第4章　Bitcoin システムのモデル化

本章では、第 2 章で述べた Bitcoin のネットワーク、そして個々のノードの振る舞いの実装の情報を素に簡略化し、モデルの作成を行う。

4.1 モデルに含める情報の限定

本研究での Bitcoin システムのモデルにおいて、ネットワークを伝搬するオブジェクトは block、及びアドレス情報という情報構造体のみに限定し、決済情報等は省く事とする。

4.2 Bitcoin システム

Bitcoin システムは、

$b_h \in B$　block とそれらの block が連結して形成する blockchain

M　参加ノード群

L:M×M　それらのノード間のリンク

A　それらのノードのアドレス情報でメッセージでもある

により構成される。

Bitcoin システムは、参加ノード群 M が個々にリンク L を確立し、block 等の情報を伝搬し、blockchain B を同期する事で単一グローバル台帳通貨として成り立っている。

Bitcoin システムでは、$h>1$ となる任意の高さ h の block b_hは、一つ前の高さの特定の block b_{h-1}に対して

$$b_h \leq b_{h-1}$$

となる半順序集合の関係が成り立ち、当半順序集合を連結的に形成する block 群により片方向連結リスト構造となり、当該 block 群の最新 block の高さを h とした blockchain B_hを成し、個々のノードが保有する。

個々のノードは、最新 block の受信、また自身の作成する block の拡散を求めて他ノードとのリンクを確立する。また自身と隣接していないノードにも接続を促す為アドレス情報メッセージの拡散による宣伝、また伝搬を行う。

4.3 オブジェクトの定義

以上の事から Bitcoin システムを構築する要素であるオブジェクト群を以下のようにモデル化し表現する。

任意のノード M_i はタプル $M_i(b,a,r,l,t)$

b :　当該ノードの保有する最新 block

a :　当該ノードのアドレス

r :　当該ノードのネットワークに占める block 作成能力の割合

l :　当該ノードの保有するリンク群で、個々にタプル $l(d,s)$となっており：

d :　リンク先の通信相手となるノード

s :　リンクを通して情報を伝達するのに要する遅延

t :　過去に接続したノード、受信したアドレス情報により構築されるアドレステーブル

特定の block 情報構造体は、 B_i タプル $B_i(p,h)$

p :　当該 block が作成される際参照された block

h :　当該 block の blockchain に占める高さ

で表現される。

そして任意のアドレス情報の構造体は　A_i タプル $A_i(e_{A_i}, \alpha_{A_i})$

e :　当該 addr の内包する α のエントリー数

α :　当該 addr の内包するエントリー群で、個々にタプル $\alpha_i(d_{\alpha_i}, s_{\alpha_i})$となっており：

d_{α_i} :　当該エントリーの指すノードのネットワークアドレス

s_{α_i} :　当該 addr 作成時に送信者が刻んだタイムスタンプ

で表す事が出来る。

4.4 パラメーター一覧

本研究におけるパラメータの一覧を以下の表 1 に表す。

表 1. パラメーター一覧

表記	内容	値
N	常時最低存在ノード数	6000 台
hogehoge	時間経過による存在ノード数の変動	http://www.bitcoinpulse.com/#/chart/bitnodes/num_nodes
これこれ	新規参入ノード数のレート	https://blockchain.info/charts/my-wallet-n-users
あれあれ	永久離脱ノード数のレート	https://blockchain.info/charts/my-wallet-n-users
何%	外部から接続を受け付けないノード数	何%
P_b	ネットワーク全体の block 作成確立	$\frac{1\text{ block}}{594.6231417\text{s}}$ ~~$\frac{1\text{ block}}{544.216\text{s}}$~~ ~~$\frac{1}{\epsilon}$~~

このような分布	Block 作成能力の全ノードに対する分布	http://organofcorti.blogspot.com.au/search/label/weekly%20network%20statistics
	Block の受信と他ノードへの送信に要する時間	平均 0.55124 秒
	Addr の受信と他ノードへの送信に要する時間	上記の 1/3

4.5 Bitcoin ノード群

Bitcoin システムにおいて、そのネットワークには常時最低 6000 台(仮、パラメータ)のノードが存在し、ある特定の時刻 t に存在するノード数 m は常に m>n な関係となり、そのノード数 m の時間の経過による値の変動は hogehoge (パラメータ)となり、fugafuga な分布を描く。

https://getaddr.bitnodes.io/nodes/

https://bitinfocharts.com/comparison/nodes-btc.html#1y

http://www.bitcoinpulse.com/#/chart/bitnodes/num_nodes

(bitcoind を改造し多くのノードと接続を行い、常時ネットワークを伝搬する addr を一定の時間観測する事で、ノード数を調べようと思います)

Bitcoin システムへの新規参入者は、これこれ(パラメータ)のレートで参入しており

https://blockchain.info/ip-log

https://blockchain.info/charts/n-unique-addresses?timespan=all&showDataPoints=false&daysAverageString=1&show_header=true&scale=0&address=

https://blockchain.info/charts/my-wallet-n-users

ノード数 m と 新規参入者 があれあれな値な為、事実上永久に離脱するノードの数はこのような値(パラメータ)となる。

4.6 リンク

Bitcoin システムにおいて、個々のノードは最新 block の取得、並びに自身の作成する block の伝搬等を求めて他ノードに対してリンクを確立、また他ノードからの接続要求を受け付ける。

各々のノードは、最大で 8 つのノードに対外的に接続を要求、確立し、最大で 117 のノードからの接続要求を受け付ける。しかし何%の(パラメータ)ノードは、外部からの接続要求を受け付けない為、ネットワーク全体のリンク数はネットワークに存在する全ノード数を M とし、それらのノードが当該時点に最大対外接続数を満たした場合：

$$\mathrm{L} = \frac{\left(\left(8 \times \mathrm{A} \times \text{何\%}\right) + \left(\mathrm{M} - \left(M \times \text{何\%}\right)\right) \times \left(rand()\%118 + 8\right)\right)}{2}$$

で表す事が出来る。

4.7 アドレステーブルと他ノードへの接続

ノードは他ノードとのリンクを求めるが、その際最初期であれば乱数的にシードノードへ接続、その他の場合、自身のアドレステーブル t を参照する事となる。このアドレステーブル t は、個々にタプル $\mathrm{t}(c,r,n,c)$となっており：

c：個々のアドレス情報のエントリー群で、個々に$c(cS,nL,nA,nR,)$となっており：

nI：後述の r、n、c、にて登録される当該エントリーの ID

cS：当該アドレスを当ノードがどの他ノードから伝えられたか

nC：接続実績有りの情報構造体に登録されているか

nR：新規アドレス情報群の中に幾つ重複登録されているか

等の管理情報があり、またこれらのアドレス情報の接続性を表す情報群として：

nL：最後にいつ当該アドレスに対して接続に成功したか

nA：最後に接続してから何度接続を試みたか

nT：アドレスの宣伝等によりその活動をいつ確認したか

等の情報を内包する。

r：アドレス情報エントリー群を乱数的に一次元配列にしたもの

n：接続実績の無いアドレス情報群をどのノードから受信したかにより分類した情報構造体

c：接続実績の有るアドレス情報群を当該アドレスのプリフィックスにより分類した情報構造体

先ず rは、特定の新規ノードがネットワークに接続を行い、自身の保有するアドレステーブルが内包するエントリー数が 1,000 未満の場合アドレス情報群を要求し、それに対する応答として転送を前提としない大容量のアドレス情報メッセージを送信する。その際乱数的に順序を入れ替えながら送信候補を選択するのが rである。

nと cだが、これらは共に対外接続ノード数が最大の 8 にみたないノードが接続先の選定に用いられる情報構造体であり、その内容は、戦術の大容量のアドレス情報メッセージ及び、ネットワークを伝搬するネットワークアドレスの宣伝を目的として伝搬しているアドレス情報メッセージにより埋められる。

4.7.1　アドレス情報の登録

他ノードに接続したノードは、自身のアドレステーブルのエントリー数が 1,000 に満たない場合アドレス情報要求を送信し、後に多量の転送を前提としないアドレス情報を受信、また他ノードが自身のネットワークアドレスを宣伝した物が伝搬されて来たものを受信する。その後当該ノードはこのアドレス情報を自身のアドレステーブルに登録する。

任意のメッセージとして受信した新規アドレス情報 α_i に対し、登録済みであるか否かを

指す指示関数 $P(\alpha_i)$、登録済みである場合のその当該アドレスのエントリーを c_{α_i} 、多重登録を行えるか否かを示す関数を $R(c_{A_i})$とすると：

$$P(\alpha_i)\begin{cases} 1 & (c_{\alpha_i}\text{が存在しない}) \\ R(c_{\alpha_i},\delta) & (c_{\alpha_i}\text{が存在}) \end{cases}$$

$$0 \le \delta < 2^{nR_{c\alpha_i}}$$

$$R(c_{\alpha_i},\delta)\begin{cases} 1 & (nC_{c_{\alpha_i}} = 0\,;\, nR_{c_{\alpha_i}} < 4\,;\, s_{\alpha_i} > nT_{c_{\alpha_i}}\,;\, \delta = 0) \\ 0 & (\text{上記以外}) \end{cases}$$

とし、$P(\alpha_i)$によりα_iが当該ノードのネットワークアドレステーブルの接続実勢のないアドレス情報群を保管する情報構造体 n において、登録もしくは多重登録され接続対象として選ばれる可能性を増す事が出来るか否かを表現出来る。

4.7.2　他ノードへの接続

ノード m

ノード m の対外接続数 $nO_m = \sum l_{mo}$

接続先設定関数に与えるバイアス値$B_m = 10 + \begin{cases} 80\ (nO_m > 8) \\ 10 * (nO_m) \end{cases}$

選択に用いられる乱数値$0 \le GRI < 2^{30}$

接続実績のあるアドレス情報のエントリー数 nCe

接続実績のないアドレス情報のエントリー数 nNe

仮の変数 1　$\mathrm{nCT} = \sqrt{\mathrm{nCe}} * (100 - B_m)$

仮の変数 2　$\mathrm{nCN} = \sqrt{\mathrm{nNe}} * B_m$

新規もしくは実績情報構造体設定関数$NC(B_m) = \begin{cases} 1\ ((\mathrm{nCT} + \mathrm{nCN}) * \frac{\mathrm{GRI}}{2^{30}}) < nCT) \\ 0 \qquad\qquad (\text{そうでなければ}) \end{cases}$

接続先設定関数$S(NC(B_m)) =$

4.8 block 伝搬のモデル

個々のノードは block を保有しており、全ノードの集合を A と定義した場合、任意の保有している最新の block の高さを b と表記した場合、ノードの集合は：

$$\mathrm{X_b} \in A$$

で、高さが異なる場合

$$\mathrm{X_b} \cap \mathrm{X_{b-1}} = \emptyset$$

となり、また相克的な blockchain fork を引き起こす同じ高さの block b' が存在する場合も：

$$\mathrm{X_b} \cap \mathrm{X}_{b'} = \emptyset$$

となる為、全ての別個の block を最新と認識している任意の集合$\mathrm{X_a}$と$\mathrm{X_b}$の和は、

$$\mathrm{X_a} \cap \mathrm{X_b} = \emptyset \ (\forall \mathrm{a} \neq \mathrm{b})$$

となる。

新規 block bが作成されてそれがネットワークに伝搬される。この際、新規に作成された block bを受信し、当該 block bの高さ以上の block を知らないノードは、隣接するノードに対して転送を行う。個々のノードは 8~125 までのリンクを保有し、時間(t)に block bをそれらのリンクを通して他ノードに伝搬する事で、送受信と採用に要する時間を 1(パラメータ)とした場合、新規に時間(t+1)に 8~125 のノードが新たに当該 block bを知ることになる。

しかし、それらの内いくつかのノードは、既に他ノードから当該 block bの事を知っている可能性や、時刻(t)に当該 block を知った複数のノードとリンクを保有している可能性がある。その為、ネットワーク内の全ノードを A とし、ネットワーク内の block bの前までの block しか持っていないノードをX_{b-1}とし、block の拡散を試みるノードが持つリンクの 75%が同時刻に当該 block を受信した一他ノードと同じ対象に繋がっているとした場合、時間(t)に block bの事を知ったノードの数 x_t は、関数$f(k)$で当該ノードの隣接ノード数を表した場合：

$$f(k)\begin{cases} 8 & (\mathrm{k}\text{ が外部からの接続を受け付けない場合}) \\ 8 \leq \mathrm{a} \leq 125 & (k\text{が外部からの接続を受け付ける場合}) \end{cases}$$

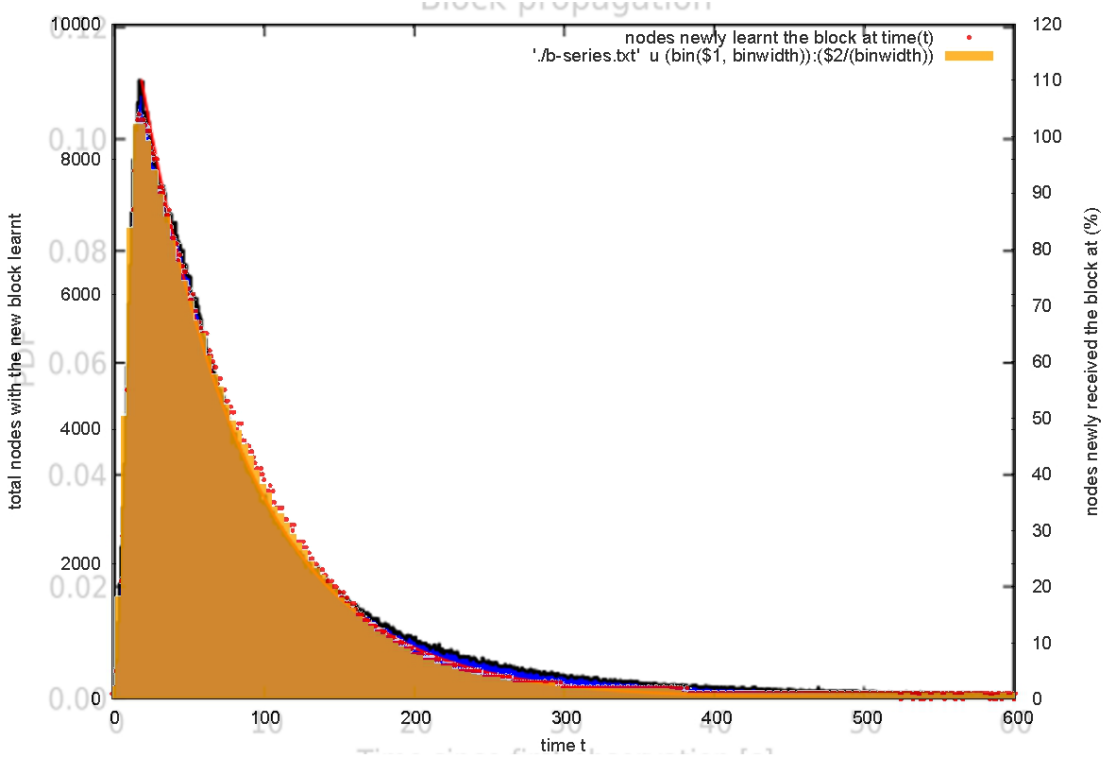
以下の級数で表され。

$$x_t = \sum_{\mathrm{k}=1}^{x_{t-1}} f(k) \times \frac{\mathrm{X}_{b-1}}{A} \times \left(\frac{1}{4}\right)^{k-1}$$

時間(t)までに block bの事を知ったノードの累計数 $\mathrm{X_{b(t)}}$ は：

$$\mathrm{X}_{bt} = \sum_{\mathrm{t}=1}^{t} \mathrm{x_t} \quad A = \sum_{\mathrm{t}=1}^{\infty} \mathrm{x_t}$$

となり、その block の伝搬の様子は全ノード数を 10000 と仮定し、外部から接続を受け付けるノードの割合が 8.21%とした上で、グラフに示すと以下の通りとなる：



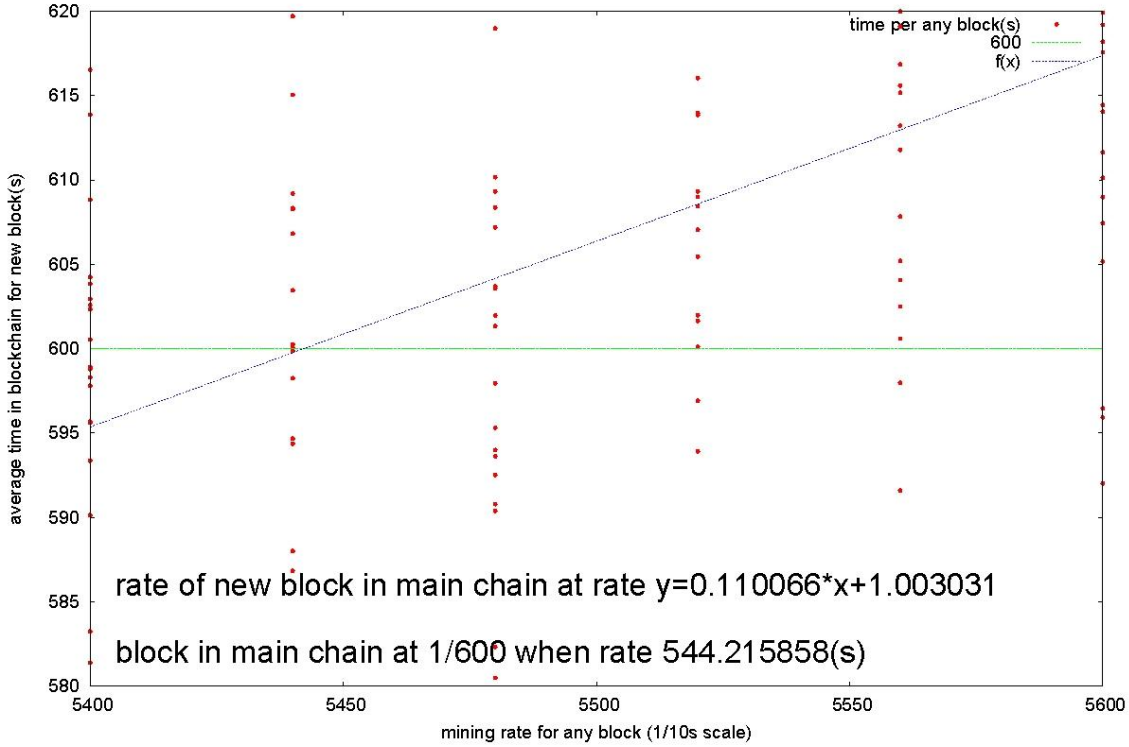
4.9 block 作成

個々のノードは、block 作成報酬を求め block 作成を試みる。実 Bitcoin システムでは、ネットワーク全体で block が作成され blockchain に登録されるレートが 600 秒に 1block となるように調整されており、その過程で作成されても不採用とされる block も存在する為ネットワーク全体で block が作成される実際の確率は 1block/600 秒よりも高いと仮定される。その為前節 4.8 にて示した block の伝搬をモデル化して得た関数を素に相克的な block も作成されるシミュレータを作成し、システム全体の blockchain に新規の block が 1block/600 秒で追記されるようになる実際の block 作成レートは：

$$P_b = \frac{1\ \text{block}}{594.6231417\text{s}} \quad \frac{\cancel{1\ \text{block}}}{\cancel{544.216\text{s}}} \quad (\text{パラメータ})$$

とである事が算定された。

https://blockchain.info/charts/n-orphaned-blocks?showDataPoints=false&timespan=1year&show_header=true&daysAverageString=1&scale=0&format=csv&address=



その為、任意のノードxが block 作成に成功する確率 $\mathrm{P}_{x\mathrm{b}}$ は、ネットワークに存在する全ノード A の block 作成能力r_Aを

$$\mathrm{P_b r}_A = \mathrm{P_b} \sum_{a=1}^{A} \mathrm{r}_a = \mathrm{P_b}$$

とし、当該ノードの Bitcoin ネットワーク全体に占める block 作成能力を：

$$0 \leq \mathrm{r}_x \leq 1$$

とした場合、シミュレーションにて個々のノードが block を作成出来る確率は、シミュレータ内の時間の 1 ユニットと 1 秒の対応を：

$$\frac{1s}{\mathrm{t}_1}$$

とし、前回の block 作成挑戦からの経過仮想時間をt_aとした場合：

$$\mathrm{P}_{x\mathrm{tb}} = \mathrm{P_b} \times \mathrm{r_x} \times \mathrm{t_a} \times \frac{1s}{\mathrm{t}_1}$$

となる。なお、ノード群の個々に保有する block 作成能力値である r は、その総和が 1 となり、そのノード毎の block 作成能力値はこのような分布となる。(パラメータ)

4.10　addr 伝搬のモデル

Bitcoin ネットワークにおいて、block の他にもう一つ伝搬する情報構造体として addr が挙げられる。addr は情報構造体であると共にノード間で送受信するメッセージでもある。addr

は、個々のノードに関する情報のエントリーが複合して成り立っており、個々のエントリーは：

・そのエントリー情報の新しさを表す時刻

・当該エントリーが示すノードとの接続に必要なアドレス情報

で出来ている。

addr はそのサイズにより 2 つにわけられ：

・多数のエントリーを内包する addr で、受信者は転送を行わない

・少数エントリー内包 addr で、受信者は隣接する 2 ノードに転送する

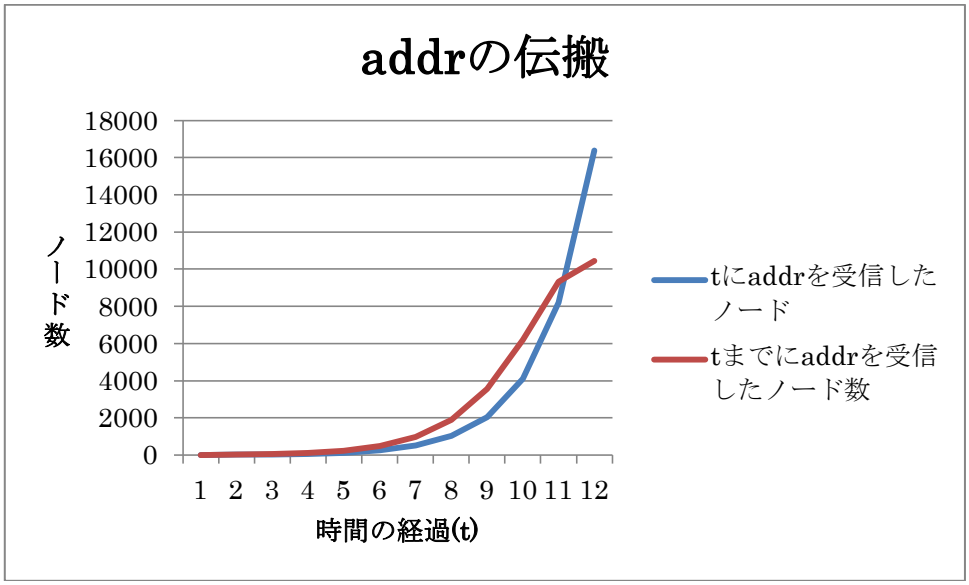
後者の伝搬モデルは、ノードの addr 受信及び他ノードへの転送に要する時間を t=1 とし(パラメータ)、特定の時間 t に当該 addr を受信したノード数をx_tとし、初期値を t=0, x_0=当該 addr 発行ノードの隣接ノード数と定義すると

$$x_t \quad = x_{t-1} \times 2$$

ネットワーク内の全ノード数を A とし、任意の時間(t)[t<3600]までに当該 addr の事を知ったノードの累計数を $X_{b(t)}$とすると、$X_{b(t)}$は：

$$X_{b(t)} = X_{b(t-1)} + x_t \times \frac{A - (X_{b(t-1)})}{A}$$

だと仮定され、以下のグラフの様に当該 addr 受信ノードが増加する物と思われる：



(これも取り敢えず仮です..)