Towards an Architecture-Based Ensemble Methods for Online Social Network Sensitive Data Privacy Protection

https://doi.org/10.3991/ijes.v9i1.20819

Felix O. Idepefo (✉), B.I. Akhigbe
Obafemi Awolowo University, Ile-Ife, Nigeria
felixidepefo@gmail.com

O.S. Aderibigbe
Lagos State Polytechnic, Lagos, Nigeria

B.S. Afolabi
Obafemi Awolowo University, Ile-Ife, Nigeria

Abstract—In 2014, the world woke up to a giant data breach that leveraged users' personal information that was taken from one of the world's biggest social network platform. Based on the literature, this was possible because of the Centralised Architectural-based Approach to protecting the privacy of users' online data. Although the literature is inundated with decentralized approaches, there is none to the best of our knowledge that uses an ensemble of methods and draws on a consensus mechanism to address the challenges caused by the Centralised Architectural-based Approach. This paper presents a decentralized approach that adopts and adapts an ensemble of methods. These methods include cryptographic, hashing, and the plenum byzantine fault tolerance algorithms that present a consensus platform, protocol, and mechanism to use the technology of blockchain in a novel manner as a significant contribution. This paper adopts the descriptive approach in its presentation as the usable implementation of the presented proposal is near completion with issues of computational overhead addressed based on preliminary results that show promise of being able to support agreement up to 75% in terms of making changes by participants in the chain.

Keywords—Sensitive data protection model, Online social network, Blockchain technology, Cryptography, Consensus mechanism

1 Introduction

The emergence of Web 2.0 and the development of the Internet introduced a new paradigm in the exchange of user-generated contents [1, 3]. Web 2.0 remain a critical network infrastructure and knowledge platform for entities - man, machine, group, and even brain-like computer - to exchange and share information, knowledge, wisdom and data [3]. One of the most remarkable phenomena that blossomed in the Web 2.0 era is Online Social Networks (OSNs) that include Facebook, MySpace, Whatsapp, Instagram, Netlog, LinkedIn, etc. [2, 3]. OSNs organize Social Interactions and Related Activities (SI&RA). It also sustains the management of SI&RA and facilitate the emergence of a new virtual societal workspace that supports sundry human, business and even scientific activities the world over [3]. As a digital communication tool; OSNs allow users to leverage their profiles, virtually represent themselves and declare explicitly their relationships (or connections) with other users [4]. The most representative Social Networking System (SNS) is Facebook. It has about 2 billion users that are active with daily user connections that are the highest among other SNS [5,6]. Twitter, is another SNS with micro-blogging potentials with over 313 million monthly active users who tweet in more than forty (40) languages [7]. OSNs provide several SNS services that offer users the opportunity to build public profile, among registered users look up new friends, establish relationships, and share contents with the possibility of growing into communities based on common interests [5]. These social communities are Open Virtual Spaces (OVS) for autonomous exchange of information. They afford millions of people of all ages and backgrounds across the globe the opportunity to connect with each other.

This virtual connectivity is open and allows personal and social expressions that span across geographical borders with the deliberate sharing of information. However, the consequences of autonomously exchanging information in OVS on users' privacy remain vague and obscure to most OSN users [8]. The Centralized Architectural Approach (CAA), which OSNs employ to provide their services are responsible for these online malicious activities. OSNs like Facebook act as a central authority and thus exercise autonomous control over users' information based on the CAA. Consequently, huge amount of users' private and sensitive data (or information) that contain confidential interactions are unprotected [4]. This highlights the major challenge with the practice of using the centralized architecture, which is the violation of users' privacy. One of such violation is the commercial gains made from the data that are stored in siloes (or centralized servers) [5]. The CAA employed by OSNs supports polices that make it possible for users to share information in such a way that the risk of censorship, surveillance and the revealing of information without the consent of users' is inevitable [4,9]. The challenges caused by the CAA, which is evident in the literature are solvable user-oriented concerns that are further exacerbated by the fact that users are rarely even aware of the amount of data and meta-data that are collected about them. What is even more disquieting is the understanding from the literature that users are oblivious of the value of their data and the sensitive information in them [10].

This paper proposes an architectural-based ensemble of methods that leverage the blockchain technology to protect the privacy of Online Social Network Sensitive Data, which the existing centralized architectural-based approach does not cater for. This objective was addressed by (i) proposing a Sensitive Data Protection Model Architecture (SDPMA), (ii) modelling the sensitive data protection scenario, and (iii) presenting implementable UML Object-Oriented (OO) formalisms of the SDPMA. This paper contributes an architectural-based ensemble of methods that operate as a consensus mechanism to use the blockchain technique to protect the privacy of users' sensitive data. This contribution uses a decentralized approach unlike the CAA and also addresses the challenges of the Free-service provisioning capability of the OSNs model that supports targeted and retargeted marketing intentions that makes it easy for malicious users to target users' sensitive data [11-13]. Therefore, it will be difficult to give the users of OSN a false hope of control over their privacy of data [5]. Data breaches such as the one that resulted in the harvesting of millions of Facebook profiles - users' personal information - at the beginning of 2014, which was stolen for political advertisement can be checked [14]. The contribution in this paper is significant since the proposed method uses low-cost users' action-oriented attributes. These attributes are personal information of users delineated in a three-some manner - personal identifiable user information, potential personal identifiable user information, and users' posts. This is easy to come by since they appear un-useful on the wall of social media platforms, but they are really the basic items that are needed to compromise and leak user data and commit identity theft [15]. The rest of the paper is structured as follows, with Sections 2.0, 3.0. 4.0, 5.0 and 6.0 dedicated to the review of literature, Methodology, results and discussion, system implementable formalism, and the paper's conclusion respectively.

2 Literature Review

2.1 State-of-the-art in online social network sensitive data protection

In the past few years, there has been significant growth and improvements in the services offered by OSNs [15, 16-17]. Some Studies [18-19] have shown that users' private life are often in jeopardy whenever they post sensitive information when on any of the OSN space. It is even more worrisome that a great number of users of OSN tool are unaware of the importance of protecting their privacy [20]. Reports have shown that shared information on OSNs can reveal contents that are meant to be private and sensitive enough not to be published since malicious users can use them to invade individual's privacy [21-23]. Both security and privacy concerns have been highlighted as a major challenge with OSNs in the literature since they are built on the centralized architectural philosophy [24-25]. Efforts are rife in literature with the aim of logically decentralizing the functionalities of OSNs and mitigate privacy issues. Based on the literature, decentralized architectures can be implemented using multiple independent and trusted servers [2, 13, 30]. Some of these efforts used federated architectures as proposed in [26]. This approach used an architectural framework that protects users' privacy by shielding users' personal posts or messages from service providers and other third-party applications who are not authorized by users to view the content. In a similar effort by [27], the federated and decentralized social network used users' profiles to help individual users decide for themselves where their information should be stored. Usually, every user-generated content were encrypted using a random key, which in turn is distributed to every authorized user. The approach that was applied by [27] and [26] using the attributes of users' personal posts and users' profiles respectively was leveraged in [28] and [29] based on open-source to offer microblogging functionalities. In these efforts, users' identity played a major role as the attribute that was employed in the formulation of their federated architectural framework. However, the Federated Architectural Approaches (FAA) are vulnerable to information leakage. The FAA is porous, intruders easily carry out data breaches and other malicious attacks and abuses from central service providers is high.

In the literature, an alternate approach using the popular P2P architecture was applied to decentralize the OSNs with trust as the main challenge that was addressed. In a recent work, [31] proposed a decentralized approach that is built upon an overlay and relies on trusted nodes to ensure the security of the network. However, this approach is weak towards the detection of unauthorized users who could use fake profiles and spam messages to initiate security breaches. In another related work by [32], a trust-aware model was developed to securely shared knowledge using the Distributed Hashing Technique (DHT) and a predecessor replication technique that rely on social trust. The model allows trusted friends to be admitted with continuous security from unsuspecting malicious nodes. The DHT with static replication technique has also been used to secure the storage of static bulk data (videos, photo albums) from their basic profile information or social glue [30]. An ensemble of Encryption, decentralization and direct data exchange has been applied to solve privacy and connectivity problems [33]. The Open-DHT is a variant of the DHT method, which implementation is suitable for look-up service prevision. Based on the literature, DHT has been a useful technique to mitigate attacks from malicious nodes [34]. Sometimes, the technique is adapted to support users by anonymizing communications as well as replicate contents and profile information to trusted nodes. With DHT, low latency and high data availability depending on the number of trusted friends within a social connection is a critical issue. The drawback with the use of DHT is that it is difficult for users who maintain few social connections to maximize its potentials [34]. Aside from these successes, some other research efforts concerning the protection of sensitive data have attempted to use both the federated and P2P methods. [35] developed an approach that uses the Ciphertext-Policy Attribute-Based Encryption (CP-ABE) toolkit and Google Drive to hold encrypted messages. A cloud-backed P2P with decentralized and encrypting capability for personalized online social networking was proposed by [36]. In a related work, an infrastructure surrogate with content key that is symmetrically random, and in turn is encrypted with a proper ABE key was developed by [25]. However, based on findings in [5], third party platform like the cloud does not guarantee satisfactory privacy of user data. The same goes for federated and hybrid P2P architectures that also rely on third party policy by cloud providers with fake privacy assurance. Though, the P2P method allow users' data to be stored on DHT and home gateway, trust still remain a concern even with high latency challenges among peers [5].

2.2 Review of blockchain techniques

Blockchain Technology (Bloc-Tech) has been used as an immutable distributed ledger to resolve the trust issue in the P2P method [37]. Available and successful research efforts are in the literature that highlights satisfactory promise concerning the use of Bloc-Tech to ensure privacy, trust and availability of sensitive data among untrusted peers in a network. The work of [38] demonstrated this by proposing a decentralized technique based on Bloc-Tech that uses Ethereum Blockchain and Proof of Work (PoW) consensus algorithm. This approach was used with a 51% success rate to ward off attacks when managing a photo group that uses a decentralized social media Web-based photo sharing application. The potential of Bloc-Tech was also exhibited in the research work of [39]. They proposed a decentralized approach that a Delegated Proof of Stake (DPoS) consensus protocol. Similarly, [40] leveraged the technique of blockchain to develop a social networking service provisioning system with irrevocable peer review records and traceable reputation structure to distribute content. It was found from these research work (e.g. [37-40] that the bitcoin-based Bloc-Tech that was used, although it uses the proof of reputation consensus algorithm, it is permission-less with public inclinations. Therefore, it is prone to weak consistency, low transaction throughput, and vulnerable to malicious attacks, that include double spending attacks, eclipse attacks, and selfish-mining. The alternate PoW consensus algorithm that the Bloc-Tech implementation already stated employs encourages high computational power wastage and it is subject to selfish mining by intelligent miners. The DPoS consensus protocol by [39] was supposed to ensure decentralization, but in reality, this was traded for scalability, which can only support few more users. The need for a consensus technique that is robust and scalable like the Bloc-Tech is thus overarching. In [45] a secured network solution that enforces data control and overcome privacy concerns, and security compromises through blockchain is proposed. The solution is a decentralized solution based on the descriptions presented. However, the technique of enforcing the solution was not provided. On the contrary, we employed the Indy genre of Hyperledger to enforce self-sovereign identity [50]. Unlike the possibility of the model in [45] to provide a dais for authorities to assist users with privacy, this current work applied the Indy technicalities to forestall this. The plausibility of the method in [45] remain cynical since no clue was presented regarding the validation of their method.

3 Methodology

3.1 The sensitive data protection model architecture

The Sensitive Data Protection Model (SDPM) is presented as an ensemble architectural-based method. This archetypical model is a Decentralised Application (DApps) that enables online users to communicate with the Blockchain to manage the state of network actors. At the backend of the Dapps, the models business logic is represented by one (or several) smart contracts that interact with Blockchain technology. The frontend is made up of decentralized storage networks, which is hosted on an inter-planetary file system. To manage cryptographic keys, a wallet is used to house the distributed identifier and the blockchain addresses (see Figure 1). The DApps interacts back and forth with the CP-ABFHE module. Here, a hybrid cipher-text policy and fully homomorphic encryption algorithm will encrypt the data of users that are stored in the Local Database (L-Dbase). Friend Recommendation (FR) is an essential part of the OSN platforms and will be implemented in the FRM. This module uses an attribute-based community detection algorithm built on community discovery and attribute dependency to satisfy the FR requirements of OSNs and to allow collaboration between trusted friends. The FRM therefore interacts with the CP-ABFHE module and the L-Dbase and through the L-Dbase with the BlockChain Module (BCM). The BCM houses the Hyperledger Indy BlockChain (HIBC). The HIBC interacts with the ChainCode (C-C), which is a "smart contract" that creates transactions while running on the peers and update the World state of the Assets (WsotA). The WsotA is in the Global Database (G-Dbase). A Secure Hashing Algorithm (SHA-256) is applied to strengthen the security of the already encrypted user's data from the CP-ABFHE module through the L-Dbase that is now stored in the G-Dbase in the HIBC module. The SHA-256 also acts as a compressing technique that reduces the size of the encrypted data stored in the G-Dbase. The choice of the SHA-256 is premised on its ability to be computationally infeasible for potential malicious nodes (or users) on a network. The role of the Plenum Byzantine Fault Tolerance (PBFT) algorithm is to provide the consensus mechanism to vote based on a consensus protocol in the HIBC module to add validated transactions to the mechanism of Blockchain. The PBFT algorithm enable validator nodes to take part in the process of voting to bring in the next block till there is a consensus. This consensus must be among more than two-thirds of the validator nodes that agree before a new block is added to the chain. The choice of the Hyperledger Blockchain technique stems from its permission-orientedness as a distributed ledger to provide tools, libraries, and reusable components that is purpose-built to allow the decentralization of identity.

3.2 Model validation of proposed method

The blockchain model from the SDPMA was validated for effectiveness by using the Evaluation Framework for Blockchain Hyperledgers (EFBH) based on the provisions in the literature [46, 47]. Following documented best practices in the literature [46, 47, 48] regarding the use of EFBH, throughput, execution time (during query and invoke transactions), and block size was evaluated. Transactions up to 10,000 was experimented with. The transactions in the simulation using a version of the hyperledger caliper that is modified [49] was measured based on the submission from the consensus transaction by simulated peers. The execution time covered the time that is required to successfully execute a transactions after it is added. The throughput of the model is meant to capture the amount of successful transaction(s) for each (or per) second [46, 47]. The evaluated blockchain (or block) size is meant to capture the number of transactions usually per second, which is an important design parameters [47]. With the execution time it is possible to observe the behaviour of the model during query and invoke transaction vis-à-vis the chain code during, which is important since the Indy distributed ledge of the Hyperledger project to enforce a better decentralized ledger solution that support self-sovereign identity.

3.3 Sensitive data scenario protection modelling

Sensitive data are data that must be protected against unwanted disclosure. Therefore, protecting it from unauthorized access to safeguard its privacy and security is of paramount important. This conception guided the modelling of the sensitive data scenario. Given the scenario (i.e., situation); let the sensitive data be 'a'; where 'a' is assumed earlier postulated to be made up of users' Personal Information (P-Inf) that are on the on the wall of social media users' platform. For modelling purpose, we use the P-Inf components: The Personally Identifiable Information

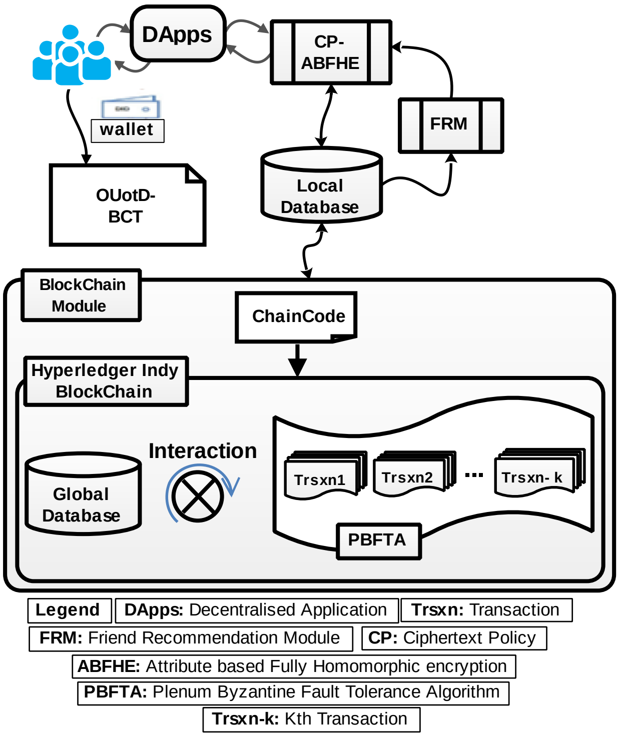

Fig. 1. The proposed sensitive data protection model architecture

(PII), Potentially Personally Identifiable Information (PPII) and Posts (P). Formally, this conception is stated as;

$$a = \{PII, PPII, P\} \tag{1}$$

The sensitivity of a certain data attribute was therefore defined as a function

$$s \rightarrow (a, r) \rightarrow [0, 1] \tag{2}$$

Where
a = a vector of data attribute
r = the recipient (other users)

Formally, a = *(a_1, $a_{2,.}$ . . a_n)* which is a vector of data attributes (e.g.. name, address, posts, etc.) that can possibly be requested by recipient '*r'* in order to create a relationship(s) or interaction(s) such that *$s(a, r) \in [0, 1]$* is a user-specified level of sensitivity of sharing information that relates to a j^{th} data attribute with a recipient *r*.

This consists of
$r_j = 1$; if the j^{th} data attribute is requested by a recipient *r* and $r_j = 0$ otherwise.

Mathematically, the SDPM (sd) was represented as a 6-tuple which is define as shown in Equation (iii) as follows;

$$sd = \{u, a, e, l, f, \beta) \qquad (3)$$

Where
u = User
a = Sensitive data
e = Encryption algorithms
l = Local database
f = Friend recommendation algorithm
β = Blockchain that is based on Hyperledger Indy
framework

Additionally, the Blockchain (*β*) is a 3-tuple as shown in Equation (iv) as follows;

$$B = \{c, p, s1\} \qquad (4)$$

Where
c = chaincode
p = consensus mechanism
s^1= Hashed of the encrypted data

Whenever a user communicate with the SPDM, the DApps is downloaded and setup with user registration to create their profile (for new users), while existing users would login to perform interactions (i.e. transactions) such as posts, likes, follows, comments, etc. The downloaded DApps would contain both the L-Dbase and the HIBC while each user owns a wallet that contains the decentralised identification that enable them to generate private keys using the public key in their wallet.

4 Results and Discussion

It was important to choose these metrics - throughput, execution time and block size since the proposed method is novel in that it provides not just a decentralized solution as descriptively presented in [45] but a user-centric self-sovereign identity-based solution. From the preliminary result obtained it was observed that on all fronts – execution time, throughput and block size results interestingly follow the pattern documented in the literature [46-50]. For instance, The Tables 1, 2, 3, and 4 show the simulation results of the throughput, block size, and execution tine for query and invoke of the method suggested using the SDPMA presented as follows.

Table 1. Throughput Result

NoT	1	10	50	100	500	1000	2000	3000	5000	10000
TpS	30	70	180	300	270	250	230	210	200	160

NoT (Number of Transaction); TPS (Transaction per Seconds)

Table 2. Block size (Using # transactions)

NoT	1	10	50	100	200	300	400	500	800
TpS	200	320	380	350	355	350	345	340	320

NoT (Number of Transaction); TpS (Transaction per Seconds)

Table 3. Execution time (Query)

NoT	0	10	100	1000	10000
TpS	0	10	20	30	40

NoT (Number of Transaction); TpS (Transaction per Seconds)

Table 4. Execution time (Invoke)

NoT	0	10	50	100	500	1000	5000	10000
TpS	0	10	20	30	40	50	60	70

NoT (Number of Transaction); TpS (Transaction per Seconds)

Similarly, the results in Tables 1 to 4 of the throughput, block size, and execution time derived from simulating the model is also presented graphically as follows in Figures 2 to 5.

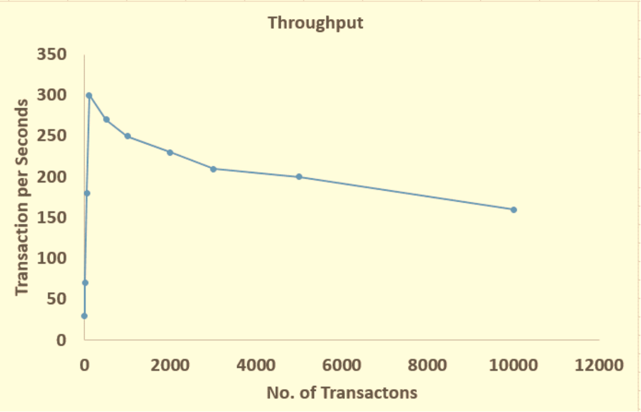

Fig. 2. Throughput

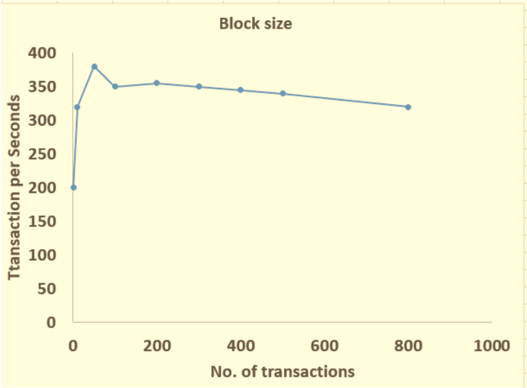

Fig. 3. Block size

(No. of transaction vs transaction per seconds) (no. of transaction vs transaction per seconds)

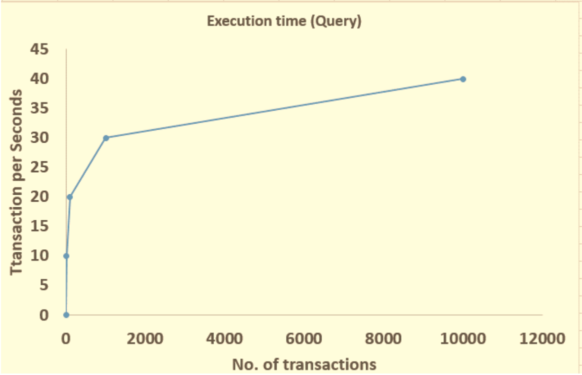

Fig. 4. Execution time (query)

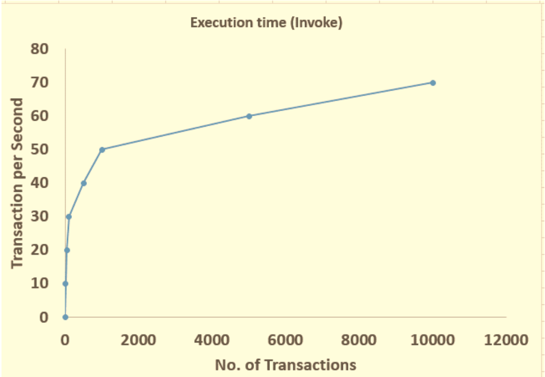

Fig. 5. Execution time (invoke)

(No. of transaction vs transaction per seconds) (no. of transaction vs transaction per seconds)

From Figure 2, the average throughput of the model is observed as clearly higher since it processed up to 300 transactions per second when compared to 40 transactions per second obtained in previous work [51]. As shown in Figure 3, only few transactions per block were identified to have undesirable influence on throughput. Though, there was a quick increase of throughput at 10 transaction per block, increase of performance was observed to diminish. The maximum throughput of 350 transactions per second plots around 100 transactions per block and thus did not exceed the recommended block size, generation, and mining time, which is consistent with highlighted requirements in [47, 52, 53]. The same pattern of using more time for more transactions that is found in the literature [46, 51] can be observed in Figures 4 and 5. This informs and validate the proposed method in this paper as plausible. Since execution time is the time required for a method like the one presented in this paper to execute a transaction after adding one successfully [46], the result in Figures 4 and 5 is consistent with what obtains in the literature [46, 51] and highlight a good consensus provision. It can be inferred based on the provisions in [52] that the proposed method would show capability in respect of supporting agreement up to approximately 75% regarding making changes by participants in the chain. This is a good performance and consistent with the behaviour described of Hyperledger-based model solutions [46, 48, 51, 52, 53].

5 System Implementable Formalism

This section presents implementable UML OO formalisms to contemplate in the implementation of SDPMA. Three of this formalism are presented as shown in Figure 6, 7, and 8. The rational for this is to present varieties of algorithms, which can be complex, difficult to present and understand in an easy and simplified way. This inadvertently ensures when an ensemble of methods is proposed as done in this paper. Cognizance of this, in Figure 2, the block diagram is presented to show relevant modules and their description showing tier role vis-à-vis their responsibilities. The block diagram helps to visualize the detail flow and communication between existing components and show the convenience in implementing the proposed processes involved in protecting the privacy of users' sensitive data. Similarly, the Activity Flow Diagram (AFD) was applied to show both the user-based functionalities and the blockchain operation in the SDPMA. Both models in Figures 7 and 8 show the flow of control from activity to activity, thus shifting the focus of the flow of control from object-orientation as shown using the model in Figure 6 to specific activities as shown in the models in Figures 7 and 8. The dynamic nature of the Executable and Implementable System (E&IS) from the architecture presented in Figure 1 is presented using the AFD (see Figures 7 and 8). The behaviour of the E&IS in dynamic terms showing the concurrent as well as sequential processes are shown using the model in Figure 7.

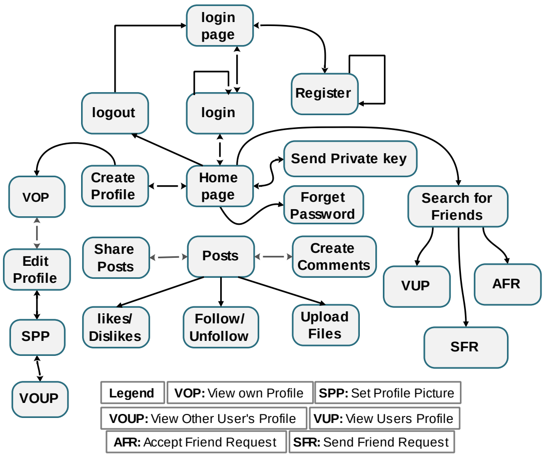

Fig. 6. Block diagram of the SPDM

6 Conclusion

The main goal of this research work is to develop a Sensitive Data Protection Architecture-based Model (SDPA-bM) that delivers secure solutions. This aim was achieved by proposal presented on a sensitive data protection model-based architecture that preceded the modelling of a sensitive data protection scenario and presentation of OO-UML- based formalisms to implement the proposed SDPA-bM. This paper contributes an architectural-based ensemble of methods that uses the blockchain technique to protect the privacy of users' sensitive data. The architectural-based ensemble of methods is used to integrate trust in the network itself to enable identity owners have sovereignty of their identity and control access to their records while ensuring integrity and content availability. The ensemble of method, which approach is presented applies a fully distributed and secure methodology to offer high-quality services with no operational cost, despite running on unreliable, unsecure and sometimes malicious user devices. The paper employs a novel approach that uses the technology of Blockchain in synergy with cryptographic techniques, hashing and consensus mechanism to enforce privacy, trust and availability of data among untrusted peers on OSNs. However, the research work that resulted in the proposal reported in this paper is still ongoing with implementation of the prototype model for deployment in a real social network environment already at advanced stage. Based on preliminary result, the computational overhead incurred by applying the ensemble method is significantly less. This is consistent with the belief in literature (e.g. [41-44] that ensemble methods are computationally feasible with the use of less resources and computational cost since the computational time scales linearly.

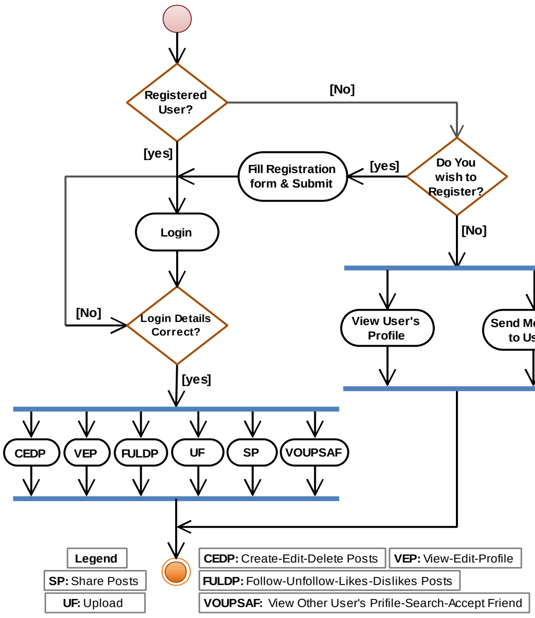

Fig. 7. Activity flow diagram for User's functions in the SPDM

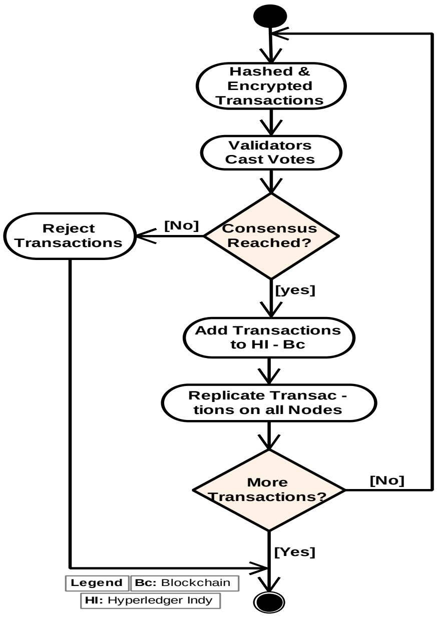

Fig. 8. Activity flow diagram for Blockchain Operations in the SPDM

7 References

[1] Kodad, M. (2020). Engagement and Performance Studies of Media Agencies Publications on Social Networks. *Int'l Jour. of Rec. Contr. from Engr. Sc. & IT (iJES)*, *8*(3), 4 - 19. https://doi.org/10.3991/ijes.v8i3.16949

[2] Lerro, F., Orduña, P., Marchisio, S., & García-Zubía, J. (2014). Development of a remote laboratory management system and integration with social networks. *International Journal of Recent Contributions from Engineering, Science & IT (iJES)*, *2*(3), 33-37. http://dx.doi.org/10.3991/ijes.v2i3.3821

[3] Zhang, Z., and Gupta, B. B. (2018). Social media security and trustworthiness: overview and new direction. *Future Generation Computer Systems*, *86*, 914-925. https://doi.org/10.1016/j.future.2016.10.007

[4] De Salve, A., Mori, P., and Ricci, L. (2018). A survey on privacy in decentralized online social networks. *Computer Science Review*, *27*, 154-176. https://doi.org/10.1016/j.cosrev.2018.01.001

[5] Guidi, B., Conti, M., Passarella, A., and Ricci, L. (2018). Managing social contents in Decentralized Online Social Networks: A survey. *Online Social Networks and Media*, *7*, 12-29. https://doi.org/10.1016/j.osnem.2018.07.001

[6] Zlatolas, L. N., Welzer, T., Hölbl, M., Heričko, M., and Kamišalić, A. (2019). A Model of Perception of Privacy, Trust, and Self-Disclosure on Online Social Networks. *Entropy*, *21*, 772. https://doi.org/10.3390/e21080772

[7] Kayes, I., and Iamnitchi, A. (2017). Privacy and security in online social networks: A survey. *Online Social Networks and Media*, *3*, 1-21. https://doi.org/10.1016/j.osnem.2017.09.001

[8] Ellison, N. B., and Boyd, D. M. (2013). Sociality through social network sites. In *The Oxford handbook of internet studies*. DOI: 10.1093/oxfordhb/9780199589074.013.0008

[9] Alvarado, C., Devadoss, N., Rivens, R., and Engels, D.W. (2018). It's Your Data: A Blockchain Solution to Facebook's Data Stewardship Problem. *SMU Data Science Review*, *1*(4), 2. Available at: https://scholar.smu.edu/datasciencerevie- w/vol1/iss4/2

[10] Bartsch, M., and Dienlin, T. (2016). Control your Facebook: An analysis of online privacy literacy. *Computers in Human Behavior*, *56*, 147-154. https://doi.org/10.1016/j.chb.2015.11.022

[11] Scott, D. M. (2015). *The New Rules of Marketing and PR.: How to Use Social Media, Online Video, Mobile Applications, Blogs, News Releases, and Viral Marketing to Reach Buyers Directly*. John Wiley & Sons. ISBN: 978-81-265-6004-2.

[12] Rathore, N. C., and Tripathy, S. (2019). InfoRest: Restricting Privacy Leakage to Online Social Network App. arXiv preprint arXiv:1905.06403. Available at: https://arxiv.org/abs/1905.06403

[13] Bahri, L., Carminati, B., and Ferrari, E. (2018). Decentralized privacy preserving services for online social networks. *Online Social Networks and Media*, *6*, 18-25. https://doi.org/10.1016/j.osnem.2018.02.001

[14] Cadwalladr, C., and Graham-Harrison, E. (2018). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The guardian*, *17*, 22. Availabe at: https://www.theguardian.com/news/2018/mar/17/

[15] Onik, M. M. H., Kim, C. S., Lee, N. Y., and Yang, J. (2019). Privacy-aware blockchain for personal data sharing and tracking. *Open Computer Science*, *9*(1), 80-91. https://doi.org/10.1515/comp-2019-0005

[16] Hudson, S., Huang, L., Roth, M. S., and Madden, T. J. (2016). The influence of social media interactions on consumer–brand relationships: A three-country study of brand perceptions and marketing behaviors. International Journal of Research in Marketing, 33(1), 27-41. https://doi.org/10.1016/j.ijresmar.2015.06.004

[17] Aljably, R., Tian, Y., Al-Rodhaan, M., and Al-Dhelaan, A. (2019). Anomaly detection over differential preserved privacy in online social networks. *PloS one*, *14*(4), e0215856. https://doi.org/10.1371/journal.pone.0215856

[18] Fire, M., Goldschmidt, R., and Elovici, Y. (2014). Online social networks: threats and solutions. *IEEE Communications Surveys & Tutorials*, *16*(4), 2019-2036. DOI: https://doi.org/10.1109/COMST.2014.2321628

[19] Gao, H., Hu, J., Wilson, C., Li, Z., Chen, Y., and Zhao, B. Y. (2010). Detecting and characterizing social spam campaigns. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement* (pp. 35-47). ACM. https://doi.org/10.1145/1879141.1879147

[20] Hallinan, D., Friedewald, M., and McCarthy, P. (2012). Citizens' perceptions of data protection and privacy in Europe. *Computer law & security review*, *28*(3), 263-272.

[21] Greschbach, B., Kreitz, G., and Buchegger, S. (2012). The devil is in the metadata—new privacy challenges in decentralised online social networks. In *2012 IEEE International Conference on Pervasive Computing and Communications Workshops* (pp. 333-339). IEEE. https://doi.org/10.1016/j.clsr.2012.03.005

[22] Shozi, N. A., and Mtsweni, J. (2017). Big data privacy in social media sites. In *2017 IST-Africa Week Conference (IST-Africa)* (pp. 1-6). IEEE. https://doi.org/10.23919/ISTAFRICA.2017.8102311

[23] Ali, S., Islam, N., Rauf, A., Din, I., Guizani, M., and Rodrigues, J. (2018). Privacy and security issues in online social networks. *Future Internet*, *10*(12), 114. 114; https://doi.org/10.3390/fi10120114

[24] NaliniPriya, G., and Asswini, M. (2015). A survey on vulnerable attacks in online social networks. In *International Confernce on Innovation Information in Computing Technologies* (pp. 1-6). IEEE. DOI: https://doi.org/10.1109/ICIICT.2015.7396102

[25] Koll, D., Li, J., and Fu, X. (2017). The good left undone: Advances and challenges in decentralizing online social networks. *Computer Communications*, *108*, 36-51. https://doi.org/10.1016/j.comcom.2017.04.008

[26] Sharma, S., and Sodhi, J. S. (2016). EncryptPost: A Framework for User Privacy on Social Networking Sites. In *Proceedings of First International Conference on Information and Communication Technology for Intelligent Systems: Volume 1* (pp. 51-60). Springer, Cham. https://doi.org/10.1007/978-3-319-30933-0_6

[27] Dwyer, C. (2011). Privacy in the Age of Google and Facebook. *IEEE Technology and Society Magazine*, *30*(3), 58-63.: https://doi.org/10.1109/MTS.2011.942309

[28] Oukemeni, S., Rifà-Pous, H., and Puig, J. M. M. (2019). Privacy Analysis on Microblogging Online Social Networks: A Survey. *ACM Computing Surveys (CSUR)*, *52*(3), 60. https://doi.org/10.1145/3321481

[29] Raman, A., Joglekar, S., Cristofaro, E. D., Sastry, N., and Tyson, G. (2019). Challenges in the Decentralised Web: The Mastodon Case. In *Proceedings of the Internet Measurement Conference* (pp. 217-229). ACM. https://doi.org/10.1145/3355369.3355572

[30] Paul, T., Lochschmidt, N., Salah, H., Datta, A., and Strufe, T. (2017). Lilliput: A storage service for lightweight peer-to-peer online social networks. In *2017 26th International Conference on Computer Communication and Networks (ICCCN)* (pp. 1-10). IEEE. https://doi.org/10.1109/ICCCN.2017.8038443

[31] Ding, D., Conti, M., and Figueiredo, R. (2019). SAND: Social-aware, network-failure resilient and decentralized microblogging system. *Future Generation Computer Systems*, *93*, 637- 650. https://doi.org/10.1016/j.future.2018.11.007

[32] Aderibigbe, S. O., Akhigbe, B. I., Afolabi, B. S., and Adagunodo, E. R. (2017). A friend-to-friend approach for secured knowledge sharing. In *Systèmes d'organisation des connaissances et humanités numériques: Actes du 10ème colloque ISKO France 2015* (p. 172). ISTE Group. ISBN: 978-78406-214-9.

[33] Buchegger, S., and Datta, A. (2009). A case for P2P infrastructure for social networks-opportunities & challenges. In *2009 Sixth International Conference on Wireless On-Demand Network Systems and Services* (pp. 161-168). IEEE. https://doi.org/10.1109/WONS.2009.4801862

[34] Cutillo, L. A., Molva, R., and Onen, M. (2011). Analysis of privacy in online social networks from the graph theory perspective. In *2011 IEEE Global Telecommunications Conference-GLOBECOM 2011* (pp. 1-5). IEEE. https://doi.org/10.1109/GLOCOM.2011.6133517

[35] Shinjo, Y., Naito, S., Kunyao, X., and Sato, A. (2017). ABnews: A fast private social messaging system using untrusted storage and attribute-based encryption. In *2017 15th Annual Conference on Privacy, Security and Trust (PST)* (pp. 317-31709). IEEE. https://doi.org/10.1109/PST.2017.00045

[36] Klukovich, E., Erdin, E., and Gunes, M. H. (2016). POSN: A privacy preserving decentralized social network app for mobile devices. In *Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* (pp. 1426-1429). IEEE Press. DOI Bookmark: https://doi.ieeecomputersociety.org/10.1109/ASONAM.2016.7752436

[37] Elisa, N., Yang, L., Chao, F., and Cao, Y. (2018). A framework of blockchain-based secure and privacy-preserving E-government system. *Wireless Networks*, 1-11. https://doi.org/10.1007/s11276-018-1883-0

[38] Mbinkeu, R. C. N., & Batchakui, B. (2015). Reducing Disk Storage with SQLite into BitCoin Architecture. *Int'l Jour. of Rec. Contr. from Engr., Sc. & IT (iJES)*, *3*(2), 10-14. http://dx.doi.org/10.3991/ijes.v3i2.4490

[39] Li, C., and Palanisamy, B. (2019). Incentivized Blockchain-based Social Media Platforms: A Case Study of Steemit. *arXiv preprint arXiv:1904.07310*. https://doi.org/10.1145/3292522.3326041

[40] Qin, D., Wang, C., and Jiang, Y. (2018). RPCHAIN: a blockchain-based academic social networking service for credible reputation building. In *International Conference on Blockchain* (pp. 183-198). Springer, Cham. https://doi.org/10.1007/978-3-319-94478-4_13

[41] Dutta, H., and Srinivasan, A. (2018). Consensus-based modeling using distributed feature construction with ILP. *Machine Learning*, 107(5), 825-858. https://doi.org/10.1007/s10994-017-5672-2.

[42] Rathore, H., Mohamed, A., and Guizani, M. (2020). A Survey of Blockchain Enabled Cyber-Physical Systems. *Sensors*, *20*(1), 282. https://doi.org/10.3390/s20010282

[43] Zhang, X.X., Ge, B.F., and Tan, Y.J. (2016). A consensus model for group decision making under interval type-2 fuzzy environment. *Frontiers of Information Technology & Electronic Engineering*, 17(3), 237-249. https://doi.org/10.1631/FITEE.1500198

[44] Zhan, H., Gomes, G., Li, X. S., Madduri, K., Sim, A., and Wu, K. (2018). Consensus ensemble system for traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*, 19(12), 3903-3914. https://doi.org/10.1109/TITS.2018.2791505

[45] Virmani, C., and Choudhary, T. (2020). Blockchain-Based Social Network Infrastructure. In *Strategic System Assurance and Business Analytics* (pp. 517-528). Springer, Singapore.

[46] Nasir, Q., Qasse, I. A., Abu Talib, M., and Nassif, A. B. (2018). Performance analysis of hyperledger fabric platforms. *Security and Communication Networks (Wiley/Hindawi), 2018*. https://doi.org/10.1155/2018/3976093

[47] Fan, C., Ghaemi, S., Khazaei, H., and Musilek, P. (2020). Performance Evaluation of Blockchain Systems: A Systematic Survey. *IEEE Access*, *8*, 126927-126950. DOI: *10.1109/ACCESS.2020.3006078.*

[48] Pongnumkul, S., Siripanpornchana, C., and Thajchayapong, S. (2017). Performance analysis of private blockchain platforms in varying workloads. In *2017 26th International Conference on Computer Communication and Networks (ICCCN)* (pp. 1-6). IEEE.

[49] Hyperledger/caliper [Online]. Available: https://github.com/hyperledger/caliper. [Acces sed: 12-Jan-2020].
[50] Ferdous, M. S., Chowdhury, M. J. M., Hoque, M. A., and Colman, A. (2020). Blockchain Consensus Algorithms: A Survey. [Online]. Available: http://arxiv.org/abs/2001.07091 [Assessed: 13-Jan-2020].
[51] Memon, R. A., Li, J. P., and Ahmed, J. (2019). Simulation model for blockchain systems using queuing theory. *Electronics*, *8*(2), 234. DOI: 10.3390/electronics8020234.
[52] Ajao, L. A., Agajo, J., Adedokun, E. A., and Karngong, L. (2019). Crypto hash algorithm-based blockchain technology for managing decentralized ledger database in oil and gas industry. *J—Multidisciplinary Sc. Journal*, *2*(3), 300-325.
[53] Memon, R.A., Li, J.P., and Ahmed, J. (2019). Simulation model for blockchain systems using queuing theory. *Electronics*, *8*(2), 234. DOI: 10.3390/electronics8020234.

8 Authors

Felix O. Idepefo is a Postgraduate student of Obafemi Awolowo University Ile-Ife, Osun State, Nigeria. His research work focuses on Sensitive Data Protection on Online Social Networks. He has MSc. Computer Science from University of Lagos, Akoka, Lagos State and MSc. Computer Science (Software Engineering Option) from University of Ilorin, Kwara State, Nigeria. His research interest include Cyber/Information security, Biometrics, Software Engineering and Information Management. He is a member of NCS, IAENG, IACSIT, ISOC and SDIWC. (email: felixidepefo@gmail.com).

Bernard Ijesunor Akhigbe holds a Ph.D. in computer science and he is a Senior Lecturer in the Department of Computer Science and Engineering, Obafemi Awolowo University, Nigeria. He researches generally in Information system and Software engineering with interest in IoT and Blockchain for management gains. He has published widely and attended learned workshop and conferences both locally and internationally. He is a member of ISRG, ISKO France, NCS and CPN.

Ojo Stephen Aderibigbe holds a Ph.D in Computer Science and He is a Senior Lecturer in Computer Science Department of the Lagos State Polytechnic, Ikorodu, Lagos State, Nigeria. He is a member NCS, CPN, and ISKO. His research focus on Information storage and Retrieval, and Collaborative Trust-Aware models. He has presented papers at both local and international conferences.

Babajide Samuel Afolabi is a Professor in the Department of Computer Science and Engineering, Obafemi Awolowo University, Nigeria, who holds a Ph.D. from Université Nancy 2, Nancy, France. He is focused on developing applications for enhanced Living. He is currently the Director of the Obafemi Awolowo University Computer Centre. He researches in Information system and Software engineering. He is well published and has attended learned conferences both locally and internationally. He is a member of ISRG, ISKO France, NCS and CPN.

Article submitted 2020-12-30. Resubmitted 2021-01-28. Final acceptance 2021-02-05. Final version published as submitted by the authors.