

MADEによるパンデミックシミュレーション: マルチエージェントと微分方程式モデルの組み合わせ

豊坂祐樹[†] 廣瀬英雄^{††}

パンデミックシミュレーションの手法には主にMASとSEIRモデルの2つが知られている。MASは社会ネットワークモデルの標準的なモデルである。シミュレーションにおいて境界条件や初期条件など詳細な設定が可能ではあるが、計算コストが高い。一方、SEIRモデルは使用するパラメータも少なく全ての人間を均質に扱う微分方程式モデルであり、計算コストは低い。しかし、詳細な設定が困難なため、使用する際の汎用性に乏しい。そこで、本研究では2つのモデルを組み合わせた新しいモデル作成を提案する。シミュレーションの初期段階ではMASを使用し、その段階で得られた計算結果からSEIRモデルのパラメータを求め、それ以降をSEIRモデルのシミュレーションに切り替える手法である。この手法により、計算コストが低く、実際の社会ネットワークに似せたパンデミックシミュレーションを行うことが可能になる。

Pandemic simulations by MADE: a combination of multi-agent and differential equations

Yuki Toyosaka[†] and Hideo Hirose^{††}

Two pandemic simulation approaches are known: the MAS and the SEIR. The MAS can deal with detailed simulations under a variety of initial and boundary conditions with standard social network models; however, the computing cost is high. The SEIR can quickly deal with simulations for homogeneous populations with simultaneous ordinary differential equations and a few parameters; however, it lacks versatility in its use. We propose a new method combining these two methods, such that we use the MAS in the early stage in a simulation to determine the parameters that can be used in the SEIR, and then use the SEIR in the subsequent stage. With this method, we may deal with pandemic simulations for real social structures with lower computing costs.

1. はじめに

人間社会において人が未知の伝染病に感染した場合、マルチエージェントシミュレーション(MAS) [2][4][5]等の手法を用いて、感染拡大規模(感染の範囲、感染者の数等)の推定を行う方法がある[1]。MASとは標準的な人間社会のネットワークをモデル化したものであり、境界条件や初期条件などの詳細な設定が可能なシミュレーションではあるが、計算コストはシミュレーションで設定された総人口が増加するに従って指数関数的に高くなる欠点をもつ。もう1つのパンデミックシミュレーションの手法として、SIRまたはSEIRモデルと呼ばれる常微分方程式モデルを用いる方法がある。SEIRモデルはSIRモデルを拡張したモデルである。このモデルは閉鎖された空間内の集団における伝染性感染症の感染者数推定を行うことができる。SEIRモデルの S 、 E 、 I 、 R はそれぞれ susceptible (感染する可能性のある者)、exposed (感染してまだ他人に病気を感染させない潜伏期間中者)、infected (発症し他人に病気を感染させる者)、removed (回復又は死亡の除外者)の遷移状態を表し、微分方程式といくつかのパラメータを用いて全ての人間が均質な集団での感染者数推定を行うモデルである。このモデルは計算コストが低いという利点と詳細な状況設定が困難という欠点が存在する。

本論文では、前述の2つのモデルを組み合わせた新しいモデルMADE (multi-agent and differential equations)を提案する。このモデルはSEIRモデルに使用する近似パラメータを決定するために、初期段階にMASによるパンデミックシミュレーションを行い、それ以降をSEIRモデルに切り替えてシミュレーションする手法である。SEIRモデルに使用するパラメータは、SEIRモデルの微分方程式を差分形に変形したパラメータの算出式にMAS計算結果の数値(日毎の S 、 E 、 I 、 R)を代入して求められる。なお、SEIRモデルの解は4次のルンゲクッタ法によって求められる。提案するMADEモデルを使用するためには、モデルの妥当性を示す必要がある。そこでMASの計算結果から得られたパラメータを用いて計算したSEIRモデルの計算結果と、元となるMASの計算結果を比較することで整合性を検証する。そのことによって初めて計算コストがより低く現実世界に近い設定のシミュレーション(MADEモデル)が可能となる。MADEは不完全データによる統計的パラメータの推定結果と違って、様々な状況に対応可能な汎用性がある。また、パンデミックの早い段階にMASによって得られたパラメータであったとしても、MADEの最終除外者数の計算結果はMASのみでシミュレ-

[†] 情報処理学会
Information Processing Society of Japan

^{††} 情報処理学会
Information Processing Society of Japan

シミュレーションした計算結果に近い値となる。

2. MADE に使用する 2つのモデルについて

ここでは MADE に使用する MAS の設定と SEIR モデルについて記述する。

2.1 MAS (multi-agent simulation)

MAS はコンピュータ上に仮想の都市を設定し、都市内の住人を離散時間ごとに動作させることによって現実の社会に近い人間のネットワークを構築する個人ベースのモデルのことを指す。対象となる感染症が同じ感染空間内にいる感染者から非感染者へと病気が伝搬していくと仮定する。その場合、感染蔓延は人口の規模(人口密度等)や人々の接触する回数などに大きく影響を受ける。本研究では、擬似的な人間のネットワークを構築する MAS を使い、人間のネットワークを介して伝搬する感染症蔓延の予測モデルを作成した。以下に、本研究で作成した MAS における設定を記述する。現実の町を模擬した小さな仮想空間を構築するにあたって、図1のように町の建物、人間の動作等を仮定した。まず、町の住人には、社会人、主婦、生徒の3種類の人間が存在し、夜は全員自宅に帰宅して、就寝する。朝の活動時には徒歩、車、電車の移動手段を用い、会社、スーパー、学校へ行く。各人間の動作時間や立ち振る舞いの設定に関しては前回の研究[13][14]と同じとする。また、図1の建物や交通機関のうち、楕円形の線で囲まれた場所が人から人への感染伝搬が発生する感染エリアであり、そこでの離散時間(単位:分)ごとの感染率を β とする。

本研究では、MAS と MADE 間の整合性や相関関係を焦点としているため、人口は 10000 人(社会人 7000 人、主婦 1500 人、生徒 1500 人)に固定する。MAS のシミュレーションは計 18 通り(初期感染者数 100, 10, 1 人、感染率 0.095, 0.09, 0.085, 0.08, 0.075, 0.07%)でそれぞれ 10 回ずつ計算を行う。

2.2 SEIR モデル

SEIR モデルは(1)式の常微分方程式を用いて計算を行う。 S, E, I, R はそれぞれ susceptible, exposed, infected, removed, パラメータ λ, σ, γ は感染率, 状態遷移率, 除外率を表している。人間の集団は感染率 λ の割合で susceptible から exposed へ, 状態遷移率 σ の割合で exposed から infected へ, 除外率 γ の割合で infected から removed へとそれぞれ毎時間ごとに移行する。

$$\begin{aligned}
 S'(t) &= -\lambda S(t)I(t) \\
 E'(t) &= \lambda S(t)I(t) - \sigma E(t) \\
 I'(t) &= \sigma E(t) - \gamma I(t) \\
 R'(t) &= \gamma I(t)
 \end{aligned}
 \tag{1}$$

また、一旦 removed になった人間は susceptible になることはなく、二度と感染しないと仮定する。(1)式の微分方程式の解法には 4 次のルンゲクッタ法を用いる。

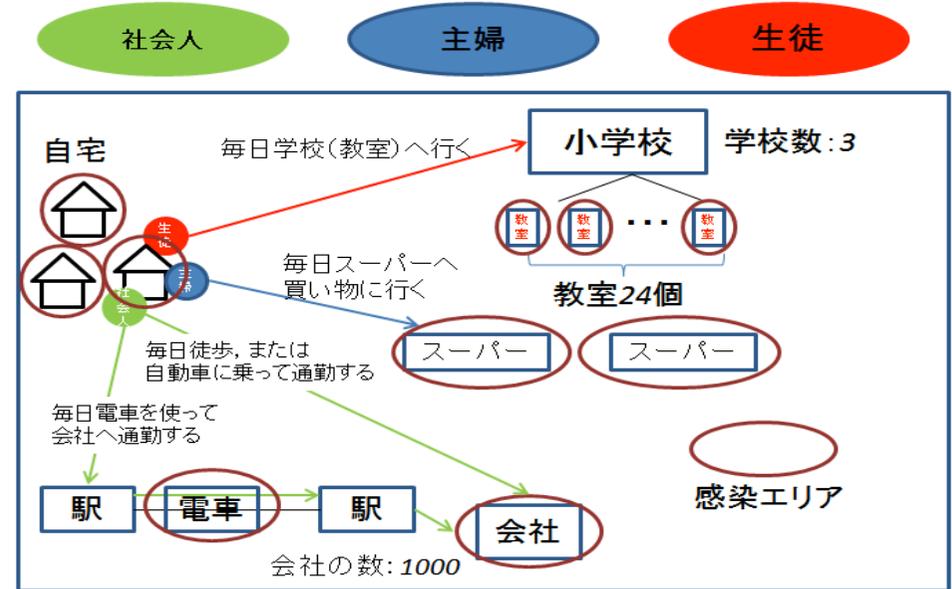


図1 MAS の概要 (仮想都市の設定, 各人間の動作設定)

3. SEIR モデルに使用するパラメータの算出法

MAS のシミュレーション結果から、1 日ごとの S, E, I, R の人数を得ることができる。SEIR モデルに使用する時間 t (日) におけるパラメータ $\lambda(t), \sigma(t), \gamma(t)$ は (2) 式によって近似的に求めることができる。

$$\begin{aligned}
 \lambda(t) &= \frac{S(t) - S(t+1)}{S(t)I(t)}, \\
 \sigma(t) &= \frac{\{E(t) - E(t+1)\} + \{S(t) - S(t+1)\}}{E(t)}, \\
 \gamma(t) &= \frac{R(t+1) - R(t)}{I(t)}.
 \end{aligned}
 \tag{2}$$

シミュレーションに使用するパラメータ λ , σ , γ の値は比較的安定した期間の平均値を使用する。

4. MASの結果から得られたパラメータを使用したSEIRモデルの計算

MASのシミュレーション結果から算出されたパラメータの値を使用して、SEIRモデルのシミュレーションを行う。図2はMASとSEIRモデルでの、日ごとの exposed, infected, removed の人数の一例を示した図である。初期感染者数 n_0 が100人、感染率 β が0.08%の条件で計算した結果であるが、2つのモデルの E , I , R 成長曲線はかなり接近しているのが分かる。しかし、これでは2つのモデルの数量的な違いをみる事ができない。そこでMASとSEIRモデルの誤差を調べるために誤差率 $e(t)$ を(3)式として定義する。

$$e(t) = \frac{R(t)_{SEIR} - R(t)_{MAS}}{R(t_{final})_{MAS}} \quad (3)$$

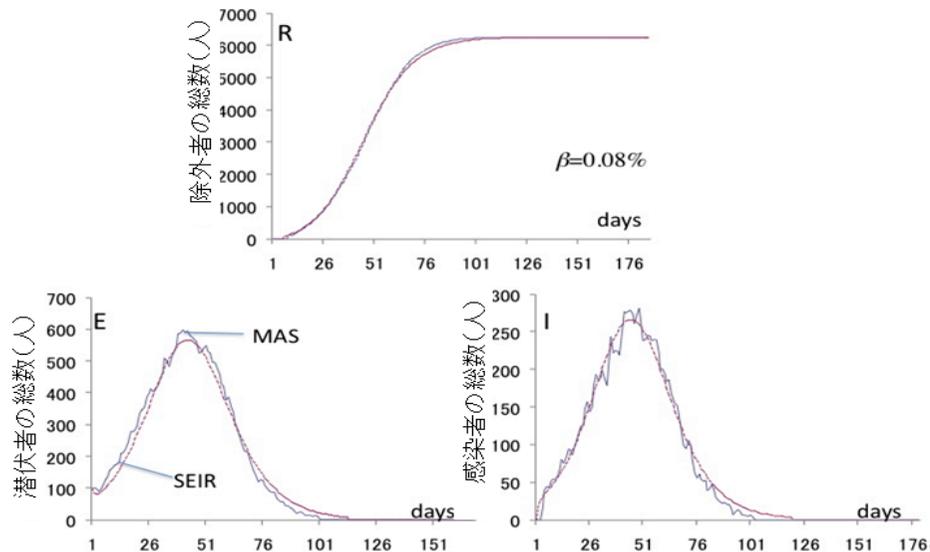


図2 MASとSEIRモデルの E , I , R 計算結果比較

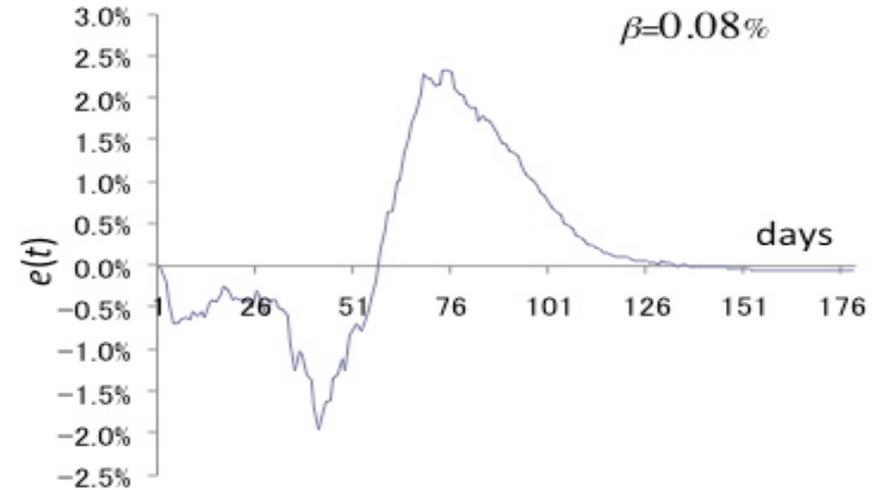


図3 除外者誤差率 $e(t)$ の推移

図3はシミュレーションによって得られた2つのモデルの日ごとの除外者数の数に対して(3)式を適用して算出した1日ごとの誤差率 $e(t)$ を示している。また、最大誤差率 e_{∞} を

$$e_{\infty} = \sup_t |e(t)| \quad (4)$$

と定義する。次に各10回ずつシミュレーションを行った計算結果の誤差率平均値と標準偏差を表1に示す。表中の n_0 は初期感染者数である。表1を見ると各条件において、最大誤差率 e_{∞} は非常に小さい値となっていることがわかる。

この結果はSEIRモデルとMAS間の計算結果に整合性があることを示している。SEIRモデルはMASからパラメータを求める情報が得られれば、MASの代用として使用することが可能である。従って、これらの2つのモデルを組み合わせる手法を検討することも可能となる。

表 1 各条件における最大誤差率 e_{∞} (上の数値は平均値, 下の数値は標準偏差)

n_0	β (%)					
	0.095	0.09	0.085	0.08	0.075	0.07
100	1.61	1.62	1.72	1.85	2.11	2.31
	0.33	0.35	0.39	0.53	0.47	0.51
10	2.25	2.32	2.76	2.29	2.89	3.29
	1.05	0.87	1.07	0.64	1.46	1.14
1	2.83	3.57	2.92	3.44	3.74	3.92
	0.92	1.60	1.19	1.23	1.65	1.41

5. MADE の計算

MADE はパンデミックシミュレーションの際に, SEIR モデルに使用するパラメータ値を決定するために初期段階では MAS を使用し, その後の推定を SEIR モデルに切り替えてシミュレーションを行うという 2 つのモデルを組み合わせた手法である. MAS の初期段階のデータを用いて得られたパラメータを使用し, それ以降のシミュレーションでは SEIR モデルを使用する. SEIR モデルに使用するパラメータの算出方法は前述のとおり実行される. ここで, MAS から SEIR モデルに切り替わる時間 (日) τ を定義する. 今回 τ を 4 つ ($\tau=1/4$, $\tau=1/3$, $\tau=1/2$, $\tau=2/3$) 用意した. ($\tau=1$ は感染蔓延の終息日とする)

図 4 は MADE を使って推定した結果の一例 (初期感染者 100 人感染率 $\beta=0.095\%$, $\tau=1/4$, $1/3$, $1/2$, $2/3$ の 1 ケースの除外者数比較) を示したものである. 図 4 では MAS と MADE の計算結果の差が小さいことがわかる. 誤差の最大値 e_{∞} も 3% 程度に収まっているので, この程度の小さい誤差ならば MADE モデルは十分に使えると判断できる. 表 2 は各シミュレーション $n_0=100, 10, 1$ $\beta=0.095, 0.09, 0.085, 0.08, 0.075, 0.07$ $\tau=1/3, 1/2, 2/3$ の計 54 パタンのシミュレーションを 10 回ずつ行った結果の最大誤差値 e_{∞} の平均値と標準偏差を表している.

6. MADE の精度について

SARS のような不完全データの統計解析では, 初期段階でデータが少ない場合, パラメータの算出や推定は難しい. [6]や[7]によると, 接続時間が最終的に除外者数の安定した段階 (日) の半分より前の段階だと最尤パラメータを得ることは困

難である. [6]の図 2 では, 2003 年 5 月 23 日 (この日は回復者と死亡者の発生が高くなるピーク時を過ぎ, また感染者数もほとんど最終の安定期に入った時期に当たる) 以前はパラメータが安定して得られないことがわかる. これは truncated model または truncated model のような統計的な推定手法の場合, データの少ない初期の時間に対してとても敏感であることを示している.

これに対して, MADE はたとえ $\tau=1/3$ の早い段階だとしても, 除外者数に関してある程度正確な値を供給することができる. なぜそうなるのか原因は解明できていない. 考えられる理由の 1 つとしては方程式の数 ($S(t)$, $E(t)$, $I(t)$, $R(t)$) に比べて不明なパラメータ数が 3 つのみで少ないことが考えられる. [6]の統計モデルの場合, 3 種類の観測値 (infected, cured, dead persons) とパラメータの総数が 6 つも存在する. もう 1 つの理由として SARS のケースとは違い, MADE モデルにおける発生初期段階を知っているからかもしれない. ここで最終的な除外者数 $R(t_{\text{final}})_{MS}$ と $R(t_{\text{final}})_{MAS}$ の誤差率 r を (5) 式として定義する.

$$r = \frac{R(t_{\text{final}})_{MS} - R(t_{\text{final}})_{MAS}}{R(t_{\text{final}})_{MAS}} \quad (5)$$

表 3 は (5) 式を用いて算出した各シミュレーションの r の結果を示しているが, τ が $1/3$ の早い段階でも初期感染者数 100 人の場合は r の値は小さく, 総人口の 3~5% 程度の誤差であった.

7. 計算コストについて

2 つのモデルの計算コストについて検証する. 総人口が 1000 人, 初期感染者 10 人感染率 0.095% で最終的に感染蔓延が終息するのが 200 日程度のケースでは, MAS の計算コストは 9 秒ほどである. それに対し, SEIR モデルでの計算コストは同じ人口 1000 人のケースでも約 0.031 秒と非常に小さい. MAS に対する SEIR モデルの計算コストの減少率は約 3.4×10^{-3} となる. 次に総人口が 10000 人の場合, 初期感染者 $n_0=10$ 人, 感染率 $\beta=0.095\%$ で最終的な感染蔓延終息が 200 日程度と仮定して計算すると, その計算コストは前回 (人口 1000 人) の 90 倍近く 804 秒まで大きくなる. それとは対照的に SEIR モデルの計算コストは 0.031 秒前後のまま変わらない. その計算コストの減少率は 3.9×10^{-5} である. ここでは 3GHZ Pentium(R) 4CPUs with 1.5GB RAM を使用した. SEIR モデルの計算コストはネットワークの関係で $N(N-1)/2$ に比例して計算コストが高くなる MAS とは違い, 総人口 N に影響を受けない. 福岡市のような 100 万人以上の人口規模を誇る主要な

大都市のケースに取り組む際には、MAS の計算コストは N^2 に比例して増加するため、MADE は計算コストを減少させるという点で重要な意味をもつ。

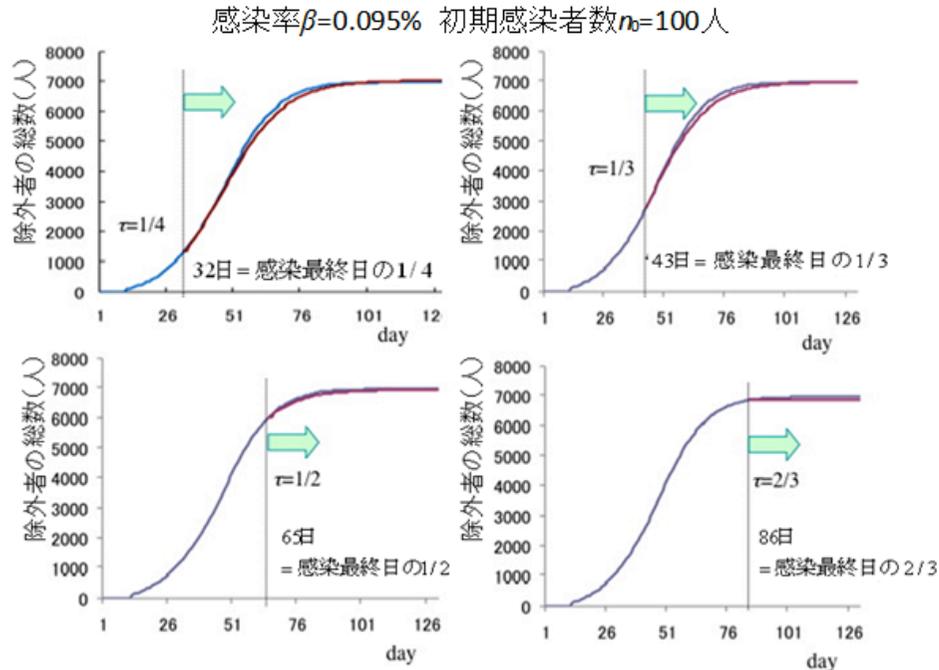


図 4 各 τ における MADE 推定結果と MAS 計算結果 (一例) の比較

8. おわりに

都市の中で何人かが感染病に感染したと仮定し、都市内で発生する病気の伝搬を MAS で表現すると、感染病蔓延の結果 (感染者数などの推定値) を得ることができる。またその結果に基づき、微分方程式に使用する SEIR モデルのパラメータを得ることができる。そのことで SEIR モデルの推定計算を行い、元となる MAS のデータと比較できる。検討の結果、SEIR モデルによる計算の結果は MAS の計算結果に近いということがわかった。そのことは 2 つのモデルの間に整合性が存在するかどうか判断するための 1 つの指標となり得る。次の段階として、パ

ンデミックの最終段階を推定するために MAS と SEIR モデルを組み合わせた MADE モデルを実際に使用し、元となる MAS の計算結果との比較も行った。

MADE は近似パラメータを決定するためにパンデミックシミュレーションの初期段階で MAS を活用し、初期段階で得られたパラメータをその後の SEIR モデルのシミュレーションに適用することによって、より少ない計算コストと小さい計算誤差の除外者人口やその他のデータを得ることが可能となる。

MADE は不完全データを条件とする統計的パラメータ推定結果とは違い、様々な状況下で推定が可能で汎用性が高い。そのようなモデルによる計算結果のおかげで、新型の感染症の発生やバイオテロなどの危険なパンデミックの脅威に対して備えることができる。

表 2 τ ごとによる最終値の誤差率 r の平均値と標準偏差
 (上の数値が平均値, 下の数値が標準偏差, シミュレーション回数: 各 10 回)

τ	n_0	β (%)					
		0.095	0.09	0.085	0.08	0.075	0.07
1/3	100	2.642	3.845	4.728	3.757	3.428	5.144
	10	1.909	1.939	2.901	2.437	3.512	4.687
	1	6.655	8.402	8.177	6.526	9.814	17.34
1/2	100	3.693	4.502	5.642	4.765	7.617	12.23
	10	3.506	6.127	8.127	8.112	13.26	14.63
	1	9.515	9.813	10.08	11.46	13.27	23.23
2/3	100	0.304	1.089	1.059	0.333	0.589	1.557
	10	1.088	0.447	1.12	1.316	0.971	2.655
	1	0.785	2.13	2.637	1.318	2.608	8.209
1	100	1.799	2.393	2.673	1.813	3.577	6.127
	10	0.415	3.201	2.592	2.801	8.828	12.52
	1	4.714	4.137	4.502	4.501	7.51	9.431
1	100	0.082	0.008	-0.005	0.272	0.152	0.205
	10	0.227	0.186	0.299	0.293	0.537	0.614
	1	0.031	0.015	0.199	0.017	0.741	3.446
1	100	0.337	0.25	0.841	0.659	1.602	4.55
	10	-0.217	0.236	0.414	-0.097	1.814	1.229
	1	0.867	1.267	1.249	0.946	3.487	4.926

expressed by percentage, upper: mean, lower: s.d.

謝辞 本論文の作成にあたって、ご意見をいただき、またご協力頂いた藤尾准教授に感謝の意を表します。

参考文献

- 1) R. Anderson and R. May, Infectious diseases of humans: Dynamics and control, Oxford University Press, 1991.
- 2) C. L. Barrett, S. G. Eubank and J. P. Smith, If smallpox strikes Portland, Scientific American, vol. 292, pp. 54-61, 2005.
- 3) F. Brauer, P. van den Driessche and J. Wu (ed.) Mathematical Epidemiology, Lecture Notes in Mathematics, Springer, 2008.
- 4) L. R. Elveback, J. P. Fox, E. Ackerman, et al. An influenza simulation model for immunization studies, American Journal of Epidemiology. vol. 103, pp. 152-65, 1976.
- 5) S. Eubank, Scalable, efficient epidemiological simulation, Proceedings of the 2002 ACM symposium on Applied computing, pp. 139-145, 2002.
- 6) H. Hirose, The mixed transored model with applications to SARS, Mathematics and Computers in Simulation, vol. 74, pp. 443-453, 2007.
- 7) H. Hirose, Estimation for the size of fragile population in the transored and truncated models with application to the confidence interval for the case fatality ratio of SARS, Information, vol. 12, pp. 33-50, 2009.
- 8) C.Y. Huang, C.T. Sun, J.L. Hsieh, Y.M.A. Chen and H.L. Lin, A novel small-world model: Using social mirror identities for epidemic simulations, Simulation, vol. 81, pp. 671-699, 2005.
- 9) W.O. Kermack and A.G. McKendrick, Contributions to the mathematical theory of epidemics-III. Further studies of the problem of endemicity, Proceedings of the Royal Society, vol. 141A, pp. 94-122. 1933.
- 10) I.M. Longini, Jr., M.E. Halloran, A. Nizam and Y. Yang, Containing pandemic influenza with antiviral agents, American Journal of Epidemiology, vol. 159, pp. 623-633, 2004.
- 11) M.I. Metzger, N.J. Cox and K. Fukuda, The economic impact of pandemic influenza in the united states: priorities for intervention, Emerging Infectious diseases, vol. 5, pp. 659-671, 1999.
- 12) S. Milgram, The small world problem, Psychology Today, vol. 2, pp. 60-67, 1967.
- 13) Y. Toyosaka and H. Hirose, The consistency between the two kinds of pandemic simulations of the SEIR model and the MAS model, 9th International Conference on Computers, Communications and Systems (ICCCS 2008), Nov. 7, 2008.
- 14) Y. Toyosaka and H. Hirose, The consistency of the pandemic simulations between the SEIR model and the MAS model, IEICE Transactions on Fundamentals, Vol. E92-A, No. 7, to appear.