

# Protecting (Young) Victims of Cryptocurrency Fraud

## 1 Background

Cryptocurrency has evolved from a niche application developed by activists to a wide-scale form of payment. This trend is likely to accelerate as a result of banks' interests, e.g., Central Bank Digital Currencies (CBDCs), UK government support and plan to become a "global hub" for the cryptocurrency industry [1], and industry initiatives to embed cryptocurrency payments in popular applications, e.g., Revolut.<sup>1</sup> However, cryptocurrencies have also drawn the attention of criminals who want to steal digital currency. According to the UK National Fraud and Cyber Crime Reporting Centre over **£146 million was lost to cryptocurrency fraud** in 2021 [2], a 30% increase since 2020. Santander reports an 87% increase in the volume of cases of this type of fraud in 2022, compared to 2021 [3]. In the UK, victims in the age range of 18–25 account for the highest percentage of reports related to cryptocurrency fraud, i.e., 11%, [2]. The harm resulting from cryptocurrency fraud is not unique to the UK. In the US, the Federal Trade Commission suggested that more than 46,000 people reported losing over \$1 billion in cryptocurrency to fraud, from January 2021 to March 2022 with 34% of the victims in the age range of 18–29 [4]. The true cost of cryptocurrency fraud often extends beyond the immediate financial loss, including law-enforcement time and dealing with the emotional fallout of the fraud. Unfortunately, there have been reports of **suicides** of cryptocurrency fraud victims [5], [6], [7], [8], [9].

## 2 Critical Limitations of State-of-the-art

Cryptocurrencies have several features that distinguish them from traditional payment systems. They are decentralised, allowing competition between different providers to drive down costs to users. They support smart contracts to reduce the trust that needs to be put in the cryptocurrency operator. They have irrevocable payments to give certainty to recipients of funds. While these features have considerable advantages, they also create new risks; for instance, **there is no (central) authority to facilitate recovery from fraud**, whether the result of a flawed smart contract or a malicious party.

Our research project will address this critical limitation by **developing decentralised fraud recovery mechanisms**, whereby victims can obtain redress

while retaining the unique strengths of cryptocurrencies. The scheme will mirror the insurance facility in traditional payment systems, which are paid for by customers through transaction fees, whereby banks will in certain circumstances reimburse fraud victims. In our proposal, the insurance scheme will be decentralised and allow competition between providers and so reduce transaction fees, with smart contracts enforcing fair and transparent reimbursement policies.

This will complement existing research on detecting fraud in cryptocurrencies mainly through methods based on machine learning [10], [11], [12] which (i) lack generality as they cannot support all types of cryptocurrencies, (ii) have been designed to detect specific types of fraud, e.g., Ponzi or pump-and-dump schemes, and (iii) cannot deal with all types of fraud nor defend against human errors. On the other hand, the insurance market's offerings to cover cryptocurrency fraud have not been established yet [13]. Since cryptocurrency is still in its infancy, insurers have been unable to price the risk. The lack of a clear regulatory framework also makes it challenging to unambiguously exclude cryptocurrency-related risks from businesses' insurance policies, potentially leading to losses for insurers [14]. Therefore, to date, (cyber) insurers have had little appetite to cover cryptocurrency.

**Critical Research Gap:** *There exists no scientific study to understand how to devise secure mechanisms that can help victims of cryptocurrency fraud receive compensation for their financial losses.*

Hence, due to the significant amount of money lost to cryptocurrency fraud and the lack of solutions, there is a pressing need to fill the above gap.

## 3 The Research Programme

### 3.1 The Research Aim and Questions

The project will design a novel solution to mitigate cryptocurrency fraud, addressing the question:

**Main Question:** *How can we devise a generic mechanism that can compensate cryptocurrency fraud victims without having to build new cryptocurrency payment systems from scratch?*

This question is of critical importance for three reasons: (a) there exists no technical mechanism to compensate victims of cryptocurrency fraud, (b) different

<sup>1</sup> To save space, all references in this document have been hyperlinked.

cryptocurrency payment systems follow different design principles that affect their capabilities (e.g., only some support arbitrary smart contracts); thus, it is vital to have a mechanism that is generic enough to support victims of cryptocurrency fraud on different platforms, and (c) building new secure cryptocurrency payment system is error-prone and time-consuming.

### 3.2 The Research Objective and Methodology

The research’s main objective is to dramatically improve the protection against cryptocurrency fraud if fraudsters succeed. To achieve it, the research will rely on the following hypothesis.

**Main Hypothesis:** *An insurance-like mechanism can help victims receive reimbursement for their financial losses to cryptocurrency fraud.*

To verify the hypothesis, the research will take three complementary approaches: (1) formal modelling, (2) devising provably secure security protocol, and (3) implementation and evaluation. Also, the research will explore the application of the devised solution beyond protecting victims of cryptocurrency fraud. This project embodies cross-disciplinarity by relying on tools, techniques, theorems, and the project partner’s expertise in consumer protection, human factors, computer science, and mathematics. The research will be organised into the following four Work Packages (WPs).

#### WP1:

**Formal Modelling (month 1–7).** *This WP’s objective is to establish a scientific formal foundation for the core security guarantees that a protocol must offer to reimburse cryptocurrency fraud victims.*

This WP involves developing an accurate mathematical model (task 1). The mathematical model developed in this WP will be based on a **novel combination** of (i) the models and theories used in the insurance industry, (ii) game theory, and (iii) a formal simulation-based paradigm [15], to ensure that any solutions that fit this model can reimburse cryptocurrency fraud victims. The use of models and theories employed in the insurance industry will allow honest victims of (cryptocurrency) fraud to receive compensation for their financial losses to the fraud, via charging interacting parties a certain amount of **premium**. The use of game theory will ensure that the proposed model will capture the real-world settings in which adversarial and fraudulent behaviours are motivated by financial incentives; it will play a vital role in the calculation

of premiums. Also, standard security/cryptographic models (such as the simulation-based paradigm) will ensure that any protocol that realises them would remain secure regardless of adversaries’ strategies.

The research will ensure that the model will be generic and solution-agnostic, so it can be used as a reference point by future researchers who want to develop enhanced provably secure solutions (that can realise the model).

• **Outcome:** It will be the first generic mathematical model for any mechanism that must reimburse victims of cryptocurrency fraud.

#### WP2:

**Devising Security Protocol (month 6–18).** *This WP’s objective is to develop a provably secure protocol that matches the model and allows cryptocurrency fraud victims to receive compensation.*

Briefly, this WP includes two main tasks; developing a protocol (task 2.1) and proving the protocol’s security (task 2.2). Specifically, the research will devise a novel security protocol (i.e., a set of accurate mathematical procedures) that matches the model developed in WP1 and can be used in practice without having to change existing cryptocurrency payment systems. For it to be useful in the real world, the research will ensure that the protocol will satisfy the following fundamental requirements, (a) **generic**, to guarantee that it can protect a broad class of users that may use different cryptocurrencies, (b) **decentralised**, to ensure that it would not negate the decentralisation feature offered by cryptocurrency, (c) **secure**, to ensure that the validity of any computation can be verified, (d) **privacy-preserving**, to ensure that the privacy of those parties who make subjective decisions (e.g., auditors) and users of the system is preserved, and (e) **efficient**, to ensure it does not impose high (computation and communication) costs on users and can scale when the number of users grows. To design a protocol with the above features, the research must address several challenges, outlined below.

1. **Cryptocurrencies Vary in Capabilities.** Each cryptocurrency has different capabilities of supporting computations on transactions and on data stored on them, which significantly affect the way new security features can be integrated into them, without having to rebuild the entire cryptocurrency system from scratch. For instance, Ethereum by supporting arbitrary smart contracts can support (almost) any computation on transactions, whereas Bitcoin supports very limited computa-

tion. To address this challenge and develop a generic mechanism, the research will devise an off-chain protocol that will be run on powerful (but potentially untrusted) cloud computing servers that will need to only read the cryptocurrencies' content, and execute required computations *locally*.

2. **Lack of Transparent Logs.** Currently, messages exchanged between a client and insurance are logged by the insurance and are not accessible to the client without the insurance's collaboration. Even if the insurance provides access to the transaction logs, there is no guarantee that the logs have remained intact. Due to the lack of a transparent logging mechanism, a client or insurance can wrongly claim that (a) it has sent a certain message or (b) it has never received a certain message. Thus, it would be hard for an honest party to prove its innocence. To address this challenge, the research will use a public tamper-evident log to which parties send their messages.
3. **Preserving Privacy.** Although the use of a public logging mechanism is essential in resolving disputes transparently if it does not use a privacy-preserving mechanism, then parties' privacy would be violated. To protect the privacy of parties (from cloud computing), the research will use the efficient "Statement Agreement Protocol" (SAP) developed in [16]. SAP lets parties provably agree on encoding decoding tokens with which they can encode their messages. Later, a party can provide the token to a third party which checks the token's correctness, and decodes the messages. To protect the privacy of independent auditors from other parties, the scheme will ensure that only the final verdict (but not each individual vote) will be revealed. Thus, nobody can link a vote to a specific auditor. To this end, the research will use threshold e-voting protocols developed in [16].
4. **Security.** Although the use of (multiple powerful servers in) the cloud could allow generic and scalable protocols, the cloud itself cannot be trusted with the correctness of computations it runs [17]. The cloud's misbehaviour can have serious repercussions, e.g., can switch the final verdict against a certain client to indicate that it should not receive reimbursement. To address this challenge, the research will use Verifiable Computation (VC) to enforce the cloud to prove the correctness of the computations it runs. To ensure the protocol will remain secure in the case where parties collude with

each other to exploit the system, the research will use the counter-collusion mechanism in [18] that creates distrust between colluding parties.

Overall, the research will use a **novel combination** of cloud computing, e-voting scheme, threshold signature scheme, insurance-like mechanism, game theory, tamper-evident log, SAP, and VC protocol to devise the protocol. The protocol will involve five types of parties; namely, (a) **servers**, each of which is a service provider which accepts cryptocurrency in exchange for the service it provides (e.g., investments in cryptocurrency, exchange of fiat currency with cryptocurrency, or selling items for cryptocurrency), (b) **clients**, each of which is a customer of a server (c) **a set of Cloud Servers ( $\mathcal{CS}$ )**, (d) **a committee of auditors**, consisting of trusted third-party auditors that compile complaints and provide their verdicts, and (e) **an insurance operator ( $O$ )**, a third party whose main role is to register the servers, clients, and auditors with the  $\mathcal{CS}$ .

The idea behind the protocol design will be that each time a client sends digital money to a server, it needs to pay a certain amount of premium to cover the transaction. Later, when it finds out it has been defrauded by the server, it raises a dispute by sending a complaint to  $\mathcal{CS}$ ; the auditors compile the complaint and reimburse the victim if its complaint is valid.

At a high level, the protocol will work as follows. First,  $O$  registers a set of servers, clients, and auditors.  $O$  also fixes a set of public parameters (which will be used to determine insurance premiums) and sends them to  $\mathcal{CS}$ . All data are recorded in a tamper-evident log (e.g., through Proofs of Data Retrievability (PoR) [19]) maintained by  $\mathcal{CS}$ , to ensure the data integrity is protected from  $\mathcal{CS}$ .

Next, each client and  $O$  run SAP to provably agree on a secret key,  $k$ . Also,  $O$  and the auditors jointly generate a public and private key pair for a threshold signature scheme. They do that for each cryptocurrency payment system (e.g., for Ethereum and Bitcoin). Certain threshold signature schemes (e.g., in [20], [21]) let parties (1) generate the signing key without any party learning the key and (2) generate a signature on a message only if at least a certain threshold of the parties signs the message.

After that,  $O$  tags each public key, say  $pk_i$ , with the related cryptocurrency's name, e.g.,  $(pk_i, \textit{Bitcoin})$ . It stores the tagged public keys in the log maintained by  $\mathcal{CS}$ . Any time a client wants to send a certain amount of money to a registered server via a certain cryptocur-

rency, it: (1) retrieves the related public key from  $\mathcal{CS}$  and verifies its correctness, (2) sends the amount to the server via a transaction, say  $t_j$ , and (3) sends a premium to the (account related to the) above public key via a transaction; in this transaction, the client includes  $t_j$  too. In this protocol, the amount of premium will be calculated by  $\mathcal{CS}$  and will be a function of various parameters/factors, e.g., the amount the client wants to send, the amount of coverage that it wants, and the server’s reputation. The research will investigate and take into account other influential factors. The calculation of the premium amount will be determined by the theories and models used in the insurance industry (e.g., the Poisson process and ruin theory) and game theory (e.g., expected utility theory). Also, the protocol will rely on a VC protocol (e.g., [22]) to allow the client to efficiently verify the correctness of the premium amount calculated by  $\mathcal{CS}$ .

When a client realises it has been defrauded by one of the registered servers, it raises a dispute, by sending an encrypted complaint to  $\mathcal{CS}$ , where  $k$  is the key used to encrypt the message. The client sends directly to the auditors proof asserting that a correct key has been used to encrypt the message. The client can include in the complaint pieces of evidence too, e.g., details of off-chain interactions/transactions it had with it the server, and the details of the transaction about the premium it paid.

Each auditor verifies the proof. If the verification is passed, then it decrypts and compiles the complaint. The auditor (i) checks whether the client has paid the premium, (ii) generates a verdict, (iii) encodes the verdict (using the efficient e-voting that we developed in [16]), and (iv) sends the result to  $\mathcal{CS}$  which can generate a transaction signed by a predefined threshold of the auditors (using a threshold signature) if the threshold of them voted to compensate the client. In this case,  $\mathcal{CS}$  sends the signed transaction to the cryptocurrency network and accordingly the client will receive compensation.

To ensure that the protocol will remain secure if a client and server/auditor collude with each other to exploit the system and increase their utility, the research will use the game theory-based counter-collusion mechanism in [18] which creates distrust between the colluding parties and incentives a party to betray its counter-party for a higher payoff.

The research will **prove** the security of the devised protocol and formally show that it will fit the model in WP1.

• **Outcome:** It will be the first secure generic protocol that will help victims of cryptocurrency fraud receive compensation for their financial losses to the fraud.

### WP3:

#### Implementation and Evaluation (month 16–28).

*This WP’s objective is to implement the protocol that will be devised in WP2 and analyse the protocol’s concrete costs.*

This WP includes two primary tasks: implementing the protocol (task 3.1) and evaluating the protocol’s performance (task 3.2). Specifically, the research will implement the protocol that will be devised in WP2, for evaluation and establishing concrete parameters of the protocol. The implementation will be developed in C++, as there exist various cryptographic libraries written in this programming language. The implementation will utilise the “Cryptopp” library for cryptographic primitives. We have already implemented SAP, tamper-evident logging, and threshold e-voting protocols (see [23], [24], and [25] respectively for the source code). However, other sub-protocols (e.g., threshold signature, client, and server) and the main (wrapping) protocol will be implemented. The protocol will be implemented in the form of two packages (thus each task will be split into two subtasks).

In the first package (task 3.1.1), the research will use (1) local servers, i.e., the UCL Computer Science High-performance Clusters, instead of actual cloud servers and (2) a cryptocurrency test net, instead of using an actual cryptocurrency system. This approach will allow conducting of various tests and refinements for free without having to use the actual cloud and cryptocurrency network. The research (in task 3.2.1) will evaluate the run-time of different parties when the number of victims is low and high (e.g., in the range of [1, 1000]). It is expected that auditor-side computation (in particular executing threshold signature scheme) will be a bottleneck when the number of victims is high at any given time, as it would involve modular exponentiations that are usually computationally expensive. The research will perform a concrete cost evaluation to verify the above hypothesis and remove the bottleneck, e.g., by finding an optimal number of auditors and/or servers.

In the second package (task 3.1.2), the research will run the improved implementation using actual cloud servers and cryptocurrency and evaluates parties’ costs in a real-world setting (task 3.2.2).

- **Outcome:** WP3’s outcome will be (i) two open source packages that implement the protocol, and (ii) performance evaluations of the protocol.

#### WP4:

##### **Exploring Further Applications (month 27–36).**

*This WP’s objective is to explore other applications of the protocol.*

The research will explore further applications of the protocol (from WP2); specifically, the research will investigate (i) insuring the VS: Verifiable Service (task 4.1), and (ii) insuring the MPC: secure Multi-Party Computation (task 4.2). The idea is that for each run of VS/MPC, the participants of VS/MPC pay a premium. They will be compensated if they can prove that they acted honestly but their counter-parties acted maliciously.

A Verifiable Service (VS) is a two-party (client-server) protocol in which a client chooses a function,  $\mathcal{F}$ , and provides (an encoding of)  $\mathcal{F}$ , its input  $u$ , and a query  $q$  to a server potentially malicious. The server is expected to evaluate  $\mathcal{F}$  on  $u$  and  $q$  (i.e.,  $\mathcal{F}(u, q) = o$ ) and respond with the output  $o$  and proof  $\pi$  asserting that the output was generated correctly. Given  $(o, \pi)$ , the client verifies that the output is indeed the output of the function computed on the provided input. In VS, either the computation (on the input) or both the computation and storage of the input are delegated to the server. “Proofs of Retrievability” and “Verifiable Computation” are two examples of VS’s instantiations. A serious limitation of state-of-the-art VS is that a client receives no compensation when the server does not deliver the service. However, this is not suitable for **mission-critical data or computation**, as other crucial services may depend on the result that an honest server would provide.

The research will adjust the protocol, devised in PW2, to let an honest client in VS receive compensation if the server does not deliver the promised service. To do so, the research will (a) use publicly verifiable VS that is also secure against a malicious client, i.e., Verifiable Service with IDentifiable abort (VSID) in [26], and (b) require the server in VS to sign the pair  $(o, \pi)$  that it sends to the client. In this modified protocol, when the client rejects the server’s proof/output, it forwards the signed  $(o, \pi)$  to the auditors (of the WP2’s protocol) that verify the output’s correctness, ensure the client has acted honestly in VS, and decide whether the client should be compensated. Thus, the research will modify the auditor-side verification

algorithm of the protocol in WP2 and uses an appropriate VS (i.e., VSID) to ensure that (1) an honest client will receive compensation if the server provides invalid proof, and (2) a malicious client cannot exploit the protocol and receive compensation that it does not deserve. The research will formally prove the security of the resulting protocol (task 4.3).

Furthermore, the research will explore a use case of WP2’s protocol in secure Multi-Party Computation (MPC), which has been drawing considerable attention from researchers and industries. MPC is a cryptographic protocol that lets parties jointly run a certain computation on their private inputs without being able to learn anything beyond the result. It is known that during the execution of MPC some parties may misbehave and prevent honest parties from learning correct results. To date, the only mechanism that lets an honest party receive compensation (for not receiving a correct result) in MPC is the “deposit paradigm” [27]. The deposit paradigm requires all parties to deposit a certain amount of cryptocurrency before the execution of MPC. This means that (i) all parties must have a certain amount of cryptocurrency, (ii) all parties must deposit the required amount, and (iii) the amount each party deposits must be proportionate to the total number of parties participating in the protocol (which can be high when the number of participants is high). Nevertheless, these strong requirements limit MPC’s real-world applications.

The research will modify the protocol of PW2 to let parties in MPC receive compensation if they do not receive correct results, without requiring them to deposit any form of money. The main change will be made to the auditor-side verification algorithm of the protocol in WP2. In the modified protocol, each auditor will check the validity of the proofs often parties in MPC provide to each other to prove that they have acted honestly. Specifically, each auditor will check if the claimant parties acted honestly, whereas their counter-parties have not. The research will also prove the security of the resulting protocol (task 4.4).

- **Outcome:** The outcome of WP4 will be two new applications of the protocol (developed in WP2). Specifically, the WP2’s protocol will be adjusted and then applied to VC and MPC schemes to allow an honest result recipient to receive compensation if it could not approve result correctness.

### 3.3 Novelty

The proposed work presents a significant extension for cryptocurrency which currently lacks any mechanism

to protect victims of cryptocurrency fraud. Our prior work [16] established a scientific foundation to protect customers who fall victim to “Authorised Push Payment” (APP) fraud. We can now develop mechanisms to help victims of cryptocurrency fraud be compensated for their losses.

Concretely, first, this work presents a novel mathematical model for reimbursing cryptocurrency fraud victims by relying on a unique combination of (i) the models and theories used in the insurance industry, (ii) game theory, and (iii) a formal simulation-based paradigm. Second, the design of the generic security protocol, that can work with any cryptocurrency, is very novel. The protocol will rely on a unique combination of cloud computing, e-voting scheme, threshold signature scheme, insurance-like mechanism, game theory, tamper-evident log, PETs, and VC to satisfy security and efficiency requirements. Such a combination, to our knowledge, has not been seen before.

Third, the idea of applying the WP2’s protocol to VC and MPC is novel too; as it will let an honest client receive compensation for not receiving valid results/proofs, without requiring any deposit. The existing scheme that lets a party receive compensation, in MPC, requires every party to deposit cryptocurrency proportionate to the total number of parties participating in the protocol, which significantly limits its real-world application and adoption.

## 4 National Importance and Impacts

### 4.1 Academic Importance and Impact

This research will significantly improve the protection level of cryptocurrency fraud victims. Therefore, the research community in cryptocurrency and cryptography will directly benefit from it.

The mathematical model, in WP1, will serve as a solid basis for systematic evaluations of victim protection levels of other payment systems. The security protocols, in WP2, will enable researchers to understand which tools, techniques, and computational hardness assumptions must be relied upon to build a generic protocol that can help victims of cryptocurrency fraud to receive compensation. The implementation/benchmark, in WP3, presents a contribution to software engineers, as they will be provided with a basis for building other related protocols in the future. The protocols designed in WP4 for the first time will

show how to insure the verifications’ outputs, in VS and MPC. To ensure the research will have a maximum academic impact, the researchers will: (a) **publish and present research findings** at scholarly conferences, and (b) **maintain an online anti-fraud scientific database** on the project’s website, acting as a central hub for the latest related publications.

### 4.2 Societal Importance and Impact

The result of this research (when adopted) will benefit UK residents from financial and mental health perspectives, by protecting victims of fraud. To share the research findings with the public, the researchers will (1) **maintain public-facing communication channels**, via creating social media posts, and (2) **deliver webinars at schools and universities**, to inform young people about online payment fraud.

### 4.3 Economic Importance and Impact

This research (when adopted) will directly benefit the UK economy, as people will lose far less to fraud. It is expected this research to yield **new Fintech insurance startups** that will provide users of cryptocurrency with protection against cryptocurrency fraud. This has the potential to make the UK a base for international investment in this *new line of the insurance industry*. To ensure the research will have adequate economic impacts, the researchers will (1) **share the research findings with UK regulators** (e.g., FCA) and (2) **collaborate with UCL Public Engagement and Media**, to seek effective ways to draw investors’ attention to the findings.

### 4.4 Strategic Fit

The research programme has the potential for substantial impact across multiple EPSRC’s priority areas; such as Information and Communication Technologies (ICT) and Mathematical Sciences themes. It is aligned with UKRI’s Strategic Priority called “Protecting Citizens Online” and with **Priority 5.1 of the UKRI Strategy 2022 to 2025** which aims to build a secure world by **enhancing national security across virtual spaces**. It is also aligned with **Objective 5: world-class impacts of the UKRI Strategy 2022 to 2025**, on investing to secure a competitive advantage in emerging technologies and create opportunities for UK businesses in expanding global markets including **Fintech**.