```
BIP: 126
Title: Best Practices for Heterogeneous Input Script Transactions
Author: Kristov Atlas <kristov@openbitcoinprivacyproject.org>
Comments-Summary: No comments yet.
Comments-URI: https://github.com/bitcoin/bips/wiki/Comments:BIP-0126
Status: Draft
Type: Informational
Created: 2016-02-10
License: PD
```

## Abstract

When a Bitcoin transaction contains inputs that reference previous transaction outputs sent to different Bitcoin addresses, personally identifiable information of the user will leak into the blockchain in an uncontrolled manner. While undesirable, these transactions are frequently unavoidable due to the natural fragmentation of wallet balances over time.

This document proposes a set of best practice guidelines which minimize the uncontrolled disclosure of personally identifiable information by defining standard forms for transactions containing heterogenous input scripts.

## Copyright



## Definitions

- **Heterogenous input script transaction (HIT)**: A transaction containing multiple inputs where the scripts of the previous transaction outputs being consumed are not identical (e.g. a transaction spending outputs which were sent to more than one Bitcoin address)
- **Unavoidable heterogenous input script transaction**: A HIT created as a result of a user's desire to create a new output with a value larger than the value of his wallet's largest existing unspent output
- **Intentional heterogenous input script transaction**: A HIT created as part of a user protection protocol for reducing uncontrolled disclosure of personally-identifying information (PII)

Throughout this procedure, when input scripts are evaluated for uniqueness, "input script" should be interpreted to mean, "the script of the previous output referenced by an input to a transaction".

## Motivations

The recommendations in this document are designed to accomplish three goals:

1. Maximise the effectiveness of user-protecting protocols: Users may find that protection protocols are counterproductive if such transactions have a distinctive fingerprint which renders them ineffective.
2. Minimise the adverse consequences of unavoidable heterogenous input transactions: If unavoidable HITs are indistinguishable from intentional HITs, a user creating an unavoidable HIT benefits from ambiguity with respect to graph analysis.
3. Limiting the effect on UTXO set growth: To date, non-standardized intentional HITs tend to increase the network's UTXO set with each transaction; this standard attempts to minimize this effect by standardizing unavoidable and intentional HITs to limit UTXO set growth.

In order to achieve these goals, this specification proposes a set of best practices for heterogenous input script transaction creation. These practices accommodate all applicable requirements of both intentional and unavoidable HITs while maximising the effectiveness of both in terms of preventing uncontrolled disclosure of PII.

In order to achieve this, two forms of HIT are proposed: Standard form and alternate form.

## Interaction with Other Procedures

Applications which wish to comply both with this procedure and BIP69 should apply this procedure prior to applying BIP69.

## Standard form heterogenous input script transaction

### Rules

A HIT is Standard form if it adheres to all of the following rules:

1. The number of unique output scripts must be equal to the number of unique inputs scripts (irrespective of the number of inputs and outputs).
2. All output scripts must be unique.
3. At least one pair of outputs must be of equal value.
4. The largest output in the transaction is a member of a set containing at least two identically-sized outputs.

### Rationale

The requirement for equal numbers of unique input/output scripts instead of equal number of inputs/outputs accommodates user-protecting UTXO selection behavior. Wallets may contain spendable outputs with identical scripts due to intentional or accidental address reuse, or due to dusting attacks. In order to minimise the adverse consequences of address reuse, any time a UTXO is included in a transaction as an input, all UTXOs with the same spending script should also be included in the transaction.

The requirement that all output scripts are unique prevents address reuse. Restricting the number of outputs to the number of unique input scripts prevents this policy from growing the network's UTXO set. A standard form HIT transaction will always have a number of inputs greater than or equal to the number of outputs.

The requirement for at least one pair of outputs in an intentional HIT to be of equal value results in optimal behavior, and causes intentional HITs to resemble unavoidable HITs.

## Alternate form heterogenous input script transactions

The formation of a standard form HIT is not possible in the following cases:

1. The HIT is unavoidable, and the user's wallet contains an insufficient number or size of UTXOs to create a standard form HIT.
2. The user wishes to reduce the number of utxos in their wallet, and does not have any sets of utxos with identical scripts.

When one of the following cases exist, a compliant implementation may create an alternate form HIT by constructing a transaction as follows:

### Procedure

1. Find the smallest combination of inputs whose value is at least the value of the desired spend.
   (a) Add these inputs to the transaction.
   (b) Add a spend output to the transaction.
   (c) Add a change output to the transaction containing the difference between the current set of inputs and the desired spend.
2. Repeat step 1 to create a second pair of outputs, where one output has the same value as the spend output of the previous step.
3. (optional) Repeat step 2 until the desired number of inputs have been consumed and/or the desired number outputs have been created.
4. Adjust the change outputs as necessary to pay the desired transaction fee.

Clients which create intentional HITs must have the capability to form alternate form HITs, and must do so for a non-zero fraction of the transactions they create.

### Rules

An HIT formed via the preceding procedure will adhere to the following conditions:

1. The number of unique inputs scripts must exceed the number of output scripts.
2. All output scripts must be unique.
3. At least one pair of outputs must be of equal value.
   (a) "Standard outputs" refers to the set of outputs with equal value

   (b) "Standard value" refers to the value of the standard outputs
   (c) "Change outputs" refers to all outputs which are not standard outputs
4. For a HIT containing n standard outputs, there must exist at least one possible way to organize the inputs and outputs into n sets, where all sets satisfy the following:
   (a) The set contains one or more inputs, exactly one standard output, and exactly one change output
   (b) An input or output that appears in one set must not appear in any other set
   (c) The sum of the inputs in the set minus the value of the change output is equal to the standard value with a tolerance equal to the transaction fee.
   (d) Change outputs with a value of zero (virtual change outputs) are permitted. The are defined for the purpose of testing whether or not a HIT adheres to this specification but are not present in the version of the transaction which is broadcast to the network.

## Non-compliant heterogenous input script transactions

If a user wishes to create an output that is larger than half the total size of their spendable outputs, or if their inputs are not distributed in a manner in which the alternate form procedure can be completed, then the user can not create a transaction which is compliant with this procedure.

## Reference

- BIP69 - Lexicographical Indexing of Transaction Inputs and Outputs