



Could you design a model which represents a firewall installed in a company?
企業内のファイアウォール（情報隔壁）のモデルを設計してみませんか？

November 16 - 20, 23, 2020

Study Group Workshop 2020
Institute of Mathematics for Industry, Kyushu University



Shinto Teramoto
Professor of Law, Kyushu University, Japan
teramoto.shinto.717@m.kyushu-u.ac.jp
shin.teramoto@terrara.net

この問題の背景



The reason why PHR service providers should establish internal firewalls

October 23-24, 2020

GKA TECHNO 2020

10th International Conference on Science, Technology and Society



Shinto Teramoto
Professor of Law, Kyushu University, Japan
teramoto.shinto.717@m.kyushu-u.ac.jp
shin.teramoto@terrara.net

Sharing Medical and Health Records through PHR



- It is generally agreed that sharing medical and health records of patients between them and medical institutions or healthcare service providers in charge of their care is likely to promote efficient and effective medical and healthcare services .
- However, it is not necessarily practical to have EHR (Electronic Health Records) of medical institutions accessible to other medical institutions and healthcare service providers.
- Limited interoperability of Legacy EHRs
- The security level of all of the institutions mutually connected with one another is likely to be subject to the standard of security implemented at the institution that is managing the security of EHR in the poorest manner.

- PHR (Personal Health Record) are likely to provide a solution to the said problem.
- Suppose that a patient's records in the EHR of a medical institution is duplicated to her PHR.
- It is the responsibility of the patient herself to permit future physicians and institutions caring her to access her records in her own PHR.
- By this way, we can ensure that each patients has autonomous control over access to their health and medical records, while relieving physicians and medical institutions from possible liability caused by the misappropriation of patients' records by other physicians and institutions.



Electronic Healthcare Records

EHR

Personal Healthcare Records (PHR)



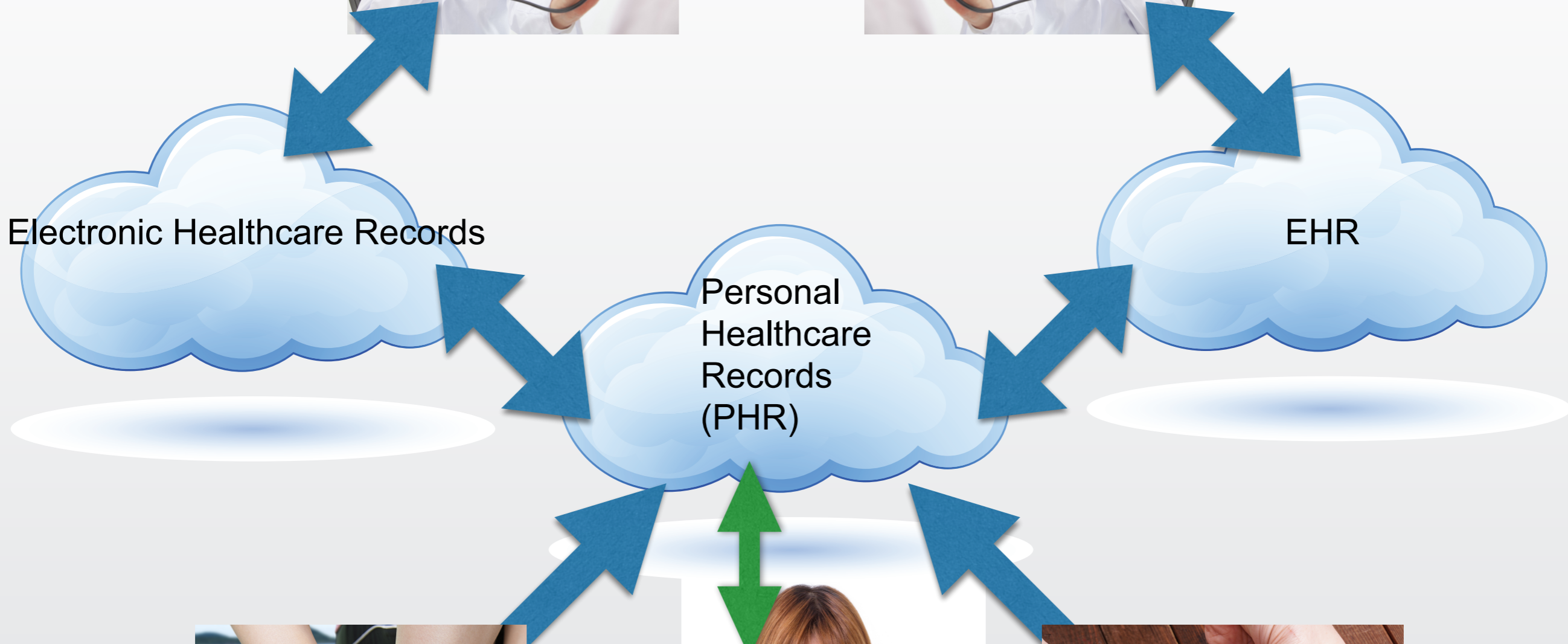
Personal Healthcare Device



7



PHD



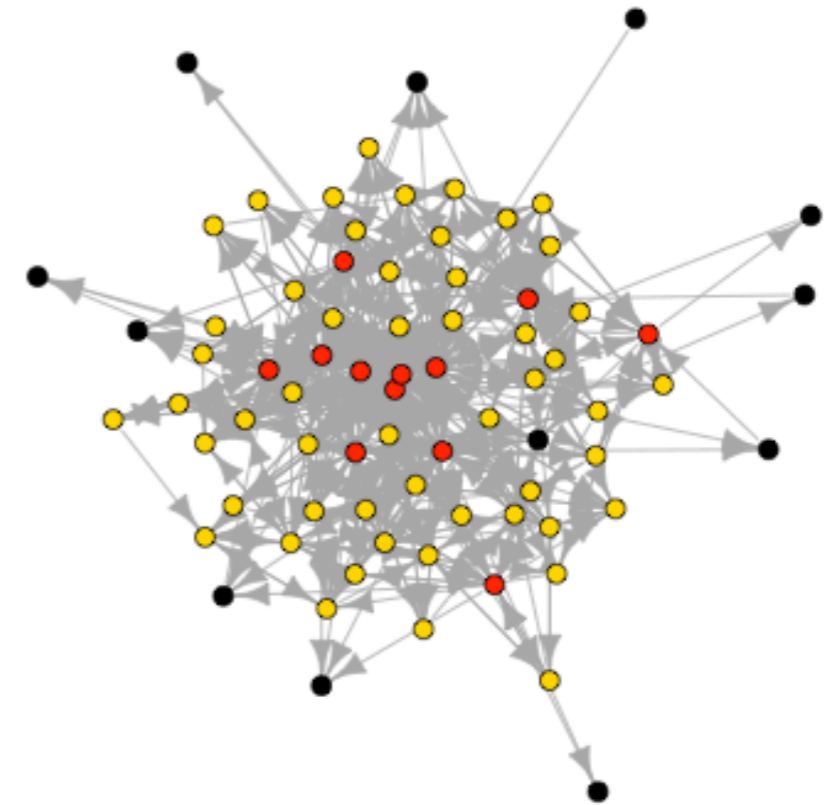
Possible but Significant Concern of the Subscribers to PHR Services



- The medical and health information of the subscribers could be divulged from the PHR, or misappropriated by the PHR service providers or their corporate customers purchasing personal information.
- Even if neither PHR service providers nor their corporate customers misappropriate subscribers' medical and health information, it is understandable that the subscribers will be concerned about such risk.

- These concerns are likely to make us hesitate to use PHR and discourage wider usage of PHR in the society, and endanger the sustainable development of PHR services.
- The possible concerns of PHR subscribers are tightly linked to the dissemination of information through the social network surrounding PHR service subscribers, providers, and customer companies purchasing information.
- It makes sense to seek a solution to mitigate the concerns of subscribers by using a social network model.

Discussion using Social Network Models



The components of a network model

- Actor: The model used here (hereinafter, the “Model”) assumes that a society consists of multiple actors and their mutual relationships.
- Vertex: Each actor is represented by a vertex, which is denoted by a dot or a small circle on a graph.
- Arc: An arc (namely, a directed tie) denotes that the sender of the arc depends on the receiver thereof in order to access one or more pieces of information held by the receiver.


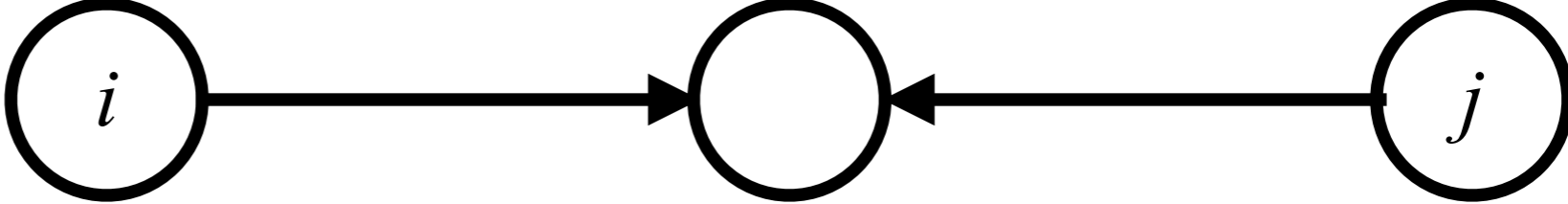

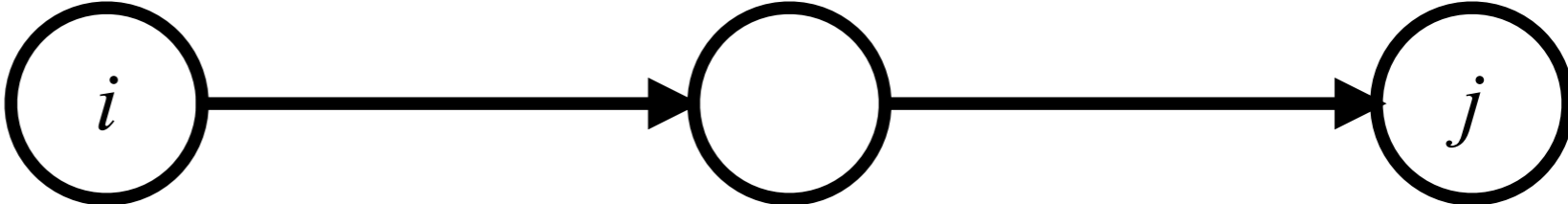
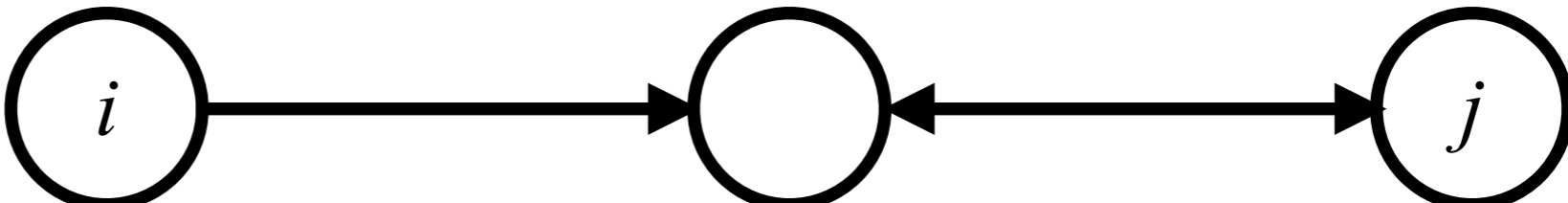
The tools to examine a network model

- Indegree Centrality

- Indegree centrality of a vertex is the ratio of the actual number of incoming arcs against the maximum possible number of incoming arcs received by the same vertex in the network to which the vertex belongs.
- Suppose that $vertex_i$ belongs to a network G , whose number of vertices is n . Suppose also that the number of incoming arcs received by $vertex_i$ is k .
- The maximum possible number of incoming arcs received by $vertex_i$ is $n-1$.
- Accordingly, the indegree centrality of $vertex_i$ is $Id_i = \frac{k}{n-1}$.
- Indegree centrality of $vertex_i$ represents how many vertices depend on $vertex_i$ to access information held by $vertex_i$.

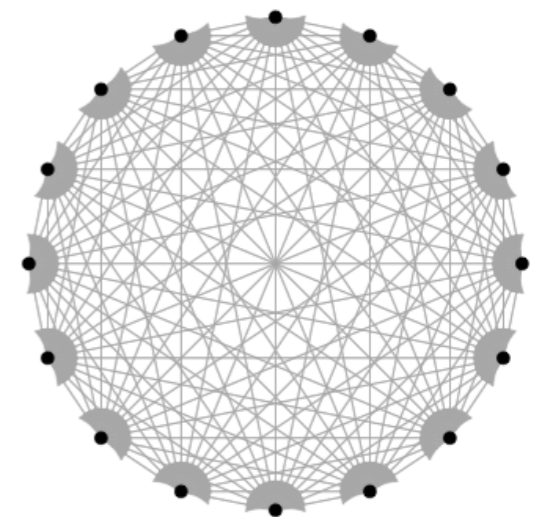
- Outdegree Centrality
 - Outdegree centrality of a vertex is the ratio of the actual number of outgoing arcs against the maximum possible number of outgoing arcs sent by the same vertex in the network to which the vertex belongs.
 - Suppose that $vertex_i$ belongs to a network G , whose number of vertices is n . Suppose also that the number of outgoing arcs sent by $vertex_i$ to other vertices is k .
 - The maximum possible number of outgoing arcs received by $vertex_i$ is $n-1$.
 - Accordingly, the outdegree centrality of $vertex_i$ is $Od_i = \frac{k}{n-1}$.
 - Outdegree centrality of $vertex_i$ represents from how many vertices $vertex_i$ collects information.

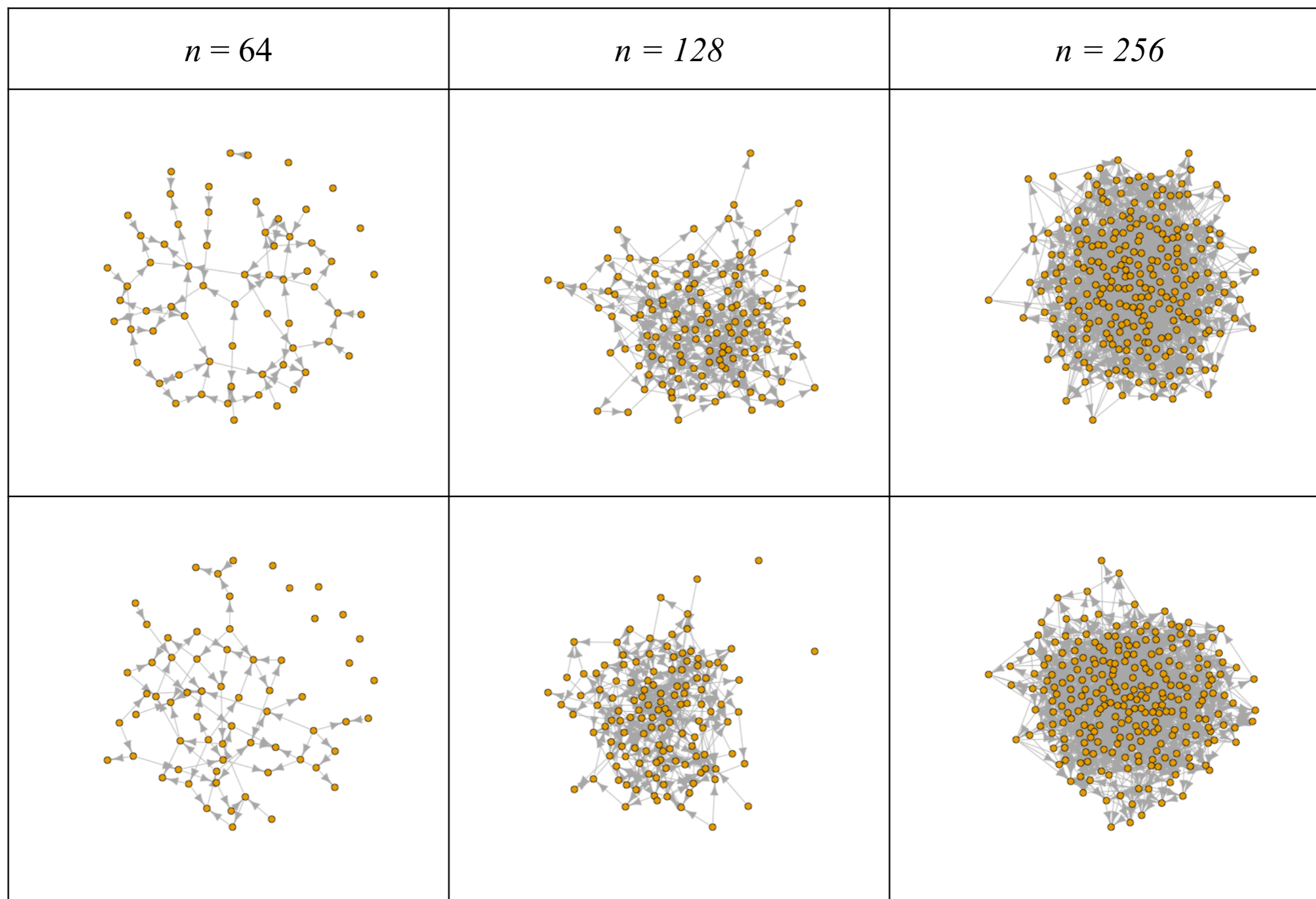
Distance

The distance from $actor_i$ to $actor_j$ is infinite.	
The distance from $actor_i$ to $actor_j$ is infinite.	
The distance from $actor_i$ to $actor_j$ is 1.	
The distance from $actor_i$ to $actor_j$ is 2.	
The distance from $actor_i$ to $actor_j$ is 2.	

Average Distance

- The average distance of a network is pertinent to the issues here, because a shorter average distance is likely to accelerate the dissemination of personal information through the society, once it is divulged.
- For example, in a complete graph (its *average distance* = 1), any information transmitted by any actor is directly delivered to every actor.





Examples of the beginning conditions of the Model, which is a directed random graph having 64 , 128 , or 256 vertices.

The Rules Implemented in the Model to Develop the Network

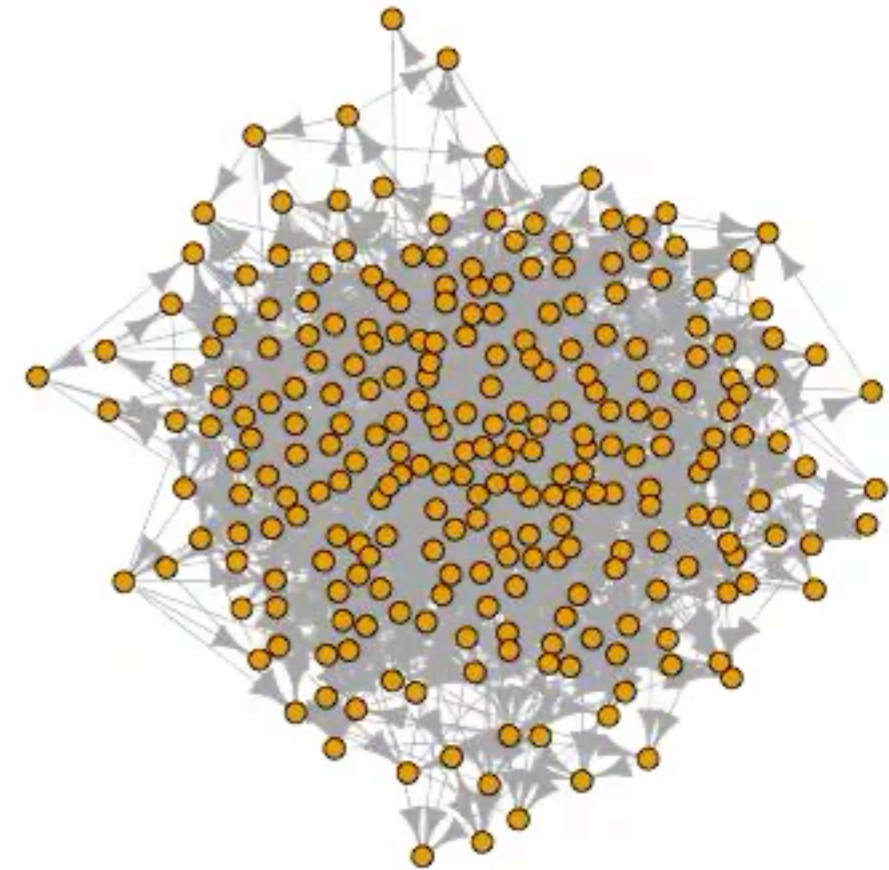


- The Model assumes that every actor behaves according to a very simple rule, which is very likely in our actual society, and predicts that the social network will develop as a result of the repeated behaviors of the actors.

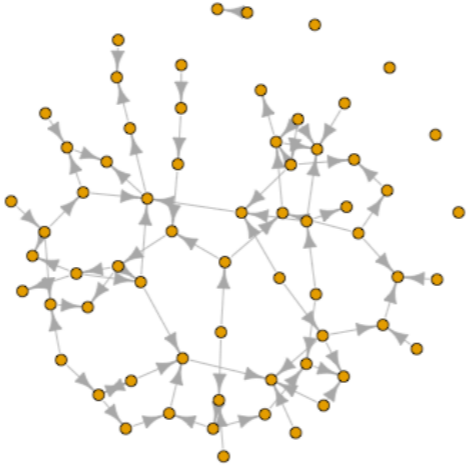
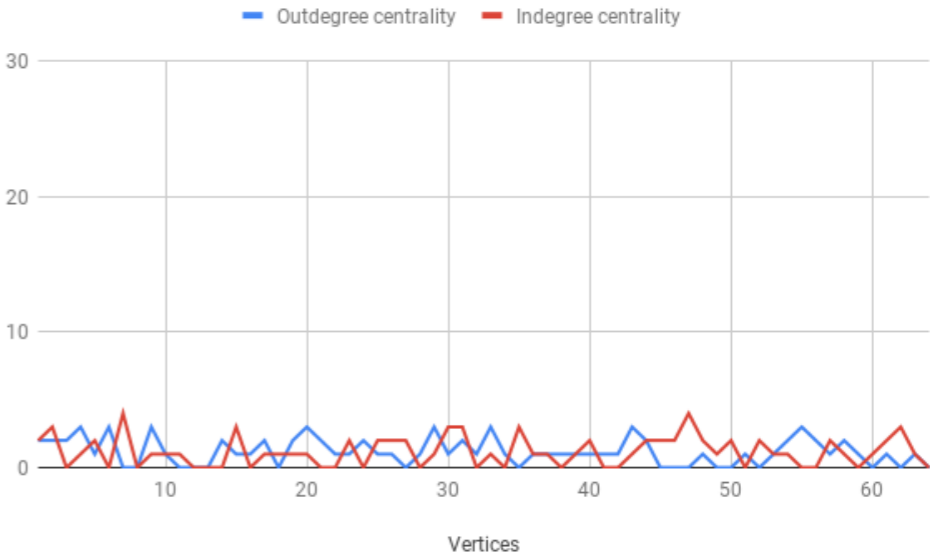
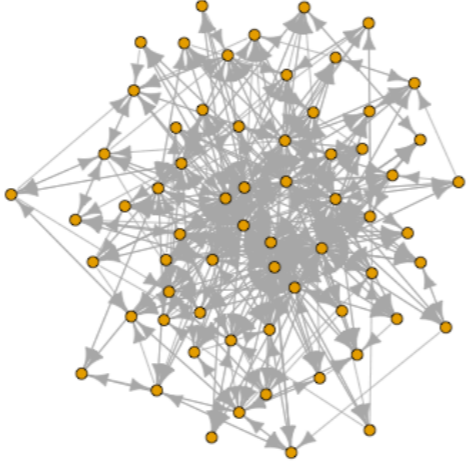
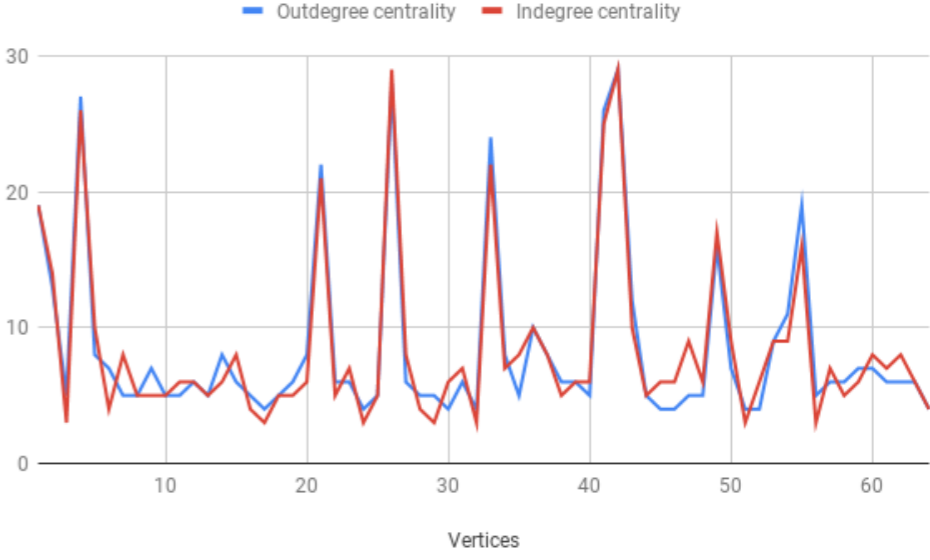
- Assumptions:
 - An actor is likely to want to know the personal information of as many as possible other actors, while she wants to minimize the cost to be borne by her to know the personal information of other actors;
 - An actor easily permits another actor to access her own personal information in exchange for her having access to the personal information of other actor(s) granted by the latter actor irrespective of whether such personal information is that of the latter actor herself or of a third party;
 - An actor cannot know the personal information of another actor unless she accesses such actor, or accesses a third actor who holds the personal information of the said another actor;
 - The cost that is borne by an actor to send an arc to another actor is the same for any actor and any arc;
 - Considering the above factors, it is preferable for any actor to access the personal information of as many as possible other actors by sending an arc; and
 - Accordingly, an actor prefers to send an arc to an actor having higher outdegree centrality than herself (in the Model, greater than twice of the outdegree centrality of herself).

- Rules:
 - Every individual actor (referred to as “ $actor_i$ ”) chooses one actor (referred to as “ $actor_t$ ”) arbitrarily from among the other actors;
 - The said $actor_i$ assesses whether the outdegree centrality of $actor_t$ is greater than twice of the outdegree centrality of $actor_i$;
 - Only when the answer is positive, $actor_i$ sends an arc to $actor_t$; and
 - In exchange, $actor_i$ receives an arc sent by $actor_t$.
- Hereinafter, the social network in the Model is developed through multiple rounds (k), at each of which every actor behaves in accordance with the said rule.

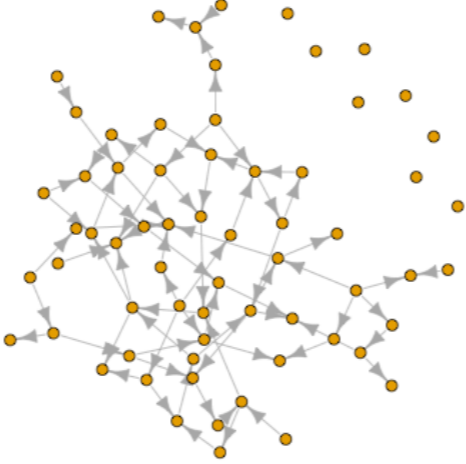
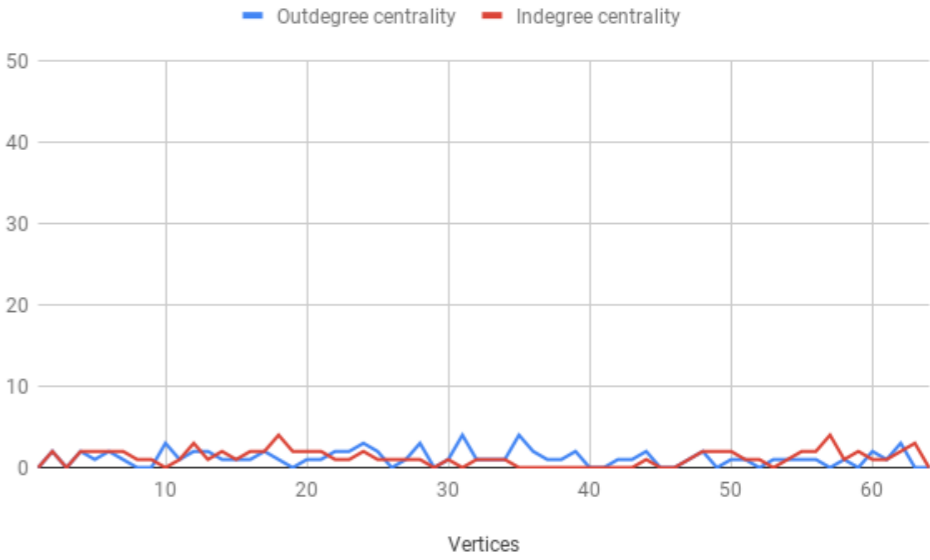
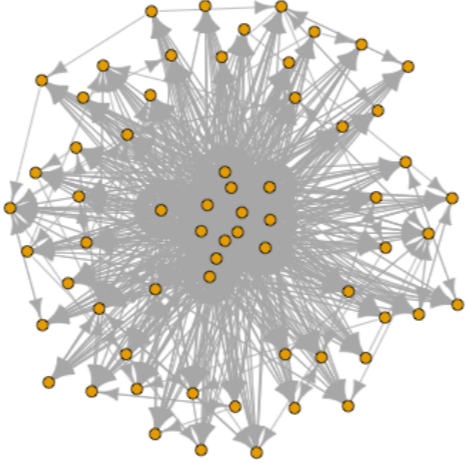
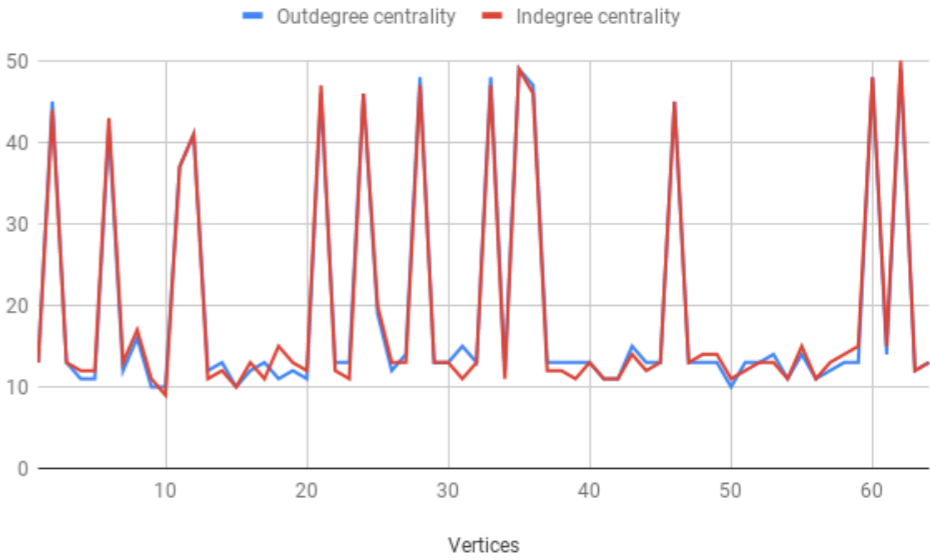
The Development of the Social Network



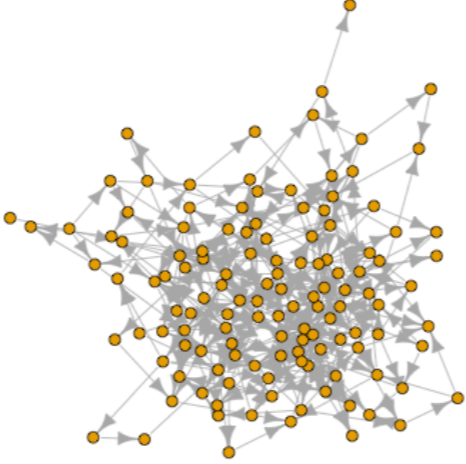
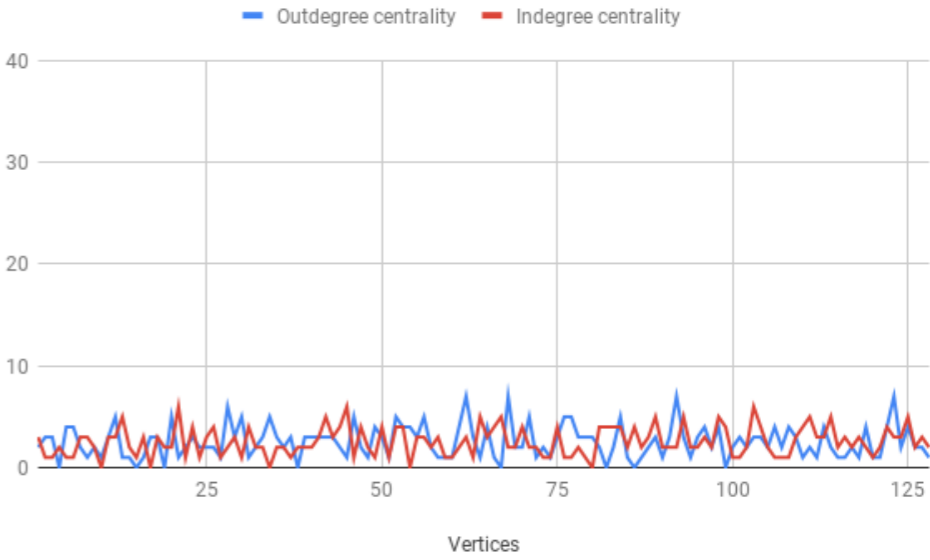
- In the Model, the social network continues to develop according to the actions of individual actors who behave subject to the said rules.

$n=64, k=32$	Graph	<i>Outdegree and indegree centralities of the respective vertices</i>
<i>At the beginning condition</i>		
<i>At the final condition</i>		

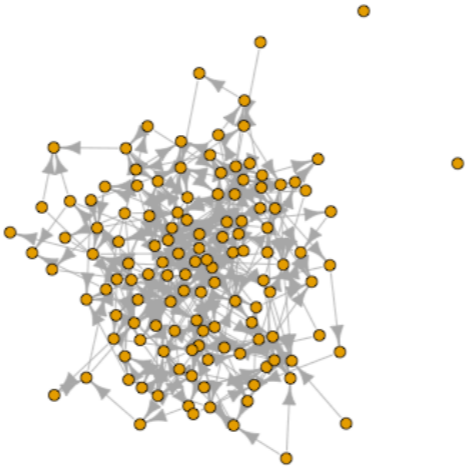
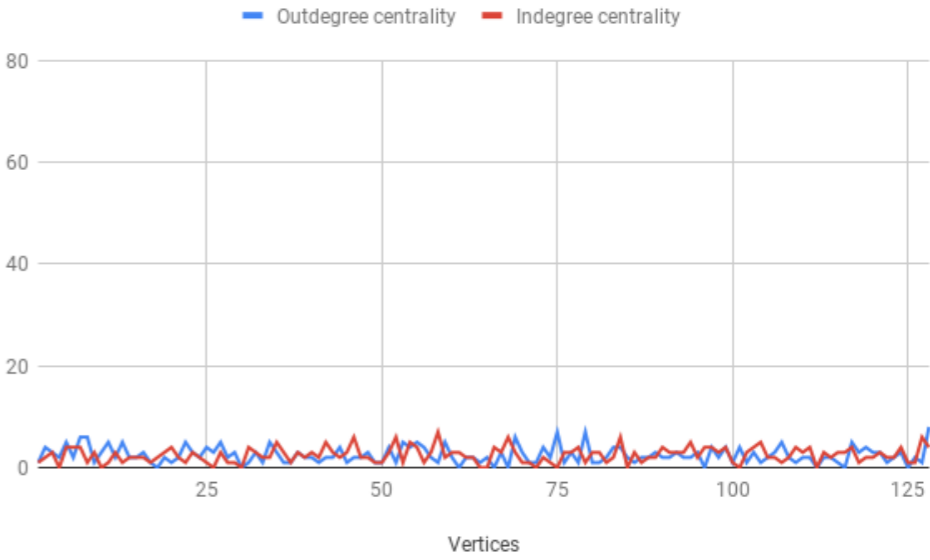
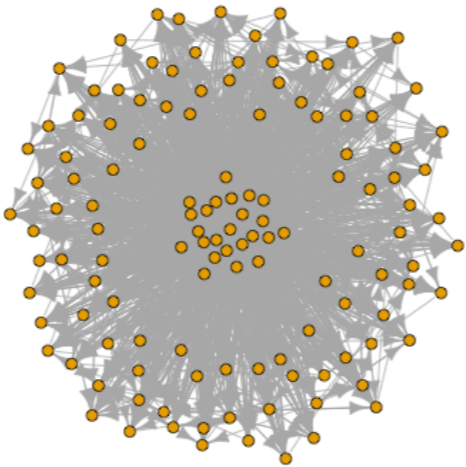
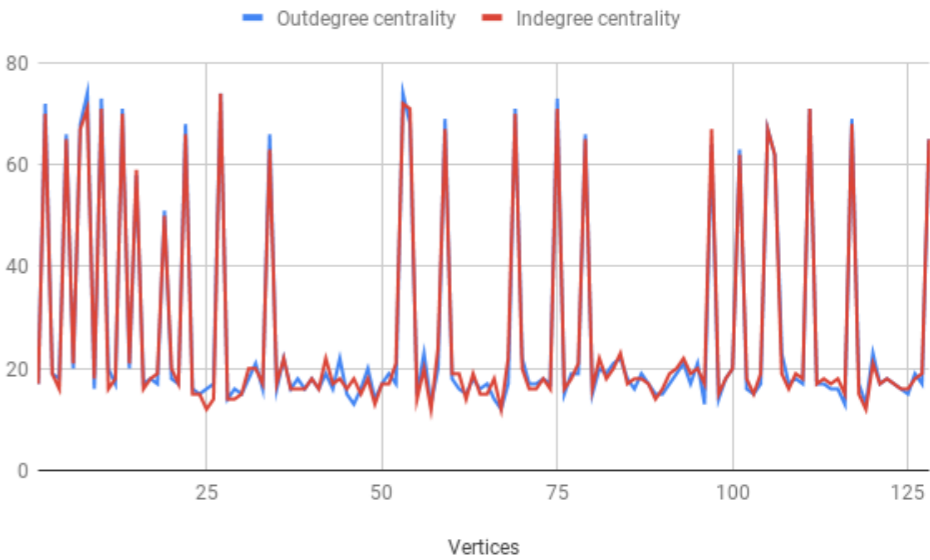
Examples of the development of a social network

$n=64, k=128$	Graph	<i>Outdegree and indegree centralities of the respective vertices</i>
<i>At the beginning condition</i>		
<i>At the final condition</i>		

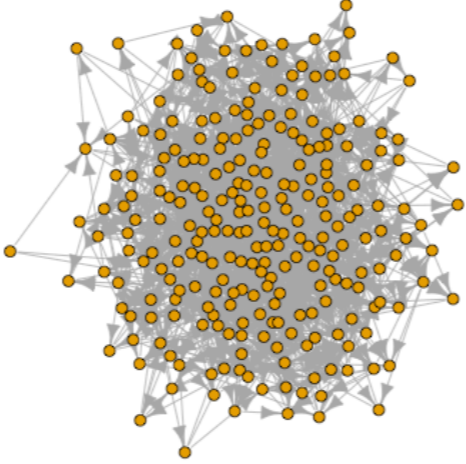
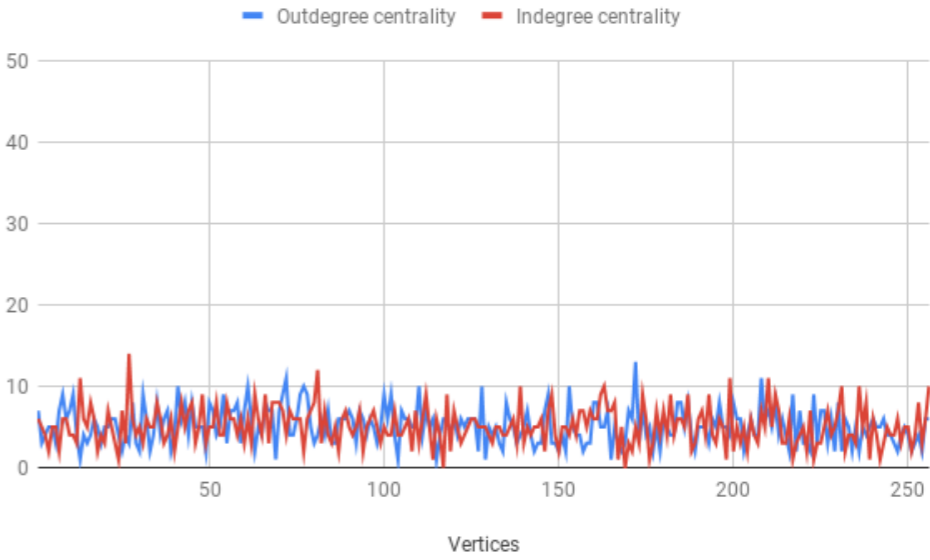
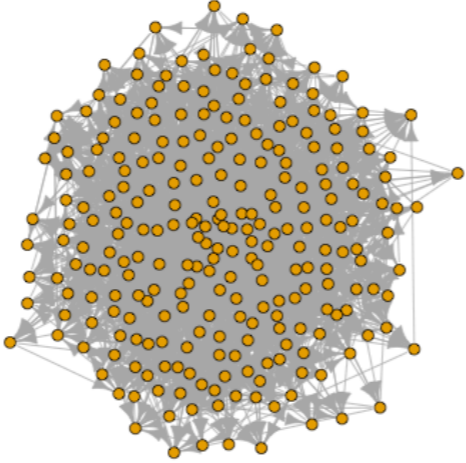
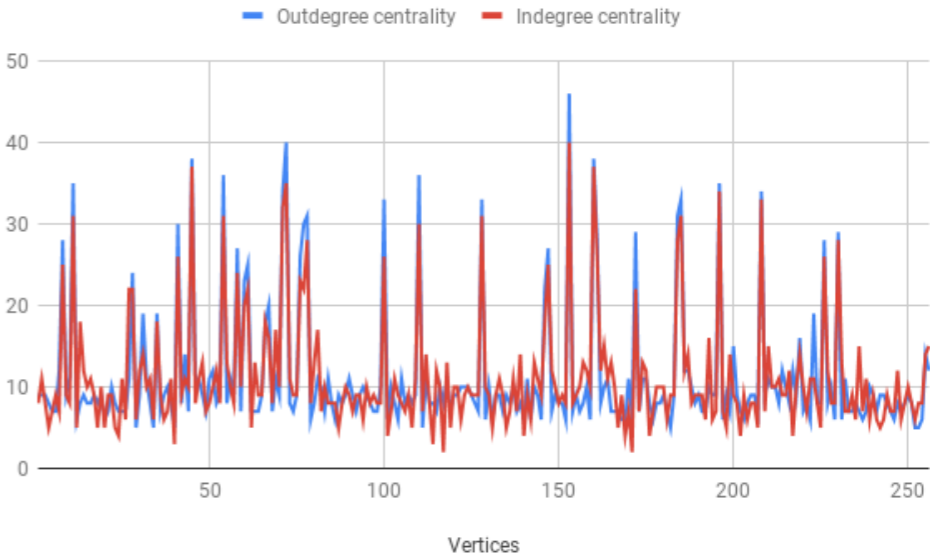
Examples of the development of a social network

$n=128, k=32$	Graph	<i>Outdegree and indegree centralities of the respective vertices</i>
<i>At the beginning condition</i>		
<i>At the final condition</i>		

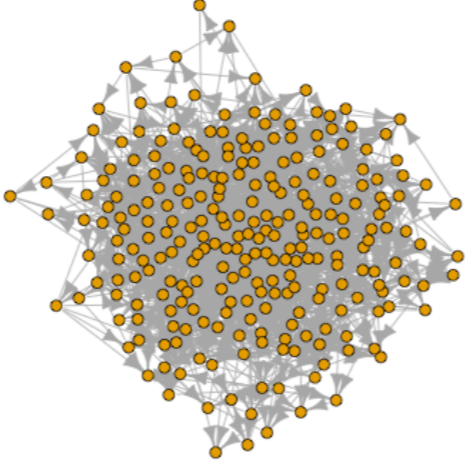
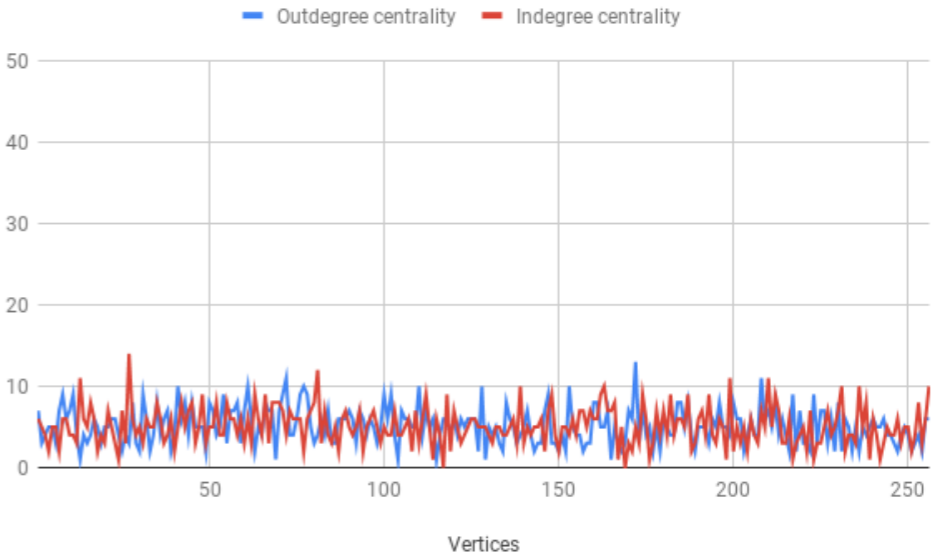
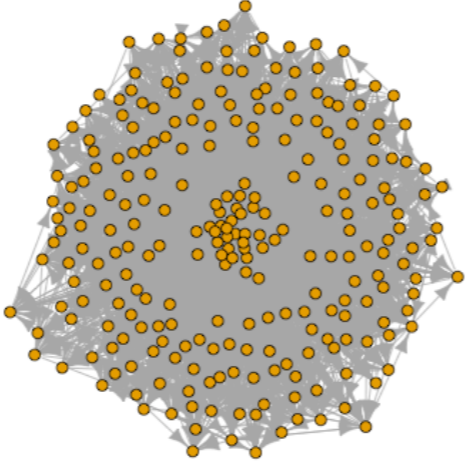
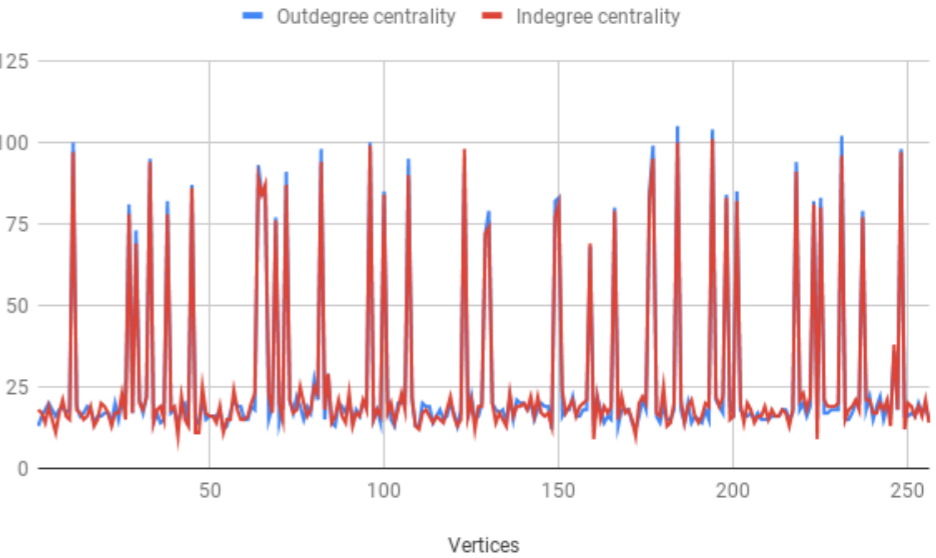
Examples of the development of a social network

$n=128, k=128$	Graph	<i>Outdegree and indegree centralities of the respective vertices</i>
<i>At the beginning condition</i>		
<i>At the final condition</i>		

Examples of the development of a social network

$n=256, k=32$	Graph	<i>Outdegree and indegree centralities of the respective vertices</i>
<i>At the beginning condition</i>		
<i>At the final condition</i>		

Examples of the development of a social network

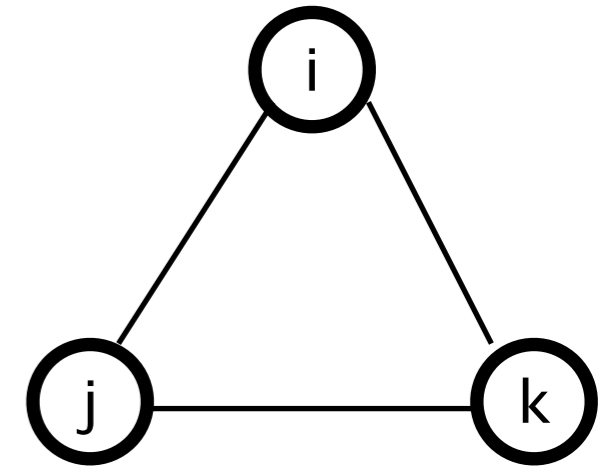
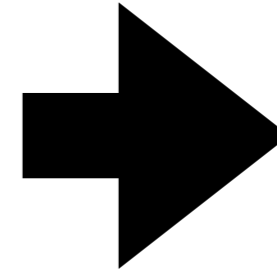
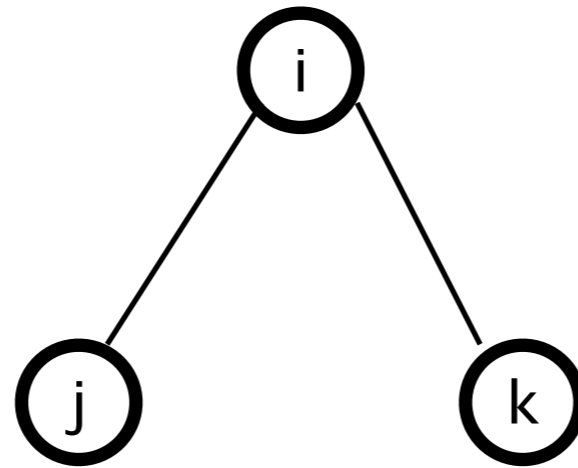
$n=256, k=128$	Graph	<i>Outdegree and indegree centralities of the respective vertices</i>
<i>At the beginning condition</i>		
<i>At the final condition</i>		

Examples of the development of a social network

- These examples suggest that an actor having very high outdegree centrality also has very high indegree centrality, and *vice versa*.
- This accords with our experience that a very limited number of ICT (Information and Communications Technology) giants collect increasingly more personal information.
- Also, this suggests that the medical and health records of a large number of people are collected and managed by a very limited number of PHR service providers.

- Suppose that every actor is connected with one another from a very short distance, that is, the social network is dense, or very close to a complete graph.
- It is very natural for the subscribers to be concerned that if any of the PHR service providers divulge or misappropriate the records of subscribers, their medical and health information is likely to be disseminated through the society very quickly.
- In contrast, if the social network is sparser, such dissemination of information will take longer and the PHR service providers or the government may be able to take counter measures to prevent such dissemination.

- In consideration of this, if we can prevent the social network from becoming denser (that is, in the context of social network analysis, to reduce the *transitivity* of the social network), such measure to mitigate the concerns of the potential and actual subscribers of PHR services

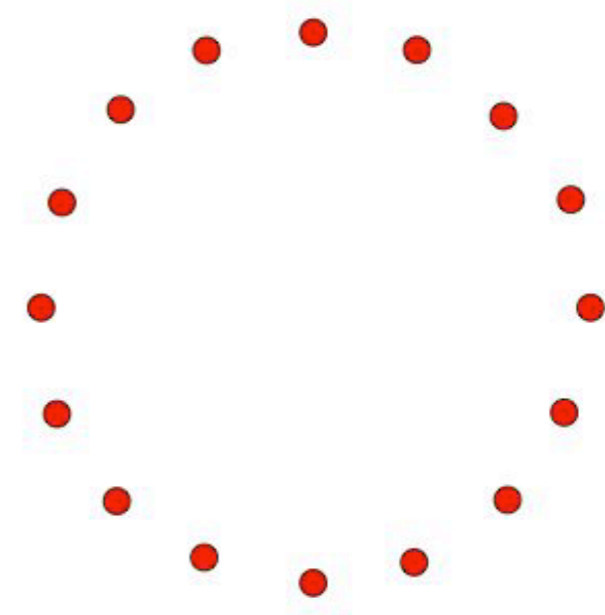


- Transitivity

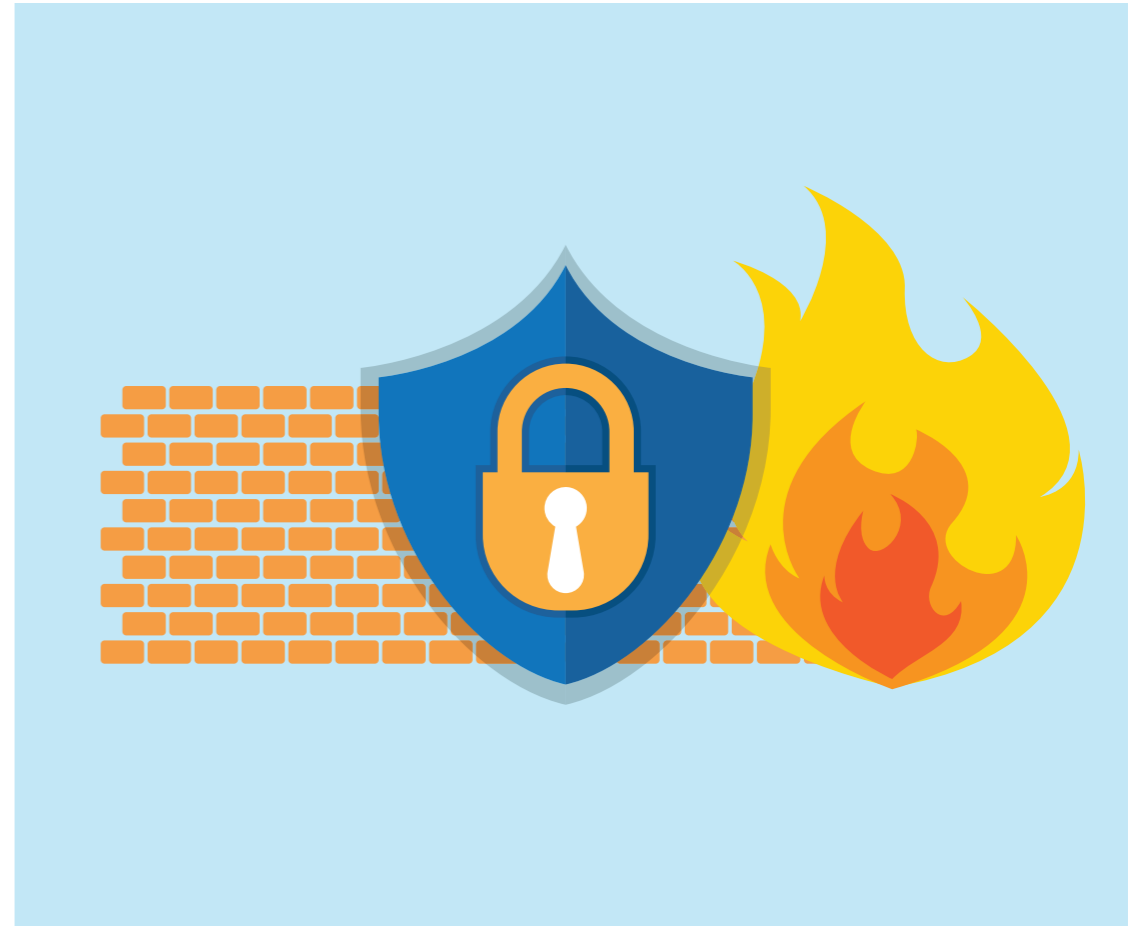
- Among three *vertices* i, j, k , if *vertex* i is connected with *vertex* j , which, in turn, is connected with *vertex* k , then it is very likely that *vertex* i will soon be connected with *vertex* k .

- In a network, once a triangle (three vertices connected with one another) is developed, triangles will increase, and the network will become incrementally closer to a complete graph.

- グラフ全体をみたときに, Distance = 2 の path (上の図だと, j と k をつなぐ path) のうち, それらの vertices の間に直接の tie (上の図だと, j と k を結ぶ tie) が存在するものの比率。

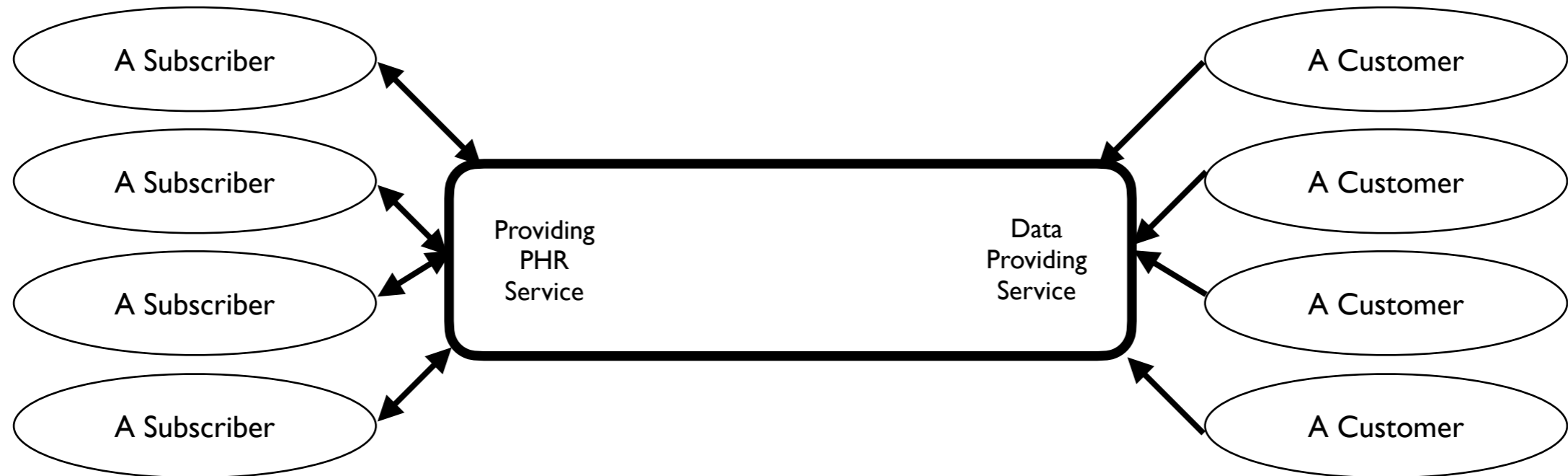


Installing a Firewall in the PHR Service Providing Company

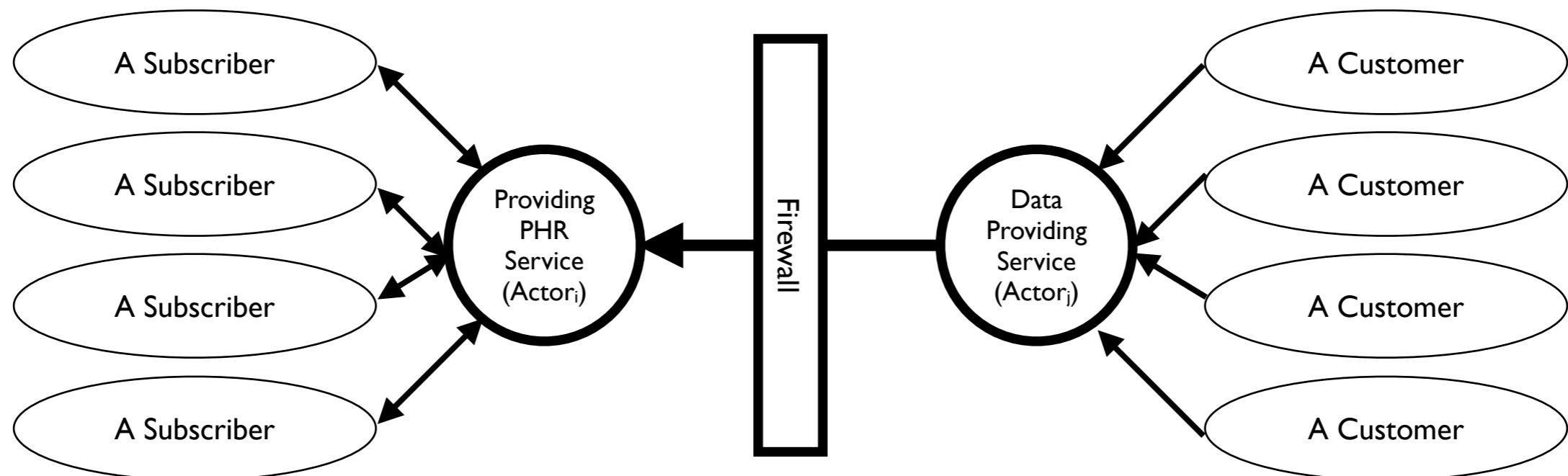


- Suppose that a PHR service providing company establishes an in-house firewall to restrict the exchange of personal information between the division providing the PHR service and other divisions.
- Suppose also that such a firewall can be represented by a slight change to the said resulting social network.
- If such change prevents the social network from becoming denser, it is likely to suggest that a firewall will mitigate the concerns of the subscribers of the PHR service.

Before installing a firewall.



After installing a firewall.

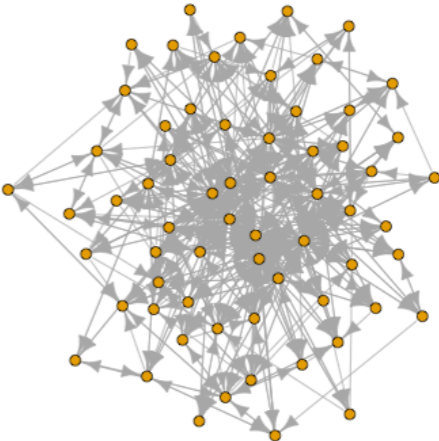
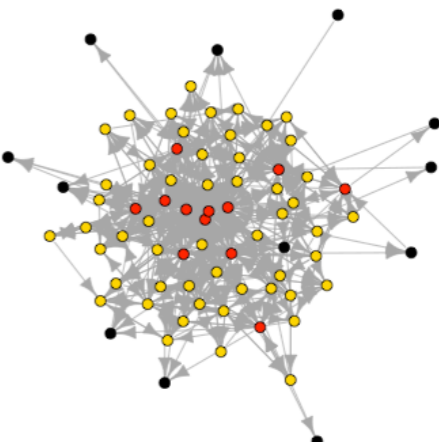


The direction shows the dependence of an actor on another actor.

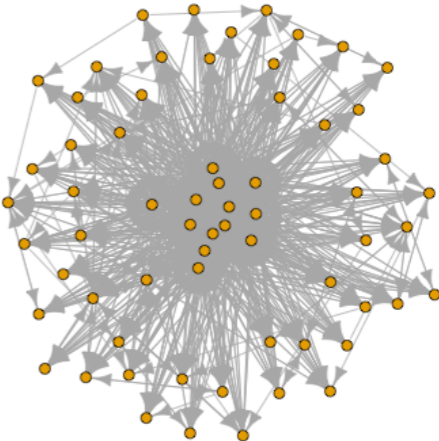
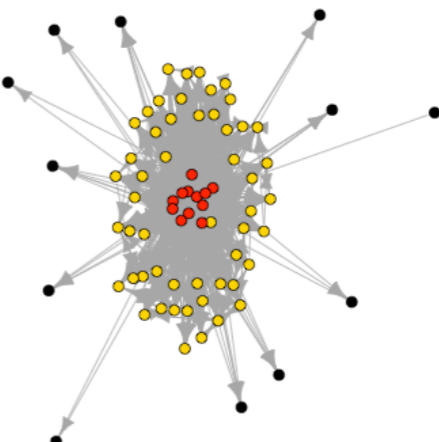
- A model of a firewall.
 - Suppose that $vertex_p$ represents a PHR service providing company.
 - Substitute $vertex_p$ by $vertex_i$ (representing the division providing PHR service) and $vertex_j$ (representing the other divisions).
 - Every bilateral tie connecting $vertex_p$ and other vertices is assigned to $vertex_i$, deeming that such ties are likely to represent the relationship between a PHR service provider and subscribers.
 - Every unilateral arcs sent to $vertex_p$ is assigned to $vertex_j$, deeming that such arcs are likely to be sent by commercial companies who purchase information derived from the medical and health information stored in PHR.
 - The said $vertex_j$ sends an arc to $vertex_i$ to access anonymized information generated from the medical and health information stored in PHR.

- Presumably, a PHR service providing company is well represented by a vertex having comparatively high indegree and outdegree centralities, because it collects medical and health information from a large number of subscribers, who, in turn, depend on PHR to access their records.
- Also, a commercial company collecting personal information is represented by a vertex having comparatively high outdegree centralities.
- In the network model used here, we cannot distinguish vertices representing PHR service providing companies from vertices representing other commercial companies collecting and/or purchasing personal information.

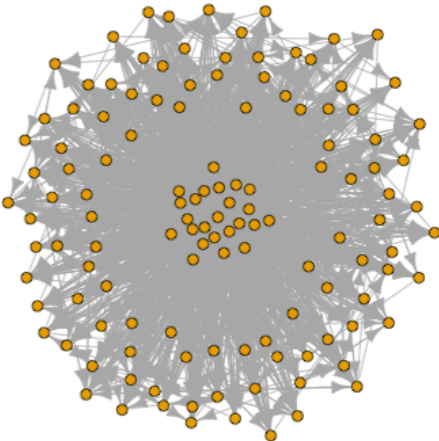
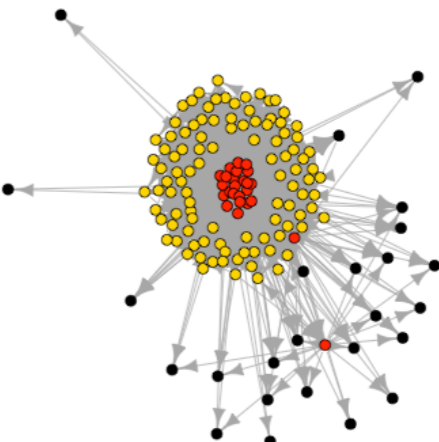
- For the purpose of convenience, in the Model discussed here, firewalls have been installed in the vertices whose outdegree centrality falls within the highest 20% of all the vertices.

$n=64, k=32$	<i>Average distance</i>	<i>Transitivity</i>	<i>Graph</i>
<i>At the beginning condition (a random graph)</i>	2.730215827	0.05960264901	
<i>Before installing firewalls</i>	2.020833333	0.1550802139	
<i>After installing firewalls</i>	2.325688889	0.07040229885	
	<ul style="list-style-type: none"> • <i>Each red dot denotes a division providing PHR service.</i> • <i>Each blue dot denotes the divisions of the same company but outside the firewall.</i> 		

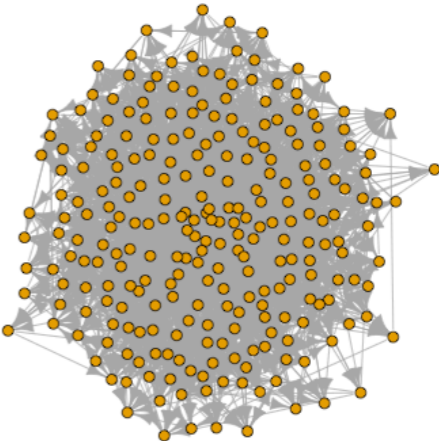
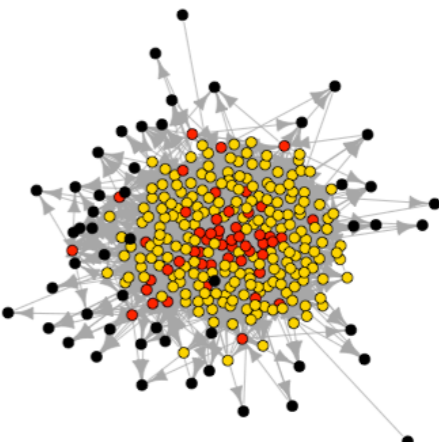
An example of the change in the average distance and transitivity as a result of installing the firewalls

$n=64, k=128$	<i>Average distance</i>	<i>Transitivity</i>	<i>Graph</i>
<i>At the beginning condition (a random graph)</i>	4.319554849	0.01851851852	
<i>Before installing firewalls</i>	1.694196429	0.2200397389	
<i>After installing firewalls</i>	1.879466667	0.1125309832	 <ul style="list-style-type: none"> • <i>Each red dot denotes a division providing PHR service.</i> • <i>Each blue dot denotes the divisions of the same company but outside the firewall.</i>

An example of the change in the average distance and transitivity as a result of installing the firewalls

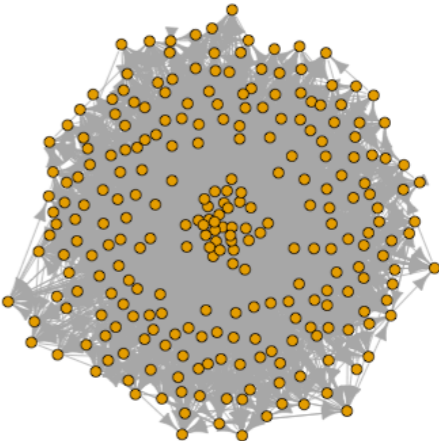
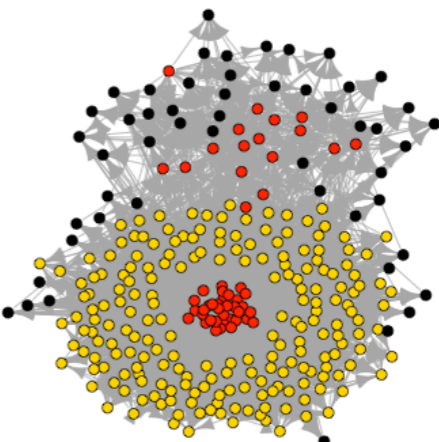
$n=128, k=128$	<i>Average distance</i>	<i>Transitivity</i>	<i>Graph</i>
<i>At the beginning condition (a random graph)</i>	4.976831091	0.04322200393	
<i>Before installing firewalls</i>	1.792937992	0.1507179904	
<i>After installing firewalls</i>	2.046525628	0.07946548589	
	<ul style="list-style-type: none"> • <i>Each red dot denotes a division providing PHR service.</i> • <i>Each blue dot denotes the divisions of the same company but outside the firewall.</i> 		

An example of the change in the average distance and transitivity as a result of installing the firewalls

$n=256, k=32$	<i>Average distance</i>	<i>Transitivity</i>	<i>Graph</i>
<i>At the beginning condition (a random graph)</i>	3.59036591	0.03777351248	
<i>Before installing firewalls</i>	2.464690564	0.06857625557	
<i>After installing firewalls</i>	2.786231651	0.04484947778	

- *Each red dot denotes a division providing PHR service.*
- *Each blue dot denotes the divisions of the same company but outside the firewall.*

An example of the change in the average distance and transitivity as a result of installing the firewalls

$n=256, k=128$	<i>Average distance</i>	<i>Transitivity</i>	<i>Graph</i>
<i>At the beginning condition (a random graph)</i>	3.600384468	0.04077148438	
<i>Before installing firewalls</i>	1.905238971	0.1147187046	
<i>After installing firewalls</i>	2.253752315	0.05787321271	<ul style="list-style-type: none"> • <i>Each red dot denotes a division providing PHR service.</i> • <i>Each blue dot denotes the divisions of the same company but outside the firewall.</i> 

An example of the change in the average distance and transitivity as a result of installing the firewalls

- Results
 - Installing firewalls in the actors having considerably high outdegree centrality (such as PHR service providers) decreases the average distance of the social network. That is, the firewalls make the social network sparser, which is likely to deter the dissemination of the information through the society.
 - Also, the installation of firewalls remarkably reduces the transitivity of the social network. That is, the firewalls reduce the likelihood of the social network becoming denser.
 - By this way, a firewall in a PHR service providing company is likely to mitigate the concerns of potential and actual subscribers of PHR services concerning the dissemination of their respective personal information through the society.

今回の問題

- 以上のモデルには、次のような、無視できない重大な欠点があります。これらを補うモデルを作ることを試みてください。
- あるアクター ($actor_i$) の、他のアクター ($actor_j$) に対する、依存関係について、 $actor_i$ が $actor_j$ 自身の個人情報にアクセスするための依存なのか、 $actor_j$ が有する第三者の個人情報にアクセスするための依存なのか、区別できていない。
- このモデルは、個人情報へのアクセスと、個人情報から得られた二次的な情報へのアクセスとを、区別できていない。例えば、PHRサービス事業者の、ファイアウォールの外側にある部門から情報を買う顧客企業があるとして、彼らが、個人情報そのものを購入しているのか、匿名化されたり、統計的な処理をほどこされたりした情報を購入しているのかが、わからない。

- The Model does not distinguish an actor's dependency on another actor to access the personal information of the latter actor from the former's dependency on the latter to access the personal information of the third parties collected by the latter actor.
- Also, the Model cannot distinguish an access to personal information from an access to generic information derived from personal information. For example, the Model does not tell us whether a firewall installed in a PHR service providing company permits the divisions outside the firewall to access personal information stored in PHR or generic or anonymized information derived from personal information.
- These drawbacks of the Model are likely to constitute the limitations of the discussions using the Model.

Thank you.

