

Interactive Demo: Visualization for Bitcoin Mining Pools Analysis

Natkamon Tovanich
IRT SystemX and
Université Paris-Saclay, LISN
Palaiseau, France
natkamon.tovanich@irt-systemx.fr

Nicolas Soulié
Université Paris-Saclay,
Univ Evry, IMT-BS, LITEM
Evry-Courcouronnes, France
nicolas.soulie@imt-bs.eu

Nicolas Heulot
IRT SystemX
Palaiseau, France
nicolas.heulot@irt-systemx.fr

Petra Isenberg
Université Paris-Saclay,
CNRS, Inria, LISN
Gif-sur-Yvette, France
petra.isenberg@inria.fr

Abstract—We demonstrate an interactive visualization tool to analyze Bitcoin mining pools. The tool allows analysts to see the evolution of mining pools distribution over time and relationships with external variables, i.e., Bitcoin statistics and news headlines. Moreover, we also display information about pool hopping among mining pools to help understand the internal dynamics of miners.

Index Terms—Bitcoin, Bitcoin mining, mining pools, pool hopping, visual analytics

I. INTRODUCTION

Bitcoin mining is a critical activity that keeps the whole Bitcoin system secure, valid, and stable. Mining pools have emerged to ensure that the Bitcoin system keeps running while allowing miners to benefit from their service through a more stable and predictable income. We developed an interactive visualization tool to help economists explore the emergence and evolution of mining pools. The tool has several new features of potential interest to both Bitcoin practitioners and researchers.

II. DATA PREPARATION

For our tool, we extracted multiple measurements about individual mining pools, mining reward distributions, and external data from difference data sources. The data is publicly available at [1] (except 3 and 4).

- 1) *Mining pools*: We tagged each coinbase transaction with a mining pool based on the coinbase string patterns and known pool addresses from [2]–[4]. We aggregated each mining pool’s monthly statistics, e.g., market share, hash rate, and total reward in BTC and USD.
- 2) *Mining pool descriptions*: We tracked the characteristics of mining pools from the edit history in the Bitcoin Wiki [5]. We then manually validated the data with the discussion of mining pools in the Bitcointalk forum.
- 3) *Bitcoin network statistics*: We obtained the Bitcoin statistics data (e.g., market price, transaction volume, and total hash rate) from Blockchain.info and averaged the statistics for each month.

This research work has been carried out under the leadership of the *Institute for Technological Research SystemX*, and therefore granted with public funds within the scope of the French Program *Investissements d’Avenir*.

- 4) *Bitcoin news*: We crawled the list of news headlines in the Press forum of Bitcointalk. We used spaCy text processing and STTM short text topic modeling [6] libraries to group news into distinct topics.
- 5) *Pool hopping*: We derived a method to track the number of miners who participated in multiple mining pools (“cross-pooling” miners) and calculated the reward they received for each pool [7]. We calculated the total reward of cross-pooling miners associated between mining pools over a time interval.

III. INTERACTIVE VISUALIZATION TOOL

The interactive visualization tool, as displayed in Fig. 1, consists of five coordinated components: (A) a time filter view, (B) a mining statistics view, (C) a detail view, (D) a Bitcoin news view, and (E) a cross pooling view.

A. *The time filter view* is a line chart showing Bitcoin statistics over time from 2011 when the first known mining pool was detected. Different statistics can be selected and shown and then the time filter can be used to drill down to a time period of interest. Once a time range has been selected, the tool will update all other views to the selected time interval.

B. *The Bitcoin mining distribution view* is the main view for exploring the emergence and evolution of the Bitcoin mining pools. It presents the temporal distribution of mining pools as a ribbon chart—a stacked bar chart showing with ribbons connecting the same mining pool over time. Each mining pool is represented as a bar whose height is relative to the measurement (i.e., hash rate, market share, or total reward) in a month. The color indicates categorical data (i.e., mining pool, payout scheme, or location). Users can select the configurations of the chart on the top panel. The bars are sorted by the total measurement of the category from the highest to the lowest value for each month. Within each category, pools are sorted by their measurement. This sorting helps us to identify large mining pools or categories that dominated mining power. Ribbons help to track a mining pool’s ranking connection between months. Analysts can select mining pools on the left panel or drag the brush along the x and y axes to highlight a range of time and mining pools.

C. *The detailed view* shows the mining statistics as separate bar charts for each mining pool. This panel allows analysts to

