THEORY

The DeFi Prediction Protocol

Whitepaper authors:

James Levy | james@poolside.finance

Last Revision: December 30, 2020

Overview

Theory is an open protocol that transforms decentralized finance activity into predictions and recommendations.

While the Graph protocol is designed for curating and indexing blockchain data, Theory is designed for producing useful predictions from this data. Just as all major cloud computing platforms now offer prediction API products to augment their hosted database products and these two types of products achieve operational efficiency through a shared backend, Indexers providing services in the Graph network can adopt Theory to add a prediction API product to their own suite of product offerings, achieving operational efficiency through a shared backend. The indexing infrastructure they are already operating for the Graph becomes an underutilized fixed asset that is repurposed to provide additional value and generate additional revenue in an emergent cloud computing ecosystem built around decentralized finance.

Theory is using an open approach, allowing each service provider on the network to use their own choice of implementation and by governing the protocol itself through a decentralized autonomous organization.

System Architecture

The Theory network ecosystem consists of Indexers, Consumers and Delegators.

Indexers

Indexers provide Theory API services to consumers. While not required, most Theory Indexers are likely to also be Graph Indexers, as the indexing infrastructure for being a service provider on the Graph can be reused for Theory. Indexers earn fees through queries and a rebate pool.

Consumers

Consumers pay Indexers for their API service. For each request made, consumers must pay an Indexer-specified fee.

Delegators

Theory token holders will be able to delegate to Indexers, and earn a commission amount specified by Indexers. A bonding curve mechanism may be used to assign each Delegator a stake in their delegated Indexer. This will reward delegators who are early to identify Indexers with highly effective Theorize implementations or other valuable attributes. High delegation amounts may help Consumers identify high quality Indexers.
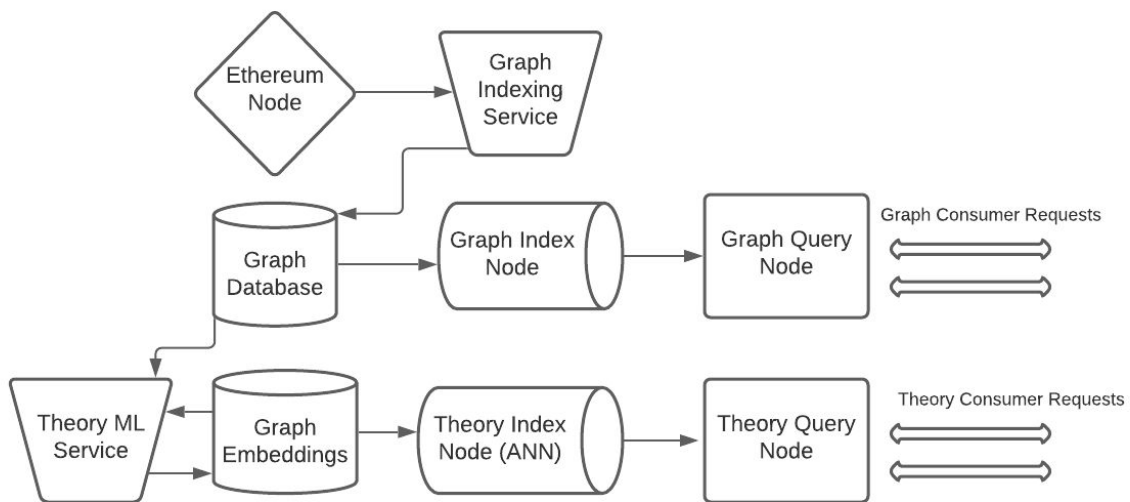
Token Recommendation API

The first API offered by Indexers will be the Token Recommendation API, which ranks the predicted relevancy of tokens to a DeFi user. Implementations will typically use a version of a similarity matching system and approximate nearest neighbor search.

The Token Recommendation API takes as input a list of addresses. For each of these specified addresses, the API returns a ranking of tokens by their predicted relevancy, with a predicted relevancy score provided for each of the ranked tokens. If no additional arguments are provided, the top rankings for each input address is returned. If an optional list of tokens to rank is provided, the score for each of the specified tokens is returned. Indexers charge a customizable fee for each request.

Theorize

Theorize is the system used by Indexers to generate predictions. Unlike the Graph where there is only one way to correctly index data, predictions are far more subjective in nature. Because of this, while there may be official Theorize implementations, Indexers will be able to use their choice of implementation. This may allow Indexers to compete with one another based on the quality of their predictions.

This diagram indicates how a Graph indexer can modify their system to run service nodes for both the Graph and Theory:

The following steps describe in more detail how the Token Recommendation API may be implemented to predict the relevance between any token and any user:

1. Periodically, subgraphs for decentralized exchanges and other apps are queried for recent transactions.

2. Transactions that appear to be from arbitrage bots and other sources of noise are filtered out.

3. Tokens from these transactions are embedded as vectorized nodes in a low-dimensional space in which tokens transacted with by the same user are placed closer to one another. The result is a token preference embedding.

4. Additional dimensions may be used for attributes of the transaction such as volume, attributes of the token (blue chip, microcap, etc.) and the category of the token protocol.

5. Embeddings for tokens and users are learned with ML techniques that may include matrix factorization, graph convolutional neural networks, and graph matrix completion.

6. An approximate nearest-neighbor (ANN) algorithm is used to build an index of the token embeddings to improve performance at runtime. This may utilize one of a number of highly performant open source ANN tools, including libraries from Spotify (Annoy), Google (ScaNN), and Microsoft (SPTAG).

Poolside: Personalized DeFi News

The first consumer of Theory will be Poolside - a product that will generate personalized DeFi updates powered by Theorize recommendations. To get started using Poolside, a new user simply needs to provide their preferred mode of contact (email, Telegram, Discord, etc.) and can sign in with their wallet at Poolside.finance so that their activity history can be used to generate recommended content.

New items will be derived from several sources, including Twitter, RSS/ATOM feeds from relevant news sources, and updates may be derived directly from queries of certain DeFi subgraphs.

Each news item is tagged with tokens or apps it is referencing. This topic classification will initially be implemented using simple techniques such as matching on relevant substrings such as the name or ticker symbol of a DeFi token found in the title or body of a news feed item, and may later incorporate more sophisticated topic classification techniques. News items will be filtered based upon the calculated Token Recommendation score of matched topics for each user.

In addition to Poolside demonstrating the capabilities of Theory, and bootstrapping usage of the network, it will also be possible to monetize Poolside and subsidize its operational

cost through promoted content added to its news digests, and potentially through subscription fees.

Roadmap

Prediction System Enhancements
Theorize recommendations could be improved upon to better be able to predict the relevancy of various types of non-fungible tokens, and better handle "cold start" users without a long activity history. It would also be possible to target trading use-cases where users are mainly concerned with predicting upcoming price action and determining their trade strategies.

Payment Escrow and Indexer Staking
A simple escrow system will be introduced that will confirm that an Indexer has returned the correct number of properly formatted scores before releasing payment for a request.

The evolution of this escrow would be to transition our delegation feature into a staking mechanism that may actively penalize Indexers for downtime, returning improperly formatted responses or the incorrect number of scores in a response,

There may be a way to incorporate into a staking system a checksum attestation or other means of verifying the integrity of Indexer responses created with official Theorize models.

Attribution Modeling
Within the context of advertising or recommendation systems, attribution for a user action, either directly after an impression

(direct conversion), or within a window of time (view-through conversion) is used to assign credit to the advertisement or recommendation. The modeling of attribution can be incorporated into the economics of an advertising system via cost per conversion metrics, and can also be used to provide training data to machine learning models that are generating the predictions of what recommendations or advertisements will result in conversion. An attribution modeling mechanism could help Theory Indexers and Consumers better understand how effective generated predictions and recommendations are.

Theory for Apps

Initial tooling for apps will include Theory.js, a javascript library that makes it easy to incorporate Theory recommendations into an app. It would also be possible to develop tools that utilize Theory-generated insights for product intelligence and developing campaigns for customer retention and acquisition, especially if Theory incorporates attribution modeling. These acquisition and retention campaigns may utilize messages and offers sent to Poolside subscribers who have opted into these types of updates.

Thanks

A thank-you section will be added here for people who have contributed feedback or have otherwise been helpful.