PAI BLOCKCHAIN PROTOCOL

A DECENTRALIZED ARTIFICIAL INTELLIGENCE NETWORK TECHNICAL OVERVIEW

October 2017
ProjectPAI.com

ABSTRACT

Personal Artificial Intelligence, or PAI, gives humans the ability to augment themselves. A decentralized network can provide public infrastructure so that Personal Artificial Intelligences can improve and interact with one another. Decentralization minimizes the need for counterparty trust with incentives to optimize individual freedoms rather than centralized profit. Digital signatures are used to sign and provision data. Personal data has a chain of custody with tools for permission and revocation. The network seeks to protect personal information while also providing channels for communication and transaction.

1. INTRODUCTION

Data is flooding the world. And the data being generated in ever-increasing quantities every day, all around the globe, has tremendous value: it serves to train, inform, operate, and improve the new technologies that are changing people's relationships not only with businesses and governments but with each other—and even with themselves. Data is the lifeblood that fuels the exponential growth of sophisticated software. It breathes life into algorithms.

The advent of Personal Artificial Intelligence (PAI)[1] holds out the prospect of a bold new future with many distinct advantages, but one that is not lacking in challenges. A future in which our PAIs respond to emails, organize our calendars, teach us a new language, and answer questions on our behalf requires robust infrastructure. Many of the advantages and challenges of building this infrastructure involve data.

The data fed to a PAI typically originates with a human author. It informs the PAI to behave like—and take on the persona of—its human counterpart. A PAI learns how to model its behavior by training on the personal data provided to it. Since this data is the essence or spirit of the PAI, and a PAI is the closest digital approximation of a human being, the value of this PAI-specific data is arguably higher than that of ordinary web-interaction data.

This personal data also has value outside of the individual Human:PAI relationship. For example, if someone's PAI becomes particularly good at image recognition, other PAIs in the ecosystem can benefit

[1] Link to Project PAI Whitepaper

from that learned behavior. Owners can be compensated for the effort that they put into training their PAIs through a data-transfer marketplace. The specific algorithms and data that evolves from training can be sold on a market where other PAIs reap the benefits and where the original creators are compensated when others make use of their work. In this way, PAI owners are incentivized to improve their PAIs. The larger PAI community benefits when one individual becomes exceptionally good at a specific task—and that individual benefits by selling the fruits of their labor.

The current trajectory for the information economy is not sustainable. Surrendering our personal data to others, without compensation, has unfortunately become the status quo. Technology companies today vacuum up vast quantities of user data, which they deploy for advertising, recommendations, and many other features. This would be less deplorable if users were aware of what and how much personal information they were giving up, but in many cases they haven't knowingly given permission for their data to be harvested and used in these ways. As it stands, the ecosystem for the sharing of personal data lacks transparency and, therefore, accountability. The quality of data naturally increases when its creators are appropriately compensated, yet the existing community of web-based mega-organizations takes users' data without compensating them directly. By hiding these costs, companies are able to benefit from information asymmetry.[2] The hidden expense is incurred by people using services such as social networking, search, and email.

In this current ecosystem, individuals are unable to capture the value of their data, because they have lost control of its dissemination. Its value accrues instead to tech firms, to banks, to advertisers, and—it hardly needs saying—to malware authors and identity thieves. People's personal data is a valuable commodity—but not one they can cash in on. This existing model is unsustainable because it concentrates the data economy in the hands of the few and leaves the general public out in the cold.

PAI Blockchain will serve as the public infrastructure necessary to upend this status quo, enabling people to own, provision, revoke access permission for, and conduct commerce with their personal data. They will be able to do this transparently and without having to rely on trusted central parties that may or may not be worthy of trust.

PAIs learn to think and behave like people because each one learns from the data of a specific person as well as from the data of that person's community. PAIs excel when their ecosystem is driven by metrics that benefit the rights of individuals and the communities to which they belong. This starts with placing privacy, ownership of data, and permissioned access above centralized profit.

The blockchain industry has enjoyed a huge infusion of capital in recent months. Most of the new funds have come from cryptocurrency users themselves—so-called retail investors—participating in

[2] Lanier, Jaron. Who Owns the Future? Penguin Books, 2014.

crowdfunded token sales, which not only provide a way for open-source projects to raise money to finance the development of their networks but also link their success with the participation and profit of their user-constituents. From January through September 2017, initial coin offerings[3] raised a total of $2.2 billion for various blockchain projects, more than three times as much as venture capital firms invested in the industry during the same period, according to Novum Insights. Under the old model, represented by venture capital firms, the largest profits accrue to a few wealthy early investors. The old model makes users second-class citizens. But the blockchain industry is breaking that model.

As blockchains move old industries to on-chain dependencies, money and data will become synonymous with one another and their flows will become increasingly transparent. This will incentivize entrepreneurs to work more closely with their customers, cutting the strings that come from third-party capital and creating an overall higher level of accountability to their users. If caring about the success of a startup could be quantified as a variable coefficient, that coefficient would increase as users were directly incentivized to participate in a company's growth.

An information economy driven by proper incentives and compensation will create new frontiers for entrepreneurship. For many businesses, a primary objective will be the amalgamation of comprehensive datasets. Lucrative opportunities await those companies that can facilitate the secure storage of data and broker it responsibly. PAI augmentation itself will create entirely new markets.

The futuristic nature of PAI technology has inspired infrastructure which facilitates that future—while transcending the norms of data appropriation. Not only individuals but companies could benefit from new methods of data provision and compensation, from shifting to a radical and sustainable PAI-based model.

2. METHODOLOGY

PAI Blockchain operates through the combination of three distinct modules:

(1) The Authentication Layer controls relationships that exist in the PAI ecosystem.

(2) The Blockchain Layer houses a record of network transactions as well as protocols for peer-to-peer engagement.

(3) The Data Layer indexes encrypted data, peer-to-peer, with hashed references.

These modules are separated in order to silo the specific incentives and subsystems used to facilitate their purpose. The Blockchain Layer is the consensual fabric that marries people and data. ECDSA[4], Bitcoin[5], and BitTorrent[6] have been chosen as the base technologies to fulfill the current needs of the PAI

[3] Initial coin offering, or ICO, is still the most common term, although the nomenclature is changing as some projects seek to do everything possible to avoid having their tokens classified as securities by financial regulators.

[4] R.C. Merkle, "Protocols for public key cryptosystems," In Proc. 1980 Symposium on Security and Privacy, IEEE Computer Society, pages 122-133, April 1980. 2014.

[5] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System," https://bitcoin.org/bitcoin.pdf

[6] B. Cohen "Incentives Build Robustness in BitTorrent," http://bittorrent.org/bittorrentecon.pdf

Blockchain layers. The codebase is open source and encourages developers to contribute through a non-contentious affirmation process involving standard GitHub pull requests.

The scientific method informs technical decisions, in that architectural choices are treated as hypotheses which need proving. Each segment of the three layers is agnostic as to implementation. If alternatives to Bitcoin, BitTorrent, and ECC are empirically proven to be substantially beneficial, then those alternatives will be adopted in place of or alongside the three initial base technologies.

The Project PAI Foundation will put the open-source development process in motion, guided by an ethos which separates business logic from blockchain technology. The foundation is committed to a vision in which individual rights are a greater priority than business interest. Business and application development will be explicitly distinct from the PAI foundation's focus on individual rights. Maintainers of the open-source codebase will not take compensation from PAI-related businesses.

2.1 AUTHENTICATION

In recent decades, cryptographic systems based on elliptic-curve cryptography have proven to be significantly stronger—that is, more secure—than those based on first-generation public-key cryptography such as RSA. (For instance, it would take less energy to break a 228-bit RSA key than is needed to boil a teaspoon of water. The amount of energy needed to break an elliptic-curve key of the same length, by contrast, could boil all the water on earth.)[7] Bitcoin has established itself as the longest-running instance of elliptic-curve cryptography in a blockchain. PAI Blockchain builds on this success through the use of Bitcoin-standard authentication protocols.

PAI Blockchain uses secp256k1[8] EC field and curve parameters utilizing BIP32[9] hierarchical deterministic key derivation. The mobile client uses BIP39[10] mnemonic codes to streamline the user experience for private key generation and storage. The client generates a keypair chain to accommodate each specific use case. Public keys are provisioned for transmitting and receiving coins, data, permissions, and other forms of communication between network nodes.

Key derivation on consumer hardware has proven to be insecure.[11] Better methods will become available once tamper-resistant, auditable, and provably secure crypto-processors exist for ordinary users.

[7] N. Sullivan, "A (relatively easy to understand) primer on elliptic curve cryptography," https://arstechnica.com/information-technology/2013/10/a-relatively-easy-to-understand-primer-on-elliptic-curve-cryptography

[8] D. Brown "Recommended Elliptic Curve Domain Parameters," http://www.secg.org/sec2-v2.pdf

[9] P. Wuille "Hierarchical Deterministic Wallets," https://github.com/bitcoin/bips/blob/master/bip-0032.mediawiki

[10] M. Palatinus, P. Rusnak, A. Voisine, S. Bowe, "Mnemonic code for generating deterministic keys," https://github.com/bitcoin/bips/blob/master/bip-0039.mediawiki

[11] Silent Bob is Silent https://embedi.com/files/white-papers/silent-bob-is-silent.pdf

BIP11[12] M-of-N signatures can be used for escrow services as well as multi-party authorization of data and/or coin flows, thereby enforcing an approval process for granting a network participant access to data or funds.

2.2 BLOCKCHAIN NETWORK

PAI Blockchain uses a forked Bitcoin codebase, and it is designed to adopt future Bitcoin security and feature patches. But PAI Blockchain possesses some key differences from Bitcoin.

Implicit modifications are made to the Bitcoin message protocol which allow for references to the PAI Blockchain data storage layer. OP_RETURN is used to store data references via transaction scripts.

PAI coins represent value on the network and are transmitted in a fashion similar to that of standard Bitcoin transactions. A future two-way peg event[13] may allow Ethereum ERC-20 tokens and other off-chain coins to interact with the PAI Blockchain as well.

An early allocation of PAI coins has been set aside for participants in the network's initial coin generation event. Otherwise, new coins are issued through mining. Mining generates coins from coinbase transactions as well as from the proven fee-based incentives found in the Bitcoin network.

Consensus is derived in standard Bitcoin terms: mining on the longest chain and accepting transactions from that chain as valid. At any time, however, users can initiate acceptance of a new protocol by diverging from this methodology.

It is expected that changes to the difficulty-retargeting algorithm will be made after the PAI blockchain codebase becomes available for public pull requests. Attacks on this vector are mitigated by using the term "valid chain" to denote the longest chain with the least amount of difficulty variance amplitude. Merged mining and retargeting improvements will help to mitigate distribution challenges so that coin distribution follows Bitcoin's 10-minute average block time, asymptote-controlled supply, and Nakamoto step-function halving of the block reward.

2.3 DATA

The BitTorrent protocol is bootstrapped by dedicated Project PAI nodes to maintain initial availability, redundancy, and security of data. Data stored in this layer is encrypted in the client. Nodes operated by Project PAI will deliberately reject data that is not encrypted. This precedent is a best effort with the goal

[12] G. Andreesen, "M-of-N Standard Transactions", https://github.com/bitcoin/bips/blob/master/bip-0011.mediawiki

[13] A. Back, M. Corallo, L. Dashjr, M. Friedenbach, G. Maxwell, A. Miller, A. Poelstra, J. Timón, P. Wuille "Enabling Blockchain Innovations with Pegged Sidechains," https://blockstream.com/sidechains.pdf

that all data-storage nodes will apply filters in order to prevent storage of unencrypted data.

The data-storage layer is not specific to BitTorrent. Message parameters simply require designation of the network, version, data, and intended recipient. The initial implementation of BitTorrent could thus be replaced by alternative or multiple storage systems. These data references are generalized to support both streaming and static data.

References to the encrypted data are passed using Bitcoin-standard messages in which meta parameters are stored in the OP_RETURN field.

3. CONCLUSION

We are already living through the ongoing paradigm shift known as the information economy. And now a new paradigm is emerging with the advent of useful artificial intelligence. Project PAI wants to build toward a future in which the underlying technologies of both are intertwined with the people who use them. A digital grid with clusters of information, informed by the ownership of personal data, will create a new digital frontier for us to explore together. The systems described here work together to lay the foundation for the PAI ecosystem in order for it to flourish. They can be amended as necessary, while always striving to stay in sync with open source and open infrastructure initiatives.

Fundamentally, the PAI Blockchain system seeks simultaneously to provide a sustainable community in which PAIs can interact with their human counterparts and to upend the traditional notion that data doesn't belong to the individuals who produce it.

The system's amendable public infrastructure will require constant involvement and improvement from its community. We hope that you will be a part of that community, create your PAI, and join in pushing technology to new heights—while never forgetting that you are the most important element.

If we keep dreaming of worlds we may never see, they will be here sooner than we could have imagined.

ABOUT THE AUTHOR

Alex Waters is leading the development of the PAI Blockchain Protocol. He was part of the original development team for Bitcoin Core and has founded several cryptocurrency and Bitcoin-focused initiatives. Currently, Alex is CEO and co-founder of Coin Apex, a NYC-based software, technology, and cryptography incubator.