Modelling the Secure Computational Marketplace by Deriving Abstractions From Technical Implementations in Modern Cryptography

Josh Fourie

State of the art cryptographic routines have historically exposed unforeseen economic opportunities for the market, and contemporary advancements in zero-knowledge, and secure multi-party computation primitives will continue to reveal similar and exciting prospects in the future. The schemes discussed here are all commonly capable of uniquely guaranteeing the integrity of computational processes without compromising, or otherwise reducing, the privacy of market participants. This paper introduces a novel conception of the *secure computational marketplace*, which refers to a collection of abstract observations and qualities inherent to these technologies, and then coalesced as an impression of a descriptive, pseudo-economic model. The marketplace operates as an extension of modern cloud-computing architecture to manifest as a solution for the development of an interim bridge apprehending the realisation of, and transition into,the homomorphic data-economy.

1 Introduction: Cryptography

Cryptography encodes security in information systems to ensure that the content of data-stores or lines of communication are protected from unauthorised access. The primitives provide security through encapsulating or hiding 'secret keys' within mathematically 'hard' problems, which are defined as the subset of queries that cannot be solved efficiently on a classical or quantum computer. A party has proven authorisation wherever they are capable of 'providing' or demonstrating a knowledge of the solution. Importantly, cryptographic routines adapt alongside the development of computing technology to retain effectiveness, and methodologies may be deprecated accordingly.

RSA. The surprising effects of novel cryptography are evident in the implementation of the RSA signature-scheme [SOURCE]. The protocol was introduced as a response to a paper from Diffie and Hellman [SOURCE] to provide for secure communication in the presence of an adversary by implementing an encryption/decryption scheme based on the hardness of exponents problem [CONFIRM]. Consequentially, any party was empowered to verify the originating identity of a message with strong guarantees, which is an essential component of commercial processes in the data-economy.

Contemporary advancements in the cryptographic tripartite of zero-knowledge, multi-party and homomorphic computations promise to embed privacy-as-a-default within cyber-infrastructure in a similarly disruptive way. [1] The protocols are commonly motivated by the need to verifiably compute on data without transferring or otherwise exposing the information along vulnerable lines of communication, as well as to minimise the authorised circle of access wherever possible. Typically, schemes within these distinctions are differentiable with regards to the underpinning mathematical problem or assumption, such as a choice of lattice or number-theoretic cryptography. Abstractly, the schemes are economically valuable as resolution of the tension between privacy and computational integrity currently inherent to localised or privacy-preserving machine processes.

Homomorphism. Homomorphic encryption is the 'Goldilocks' of modern-Cryptography enabling any party to compute on an encrypted data-set without decryption. [FILL OUT INTRO.] The realisation of homomorphic technology intuitively disrupts the economic infrastructure of the data-economy.

Multi-Party. Secure multi-party computation was introduced by Yao [SOURCE] to resolve the motivating dilemma where a collection of parties wish to compute the aggregate analysis of their data without sharing the information amongst themselves. [FILL OUT INTRO.] Similarly, secure multi-party computation would provide significant economic opportunities for cases such as decentralised aggregate analysis of a population.

Zero-Knowledge. Zero-knowledge is used loosely within this paper to describe the technologies deriving from GWR [SOURCE], where a single party is capable of proving the integrity of a computation without revealing the underlying information. [2] Generically, a zero-knowledge scheme requires that $\mathcal{P}$ initiates the routine by computing a pre-agreed circuit over a set of private '*witness*' inputs (w). [3] The computation is subsequently *blinded* to hide w, and then abstracted to an alternate representation such as 3-SAT or 'multi-party computation in the head' [SOURCE]. The counter-party subsequently verifies the correctness of the abstraction either interactively or non-interactively, according to the implementation. Contrary to traditional cryptographic routines, there is no authorised party external to the original computer that is capable of reversing the hiding process, which satisfies the 'zero-knowledge' component. The remaining guarantees relate to the *soundness* and *completeness* of the protocol, whereby a verifier will correctly identify an honest or dishonest proof within negligible error bounds.

The remainder of the paper is concerned with explicating the architectural consequences of these technologies, evaluating the state-of-the-art, and then extracting

[1] The relevant primitives may be a cross-blend of these distinct but highly inter-related routines.

[2] The reader should note that the term nevertheless retains meaning as a generic condition that homomorphic and multi-party routines often self-prescribe.

[3] A witness is merely a solution to a computationally hard problem that can be efficiently checked to determine correctness.

the yet-unexplained secure computational marketplace model.

2 Introduction: Cloud-Computing

The appreciation for the significance of modern cryptography presupposes an understanding of the limitations in the contemporary architecture of the data-economy, with regards to privacy and security. It is therefore important to briefly detour through a contextual analysis of existing and future cloud-computing systems.

Cloud-computing refers to network structures that transmit information to a 'central' repository for computation, rather than localised processing. [4] The motivation for these structures are generally related to the either the inherent character or computational requirements of the task, or the need for verifiability and integrity in the computational processes. These systems are increasingly distributed, and the computation is executed amongst various layers of localisation, namely, the 'Edge' and the 'Fog'. The essential character, however, remains identical, whereby data is securely transmitted and then decrypted for analysis, rather than locally computed.

Collectors. [TODO]

IoT. The 'revolution' accompanying the Internet of Things engendered within cloud-technology exposes relevant yet distinct opportunities for the matter at hand. There is an evident motivation to extract the data processed by devices positioned in the household that are capable of measuring and extrapolating previously inaccessible information. There exists, however, a parallel concern that a cloud-computing model reliant on lines of communication for that information would embed a security risk in the architecture. Moreover, there may be a societal resistance to accurately sharing the collected information without strong privacy guarantees. The implications of these privacy and security concerns are influential amongst industry, but a concrete and substantial resolution appears unlikely.

AI. Though not necessarily a consequence of cloud-computing, Artificial Intelligence relies predominantly on layered cloud-architecture for computing machine-learning routines and other essential processes. The utility of AI technology is, furthermore, deeply and inherently connected to 'Big Data' analysis consuming sensitive information for the development of training models, typically executed on the cloud-layer. The relevant dilemma relates to the AI's dependence on the integrity of the computations responsible for analysing data-sets and deriving learning models, which resists the privacy-preserving mechanisms required to ensure the legitimacy and granularity of the collected data.

[4] The 'central' system may itself be distributed, but the operative consideration is the existence of a 'single' endpoint.