Tokenization of Loan Portfolios in Numbers
(Early Research Report)

Student Name: Waralak Pariwatphan
Advisor Name: Yury Yanovich

1. Introduction

1.1. Overview

In quantitative finance, portfolio optimization based on Markowitz's framework is well recognized. The fundamental concept is to diversify financial assets in a portfolio in order to decrease portfolio risk and choose the most efficient portfolio. Blockchain is a notable technology that allows for the distribution of digital ledgers without the requirement for a central entity to keep its state secured. Tokenization, which can fractionize assets into tokens and also transform ownership asset rights into that digital token, is the important characteristic.

Decentralized Finance (DeFi) is a financial technology based on blockchain that eliminates the need for intermediary and centralized financial organizations in financial transactions. MakerDao is the world's leading DeFi that enables the generate stable coin Dai and allows users to lock their cryptocurrencies as collateral and borrow Dai. According to these blockchain-based loan portfolios, we aim to investigate the optimal portfolio tokenization by implementing tokenization algorithms and evaluating the tokenized asset fraction in the portfolio package.

1.2. Aim and purpose

The aim of early research is to analyze the MakerDao dataset and apply tokenization to real loan portfolio dataset. Moreover, we also study the workflow of smart contracts and run pay out functions.

1.3. Methodology

First, we used BigQuery to collect peer-to-peer lending data from MakerDAO. The MakerDao dataset was stored in Ethereum transactions, and we could access it using smart contracts (Author, 2022). Because the data was not in human-readable format, we had to transform the raw data and utilize Web3 in Python to extract the important information (Huang, 2021). Second, we estimated the MakerDAO dataset's economic parameters, which included the Loss Given Default (LGD), Effective Interest Rate (EIR), and Probability of Default (PD). Furthermore, we tokenized the dataset by using tokenization algorithms (portfolio sold-out problem) and implemented the results. Finally, we demonstrated a tokenization workflow on a blockchain using the Ethereum Testnet and the tokenization algorithm results as input. The tokenization method was then applied to smart contracts.

2. Literature review

2.1. Geometric Brownian motion

Geometric Brownian motion is a stochastic process used to represent random motions. It is similar to a random walk. It is commonly applied in the study of financial markets, such as stock prices, foreign exchange, financial assets, and so on. A Wiener stochastic process is the standard Brownian motion (Pishro-Nik, 2014). Let $W(t)$ is the value in $\mathbf{R}$ when $t \in [0, \infty)$, the properties of Wiener process are

1. $W(0) = 0$.

2. for all $0 \leq t_1 < t_2, W(t_2) - W(t_1) \sim N(0, t_2 - t_1)$.

3. $W(t)$ has independent increments. For all $0 \leq t_1 < t_2 < ... < t_n$ the random variables

$$W(t_2) - W(t_1), W(t_3) - W(t_2), ..., W(t_n) - W(t_{n-1})$$

are independent.

4. $W(t)$ paths are continuous.

First passage times of levels (Shiryaev, 1999) are applied in a various applications. The time when the process hits an arbitrary threshold a for the first time is a random variable, according to Brownian Motion theory. It is possible to assume that the process begins at a certain position $x_0 > 0$ and is bounded to the positive semi-axis by an absorbing barrier x = 0. Assume that $a > 0$ and let $T_a = \inf \{t \geq 0 : B_t = a\}$. Then we will have

$$\mathrm{P}\left(T_a < t\right) = \mathrm{P}\left(\sup_{s \leq t} B_s > a\right).$$

Applying the reflection principle of D. Andre, it follows that

$$\mathrm{P}\left(T_a < t\right) = 2\mathrm{P}\left(B_t \geq a\right).$$

Since

$$\mathrm{P}\left(B_t \geq a\right) = \frac{1}{\sqrt{2\pi t}} \int_a^{\infty} e^{-\frac{x^2}{2t}} dx,$$

So,

$$P\left(T_a < t\right) = \int_0^t \frac{a}{\sqrt{2\pi s^3}} e^{-\frac{a^2}{2s}} ds$$

2.2. Homogeneous tokenization algorithm

Homogeneous systems have useful properties for analysis and design, such as stabilization, convergence rates, and trajectory scalability (Chaleenutthawut et al., 2021). The algorithms' outputs are the number of created tokens M and the composition matrix $C_M = (\vec{c}_1 | \ldots | \vec{c}_M)$ with M columns and N rows. The amount of assets that make up into the packages $\bar{a} = \sum_{m=1}^{M} \vec{c}_m$.

2.2.1. Discrete algorithm

For discrete homogeneous tokenization, there are the set of M packages

$$C_M = (\vec{c}_1 | \ldots | \vec{c}_M) \in \left\{0, \frac{1}{k}\right\}^{N \times M}$$

with the variance reduction parameter k of the portfolio ($\vec{A}$, K) if

$$\sum_{m=1}^{M} \vec{c}_m \leq \vec{A}$$

and

$$\forall m \in \bar{M} : \left\|\vec{c}_m\right\|_1 = 1.$$

Table 1. Descriptive data for MakerDao dataset

STATISTICS	DESCRIPTIVE DATA
DATES	15/11/2019 - 22/06/2022
NUMBER OF ASSETS	54
NUMBER OF DEBTS	22,830
NUMBER OF BORROWERS	18,142
DAI BORROWED AMOUNT (USD)	879,674,958.44
ASSET LOCKED AS COLLATERAL (USD)	2,010,197,300.42

2.2.2. CONTINUOUS ALGORITHM

For continuous homogeneous tokenization, there are the set of M packages

$$C_M = (\vec{c}_1|\dots|\vec{c}_M) \in \mathbf{R}^{N\times M}$$

with the variance reduction parameter k of the portfolio $(\vec{A},\ \mathrm{K})$ if

$$\forall m \in \bar{M} : \|\vec{c}_m\|_2 \leq \frac{1}{k} \wedge \|\vec{c}_m\|_1 = 1$$

2.2.3. COMPACT CONTINUOUS ALGORITHM

The portfolio sold-out problem formulation is

$$M \to \max_{M,\mathrm{C}_M} .$$

The constraints are:

$$\vec{c}^T\,\mathrm{K}\vec{c}_m \leq \sigma^2,$$

$$\sum_{m=1}^{M} \vec{c}_m \leq \vec{A},$$

$$\|\vec{c}\|_1 = 1.$$

There is a proof that an optimal matrix $\bar{C}_M$ with $\vec{c_1} = \vec{c_2} = ... = \vec{c_M} = constant$, where $\vec{\bar{c}} = \frac{1}{M}\sum_{m=1}^{M}\vec{c}_m$. Thus, original portfolio sold-out problem can be reformulated as

$$M \to \max_{\mathrm{C}_M}$$

with the constraints

$$\vec{a}^T K \vec{a} \leq \sigma^2 \cdot \|\vec{a}\|_1^2,$$

$$\vec{0} \leq \vec{a} \leq \vec{A}.$$

The problem was solved an open source python package MEALPY(Thieu & Mirjalili, 2022) which allow user to specify the convex problem in intuitive syntax.

3. Results

3.1. MakerDao overview data

We collected the data from 15 November 2019 to 22 June 2022. After analyzing data, we found that the assets utilized as collaterals totaled 54 assets (cryptocurrencies). There were 22,830 debt incidents reported by 18,142 borrowers. The entire amount borrowed was 879,674,958.44 US dollars, while the collateral assets were 2,010,197,300.42 US dollars.

3.2. MakerDao financial statistics

As illustrated in Figure 3.1, the probability of default increases as the DAI/ETH rate falls. Borrowers' portfolios will be liquidated as default status if the value of their portfolios is less than the debt plus collateral ratio, and the average time until default is 100.88 days.

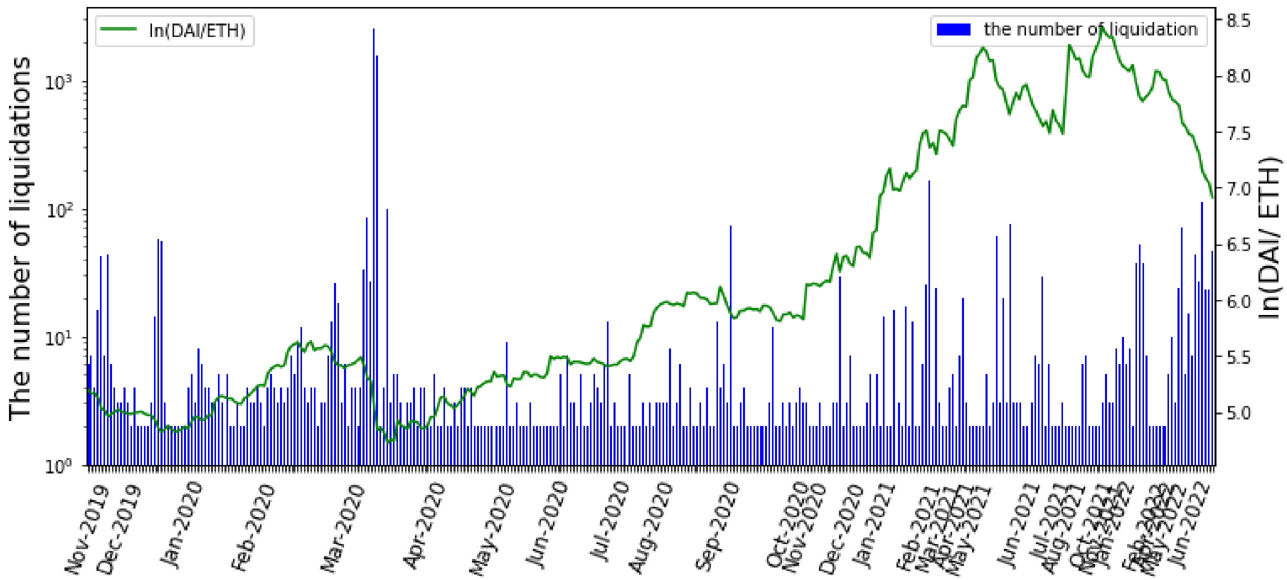

Figure 1. Number of liquidations and DAI/ETH rate

3.3. Tokenization algorithms

Figure 3.2 shows that the tokenized fraction from the compact continuous tokenization algorithm outperformed the discrete and continuous algorithms at the same risk level.

3.4. Deploying in smart contract

The smart contract was deployed, which included the mint package, send package, and call pay out function.

4. Conclusion

The dataset for real loan portfolios is gathered by matching MakerDao contracts with functions that hold the information. Data is supported with economic parameters once it has been analyzed. The probability of default increases while the DAI/ETH rate falls. Furthermore, the data is tokenized, and the tokenized assets fraction graph demonstrates that the higher the proportion, the better. Finally, the tokenization process is described on blockchain, consisting of four major components: smart contract deployment, mint package, send package, and call pay out function.

5. Future Plan

To find additional methods for numerical optimization or create smart contracts for tokenization (payout function).

References

Author, N. N. Vat - detailed documentation, the maker protocol's core accounting system, 2022.

Chaleenutthawut, Y., Davydov, V., Kuzmin, A., and Yanovich, Y. Practical blockchain-based financial assets tokenization. *ICBTA 2021: 2021 4th International Conference on Blockchain Technology and Applications*, pp. 51–57, 2021.

Huang, Y. Decoding ethereum smart contract data, 2021.

Pishro-Nik, H. (ed.). *Introduction to Probability, Statistics, and Random Processes*. Kappa Research, 2014.

Shiryaev, A. N. (ed.). *Essentials of Stochastic Finance: Facts, Models, Theory*. World Scientific, 1999.

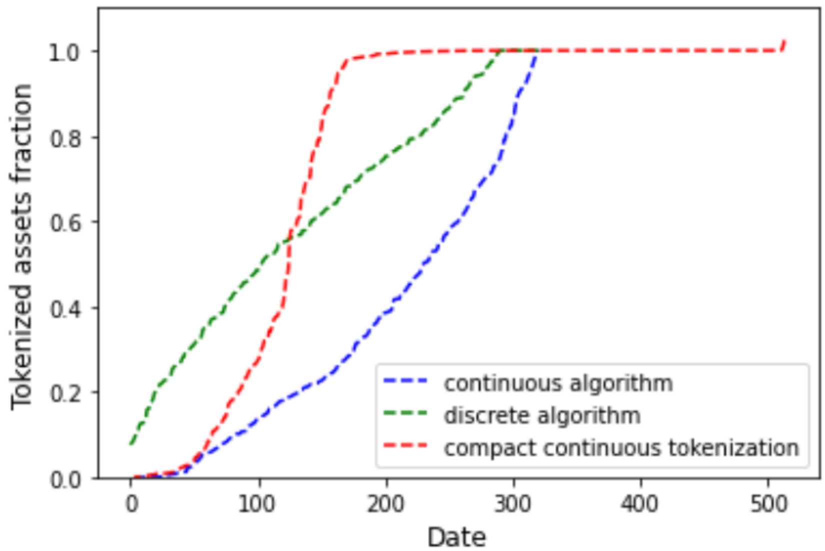

Figure 2. Tokenized assets fraction for MakerDao data

Thieu, N. V. and Mirjalili, S. MEALPY: a Framework of The State-of-The-Art Meta-Heuristic Algorithms in Python, June 2022. URL `https://doi.org/10.5281/zenodo.6684223`.

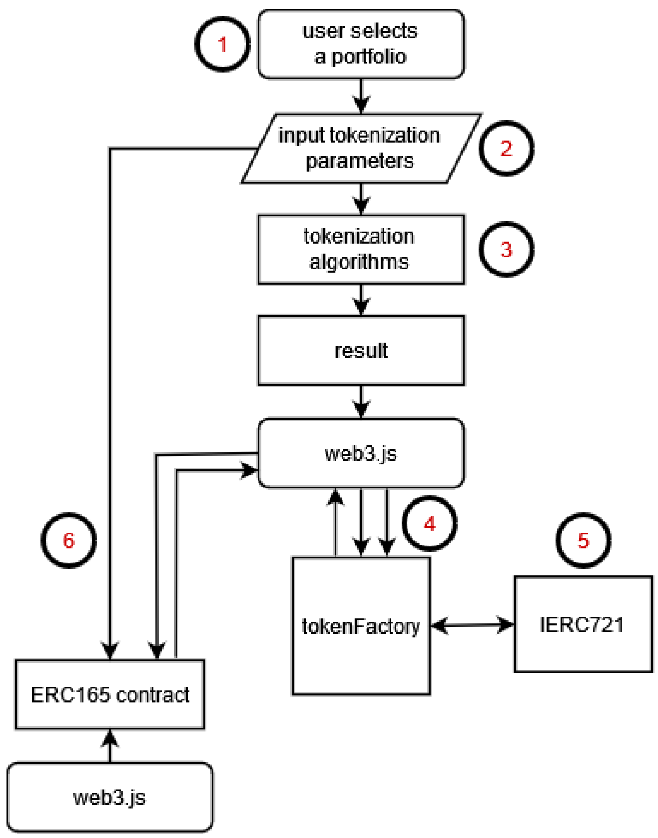

Figure 3. Tokenization ans smart contract workflow