

セルオートマトンによる相互結合網の間欠的な輻輳の解析

横田 隆史[†] 大津 金光[†]
古川 文人^{††} 馬場 敬信[†]

独立動作する多数のルータの集合で構成される大規模相互結合網では、輻輳の制御が大きな鍵となる。輻輳を上手く制御するには、対象となる系で発生する輻輳のメカニズムや性質を明らかにすることが第一となる。しかし大規模相互結合網では、個々のルータの構成は単純であっても全体として複雑な挙動を示すため、メカニズムの解明は困難だった。本稿では、対象を簡略化して得られるセルオートマトンモデルを用いて、間欠的に観測される輻輳の解析を試みる。輻輳の発生条件を解析的ならびに実験的に求めるとともに、輻輳の発生・消滅に関する時間的な挙動を解析する。また一時的・局所的なゆらぎに起因する輻輳発生条件の解明も試みる。さらに、セルオートマトンモデルで得られた結果と同様のことが、相互結合網シミュレータ上でも観測されることを示す。

Analysis of Intermittent Congestion in Large-Scale Interconnection Networks

TAKASHI YOKOTA,[†] KANEMITSU OOTSU,[†] FUMIHITO FURUKAWA^{††}
and TAKANOBU BABA[†]

One of the most important issues in large-scale interconnection networks is appropriate control of congestion. Clarifying congestions' mechanisms and properties should be the first step to the problem. However, its actual behavior is too complicated to unveil the essential part. Our approach to this problem is to use a cellular automata (CA) model. This paper represents analysis of intermittent congestion by means of CA. We discuss critical congestion condition in both analytical and simulation ways. Our interconnection network simulator reveals quite similar phenomena that will be applied to actual networks.

1. はじめに

大規模相互結合網では、輻輳の制御が重要な鍵となる。輻輳の発生を上手く制御することにより、相互結合網をその最大能力が得られる状態で使用したい。このため我々は均一なトポロジを持つ大規模相互結合網を対象として、輻輳の発生・消滅のメカニズムやその性質について研究している。

一般に大規模相互結合網は、自律的に動作するルータの集合からなり、基本的に集中的な制御メカニズムを持たない。個々のルータの構成や動作は比較的単純であっても、それらが独立動作することにより、全体として複雑系としての挙動をなすことがある。特に、ランダム転送など不規則性を伴う転送パターンでは、複雑系ゆえに扱いが難しい。

複雑な挙動をする対象の動作・性質を解明するには、

本質を損なわない範囲でできるだけ簡略化を行うことが、ひとつの有力な方法である。そのために、議論したい内容に関する事項のみを注意深く残し、他の部分は大きく簡略化する。この簡略化の手段として、我々は、セルオートマトンの考え方¹⁾をとり入れたアプローチをとっている。

これまでに我々は、2次元トラス網を単純化することで得られたセルオートマトン・モデルを利用し、輻輳が発生・成長・消滅する基本的なメカニズムを解明するとともに、系の輻輳の度合いを表すための尺度としてエントロピーを提案している^{2)~4)}。エントロピーは、熱力学からのアナロジーによりパケットの移動度をもとに定義され、系の輻輳の度合いを適切に表すことができる。

また我々は先行研究の過程において、ネットワークの転送特性が飽和する臨界の転送負荷の付近で、輻輳が間欠的に発生・消滅する現象を観測している。本稿では、こうした間欠的な輻輳の発生・消滅に関してより詳細な解析を試みる。

[†] 宇都宮大学
Utsunomiya University
^{††} 帝京大学
Teikyo University

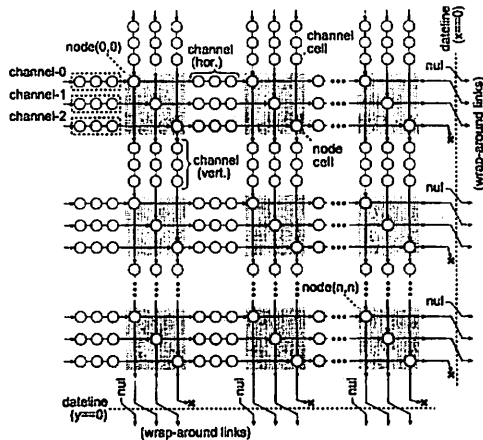


図1 セルトマトン・モデル

2. セルトマトン・モデルと間欠的な輻輳

2.1 セルトマトン・モデル

二次元トラス網の packets 進行方向を、各次元ごとに単方向に制限する。これにより、もともになる相互結合網の構成を簡略化しうえで、packet バッファをセルの列で置き換え、クロスバスイッチと演算ノードをチャンネルごとに割り当てた1個のセルで置換する。その結果得られる相互結合網のセルオートマトン・モデルを図1に示す。

ここで、packet は長さ1とし、1つのセルを占める。packet は、進行方向のセルが空いている場合のみ移動できる。縦・横のチャンネル・セルからノードセルに移動するとき packet 間の競合が発生する。あらかじめノードにクロックごとに1/0の状態をトグルする「信号」を設けておき、競合した場合はその信号の状態によってどちらを優先させるか決める。

2.2 エントロピー測定

packet を R 次元空間を移動する粒子として考える。相互結合網の転送負荷が低い状態では、packet 密度が低く（まばらで）packet 間の衝突が多く発生しないが、転送負荷が高いときは、逆に、packet 同士が頻りに衝突している状況になる。

こうした状況を熱力学からのアナロジーで考え、エントロピーにより定量的に表現する。すなわち、packet の存在密度が低く衝突が減るに派生しない状況ではエントロピーが高く、頻りに packet の衝突が発生する状況ではエントロピーが低いものとする。熱力学のエントロピーをもとにして、以下のようにエントロピーを定義する。時間 Δt に h ホップ進行した packet の数を n_h として、以下のようにエントロピーを定義する。なお N は系内にある packet の総数である。

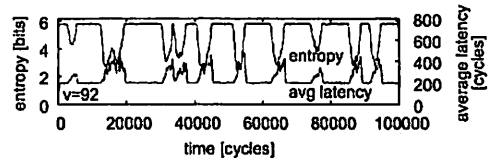


図2 間欠的な輻輳の様子

$$H_{\Delta t}(S) = \frac{1}{N} \sum_{h=1}^{\Delta t} n_h \log_2 h \quad (1)$$

2.3 間欠的な輻輳

一般の相互結合網での評価と同じように、各ノードからの packet 送出間隔をパラメータとして、与える転送負荷に対して得られたスループット、平均レイテンシを求めると、相互結合網と同様の結果が得られる³⁾。packet は必ず仮想チャンネル番号0の位置にあるノードで生成され、送信先はランダムに決められる。送信先に到達した packet は系内から除去される。なお、上記では packet 送出間隔をパラメータとしたが、系内の packet の総数をパラメータとすることもできる。この場合、送出元のノードは、仮想チャンネル0の位置にあるノードであり、しかも packet を持っていないものの中から、packet 総数が指定された一定値になるようにランダムに選択される。

転送負荷を、転送性能が飽和する前後の値にすると、系内に輻輳が間欠的に発生することが観測される。図2はこうした間欠的な輻輳が発生している様子を、式(1)によるエントロピーと平均レイテンシにより表している。同図は、 $\Delta t = 100$ サイクルごとに測定したエントロピーと平均レイテンシの変化の様子を時系列に表示している。

3. セルトマトンによる解析

3.1 平均場近似による平衡条件の導出

$L \times L$ ノード、チャンネルセルの数 P の系に N 個の packet があるものとする。系に輻輳領域があり、しかも大きさが安定している状況を考える。

領域を一边 ω ノードの正方形で近似し、領域内には隙間なく packet があるものとする。また輻輳領域の外部では、packet が衝突なく移動しているものと仮定する。輻輳領域外部で packet が均一に分布しているものと近似すると、平均場近似により輻輳領域の平衡条件を求めることができる。

図3をもとに輻輳領域に吸収される packet の数を考える。輻輳領域内にある packet の数は $2P \times \omega^2$ であり、領域外にある packet の密度は

$$\xi = \frac{N - (2P + 1)\omega^2}{(2P + 1)L^2} \quad (2)$$

となる。輻輳領域外で packet は自由に移動するから

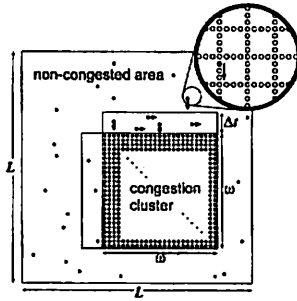


図3 輻輳領域に吸収されるパケット

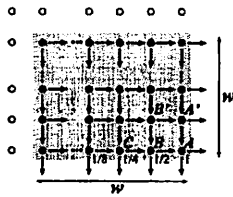


図4 輻輳領域から開放されるパケット

Δt 時間で Δt セル分だけ移動する。x 軸方向、y 軸方向に進行するパケットの数と同じとするならば、 Δt 時間にこの輻輳領域に吸収されるパケット数は $\xi\omega^2\Delta t$ となる。したがって単位時間あたりに輻輳領域に吸収されるパケット数の期待値は

$$\gamma_{in}(\omega) = \frac{N - (2P + 1)\omega^2}{2(P + 1)L^2} \omega \quad (3)$$

となる。

一方、輻輳領域から開放されるパケットはどうか、図4をもとに考える。上述のように矩形的輻輳領域を仮定すると、この領域からパケットが開放されるのは、右・下の辺縁部のみとなる。領域の右上上のノードセルにあるパケットは、その進行方向がx軸の場合に限りそのまま開放される。しかし、進行方向がy軸であった場合ブロックされる。辺縁部のノードセルで唯一ブロックされないのは、領域の右下の頂点に位置するノードAのみであり、この点が明確なボトルネックとなる。単位時間にノードAから放出されるパケット数は1であるから、輻輳領域から放出されるパケットは、結局、

$$\gamma_{out}(\omega) = \sum_{i=1}^{\omega} \left(\frac{1}{2}\right)^{i-1} \quad (4)$$

と算出される。

3.2 臨界パケット密度の導出

3.2.1 平均場近似から

3.1節で輻輳領域に吸収/開放されるパケットの期待値を求めた。系の構成パラメータ L, P およびパ

表1 $\gamma_{in} = \gamma_{out}$ バランス点

L	P	N(analysis)	ω	N(simulation)
16	5	1440	6	650
32	5	3660	10	2300
64	5	9230	16	7500
128	5	23250	25	28500
256	5	58530	40	110000
512	5	147460	64	425000
1024	5	371580	102	(N/A)

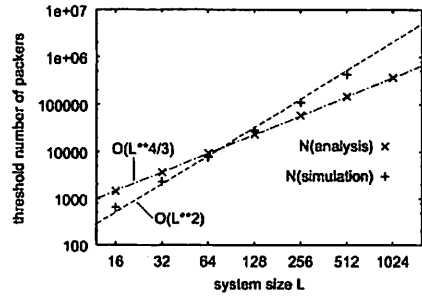


図5 輻輳が発生する臨界の総パケット数

ケットの総数 N が与えられれば、式(3)、(4)から $\gamma_{in}(\omega) = \gamma_{out}(\omega)$ として輻輳領域のサイズ ω を求めることができる。ここで、 ω が正しく求められない場合^{*}、3.1節で行った仮定が成立しない、すなわち、輻輳領域が発生しないと考えることができる。

式(3)、(4)をもとに $\gamma_{in}(\omega) = \gamma_{out}(\omega)$ とすれば、輻輳が発生する N の条件を求めることができる。この式をもとに ω の求解条件から数値的に N の臨界値を求めた結果を表1に示す。この表で $N(\text{analysis})$ は系内に輻輳が発生する臨界の総パケット数を示しており、 ω はそのときの輻輳領域のサイズを表している^{**}。

3.2.2 セルオートマトンから

さまざまな規模 (L) のセルオートマトンモデルに対して、系内の総ノード数 (N) を変化させ転送性能を測定した。その結果から、輻輳が発生する総ノード数の臨界値 (概数) を求めた。結果を表1の $N(\text{simulation})$ に示す。

図5は表1をグラフ化したものである。 $N(\text{analysis})$ はおおよそ $O(L^{4/3})$ の直線に乗り、 $N(\text{simulation})$ は $O(N^2)$ の直線に乗ることが分かる。両者の差は、近似によるものと考えられる。 $N(\text{simulation}) \propto L^2$ であるから、つまり、輻輳の臨界状態でのパケット総数はノード数 (L^2) に比例する。

3.3 輻輳挙動の応答時間

輻輳が間欠的に発生する状況では、図2に示したように、輻輳の発生時点に対応する箇所 (下向きの) ピークとなる時系列変化が現れる。そこで、ピーク

* 正の実数値にならない場合

** ω は四捨五入により整数値としている

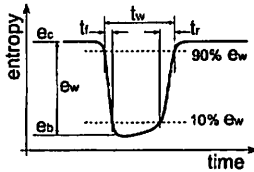


図6 タイミング定義

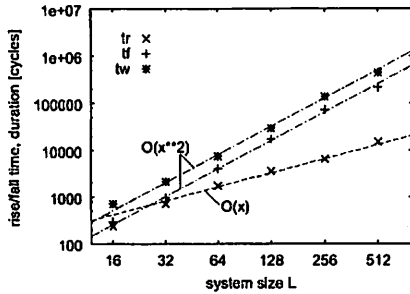


図7 立ち上がり/立ち下がり時間, 持続時間

1回あたりの立ち下がり時間 (t_f , fall time), 立ち上がり時間 (t_r , rise time), ピーク幅 (t_w , duration) を図6のように定義した。エントロピーの天井値 (ceiling value, e_c) は、シミュレーション時間中に観測された非輻転時間でのエントロピーの最尤値である^{*}。ピークの下限值 (bottom value, e_b) は、ピークごとに求めている。ピークの振れ幅は $|e_c - e_b|$ であり、その10%値, 90%値の箇所を t_f , t_r , t_w の測定のために使っている。

$L = 16, 32, 64, 128, 256, 512$ の各場合で測定された t_f , t_r , t_w の平均値を図7に示す。測定に用いたシミュレーション時間は、 $L = 16, 32$ の場合は100,000サイクル、 $L = 64, 128$ の場合1,000,000サイクル、 $L = 256$ の場合は1,000,000サイクルのシミュレーションを2回分、 $L = 512$ では2,000,000サイクルのシミュレーション8回分である。複数回数のシミュレーションを行っているのは、シミュレーションに要する時間が長くなるためである。

間欠的な輻転の挙動、すなわち、輻転が発生・成長・移動・消滅する一連のプロセスと、それら各段階でのエントロピー値の変化を観測した結果から、以下のことが言える。 t_f (立ち下がり時間) は、系内に輻転が発生し、それが系全体に拡がるまでに要する時間を表している。つまり t_f は輻転領域の拡大速度である。逆に t_r は、輻転領域が解消の局面になってから、実際に解消しきるまでの時間に対応する。 t_f から t_r までの期間、輻転領域は形をほぼ保ちながら移動している。 t_w (持続時間) は、輻転が発生してから消滅するまでの寿命に相当すると考えれば良い。

* 最大値ではない

図7から、 t_f がほぼ $O(L)$ で、他の測定は $O(L^2)$ となっていることがわかる。結合網の規模が大きいくほど (L が大きいほど) 持続時間 (t_w) や解消速度 (t_r) が $O(L^2)$ で長くなるが、一方、輻転の成長速度 (t_f) は $O(L)$ である。

- 輻転の持続時間は網の規模 (L^2) に比例したスケールな関係になっているが、
- 成長速度は網の規模の1/2乗で効いてくる。つまり大規模システムほど、相対的に輻転の進行速度が速くなる。

と言えよう。また、上記とは逆の観点から、大規模システムでは一旦輻転状態が発生すると解消するまでに長時間要する、ということもできる。

4. 発火条件

セルオートマトンモデルでは、系内のパケット密度が一定のレベル以上になると、局所的な輻転が解消しきれずに周囲のパケットを吸収しながら成長する現象が観測される。前節では、この閾値を系内のパケット総数として求めた。しかし、図2に示されているように、輻転の発生は間欠的であることから、実際に輻転が発生するか否かは、パケット密度の平均値 (パケット総数) ではなく、局所的な密度 (の最大値) で考えるべきである。

一辺 ω (ノード) の微小領域と、その中にあるパケットの密度を考える。系内の任意の微小領域についてパケット密度を調べ、それらの最大値を局所最大密度と呼ぶ。

局所最大密度がある一定の閾値を超えた場合に、その領域の周囲のパケットを吸収して成長する輻転が発生するものと仮定する。つまり、局所最大密度が一定以下であれば大規模に成長する輻転が生じないが、逆に、一定値以上になった場合は大規模に成長する境界条件があるものと仮定する。

この条件を発火条件と呼ぶことにする。

輻転がない状態では、エントロピーは高く安定している^{**}。この非輻転状態でも、パケット転送のゆらぎにより局所最大密度は変動する。輻転が局所的・小規模なものですぐに解消する場合は、エントロピーは大きな影響を受けない。この場合、局所最大密度は一時的に大きくなるものの、輻転状態の解消とともに元の低い値に戻る。しかし、輻転の種が解消せずに成長を始めると、局所最大密度はさらに大きくなり、同時にエントロピーの低下を招く。

そこで、以下のようにして輻転を生じない最大の局所最大密度を求める。上述のように、システムが非輻転状態にあるときでも局所最大密度はある程度の変動が観測される。局所最大密度が大きくなると局所的な

** システムのサイズによる。規模が小さいほどばらつきが大きくなる。

輻転が発生したと推測できるが、ある時点でピークとなりその後元の基底値に戻った場合は、局所的な輻転がそれ以上成長せずに解消したと判断できる。このため、十分に長い時間、非輻転状態にあるときの局所最大密度を観測し、とり得る最大値を求めれば、その値が局所最大密度の閾値と判断できる。

ただし、実際のエントロピーや局所最大密度がなかなか変化するとに限らず、高周波成分（ノイズ）を含む。図8をもとにその方法を説明する。まず、シミュレーションにより各測定値の時系列変化を記録する。この結果をもとに、エントロピーの天井値（ceiling value, e_c ）および天井値の98%値（98% e_c , e'_c ）、最大密度の基底値（baseline value, d_b ）を求めると。

エントロピー値が e'_c を超えたとき（図8中(a)）に非輻転状態が開始したと判断する。ただしこの時点では、直前の輻転状態の原因だった輻転領域が完全に解消していない可能性が高いため、さらにエントロピー値が天井値 e_c になるまで待つ（図8中(b)）。このとき、直前の輻転は解消されていると推測できるが、さらに局所最大密度が基底値 d_b に達するまで待つ（図8中(c)）。図8中の(b), (c)の時間は、実際には前後する可能性がある。図8中(c)以降の局所最大密度を閾値測定の対象とする。

上述の方法で測定開始点を決められるが、終了時点の決定、つまりどのタイミングまで局所最大密度を測定すればよいのか、が問題となる。輻転の発生は、エントロピー値の低下により検出するしかない。エントロピーが天井値の98%値 e'_c を下回った時点（図8中(d)）で系が輻転状態になったと判断する。この時点では、局所最大密度は既に閾値を超えているため、図8中(d)の時点から遡って閾値測定の終了時間を特定する必要がある。

輻転が発生する直前、局所最大密度は乱高下を伴いながら全体的には急上昇する（図8にも示されている）。この乱高下のなかで閾値を求めることは不可能であるから、測定対象から除外する。このため、輻転状態を検出した時点（図8中(d)）以前で、局所最大密度が基底値になった最後の時点（図8中(e)）を測定終了時間とする。こうして求めた測定期間(c)-(e)で最大の局所最大密度を求めれば良い。

シミュレーション時間を十分に長くすれば、図8のような輻転状態・非輻転状態は何度も繰り返す。シミュレーション時間全体で最大の局所最大密度を求めれば、その値が本節で求めようとしている閾値（にきわめて近いもの）となる。

こうして求められた発火閾値の前後のエントロピーおよび局所最大密度の様子を図9に示す。局所最大密度は $w=2$ および $w=3$ の場合のみ示している。各 L, N で求められた局所最大密度の発火閾値を、 w ごとに表2に示す。

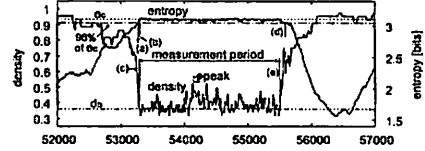


図8 最大パケット密度の測定

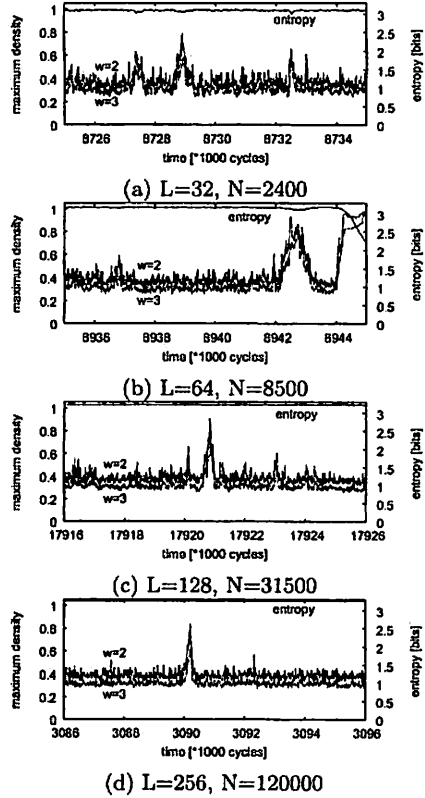


図9 発火閾値前後でのエントロピー、局所最大密度の変化の様子

表2 各システムサイズ (L) に対する局所最大密度の発火閾値

L	N	局所最大密度の発火閾値			
		w=2	w=3	w=4	w=5
16	660	0.6364	0.4848	0.3977	0.3491
32	2400	0.7955	0.5859	0.4773	0.4145
64	8500	0.9318	0.8182	0.6989	0.5855
128	31500	0.9091	0.7071	0.5568	0.4582
256	120000	0.8409	0.6162	0.4716	0.3855

5. 相互結合網での間欠的な輻転

前節までに得られた知見が、実際の相互結合網でどの程度適用できるのか、評価しておく必要がある。さまざまな規模の相互結合網において、臨界転送負荷付

近での系の挙動を、エントロピーの時系列変化として測定した^{*}。その結果を図 10 に示す。各グラフ中で、 v は各ノードからのパケット生成間隔 (単位: サイクル) を示す。この値が小さいほど投入転送負荷が高くなる。

投入転送負荷に対してエントロピー値は全体的にはなだらかに増減することがわかる。各グラフの中段のプロットでは、やや大きな下向きのピークが観測されている。これらは、さらに数段下方にあるプロットまで突き抜けており、これらのピークのときに一時的な (しかし系全体にわたる) 輻輳が発生していることがわかる。すなわち、セルオートマトンで観測されたのと同様に、間欠的な輻輳が生じている。さらに、図 10 に示されている各ピークの深さ、幅などを比較すると、結合網の規模が大きいほど顕著に現れることがわかる^{**}。これ以上の詳細な検討は、今後の課題である。

6. おわりに

本稿では、二次元トラス網を簡略化して得られるセルオートマトン・モデルを使用することで、結合網の性能が飽和する臨界付近の投入負荷で発生する間欠的な輻輳のメカニズムと諸性質を明らかにした。平均場近似の手法を用いることで簡略化した形で解析的な議論を行うことができる。またセルオートマトンの動作 (シミュレーション) から、輻輳状態を観測することができる。これらの結果から、輻輳が発生する臨界での系内のパケット総数を求めた。システムの一辺のノード数 L に対し、臨界パケット総数は解析値では $O(L^{\frac{1}{2}})$ 、シミュレーションでは $O(L^2)$ の関係にあることがわかった。また、システムが大規模化するほど、輻輳の成長速度に対して、持続する時間が長くなることが明らかになった。成長速度は $O(L)$ であるが、持続時間や解消時間は $O(L^2)$ である。

さらに、セルオートマトンで観測された間欠的な輻輳の挙動は、相互結合網シミュレータでも同様に観測されることを示した。詳細な解析は今後に譲るが、セルオートマトンで得られた知見の多くが適用できる期待がある。

謝辞 本研究は、一部日本学術振興会科学研究費補助金 (基盤研究 (B)18300014, 同 (C)16500023, 若手研究 (B)17700047) の援助による。

参考文献

- 1) T. Sakakibara, Y. Honda and T. Horiguchi, "Effect of obstacles on formation of traffic jam

^{*} 32×32 二次元トラス網。次元順ルーティング、パケット長 8 フリット、パケットバッファ容量 16 フリット、virtual cut-through ルーティング、virtual channel 数 3

^{**} 図中の各グラフでは時間軸のスケールが異なっていることに注意されたい。

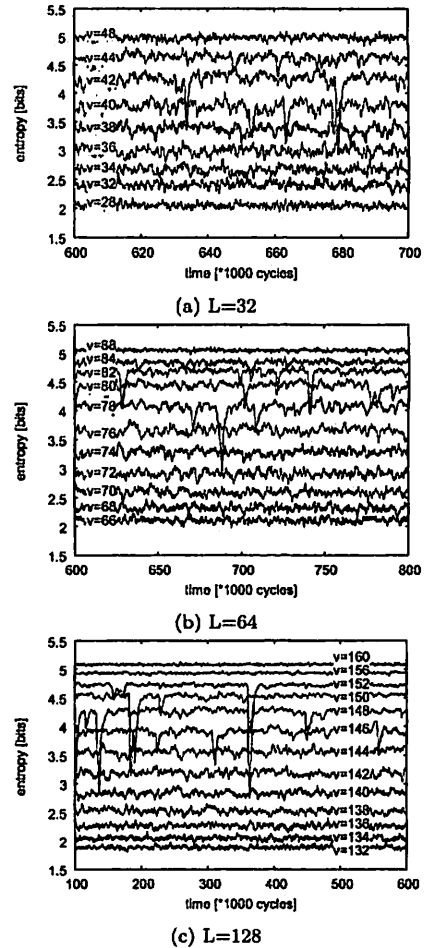


図 10 エントロピーの時系列変化の様子

in a two-dimensional traffic network," *Physica A*, Vol.276, pp.316-337, Elsevier, 2000.

- 2) 横田 隆史, 大津 金光, 古川 文人, 馬場 敬信, "セルオートマトンによる大規模相互結合網シミュレーションの試み", 信学技報, CPSY2005-32, 2005 年.
- 3) 横田 隆史, 大津 金光, 古川 文人, 馬場 敬信, "セルオートマトンによる相互結合網の輻輳の解析", ACS-14, pp.21-42, 2006 年.
- 4) 横田 隆史, 大津 金光, 古川 文人, 馬場 敬信, "エントロピー・スロットリング: 相互結合網のパケット移動度に着目した輻輳制御手法", ACS-15 (掲載予定).