What Makes a Firewall Fail?

寺本, 振透
Faculty of Law, Kyushu University : Professor

KYUSHU UNIVERSITY

What Makes a Firewall Fail?

2022 International Conference on Big Data and Artificial Intelligence (ICBDAI 2022)

September 24, 2022

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

Building big data is essential to enable A.I. to augment lawyers, governments, physicians, public health experts, etc.

- A.I. is the abbreviation of both Artificial Intelligence and Augmented Intelligence.

- Lawyers and physicians often emphasize that the final decisions are to be made by themselves and prefer to use Augmented Intelligence.

- Collecting individual cases, building big data consisting of them, and sharing such big data is essential to enable lawyers and governments, to design plausible laws, regulations or contracts, and practices thereunder.

- Building and sharing big data consisting of medical and health records of individual patients and citizens is essential in medical science and practice, as well as design and implementation of public health policy.

- Likewise, in order for A.I.s to become able to augment the considerations and decisions of lawyers, governments, physicians and public health experts, A.I.s have to learn big data.

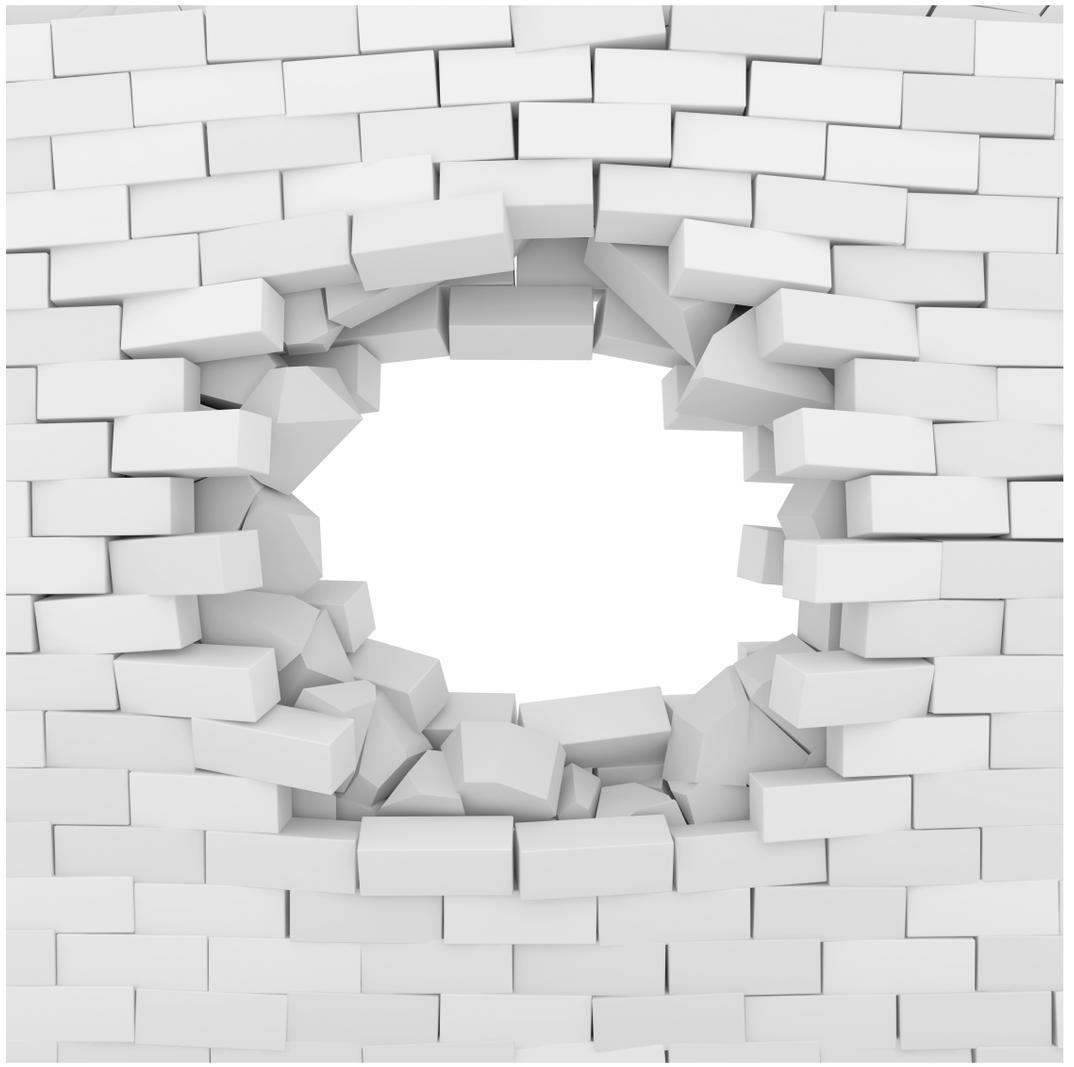
Building big data is likely to conflict with the protection of sensitive or personal information

- Raw records which are to be incorporated in big data often contain sensitive or personal information.

- In many jurisdictions, laws and regulations regulate the sharing of sensitive or personal information through multiple entities.

- Also, laws often entitle individuals to demand the governments and private entities not to share their personal information with other entities.

Firewalls are installed everywhere; and
Failure of a firewall is found everywhere



- In order to comply with laws and regulations, and to avoid liabilities to compensate for the damages suffered by the subject of sensitive or personal information, governmental and private entities install firewalls inside them.

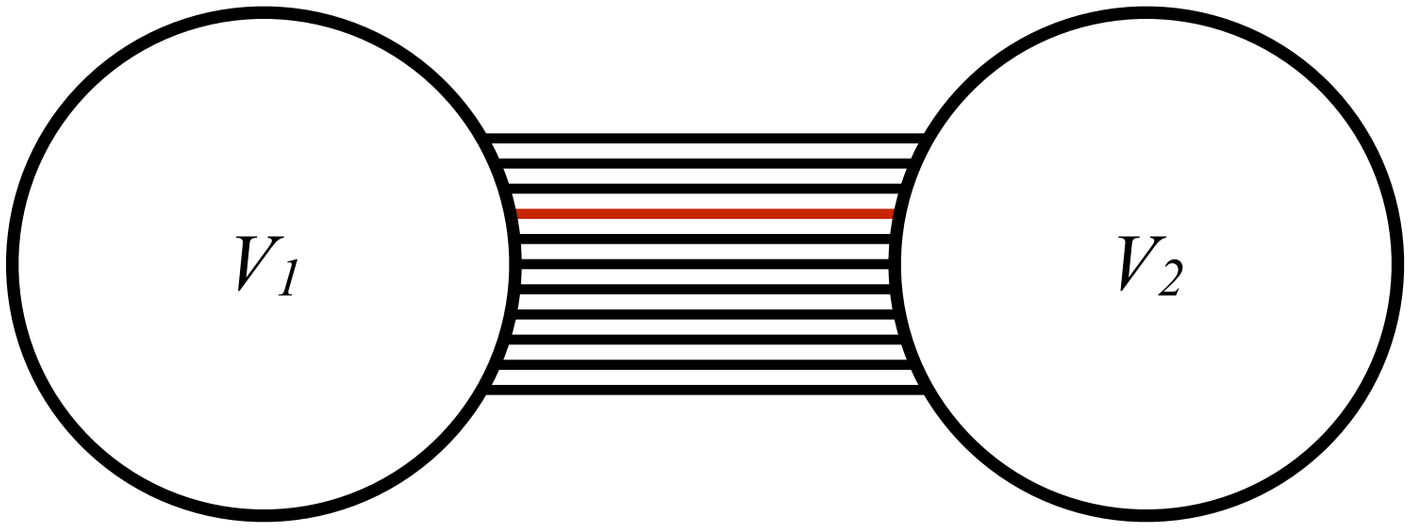
- By installing firewalls, we expect to prevent such information from incidental or intentional sharing with those who are not authorized to access.

- Unfortunately, we often experience failure of a firewall and resulting data leak, irrespective of whether the individual nodes inside and outside the firewall are human or machine.

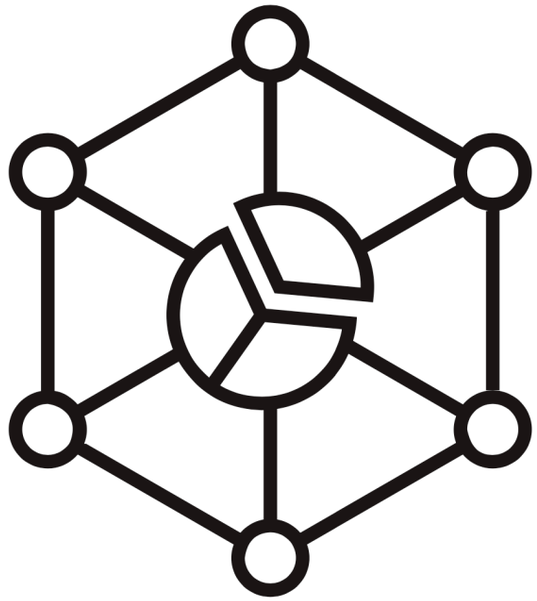
- A firewall is designed to cut off the edge (the red line, below) delivering or accessing sensitive or personal information from the nodes inside to the nodes outside the firewall.

- However, friendships, comradeships, and other edges between them which are not intended to deliver sensitive or personal information remains or are newly established even after the firewall is installed.

- It is very likely that an innocent edge may restore the cut-off edge or may be shifted to a harmful edge, and make the firewall fail.

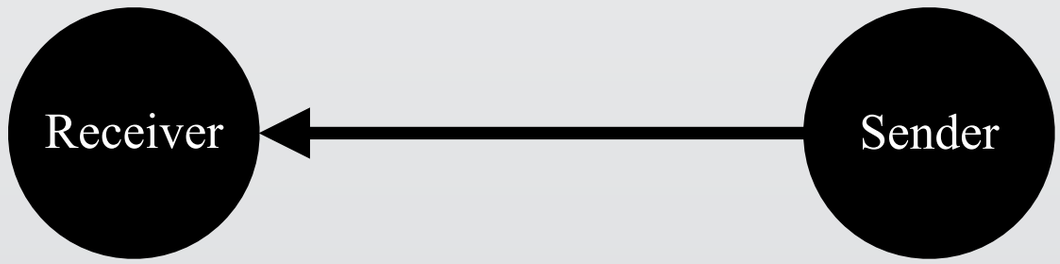


Examination by means of a simple model



- We assume the following to design our models to represent a social network in which a firewall is installed.
- These are very ordinary means to represent a network, although there are many other ways.
    - A model is represented by a graph and a matrix. A graph can be exchanged for a matrix, and vice versa.
    - The said graph is comprised of vertices (nodes) and arcs (i.e., directed edges).
    - A vertex denotes an individual actor acting in the social network.
    - An arc sent by a vertex and received by another vertex denotes the relationship of a pair of actors, one of which is represented by the sender, and another of which is represented by the receiver.

- An arc is a directed line connecting a pair of vertices.
- The actor denoted by the vertex sending an arc is referred to as the “sender,” and the actor denoted by the vertex receiving the same arc is referred to as the “receiver.”
- Assume that the direction of an arc represents the dependency of the sender on the receiver. That is, the sender depends on the receiver to access the specific information held by the receiver.

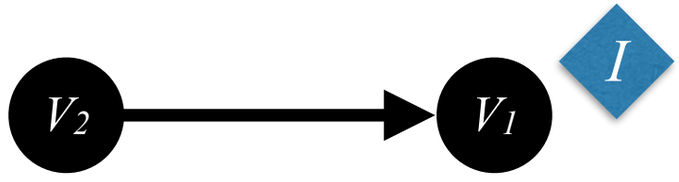


- A firewall is installed to cut off an existing arc between a pair of actors to prevent them from sharing specific kinds of information initially owned by only either of them.
- However, a firewall is not necessarily successful in cutting off a targeted arc.
- In order to represent the vulnerability of a firewall, we assume that each arc, which should have been cut off by a firewall, survives at the probability of p.

Event	Probability
The targeted arc survives the installation of the firewall.	p
The targeted arc is cut off by the firewall.	$1 - p$

Model 1

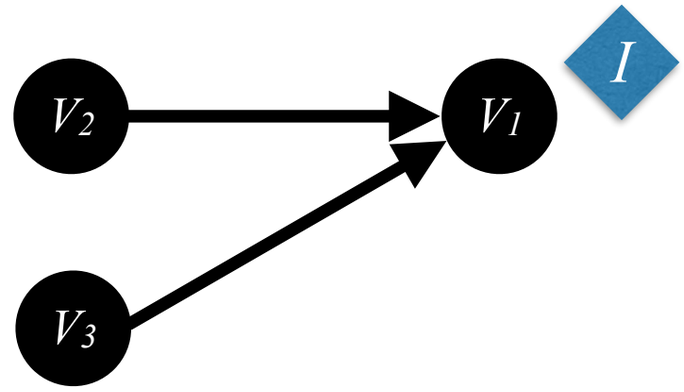
- Assumptions:
  - V_1 holds a piece of information ("I").
  - V_2 is sending an arc to V_1.
  - A firewall is installed to prohibit V_2 from accessing I by depending on V_1.
    - The probability that an arc survives after the installation of the firewall is p.
- The probability that I is available to V_2 by depending on V_1 after the firewall is installed is p.



Event	(V_2, V_1) survives	(V_2, V_1) is cut off
Probability	p	$1 - p$
Result	I is available to V_2 (failed firewall)	I is not available to V_1 (successful firewall)
Probability	p	$1 - p$
Suppose that $p = 0.2$	0.2	0.8

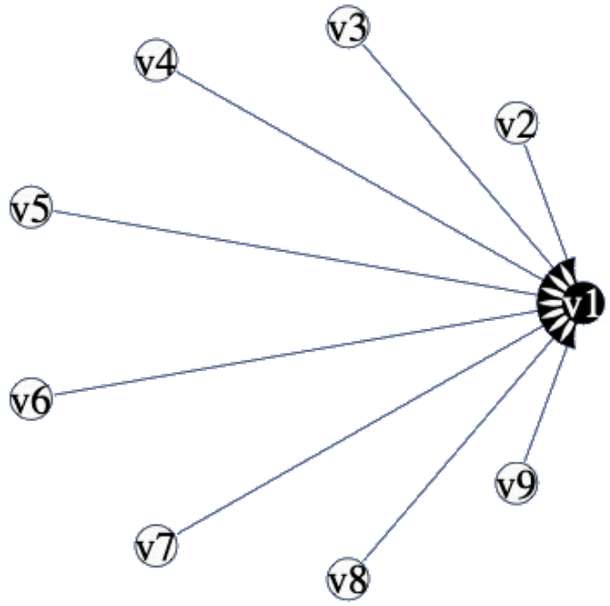
Model 2

- Assumptions:
  - V_1 holds I.
  - Each of V_2 and V_3 is sending an arc to V_1.
  - A firewall is installed to prohibit V_2 and V_3 from accessing I by depending on V_1.
  - The probability that an arc survives after the installation of the firewall is p.
- The probability that I is available to either or both of V_2 and V_3 by depending on V_1 after the firewall is installed is $1 - (1 - p)^2$.



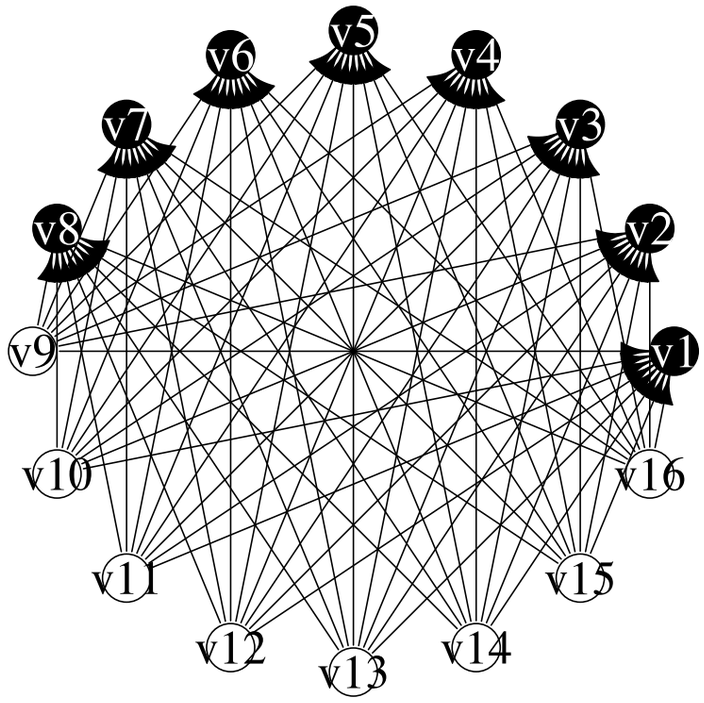
Event	(V_2, V_1) survives	(V_2, V_1) is cut off
Probability	p	$1 - p$
Event	(V_3, V_1) survives	(V_3, V_1) is cut off
Probability	p	$1 - p$
Result	I is available to V_2 and/or V_3 (failed firewall)	I is not available to V_2 and V_3 (successful firewall)
Probability	$1 - (1 - p)^2$	$(1 - p)^2$
Suppose that $p = 0.2$	*0.36*	*0.64*

Model 3



Result	failed firewall	successful firewall
Probability	$1 - (1 - p)^8$	$(1 - p)^8$
Suppose that $p = 0.2$	*0.832*	*0.167*

Model 4



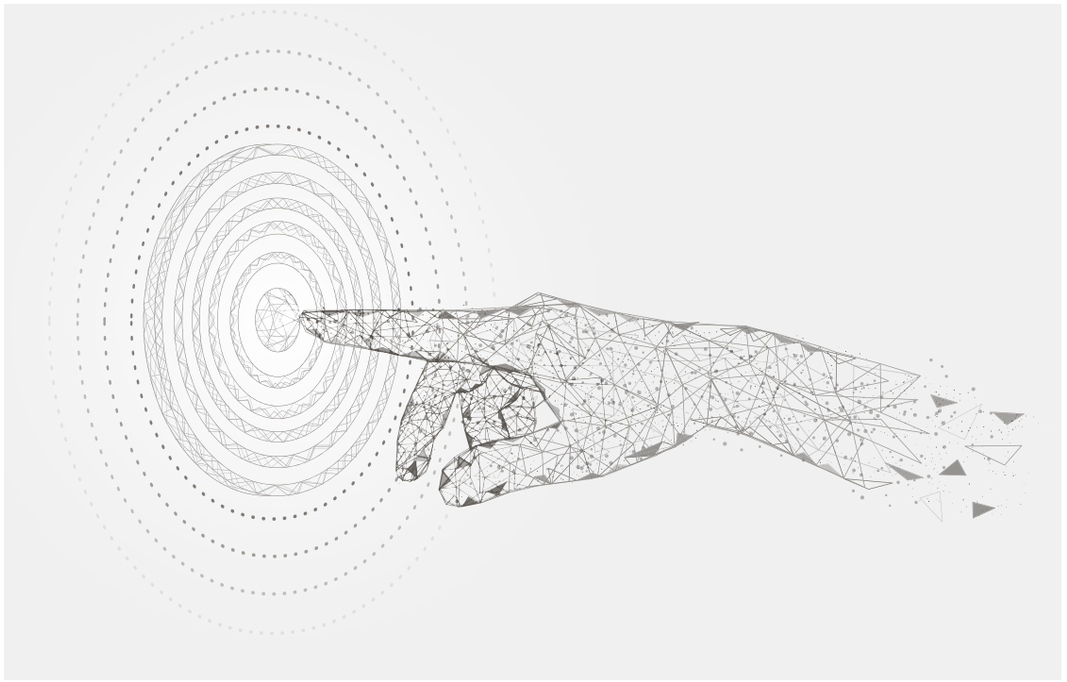
Result	failed firewall	successful firewall
Probability	$1-(1-p)^{8\times 8}$	$(1-p)^{8\times 8}$
Suppose that $p = 0.2$	0.999	6.27×10^{-7}

- The foregoing models have assumed that the vertices in the social networks are connected with one another before a firewall is installed.
- However, in reality, some of the vertices in and out of the firewall are connected before the installation of the firewall, while others are not connected.
- The foregoing models don’t take into account this reality.

- Presumably, the edges connecting between the individual nodes inside and outside of the firewall are one of the material causes which make the firewall fail.

- However, developing and maintaining mutual ties are inherent characteristics of intelligent nodes, human or machine.

The impact of the information held by the POIs



- Each of the nodes inside the firewall connected with one or more nodes outside the firewall can be deemed as a Point of Interface (“POI”).

- Suppose that we can minimize the probability that individual POIs have pieces of information which are usable by or meaningful to the nodes outside the firewall.

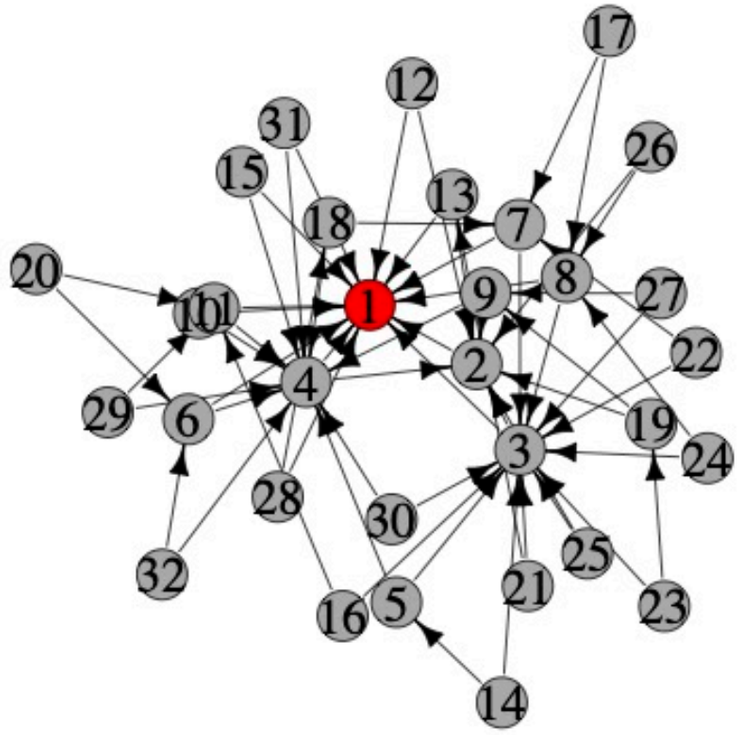
- Suppose also that a node outside the firewall has a scheme to access and utilize the pieces of sensitive or personal information maintained inside the firewall.

    - In order to realize his or her scheme, the node outside has to access as many nodes inside as possible, because individual nodes insides have only minimum pieces of information.

    - Or, in turn, the nodes outside having such a scheme have to cooperate with one another and aggregate the pieces of information which each of the nodes outside could collect from the nodes inside.

    - Both measures are likely to increase the cost born by the nodes outside significantly.

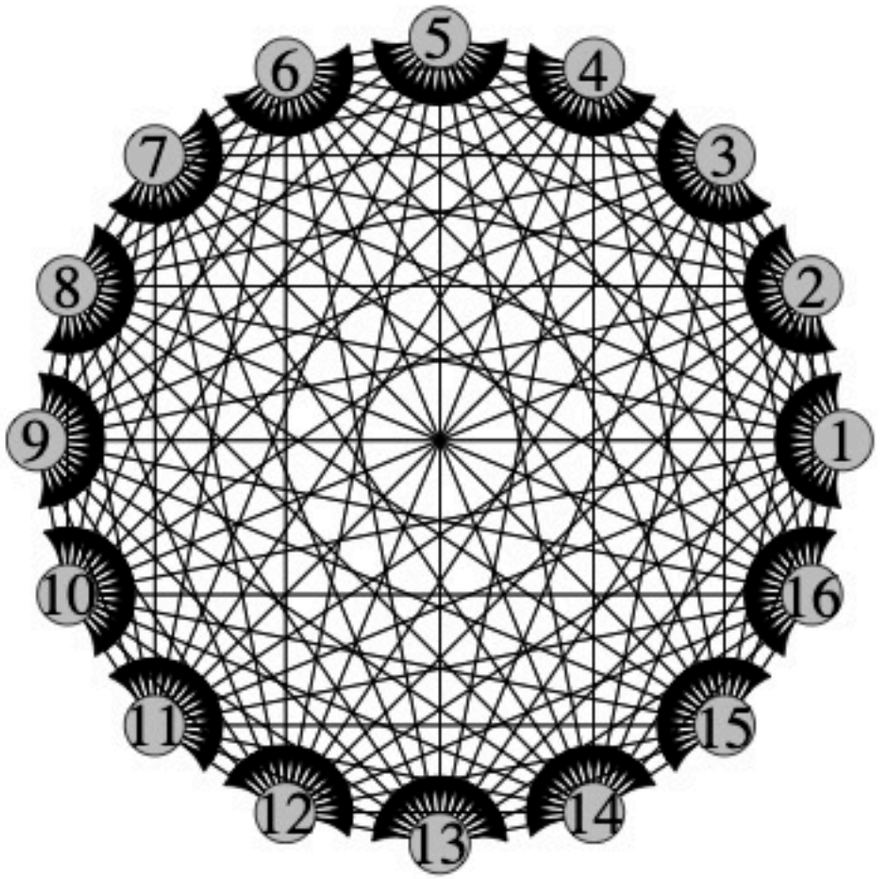
Attention to the network structure inside the firewall (ongoing study)



- Suppose that the pieces of information held by individual POIs can be effectively regulated.

- It may enable us to reduce the probability of the failure of a firewall.

- We anticipate that the geometric structure of the network among the vertices inside the firewall may affect the pieces of information held by individual POIs.



- A dense network (Left) is likely to help each POI to have more pieces of information, while a sparse network (Right) is likely to inhibit each POI from efficiently gaining information.

- However, the networks within the private or governmental entities are not so simple.

- Under this assumption, we have started examining networks of various geometric structures to find out which structure effectively regulate the information held by individual POIs, as well as which node(s) are likely to cause the failure of a firewall.

- Two mathematicians, Prof. Shizuo KAJI, Kyushu University, and Assistant Prof. Shota OSADA, Kagoshima University are joining our study.

Thank you.