
ANÁLISIS COMPARATIVO DE COMUNIDADES EN ETHEREUM: DE POW A POS

TRABAJO PROFESIONAL

Ignacio Argel
FIUBA
iargel@fi.uba.ar

Manuel V. Battan
FIUBA
mvbattan@fi.uba.ar

Lucas Bilo
FIUBA
lbilo@fi.uba.ar

Nahuel H. Spiguelman
FIUBA
nspiguelman@fi.uba.ar

ABSTRACT

El cambio de Proof of Work a Proof of Stake en la blockchain descentralizada de Ethereum tiene un impacto directo en la función desempeñada por los mineros, ya que son reemplazados por Stakers como nuevos validadores de la red. La disparidad en las dinámicas de este nuevo protocolo plantea interrogantes sobre si ha ocasionado modificaciones en el comportamiento de los usuarios que operan en él, ya sean stakers o participantes frecuentes de la red.

Las investigaciones previas en Ethereum han priorizado aspectos técnicos, dejando un vacío en el análisis de interacciones entre cuentas y la formación de comunidades. Esto es especialmente notorio en la ausencia de estudios que aborden cómo el comportamiento de usuarios y comunidades ha cambiado con la migración de PoW a PoS.

Para este análisis utilizaremos un modelo cuya finalidad es crear embeddings de nodos que representen, de una manera inteligente, las relaciones entre addresses en la red. A su vez, este modelo es superior en términos de limitaciones computacionales y de utilización de features, que algoritmos tradicionales para analizar grafos no tienen en cuenta por defecto.

Keywords Ethereum · Proof of Stake · Proof of Work · Networks · Community Analysis · Pattern Detection

1 Introducción

La blockchain Ethereum ha sido una de las innovaciones más significativas en el mundo de la tecnología financiera en la última década. Desde su lanzamiento en 2015, Ethereum ha operado bajo un protocolo de consenso conocido como Proof of Work (PoW), el cual ha demostrado ser robusto y confiable en la validación de transacciones y la seguridad de la red, pero presentando problemas como el alto consumo de energía y la centralización de la minería de bloques. Debido a estos problemas, Ethereum ha migrado hacia el protocolo de consenso conocido como Proof of Stake (PoS), el cual incluye varios cambios en la manera de garantizar la integridad durante las transacciones, mejorando la escalabilidad y la eficiencia de la red.

Algunas preguntas significativas que surgen a partir del cambio de paradigma: ¿Existe un cambio en el comportamiento de los usuarios debido a esta transición? ¿Cuáles son los comportamientos invariantes o que siguen siendo vigentes después del cambio? ¿Cómo variaron las relaciones entre usuarios y comunidades?

El presente proyecto tiene como objetivo principal contestar estas preguntas. A través de un análisis exhaustivo, buscamos proporcionar una visión clara de los patrones de comportamiento de quienes operan sobre la red de Ethereum en sus dos protocolos de consenso principales, PoW y PoS. Al hacerlo, esperamos contribuir a la comprensión de las implicaciones de esta transición para la comunidad Blockchain, los desarrolladores y los usuarios de Ethereum.

2 Representation learning sobre addresses de la red de Ethereum

Las limitaciones con las que nos enfrentamos al utilizar distintos métodos de búsqueda de comunidades nos llevó a explorar otras opciones. Estas limitaciones estuvieron presentes tanto en el plano analítico, como en el computacional, ya que varias de las opciones exploradas, como Gephi, Node2Vec en conjunto con clustering o el algoritmo de Louvain no presentaban buena performance sobre una red tan grande y compleja. Por eso decidimos incursionar en el desarrollo de una red neuronal siamesa, cuyo objetivo sea encontrar la mejor representación de una address teniendo en cuenta su relación con los demás.

2.1 Triplet Loss personalizada

Para la generación de estos embeddings, empleamos una triplet loss, diseñada para calcular la distancia entre dos pares de embeddings de addresses: $d(Anchor, Positive)$ y $d(Anchor, Negative)$. El anchor representa una address que ha realizado una o más transacciones cuyo destinatario es el positivo. Mientras que el negative representa una dirección que nunca ha sido ni emisor ni receptor en una transacción que involucre al anchor.

De esta manera, a través del cálculo que involucra ambas distancias, la red neuronal se entrena para acercar el anchor al positive y alejarlo del negative, optimizando así la representación de las addresses en un espacio N -dimensional.

$$\mathcal{L}(A, P, N) = \max((\omega \cdot \| (A_e) - (P_e) \|_2) - \| (A_e) - (N_e) \|_2 + \alpha, 0) \quad (1)$$

Como resultado, se espera que estos embeddings de N -dimensiones conserven una cantidad significativa de información contextual sobre la address y sus transacciones asociadas. Esto facilita la aplicación de algoritmos de clustering, con el objetivo de identificar conjuntos de direcciones altamente relacionadas entre sí.

2.1.1 Weight: factor ponderante

Dentro de la loss, ω es usado para relacionar el anchor y el positive con mayor intensidad en casos en los que las transacciones son muy recurrentes o de altos montos. Este se calcula de la siguiente manera:

$$\omega = w_{\#T_{rx}} \frac{|T_{AP}| - \mu_{|T|}}{\sigma_{|T|}} + w_{value} \frac{\bar{V}(T_{AP}) - \mu_{V(T)}}{\sigma_{V(T)}} \quad (2)$$

Al ser una función creciente en cuanto a la cantidad de transacciones y al valor promedio por par, esto incrementa el valor resultante de la loss. Como consecuencia, la loss asigna un valor más alto a los pares donde el anchor y el positivo tienen un volumen, ya sea en cuanto cardinalidad o valor, mayor. Esto resulta en una mayor ganancia al acercarlos abruptamente, garantizando que estos pares altamente relacionados permanezcan espacialmente cercanos.

3 Examples of citations, figures, tables, references

3.1 Citations

Citations use natbib. The documentation may be found at

<http://mirrors.ctan.org/macros/latex/contrib/natbib/natnotes.pdf>

Here is an example usage of the two main commands (`citet` and `citep`): Some people thought a thing [Kour and Saabne, 2014a, Keshet et al., 2016] but other people thought something else [Kour and Saabne, 2014b]. Many people have speculated that if we knew exactly why Kour and Saabne [2014b] thought this...

3.2 Figures

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetur odio sem sed wisi. See Figure 1. Here is how you add footnotes. ¹ Sed feugiat. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Ut pellentesque augue sed urna. Vestibulum diam eros, fringilla et, consectetur eu, nonummy id, sapien. Nullam at lectus. In sagittis ultrices mauris. Curabitur malesuada erat sit amet massa. Fusce blandit. Aliquam erat volutpat. Aliquam euismod. Aenean vel lectus. Nunc imperdiet justo nec dolor.

¹Sample of the first footnote.

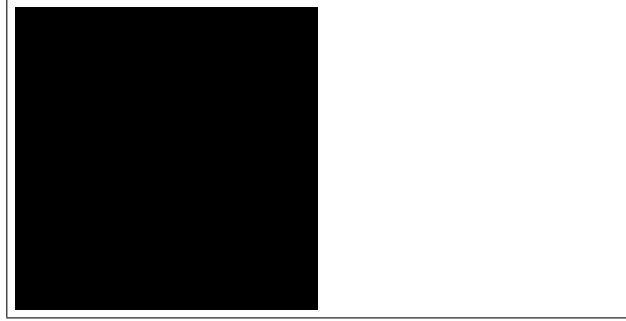


Figure 1: Sample figure caption.

Table 1: Sample table title

Part		
Name	Description	Size (μm)
Dendrite	Input terminal	~ 100
Axon	Output terminal	~ 10
Soma	Cell body	up to 10^6

3.3 Tables

See awesome Table 1.

The documentation for booktabs (‘Publication quality tables in LaTeX’) is available from:

<https://www.ctan.org/pkg/booktabs>

3.4 Lists

- Lorem ipsum dolor sit amet
- consectetur adipiscing elit.
- Aliquam dignissim blandit est, in dictum tortor gravida eget. In ac rutrum magna.

References

- George Kour and Raid Saabne. Real-time segmentation of on-line handwritten arabic script. In *Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference on*, pages 417–422. IEEE, 2014a.
- Renato Keshet, Alina Maor, and George Kour. Prediction-based, prioritized market-share insight extraction. In *Advanced Data Mining and Applications: 12th International Conference, ADMA 2016, Gold Coast, QLD, Australia, December 12-15, 2016, Proceedings 12*, pages 81–94. Springer, 2016.
- George Kour and Raid Saabne. Fast classification of handwritten on-line arabic characters. In *Soft Computing and Pattern Recognition (SoCPaR), 2014 6th International Conference of*, pages 312–318. IEEE, 2014b. doi:10.1109/SOCPAR.2014.7008025.