Proof-of-Burn

Kostis Karantias[1], Aggelos Kiayias[1,3], and Dionysis Zindros[1,2]

[1] IOHK
[2] University of Athens
[3] University of Edinburgh

Abstract. *Proof-of-burn* has been used as a mechanism to destroy cryptocurrency in a verifiable manner. Despite its well known use, the mechanism has not been previously formally studied as a primitive. In this paper, we put forth the first cryptographic definition of what a proof-of-burn protocol is. It consists of two functions: First, a function which generates a cryptocurrency address. When a user sends money to this address, the money is irrevocably destroyed. Second, a verification function which checks that an address is really unspendable. We propose the following properties for burn protocols. *Unspendability*, which mandates that an address which verifies correctly as a burn address cannot be used for spending; *binding*, which allows associating metadata with a particular burn; and *uncensorability*, which mandates that a burn address is indistinguishable from a regular cryptocurrency address. Our definition captures all previously known proof-of-burn protocols. Next, we design a novel construction for burning which is simple and flexible, making it compatible with all existing popular cryptocurrencies. We prove our scheme is secure in the Random Oracle model. We explore the application of destroying value in a legacy cryptocurrency to bootstrap a new one. The user burns coins in the source blockchain and subsequently creates a proof-of-burn, a short string proving that the burn took place, which she then submits to the destination blockchain to be rewarded with a corresponding amount. The user can use a standard wallet to conduct the burn without requiring specialized software, making our scheme user friendly. We propose burn verification mechanisms with different security guarantees, noting that the target blockchain miners do not necessarily need to monitor the source blockchain. Finally, we implement the verification of Bitcoin burns as an Ethereum smart contract and experimentally measure that the gas costs needed for verification are as low as standard Bitcoin transaction fees, illustrating that our scheme is practical.

1 Introduction

Since the dawn of history, humans have entertained the defiant thought of money burning, sometimes literally, for purposes ranging from artistic effect [9] to protest [23], or to prevent it from falling into the hands of pirates [22,12]. People did not shy away from the practice in the era of cryptocurrencies. Acts of money burning immediately followed the inception of Bitcoin [24] in 2009, with the first recorded instance of intentional cryptocurrency destruction taking place on August 2010 [29], a short three months after the first real-world transaction involving cryptocurrency in May 2010 [8]. For the first time, however, cryptocurrencies exhibit the unique ability for money burning to be provable retroactively in a so-called *proof-of-burn*.

First proposed by Iain Stewart in 2012 [28], proof-of-burn constitutes a mechanism for the destruction of cryptocurrency irrevocably and provably. The ability to create convincing proofs changed the practice of money burning from a fringe act to a rational and potentially useful endeavour. It has since been discovered that metadata of the user's choice—a so-called *tag*—can be uniquely ascribed to an act of burning, allowing each burn to become tailored to a particular purpose. Such protocols have been used as a consensus mechanism similar to proof-of-stake (Slimcoin [25]), as a mechanism for establishing identity (OpenBazaar [26,34]), and for notarization (Carbon dating [13] and OpenTimestamps [30]). A particularly apt use case is the destruction of one type of cryptocurrency to create another. In one prolific case, users destroyed more than 2,130.87 BTC ($1.7M at the time, $21.6M in today's prices) for the bootstrapping of the Counterparty cryptocurrency [1].

While its adoption is undeniable, there has not been a formal treatment for proof-of-burn. This is the gap this work aims to fill.

Our contributions. A summary of our contributions is as follows:

(i) **Primitive definition.** Our definitional contribution introduces proof-of-burn as a cryptographic primitive for the first time. We define it as a protocol which consists of two algorithms, a burn address *generator* and a burn address *verifier*. We put forth the foundational properties which make for secure burn protocols, namely *unspendability*, *binding*, and *uncensorability*. One of the critical features of our formalization is that a tag has to be bound cryptographically with any proof-of-burn operation.
(ii) **Novel construction.** We propose a novel and simple construction which is flexible and can be adapted for use in existing cryptocurrencies, as long as they use public key hashes for address generation. To our knowledge, all popular cryptocurrencies are compatible with our scheme. We prove our construction secure in the Random Oracle model [6].
(iii) **Bootstrapping mechanism.** We propose a cryptocurrency proof-of-burn bootstrapping mechanism which for the first time does not require target blockchain miners to connect to external blockchain networks.