

获取跨链项目交易数据

1 整体思路

首先熟悉跨链项目的合约细节，根据相关的主要功能函数，爬取相关交易事件，结构化存储，确定并且填充所有字段。共计 12 个字段

```
{
    contractAddress : 项目合约地址
    txhash           : 交易 hash
    blockno          : 区块号
    timestamp        : 对应 utc 时间戳
    srcUser          : 发送方
    srcID            : 源链
    srcToken         : 发送 token
    srcAmount        : 发送数量
    dstUser,         : 接受用户
    dstID            : 目标链
    dstToken         : 接收 token
    dstAmount        : 接受数量
}
```

研究对象：位于 ETH, BSC, Polygon 三条链上的项目。

研究的时间范围为，从初始项目至 2022 年 8 月 31 号，各个链条对应的截至区块：

	ETH	BSC	Polygon
截至区块高度	15449611	20936752	32534395
交易数量	417618	1695016	810375

2 获取数据

2.1 获取交易 hash

代码：

`getBlockTest.py` 根据 event 获取所有的事件
`dataFill_portal.py` `dataFill_stargate.py` 分别用来填充相关项目的字段
`getTokenChain.py` 用来做一些数据的处理，功能较多，比较凌乱。

2.1 获取价格

Workdir : 位于 getPrice 目录之下

`getPriceFromCex.py` 从 cex 中获取价格, 使用官方提供的免费的存档节点

例如 ETH: univ2 BSC:pancake v2 Polygon: quick v2

`getPriceFromCex.py` 该代码中有两个功能, 从币安交易所中获取 5min 价格数据, 使用 eisberg 供应商提供的 位于 ETH 的 uniswap v2 的数据, (目前接口已经失效)

`get_data_from_uni_v2.py` univ2 官方提供的获取数据索引方式 (仅调试, 未实际应用)

`/token` 该目录下存放的是涉到项目的相关 token

获取价格的核心思路 (比较耽误时间)

以 cex 为主, dex 作为辅助

获取数据的渠道有几种方式。

CEX : 中心化交易所, 类似于 bnb, 等。

数据厂商: 类似于 dune ,,,

Dex : Uniswap, sushiSwap, pancakeSwap

DEX: 部署个人的 the graph 节点爬取数据

(价格说明: 价格来源于 pair 合约中的 pair 数量的比值, 注意, 采用 uni 的 swap 事件获取的价格, 不够准确, 在 swap 过程中, 因为滑点会造成价格不准确)

最终的数据来源是 bnb 的交易数据, 下载 5min 交易数据, 再加上各个 dex 爬取的数据

数据分为三类:

	cex	dex		total
		Symbol	No symbol	
ETH	296122	121254	242	417618
BSC	1390802	295552	8662	1695016
Polygon	727610	82672	93	810375

注: Symbol 指 erc20 token 有 symbol, 类似于 ETH, DAI 这些

No symbol 指的是 token 中无 symbol, 有些不具备 symbol 以及 decimal , 需要后期具体分析

具体操作核心步骤:

将不同 symbol 归一化

步骤一 :

根据交易确定, 所有的交易 token, 大多数 token 是封装的,

DOGIRA < anyDOGIRA, 分析每笔交易的 logs, 确定 token 的对应关系, 得到 real_{chain}_token_block.csv

参考目录: `/token`

步骤二:

合并 symbol 相同的 token, 生成类似于

```

1446 BTC,BTC,nuBTC,gBTCB,BTCBR
1447 BIT,BITD,BitANT,BITMATIC,SenkaBit
1448 WIN,WINS,XWIN,WING,SWINGBY
1449 HARE,SHARE,PSHARE,DSHARE,MSHARE
1450 ADA,NADA,ADAL,RADAR,STADA,ADAPAD
1451 ETH,ETH,nETH,ETH0,ETHPAD,FomoETH,50WETH-50WMATIC
1452 POL,POLP,POLC,POLS,POLAR,POLARv3,POLYPAD
1453 SOL,SOL,SOLO,SOLT,MSOL,SOLX,SOLCH
1454 ART,ARTH,MART,CART,ARTEM,START,NFTART,ARTH.bsc
1455 USD,USD,USDS,DUSD,AUSD,USDN,iUSDS,USDibs,terrausd
1456 DOG,DOGE,DOGO,SDOG,Dogk,SDOGE,BabyDoge,METADOGE,PAPADOGE
1457 SHI,SHIB,YOSHI,SUSHI,SHILL,SHIBG,Sushi,RYOSHI,SHINJA,bioShib,SAUDISHIB

```

手动检查，将相同 symbol 的进行合并。

确定一组稳定币，

```

["USDC","USDT","DAI","yvDAI","BUSD","TUSD","USD","USDD","USDK","LUSD",
"PUSD","cUSD","a1USD"]

```

将他们看作是等的稳定币。

筛选一：

将同名 symbol 与 cex 中同名 的价格 进行兑换。

相同 symbol 的价格兑换之后，(大部分手动)对比相同 token 之间的价格，确认相同的 token 是价格一致的，筛选出了一些不相等的价格的 token，将这些筛选出来，在 dex 中查找价格

筛选二：

symbol 相同的 Token，可能存在不相等的情况，需要分别查找，筛选。同一个项目方同 symbol，其价格可能也不相同

剩下的币种需要从各个链的 dex 中去获取。大多数相同 token 需要去浏览器查看交易情况。

最终筛选出 9 个文件，在 dex 中进行价格查询

```

BSC_no_price.csv    ETH_no_price.csv    Polygon_no_price.csv
BSC_no_priceSymbol.csv  ETH_no_priceSymbol.csv  Polygon_no_priceSymbol.csv
BSC_have_price.csv   ETH_have_price.csv   Polygon_have_price.csv

```

#dex 数据获取

1: 最直接，最全面，所依赖本地节点

根据 factory 合约获取所有的 swap pair address，存储相关的 token，计算价格

参考网站：

币安交易所

https://www.binance.com/zh-CN/trade/USTC_BUSD?theme=dark&type=spot

<https://data.binance.vision/?prefix=data/spot/daily/klines/> 币安的历史数据

抹茶交易所

<https://www.mexc.com/zh-CN/markets/spot>

最打最完整得第三方平台

<https://coinmarketcap.com>

以上 3 个中数据都不够全面

* 后期从 dex 获取数据之后，注意一些市值比较小的币种，防止池子过小，造成价格过大。

Dune(有些小币种不存在)，bsc 等第三方交易所平台，数据长度不够，数据量也不够。

取 UniswapV2 的数据，uniswapv3 上大币种一般比较多，综合来看，取 uniswapv2 的币。

查看价格：

[https://www.dex-tools.io/app/cn/bnb/pair-](https://www.dex-tools.io/app/cn/bnb/pair-explorer/0x4b729d5d871057f3a9c424792729217cde72410d)

[explorer/0x4b729d5d871057f3a9c424792729217cde72410d](https://www.dex-tools.io/app/cn/bnb/pair-explorer/0x4b729d5d871057f3a9c424792729217cde72410d)

<https://tw.tradingview.com/symbols/BTCUSD/>

<https://tradingstrategy.ai/trading-view>

<https://coinmarketcap.com/tokens/views/all/> #可以看市值前 100 的币种

<https://graphql.bitquery.io/user/account> #收费

相关获取价格的博客

<https://blog.windisco.com/calculate-history-price-of-token-according-to-dex-trades/>