identify and trace prescription drugs as they are distributed in the United States. In China, for approximately 8 years, stakeholders above had been required to upload the drug information of individual pharmaceutical products to the official designated IT system whenever there are drugs in or out of their warehouses (which was suspended in Feb 2, 2016 for the stakeholders' concern about their data privacy and the fairness of the system operated by a company which also has its pharmaceutical business [31])

Generally, a drug traceability system should be able to keep track or trace of the drug transaction flow through different stakeholders along the drug supply chain. It should provide reliable information about the flow for stakeholders and patients, especially that of drug production origin for anti-counterfeit purpose. Or, at least it could be used to bind the responsibility of drug security to the relevant stakeholders for government regulation. Furthermore, the privacy of traceability data in the system is required to be protected as much as possible, especially that of statistical information (e.g. productivity, quantity of sale, etc.) of drugs that have been past the stakeholder. Moreover, the system needs to provide useful APIs for stakeholders to flexibly integrate into their own ERP (Enterprise Resouce Planning) system.

Currently, the mainstream technical solution (e.g. [1, 26, 36]) adopts a centralized client-server architecture similar to what was done in China. And patients or stakeholders query the server to verify the authenticity of the drug they bought or to track or trace the information flow along the drug supply chain. Several problems existed in this centralized scheme. First, the raw drug information uploaded to the server could be easily obtained and modified by the service operating entity, which directly impairs the stakeholders' business privacy and the data authenticity. Second, the centralized server is more vulnerable to Denial of Service attack and the resilience heavily depends on the technical ability of operating entity. Third, in this architecture, integrating the drug traceability system with the stakeholders' own ERP system is not flexible with limited and specific APIs provided by the service provider.

This paper presents Drugledger, a fully scenario-oriented blockchain system for drug traceability and regulation. It reconstructs the whole service architecture, ensures the both authenticity and privacy of traceability data, and meanwhile achieves a finally stable blockchain storage with time going by. Algorithms reflecting the practical workflow of drug supply chain have also been presented.

The design of Drugledger follows several principles. First, Drugledger blockchain system should truly reflect the practical drug transaction logic of the supply chain, especially that of drug package, repackage, unpackage, and order cancelling, etc. Second, Drugledger should guarantee both authenticity and privacy of the stakeholders' traceability information , since many untrusted parties from different drug supply chains coexist in the same blockchain network. And this should be

achieved without losing resilience of the blockchain system. Third, the data storage of the blockchain system should not be continuously increasing without end, which guarantees the scalability in storage with time going by. Finally, Drugledger, should be able to counter Sybil attacks [8].

Following these principles, Drugledger is designed. First, to better reflect the cyber-physical characteristics of the track-and-trace system, Drugledger uses an expanded UTXO-based transaction model tailored to the drug supply chain to construct the whole workflow, including drug package, repackage, unpackage, and drug transaction cancelling, manufacturing, drug arriving and leaving, etc. Specially, Drugledger implements the corresponding mechanisms of package, repackage, and unpackage, which is of great significance not only for traceability in a drug level but for tracing the whole drug supply chain originating from the pharmaceutical raw materials.

Second, Drugledger seperates the service provider into three independent roles, namely, certificate service provider (CSP), query service provider (QSP), and anti-attack service provider (ASP). Stakeholders in Drugledger compute a proper one-way cryptographic function (e.g. cryptographic hash of SHA-256) of the properly encoded raw information certain level packages of drugs so as to get the irreversible metadata. This metadata will be added to the Drugledger blockchain in the form of blockchain transaction across every replica as both proof of existence of physical drug transaction and more importantly traceability index in Drugledger. Furthermore, indexes of different package levels could be correlated based on the aforementioned UTXO workflow, especially that of package, repackage, and unpackage. This way, Drugledger manages to separate drug traceability query service from its data modification while ensuring data authenticity and privacy. Besides, Sybil attacks could also be effectively countered.

Third, in Drugledger, we do not seek for general purpose schemes for storage pruning, but only focus on pruning performance of the specific scenario, which we believe will be better optimized. After being fully acquainted with the drug supply chain scenario, we propose to prune the blockchain based on the drug expiration date and give the corresponding algorithm so that the blockchain could finally achieve a stable storage, which seems to have been ignored in the blockchain academic community which cares more about universality than usability and feasibility.

This paper is organized as follows. Related works on bockchain-based drug traceability system are first presented in Section II, where the blockchain background is also given. Sections III then describes detailed technical design of Drugledger from four aspects, namely, service provider seperation, expanded UTXO-based workflow, drug package, repackage and unpackage, and scenario-oriented storage pruning. After that, Section IV discusses implementation and evaluation of Drugledger. Finally, we conclude the paper in Section V.

II. Related Work

A. Blockchain Background

Blockchain, which is also called distributed ledger, originates from the famous decentralized cryptocurrency of bitcoin [25] where it serves as the underlying core technology. Generally, a blockchain system could be viewed as a distributed system implemented with a group of replicated state machines [28] in a peer-to-peer network. The transaction is the basic behavior unit of blockchain, which could be generally defined as the transfer of some digitalized asset that maybe have its meaning in reality. Historical agreed transactions form the current configurations of the replicated state machine. New transactions can be replicated and finally delivered to the blockchain system by executing the corresponding group of replicated state machines, usually including executions of some consensus algorithm and external validity of transactions [6]. Organizations of transactions in a local replica can vary depending on concrete implementations and consensus mechanisms, the mainstream of which is in the form of linked list of blocks where a block contains transactions in a period. Another recently remarkable form of transaction organization is to use the direct acyclic graph (i.e. DAG) [17].

Academic research on blockchain has not come so long though the field has attracted an increasing number of excellent researchers. Currently, works on blockchain are mainly focused on the security, efficiency, and scalability of consensus protocols with an emphasis on the application of cryptocurrency or general purpose blockchain, and have increasingly treated it in a more formal way learned from established practice in areas of distributed systems and cryptography [13, 18, 27]. In the industry, more self-claimed general purpose blockchain platforms are being developed attached with the so called smart contact programmable module, such as Ethereum [9], Hyperldger Fabric [16], etc.

Blockchain is promising to bring several beneficial changes to the drug track-and-trace industry. First, drug traceability system based on blockchain could efficiently prohibit the modification of data, ensuring the authenticity. Second, the security and resilience of the system could be enhanced for utilizing such a p2p architecture. Finally, blockchain makes it easy for stakeholders to integrate necessary information into their ERP system, since they will have full access to their data flow.

B. Traceability of Drug Supply Chain

Currently, few technical and practical works could be publicly found on combining drug traceability system with blockchain, but there are some general discussions. For example, Mettler *et al*. [24] generally proposed the possibility of using blockchain to fight counterfeit drugs in the pharmaceutical industry. Kurki [19] disscussed the benefits and guidelines for utilizing blockchain in the drug supply chain. Archa *et al*. [3] shared their insights of combining the GDP IoT framework with Tendermint [32] blockchain for drug supply chain traceability but gave few technical details on

how they implement the practical workflow with blockchain. Bocek *et al*. [4] said they had built a prototype of drug supply chain traceability system based on Ethereum smart contract [9], without giving concrete design of the worklfow.

C. Traceability of Universal Supply Chain.

Despite of few works on the specific drug traceability area, traceability research for universal supply chains or other specific supply chain scenarios is available, which is necessary to be introduced here. Tian [33, 34] proposed to use blockchain to track and trace agri-food with a conceptual framework given and make comparative analysis of the advantages and disadvantages to the traditional solutions of centralized architecture. Lu *et al*. [23] shared their experience of developing OriginChain with a consortium architecture based on smart contracts for product traceability, in which some valuable general design questions are discussed (e.g. storage of data off-chain or on-chain, data privacy, regulation compliance, etc.). But it gives few concrete details on how they implement their supply chain workflow, which is addressed in Drugledger. Li *et al*. [21] proposed a framework based on hybrid p2p network and blockchain data model for tracking the real-time status of the shipment with a focus on the traceability of truckload transportation stage. To be strict, it is not a blockchain based system, since the p2p network (sub-network) for producing a private ledger was created on demand for once only between two end-to-end nodes of cooperating stakeholders and would then be terminated after the cargo movement was completed. Toyoda *et al*. [35] proposed a product ownership management system for anti-counterfeits in the post supply chain based on Ethereum smart contract, where it designed Manufacturers-Manager Contract for managing the information of manufactures like enrollment of a new company and ProductsManager Contract for managing products information like enrollment of a new product. As a general framework, it gives the basic product ownership logic at the Ethereum smart contract level for proof of concept, without considering much the practical needs of various scenarios (like package and unpackage), and how to protect the commercial data privacy when using public smart contract and reduce the system storage, which are otherwise addressed in Drugledger.

III. Design of Drugledger

A typical drug supply chain scenario can be seen in Fig.1. The basic components include supplier, manufacturer, wholesaler distributor, packager, pharmacy, hospital, and also patients. Here, for accuracy, we define drug transaction in a drug supply chain as the physical flow of drugs from one stakeholder to another stakeholder or from one station to another station of the same stakeholder (e.g. drug flow from manufacturer to wholesaler distributor, or from agency *A* to agenecy *B* of the same wholesaler distributor). Accordingly, we still define Drugledger transaction as transfer of digitalized asset as in Section II.A. Here, the digitalized asset corresponds to the physical drug and Drugledger transaction maps to the drug transaction in a physical suppy chain.

Drugledger is a system for drug traceability and regulation that integrates blockchain with drug supply chain. The cyber-physical characteristic should be considered when designing such a system. That is, there are two different and asynchronous flows of drugs that should be taken into account when designing Drugledger: physical flow of real drugs along the drug supply chain, and corresponding information flow that goes through the Drugledger network in the form of blockchain transactions.

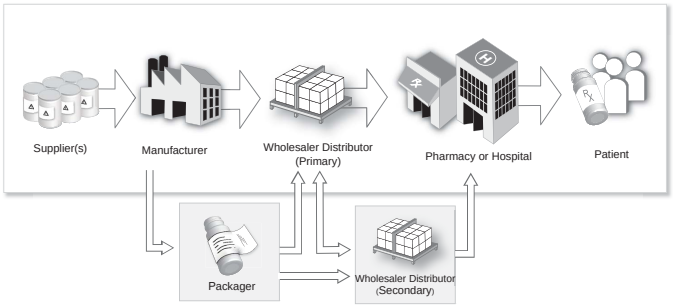

Fig. 1. A drug supply chain example [37]

In this section, we will first present the service provider separation mechanism which constructs the whole service architecture of Drugledger and ensures traceability data authenticity and privacy. Then the basic workflow in Drugledger is detailed and the corresponding algorithms are given. After that, we address how Drugledger implements mechanisms of package, repackage, and unpackage, which are of great significance for the real transaction logic and complete traceability. Finally, a scenario-oriented storage pruning algorithm is discussed, which achieves a finally stable and acceptable blockchain storage, enhancing its scalability. A basic illustration of Drugledger on its service architecture and basic workflow could be found in Fig. 2, where *Other Nodes* refer to other stakeholders not in the same drug supply chain as the one illustrated, and which will be further detailed in the next subsections.

A. Service Provider Separation

For traditional centralized server-client solutions provided by some third-party service provider (i.e. some technical operating company), one of its drawbacks is the inability to separate the function of query from that of data access and modification, which naturally raises stakeholders' concern about their data privacy and the authenticity of track-and-trace information. The key lies in the double roles of the only service provider, who is on one hand the query service provider for both the stakeholders and also the patients, and on the other hand the unique administrator which could obtain the complete production information and also modify the data uploaded by the stakeholders.

Drugledger reconstructs the service architecture by separating the service provider into three independent parts, namely, certificate service provider (CSP), query service provider (QSP), and anti-attack service provider (ASP), as seen in

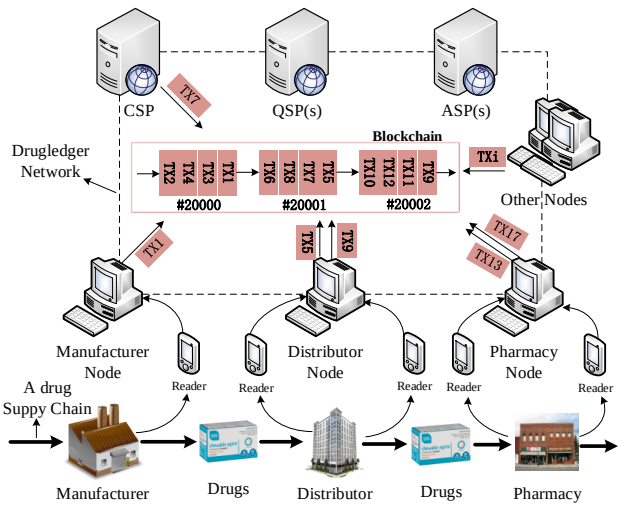

Fig. 2. Drugledger service architecture and basic workflow

Fig. 2. This seperation ensures the authenticity and privacy of stakeholders'traceability data, without sacrificing resilience of the blockchain system.

1) Certificate Service Provider (CSP): An PKI (Public Key Infrastructure) service operated by CSP is embedded into Drugledger which requires only nodes with valid certificate could participate in Drugledger. It could be used to counter Sybil attacks where malicious nodes may create quantities of fake names and execute many meaningless or malicious transactions. Furthermore, illegal pharmaceutical stakeholders are also banned to participate in Drugledger. The stakeholder's real identity information is not included in the Drugledger certificate. Instead, its public key in the Drugledger network will be recorded in the certificate. For specific participant in Drugledger, it could have multiple authenticated pulic keys, which increases the difficulty of possible malicious commercial analysis. Moreover, concrete roles of the stakeholders (manufactures, wholesaler distributors, etc.) are also verified and added to the certificate for supporting diverse potential scenarios, such as countering attacks of flooding quantities of fake drug production transactions in Drugledger network and forcing stakeholders to deal with these meaningless transactions.

Participants' identity privacy is protected in the sense that other parties not cooperating with them cannot manage to get the real identity information through Drugledger network. Besides, stakeholders in the same supply chain do not or not need to rely on this to build any mutual trust (consider the physical scenario of drug supply chain business where cooperators apparently know each other offline). The certificate service provider could be the drug supply chain regulator (i.e. FDA in U.S., CFDA in China) or some other designated operating entity, and this could be embedded in a modular way for Drugledger.

2) Query Service Provider (QSP): In Drugledger, every node can be a query service provider and stakeholders in the same drug supply chain are free to register one or more of the many query service providers they want to cooperate with. The QSP could thus correlate stakeholders' public keys with physical identities in real world (e.g. a hospital name). First, for packages of drugs, they all have been encoded with necessary drug information (i.e. raw data) into the barcode or the RFID label attached to them, according to some standards (e.g. GS1 standard [15]). Second, for stakeholders in Drugledger, they are required to add the cryptographic hash (e.g. SHA-256) of the above encoded raw data (obtained by scanning the barcode or the RFID label) as metadata to Drugledger transaction, whenever a corresponding drug transaction happens. The metadata will be used as future index for quering the service provider for the whole track and trace information of drugs relating to certain stakeholders. Finally, the QSP could thus provide the traceability service for certain package of drugs by first searching the Drugledger replica for all transactions relating to the package metdata, and then interpreting them according to the Drugledger transaction workflow discussed in later sections. Therefore, the traceability data privacy is protected while guaranteeing the commercial service delivery and not losing resilience of the Drguledger blockchain system.

3) Anti-attack Service Provider (ASP): Drugledger sets the role of anti-attack service provider for supervising exceptional activities in the network to keep it functioning properly. For instance, an ASP may detect that some stakeholders in Drugledger have been abusing transactions that apparently go beyond its business scale. Different from certificate service provider, ASPs do not have the right to directly revoke suspicious stakeholders' access to Drugledger or make possible economic punishment, but will report relevant details to CSP which will decide on the following actions. In this sense, ASP is more like anti-virus software for a operating system (i.e. Drugledger system) of specific version in a modern computer. Furthermore, concrete detecting mechanisms are very adaptive and will keep evolving according to Drugledger dynamics and the corresponding possible attacks on Drugledger network. More importantly, Drugledger do not aim to be some decentralized platform but focuses on being a feasible drug traceability system combined with necessary technologies or techniques. Setting the role of ASP is innovative and significant, which contributes to the sustaining availability of Drugledger. Anti-attack services could be operated by competent technical companies which provide technical support services and could charge for them.

B. Expanded UTXO-Based Workflow

The most important aspect that should be considered when designing such a system is the cyber-physical characteristic of the drug traceability system, which shares both characteristics of the drug track-and-trace information flow through the Drugleder network and also drug transaction flow of real world. In terms of this, Drugledger uses the expanded UTXO

transaction model [2] to construct the whole workflow.

In bitcoin system, a transaction is a transfer of bitcoin from one owner to the other, which is implied by the transaction data structure in Fig. 3. The field of transaction output signifies

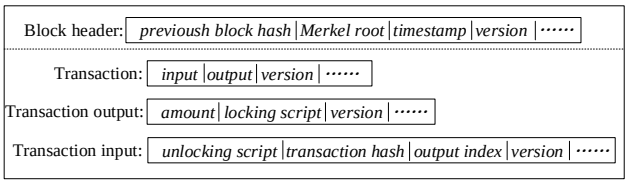

Fig. 3. Bitcoin transaction and block header [2]

how many bitcoins have been transafered to which owners (i.e. public keys), in which the right to spend the bitcoin is expressed in locking script. Only the one being able to produce the right unlocking script (e.g. public key verification script) can spend the bitcoin value existing in the unspent transaction output (UTXO). The locking script and the pointer to the transaction containing the UTXO should thus be included in the field of transaction input.

Drugledger expands the transaction model above to the drug supply chain, as illustrated in Fig. 4. The fields of *metadata, type, lifetime, block lifetime* are newly added, in which *lifetime* and *block lifetime* will be detailed in subsection *D*. The field of *type* refers to different kinds of Drugledger transactions which maps to corresponding drug transactions in drug supply chain. Drugledger tailors the expanded UTXO model to drug supply chain scenario in several ways.

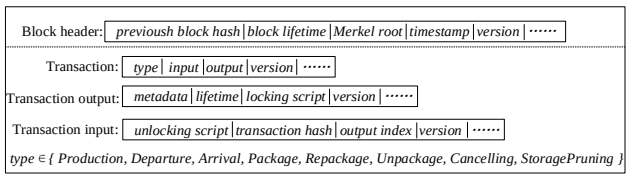

Fig. 4. Drugledger transaction and block header

First, most transactions in Drugledger blockchain are corresponding to the physical drug transaction process, as seen in Fig. 2. It records the corresponding ownership transfer of drugs of different-level packages. Stakeholders are required to initiate Drugledger transactions which take as input the previous "unspent" transaction output belonging to corresponding stakeholders, whenever there are drug transactions accordingly (e.g. reaching or leaving the warehouse). The Drugledger transactions can only be initiated again by the authenticated stakeholders which have valid private keys.

Second, the transactions above are initiated for two main purposes. One is for declaring that the drug has left or arrived the warehouse (e.g. TX5 and TX9 in Fig. 2). This is because the information flow in the Drugledger is not synchronized with the physical flow of drugs. And the other is for the first-time production (i.e. TX1 in Fig.2) from an authenticated manufacturer where drugs are packaged in a minimum level. The basic workflow for production, departure and arrival of drugs is described in Algorithm 1.

Algorithm 1 Basic Drugledger transaction procedure for drug production, arrival, and departure along the drug suppy chain.

procedure INITIATEBASICTX
  type ← TYPE
  // TYPE could be Production, Arrival, or Departure
  $utxos \leftarrow$ *SynchronizeUTXO()*
  // stakeholders synchronize the node to get UTXOs
  // availabe to "spend"
  $rawdata \leftarrow$ *ReadDrugPackage()*
  // read the encoded drug package to get the raw data
  $metadata \leftarrow$ *SHA256(rawdata)*
  // get the metadata in possible Drugledger transactions
  // by cryptographic hash of SHA-256
  $\langle utxo, utxoptr \rangle \leftarrow$ *GetUTXO(utxos, metadata)*
  // check if there is corresponding utxo to metadata
  // in above utxos
  if $\langle utxo, utxoptr \rangle == \langle null, null \rangle \&\&$ $type\ != Production$ **then**
  return false
  // *utxo* == *null* means that the package just read is not
  // in the Drugledger system. if it is not for the first
  // time production, then wrong operation, return
  $isvalid \leftarrow$ *ValidQuery(metdata,pidman,pidpre,QSP)*
  if $isvalid ==$ *false* **then return** false
  // if the package of drugs exists, query for QSP
  // to verify the manufacturer physical identity *pidman*
  // and the previous stakeholder physical identity *pidpre*
  // this is optional, if cooperators are trusted
  $tx \leftarrow$ *CreateTX(metadata,utxo,utxoptr,viddl,type)*
  // if all is well, "spend" the utxo and create the
  // transaction as in Fig 4, *viddl* is the public key of
  // next stakeholder in Drugledger network
  $Gossip(tx)$ // gossip the transaction for consensus
  return true

Third, misbehaviors on the transaction in such a permissioned Drugledger system can be cancelled, by adding an additional "payee" (i.e. current owner of the drugs) in the transaction. This happens if and only if it has not been "spent" by the original "payee" which has initiated transaction to verify the ownership transfer of drugs. This is very realistic and important when carefully considering the practical business scenario of the drug supply chain, for example, when the public key of the next owner is lost or wrong public keys are added to the transaction.

Finally, in some circumstances where drugs need to be packaged, repackaged or unpackaged, the workflow above cannot work and new mechanisms need to be introduced to solve the problem, which are described in the next subsection.

C. Drug Package, Repackage, and Unpackage

In a practical drug supply chain scenario, the drug track-and-trace workflow is more complex than what is described in

Fig. 2, since there are different levels of packages of drugs. For example, from the manufacturers' production of drugs with a minimum package level, to the patients' consumption of drugs also with a minimum package level, there are processes of drug package, repackage, and unpackage, which make the basic workflow disscussed in subsection *B* more complicated, as illustrated in Fig. 5. More importantly, it also makes the track and trace more difficult, since the same drug can be expressed in different Drugledger transactions with different metadata standing for various level packages.

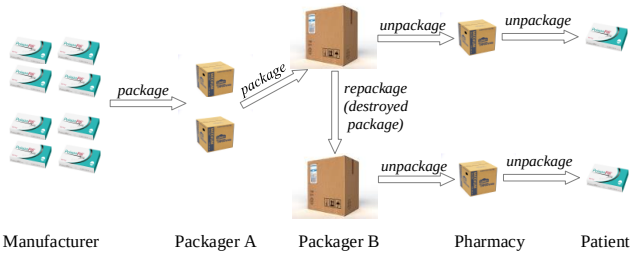

Fig. 5. Package, repackage, and unpackage in drug supply chain.

In Drugledger, mechanisms of package, repackage and unpackage are proposed to keep the UTXO-based basic workflow still effective. The basic idea is to realize processes of package, repackage and unpackage as the inherent Drugledger transactions which transfer drugs to the current stakeholder itself (which we call package transaction, repackage transaction and unpackage transaction). After that, the basic UTXO workflow discussed in subsection B can still function as before. The relations between different metadata of various level packages are recorded in the Drugledger blockchain, so that we could still keep track or trace of the whole drug supply chain.

Since drug package, repackage or unpackage happens after the arrival of drugs of current package level (i.e. Arrival transaction in Drugledger), we do not need to repeat the validation steps executed in Alogrithm 1 when initiating corresponding transactions. The metadata of next level package (computed as in Section III.A) is added to the corresponding transaction outputs.

For package transaction, where drugs of certain package level need to be reorganized as a whole in a higher-level package, it is created by spending the multiple "unspent" transaction ouputs which stand for the multiple current-level packages of drugs and setting the next owner to the packager itself in the transaction output, as illustrated in simplified Fig. 6. $UTXO^{i}_{pl}$ and $UTXO^{j}_{pl}$ refer to current unspent transaction outputs corresponding to package *i* and package *j* of drugs with package level pl. TXO^{k}_{pl+1} refers to the transaction output corresponding to the next level package *k* of drugs with package level $pl+1$. Therefore, the information flow can be traced with the interrelation reflected in the package transaction.

For repackage situation, where there are some destroyed packages of drugs, the repackage transaction is realized by spending the UTXO of detroyed package of drugs to the

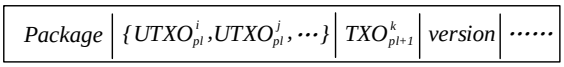

Fig. 6. Drugledger package transaction (simplified)

repackager itself, with a new metadata added to the repackage transaction output. The new metadata is obtained by comptuting cryptographic hash (SHA-256) of the newly added package raw data. Besides, it is necessary to check whether there was a package transaction which packaged drugs of previous level package into the current destroyed package. This is to avoid possible manmade mistakes, like mixing low-level packages not in the same current detroyed package. The repackage procedure is described in Algorithm 2.

Algorithm 2 Repackage under the situation of ruined package of current level.

```
1:  procedure REPACKAGE
2:      type ← TYPE
3:      // TYPE = Repackage
4:      utxos ← SynchronizeUTXO()
5:      rawdatacur ← ReadDrugPackage()
6:      // read current destroyed package
7:      metadatacur ← SHA256(rawdatacur)
8:      ⟨utxo, utxoptr⟩ ← GetUTXO(utxos, metadata)
9:      rawdatapres ← ReadDrugPackage()
10:     // read all packages of previous level in the
11:     // destroyed package, obtaining corresponding raw data
12:     metadatapres ← SHA256()
13:     result ← IsCorrelated(metadatacur, metadatapres)
14:     // verify correlation between packages in Drugledger
15:     if result == false then return false
16:     CreateTX(metadatanew, utxo, utxoptr, viddl, type)
17:     // metatanew is caculated from the newly added
18:     // package that replaces the current destroyed package
19:     return true
```

In addition, the unpackage process also influences the function of track-and-trace in Drugledger, since the metadata of the lower level package can only be traced before the unpackage stage based on the expanded UTXO transaction model. Unpackage transaction is thus proposed to solve the problem. Similar to package transaction, unpackage transaction is created by spending the "unspent" transaction output corresponding to currentl level packages of drugs and producing the multiple transaction outputs with metadata of lower-level packages of drugs in the current level package, as illustrated in simplified Fig. 7. Similar to repackage transaction, unpackage

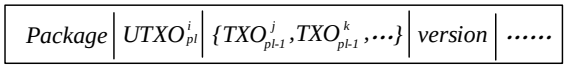

Fig. 7. Drugledger unpackage transaction (simplified).

transaction will also go through the same validation process to verify the correlation between metadata of transaction input and those of transaction outputs.

D. Scenario-Driven Storage Pruning

The problem of continuous increasing storage has been a general bottleneck of blockchain system, for which valuable approaches (like bitcoin SPV (Simplified Payment Verification [25], and Ethereum state tree pruning [10]) have been proposed. These approaches, however, still cannot stop the increasing storage.

In Drugledger, the practical scenario of drug supply chain is carefully investigated for potential optimization, where the expiration date of drugs is taken into account to optimize the workflow. It is shown that the drug expiration time (shelf life) is usually less than 5 years as a result of business reproduction and government regulation requirement [12] [30] [38], which means the transaction in Drugledger could have its *lifetime*. Considering the length of the blockchain is linearly growing with time, the *block lifetime* in Drugledger could then be optimized to the maximum *lifetime* of production transactions (*type* = *Production*) in it. For blocks without production transactions in it, the *block lifetime* is thus defined as 0. Therefore, the block can be pruned after the expiration or some time long after that if necessary, which contributes to a finally stable and acceptable blockchain storage. There are circumstances where the expiration date of a specific drug is extremly long (like some kinds of Chinese traditional drugs), which greatly increases the block lifetime. In this situation, we currently choose to ban transactions of those drugs in Drugledger or just to simply ignore these drugs when pruning Drugledger.

Algorithm 3 Storage pruning in Drugledger.

```
1:  procedure PRUNESTORAGE
2:      type ← TYPE
3:      // TYPE = StoragePruning
4:      timestampgen ← GetTimestamp(genesisblock)
5:      // timestamp of the current begining block
6:      timestampnew ← GetTimestamp(latestblock)
7:      // timestamp of the newly added block
8:      // or some blocks away from it for security
9:      lifetimegen ← GetLifetime(genesisblock)
10:     // get the block lifetime in block header
11:     blockage ← (timestampnew - timestampgen)
12:     if blockage < lifetimegen then return false
13:
14:     CreateTX(metadata, null, null, null, type)
15:     return true
```

The pruning transaction in Drugledger will be recorded in the blockchain before the local block pruning operation is executed. The security and resilience of Drugledger are still maintained, since those blocks are pruned forward from the beginning and the transactions (i.e. the past Drugledger transactions relating to expired drugs) in the pruned block will not be used again in the future. The following block then becomes the oldest block ("genesis block"), which is different from cryptocurrency like bitcoin in which there is only one kind of transaction and every transaction may be used in the future for either balance calculation or validation and thus has an indefinite lifetime. The Drugledger pruning transaction (i.e. TX7 in Fig.2) is initiated only by the CSP with certificate verification once a block is expired or sometime long after that. After the local pruning operation, those expired drugs cannot be traced in Drugledger with a complete information flow and will thus be taken invalid. The process is detailed in Algorithm 3.

IV. Implementation and Evaluation

A. Implementation

Currently, we are developing a prototype of Drugledger mainly with C++ in Ubuntu 16.04 LTS. The main modules of Drugledger include certificate module, storage module, communication module, consensus module and Drugledger transaction logic module. Some popular open source libraries are used in the current working prototype, like Crypto++® [7], LevelDB [20], libevent [22], Boost [5], etc. The communication module uses gossip to implements a p2p network. The consensus module is implemented temporarily with algorithms from Algorand cryptocurrency [14], where it limits the participation of final byzantine consensus process (i.e. $BA\star$) to a small committee selected by cryptographic sortition and thus improves its transaction throughoutput. Users in Algorand are weighted by their balance, whereas Drugledger weights its user based on the number of valid transactions in the past. This is reasonable in the sense that stations of large companies process more pharmaceutical transactions regularly, take more risks, and thus should share more reputation in proposing a block.

B. Evaluation

1) Practicality: The first and foremost requirement for drug supply chain traceability system should be its practicality, which seems to have been ignored in related works in Section II. Drugledger is designed for practical drug supply chain traceability and regulation. First, Drugledger gurantees the commercial sustainable service delivery by reconstructing the service architecture while protecting the data privacy and authentication. Second, as a blockchain solution, Drugledger first raises and tackles problems of package, repackage, and unpackage which exert great influece on the traceability of drugs in a practical drug supply chain. Finally, Drugledger first proposes the scenario-orientied optimization in decreasing blockchain storage with drug expiration date, which achieves a finally stable and acceptable storage.

2) Security: Drugledger is designed to be a permissioned blockchain system for countering Sybil attacks. Certificate service provider controls the access of participants to Drugledger. Nodes not in a drug supply chain or stakeholders with bad commercial reputation in drug supply chains are therefore banned to access the network. Since Drugledger is based on

peer to peer architecute, it is less prone to DoS attacks compared to traditional server-client solutions. Formal analysis of secure multi-party computation in Drugledger may require an extra large section describing the concrete consensus protocol, which itself would already be an independent work and thus be ignored here.

3) Efficiency: The efficiency (e.g. transaction throughoutput, latency, etc.) of blockchain is mainly determined by the consensus protocol. Drugledger does not seek for absolute improvement in efficiency, but only focus on the practical requirement of drug supply chain. In Drugledger, the communication overhead with service provider (i.e. CSP, QSP), efficiency of consensus algorithm, and the practical throughput requirement of drug supply chain together determine whether it could be applied in production system. Quantitative assessment is presented after the implementation of current prototype, which will be discussed in future work.

V. Conclusion

In this paper, we propose Drugledger, a practical blockchain system for drug traceability and regulation. Drugledger reconstructs the whole service architecture by service provider separation, guaranteeing the authenticity and privacy of traceability data while ensuring the service delivery. Service providers (CSP, ASP, and QSP), stakeholders, and also patients as a whole form a healthy and sustainable business ecology. Drugledger completes its workflow based on the expanded UTXO data structure, especially that of package, repackage, and unpackage. Expiration date of drugs is utilized to prune blockchain storage, achieving a finally stable and acceptable storage.

References

[1] AliHealth. Available: http://www.mashangfangxin.com

[2] A. M. Antonopoulos, "Mastering bitcoin: unlocking digital cryptocurrencies", *1st edition, O'Reilly Media, Inc*, 2014.

[3] Archa, B. Alangot, and K. Achuthan, "Trace and track: enhanced pharma supply chain infrastructure to prevent fraud", *in Springer International Conference on Ubiquitous Communications and Network Computing*, Aug. 2017, pp. 189-195.

[4] T. Bocek, B. B. Rodrigues, T. Strasser, B. Stiller, "Blockchains everywhere - a use-case of blockchains in the pharma supply-chain", *in IEEE IFIP/IEEE Symposium on Integrated Network and Service Management (IM)*, May. 2017, pp. 772-777.

[5] Boost. Available: http://www.boost.org

[6] C. Cachin and M. Vukolić, "Blockchains consensus protocols in the wild", *arXiv preprint*, Jul. 2017, arXiv:1707.01873.

[7] Crypto++®. Available: https://www.cryptopp.com

[8] J. R. Douceur, "The sybil attack", *in Springer International Workshop on Peer-to-Peer Systems*, Mar. 2002, pp. 251-260.

[9] Ethereum. Available: https://www.ethereum.org

[10] Ethereum state tree pruning. Available: https://blog.ethereum.org/2015/06/26/state-tree-pruning

[11] European Medicines Agency, "Falsified Medicines Directive". Available: https://ec.europa.eu/health/human-use/falsified_medicines_en

[12] European Medicines Agency, "Guidelines on stability testing for applications for variations to a marketing authorisation". Available: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2014/04/WC500164972.pdf

[13] J. A. Garay, A. Kiayias, and N. Leonardos, "The bitcoin backbone protocol: analysis and applications", *in Springer Annual International Conference on the Theory and Applications of Cryptographic Techniques*, Apr. 2015, pp. 281-310.

[14] Y. Gilad, R. Hemo, S. Micali, G. Vlachos, and N. Zeldovich, "Algorand: Scaling byzantine agreements for cryptocurrencies", *in ACM Proceedings of the 26th Symposium on Operating Systems Principles*, Oct. 2017, pp. 51-68.

[15] GS1 standards in healthcare. Available: https://www.gs1.org/healthcare/standards

[16] Hyperledger Fabric. Available: https://www.hyperledger.org/projects/fabric

[17] IOTA white paper. Available: https://iota.org/IOTA_Whitepaper.pdf

[18] A. Kiayias, A. Russell, B. David, and R. Oliynykov, "Ouroboros: A provably secure proof-of-stake blockchain protocol", *in Springer Annual International Cryptology Conference*, Aug. 2017, pp. 357-388.

[19] J. Kurki, "Benefits and guidelines for utilizing blockchain technology in pharmaceutical supply chains: case Bayer Pharmaceuticals", *Bachelor thesis, Aalto University*, 2016.

[20] LevelDB. Available: https://github.com/google/leveldb

[21] Z. Li, H. Wu, and J. Wassick, and J. Tazelaar, "On the integration of event-based and transaction-based architectures for supply chains", *in IEEE 37th International Conference on Distributed Computing Systems Workshops (ICDCSW)*, Jun. 2017, pp. 376-382.

[22] Libevent. Available: http://libevent.org

[23] Q. Lu and X. Xu, "Adaptable blockchain-based systems: A case study for product traceability", *IEEE Software*, vol. 34, no. 6, pp. 21-27.

[24] M. Mettler, "Blockchain technology in healthcare: The revolution starts here", *in IEEE 18th International Conference one-Health Networking, Applications and Services (Healthcom)*, Sep. 2016, pp. 1-3.

[25] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system", http://bitcoin.org/bitcoin.pdf, 2008.

[26] OPTEL track and trace. Available: https://www.optelpharmaceutical.com

[27] R. Pass, L. Seeman, and A. Shelat, "Analysis of the blockchain protocol in asynchronous networks", *in Springer Annual International Conference on the Theory and Applications of Cryptographic Techniques*, Apr. 2017, pp. 643-673.

[28] F. B. Schneider, "Implementing fault-tolerant services using the state machine approach: A tutorial", *ACM Computing Surveys (CSUR)*, vol. 22, no. 4, pp. 299-319, Dec. 1990.

[29] State Food and Drug Administration of China, "Notice on questions of putting the electronic supervision system in practice", Apr. 2008. Available: http://www.sda.gov.cn/WS01/CL0055/29178.html

[30] State Food and Drug Administration of China, "Guidelines for drug registration". Available: http://www.sda.gov.cn/WS01/CL0053/24529_3.html

[31] State Food and Drug Administration of China, "On suspension of drug electronic supervision system", Feb. 2016. Available: http://www.sda.gov.cn/WS01/CL0051/144782.html

[32] Tendermint. Available: https://tendermint.com

[33] F. Tian, "A supply chain traceability system for food safety based on HACCP, blockchain and Internet of things", *in IEEE International Conference on Service Systems and Service Management (ICSSSM)*, Jun. 2017, pp. 1-6.

[34] F. Tian, "An agri-food supply chain traceability system for China based on RFID and blockchain technology", *in IEEE International Conference on Service Systems and Service Management (ICSSSM)*, Jun. 2016, pp. 1-6.

[35] K. Toyoda, P. T. Mathiopoulos, I. Sasase, and T. Ohtsuki, "A bovel blockchain-based product ownership management system (POMS) for anti-counterfeits in the post supply chain", *IEEE Access*, Jun. 2017.

[36] Tracelink. Available: https://www.tracelink.com

[37] U.S. Food and Drug Administration, "A drug supply chain example". Available: https://www.fda.gov/downloads/Drugs/DrugSafety/DrugShortages/UCM277651.pdf

[38] U.S. Food and Drug Administration, "Drug stability guidelines", Dec. 2008. Available: https://www.fda.gov/downloads/AnimalVeterinary/GuidanceComplianceEnforcement/GuidanceforIndustry/ucm051556.pdf

[39] U.S. Food and Drug Administration, "Drug Supply Chain Security Act". Availabe: https://www.fda.gov/Drugs/DrugSafety/DrugIntegrityandSupplyChainSecurity/DrugSupplyChainSecurityAct