Nectar for Cyber Security Telemetry

PolySwarm

2021-05-17

Abstract

We introduce a new utility use for PolySwarm's Nectar token: distributing rewards for security relevant telemetry data encountered by everyday users. Telemetry Providers (i.e. users) will now be rewarded for the value they provide, and by contributing to a worldwide telemetry network gain access to the collective protection of the global PolySwarm network. Stakers and aggregators participate in the network to help determine the most useful sources of telemetry and help reduce the threat of spam on the network, for which they also earn a portion of NCT. Finally, Telemetry Consumers will now be able to access a truly worldwide network to find the security data they need to identify and fight emerging threats.

Introduction

In this whitepaper we introduce a new utility use for PolySwarm's Nectar token for average users: distributing rewards for security relevant data about TLS certificates, DNS resolutions, and potentially malicious files encountered in daily computer use. Many of these telemetry sources are already collected from user devices by Antivirus (AV) providers, but there are a number of serious issues with how they are collected, how users are compensated for their information, and how these results are shared. By re-imagining how this marketplace works, we can increase collection transparency, fairly compensate all participants in the marketplace, and, most importantly, create a more unified source of security telemetry that will better protect users worldwide.

Project Goals

Unified Telemetry Source

In our original whitepaper, we discussed the fragmentation of the AV market and how, in its current form, this fragmentation leads to worse outcomes for users in the marketplace. However, this fragmentation is not limited simply to the world of scanner providers: it affects many other parts of the security industry as well. As we built the PolySwarm marketplace, we realized that many of the disparate pieces of security information our customers were trying to connect are often ones that exist, but in practice are inaccessible due to the fragmentation of the market.

To attack this problem, we are building a unified marketplace for security telemetry. We envision a single platform in which users, security companies, and hackers worldwide can share and access this telemetry securely, safely, and fairly.

Fair Compensation

As the market exists today, users receive very little, if anything, for providing telemetry to their security providers. In fact, many people are unaware that this data is even being collected in the first place. Companies use this information to detect potential emerging threats, but their visibility is fundamentally limited by the number and location of subscribers they have, which requires them to either purchase this information from resellers or try to work around their blind spots.

By creating a market in which users actually get rewarded for the useful information they provide and the other users they protect, we can greatly improve both the availability of security telemetry worldwide and make sure that users are fairly compensated for the valuable contributions they are making. Security providers also benefit greatly from this arrangement, however, as there is now a way for them to expand their threat visibility without having to deal with middlemen.

Collection Transparency and Security

Ethical collection is an absolute requirement for any system dealing with the collection of user data. Our design of the marketplace pays heavy attention to the privacy and security of the data being sold to telemetry Aggregators, and ensures that the data being provided is properly anonymized and protected. By giving users more control over the data they provide, and compensating them for the data they feel safe sharing, we hope to create both a more open and more effective data source for the security community at large.

Participants

Telemetry Providers

Telemetry providers are, in short, normal users, and can join the network simply by installing a small browser plugin (v1) or monitoring daemon (v2). These users agree to provide Aggregators of their choice with telemetry of their choosing. In return, these individual providers are compensated by the usefulness of their contributions to the network. They also gain early warning of potentially malicious artifacts they've encountered.

Telemetry Consumers

Telemetry Consumers are the main buyer of data from providers and insights from Aggregators. They consistently leverage NCT to buy queries on or feeds of telemetry data from one or more Aggregators.

Cyber security focused consumers of the PolySwarm platform already show interest for additional telemetry based insights on the 750K+ daily malware samples seen on the platform. Traditionally, this data has been fragmented and distributed across anti-virus and threat intelligence companies that do not have a central marketplace for sharing. PolySwarm gives consumers better access to data that highlights the origins and behavior of malware: Observed distribution URLs/IPs and other infection vector spotting Coarse geolocation of targeted populations Passive DNS-like data that highlights new malicious domains and other "fast-flux" infrastructure

Aggregators

Aggregators are the center of the market: they accept telemetry from Providers, and in turn provide Consumers with both an aggregate view of telemetry data and valuable insights into new threats gleaned from this data. Employing PolySwarm's decentralized telemetry model gives Aggregators the ability to assemble a wider view of the threat landscape cheaply and directly, resulting in better intelligence for consumers, and more opportunities for them.

Aggregators benefit telemetry providers by providing a single, automated, market for their data that is a largely passive income stream. Telemetry Consumers benefit from the insights contained in the data by both receiving distilled threat feeds from Aggregators and being able to query a large amount of telemetry data in one place.

Marketplace Design

Our V1 design for this new marketplace focuses around three main design goals: ease of participation, balanced incentives, and minimal on-chain transactions.

Telemetry Collection

Participating as a telemetry provider will be as simple as downloading and installing a PolySwarm developed telemetry collection application. Our V1 design is focused on providing an easy-to-use telemetry collection setup with safe and sane privacy defaults for a user. Telemetry data is collected according to user privacy preferences and made available to Aggregators in privacy-sensitive data structures. These data structures are queried and maintained by Aggregators, in a decentralized way, such that the user retains control over their data and its income while still allowing Aggregators to highlight value in telemetry data like IP, URL, and file hash sightings.

Telemetry Aggregation

Aggregators serve three primary functions: assembling a wide view of telemetry data from providers, rewarding those providers, and adding value for consumers by analyzing telemetry. This results in a deliverable that is packaged two main ways for consumption: queries and feeds. Aggregators earn NCT when consumers either (a) subscribe to a feed or (b) receive a query response from an aggregator's data.

Aggregators register themselves on the PolySwarm platform and become a destination option for telemetry provider's data within the PolySwarm collection applications. In V1, the PolySwarm team will design, build, and operate the first aggregator. We cover how Aggregators should approach telemetry compensation in the provider compensation section and plan to release open reference implementations that help other aggregators get started.

Aggregator query servicing is simple to understand and assumes that a consumer has pre-purchased a block of queries from an Aggregator in NCT. Consumers issue queries around supported telemetry data (e.g. IP address range) across one or more Aggregators. Aggregators answer the query and register the query author's interest in specific data for later per-period compensation of the telemetry user's that provided the data.

Aggregator feed subscriptions are, likewise, quite straightforward. Per-period, aggregators collect a variable amount of telemetry data which is the raw ingredient into a security relevant feed. For example, at PolySwarm we generate a daily data feed which catalogs observed ransomware and metadata that helps our users detect and block infection attempts. As an aggregator, we would add further value to this feed by (a) ranking each daily threat by infection vector sightings in the wild and (b) attaching telemetry derived and reinforced DNS and IP address information to the sample family. Ranking threats allows a consumer, especially a cyber security response team, to prioritize any alerts generated by trending ransomware releases. Attaching further IP/DNS information from telemetry data allows consumers to quickly and automatically block malware command and control traffic with confidence that the control infrastructure was observed by multiple independent telemetry providers.

While feeds provide direct value to defenders/consumers, we discuss a novel method of staking in the provider compensation section that helps bring telemetry of value to the forefront.

Telemetry Consumption

Cyber security telemetry data is valuable to many industry consumers: managed security solutions providers (MSSPs), anti-virus and EDR product vendors, and Enterprise security operations centers. Within our marketplace design we try to capture the "value" that these various consumers apportion to telemetry data in order to reward both Aggregators and the crowd of telemetry providers. Value is transferred through NCT by consumers that subscribe to Aggregator feeds and execute pay-per-use queries across Aggregators.

Provider Compensation

Telemetry Providers

Telemetry Providers will be compensated based on the quality of their submissions, as judged by the Aggregators who are in turn informed by consumption patterns. Aggregators will publish their total period compensation pool. In V1 of the network, this compensation will be calculated opaquely by the Aggregator, but in v2 we expect to design and deploy governance mechanisms that let Aggregator participants determine the rewarding formula. This is inspired partially by Uniswap's constant product formula for liquidity, but with an added twist that, in PolySwarm, telemetry and staking participants can evolve the rewards formula as they add value to the Aggregator.

As Aggregators may implement different, independent, compensation strategies, we first detail our plans for the PolySwarm Aggregator as an illustrative example of the compensation potential for Telemetry data.

The V1 PolySwarm Aggregator will be funded daily with X NCT, which will be adjusted based on the NCT/USD exchange rate and the average number of active participants. As other Aggregators join, total NCT compensation should increase.

NCT rewards are distributed to telemetry providers at the end of each time period based on the following function:

$$NCT_t = f(R_p, t, S_p)$$

Where R_p represents the relevance of telemetry provided during that time period, t represents the tenure of a provider that has been providing data, and Sp represents the amount of NCT staked on that telemetry provider. Initially, we plan to weight 75% of the daily NCT funds for telemetry towards rewards for valuable contributions. The remaining 25% is distributed across contributors to encourage the growth of telemetry data that may not be immediately valuable.

The output of the reward function is simply the percent share of the total rewards in the time period. In the first version of the network, this function will exclusively be defined by the Aggregator themselves, but we intend to pursue more decentralized governance in later versions.

Initially the relevance score (R_p) in V1 (referenced above) is defined by two properties: (1) PolySwarm consumer/customer interest and (2) severity of the threat. A user providing valuable telemetry can expect to receive NCT rewards from that pool according to the following formula:

$$R_p = \sum_{i=0}^{n} \frac{1}{N_i} \times W_i$$

Where N_i is the number of other users with a matching sighting and Wi is the weight assigned to the sighting by the Aggregator. Note that PolySwarm's Aggregator plans to calculate per sighting weight in a manner that approximates a sightings value to the customer and the community over multiple periods. Initially, this will be a combination of our PolyScore and relevance as defined by consumer queries and feed inclusion.

Multiple period evaluation is especially important in the case of feeds enhancing telemetry remains relevant over a longer period or the threat increases in severity. DarkSide, the ransomware that caused the shutdown of the Colonial pipeline, is a good example of this phenomenon. PolySwarm had detected variants as early as August of 2020 but early canary telemetry providers around this malware would have seen a heady increase in their relevance and severity weightings in May 2021.

This is the main reason that tenure as a telemetry provider factors into the NCT rewards equation: observations over the long run are valuable when acquiring true telemetric context for baselining a threats history and current, relative, severity.

Aggregator Compensation

Aggregators will be compensated by fees taken off the top from query and feed subscription values. The calculation of the fees will be transparent to the Aggregator's consumers as a proportion of each period's pool rewards. These fees incentivize the Aggregators to attract consumers of both feeds and queries with relevant, well packaged, telemetry data that have traditionally been focused on a specific cyber security vertical (e.g. banking related threats targeting the financial sector).

Staking Compensation

To help prevent spam and to better allocate rewards, it will be possible to stake NCT on individual Telemetry Providers, including your own. The NCT stake provides a crucial signal to the Aggregator as to the trustworthiness of an individual provider and informs the reward function (S_p). For this information, staking users are rewarded with a relative percentage of 20% of the Telemetry Provider's profits on every reward period payout. The risk assumed by the staker, however, is if a particular provider is found to be malicious or providing invalid telemetry: in this case, an Aggregator can propose the burning of a stake to the network, and funds may be lost along with the trust in the telemetry provider.

The staking rewards can be defined as follows:

$$NCT_s = S_t \times NCT_t \times S_s$$

Where S_t is the percentage of the total rewards that the telemetry provider wishes to give to stakers, NCT_t is the total NCT awarded to the telemetry provider in this period, and S_s is the percent of the total stake the current staker has.

Timeline & Comments

On July 19, 2021 we plan on releasing a roadmap detailing our development schedule for the high-level ideas discussed in this whitepaper. We welcome discussion and feedback on the concepts presented at info@polyswarm.io or in the PolySwarm Telegram channel.

Summary

This new marketplace provides immediate benefits to all participants. Users get more control over their data and are actually compensated for the value they provide, as well as receive early warnings about threats they have encountered. Aggregators are no longer dependent on their own install base for data, and earn NCT for providing query computation and telemetry validation. Stakers help the network determine the most useful sources of telemetry, and help reduce the threat of spam on the network, for which they also earn a portion of NCT. Finally, Consumers will now be able to access a truly worldwide network to find the data they need to identify and fight emerging threats.

At PolySwarm, it is our mission to bring the security community and users worldwide together to fight malware. By leveraging Ethereum's global, decentralized network, our new marketplace will greatly further our efforts to bring these groups together, by enabling (and incentivizing!) everyone to help solve this difficult problem.

Disclaimer

This paper is for general information purposes only. It does not constitute investment advice or a recommendation or solicitation to buy or sell any investment and should not be used in the evaluation of the merits of making any investment decision. It should not be relied upon for accounting, legal or tax advice or investment recommendations. This paper reflects current opinions of the authors and is not made on behalf of PolySwarm or its affiliates and does not necessarily reflect the opinions of

PolySwarm, its affiliates or individuals associated with PolySwarm. The opinions reflected herein are subject to change without being updated.