B. Trust and Decision-Making in Blockchains

Bitcoin, or blockchains in general, assumes all nodes are equally untrusted and that their proportion in the collective decision-making process is solely based on their computational resources (known as the Proof-of-work algorithm) [17]. In other words – for every node n, $trust_n \propto resources(n)$ (probabilistically) decides the node's weight in votes. This leads to adverse effects, most notably vulnerability to sybil attacks, excessive energy consumption and high-latency.

Intuitively, Proof-of-Work reasons that nodes which pour significant resources into the system are less likely to cheat. Using similar reasoning we could define a new dynamic measure of trust that is based on node behavior, such that good actors that follow the protocol are rewarded. Specifically, we could set the trust of each node as the expected value of it behaving well in the future. Equivalently, since we are dealing with a binary random variable, the expected value is simply the probability p. A simple way to approximate this probability is by counting the number of good and bad actions a node takes, then using the sigmoid function to squash it into a probability. In practice, every block i we should re-evaluate the trust score of every node as –

$$trust_n^{(i)} = \frac{1}{1 + e^{-\alpha(\#good - \#bad)}}, \tag{3}$$

where α is simply the step size.

With this measure, the network could give more weight to *trusted* nodes and compute blocks more efficiently. Since it takes time to earn trust in the system, it should be resistant to sybil attacks. This mechanism could potentially attract other types of attacks, such as nodes increasing their reputation just to act maliciously at a later time. This might be mitigated by randomly selecting several nodes, weighted by their trust, to vote on each block, then taking the *equally-weighted* majority vote. This should prevent single actors from having too much influence, regardless of their trust-level.

VI. Conclusion

Personal data, and sensitive data in general, should not be trusted in the hands of third-parties, where they are susceptible to attacks and misuse. Instead, users should own and control their data without compromising security or limiting companies' and authorities' ability to provide personalized services. Our platform enables this by combining a blockchain, re-purposed as an access-control moderator, with an off-blockchain storage solution. Users are not required to trust any third-party and are always aware of the data that is being collected about them and how it is used. In addition, the blockchain recognizes the users as the owners of their personal data. Companies, in turn, can focus on utilizing data without being overly concerned about properly securing and compartmentalizing them.

Furthermore, with a decentralized platform, making legal and regulatory decisions about collecting, storing and sharing sensitive data should be simpler. Moreover, laws and regulations could be programmed into the blockchain itself, so that they are enforced automatically. In other situations, the ledger can act as legal evidence for accessing (or storing) data, since it is (computationally) tamper-proof.

Finally, we discussed several possible future extensions for blockchains that could harness them into a well-rounded solution for trusted computing problems in society.

References

[1] Scaling the facebook data warehouse to 300 pb, 2014.

[2] James Ball. Nsa's prism surveillance program: how it works and what it can do. *The Guardian*, 2013.

[3] Michael Ben-Or, Shafi Goldwasser, and Avi Wigderson. Completeness theorems for non-cryptographic fault-tolerant distributed computation. In *Proceedings of the twentieth annual ACM symposium on Theory of computing*, pages 1–10. ACM, 1988.

[4] EUROPEAN COMMISSION. *Commission proposes a comprehensive reform of data protection rules to increase users' control of their data and to cut costs for businesses*. 2012.

[5] Yves-Alexandre de Montjoye, César A Hidalgo, Michel Verleysen, and Vincent D Blondel. Unique in the crowd: The privacy bounds of human mobility. *Scientific reports*, 3, 2013.

[6] Yves-Alexandre de Montjoye, Erez Shmueli, Samuel S Wang, and Alex Sandy Pentland. openpds: Protecting the privacy of metadata through safeanswers. *PloS one*, 9(7):e98790, 2014.

[7] Cynthia Dwork. Differential privacy. In *Automata, languages and programming*, pages 1–12. Springer, 2006.

[8] Jon Evans. Bitcoin 2.0: Sidechains and ethereum and zerocash, oh my!, 2014.

[9] Craig Gentry. Fully homomorphic encryption using ideal lattices. In *STOC*, volume 9, pages 169–178, 2009.

[10] Vindu Goel. Facebook tinkers with users' emotions in news feed experiment, stirring outcry. *The New York Times*, 2014.

[11] Federal Information and Processing Standards. FIPS PUB 180-4 Secure Hash Standard (SHS). (March), 2012.

[12] Don Johnson, Alfred Menezes, and Scott Vanstone. The elliptic curve digital signature algorithm (ecdsa). *International Journal of Information Security*, 1(1):36–63, 2001.

[13] Michael Lesk. How much information is there in the world?

[14] Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian. t-closeness: Privacy beyond k-anonymity and l-diversity. In *ICDE*, volume 7, pages 106–115, 2007.

[15] Ashwin Machanavajjhala, Daniel Kifer, Johannes Gehrke, and Muthuramakrishnan Venkitasubramaniam. l-diversity: Privacy beyond k-anonymity. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(1):3, 2007.

[16] Petar Maymounkov and David Mazieres. Kademlia: A peer-to-peer information system based on the xor metric. In *Peer-to-Peer Systems*, pages 53–65. Springer, 2002.

[17] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. *Consulted*, 1(2012):28, 2008.

[18] Arvind Narayanan and Vitaly Shmatikov. How to break anonymity of the netflix prize dataset. *arXiv preprint cs/0610105*, 2006.

[19] Juan Perez. Facebook, google launch data portability programs to all, 2008.

[20] Rt.com. Obama announces legislation protecting personal data, student digital privacy, 2015.

[21] K Schwab, A Marcus, JO Oyola, W Hoffman, and M Luzi. Personal data: The emergence of a new asset class. In *An Initiative of the World Economic Forum*, 2011.

[22] ScienceDaily. Big data, for better or worse: 90% of world's data generated over last two years. 2013.

[23] Adi Shamir. How to share a secret. *Communications of the ACM*, 22(11):612–613, 1979.

[24] Latanya Sweeney. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(05):557–570, 2002.