Information Distribution Analysis based on Human's Behavior State Model and the Small-World Network

2011 Workshop on Social Computing and Cultural Modeling

PAO SRIPRASERTSUK
Assistant Professor@GITS, Waseda University

Research Background

Secondary information distribution has been playing an important role in our society

↓

Our previous study does not consider the human network structure but only the statistical aspects of information distribution being consumed

↓

Human network structure can be represented by small-world network

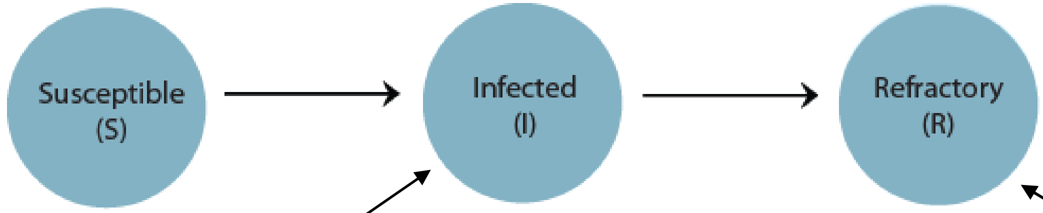
↓

Using the SIR model to analyze the effectiveness of information distribution on the small-world network

↓

Information distribution model and equations considering the human behavior in information distribution are proposed

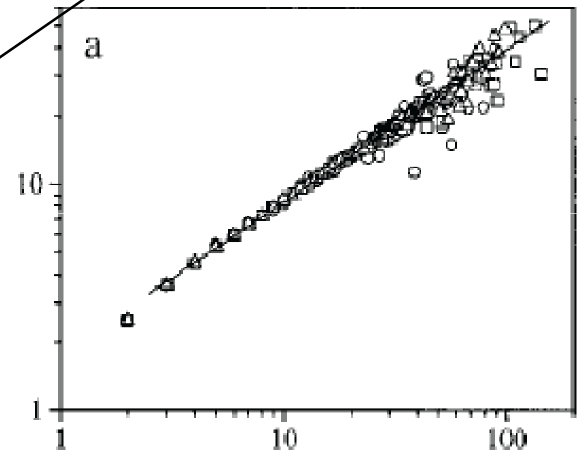
Problems of using SIR Model



Individuals who have not heard the information

Individuals who are spreading the information

Individuals who know the information but no longer spreading it



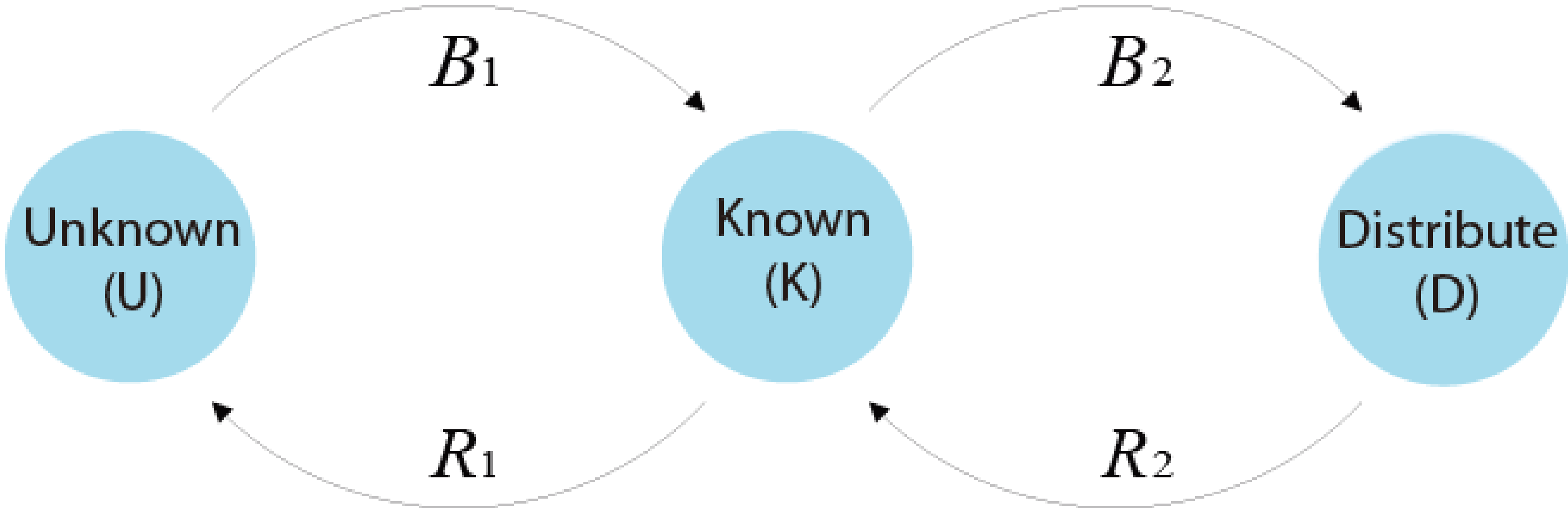
Time

Number of individuals who knows the information

- The SIR concerns, but not fully the human behavior of information distribution
- All of information propagation results reach almost all individuals when time passes
- The primary information distribution is also done during the secondary information distribution

[1] Above Figure is captured from *H. Liang, P. Kwangho and L. Ying-Chang: Information propagation on modular networks, Physical Review E, vol. 73, 035103(R), (2006).*

The Proposed Model *(UKD)*

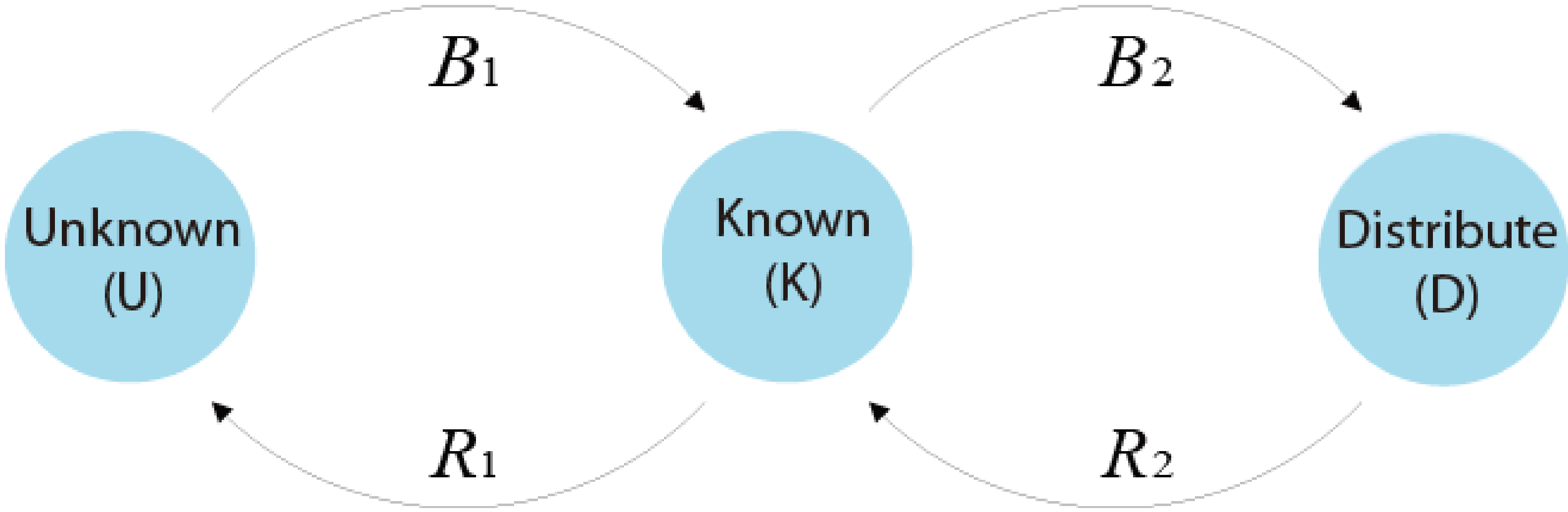


$U =$ Individuals in this state do not know information. They either do not receive information yet or they forget it.

$K =$ In this state, individuals know information but do not have any action to the information distribution

$D =$ Individuals in this state are active to distribute information. The distribution by their own intentions and other individuals requests are considered

The Proposed Model *(UKD)*



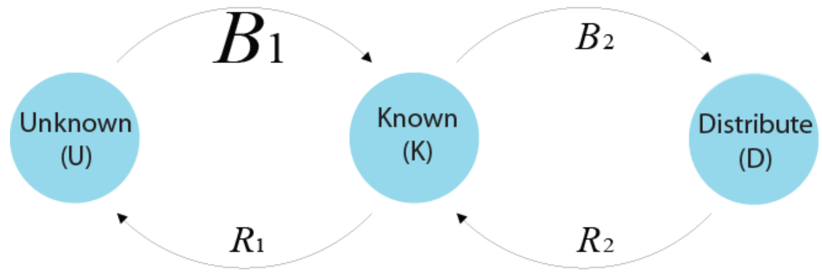
$B_1 =$ Probability of individuals in Unknown State to change into Known State.

$R_1 =$ Probability of individuals in Known State to change into Unknown State.

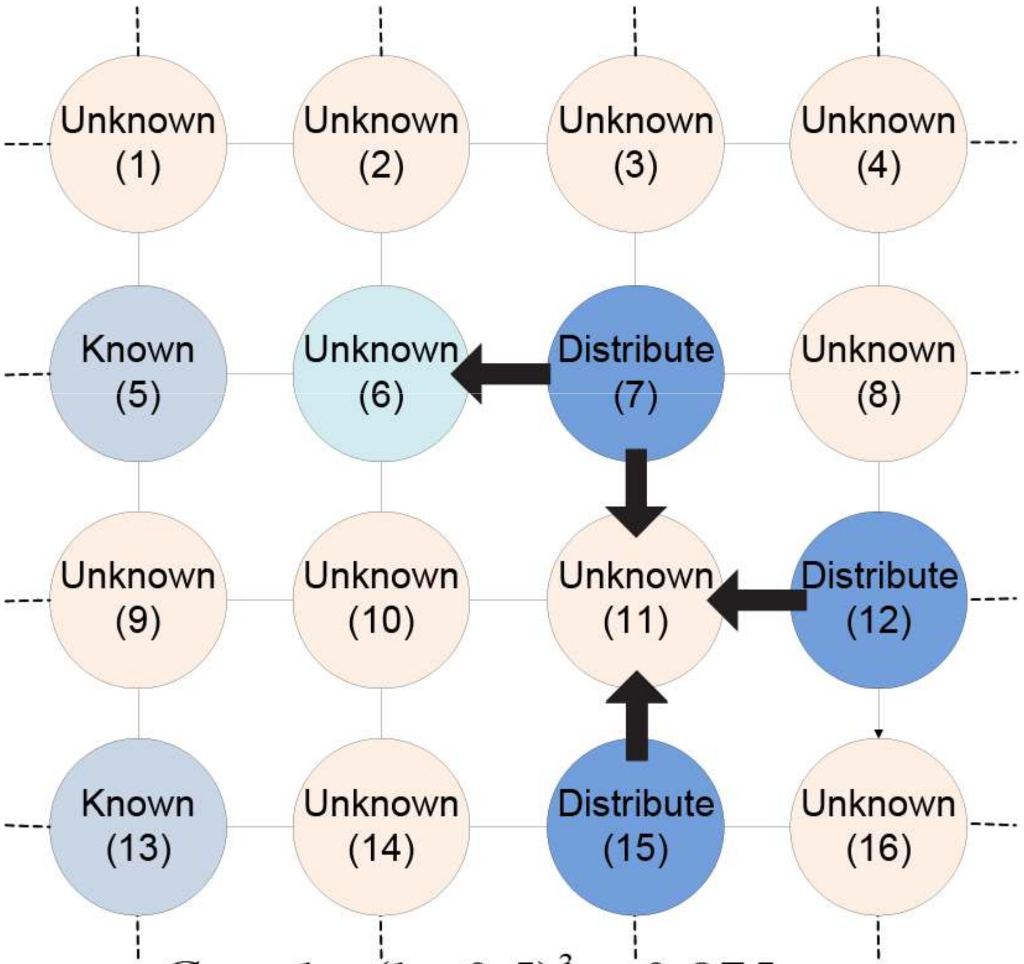
$B_2 =$ Probability of individuals in Known State to change into Distribute State.

$R_2 =$ Probability of individuals in Distribute State to change into Known State.

The Process of Becoming Known State



$$G_6 = 1 - (1 - 0.5)^1 = 0.5$$

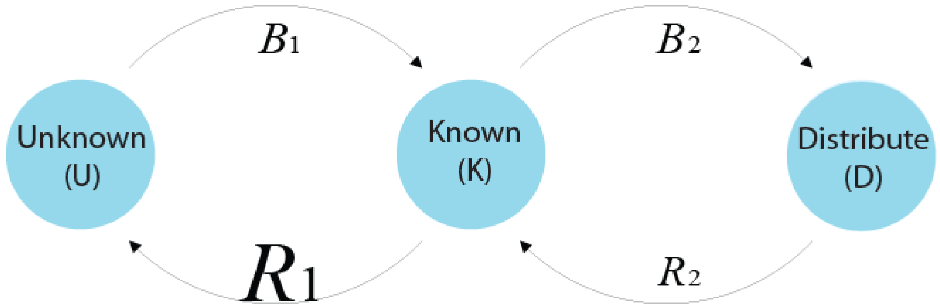


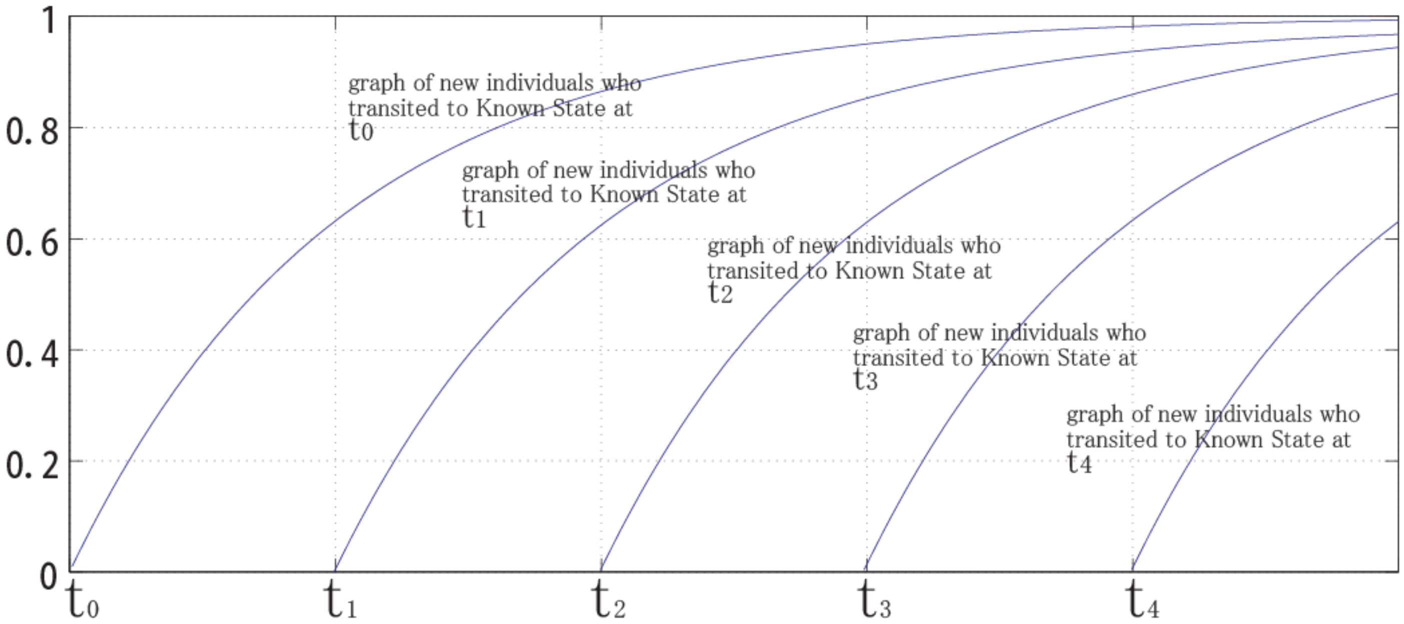
$$G_{11} = 1 - (1 - 0.5)^3 = 0.875$$

$$G_i = 1 - (1 - B_1)^n$$

n = the number of neighbors of node ***i*** in Distribute State

Dynamism in R_1



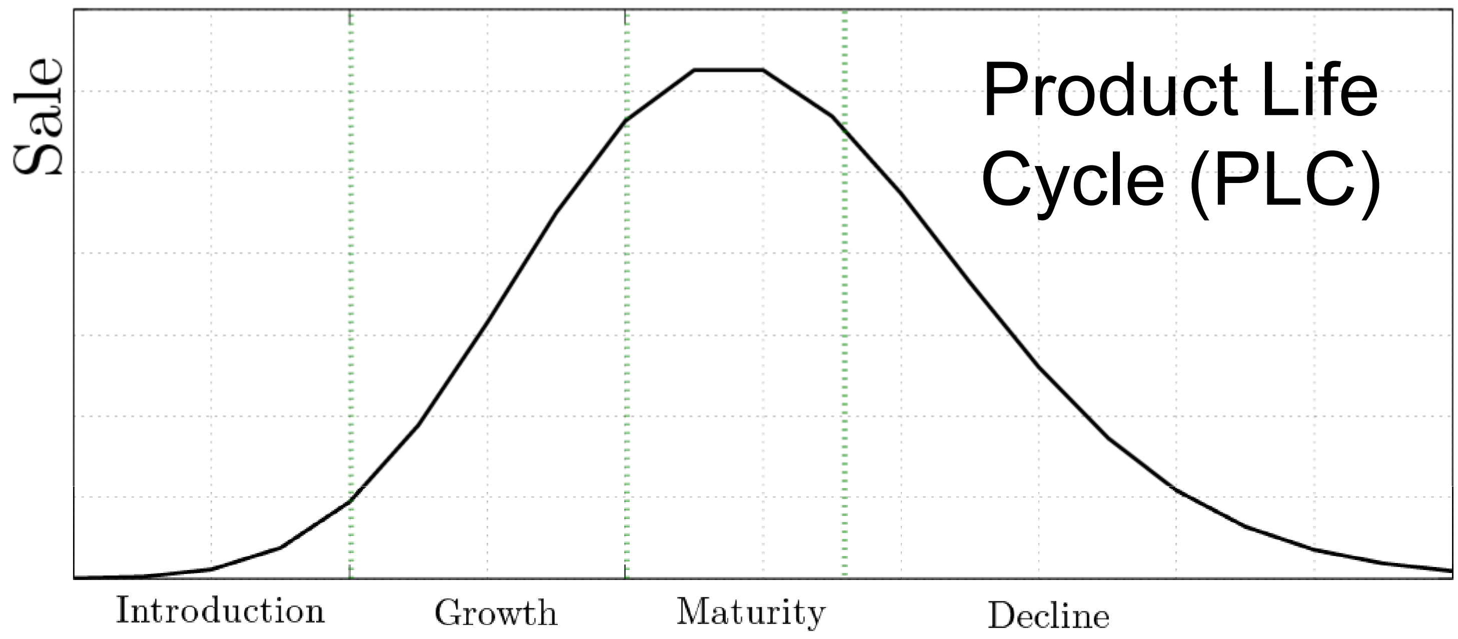


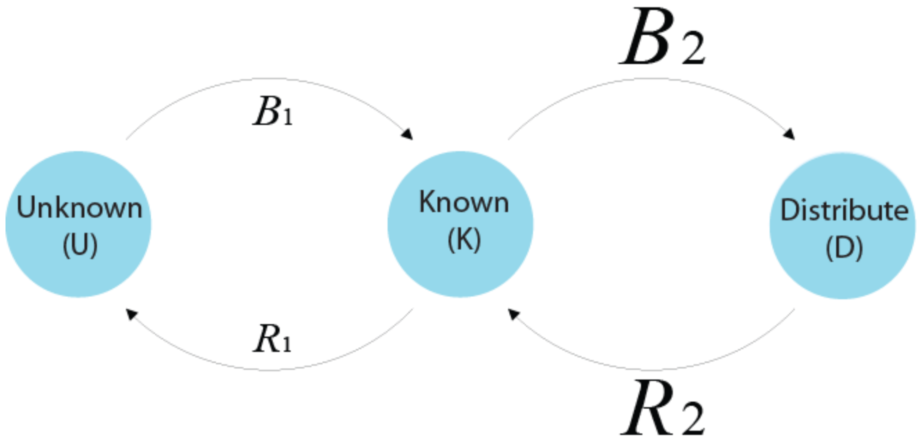
$$R_1(t)_i = 1 - e^{-\left(\frac{t - tp}{S}\right)}$$

tp = time when individual *i* transit to Know State

[2] H. Ebbinghaus: Memory: A Contribution to Experimental Psychology, Columbia University Press, New York(translated from Untersuchungen zur experimentellen Psychologie, Duncker and Humblot, Leipzig, Germany(1885), (1913).

Dynamism in B_2 & R_2

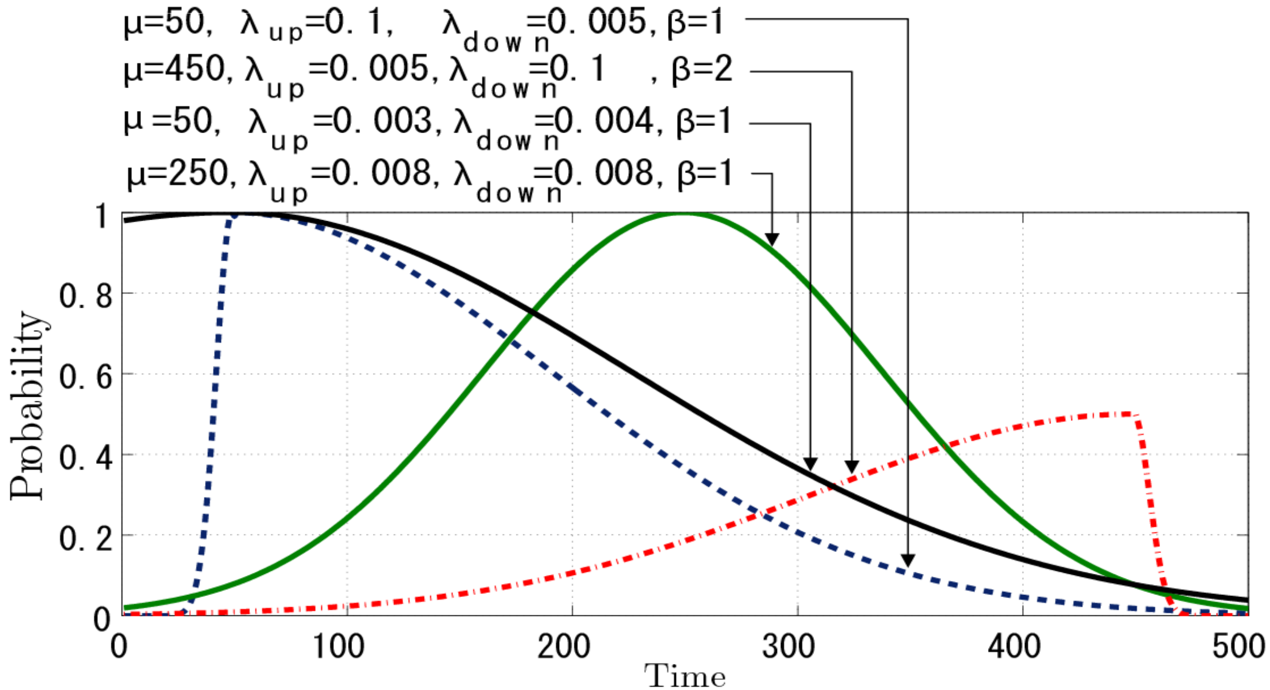




[3] T. Levitt: Exploit the product life cycle, Harvard Business Review, vol. 43, pp. 81-94, (1965).

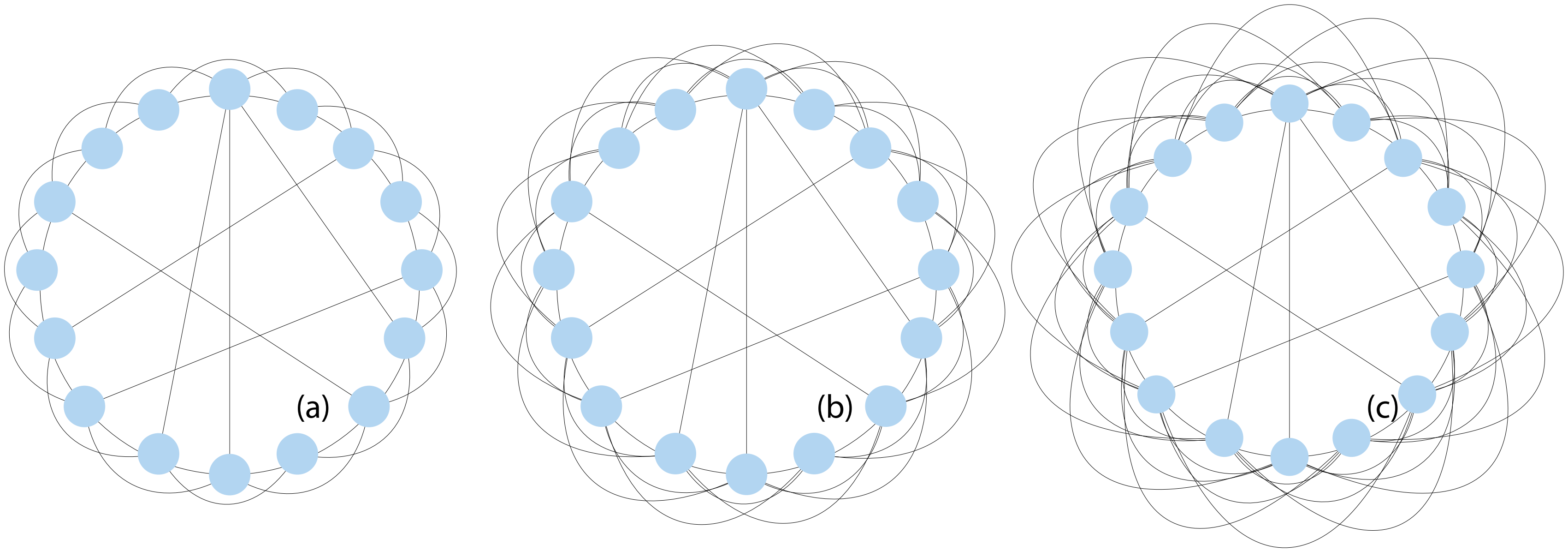
Dynamism in B_2 & R_2

$$B_2(t), R_2(t) = \begin{cases} \dfrac{e^{-(t-\mu)^2 \times \lambda_{up}}}{\beta} & (t \le \mu) \\ \dfrac{e^{-(t-\mu)^2 \times \lambda_{down}}}{\beta} & (t > \mu) \end{cases}$$



[3] T. Levitt: Exploit the product life cycle, Harvard Business Review, vol. 43, pp. 81-94, (1965).

Simulation on Small-World Network（SWN）



The illustration of a one-dimension ring lattice small-world network. (a) The model with Average Degree =4, (b) The model with Degree = 6, (c) The model with Degree = 8

SWN=Small-World Network

Average Connection＝k=Average Degree

[41] M.E.J. Newman and D.J Watts: Scaling and percolation in the small-world network model, Physical Review E60, pp. 7332-7442, (1999)
[8] D.J.Watts and S.H. Strogatz: Collective dynamics of 'small world' networks, Nature, vol. 393, pp. 440-442,(1998).

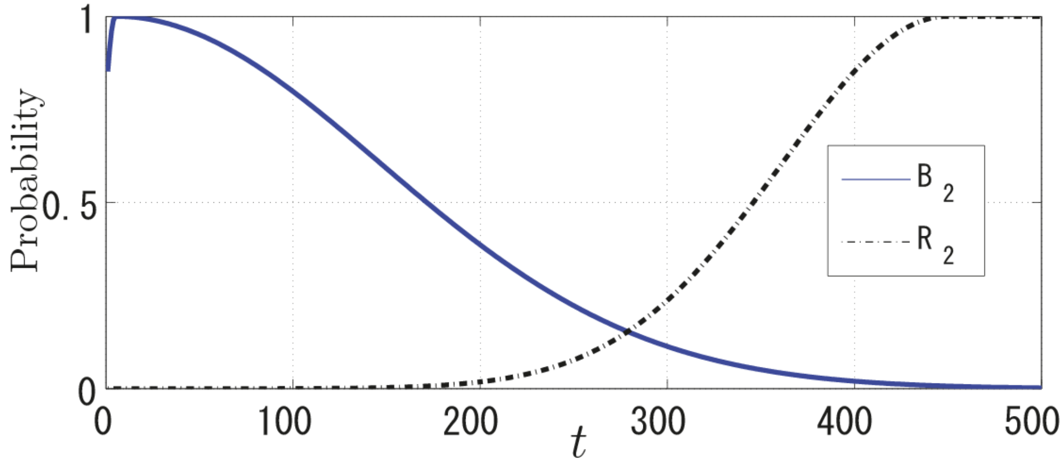
Simulation Overview

Table 5.1. Parameters Used in the Simulation.

Parameter	Explanation
$K(0)$	Number of initial Known State individuals at the beginning of information distribution. This parameter can be also considered as the result of the initial primary information distribution done in advance.
k	Average number of connections of each node in the network. In other words, k is average degree.
SC	Number of shortcuts.
Hub	Number of hubs.

Table 5.2. Default Parameter Values.

Parameter	Value
B_1	0.5
$R_1(t)$	Eq. (5.2) with $S = 100$
$B_2(t)$	Eq. (5.3) with $\mu = 5$, $\lambda_{up} = 0.1$, $\lambda_{down} = 0.005$, $\beta = 1$
$R_2(t)$	Eq. (5.3) with $\mu = 450$, $\lambda_{up} = 0.008$, $\lambda_{down} = 0$, $\beta = 1$
$K(0)$	1000
k	4
N	1,000,000
`range of` t	1 to 500
SC	100
Hub	100. Each hub has 10+k connections
$P(t)$	always 0



Simulation Overview

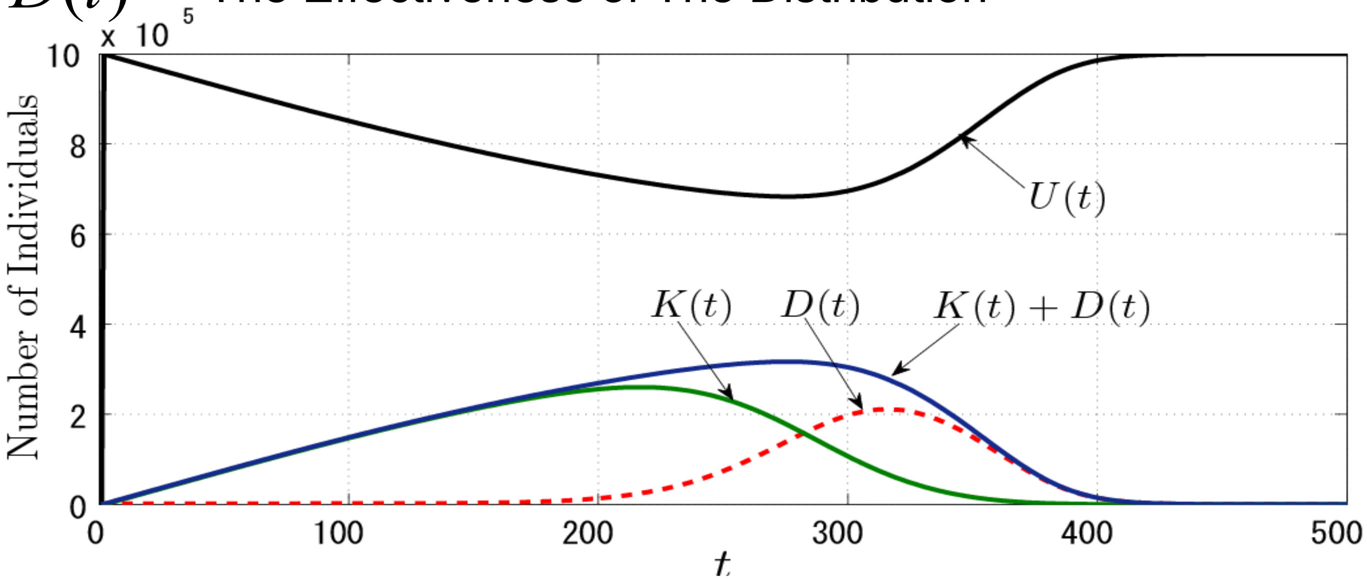
Table 5.1. Parameters Used in the Simulation.

Parameter	Explanation
$K(0)$	Number of initial Known State individuals at the beginning of information distribution. This parameter can be also considered as the result of the initial primary information distribution done in advance.
k	Average number of connections of each node in the network. In other words, k is average degree.
SC	Number of shortcuts.
Hub	Number of hubs.

Table 5.2. Default Parameter Values.

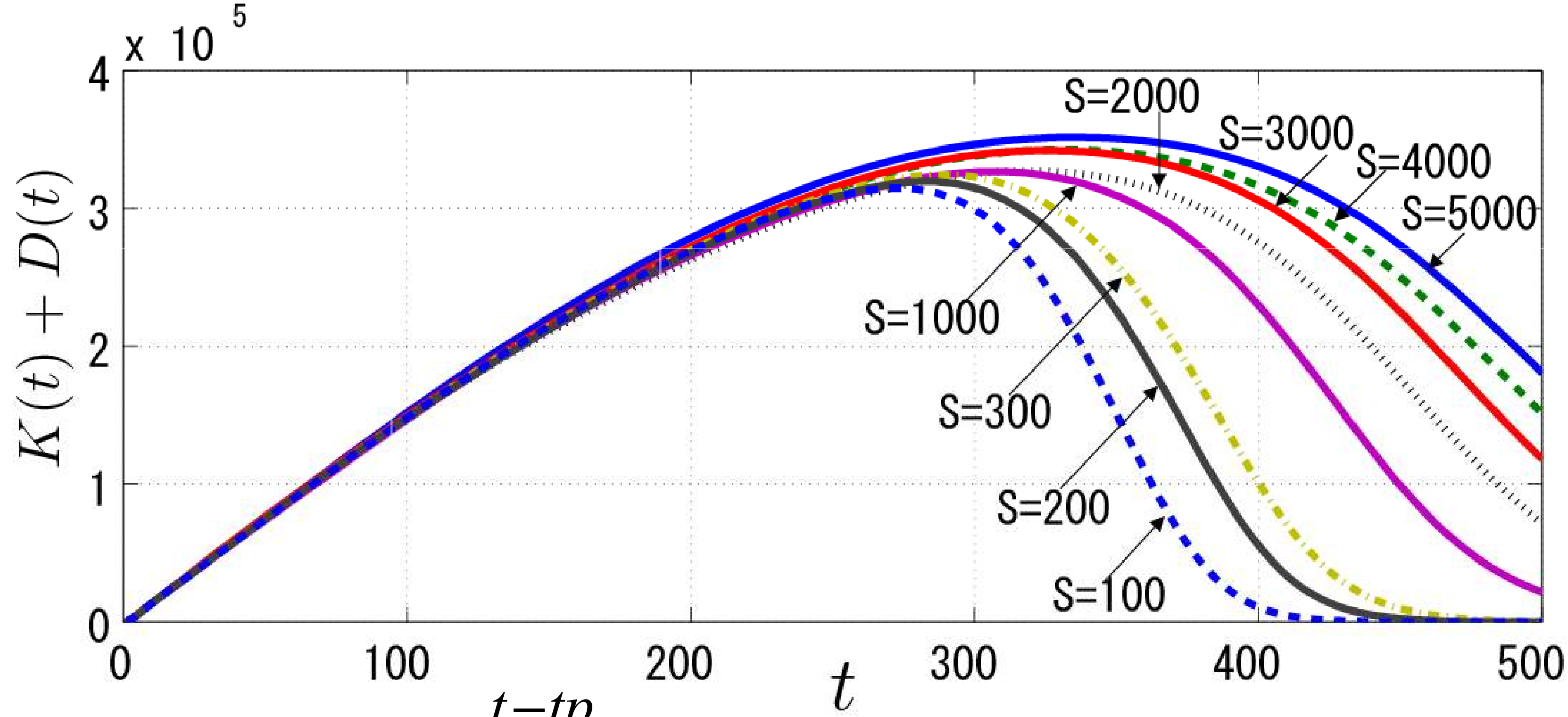
Parameter	Value
B_1	0.5
$R_1(t)$	Eq. (5.2) with $S = 100$
$B_2(t)$	Eq. (5.3) with $\mu = 5$, $\lambda_{up} = 0.1$, $\lambda_{down} = 0.005$, $\beta = 1$
$R_2(t)$	Eq. (5.3) with $\mu = 450$, $\lambda_{up} = 0.008$, $\lambda_{down} = 0$, $\beta = 1$
$K(0)$	1000
k	4
N	1,000,000
range of t	1 to 500
SC	100
Hub	100. Each hub has 10+k connections
$P(t)$	always 0

$K(t)+D(t) =$ The Effectiveness of The Distribution



The Impact of *R1*

Increasing S in below equation increases both maximum value of *K(t)+D(t)* and range of distribution effectiveness time. If value of *S* is big enough, further increasing its value does not have significant impact on the maximum value of *K(t)+D(t)*

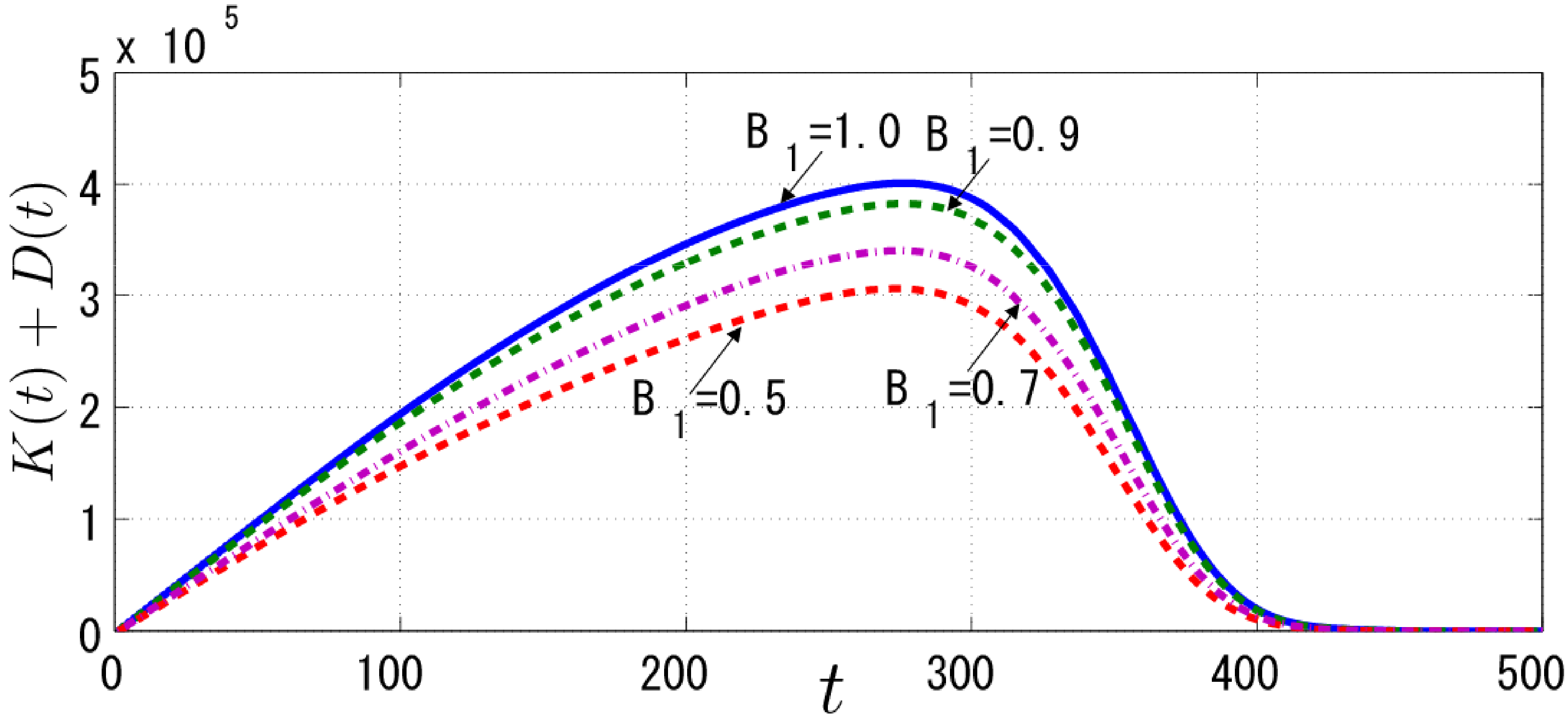


$$R_1(t)_i = 1 - e^{-\left(\frac{t-tp}{S}\right)}$$

tp=time when individual transit to Know State
S=the relative strength of memory.

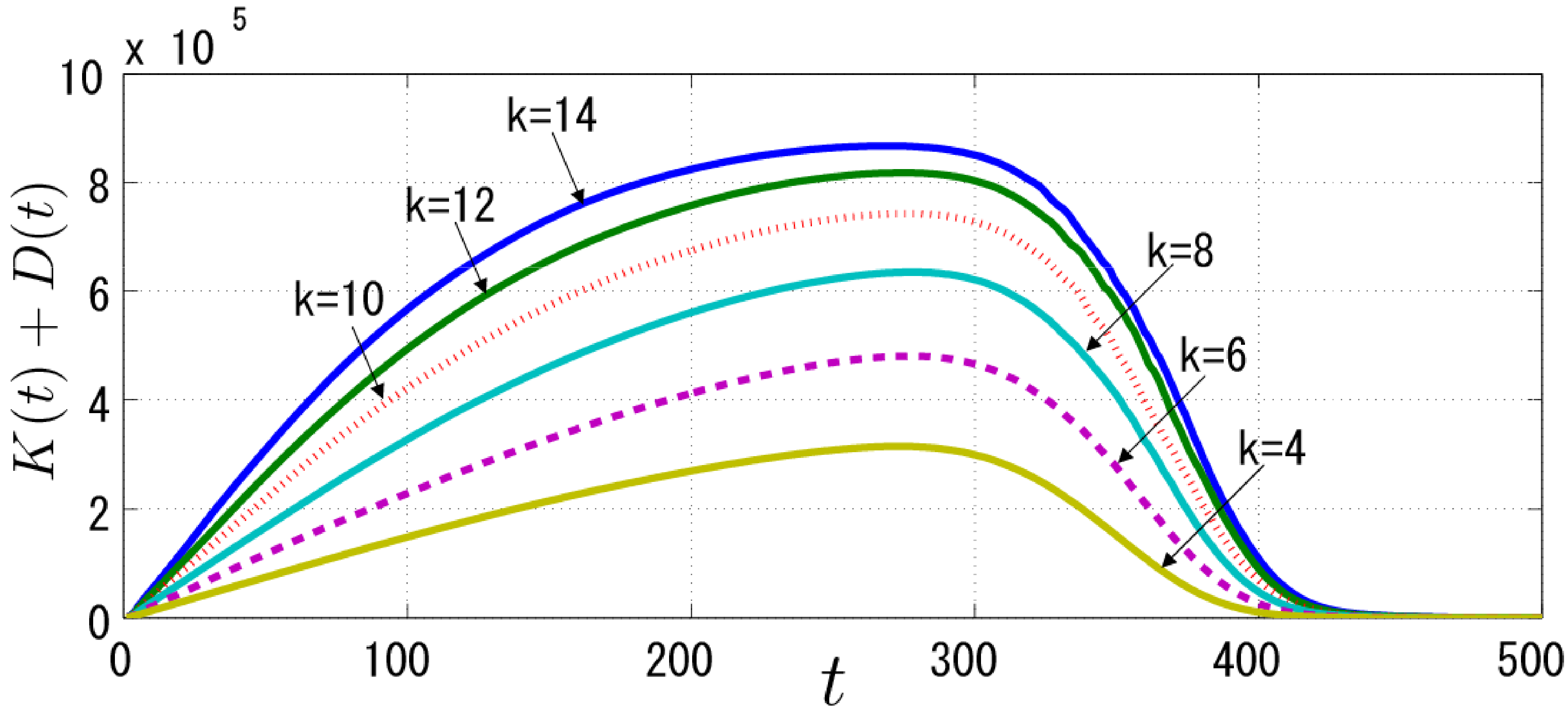
The Impact of *B1*

Using *B1* and *R1* to increase the maximum value of *K(t)+D(t)* is not effective.



The Impact of *k*

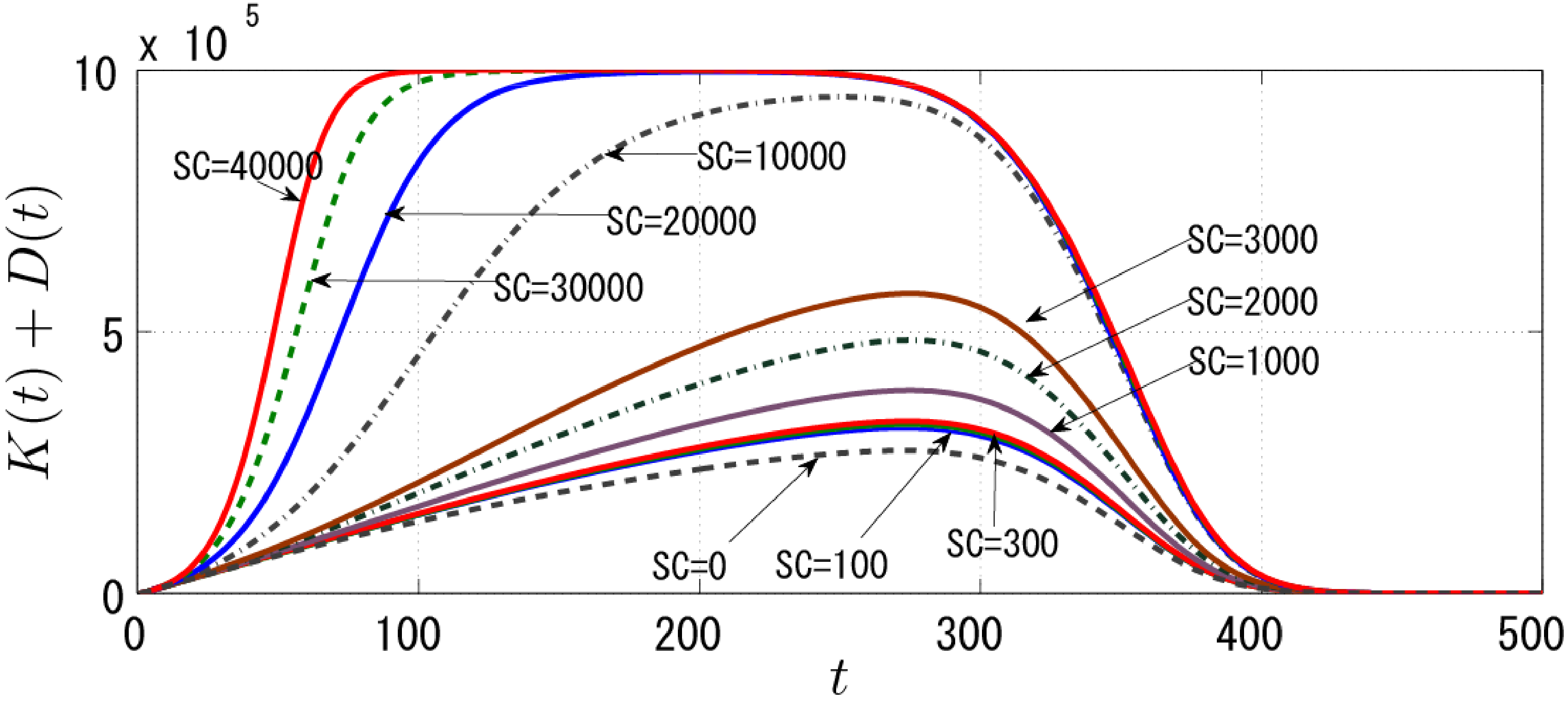
The effectiveness of information distribution on the network which has high *k* is not much improved.



The Impact of Shortcut (*SC*)

•*SC=0* (the regular ring lattice model) to *SC=100* -> Small but noticeable impact.

• If the network have very high number of shortcuts(*SC*) or *k*, it is not necessary to increase shortcuts in the networks.

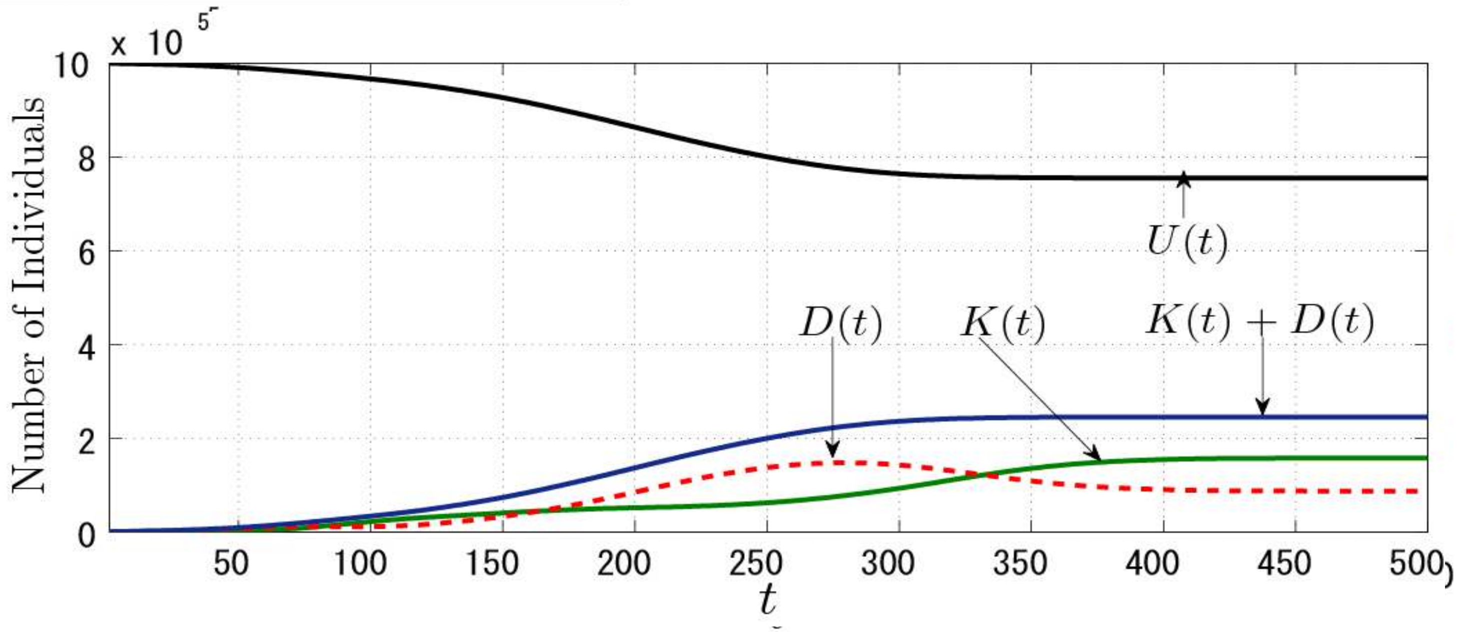


Effect of R_1, R_2 and B_2 to Distribution Result

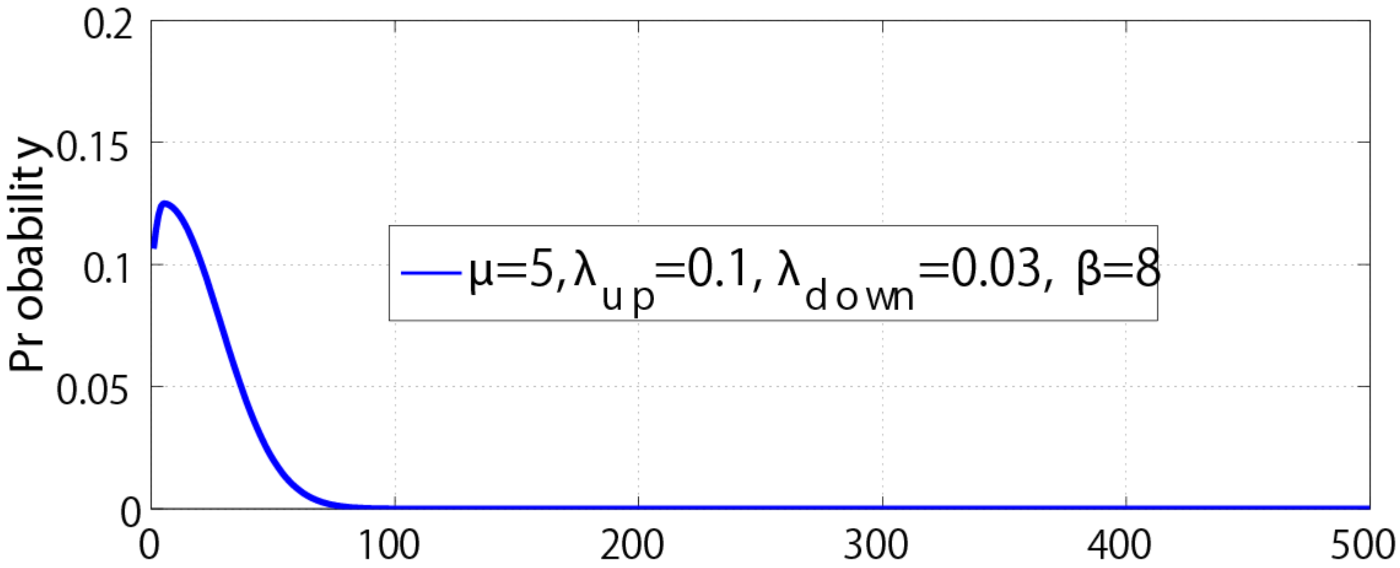
Table 5.3. Groups of R_1, R_2, B_2 and B_1 Values.

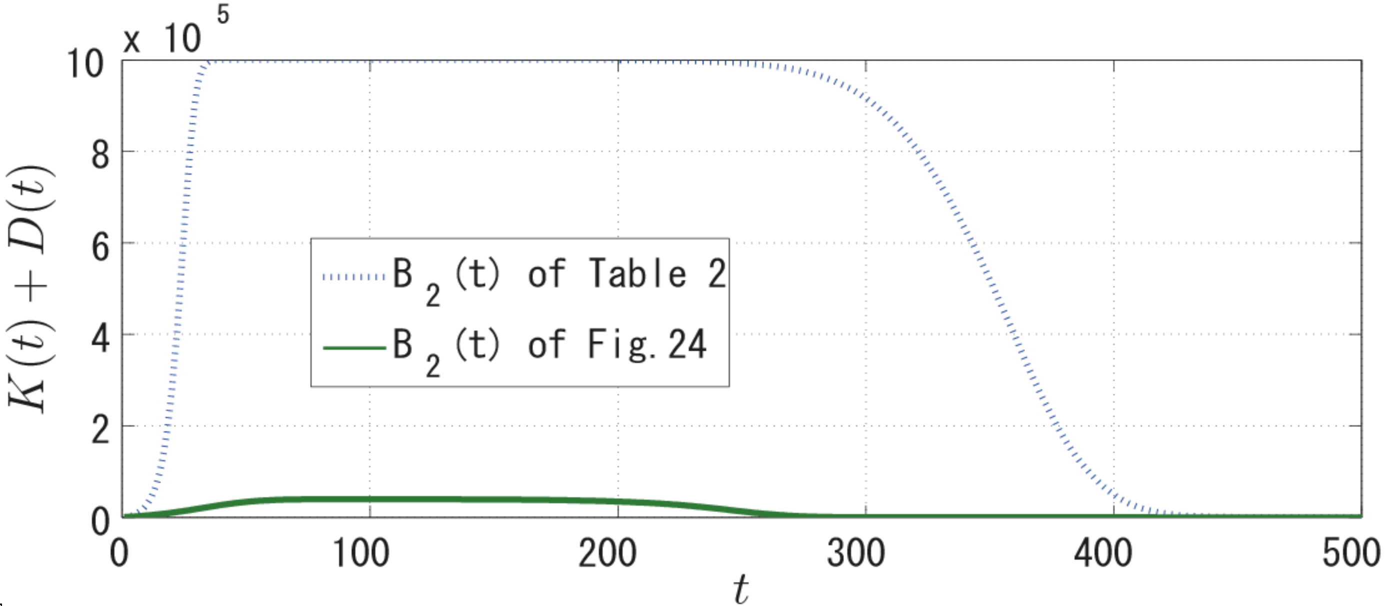
Pattern	Parameter	**Value**
Pattern1	$B_2(t)$	Eq. (5.3) with $\mu = 200$, $\lambda_{up} = 0.01$, $\lambda_{down} = 0.015$, $\beta = 1$
	$R_2(t)$	Eq. (5.3) with $\mu = 120$, $\lambda_{up} = 0.03$, $\lambda_{down} = 0.009$, $\beta = 1$
	$R_1(t)$	0
	B_1	0.5
Pattern2	$B_2(t)$	1
	$R_2(t)$	0.5
	$R_1(t)$	0.5
	B_1	0.5

Those parameters can change phase transition of the distribution results.



Consideration of $B_2(t)$





[6]P. S. Dodds, R. Muhamad, and D. J. Watts, "An Experimental Study of Search in Global Social Networks", Science, vol.301, pp.827-829, August 2003.

“Analysis of Illegal Content Distribution”

-An Application of This Research Study-

Research Background

Secondary information distribution has been playing an important role in our society

⇩

Nevertheless, the secondary information distribution has been also increasing the power of ***illegal distribution of content***.

⇩

Small-World Network (SWN) or social network is another popular way to distribute illegal distribute content

⇩

With the advanced information communication technologies, distributing content in the small-world network has been becoming ***easier and more powerful*** than before

Research Approach Introduction

The proposed model in previous study is simulated and analyzed by applying the ***SWN*** and it characteristics to ***analyze illegal content distribution***.

SWN=Small-World Network

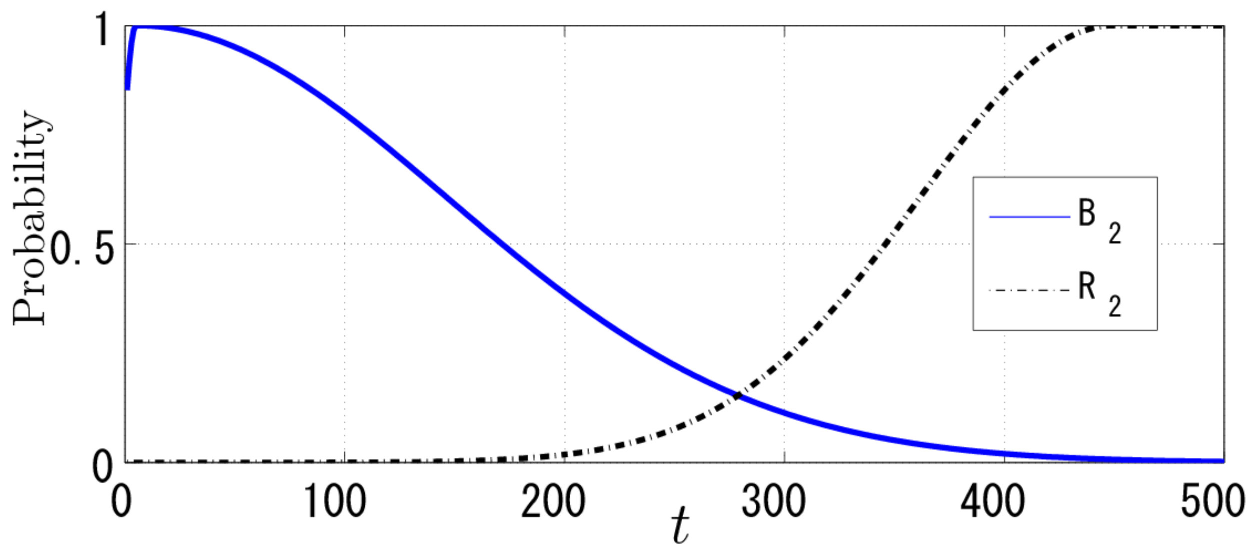
Simulation Overview

Table 6.1. Parameters Used in the Simulation.

Parameter	Explanation
$K(0)$	Number of initial Known State individuals at the beginning of information distribution. This parameter can be also considered as the result of the initial primary information distribution done in advance.
k	Average number of connections of each node in the network
SC	Number of shortcuts.

Table 6.2. Default Parameter Values.

Parameter	Value
B_1	0.5
$R_1(t)$	Eq. (5.2) with $S = 100$
$B_2(t)$	Eq. (5.3) with $\mu = 5$, $\lambda_{up} = 0.1$, $\lambda_{down} = 0.005$, $\beta = 1$
$R_2(t)$	Eq. (5.3) with $\mu = 450$, $\lambda_{up} = 0.008$, $\lambda_{down} = 0$, $\beta = 1$
$K(0)$	10
k	4
N	10,000
range of t	1 to 500
SC	50



Simulation Scenario

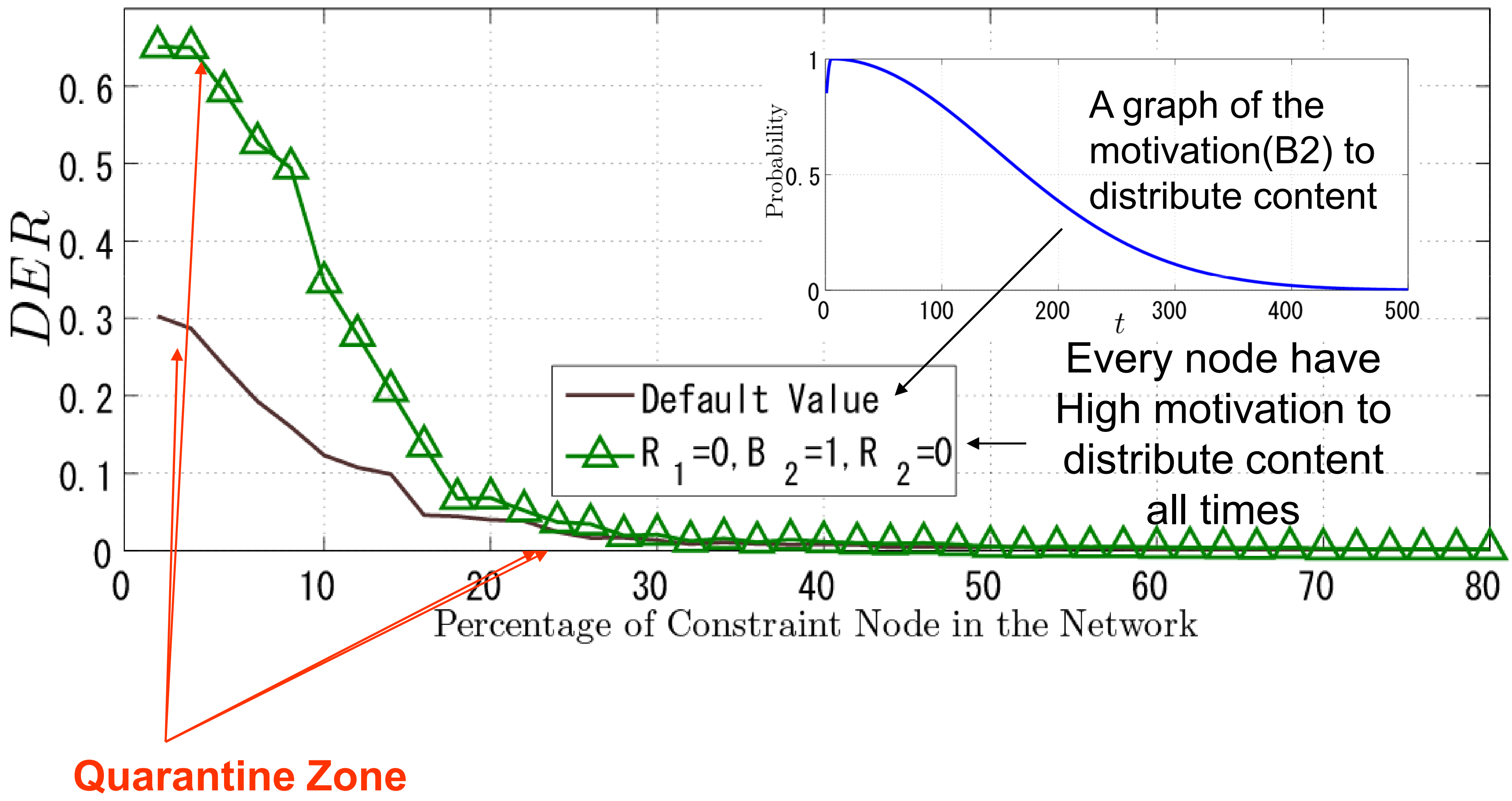
1. Nodes in the network which have ability to copy and redistribute content are defined as ***Bad nodes***, and nodes which don' t have ability to redistribute content are ***Constraint nodes***.

2. Firstly, we suppose that all nodes in the network are ***Bad nodes***.

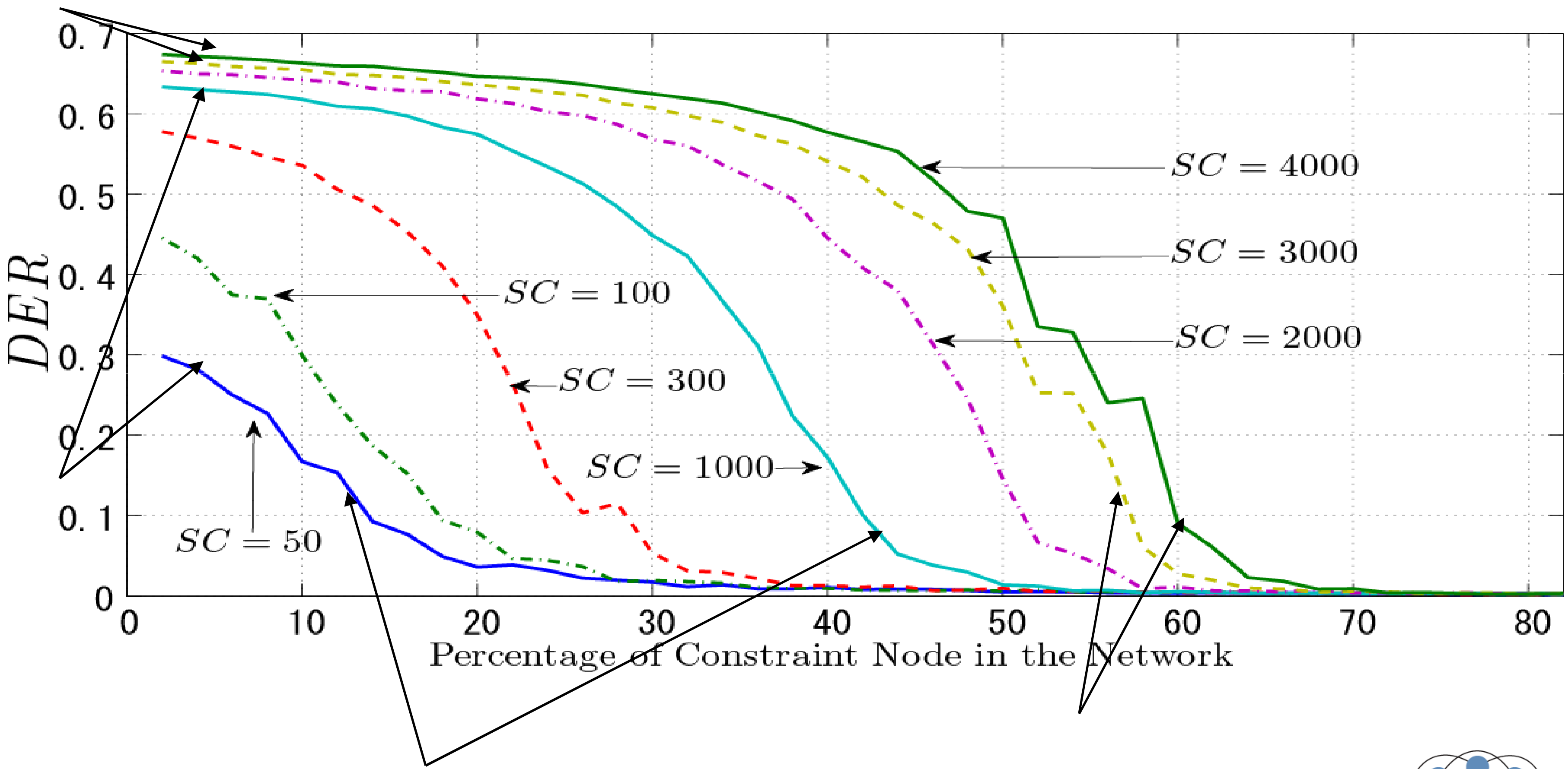
3. Subsequently, we conduct simulations by changing ***Bad nodes*** to be ***Constraint nodes*** at regular intervals and observing impact to ***DER***(*(**D**istribution **E**ffectiveness **R**ate)*).

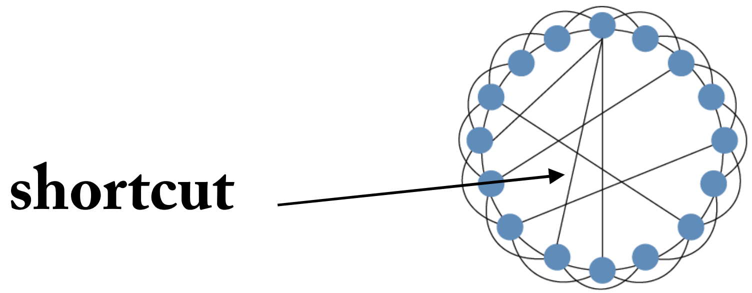
$$DER = \frac{1}{T \bullet N} \sum_{t=1}^{T} K(t) + D(t)$$

Simulation Result of Default Value

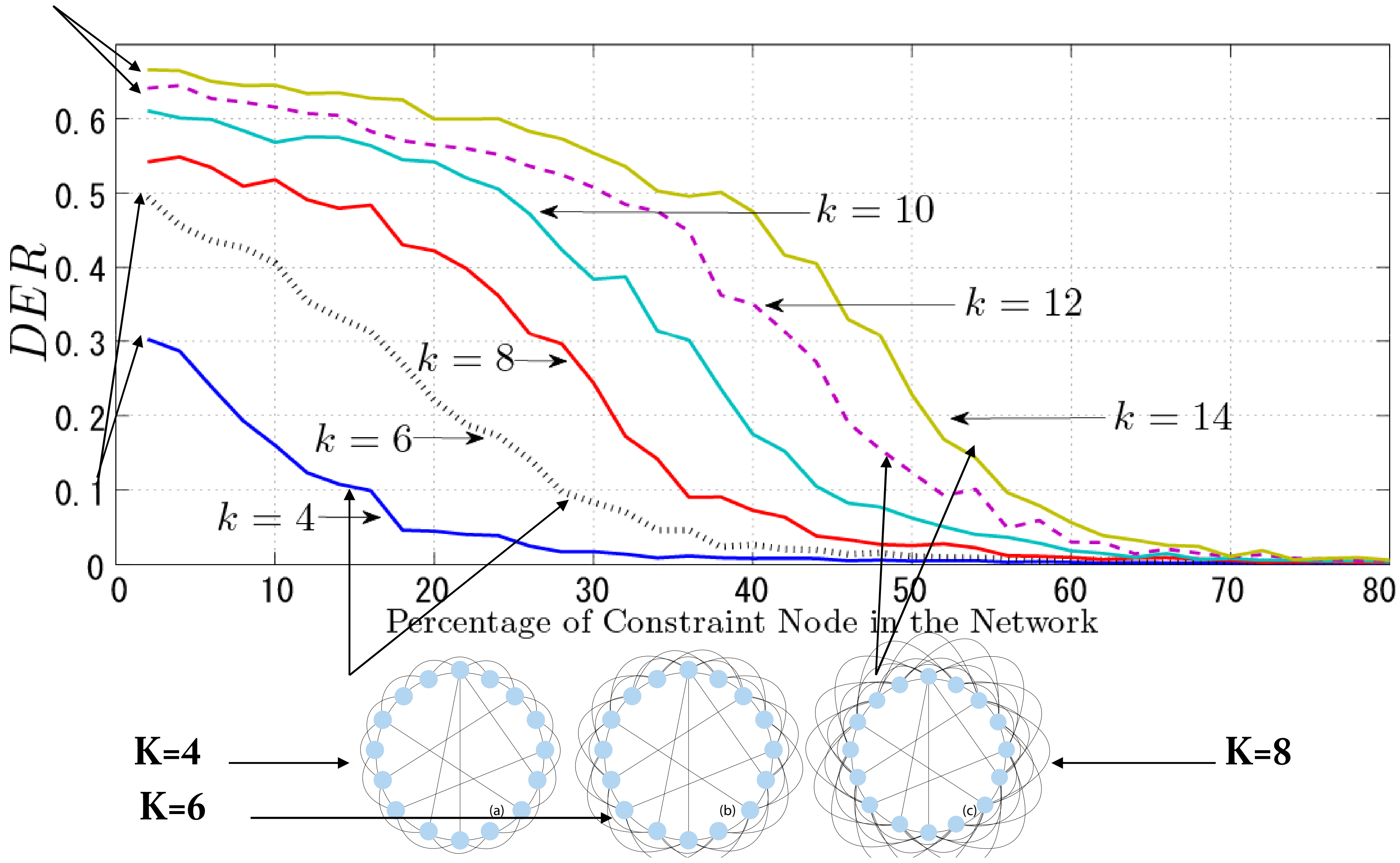


Impact of Shortcuts



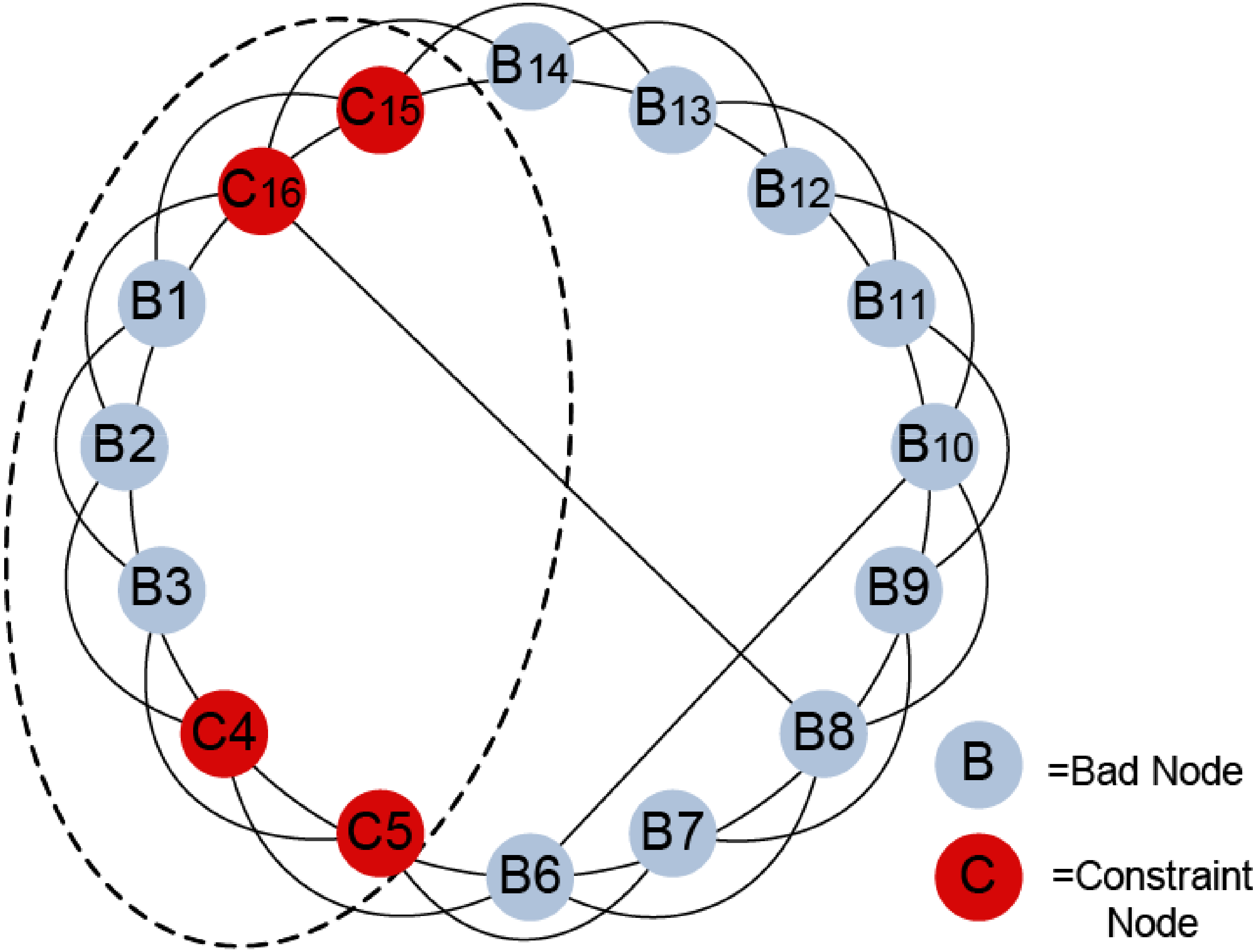


Impact of Average Degree of each Node (*k*)



Why Quarantine Zone is Happened

Confinement Phenomenon



Suppressing the number of Bad nodes
less than 80 percent Constraint Nodes

Conclusion of The Study

- Information distribution model based on human's behavior with dynamic parameters is proposed.
- Small-world network and its characteristics are applied to conduct simulations.
- The simulation results show how to increase the power of the distribution efficiently by using parameters in the proposed model, and how the model can be useful for optimizing information distribution on social networks.
- The result can apply to *optimize number of equipped DRM software* in the network in order to protect illegal distribution. *For an example, if all users in the network are bad nodes, 20% or 30 % of DRM software are sufficient to confine illegal distribution in the network which have k=4 and a few shortcuts.*

Future Works

- Find the relationship between the real-world phenomena and the parameters
- Study how to estimate the appropriate values of parameters in the proposed model for the real world social network
- Further study in analysis of *B2* and *R2.*
- Include more realistic scenarios and conduct simulations