PDPM: A Patient-Defined Data Privacy Management with Nudge Theory in Decentralized E-Health Environments

Seolah JANG†∗†††††, Sandi RAHMADIKA†∗†, Nonmembers, Sang Uk SHIN††††††††, Member, and Kyung-Hyune RHEE†††††††††††, Nonmember

SUMMARY A private decentralized e-health environment, empowered by blockchain technology, grants authorized healthcare entities to legitimately access the patient’s medical data without relying on a centralized node. Every activity from authorized entities is recorded immutably in the blockchain transactions. In terms of privacy, the e-health system preserves a default privacy option as an initial state for every patient since the patients may frequently customize their medical data over time for several purposes. Moreover, adjustments in the patient’s privacy contexts are often solely from the patient’s initiative without any doctor or stakeholders’ recommendation. Therefore, we design, implement, and evaluate user-defined data privacy utilizing nudge theory for decentralized e-health systems named PDPM to tackle these issues. Patients can determine the privacy of their medical records to be closed to certain parties. Data privacy management is dynamic, which can be executed on the blockchain via the smart contract feature. Tamper-proof user-defined data privacy can resolve the dispute between the e-health entities related to privacy management and adjustments. In short, the authorized entities cannot deny any changes since every activity is recorded in the ledgers. Meanwhile, the nudge theory technique supports providing the best patient privacy recommendations based on their behaviour activities even though the final decision rests on the patient. Finally, we demonstrate how to use PDPM to realize user-defined data privacy management in decentralized e-health environments.

key words: blockchain, decentralized e-health, data privacy management, nudge theory, smart contract, tamper-proof data

1. Introduction

The emerging technologies, such as artificial intelligence (AI), the Internet of things (IoT), blockchain, and robot technology, can result in better medical services [1], and precise and accurate medical planning that is no longer hindered by time and place. Various methods have been heavily proposed by industry, academia, and healthcare stakeholders.

Medical records of patients have become crucial parts of the modern healthcare environment. Many terms refer to patient medical record and data that can be accessed online or offline, such as electronic medical record (EMR), personal health record (PHR), continuity of care record (CCR), open electronic health records (openEHR), and so forth (see Table 1). Security techniques in accessing the medical patient’s data also support the significant increase in the data amount from several resources over time. In some cases, patient data is stored and scattered in different storage/services, making it more challenging to get full access to the data. Therefore, the blockchain-based approach is utilized as a platform that can be gathered from various sources and accessed concurrently by authorized parties within a single system. Thus, multiple aspects have begun to be further investigated, such as security, communication, and system effectiveness.

Existing e-health systems have a default data privacy policy for each patient, depending on the hospital’s policies or healthcare stakeholders. Thus, there are various types of default privacy policies in e-health environments, such as presented in [2], [3], and [4]. More precisely, the paper in [3] presented the scheme in protecting consumers’ privacy and personal data from e-health’s most common service and online reservation services. Some of the current e-health systems do not provide full flexibility for patients to update their privacy, let alone offer new privacy updates for old patients who have been registered in the system in advance. A conventional e-health system provides privacy updates for patients. However, it is still executed manually, and information is stored in a logbook (online and offline) which can be changed by irresponsible parties, leading to disputes in the future. Therefore, e-health needs an innovative arrangement that can provide recommendations for periodic privacy updates by patients themselves without any intervention. The updated version of privacy information is then recorded on the blockchain.

We propose a collaborative system called PDPM as a patient-defined data privacy management using the nudge theory concept in decentralized e-health environments to address the issues mentioned earlier. PDPM provides the best possible data privacy recommendations for patients based on accumulated data processed using the nudge theory concept. The patient fully determines the privacy management based on the suggestions given by the system. Furthermore, the selected data privacy management is executed into on-chain. Authorized entities within a confidential manner can inspect the record of every change. We leverage the Ethereum platform with smart contract feature as an
open-source and decentralized software platform.

In summary, this research provides the following contributions:

(i) We construct a secure architecture that enables patient-defined data privacy management using the nudge theory concept in decentralized e-health environments.

(ii) We provide the state of the art of our decentralized e-health system in facilitating patient-defined data privacy control.

(iii) We formulate the PDPM model and evaluate the performance based on the simulation results.

(iv) We note several requisite concerns and remarks based on our findings in modelling the PDPM system.

The road map of the paper is organized as follows. Section 2 investigates e-health’s existing model, both conventional and decentralized approaches. The prior research on nudge theory are also discussed in this section. In Sect. 3, we breakdown the technical challenges of our research. While in Sect. 4, we present the core system components as our proposed system’s backbone technology. This section includes the essentials blockchain-based smart contract, the concept of nudge theory, and decentralized e-health. Operating-system design and implementation are elaborated in Sect. 5. We also highlight several essential points, including discussions, challenges and remarks. Finally, we conclude this paper in Sect. 6.

2. Related Work

The paradigm in transforming centralized healthcare into a decentralized form has been extensively investigated by industries, healthcare providors, and academia. Methods, platforms, and goals also vary widely to achieve a decentralized healthcare system (relying on blockchain technology). Likewise, the implementation of nudge theory in decision making has been adopted in many cases, such as financial technology (FinTech), online advertising (internet advertising), insurance, open banking, and to name a few. Therefore, in this section, we focus on the prior works on decentralized e-health environment with blockchain technology and the nudge theory (e-health use cases) and its implementation.

We started with our research in 2018 [18] proposing blockchain technology for personal health information (PHI). The primary motivation is to make an efficient way in accessing the patient’s data by authorized entities. In conventional PHI, the patient’s data are scattered and stored locally at different providers over the internet. Moreover, the patients do not have full access to the data. The proposed scheme is proven to be able to overcome the issues mentioned above. Comparative approaches and objectives have also been proposed in research papers in [19], and [20].

More recently, research papers in [21] proposed a system called ssHealth as a secure, blockchain-enabled healthcare system. The authors presented a novel smart and secure healthcare system by adopting edge computing and blockchain techniques. By leveraging the proposed approach, medical data can be exchanged among healthcare entities securely. Related to this technique, a patent in [22] proposing a blockchain as a chronicle of a person’s healthcare path through life that can be shared with the authorized entities and stakeholders. While, for the systematic review, it has been extensively investigated in the following research papers [23], [24], and [25].

With regard to the data privacy management in decentralized e-health environments, numerous research leveraging conventional cryptographic primitives and designing access control models to address these concerns. The authors in [26] proposed a revocable attribute-based signature for the blockchain-based healthcare environment. The system requires pairing operations which are not relying on a central authority to protect the user’s identity. Another research paper in [27] proposed a platform called MediBchain with several cryptography protocols to manage the sensitive health data of the patients on the blockchain to achieve accountability, integrity, and security. Meanwhile, a different cloud-based approach was proposed by authors in [28]. A cloud-based system combined with cryptographic functions is used to encrypt patient data (pseudonymity).

In e-health, nudge theory is used mainly to induce behavioural changes in patients for treatment purposes. A research paper in [29], using a nudge theory to improve clinician decision-making and induce patient physical activities utilizing the loss-framing approach. A cancer patient needs to maintain a healthy life through continuous fitness and diet management. The authors applied a reward strategy, one of the nudge strategies implementations, to induce ongoing physical activity. Based on the theory that humans try to avoid losses rather than gains, the author encourages patients to engage in physical activity to maintain their rewards by giving them virtual rewards in advance, unlike what is usually assigned when the target is achieved.

| Acronym   | Description                                                                 |
|-----------|-----------------------------------------------------------------------------|
| CCR [6]   | Continuity of Care Record (Standard Specification)                          |
| CEN/TC 251 [7] | European Committee for Standardization                                 |
| DICOM [8] | Digital Imaging and Communications in Medicine                            |
| HL7/CDM/FHIR [9] | Health Level-7. Fast Health Interoperability Resources.              |
| HIPAA [10] | Health Insurance Portability and Accountability Act                      |
| ICD/ICF/ICHI [11] | Family of International Classification of Primary Care |
| ICPC [12]  | International Classification of Primary Care                             |
| ISO/TC 215 [14] | International Organization for Standardization                        |
| LOINC [15] | Logical Observation Identifiers Names and Code                           |
| openEHR [16] | Open Electronic Health Records                                           |
| SNOMED-CT [17] | Systematized Nomenclature of Medicine                                    |
These approaches have been used to improve cardiovascular disease.

### 3. Technical Challenges

This section presents the key technical challenges that arise when implementing blockchain technology in an e-health environment to user-defined data privacy management supported by nudge theory, which achieves this paper’s goals (simultaneously granting privacy and succinctness).

#### 3.1 Achieving Secure Patient-Defined Data Privacy

A dynamic patient-defined data privacy management is our primary objective in this research. Our proposed method relies on nudge theory and blockchain technology to achieve the coveted goals. In the actual implementation of e-health, patients’ data privacy management can be changed over time. Data privacy management is one of the most crucial elements in e-health environments where only certain parties can see data, such as doctors, nurses, and others. In other words, the data should be invisible publicly. Meanwhile, all authorized parties in the same private e-health system can see insensitive data (visible for all).

Before being processed by nudge theory techniques, the initial data classification also plays a vital role in managing the patient’s desired privacy management. Several types of data, such as medical treatment data, clinical research data, public institution data, lifelog data, and so on, need to be considered in advance. Therefore, achieving secure patient-defined data privacy is considered as a technical challenge.

#### 3.2 Nudge Theory in E-Health

The concept of nudge theory in e-health that enables better services for patients has begun to get a lot of attention. Our research focuses on designing the concept of nudge theory to realize better privacy management for patients. The challenge in designing nudge theory in decentralized e-health is supported by the many types of data and medical terms that must be filtered using a specific algorithm to obtain the final output. On-chain and off-chain in decision making also need to be thoroughly considered to prevent overhead in blockchain transactions.

#### 3.3 Immutable Data Privacy Management

The last technical challenges are providing unhackable, tamper-proof, and changeless data privacy information by adopting blockchain technology through smart contract features in the Ethereum platform. The smart contract’s input data is an arbitrary value of the patient’s definitive decision derived from the nudge theory provided by the e-health system. Patients have full control in determining which data can be accessed openly by e-health entities, in the sense that it is disclosed to the public. The patient independently can manage their data privacy, which certain parties allow observing the data in the clear. The technique in achieving this idea in the Ethereum smart contract blockchain is a challenge that must be addressed in the research.

### 4. Core System Components

In this section, we present the core system components of our proposed approach. We systematize core system components into three groups: blockchain-based smart contract (based on the Ethereum platform), the concept of nudge theory in general, and decentralized e-health. All core system components are united to achieve the desired objectives that we further utilize in Sect. 5.

#### 4.1 Essentials Blockchain-Based Smart Contract

Blockchain technology with unique features has revolutionized the paradigm of transacting on the internet. Transactions no longer rely solely on an intermediary or third parties in organizing the goals the parties want to accomplish. The transactions’ verification process is carried out in a decentralized manner (not centralized to one node) [30]. The ledgers are propagated across every node in the same blockchain network. Hence, every node has the exact version of the ledger state.

Blockchain transactions are time-stamped and recorded in chronological order that the parties who are granted authority can inspect the transaction. Data that has been successfully verified and stored is tamper-proof; hence, there is no possibility of altering data, data destruction, or data deletion by malicious parties. Miners or validators validate transactions that contain data through a particular consensus mechanism (see Table 2) that eliminates the risk of twofold entry, counterfeit data, or fraud. Explicitly speaking, blockchain technology provides decentralization, immutability, security, and transparency [31].

Blockchain with the smart contract feature began to be widely recognized by the public through Ethereum platforms in 2015, categorized into Blockchain 2.0 (starting from Bitcoin for Blockchain 1.0). The development has continued ever since, as evidenced in 2017 by updating the

| Benchmark     | Public Access                                      | Private Access                                  |
|---------------|----------------------------------------------------|-------------------------------------------------|
| Distributed Validation | Proof-of-work (PoW), Proof-of-stake (PoS), PoW based derivatives, Federated Byzantines agreement | Redundant Byzantine fault tolerance, Ripple consensus bilateral node-to-node (N2N), RAFT and derivatives, Delegated Proof-of-stake (DPoS) |
| Concentrated validation | Delegated Proof-of-stake (DPoS) |                                                     |
Ethereum features into Distributed Applications ICOs (categorized into Blockchain 3.0) [32]. There are various types of smart contract platforms, such as Ethereum and Hyperledger Burrow (Solidity, Serpent, Mutant, and Vipe), Hyperledger Fabric (Golang, Java, JavaScript), Quorum (Solidity), and Open Transactions (ChainScript).

Smart contracts are verified in real-time, based on conditions stated in the contract. The contracts holder can execute the contract’s functions without any interference, eliminating the long-winded attestation process. Regarding the business perspective, merging tasks and automating the contract’s function can assist streamline business services and boost profitability. The smart contracts-based application provides clear communication among the entities. Eventually, smart contracts by design are paperless, lower cost, faster run, and secure, making use of encryption at the blockchain level. With these merits, smart contracts are extensively adopted in various businesses and applications, as evidenced by the number of Ethereum addresses that have increased significantly over time, as illustrated in Fig. 1 (it reaches 34.05m addresses by January 21, 2021). Therefore, blockchain-based smart contracts are also suitable to be adopted in the healthcare environment, such as e-health, personal health information (PHI) and electronic medical records (EMR).

4.2 The Concept of Nudge Theory

Nudge theory is a concept within the field of “behavioural economics”. In traditional economics, it has often been found that it is not reasonable to make individual’s everyday choices in theory, and behavioural economics was born when economists actively accepted the findings of psychology to understand them in the framework of economics. This concept is not economics based on the premise of rational and rational economic human beings, but economics to study practical human behaviour in determining the causes and consequences of decision-making of choice behaviour.

Recently, several Western countries have made various attempts to incorporate ideas and insights from behavioural economics, including nudge, into public policy. At a time when economic incentives, a practical approach to health care policy, do not achieve much, behavioural economics, which points to errors in “economic humans” introduced by mainstream economics and explains human behaviour more plausible based on a solid study of psychology, seems to be an attractive approach.

Nudge strategies can be classified as financial and non-financial means of inducing change by providing financial means such as incentives. The financial instruments of the nudge strategy include information simplification, changes to the physical environment, default policy, and social norm. The default option to be used in this paper is a strategy to induce changes in the behaviour of policy subjects by setting the options that policy designers think are desirable as default. Changing the behaviour of policy subjects with changes in default options can be observed in many nudge policies. This default option follows the theory that “human beings usually tend to choose the first option”.

4.3 Decentralized E-Health

E-health refers to the services and activities related to the systematic provision of medical care to individuals or a community via the internet. There are many terms for e-health that refer to identical functions such as personal health information (PHI), personal health record (PHR), electronic medical record (EMR), and so on. However, all of these terms have something in common: authorized parties can access their health record via the internet provided by healthcare providers. Due to many providers involved in an e-health environment, it causes data to be fragmented and difficult to be accessed by the patient, as stated by a research paper in [34].

A decentralized healthcare system that involves many providers into a single system has begun to be formed by relying on blockchain technology with smart contract features. The objectives of decentralized e-health are spawned with a powerful idea to organize and improve healthcare services worldwide. Table 3 presents the conventional and decentralized healthcare environments in single and multiple collaborated servers. Various approaches with different techniques have been established by academia and industry to provide services that are not limited to time and place for healthcare providers, doctors, and patients.

Figure 2 illustrates the use of blockchain in multiple applications with different objectives. The objectives are divided into decentralized value transfer, gamification, transparency, interoperability, non-centralized verification, autonomous contracts, and immutable transactions. Meanwhile, healthcare use cases are divided into three categories:
Table 3  Benchmarks of the conventional and decentralized healthcare environments.

| Benchmark                  | Architecture | Security |
|----------------------------|--------------|----------|
| Hierarchical distributed EHR (HDEHR) | DE, P2P      | N/D      |
| m-Health                    | DE           | N/D      |
| Ubiquitous PHR (uPHR)       | DE           | N/D      |
| Conceptual Framework (CF)   | CS, DO       | CIA, HIPAA |
| HealthVault                 | CS           | Authentication |
| healthTicket                | CS           | CP-ABE   |
| DEPR                        | DC           | N/D      |
| My HealtheVet               | DE           | Security Policies |
| SNOW                        | DC           | Privacy Policies |

N/D: not defined; DE: distributed electronic; P2P: peer-to-peer; CS: client-server; DO: distributed object; CIA: confidentiality, integrity, and Availability; HIPAA: health insurance profitability and accountability; CP-ABE: ciphertext-policy attribute based encryption; DC: distributed components.

Fig. 2  The objective of blockchain in several use cases [35].

EHR & PHR, insurance & market, and internet-of-things & monitoring. Several healthcare providers achieve the main objectives of implementing blockchain technology using different platforms and techniques [36]. The benefits of decentralized e-health such as transparency, decentralized value transfer, interoperability, and immutable data history records are still the main goals to be attained.

5. Operating-System Design and Implementation

We now discuss the operating-system design and our approaches implementation to enable patients-defined data privacy management in a decentralized e-health environment. This section is divided into four categories. Firstly, We put forward the state of the art that presents PDPM at a high level. Secondly, the paper elaborates data privacy classification management, and nudging with collaborative filtering model is also given afterwards. Thirdly, the paper presents blockchain technology through Ethereum smart contract as immutable data storing services. Finally, concerns and remarks are concluded in the last section.

5.1 The State of the Art of PDPM

PDPM is a framework for realizing patient-defined data privacy using the nudge theory concept in decentralized e-health environments. Our approach focuses on managing the patient’s data privacy, classification and utilizing the concept of nudge theory to obtain better services and recommendations for the patient inspired by research in [37]. However, the final decision remains entirely up to the patient to organize and determine his medical data accessible by certain groups. The output data from the nudging results are then stored on-chain via an Ethereum smart contract so that authorized parties can find out the historical records of the data privacy created by the patient. Meanwhile, PDPM leverages end-to-end encryption techniques for communication, where only the communicating e-health entities can read the messages.

Figure 3 illustrates the state of the art of the PDPM framework in enabling patient-defined data privacy. The PDPM system commences by storing encrypted data by patients or other authorized entities that use cloud-based storage services. By design, the data are collected by authorized parties whose access is governed by the patient or the personal doctor representing the data owner. To deploy the PDPM framework, the authorized doctors or stakeholders roughly classify the patient’s data privacy by default, such as lifelong information, medical activities, insurance, and so forth.

The initial classification of patient data makes PDPM more manageable and convenient to use in a collaborative filtering-based algorithm so that the algorithm’s output can be more precise. Collaborative filtering protocol can be used in the PDPM framework as a filtering method used by the recommender system to make automatic predictions about patient privacy management and preferences from any inputs of data. Several inputs of data are derived from the different number of patients with their respective data that has been processed beforehand. The underlying premise of the collaborative filtering in PDPM is that if patient X has
the same interest as patient Y on a use case, patient X is more likely agreeing with Y’s selection on a different use case than that of a randomly chosen patient. Finally, the nudge theory’s concept makes the collaborative filtering output more likely that patients will perform a particular choice or behave distinctly by adjusting the conditions so that automatic cognitive rules are triggered to favour the craved outcome.

5.2 Data Privacy Classification and Nudging with Collaborative Filtering Model

Healthcare data over the internet is unique and challenging to measure. Combining e-health data from multiple sources and the interpretation of that information is essential to optimize patient care. A decent e-health data improves communication between healthcare entities with an in-depth understanding of particular health conditions, insurance, long-term planning, and so forth. Therefore, determining the correct data to be filtered before being processed using nudge theory needs to be considered, such as applying a collaborative filtering protocol.

Collaborative filtering protocol in the PDPM system can be defined into two steps. Firstly, the PDPM searching for patients who participate in the same rating patterns with the active patients (the patient whom the prediction is for). Secondly, the PDPM applying ratings from those like-minded patients observed in the earlier step to compute a forecast for the active patient. The memory-based types of collaborative learning rely on patient rating data to execute patient information similarity. Representative models of this method are neighbourhood-based CF with patient-based top-N recommendations. For instance, in patient-based procedures, the value of rating patient u gives to item i is determined as an aggregation of some similar patient’s rating of the item is \( r_{ui} = \text{aggr}_{u \in U} r_{uij} \), where \( U \) depicts the set of top N patients that are most similar to patient u who rated item i [38]. Remarkable models of the aggregation function include:

\[
    r_{ui} = \frac{1}{N} \sum_{u' \in U} r_{uij} \quad (1)
\]

\[
    r_{ui} = k \frac{1}{N} \sum_{u' \in U} \text{simil}(u, u') r_{u'ij} \quad (2)
\]

Where \( k \) is a normalizing factor defined as \( k = 1/\sum_{u' \in U} \text{simil}(u, u') \). The neighbourhood-based algorithm computes the relationship between two patients or items while producing a forecast for the patient by considering the weighted average of all the ratings. In this approach, models are developed using different data mining, machine learning algorithms to predict patients’ rating of the unrated medical data item. There are many model-based CF algorithms. Bayesian networks, clustering models, latent semantic models such as singular value decomposition, probabilistic latent semantic analysis, multiple multiplicative factor, latent Dirichlet allocation, and Markov decision process based models. One benefit of applying this method for the PDPM system is that instead of possessing a high dimensional matrix consisting of abundant missing values, we will deal with a much less matrix in lower-dimensional space. A reduced presentation could be employed for either user-based or item-based neighbourhood algorithms. Likeness computation between items or patients is essential in the PDPM system. Various models, such as Pearson correlation and vector cosine-based similarity, can be adapted to achieve a better result. Eventually, encrypting the non-disclosure patient information can be described as follows:

1. Input the computed disclosure schemes include the default schemes.
2. Output is encrypted patient’s information.
3. Check the computed disclosure schemes, if it equals non-disclosure, then go to step 2, otherwise go to step 4.
4. Search the nearest smart ledger of blockchain, if computed disclosure schemes are changed then go to step 3, otherwise go to step 4.
5. Using the blockchain proprietary key to encrypt the patient’s information.
6. Update patient’s information to the chain for consensus confirmation.

Medical data can be classified into several types as shown in Table 4. There are several types of healthcare-related data. Data such as patient clinical records information, genetic analysis information, biometric data collected through smart devices, family history related to diseases, and personal health and examination information stored in public institutions are collected and stored by various management entities. In particular, with the advent of

| Data type | Examples |
|-----------|----------|
| Medical treatment data | EMR, EHR, prescription information, hospitalization and discharging the hospital, medical image data (CT, MRI, CR, etc.) |
| Clinical research data | Drug clinical trial data, device clinical trial data, genetic study data, human origin study data, survey observation data, research data directly or indirectly using personal information |
| Public Institution data | Data collected, stored and managed by public institutions, such as insurance premium-related data, medical treatment details, health examination results, death information, etc.; |
| IoT based data | Medical devices data and patient monitoring device-based data |
| Omics data | Molecular level data like Genome, Transcriptome, Proteome, Metabolome, Epigenome, Lipidome |
| Lifelog data | Personal record of one’s daily life from wearable device (weight, heart rate, blood sugar, personal eating habits, medication, behavior and mental data) |
| Mobile application and social media data | Various healthcare-related data collected from social media |

---

Table 4: Healthcare big data types [39].
smartphones and various wearable devices, large amounts of data have been pouring out in unstructured forms, and the type, amount, and generation speed of data have been rapidly increasing. Input the computed disclosure schemes include the default schemes.

5.3 The Embodiment of Immutable Data Privacy Management

This section describes the Ethereum smart contract’s application to store the latest settings of personal patient data obtained from the nudging technique with a collaborative filtering model. We use a private Ethereum blockchain where only authorized entities can view transactions made in the e-health system. The information stored on-chain is not sensitive to the patient, so stakeholders or doctors can use the data for further processing such as analysis with artificial intelligence (AI) technology, insurance recommendation, a description of the following medical action, etc.

We utilize the Ethereum virtual machine (EVM) as the heart of the Ethereum protocol and operation that acts as a computation engine in executing the blockchain transactions. At the high level, the EVM has a stack-based architecture with a distinct state transaction function shown in Fig. 4, which can store all in-memory values on a stack. The EVM operates with a word size of 256 bits that support native hashing function and elliptic curve operations to secure the transactions. Eventually, the EVM has different addressable data components as follows:

(i) A tamper-proof program code read-only memory (ROM), as a type of non-volatile memory in the Ethereum platform. It is loaded with the bytecode of the arbitrary value functions to be executed.

(ii) A volatile memory, with each location, notably initialized to zero.

(iii) Perpetual storage that is an element of the Ethereum state. This sort of storage is also zero-initialized.

For ease of presentation, we leverage the format of electronic health record (EHR) data in clinical practice (see Fig. 5) as arbitrary values input in smart contracts. This EHR’s format is the patient’s final decision after determining without intervention for the types of private and public data that the public can access. The format adopted of patient’s private data can vary depending on each system that implements decentralized e-health as presented in Table 1 and Table 3. This sort of arbitrary data becomes an input in the smart contract to be stored and seen publicly. Furthermore, to implement Ethereum smart contracts, we use the Ganache-Truffle suite to run tests, execute commands, and investigate the state of blockchain transactions (blocks information, transactions, and logs).

The account addresses of the e-health entities are created by Ganache graphical user interface (GUI) that consists of the public address and secret key. Ganache leverages ethereumjs to simulate full client commands. We use the default Ganache setting running on an RPC server HTTP://127.0.0.1:7545, with a network ID of 5777, and the mining status is in “automining” mode. For gas prices (20000000000 units), gas limits (6721975 units), and Ether’s balance (100.00 ETH) for each entity are automatically governed by Ganache GUI. To manage entities’ wallets more efficiently, we use MetaMask (metamask.io) as a crypto wallet and gateway to a decentralized e-health application.

Figure 6 depicts the activities information on e-health transactions carried out by patients on the blockchain network through smart contracts that have been installed on their respective devices. We divide the e-health entity into several various roles, namely hospital (Hs A) as the smart contract owner, patient A (Px A), and patient B (Px B). The first transaction recorded is a transaction for contract migration to the blockchain so that each entity can use the contract provided by Hs A. Each smart contract that has been migrated has its own identity to differentiate between one contract and another within the same blockchain networks.
Our smart contract has a specific function to accommodate arbitrary input values in the form of the patient’s EHR data. This data is the patient’s final decision after getting the best recommendations from nudge theory and collaborative filtering protocol. The patient is candidly to accept, change, or even delete the potential recommendations provided by PDPM without any intervention from other e-health entities. Meanwhile, Fig. 7 is another perspective of our e-health transaction information. The “created contract” function transaction requires 225213 units of gas, which occurs in block number 1. Meanwhile, the patient’s transaction process is recorded in block number 2 with the use of gas amounting to 42363 units. The amount of gas consumed is relatively low since our e-health smart contract is a GUI that can hold arbitrary values from patient EHR data without any complex commands functions.

The Ethereum gas usage in a transaction is essential to take into account because it relates to the number of costs to be paid by the transactor. We recorded as many as twenty transactions for two patients (Px A and Px B) and ten transactions for the healthcare provider (Hs A), who also acted as the patient’s incentive provider. In this sense, the patients are rewarded with a certain amount of Ether if successfully storing their latest public information to the on-chain. In Px A’s case, ten transactions were conducted with insignificant information changes (addition and deletion of arbitrary inputs). The first transaction was the smallest amount of input. In the second transaction, we added additional input and so on. Thus, the transaction with the most significant information occurs on the last transaction (Tx-10th).

A summary of all transactions that occurred can be seen in Fig. 8. We recorded the amount of gas usage by Px A, Px B, and Hs A. For Px A and Px B, the data used were identical to the number of arbitrary inputs, which did not differ much from each other. In the first transaction, we noted that Px A required 96371 units of gas to update the settings from PX A information. The smallest gas usage occurred in the first transaction, with the most gas usage occurring in the last transaction, which consumed 158426 units with an average usage of 125309 units. Meanwhile, the transactions carried out by Px B are not much different from Px A. It is recorded that the minimum gas usage amount occurred in the first transaction, which is 96208 units, and the maximum occurs in transactions to Tx-10th with an average gas of 66141 units. The difference is very significant if the standard format of e-health used as input data in the smart contract is distinctive.

Figure 9 represents the amount of incentive distribution...
for patients in cryptocurrency, which is embodied in Ether. Cryptocurrency is given to patients to motivate each patient to provide the best data management. In other words, the patients jointly contribute to improving the PDPM system. The number of incentives awarded varies according to the contribution that has been given. Incentive policies are absolute rights of the system provider or smart contract owner, and the transaction process will continue as long as the owner does not revoke the contract. Overall, Ethereum smart contracts have been implemented to support the PDPM system. A manageable yet efficient smart contract is needed in the e-health system to minimize transaction costs.

5.4 Concerns and Remarks

At a high level, PDPM is a platform that preserves patient-data privacy management with nudge theory applied in decentralized e-health environments. In the previous sections, we have outlined the advantages of PDPM to be implemented in the e-health system. However, our approach is in separable from several challenges and concerns both a technical and theory perspective.

5.4.1 Large-Scale Medical Data

Implementing the PDPM system requires a massive number of e-health data from various patients, stakeholders, or other e-health entities with multiple categories. Meanwhile, collaborative filtering works properly with sufficient information collected from various users and healthcare stakeholders. However, the PDPM system is a new recommendation system, which can cause cold-start problems due to a lack of data. In other words, it is a challenge in forming new user trends before observing the effectiveness of filtering. The combination of data is used in the collaborative filtering protocol and nudging algorithm to provide the best patient recommendations regarding data privacy settings. Data collection techniques are also a challenge in healthcare which is also envisioned by advancing the internet-of-things with affordable sensors and devices. To effectively collect sensor data, a decent model is needed to consistently handle the access of heterogeneous sensor devices and various sensor data types. Sensor data collection is a process that must precede the construction of environmental information. In this research, we assume data collection techniques are safe without using a particular protocol. Therefore, the PDPM system requires efficient, stable, and secure data collection techniques to support the goals in the near future.

5.4.2 Nudging Data and Collaborative Filtering

There are also issues related to the enforcement of the nudge strategy. When the nudge strategy is used in the e-healthcare sector, it is necessary to consider ethics and personal choice issues first carefully. Since nudge interventions are designed to affect automatic systems, attempts to change someone’s behaviour can be regarded as challenging role responsibilities for medical personnel or system designers and can be perceived as manipulative. Therefore, a nudge strategy can cause limited bias in human behaviour. In some cases, certain policy objectives are attempted to be achieved using bias, and in some cases, the bias itself is corrected to help the subject make the right choice from the perspective of the policy designer. Those who design nudge policies cannot be free from limited rationality and bias; thus, they cannot guarantee that nudge policies help people make choices and improve their well-being. After all, it is not easy to understand that imperfect people correct their behaviour with policies designed by imperfect people. It is also a counterargument to randomized controlled trials (RCTs) used in the development phase of the nudge strategy. This means that the responses of subjects to nudge policies do not reflect reality properly. Also, most of the nudge strategy experiments published in Western economics are based on the high educational standards, industrialization and democratic political system of Western countries, raising the question of whether this conclusion of behavioural economics can be applied to an international environment. Eventually, the improvement schemes of nudge theory in the medical area can be seen in [41] and [42].

5.4.3 Transparency Concerns

One of the prominent features that blockchain offers is transparency. The public blockchain realizes transactions that are open to anyone to see the activities that occur. Transparency is one of the concrete features of the SC blockchain. However, this feature is not desirable for various cases, such as for sensitive PDPM data. Addresses representing the digital asset owner (pseudonymous) can be seen by the public freely without being directly involved in the system. This feature is beneficial for knowing where cryptocurrency comes from, to which address cryptocurrency is spent in the blockchain. Thus, there is no ambiguity for the user in conducting transactions. However, the advantage of transparent transactions is undesirable in various cases, such as in several issues relating to sensitive and confidential data like most cases in e-health environments. Therefore, transparency data in decentralized e-health needs to be investigated further before blockchain is utilized in the PDPM system, as elaborated in [43] and [44].

5.4.4 Scalability Issues

PDPM utilizes Ethereum smart contract platform to store the record of patient’s decisions about their respective data privacy. The use of the Ethereum platform is still inseparable from scalability issues. The miners in the Ethereum network have to compete to find the nonce to satisfy the target difficulty. At the same time, each node must verify that the miner’s work is valid and maintain a complete copy of the current network state. This process dramatically limits the Ethereum transaction process capability and
throughput, where Ethereum can only process roughly around 12-15 transactions per second. Moreover, the scalability issues cannot be solved straightforwardly by directly increasing the block size and the consensus difficulty since it affects the security and decentralization of blockchain. This concern is also known as blockchain scalability trilemma. This concern also affects e-health environments. However, these concerns can be addressed by adopting the strategy stated in [45]. The estimated amount of patients with their corresponding data must be carefully calculated to prevent overhead and bottleneck issues.

6. Conclusions

We have presented the PDPM as a patient-defined data privacy management using nudge theory in decentralized e-health environments. Our objective focuses on designing state of the art PDPM empowered by nudge theory and blockchain technology in providing immutable patient’s privacy management in an e-health system. The PDPM system we have designed is the initial stage of the overall design we want to achieve. This research converges on the opportunities and drawbacks of combining several technologies to accomplish the PDPM’s goals. As part of the embodiment of immutable data privacy management, Ethereum smart contracts promise to be implemented in the PDPM system due to the low transaction costs yet still having exemplary services. However, the blockchain scalability trilemma must be considered more profoundly to balance security, decentralization, and scalability. Thus, we need a detailed estimate of the number of entity members of the PDPM system for future work. More precisely, our future work will be focusing on nudging real patient data with a variant of collaborative filtering protocol and designing a more efficient Ethereum smart contract by minimizing the arbitrary inputs in the smart contract.

Acknowledgements

This research was supported by the Republic of Korea’s MSIT (Ministry of Science and ICT), under the High-Potential Individuals Global Training Program (2020-0-01596) supervised by the IITP (Institute of Information and Communications Technology Planning & Evaluation) and partially supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IOTP-2020-0-01797) supervised by the IITP (Institute of Information & Communications Technology Planning & Evaluation).

References

[1] R. Kumar, W. Wang, J. Kumar, T. Yang, A. Khan, W. Ali, and I. Ali, “An integration of blockchain and ai for secure data sharing and detection of ct images for the hospitals,” Computerized Medical Imaging and Graphics, vol.87, p.101812, 2021.
[2] S. Becher, A. Gerl, B. Meier, and F. Bölz, “Big picture on privacy enhancing technologies in e-health: a holistic personal privacy workflow,” Information, vol.11, no.7, p.356, 2020.
[3] Y. Hong, T.B. Patrick, and R. Gillis, “Protection of patient’s privacy and data security in e-health services,” 2008 international conference on biomedical engineering and informatics, vol.1, pp.643–647, IEEE, 2008.
[4] C. Butpheng, K.-H. Yeh, and H. Xiong, “Security and privacy in iot-cloud-based e-health systems—a comprehensive review,” Symmetry, vol.12, no.7, p.1191, 2020.
[5] A. Roehrs, C.A. Da Costa, and R. da Rosa Righi, “Omniph: A distributed architecture model to integrate personal health records,” Journal of biomedical informatics, vol.71, pp.70–81, 2017.
[6] CCR, “Standard specification for continuity of care record,” https://www.ncbi.nlm.nih.gov, 2021.
[7] CEN, “European committee for standardization - cEN/TC 251 - health-informatics,” https://standards.cen.eu, 2021.
[8] DICOM, “Digital imaging and communications in medicine,” https://www.dicomstandard.org/, 2021.
[9] J.C. Mandel, D.A. Kreda, K.D. Mandl, I.S. Kohane, and R.B. Ramoni, “Smart on thir: a standards-based, interoperable apps platform for electronic health records,” Journal of the American Medical Informatics Association, vol.23, no.5, pp.899–908, 2016.
[10] HIPAA, “Health insurance portability and accountability act,” https://www.hhs.gov/hipaa/index.html, 2021.
[11] ICD, “Family of international classification,” https://www.who.int/standards/classifications, 2021.
[12] IPC, “International classification of primary care,” http://www.globalsfamilydoctor.com/, 2021.
[13] IHE, “Integrating the healthcare enterprise,” https://www.ihe.net/, 2021.
[14] ISO, “Iso/TC 215 - health informatics,” https://www.iso.org/committee/54960.html, 2021.
[15] LOINC, “Logical observation identifiers names and codes,” https://loinc.org/, 2021.
[16] openEHR, “openehr - an open domain-driven platform for developing flexible e-health systems,” https://www.openehr.org/, 2021.
[17] SNOMED-CT, “Systematized nomenclature of medicine,” https://www.snomed.org/, 2021.
[18] S. Rahmatidka and K.-H. Rhee, “Blockchain technology for providing an architecture model of decentralized personal health information,” International Journal of Engineering Business Management, vol.10, p.1847979018790589, 2018.
[19] G. Tripathi, M.A. Ahad, and S. Paiva, “2sbh-a blockchain based approach for smart healthcare system,” Healthcare, vol.8, no.1, p.100391, Elsevier, 2020.
[20] M.M.H. Onik, S. Aich, J. Yang, C.-S. Kim, and H.-C. Kim, “Blockchain in healthcare: Challenges and solutions,” Big data analytics for intelligent healthcare management, pp.197–226, Elsevier, 2019.
[21] A.A. Abdellatif, A.Z. Al-Marridi, A. Mohamed, A. Erbad, C.F. Chiasserini, and A. Refaey, “sshealth: Toward secure, blockchain-enabled healthcare systems,” IEEE Netw., vol.34, no.4, pp.312–319, 2020.
[22] N.J. Witchey, “Healthcare transaction validation via blockchain, systems and methods,” US Patent 10,340,038, July 2 2019.
[23] C.C. Agbo, Q.H. Mahmoud, and J.M. Eklund, “Blockchain technology in healthcare: a systematic review,” Healthcare, vol.7, p.56, Multidisciplinary Digital Publishing Institute, 2019.
[24] S. Khezri, M. Moniruzzaman, A. Yassine, and R. Benlamri, “Blockchain technology in healthcare: A comprehensive review and directions for future research,” Applied sciences, vol.9, no.9, p.1736, 2019.
[25] H.M. Hussien, S.M. Yasin, S. Udzir, A.A. Zaidan, and B.B. Zaidan, “A systematic review for enabling of develop a blockchain technology in healthcare application: taxonomy, substantially analysis, motivations, challenges, recommendations and future direction,” Journal of medical systems, vol.43, no.10, pp.1–35, 2019.
[26] Q. Su, R. Zang, R. Xue, and P. Li, “Revocable attribute-based signature for blockchain-based healthcare system,” IEEE Access,
A.A. Abdellatif, L. Samara, A. Mohamed, A. Erbad, C.F. Chiasserini, M. Guizani, M.D. O'Connor, and J. Laughton, "Blockchain-enabled 5G edge networks and beyond: An intelligent cross-silo federated learning approach," Security and Communication Networks, vol.2021, 2021.

B. Explorer, "Total number of recorded addresses with balance," https://www.blockchain.com/explorer, 2021.

H. Shafagh, L. Burkhhalter, S. Ratnasamy, and A. Hithnawi, "Droplet: Decentralized authorization and access control for encrypted data streams," 29th USENIX Security Symposium (USENIX Security 20), pp.2469–2486, 2020.

B. Houtan, A.S. Hafid, and D. Makrakis, "A survey on blockchain-based self-sovereign patient identity in healthcare," IEEE Access, vol.8, pp.90478–90494, 2020.

A. Farouk, A. Alahmadi, S. Ghose, and A. Mashatan, "Blockchain platform for industrial healthcare: Vision and future opportunities," Computer Communications, vol.154, pp.223–235, 2020.

H. Wang, S. Ma, H.-N. Dai, M. Imran, and T. Wang, "Blockchain-based data privacy management with nudge theory in open banking," Future Generation Computer Systems, vol.110, pp.812–823, 2020.

Wikipedia, "Collaborative filtering," https://en.wikipedia.org/wiki/collaborative_filtering, 2021.

K. M. of Health and Welfare, “6th health and medical big data policy review committee, 2021,” http://www.mohw.go.kr/, 2021.

J. Man-Ki, “Building a smart healthcare ecosystem,” https://www.korchem.net/nchem/service/economy/appl/, 2017.

S.L. Yoong, A. Hall, F. Stacey, A. Grady, R. Sutherland, R. Wyse, A. Anderson, N. Nathan, and L. Wolfenden, "Nudge strategies to improve healthcare providers’ implementation of evidence-based guidelines, policies and practices: a systematic review of trials included within cochrane systematic reviews," Implementation Science, vol.15, no.1, pp.1–30, 2020.

R.E. Glasgow, C.E. Knoepke, D. Magid, G.K. Grunwald, T.J. Glorioso, J. Waughal, J.C. Marrs, S. Bull, and F.M. Ho, "The nudge trial pragmatic trial to enhance medication adherence: study protocol for a randomized controlled trial," 2021.

P. Pandey and R. Lioriya, "Securing and authenticating healthcare records through blockchain technology," Cryptologia, vol.44, no.4, pp.341–356, 2020.

I.A. Omar, R. Jayaraman, K. Salah, M.C.E. Simsekler, I. Yaqoob, and S. Ellahham, "Ensuring protocol compliance and data transparency in clinical trials using blockchain smart contracts," BMC Medical Research Methodology, vol.20, no.1, pp.1–17, 2020.

A.A. Abdellatif, L. Samara, A. Mohamed, A. Erbad, C.F. Chiasserini, M. Guizani, M.D. O’Connor, and J. Laughton, “Medgechain: Leveraging edge computing and blockchain for efficient medical data exchange,” IEEE Internet Things J., 2021.