

Till the Cows Come Home

牛歩のごとく

Carlos Valcarcel, Satoshi Koiso

1. Introduction (20pts)

- **Make an introduction in which you describe which data set you to study in the project. Follow specific suggestions of how to build the story (if any).**
- **Describe previous work on a similar or related dataset, what was the motivation of the paper, and what were their main results.**
- **Summarize what you will present in the next sessions of the paper.**

COVID-19 has raised our awareness of the risk of zoonotic disease outbreaks. This paper focuses on zoonotic diseases communicable to livestock animals which pose potentially significant risks. A pandemic of this type would cause serious issues not only in food supply chains but also in human mobility, bringing about economic downturns and hospital crises in the worst-case scenario. Thus, it is essential that we comprehend the risk based on the actual supply chain network structures and take preventive measures so we know how to act when the moment comes. Some studies assessed the livestock trade network in a region to reveal the network properties and the potential threats of disease outbreak (Mateus-Anzola et. al. 2019; Okello et. al. 2021). However, there are few studies that explored a large number of nodes and links representing a whole real supply chain network at a country level. For example, Mateus-Anzola et. al. studied the risk of influenza A virus at 260 pig farms in Lerma, Mexico, but the network connectivities assessed were experimentally generalized in simulation models. Okello et. al. analyzed the cattle trade network in Uganda to identify the most vulnerable cattle markets to *sleeping sickness*¹ transmission, but the number of nodes they analyzed was limited (n= 26 in two districts.)

This project examines the network structures of the Japanese livestock supply chain (from farm to slaughterhouse), and detects the potential risk of an outbreak of zoonotic diseases on the network. We collect cattle tracking data from the “Search service of Individual Identification Information of Cattle” provided by the National Livestock Breeding Center in Japan. Our research covers one of the most exhaustive cattle trade supply chain datasets. The paper is organized as follows. Section 2 describes the data collection and methods we used. Section 3 discusses the results of the analysis. Finally, Section 4 summaries and discusses the implications of this study and further considerations.

2. Data

Data collection

We collected cattle tracking data from the “Search service of Individual Identification Information of Cattle” provided by the National Livestock Breeding Center in Japan, where anyone can get information about individual cattle transfer data with 10-digit ID numbers attached to beef products.

¹ *Trypanosoma brucei rhodesiense*

Figure 1 shows an example of one cow's data. For each cattle identification number, we got the location with the date, prefecture, city, and location name for each movement they made. The figure is the English version of the site; we scraped data from the Japanese version of the site, which had more complete information (specifically in the “The name or title” field).

The screenshot displays the search results for a cattle identification number. The page title is "Search service of Individual Identification Information of Cattle". The search bar contains the ID "1385288462" and a "Search" button. Below the search bar, the results are organized into two main sections: "Individual Identification" and "Transfer information".

Individual Identification

Individual Identification Number	Date of birth	Gender	The Individual Identification Number of maternal parent (※1)	Breed
1385288462	Dec. 28, 2018	Female	1537209430	Japanese Black

Transfer information

Contents of transfer	Date of transfer	The prefecture where the cattle were raised		The name or title
		The prefecture	The location of the raising facilities	
1 Birth	Dec. 28, 2018	KAGOSHIMA, JAPAN	Nishinomote-shi	
2 Move-out	Oct. 06, 2019	KAGOSHIMA, JAPAN	Nishinomote-shi	
3 Carrying-in	Oct. 06, 2019	KAGOSHIMA, JAPAN	Nakatane-cho	Tanegashima livestock market
4 Sale	Oct. 06, 2019	KAGOSHIMA, JAPAN	Nakatane-cho	Tanegashima livestock market
5 Move-in	Oct. 06, 2019	KAGOSHIMA, JAPAN	Satsuma-cho	
6 Move-out	Aug. 25, 2021	KAGOSHIMA, JAPAN	Satsuma-cho	
7 Carrying-in	Aug. 26, 2021	TOKYO, JAPAN	Minato-ku	
8 Carrying-out	Aug. 27, 2021	TOKYO, JAPAN	Minato-ku	
9 Carrying-in	Aug. 27, 2021	TOKYO, JAPAN	Minato-ku	Shibaura Slaughterhouse Tokyo Metropolitan Government
10 Slaughter	Aug. 27, 2021	TOKYO, JAPAN	Minato-ku	Shibaura Slaughterhouse Tokyo Metropolitan Government

※1 As to the cattle born using the technique of embryo transfer, the cattle (recipient cattle) which delivered it is regarded as its maternal parent to be notified. In that case, the type or breed of the cattle may be different from that of its maternal parent.

Print To print this page, click this button to use the print function in the browser of the pop-up page.

[How to read the search result](#)
[Explanation about breed classification](#)

[The Beef Traceability System](#) [Individual Cattle Identification Register](#) [About This Site](#)

Privacy policy About copyright and linking
 copyright © National Livestock Breeding Center. All Rights Reserved

Figure 1. Sample data of “Search service of Individual Identification Information of Cattle”

The 10-digit ID numbers strictly follow a numbering rule to prevent the users of the “Search service of Individual Identification Information of Cattle” from registering incorrect ID numbers (Mie Prefecture, 2015). The first nine digits are divided into an odd-number digit group (1st, 3rd, 5th, and 7th digit) and an even-number digit group (2nd, 4th, 6th, and 8th digit). After summing each digit number in each group, the sum of the odd-number group is multiplied by three and added to the sum of the even-number group. The total number is divided by 10, and the remainder is subtracted from 10. The answer is the 10th digit (Figure 2).

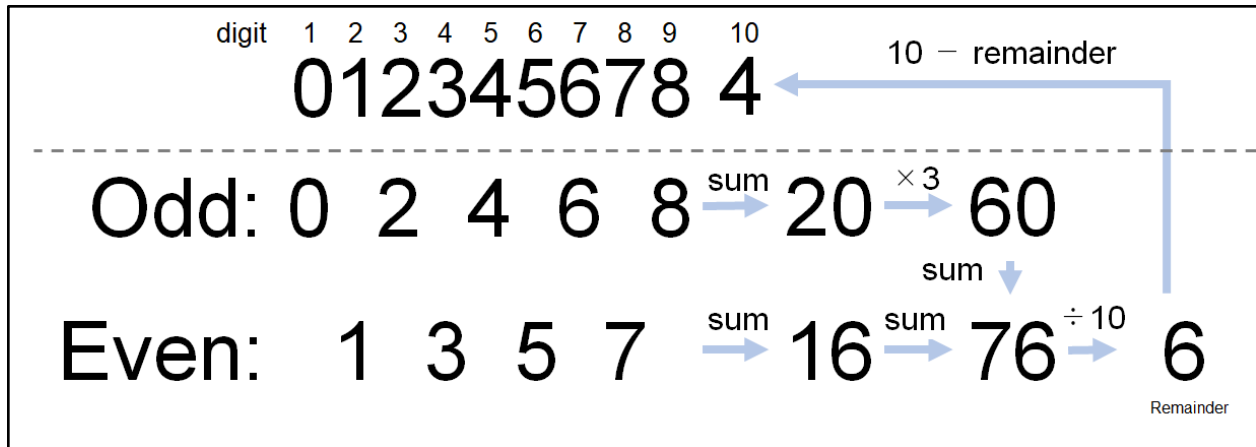


Figure 2. The numbering rule of 10-digit ID number

We randomly generated 10-digit ID numbers following the numbering rule and obtained 124,023 cows' tracking data so far. We also generated 190,859 non-existent IDs that we attempted to scrape. We tracked the number of new nodes and connections that we discovered every time we collected 5,000 real IDs, the figure below plots this progress. Once the marginal rate of information gain per ID approached zero, we understood that any additional data collected could help inform the weights across the network but rarely add new nodes or edges. We did not make use of weights in this analysis because they yielded unrealistic versions of the networks, we must collect more data until redundancies provide an accurate sample of the weights.

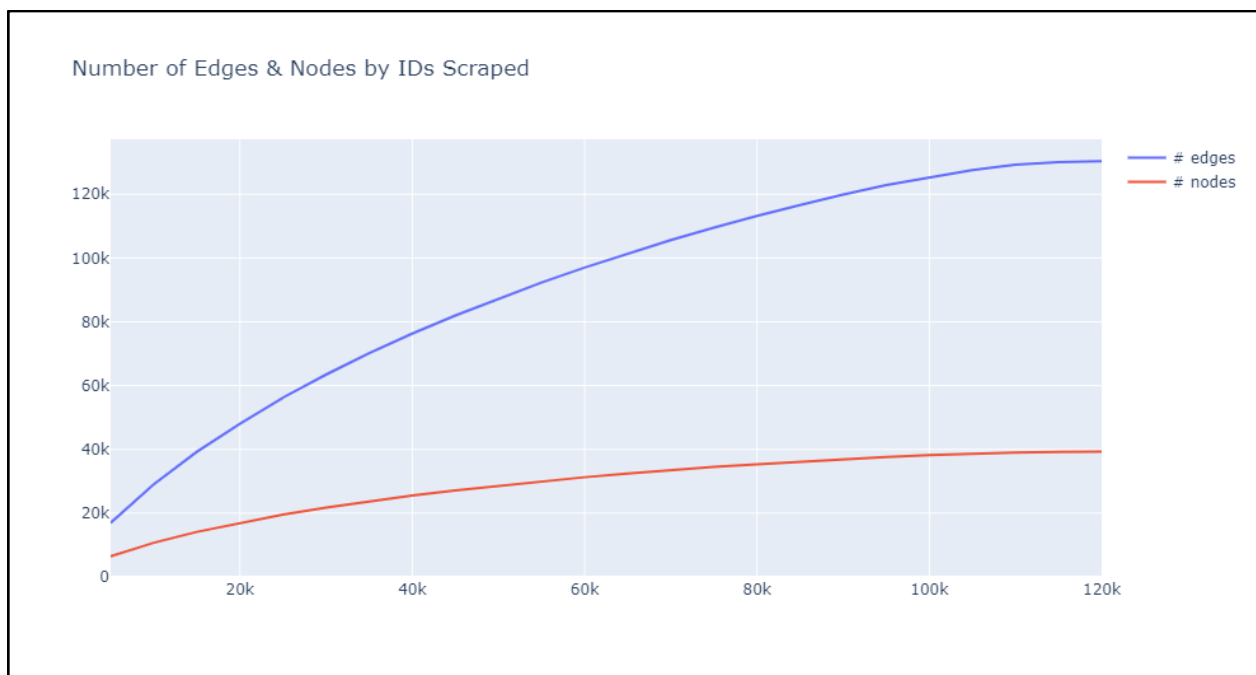


Figure 3. # Nodes (red) & # Edges (blue) plotted against # of real IDs collected

Some early checks we did on the data was check distribution of the Cows by prefecture. We confirmed that the prefecture where they started was not the prefecture where they ended. For example, few Cows start their tragic journey in Tokyo, but many end it there (Figure 4).

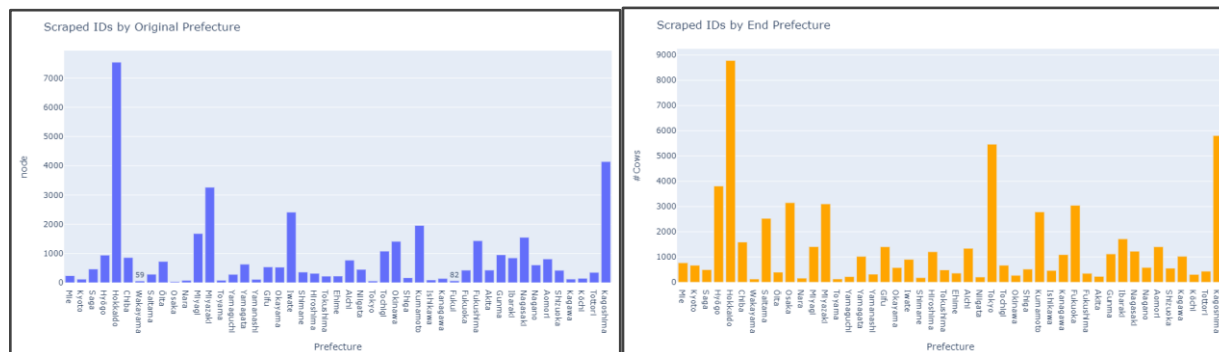


Figure 4. # of Ids by beginning Prefecture (left) vs. end prefecture (right)

3. Methods and Results (70 pts)

- Select 1 or 2 measurements that best capture your discoveries from the data. (40pts). (note: make 2 or 4 measurements for projects in groups). Submit your code.
- Describe which data properties you decided to measure and the motivation for this. Add a caption and number for each figure and make a reference to it in the text (be careful with Font and symbol sizes, and axes labels and legends). The more paper-like the results and the figures the better. (10pts)
- Fit at least one distribution or function and relate it to previous models or findings discussed in class (10pts)
- If it applies make a visualization of your dataset (10pt)

This section describes the methods and results of the supply chain network property analysis based on the collected data and the disease outbreak analysis. All methods used in this project are available at the projects Github, (<https://github.com/losDaniel/supply-chain-vulnerability>) and as a pip installable package using the command `pip install supply-chain-vulnerability`.

Sampling Method

An important decision was how we would use data that covered 20 years of activity to understand the supply chain at any given point. We created the following sampling strategies and implemented them in our project library:

- Sample by year: sample all the movements in a given set of years (e.g. [2015], [2015, 2016]) and build the network off of the nodes and connections discovered in those movements. Provides a sample of nodes and connections running in the sample time-period.
- Sample by slaughter year: sample the IDs of all cows slaughtered in a given set of years. Collect all the movements of those cows and build the nodes and connections discovered in those movements. Provides a sample of the nodes and connections traversed by the cows sold for consumption that year.

- Sample by origin year: sample the IDs of all cows born in a given set of years. Collect all movements of those cows and build the nodes and connections discovered in those movements. Provides a sample of the nodes and connections traversed by the cows sold for consumption that year.
- Sample by percentage: this method allowed us to sample the same number or percentage of cows or movements (in the case of sample by year) whenever we sampled. This was desirable because the number of nodes and edges discovered each year was correlated with the number of IDs we had managed to scrape for that year. Table 1 shows how using the full sample when sampling by year yielded heterogeneously sized networks, the disparity was reduced when sampling by slaughter year but only disappeared when we sampled the same number of IDs per year.

Sampled by Year				by Slaughter Year			Uniform Sample by Slaughter Year				
	Year	NumNodes	NumEdges		Year	NumNodes	NumEdges		Year	NumNodes	NumEdges
0	2011	8022	16699	0	2011	4659	11658	0	2011	4412	11008
1	2012	9275	19469	1	2012	5595	14087	1	2012	4551	11317
2	2013	9351	20564	2	2013	7723	19592	2	2013	4858	12091
3	2014	8518	18730	3	2014	8087	20547	3	2014	4771	11873
4	2015	7604	16769	4	2015	7628	19390	4	2015	4815	12003
5	2016	6982	15272	5	2016	6534	16484	5	2016	4719	11726
6	2017	6733	14610	6	2017	6100	15382	6	2017	4706	11711
7	2018	6158	13507	7	2018	5947	15171	7	2018	4641	11652
8	2019	5712	12502	8	2019	5862	14839	8	2019	4682	11647
9	2020	3365	7317	9	2020	5516	13971	9	2020	4619	11554
10	2021	2206	4653	10	2021	4857	12267	10	2021	4605	11601

Table 1. Number of Nodes and Edges Discovered by Sampling Method

We sampled by slaughter year; year by year from 2011 to 2021; we sampled 3,000 cows each year. This was slightly less than the minimum number of cows detected which were about 3,200 in 2011 and 2021. This provided us an understanding of the network that got cows to market each year. Giving us a comprehensive historical view of the nodes and edges that affected the products reaching the market every year.

Network Lifecycle Detection

We visualized the duration of cows' journey through the supply chain by calculating the number of movements that happened each month prior to the sample year. We divided the number of movements per month by the maximum number of movements per month associated with that sample year. We did this for every year and plotted the percentages in a graph (see Figure 5). The plot revealed when the cows sold each year had been moved.

This plot clearly showed that cows were often moved at the beginning of their lives, sometimes more than once. Then, most cows destined to be sold in a given year would spend several months in a farm growing and then be sold to slaughter in relatively tighter cohorts.

This result corresponds with the average age of cows on sale (Gotoh et al. 2014). Japanese Black cattle and Wagyu beef are slaughtered at the age of 2.5 on average. Milk cows and breeding cows are slaughtered at the age of 2 and 9, respectively (Kumamoto Prefecture Livestock Society. n.d.) which explains why some of the movements sampled traced well over three years before the sample year. The presence of older cows in the supply chain over time might be reflective of the introduction of younger breeding cows into the system which would align with trends in the introduction of genetically modified food.

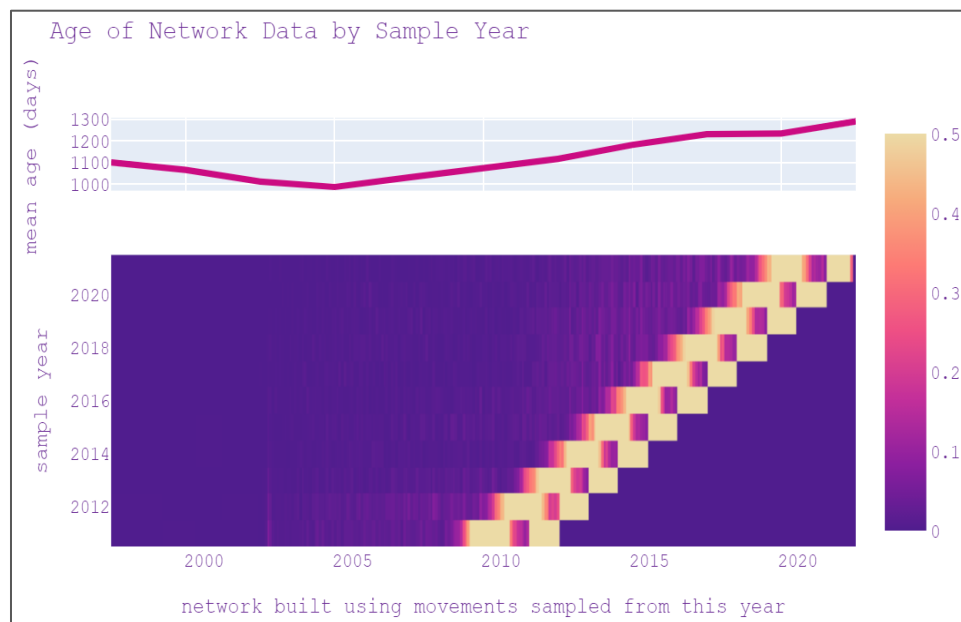


Figure 5. Age of network data by sample year

Network Node Detection

To evaluate the quality of our sample in representing the network over time we evaluated how many nodes persisted in the network year over year. Figure 6 shows the number of nodes in the network every year (persistent + vanishing). Persistent nodes were those that would be present the next year, vanishing nodes were those that would not. The incoming nodes category was added to show how many new nodes would be added each year. These proportions remained relatively constant regardless of whether we checked this relationship from one year to the next, every couple of years, or even every three years. When we checked, we found that the same nodes, those with the highest number of connections, were the ones that persisted most over time and that those vanishing or incoming nodes were usually small farms that raised cows.

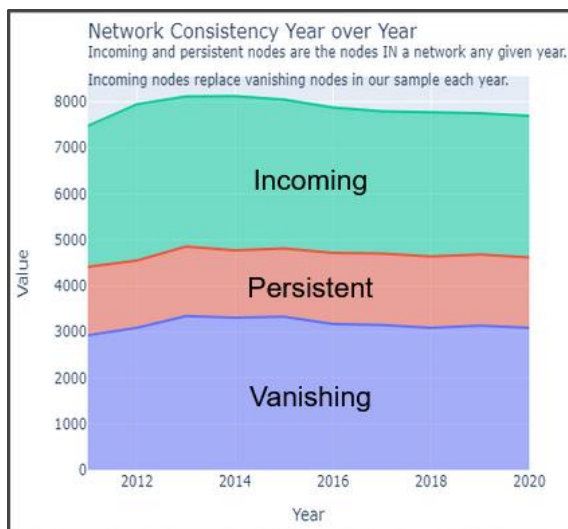


Figure 6. Network Consistency by year

Network Community Detection

We applied a Louvain community detection algorithm for each sample year in order to detect communities in the supply chain over time. Figure 7 shows that the number of communities in the network has been decreasing. In contrast, the number of nodes inside the largest community and the average number of nodes in communities have been increasing for the last several years.

This result resonated with the current Japanese livestock agriculture. There is a trend of farmers' ageing, retiring and selling their farms. This is leading to vertical integration and growing farm sizes (Ministry of Agriculture, Forestry and Fisheries. 2019).



Figure 7. Community Attributes Over Time

The Network Properties

In our network data, the nodes are cattle transfer points such as farms, distributors, and slaughterhouses. The links represent the tracking data from one transfer point to another transfer point. The weight would have been calculated by the transfer frequency of cows between nodes, however, the data for weights was too sparse so we decided to wait until more data is collected in the future to include this. The network is directed from farms to slaughterhouses which means it also covers several years as cows mature. Our network captures in-degrees and out-degrees evenly, and their distributions follow the power-law (Figure 8).

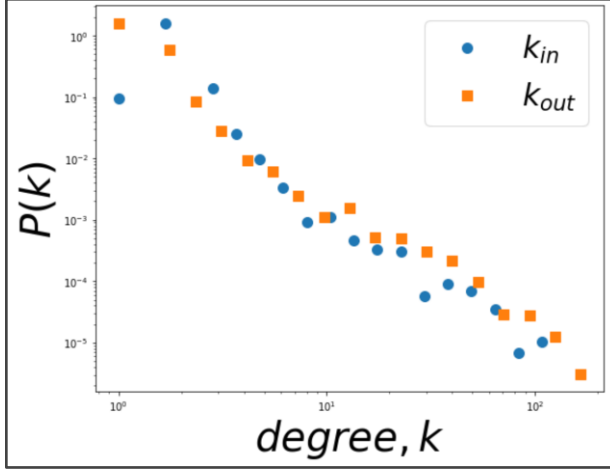


Figure 8. In-degrees and out-degrees

Based on the network properties analysis, we found that the network has properties similar to the small-world network and the Barabasi-Albert model. The network has a high clustering coefficient and a low average shortest path, which is the characteristic of the small-world network (Table 1). Given the log-log plot (Figure 9), the degree distribution is similar to the Barabasi-Albert model. These results show that the nodes are connected via a small number of links, and the network has few nodes connected with many nodes and many nodes connected with few nodes. In other words, this analysis shows that the supply chain network tends to reduce the number of cow transfers for cost reduction, and a few specific transfer points such as cow markets and slaughterhouses are the nodes where the most frequent transfers occur, however many nodes (farmers and distributors) enter the supply chain. Furthermore, the Barabasi-Albert degree distribution suggests growth and preferential attachment play a significant role in the network. This seems since concentration is increasing and many small producers only appear once or twice in the supply chains.

	# of nodes	# of link	<C>	<K>	<L>	disconnected	year
empirical network	4641	11652	0.037378	5.021332	4.602855	1	2018
small world (p=0.5787424185632628)	4641	9282	0.045261	4.000000	6.803894	0	2018
Barabasi-Albert (m=3)	4641	13914	0.009069	5.996122	4.051405	0	2018
Erdős-Rényi (p=0.0010821835365445913)	4641	23617	0.001002	10.177548	5.282886	0	2018

Table 2. Network properties comparison

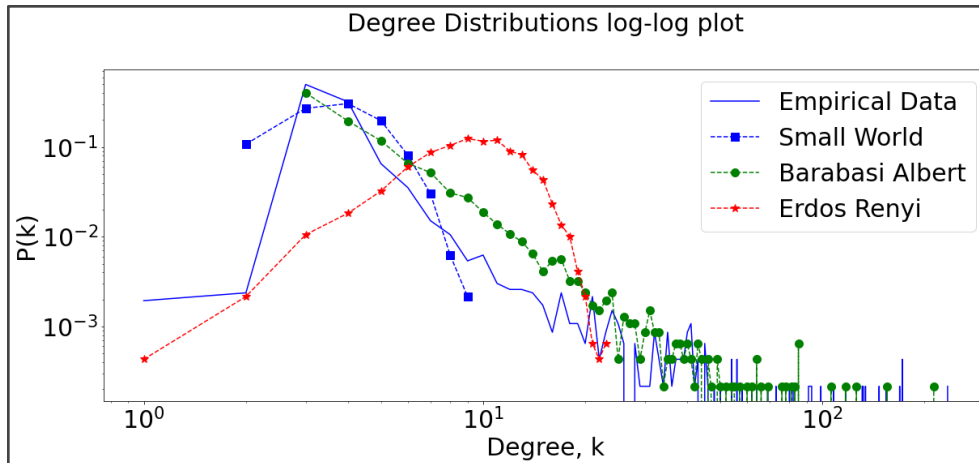


Figure 9. Log-log plot of the degree distribution of networks

Robustness Testing

We identified the top 5 nodes with the highest numbers of links and conducted robustness testing with random and targeted attacks to measure the vulnerability to disease outbreaks. Table 2 ranks the five nodes with the highest number of links. The result makes sense because the ranked places are among the biggest calf markets located in prefectures famous for their beef production, or a slaughterhouse located in the biggest beef consumption prefecture, Tokyo. (Agriculture & Livestock Industries Corporation. 2021; Ministry of Agriculture, Forestry and Fisheries. 2021).

Rank	Place	Number of Links	Prefecture	Type
1	全農岩手県本部中央家畜市場	1934	Iwate	Calf Market
2	熊本県畜産農業協同組合（熊本県家畜市場）	1868	Kumamoto	Calf Market
3	東京食肉市場（株）	1581	Tokyo	Slaughter House
4	全農みやぎ総合家畜市場	1463	Miyagi	Calf Market
5	ホクレン十勝地区家畜市場	1205	Hokkaido	Calf Market

Table 3. Top 5 nodes with the highest number of links

We tested the network robustness using random and targeted attacks to the network as of 2018. We assume that a certain number of nodes are removed in the network on the same day, defined as Day 0 of a disease outbreak (the first timing of attack with different numbers of nodes attacked). This is because, in the cow supply chain network in the real world, it is not realistic to assume a certain number of stop points are randomly damaged (infected) at each step from the geographical and epidemiological perspective although it might be possible if the disease is widely spread all over the country. Thus, we modified the model provided in the lecture to simulate the robustness with different numbers of nodes that are (randomly or purposefully) attacked only at Day0 (no step). The targeted attack was conducted in the order of the nodes with the highest number of links in increments of four.

Figure 4 shows that in the random attack from when 552 nodes are removed, there is a risk of supply chain breakdown. In the targeted attack, only 24 nodes removed can break down the supply chain.

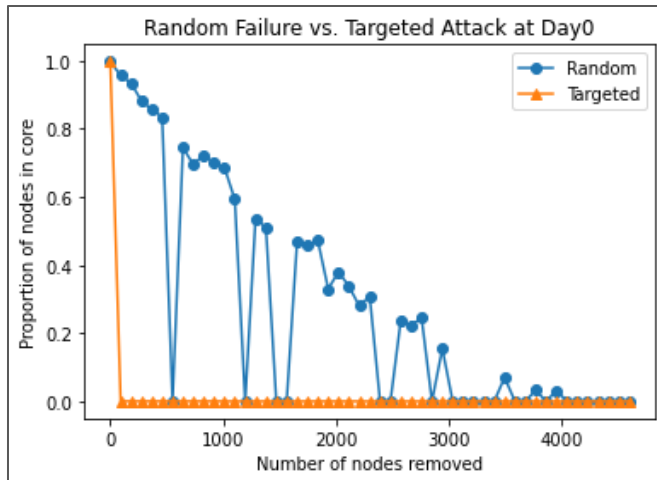


Figure 10. Robustness Testing (Random vs. Targeted attack)

SIS/SIR Model

Using the EoN package, we conducted the SIS and SIR model to the most centralized community in the network as of 2018 to simulate the disease outbreak in the community. We set R_0 as two as a realistic value, given that many infectious diseases have R -naught (R_0) ranging from 1.2 to 16.

Figure 11 visualizes the SIS and SIR model simulation results until time 6. The result of SIS shows that the infection lingers in the network for a long time, while the result of SIR shows that the infection peaks out at time 2 or 3 and wanes at time 5. This difference poses the limitation of applying the current models to the real-world network. It is arguable that the nodes (cows transfer places) are not immunized due to the ceaseless flows of different individual cows. In this sense, the SIS model is applicable. However, the sites might strengthen their preventive activities and inspection after outbreaks, which can be called “Recovered” in the SIR model. Also, we should carefully consider the meaning of time and the time lags among infection of pathogens, detection of symptoms, mobilization for disinfection. In particular, the rate of transmission between different types of nodes which we hope to address in future work. Despite these limitations, these models were built on directed graphs demonstrating that disease can widely spread inside the network.

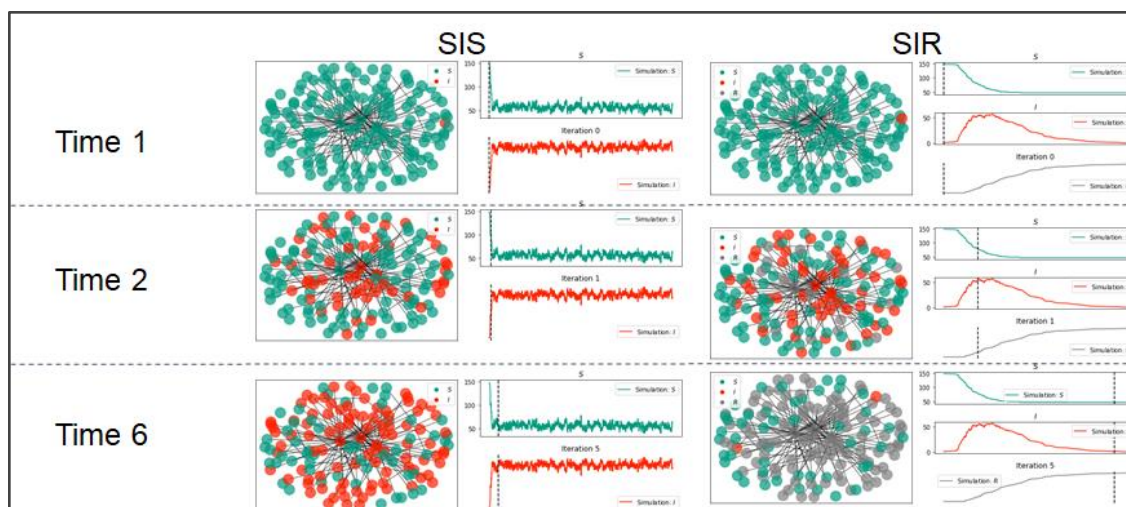


Figure 11. SIS (left) and SIR (right) model simulation to time 6

Contribution

Satoshi Koiso devised the project. Carlos Valcarcel worked out most technical details, from data collection to analysis, and developed a Python package specific to this project. The package is available online. Satoshi Koiso and Carlos Valcarcel interpreted the analysis results. Satoshi Koiso supported the implications backed by references in Japanese and English.

4. Future Work and Conclusions (10 pts)

- **Write some closing remarks motivated by your results, challenges and opportunities. Based on the related paper mention some avenues for future analysis related with the network you chose.**

Our network model well captures the current Japanese livestock agriculture industry trends. We revealed that the Japanese cattle supply chain network has the small-world and Barabasi-Albert network properties and has a significant risk of breakdown due to possible disease outbreaks. Some next steps we were able to begin but not conclude in the time afforded during this course has been the clustering of nodes by activity, duration and connections. This will allow us to evaluate how nodes with different purposes enter the supply chain at different times and customize simulations to include accurate times between transfers.

For further analysis, we will explore more realistic epidemic models such as the SEIR model, which includes the concepts of incubation periods, and might construct a new model. Also, we will investigate the correlation between the collected data and the market and census data for much detailed analysis. Possibly, we will utilize geolocation data with Google map API and Geopandas to reveal the at-risk communities on maps. Notwithstanding its limitations and room for improvement, this study is the first and the most exhaustive research on the Japanese cattle supply chain to shed light on this scientifically uncharted research area.

References

- Agriculture & Livestock Industries Corporation. (2021). *Niku-yo Koushi Torihiki Joho*. [Beef Calf Transaction Information (in English)] Retrieved December 12, 2021, from <https://www.alic.go.jp/operation/livestock/calf-report.html>
- T. Gotoh, H. Takahashi, T. Nishimura, K. Kuchida, H. Mannen. (2014). Meat produced by Japanese Black cattle and Wagyu, *Animal Frontiers*, Volume 4, Issue 4, October 2014, Pages 46–54, <https://doi.org/10.2527/af.2014-0033>
- Kumamoto Prefecture Livestock Society. (n.d.) *Ushi-no Issyou*. [Life Cycle of Cows (in English)]. Basic Information about beef cattle. Retrieved December 12, 2021, from http://kumamoto.lin.gr.jp/shokuniku/kisochisiki/kachiku_seisan/usi.html
- Mateus-Anzola, J., Wiratsudakul, A., Rico-Chávez, O., & Ojeda-Flores, R. (2019). Simulation modeling of influenza transmission through backyard pig trade networks in a wildlife/livestock interface area. *Tropical animal health and production*, 51(7), 2019–2024. <https://doi.org/10.1007/s11250-019-01892-4>
- Mie Prefecture. (2015). *Jihyo-no 10 Keta-me*. [The 10th digit of the cows' ear tags(in English)]. Retrieved December 12, 2021, from <https://www.pref.mie.lg.jp/MSHOKU/HP/86189042719.htm>
- Ministry of Agriculture, Forestry and Fisheries. (2019). FY2018 Summary of the Annual Report on Food, Agriculture and Rural Areas in Japan. Retrieved December 12, 2021 from https://www.maff.go.jp/j/wpaper/w_maff/h30/attach/pdf/index-28.pdf
- Ministry of Agriculture, Forestry and Fisheries. (2021). *Reiwa Ninen Chikusan-butsu Ryutsu Tokei*. [FY 2020 Livestock Product Market Statistics (in English)]. Retrieved December 12 2021 from <https://www.e-stat.go.jp/stat-search/files?page=1&layout=datalist&toukei=00500227&tstat=000001044816&cycle=7&year=2020&month=0&tclass1=000001044818&tclass2=000001156338>
- Okello, W. O., Amongi, C. A., Muhanguzi, D., MacLeod, E. T., Waiswa, C., Shaw, A. P., & Welburn, S. C. (2021). Livestock Network Analysis for Rhodesiense Human African Trypanosomiasis Control in Uganda. *Frontiers in veterinary science*, 8, 611132. <https://doi.org/10.3389/fvets.2021.611132>