Коэффициенты централизации и доверия блокчейна.

noise@sumus.team

1 декабря 2019 г.

Аннотация

В статье предлагается количественная оценка централизованности работы сети блокчейн в виде коэффициента, вычисленного на основе параметров сети. По значениям коэффициента оценивается работает ли сеть хаотично, централизованно или децентрализованно. Получена оценка оптимального количества узлов сети, использующей алгоритмы, основанные на достижении консенсуса, для лучшего сочетания быстродействия сети и уровня доверия к результатам её работы.

1 Введение.

Одним из главных вопросов, на которые следует ответить при оценке работы сети блокчейн является обеспечение равноправия узлов в сети и их равномерной загрузки при закрытии блоков. Такое состояние сети можно назвать децентрализованным. В идеальном случае в децентрализованной сети все работающие достаточно долго узлы закрывают одинаковое количество блоков.

Децентрализованному состоянию сети противоположны два других состояния, одинаково нежелательных. Первое — хаотичная работа сети, когда много узлов закрывают малое количество блоков каждый, но существует меньшее множество узлов, в котором каждый узел закрывает много блоков, так, что эти узлы закрывают примерно столько же блоков, сколько закрывают узлы из первого множества. Второе — централизованная работа сети, при которой большинство узлов закрывают по одному-два блока, а несколько узлов закрывают почти все блоки.

Задача состоит в построении количественной оценки сети блокчейн, позволяющей на основе простейших параметров её работы получить быстрый ответ на вопрос об уровне централизации/децентрализации/хаотичности сети на определённом интервале времени.

2 Оценка централизации работы сети блокчейн.

Множество $\aleph$ мощности α - множество блоков, закрытых на момент t^*, $B_n \subset A_N$ — множество узлов мощности n, принимавших участие в закрытии блоков из $\aleph$ [1], A_N — множество всех узлов сети.

Рассмотрим решётчатую функцию $y(i)$, где $i = 0, 1, \ldots, \alpha$, которая строится по следующему правилу: $y(1) = n_1$ — количество узлов, закрывших только по одному блоку; $y(2) = n_2$ — количество узлов, закрывших только по 2 блока, $\ldots$, $y(\left[\frac{\alpha}{n}\right]) = n\frac{\alpha}{n}$ — количество узлов, закрывших по $\left[\frac{\alpha}{n}\right]$ блоков, $y(\alpha) = n_\alpha$ — количество узлов закрывших по α блоков каждый. Очевидно, что n_α равно 0 либо 1; $y(0) = n_0$ — количество узлов, которые не закрыли ни одного блока (не работали), $i = 1, \ldots, \alpha$, как правило, $\alpha \gg n$.

Рассмотрим некоторые важные частные случаи значений функции $y(i)$.

а) Если $n_\alpha = 1$, то сеть максимально централизована и все блоки были закрыты одним узлом.

б) Если $n_\alpha = 0$, то исследование можно продолжить.

в) Если $n_1 = n$, то $n = \alpha$ и каждый узел закрыл по одному блоку, причём условие $\alpha \gg n$ не выполняется.

г) Если $n\frac{\alpha}{n} = n$, то каждый узел закрыл по $\left[\frac{\alpha}{n}\right]$ блоков, что соответствует максимальной децентрализации без хаоса. Случай в) относится также к децентрализованному состоянию сети, но не может поддерживаться длительное время.

д) Если n_0 близко к n ($n_0 < n$), то в блокчейне много неработающих узлов и надо пересматривать постановку задачи.

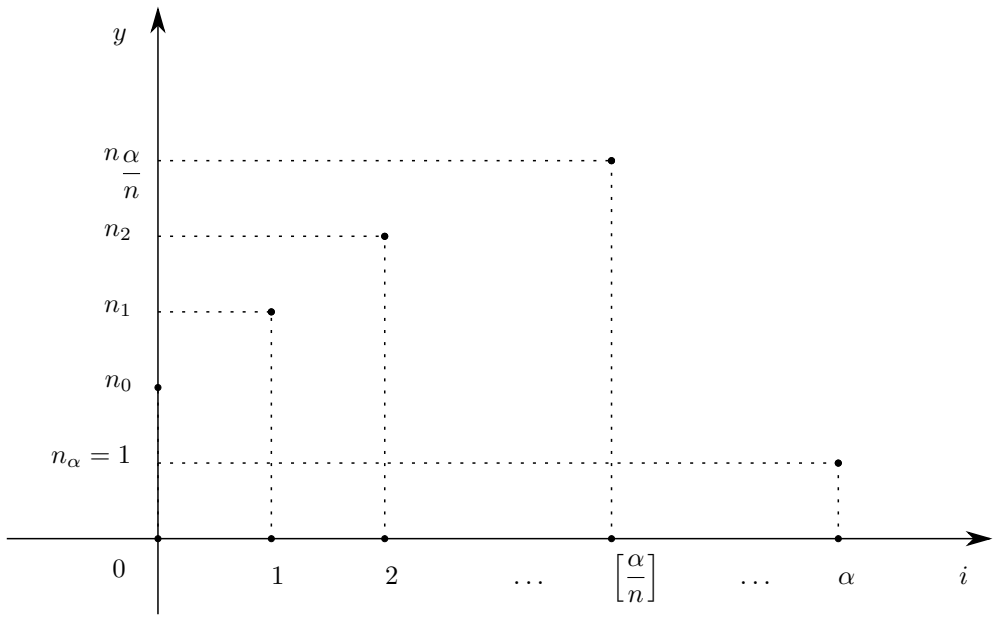

Рис. 1: Зависимость количества узлов, закрывших одинаковое количество блоков, от количества блоков, закрытых одним узлом.

Рассмотрим упрощённый вариант коэффициента централизации. Пусть $i^* = \underset{i=0,\ldots,\alpha}{\operatorname{argmax}}(y(i))$. Будем считать, что i^* единственно.

Введём коэффициент централизации

$$K_y = 1 - \frac{\alpha}{n \cdot i^*},\ K_y \in [1 - \frac{\alpha}{n}, 1] \tag{1}$$

Если $K_y < 0$, то это указывает на тенденцию к хаотичной работе системы, например, при $i^* = 1;\ \alpha \gg n$

$$K_{y\,\min} = 1 - \frac{\alpha}{n} < 0\ ,$$

при $K_y = 0$ достигается наиболее полная децентрализация без хаоса, $i^* = \frac{\alpha}{n}$. Если $y(i^*) = n$, то это соответствует полной децентрализации.

Если $K_y > 0$, то это указывает на тенденцию к централизации, при $K_{y\,\max} = 1 - \frac{1}{n} > 0$, $i^* = \alpha$, а $n_\alpha = 1$.

Если $K^* = \max\{|K_{y\,\min}|, K_{y\,\max}\}$, то нормированный коэффициент централизации есть

$$K_{y\,\mathrm{norm}} = \frac{K_y}{K^*} \tag{2}$$

Коэффициент K_y можно рассчитать, используя вместо i^* среднее количество блоков, закрытых одним узлом $\langle i \rangle$, которое определяется выражением:

$$\langle i \rangle = \underset{\tilde{i}=0,\ldots,\alpha}{\operatorname{argmin}} \left| \sum_{i=0}^{\tilde{i}} y(i) \cdot i - \sum_{i=\tilde{i}+1}^{\alpha} y(i) \cdot i \right| \tag{3}$$

$$K_y = 1 - \frac{\alpha}{n \cdot \langle i \rangle} \tag{4}$$

Использование (4) требует бóльших вычислений, но результатом будет более точная оценка централизации. Наиболее предпочтительным случаем является совпадение значений K_y, полученных по формулам (1) и (4). Если эти значения существенно различны, то сеть не может быть признана децентрализованной.

3 Оценка доверия к сети блокчейн.

Как было указано в [1],все узлы блокчейна составляют множество A_N, разбитое на K непересекающихся подмножеств B_n^i

$$A_N = \bigcup_{i=1}^{K} B_n^i \tag{5}$$

Узлы включаются в каждое B_n^i , в соответствие с гипотезой о равномерном распределении их номеров. Каждое множество B_n^i имеет одну и ту же мощность n и по своему смыслу эквивалентно множеству из раздела **2**.

Сколько надо взять узлов в B_n^i, чтобы это вызывало достаточное доверие к результатам работы сети? Чем больше n — тем лучше для доверия системе, но при этом, чем меньше n по отношению к N, тем ниже вычислительные затраты на выработку консенсуса. Тогда чем больше каждая из величин n, $\dfrac{N}{n}$, тем меньше $z(n) = \dfrac{1}{n} + \dfrac{n}{N}$. Найдя минимум по n этой функции, получим оптимальное (4).

$$n_{\text{opt}} = \sqrt{N} \tag{6}$$

Можно считать $z(n)$ коэффициентом “недоверия”. Тогда $K_{\text{dr}} = \dfrac{1}{z(n)}$ — коэффициент доверия без потери работоспособности, достигающий максимума при $n = n_{\text{opt}}$.

$$K_{\text{dr}} = \frac{n \cdot N}{n^2 + N};\ K_{\text{dr}\,\max} = \frac{\sqrt{N}}{2} \tag{7}$$

Нормированный коэффициент K_{dr} :

$$K_{\text{drn}} = \frac{K_{\text{dr}}}{K_{\text{dr}\,\max}} = \frac{2n\sqrt{N}}{n^2 + N} \in [0, 1] \tag{8}$$

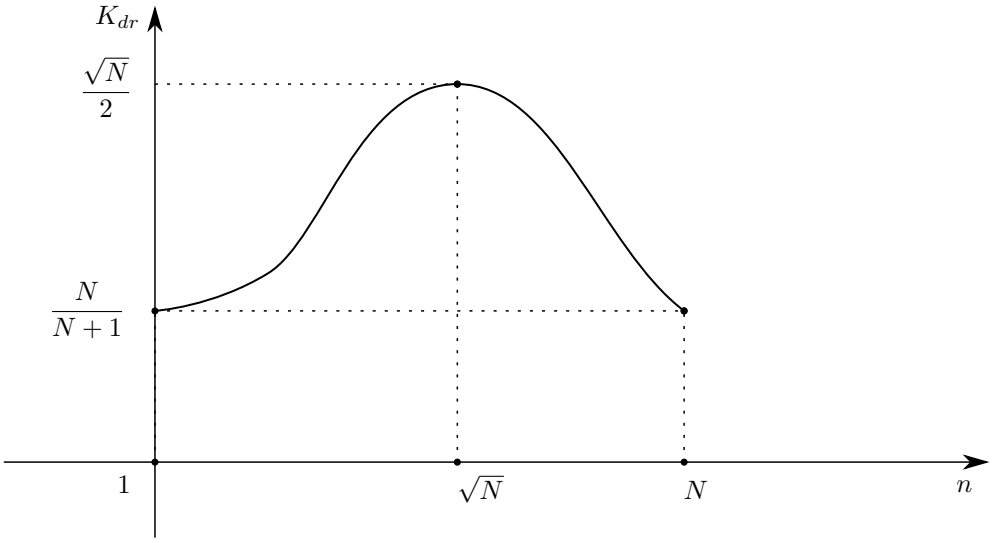

Рис. 2: Зависимость коэффициента доверия K_{dr} от количества узлов в B_n.

4 Об оптимальном числе узлов в множестве “эскорт + мастер” в сетях блокчейн, использующих алгоритм *SDBFT*

Если для блокчейна с множеством узлов B_n необходимо определить оптимальное число узлов в множестве $B_{n'}$ состоящему из мастер-узла и узлов эскорта, то можно рассуждать аналогично предыдущему разделу, заменив N на n, а n на n' в формуле (4). Тогда получим

$$n'_{\mathrm{opt}} = \sqrt{n} \tag{9}$$

Если требуется коэффициент доверия для $B_{n'}$, то с указанными заменами переменных получим аналоги формул (7), (8) с учётом ограничения алгоритма *SDBFT*.

$$n' \geq 5 \tag{10}$$

5 Пример применения коэффициентов централизации и доверия

Взяты данные работы блокчейна BITCOIN [2] за три последовательных достаточно длительных интервала времени. Для каждого из них построена функция $y(i)$ и вычислен коэффициент централизации (4).

Первый интервал: $\alpha = 210000$; $n = 186660$; $\left[\frac{\alpha}{n}\right] = 1$; $i^* = 1$; $K_y = -0,125$. Блокчейн работает практически децентрализованно, близко к случаю в). Такой режим работы не может поддерживаться сколь угодно долго, так ка требует почти полного выполнения условия $n = \alpha$.

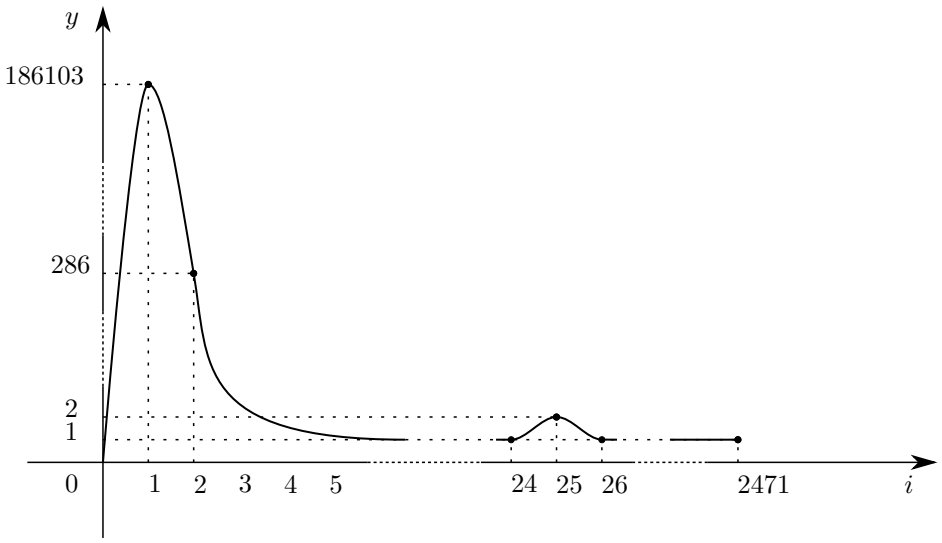

Рис. 3: Зависимость количества узлов от количества закрытых ими блоков для первого интервала.

Второй интервал: $\alpha = 210000$; $n = 9774$; $\left[\frac{\alpha}{n}\right] = 21$; $K_y = 0,996$, что соответствует централизованной работе блокчейна.

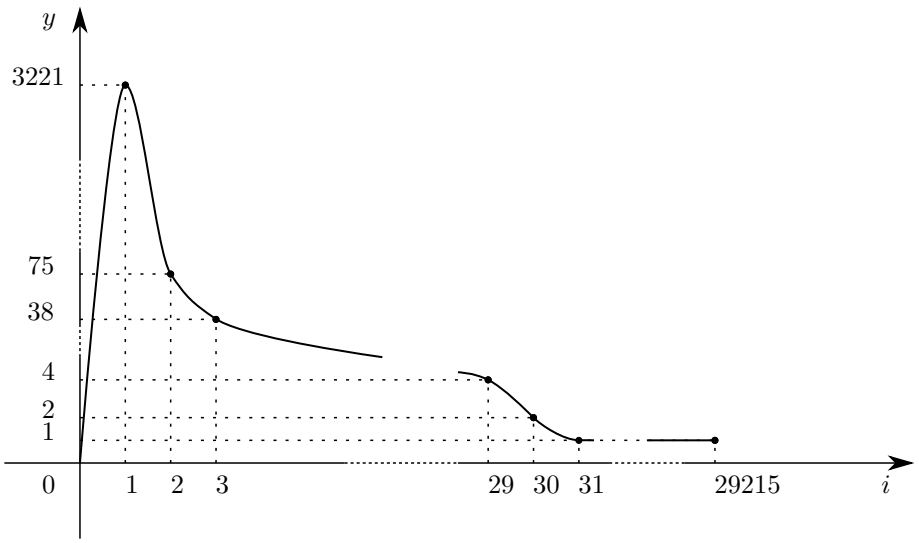

Рис. 4: Зависимость количества узлов от количества закрытых ими блоков для второго интервала.

Третий интервал: $\alpha = 188930$; $n = 512$; $\left[\frac{\alpha}{n}\right] = 369$; $K_y = 0,954$, что соответствует централизованной работе блокчейна.

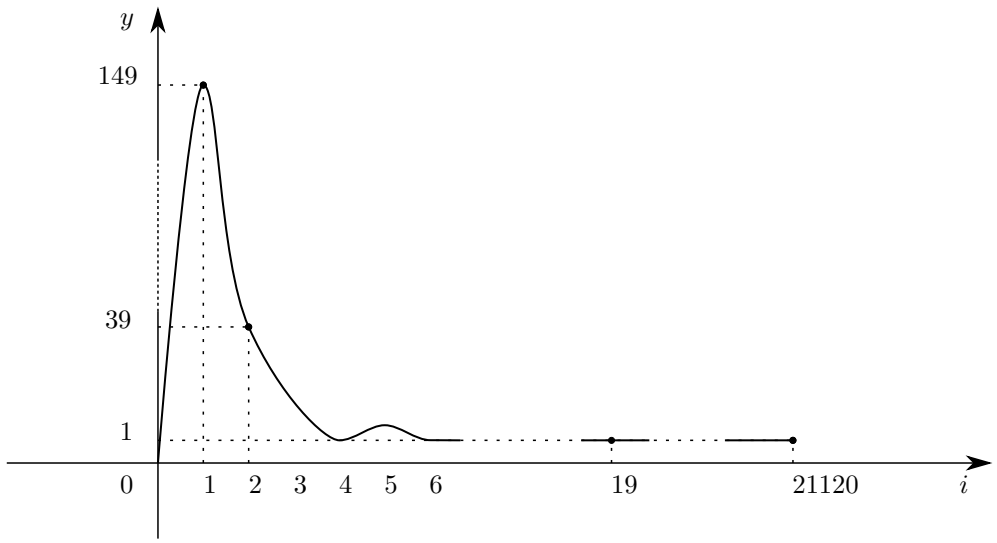

Рис. 5: Зависимость количества узлов от количества закрытых ими блоков для третьего интервала.

Заключение

В статье предложена удобная оценка равномерности загрузки узлов сети блокчейн в процессе закрытия блоков. Эта оценка достаточно проста и вычисляется как коэффициент, учитывающий количество закрытых блоков, количество узлов, закрывших определённое количество блоков и называется коэффициентом централизации.

В нормализованном виде этот коэффициент принимает значения от -1 до 1, что соответствует состоянию сети от полного хаоса (-1) до полной централизации (1), что позволяет быстро оценить соответствует ли данная сеть блокчейн на определённом интервале времени главной идее блокчейна — равноправию и независимости узлов (децентрализации) которой соответствует нулевое значение коэффициента.

Как показали исследования сети BITCOIN на трёх интервалах времени, эта сеть может достаточно долго находиться в децентрализованном состоянии при условии, что количество узлов линейно связано с количеством закрываемых ими блоков.

Однако на других интервалах времени при нарушении этой зависимости сеть приходит к централизации.

В статье также предлагается оценка доверия к работе системы, при условии её децентрализованности, исходя мощности множества узлов, обеспечивающего формирование эскорта и возможность выбора мастер-узла для выработки консенсуса и мощности множества всех узлов, включённых в сеть.

Список литературы

[1] a@sumus.team, k@sumus.team, rr@sumus.team. “Consensus Algorithm for Bigger Blockchain Networks” (2018).

[2] Satoshi Nakamoto (2009). “Bitcoin: A Peer-to-Peer Electronic Cash System”. www.bitcoin.org .