الگوي تهيه مقالات براي كنفرانس ملی سالانه محاسبات توزیعی و پردازش داده های بزرگ

صادق دوری نوگورانی1، رضا حمیدپور بدوئی2

1 استادیار، دانشکده مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس،

Dorri@modares.ac.ir

2 دانشجوی ارشد، مهندسی کامپیوتر نرم افزار، دانشگاه تربیت مدرس،

Reza.hamidpour@modares.ac.ir

چكيده

با شروع فعالیت شبکه بیت­کوین و آغاز به کار رمزارزهای دیجیتال، انقلابی جدیدی در حوزه تبادلات مالی و ساختار ارزهای سنتی ایجاد شد. بیت­کوین با فراهم سازی ویژگی­هایی همچون غیرقابل تغییر بودن تاریخچه آن، عدم نیاز به شخص ثالث درجهت تایید تراکنش­ها و ناشناس بودن باعث افزایش کاربران علاقه­مندی کاربران به این تکنولوژی نوین شده است. امروزه نیز با فزونی یافتن کاربرد تکنولوژی زنجیره بلوکی در زمینه­های مختلفی همچون اقتصاد، زنجیره تامیین و بسیاری از زمینه­های گوناگون دیگر، حوزه­های نوین تحقیقاتی را نیز با خود به وجود آورده است. یکی از مسائل بسیار پراهمیت که محقیقن حوزه­های گوناگون همچون کامپیوتر و اقتصاد امروزه بیش از پیش به آن نیازمندن داده­های نقل و انتقالات این رمزارزها در کنار ارزهای سنتی می­باشد. به گونه­ای که بتوانند میزان رشد رمزارزها ویا تاثیر گذاری قوانین و مسائل گوناگون شبکه­های زنجیره بلوکی بر روی این ارزهای نوین و همچنین ارزهای سنتی را در کنار یکدیگر مورد ارزیابی و برسی قرار دهند. در این مقاله ما قصد داریم مجموعه داده جدیدی که توسط یک بازار تبادل مالی به نام استلار ایجاد شده است و بر بستر تکنولوژی زنجیره بلوکی بنا شده است و همچنین مکانی برای تبادل رمزارزهای و همچنین ارزهای سنتی می باشد را ارائه کنیم.

كلمات كليدي

رمزارزهای دیجیتال، مجموعه داده حجیم، مجموعه داده صرافی توزیع شده، مجموعه داده رمزارزهای دیجیتال.

# مقدمه

با پیشرفت روز افزون رمزارزها1 و افزایش کارایی آن­ها در تبادلات مالی، مباحثی همچون تحلیل ساختاری شبکه­های زنجیره بلوکی2، ایجاد ثبات در تغییر نرخ تبدیل3  و همچنین شناخت رفتار کاربران4 این شبکه­ها بیش از پیش دارای اهمیت شده است. از طرفی نسل جدید تبادلات مالی که بر پایه تبادلات دیجیتالی5 شکل میگیرند نیز با پیدایش رمزارزها انقلابی را در نقل وانتقالات مالی ایجاد کرده­اند که باعث به وجود آمدن موج جدیدی از تغییرات در بازارهای اقتصادی و چشم­انداز­های اقتصادی شده است. با افزایش نیازمندی­های امروزه دانشمندان و تحلیل گران اقتصادی به تحلیل دقیق و بهتر ساختار بازار رمزارزها، نیاز به جمع آوری داده این بازارها نیز چالشی بسیار بزرگ است که دانشمندان حوزه تحلیل داده6 با آن رو به رو هستند. یکی از این محدودیت ها بدست­آوردن مجموعه داده­7 دقیق از نقل و انتقالات رمزارزها می باشد.

برای توصیف میزان اهمیت داده­های اقتصادی می­توان به این موضوعات نیز اشاره کرد که با استفاده از این داده­ها تحلیل­گران اقتصادی، سبدگردانان بازارهای سرمایه و همچنین دولت مردان می­توانند روند تغییرات یک بازار اقتصادی، میزان تاثیر گذاری قوانین بر روی آن بازار و همچنین روند رشد اقتصادی یک بازار و در امتداد آن یک جامعه را مورد برسی و تحلیل قرار دهند.

که از جمله مهم­ترین این اطلاعات می­توان به اطلاعات تبادلات مالی در یک بازار سرمایه2 اشاره کرد.

## 1-1 ادبیات موضوعی

در [1] مجموعه داده­هایی که توسط رمزارزهای پرطرفدار بیت­کوین8 و اتر9 ایجاد شده­اند را تحت عنوان مجموعه داده­های بزرگ10 معرفی میکند که در حوزه­هایی هم­چون تشخیص تقلب11، تراکنش­های غیر قانونی12، شناسایی نقاط بهرانی اقتصاد13 و حوزه­های پراهمیت دیگری که کاربرد دارند، اشاره می­کند. همچنین در این کتاب به توصیف اساسی در خصوص ویژگی­های مختلفی که می­توان از مجموعه­ داده­های یک رمزارز بدست آورد ویا ساختار کلی این نوع از داده­ها پرداخته شده است.

در مقاله [2] نیز به این موضوع اشاره شده است که زنجیره­های بلوکی توانایی بسیار زیادی برای استفاده در حوزه مالی، بازی، نقل و انتقالات مالی و زنجیره کالا دارد، اما یکی از چالش­های جدی در این حوزه تحلیل رفتاری کاربران این شبکه­ها و دسته­بندی آن­ها بر اساس رفتارشان می­باشد. در اینجا نویسندگان با دنباله­های تراکنشی کاربران شبکه بیت­کوین که از تاریخچه عمومی این شبکه دریافت شده ­است و با استفاده روش پیشنهادی خود که یک روش تحلیل رفتار شناسی می­باشد این داده­ها را مورد ارزیابی قرار داده­اند.

در مقاله [3] نیز با جمع آوری مجموعه داده­ای از بازار آتی چین17 و با استفاده از روش­های خوشه­بندی اقدام به تحلیل رفتار کاربران این بازار کرده­اند.

در این مجموعه داده با توجه به ذات این بازار در هر طرف مبادلات مالی تنها یک دارایی قرار میگیرد که می­توان این مورد را یکی از مواردی دانست که در این مقاله به آن توجهی نشده است. چرا که در نسل جدید بازارهای اقتصادی که به وسیله ساختار شبکه­های زنجیره­بلوکی شکل گرفته­اند در تبادلات مالی بین کاربران مختلف دارایی­های متفاوتی تبادل می­شود.

به عنوان مثال می­توان گفت در مجموعه داده جمع­آوری شده از شبکه استلار در یک تراکنش ممکن است بین دو کاربر دو رمزارز اتر و بیت­کوین تبادل شود، اما این درحالی است که در همان لحظه بین دو کاربر دیگر امکان تبادل دو رمزارز دیگر مانند ریپل18 و تتر19 وجود دارد.

در [4] نیز با جمع­آوری داده­های پراهمیت رمزارزهای مهمی همچون بیت­کوین، اتر، لایت­کوین20 و ریپل که بیشترین میزان سرمایه را در بازار رمزارزها به خود اختصاص داده­اند، را مورد تحلیل و برسی قرار داده­ اند. در واقع در این مقاله داده­های مربوط به تبادلات رمزارزهای مختلف را از بازار (binance.com) جمع­آوری کرده­اند و سعی در شناخت تاثیرات قوانین اعمال شده بر روی قیمت رمزارزها داشته­اند که مهم­ترین این قوانین، قوانین اعمال شده در اواخر سال 2017 و اوایل سال 2018 می­باشد. در این تحقیق با شناخت الگوی تراکنش­های این رمزارزهای مهم، سعی در شناخت هرچه بهتر این بازارها دارند.

مجموعه داده ای که در این مقاله ما جمع آوری کرده­ایم، شامل تراکنش­ها و مبادلات مالی بین کاربران شبکه استلار14 می­باشد. این شبکه با بهره­گیری از فناوری بلاکچین نسل جدیدی از صرافی­های توزیع شده را ایجاد می­کند به گونه­ای که کاربران این شبکه بدون نیاز به یک شخص ثالث اقدام به تبادل مالی با یکدیگر می­کنند.

**عملیات­ها در شبکه استلار:** عملیات­ها در شبکه­ استلار در قالب تراکنش­هایی جمع می­شوند به گونه­ای که هر تراکنش شامل چندین عملیات می­تواند باشد (عملیات­هایی همچون تغییر یک حساب کاربری، ثبت پیشنهاد خرید/فروش یک دارای، ویرایش/حذف یک پیشنهاد و انتقال مالی بین کاربران).

**دفترحساب­های کل:** دفترحساب­های کل16 نقش ذخیره­سازی تراکنش­ها در شبکه استلار را دارند، به گونه­ای که در هر دفترحساب­کل چندین تراکنش می­تواند قرار بگیرد، به طوری که هر تراکنش نیز شامل چندین عملیات می­باشد.

# مجموعه داده

همانگونه که پیش از این نیز اشاره شد با روی کار آمدن رمزارزهای دیجیتال و پر رنگ­تر شدن نقش آن­ها در دنیای اقتصاد، گرایش تحلیل­گران اقتصادی به این نوع از داده­ها نیز فزونی یافت.

شبکه­های بلاکچین با ایجاد تغییر و تحولات گسترده در سیستم انتقالات مالی، بازارهایی را ایجاد کردند که در آن تبادل­کنندگان مالی نیازمند یک شخص ثالث که وظیفه تایید و برسی انتقالات دارایی ها را برعهده داشته باشد ندارند.

پیش از این نیز به این موضوع اشاره شد که در این مقاله مجموعه داده بدست آمده، از مجموعه تبادلات مالی در شبکه استلار می­باشد.

با توجه به ساختار تراکنش­های شبکه استلار داده­های پیشنهاد خرید و فروش دارایی به 2 حالت کلی تقسیم می­شوند:

* پیشنهادات خرید یک دارایی
* پیشنهادات فروش یک دارایی

بر اساس توصیف ارائه شده توسط [6] می­توان هر پیشنهاد خرید ویا فروش توسط یک عامل در بازار را به طور کلی به یک پیشنهادر خرید تبدیل ویا بلعکس. به این صورت که اگر شخصی قصد خرید بیت­کوین را دارد و در ازای آن اتر ارائه کند، می­توان اینگونه تراکنش آن را دید که شخص مورد نظر یک درخواست فروش اتر را در شبکه ثبت کرده است.

به همین شکل اگر همان شخص مجددا قصد فروش بیت­کوین را داشته باشد درحالی که به ازای آن قصد خرید اتر را داشته باشد می­توان این بار آن را به عنوان فروشنده اتر درنظر گرفت. بر اساس این توصیف در حالت کلی تمامی درخواست­هایی که در شبکه استلار ایجاد می­شوند، در قالب یک درخواست فروش دیده­ می­شوند.

## توصیف فیلدهای داده:

در این بخش به توصیف فیلدهای داده ای هر پیشنهاد در شبکه استلار می­پردازیم.

* Source\_account: این ویژگی بیانگر ایدی شخصی است که پیشنهاد را درون شبکه ایجاد کرده است.
* Offer\_id: درصورتی که این ویژگی برابر با مقدار صفر باشد به این معنی است که این پیشنهاد به تازگی توسط کاربر ایجاد شده است و درون لیت فروش22 قرار گرفته است.
* Created\_at: بیانگر زمان ایجاد این پیشنهاد می­باشد.
* Ledger\_id: بیانگر ایدی دفتر حساب کلی می­باشد که این پیشنهاد در آن قرار گرفته است.
* Transaction\_hash: همانگونه که اشاره شد هر پیشنهاد در شبکه استلار در یک تراکنش قرار میگرد که هر تراکنش مجموعی از چندین عملیات است، در این فیلد مقدار هش23  تراکنشی که این عملیات متعلق به آن هست قرار میگیرد.

|  |  |  |  |
| --- | --- | --- | --- |
| تعداد تراکنش | نوع دارایی | تعداد تراکنش | نوع دارایی |
| 26747 | MOBI | 365124 | native |
| 21498 | EURT | 168119 | ETH |
| 18569 | CNY | 156867 | XRP |
| 16198 | MERA | 143540 | BTC |
| 15901 | FRAS | 128948 | LTC |
| 14464 | XCN | 112107 | SLT |
| 13089 | ZRX | 111378 | RMT |
| 10034 | USDM | 110495 | NRV |
| 7659 | XFF | 40854 | TERN |
| 7461 | USD | 31470 | GTN |
| 29500 | KIN | 29500 | KIN |

* Selling\_asset\_type: بیانگر نوع دارایی است که پیشنهاد دهنده قصد فروش آن را دارد. درصورتی که مقدار آن برابر با دارای بومی24 در داخل شبکه استلار باشد دو فیلد بعدی مقداری ندارند.
* Selling\_asset\_code: نماد دارایی که به فروش می­رسد.
* Selling\_asset\_issuer: ایدی شخصی است که این دارایی را در شبکه ایجاد کرده است. با توجه به ساختار شبکه استلار هر دارایی را یک کاربر می­تواند ایجاد کند و به وسیله واسط­هایی25 که در شبکه وجود دارند.
* Amount: بیانگر مقداری است که شخص پیشنهاد دهنده از دارایی مورد نظر برای فروش ارائه می­کند.
* Price\_r {n,d}: این ویژگی شامل دو مقدار n و d می­باشد و میزان ارزش دارایی است که به فروش می­رسد در مقابل دارایی که خریداری می­شود. به عنوان مثال اگر شخصی بخواهد مقدار 1 بیت­کوین را در ازای 100 اتر به فروش برساند مقدار این ویژگی برابر با {1،100} می­باشد.
* Price: این ویژگی از ضرب دو ویژگی n در فیلد price\_r در ویژگی amount به وجود می­آید.
* Buying\_asset\_type: بیانگر نوع دارایی است که شخص پیشنهاد دهنده قصد دریافت آن را دارد در مقابل دارایی که به فروش می­رساند. درصورتی که مقدار آن برابر با دارای بومی24 در داخل شبکه استلار باشد دو فیلد بعدی مقداری ندارند.
* Selling\_asset\_code: نماد دارایی که به فروش می­رسد.
* Buying\_asset\_code: کد دارایی که پیشنهاد دهنده می­خواهد دریافت کند.
* Buying\_asset\_issuer: ایدی شخصی است که در شبکه این دارایی را ایجاد کرده است و پیشنهاد دهنده از آن استفاده می­کند.

داده­های مورد استفاده در این مقاله به وسیله ای پی آی26 های ارائه شد توسط سرورهای هوریزون27 شبکه استلار جمع­آوری شده­اند.

بازه زمانی داده­های جمع­آوری شده برابر است با 01/12/2018 تا 30/12/2019 که در این بازه­زمانی 1،650،718 عملیات خرید/فروش ایجاد شده است که از این جمع عملیات 1،165،274 عملیات درخواست خرید جدید بوده است و 891،449 عملیات نیز عملیات ویرایش ویا حذف درخواست­های قبلی می­باشد. تمامی این عملیات­ها در 6،680،451 دفترحساب­کل جمع­آوری شده است. درکل این مجموعه داده 16،338 کاربر عملیات تبادل دارایی انجام داده­اند به طوری که 466 دارایی در حالت فروش قرار گرفته است و تعداد 444 دارایی نیز خریداری شده است.

جدول-1 تعداد تراکنش­ها بر اساس دارایی­هایی که در قسمت فروش قرار دارند.

جدول-1 و جدول-2 نشان دهنده این موضوع هستند که در شبکه استلار بیشترین حجم تبادلات مالی به ترتیب مربوط به دارایی­های بومی شبکه استلار، اتر و بیت­کوین می­باشند.

در مقالات مشابه که بر روی داده­های اقتصادی بازارهای مشابه فعالیت انجام داده­اند ساختار بازارها به گونه­ای بوده است که در آن به ازای دارایی­های مختلفی که هر کاربر خریداری می­کند تنها یک دارایی واحد را به ازای آن­ها پرداخت می­کند، به عبارتی دیگر در هر تبادل مالی یک سوی آن تبادل یک دارایی مشخص و واحد می­باشد، این درحالی است که در بازار تبادلاتی استلار به ازای دارایی­های مختلفی که کاربران خریداری می­کنند، دارایی­های متفاوت دیگری را ارائه می­کنند. به عنوان مثال شخصی که قصد خرید بیت­کوین را دارد می­تواند در مقابل آن دولار آمریکا ارائه کند، درحالی که شخصی دیگر در همان لحظه می­تواند پیشنهاد مشابه­ای را ارائه کند با این تفاوت که دارایی ارائه شده توسط این شخص به جای دولار آمریکا رمزارز ریپل باشد.

جدول-2 تعداد تراکنش­ها بر اساس دارایی­هایی که در قسمت خرید قرار دارند

|  |  |  |  |
| --- | --- | --- | --- |
| تعداد تراکنش | نوع دارایی | تعداد تراکنش | نوع دارایی |
| 11943 | SALT | 1172217 | native |
| 10368 | OMG | 45292 | NRV |
| 7649 | USD | 41608 | ETH |
| 7620 | REP | 33679 | XRP |
| 6813 | XFF | 25590 | BTC |
| 6254 | LTC | 25469 | TERN |
| 5986 | CNY | 22436 | GTN |
| 5398 | MTL | 21722 | SLT |
| 5238 | LINK | 19021 | MOBI |
| 4761 | BCH | 14711 | USDM |
| 4693 | CALL | 13007 | XCN |

# 3- تحلیل داده

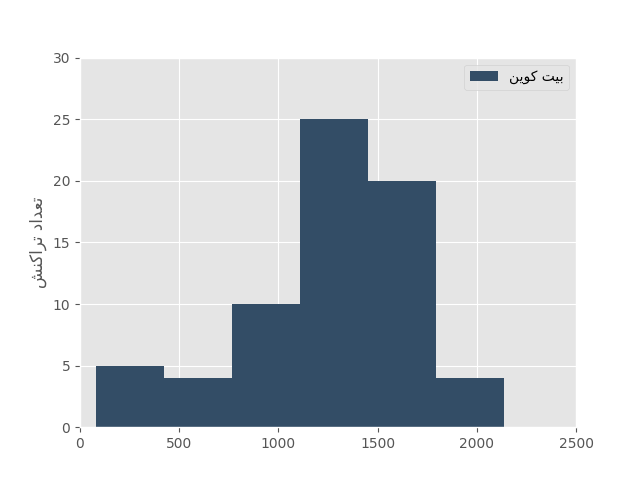
در این بخش با تمرکز بر روی دو رمزارز بیت­کوین و اتر قصد داریم توصیف بهتری نسبت به تبادلات مالی این مجموعه داده ارائه بدهیم. در همین جهت از میان عملیات­های انجام شده در این مجموعه داده­ تبادلاتی که، حداقل در یک طرف آن­ها یکی از رمزارزهای بیت­کوین ویا اتر قرار دارد را در قالب دو گروه قرار می­دهیم.

جدول-3 تعداد تراکنش در هر گروه دسته بندی

|  |  |
| --- | --- |
| تعداد پیشنهادات | نوع دارایی |
| 169،130 | بیت­کوین |
| 209،727 | اتر |

در جدول-3 تعداد تراکنش­ها به تفکیک هر یک از دارایی­های بیت­کوین و اتر قرار دارد که بیانگر تمامی عملیات­هایی می­باشد که حداقل یک طرف آن­ها بیت­کوین ویا اتر قرار دارد.

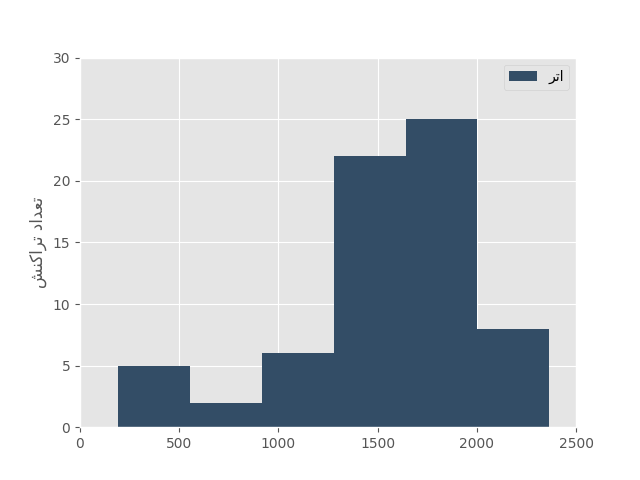
شکل-1 هیستوگرام تعداد تراکنش کاربران بیت­کوین



تعداد کاربران

بازه­های تعداد تراکنش کاربران

شکل-2 هیستوگرام تعداد تراکنش کاربران اتر

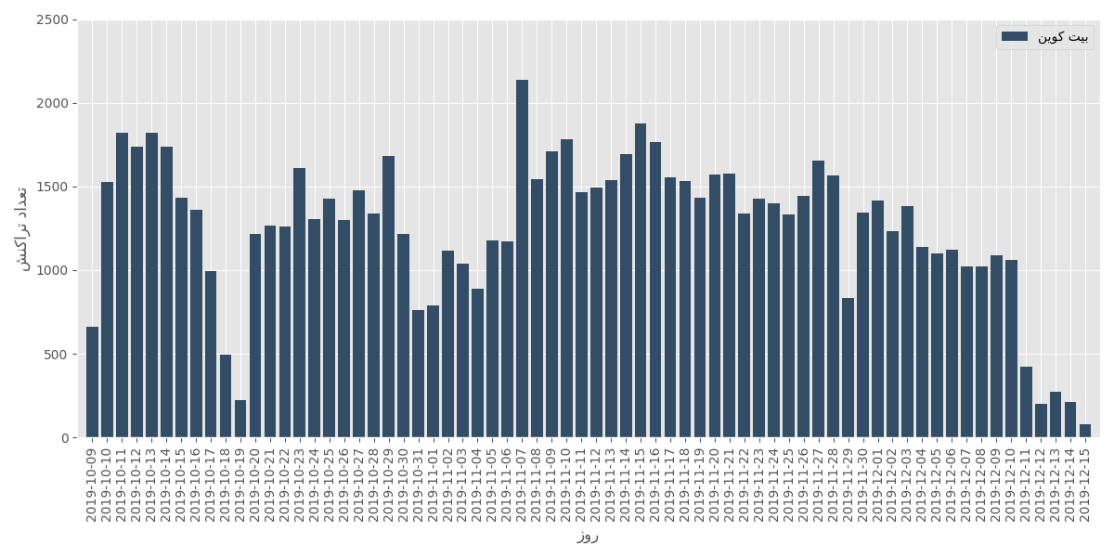


تعداد کاربران

بازه­های تعداد تراکنش کاربران

تعداد کاربران

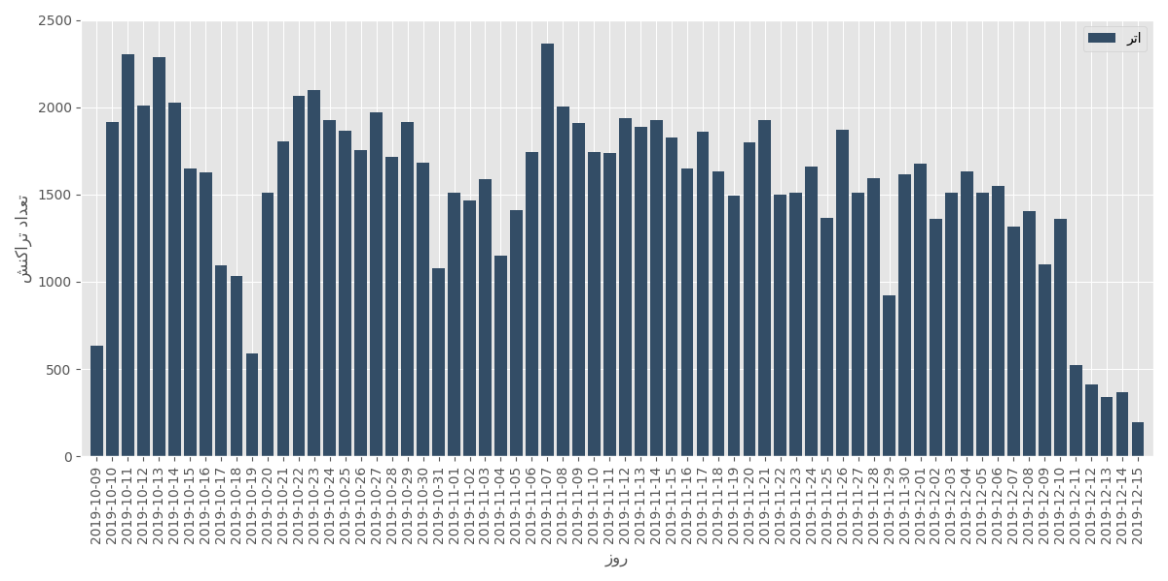
شکل-3 حجم تراکنش کاربران بیت­کوین به تفکیک هر روز



تعداد تراکنش

روز

شکل-4 حجم تراکنش کاربران اتر به تفکیک هر روز



حجم تراکنش

روز

# 4- نتیجه

مراجع

S. S. Chawathe, “Clustering Blockchain Data,” in Clustering Methods for Big Data Analytics: Techniques, Toolboxes and Applications, O. Nasraoui and C.-E. Ben N’Cir, Eds. Cham: Springer International Publishing, 2019, pp. 43–72.

B. Huang, Z. Liu, J. Chen, A. Liu, Q. Liu, and Q. He, “Behavior pattern clustering in blockchain networks,” Multimed Tools Appl, vol. 76, no. 19, pp. 20099–20110, Oct. 2017, doi: 10.1007/s11042-017-4396-4.

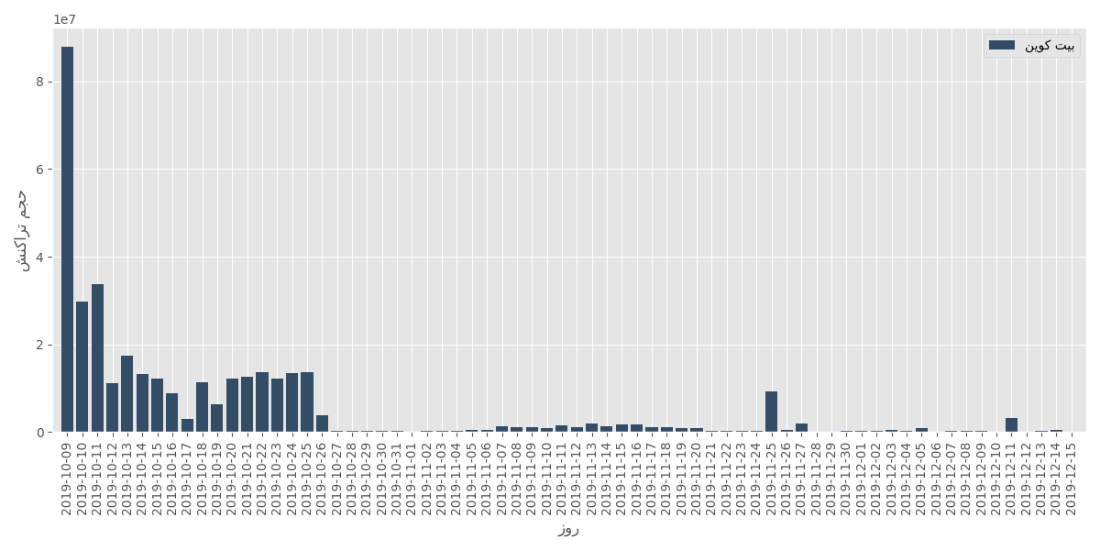
G. Shi, L. Ren, Z. Miao, J. Gao, Y. Che, and J. Lu, “Discovering the Trading Pattern of Financial Market Participants: Comparison of Two Co-Clustering Methods,” IEEE Access, vol. 6, pp. 14431–14438, 2018, doi: 10.1109/ACCESS.2018.2801263.

M. Gidea, D. Goldsmith, Y. Katz, P. Roldan, and Y. Shmalo, “Topological recognition of critical transitions in time series of cryptocurrencies,” Physica A: Statistical Mechanics and its Applications, p. 123843, Jan. 2020, doi: 10.1016/j.physa.2019.123843.

زير‌نويس‌ها

ضمایم

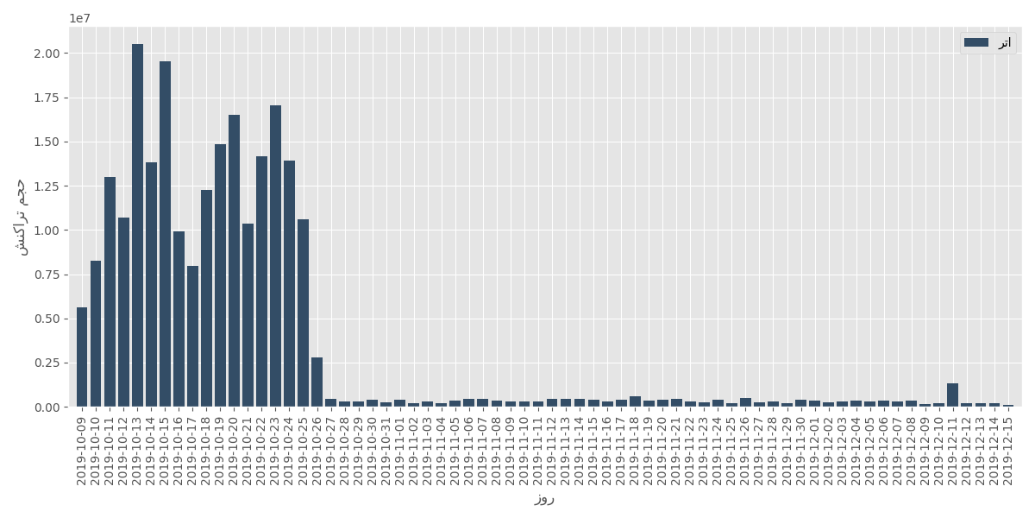
شکل-5 حجم تراکنش کاربران بیت­کوین به تفکیک هر روز



حجم تراکنش

روز

شکل-6 حجم تراکنش کاربران اتر به تفکیک هر روز



حجم تراکنش

روز