Japan-Korea Micro Geo Data International Symposium (December 1st, 2015)

Ritsu Sakuramachi, Yuki Akeyama
(Shibasaki Lab., The University of Tokyo)

outline

- What is the inter-firm transaction big data?
- Introduction of studies using the inter-firm transaction big data ① (Akeyama)
- Introduction of studies using the inter-firm transaction big data and GPS big data ② (Sakuramachi)

What is the inter-firm transaction big data?

This data contains information of

- **1,600,000 firms (1996-2013)**
- **4,300,000 transactions (2008-2013)**

Data Producer: Teikoku Databank, LTD (TDB)

(Largest credit research company of Japan)

Japanese largest transaction data developed by questionnaire surveys and interviews for firms by 1700 investigators with firm credit check

Inter-firm transaction big data

Inter-firm transaction data (2008-2013)

Database of approx. 4.3 millions Inter-firm transactions with following attributes

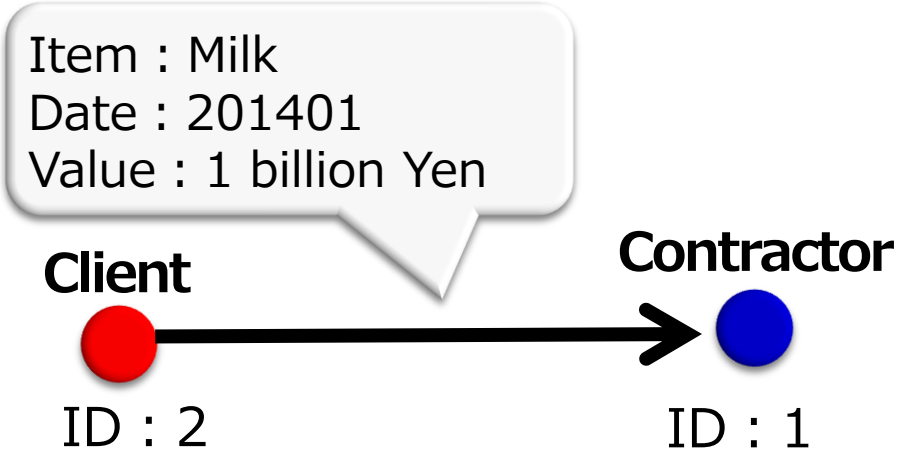
- Client and contractor firm ID
- Transaction item
- Transaction date
- Estimated transaction value[1)]

Firm data (1996-2013)

Database of approx. 1.6 million firms involved in transactions with following attributes

- Enterprise ID
- Sales amount
- Address
- Capital
- Business category

etc…



ID : 1　Sales amounts : 5 billion Yen
Address : Nemuro city, Hokkaido
Business : Milk product　etc…

ID : 2　Sales amounts : 20 billion Yen
Address : Fukuoka city, Fukuoka
Business : Wholesale trade etc…

1) Tamura, K., Miura, W., Takayasu, M., Takayasu, H., Kitajima, S., and Goto, H., 2012, "Estimation of Flux Between Interacting Nodes on Huge Inter-firm Networks", International Journal of Modern Physics: Conference Series, 16,93–104.

Inter-firm transaction big data

Inter-firm transaction data (2008-2013)

Database of approx. 4.3 millions Inter-firm transactions with following attributes

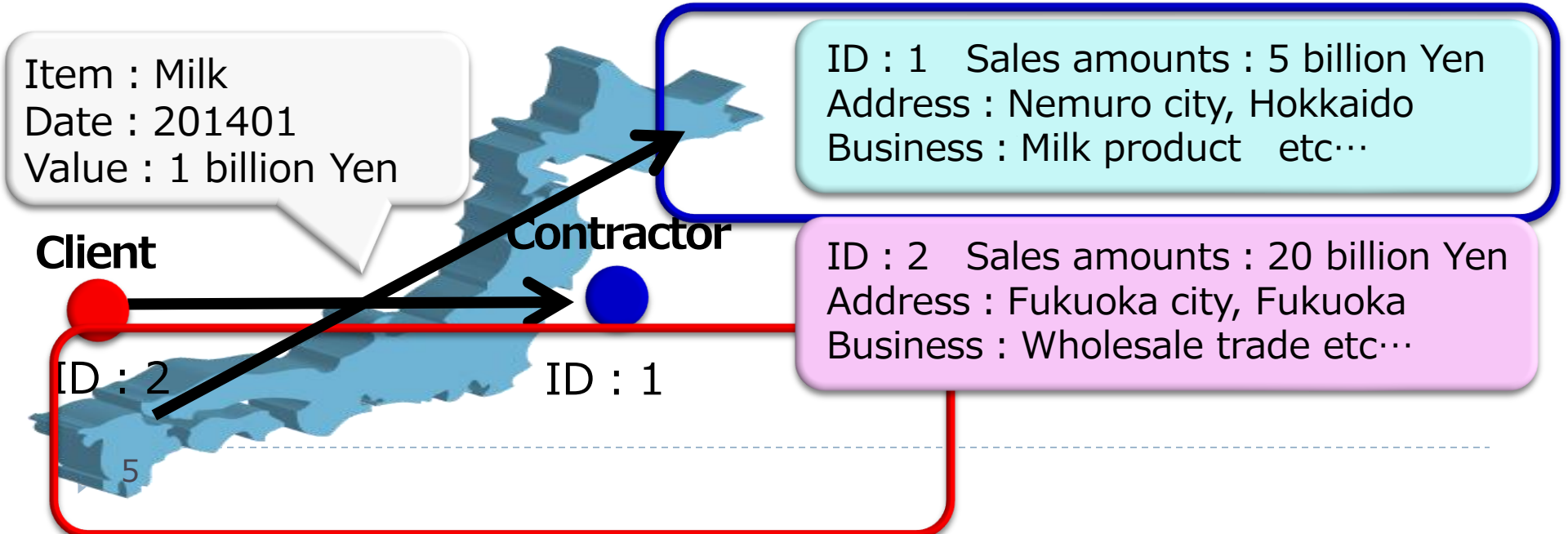
- Client and contractor firm ID
- Transaction item
- Transaction date
- Estimated transaction value[1)]

Firm data (1996-2013)

Database of approx. 1.6 million firms involved in transactions with following attributes

- Enterprise ID
- Sales amount
- Address
- Capital
- Business category

etc…



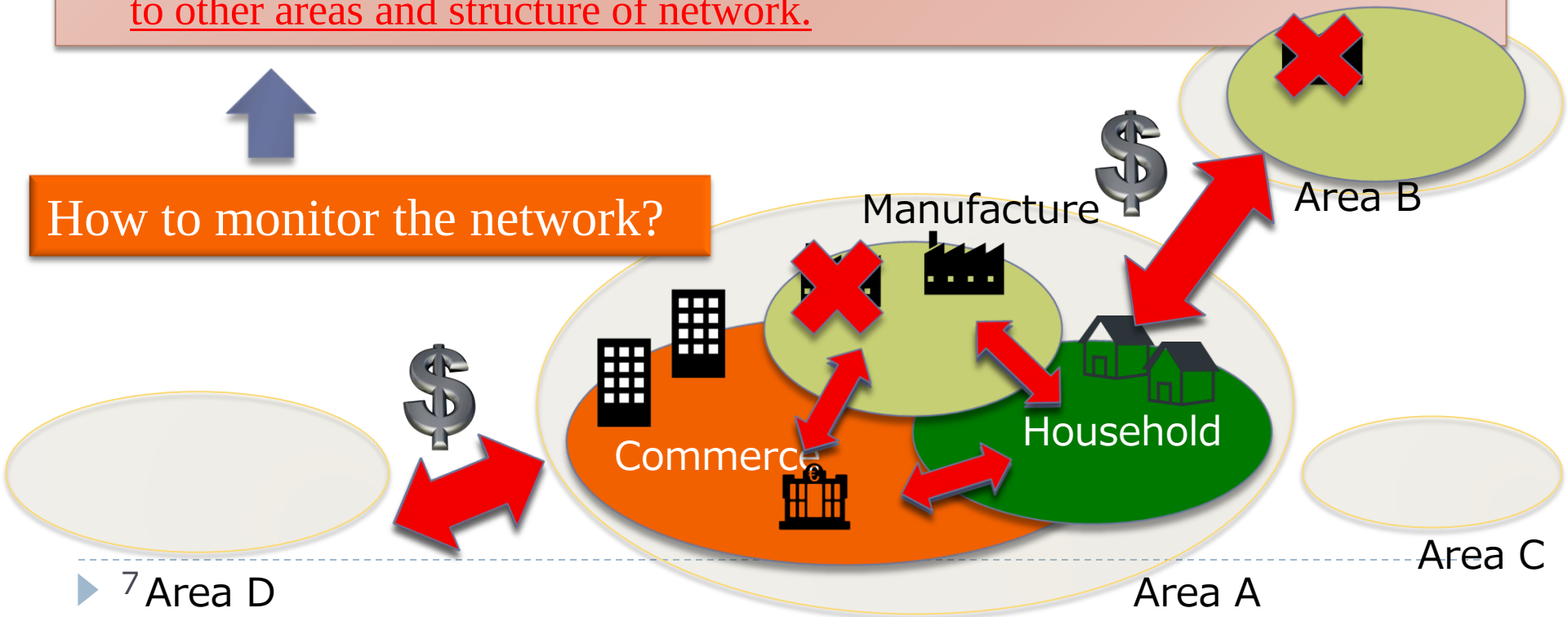
Visualization of network between regions using large-scale business transaction data

Yuki Akeyama – Civil Engineering, U.T.
toake555@iis.u-tokyo.ac.jp
Yuki Akiyama – CSIS, U.T.
Ryosuke Shibasaki – CSIS, U.T.

Background

- Increase or decrease of firm sales ⇒ impact to trading partners
  ⇒ impact to consumption of retails in the area where firms are locate
- Example：Great East Japan Earthquake(2011) happened…
  ⇒ Damage to car component suppliers
  ⇒ decline of sales of car manufactures or other component suppliers
- it's necessary to understand how each local area have economical relationship to other areas and structure of network.

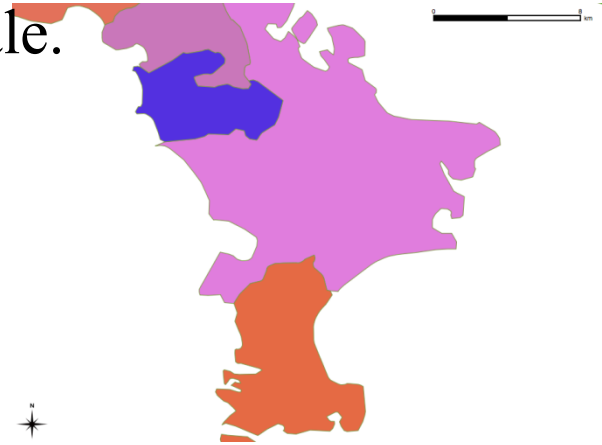
How to monitor the network?



Economic analysis using Big Data

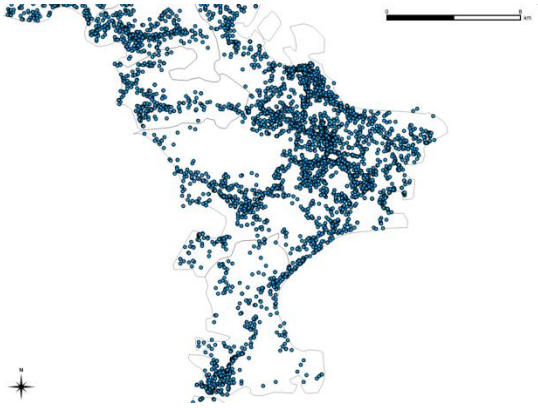
Previous study ：using governmental statistics or I/O table

Problem： the data is aggregated in administrative area.
⇒ It's difficult to analyze in area of smaller scale.



This study：using large-scale transaction data

Non-aggregated data
⇒we can analyze economy in "various scale".
（city block ・damaged area・ commutable area…）

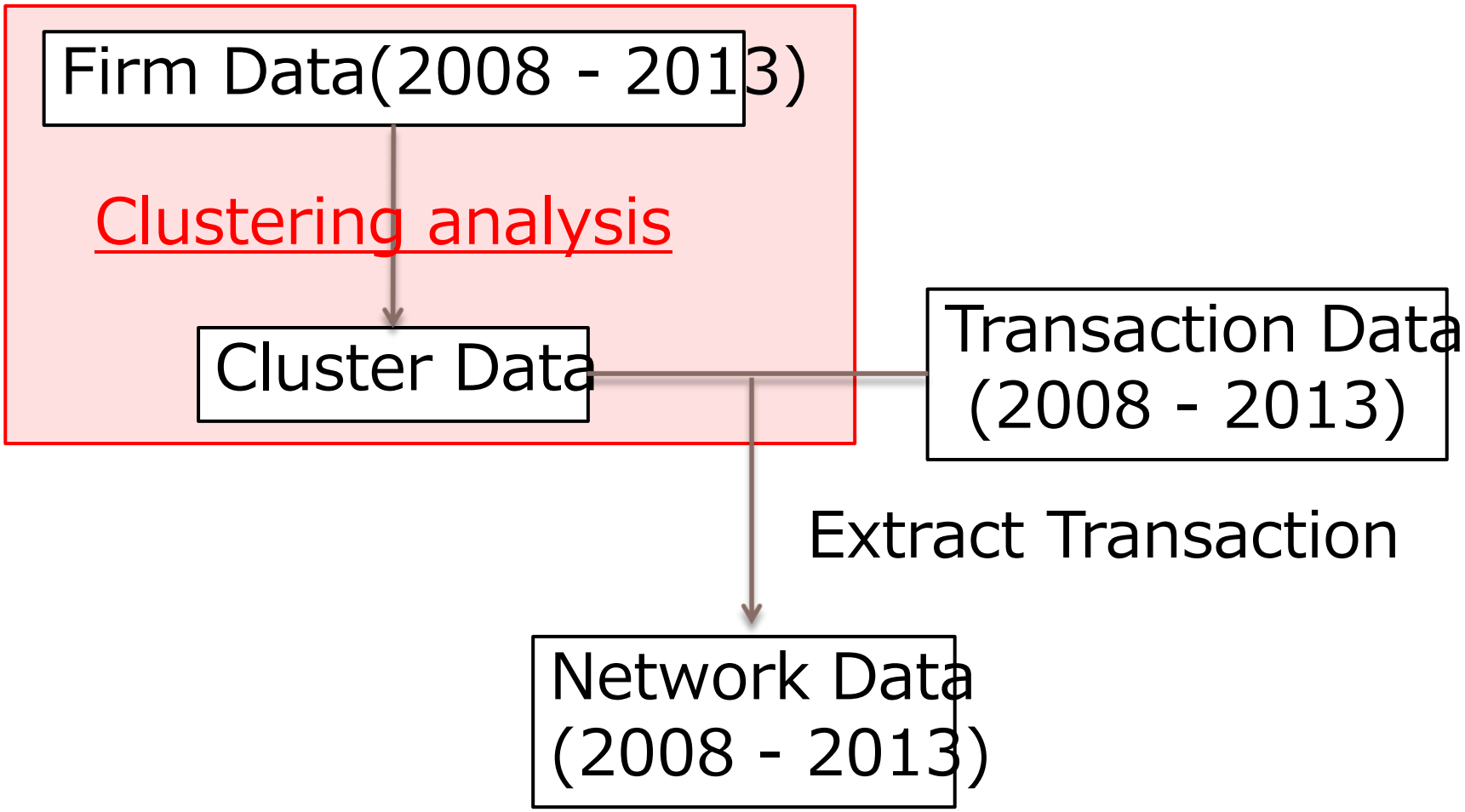


Our objective is…

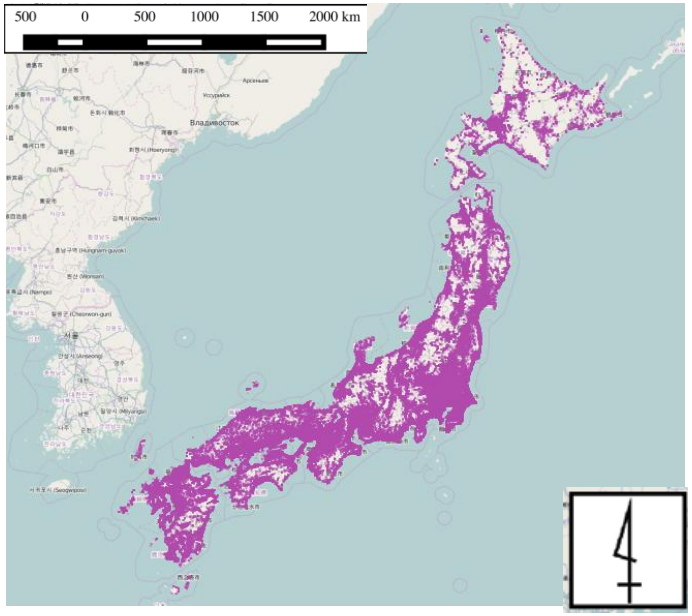
1. Defining local area with clustering method. (non-administrative division)
2. Visualizing economic network between local area in Japan using transaction data.

Understanding important counterparts for each local area.

Processing flow

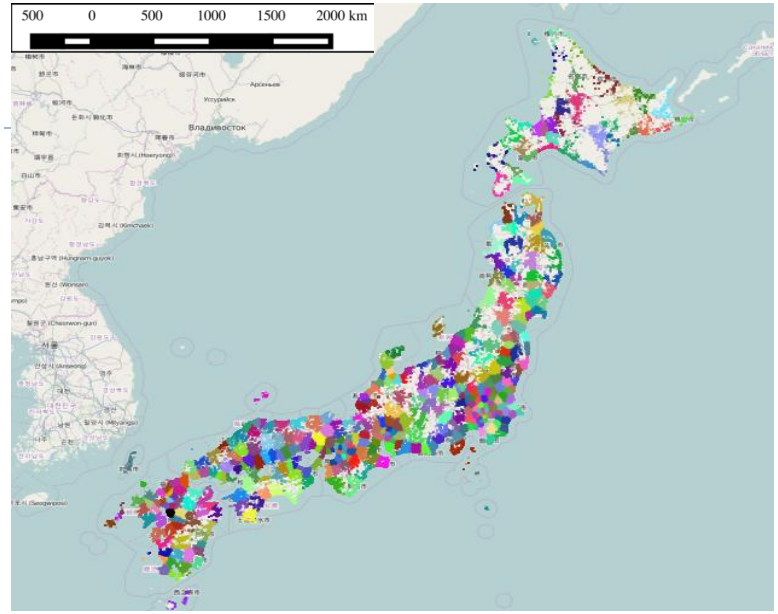


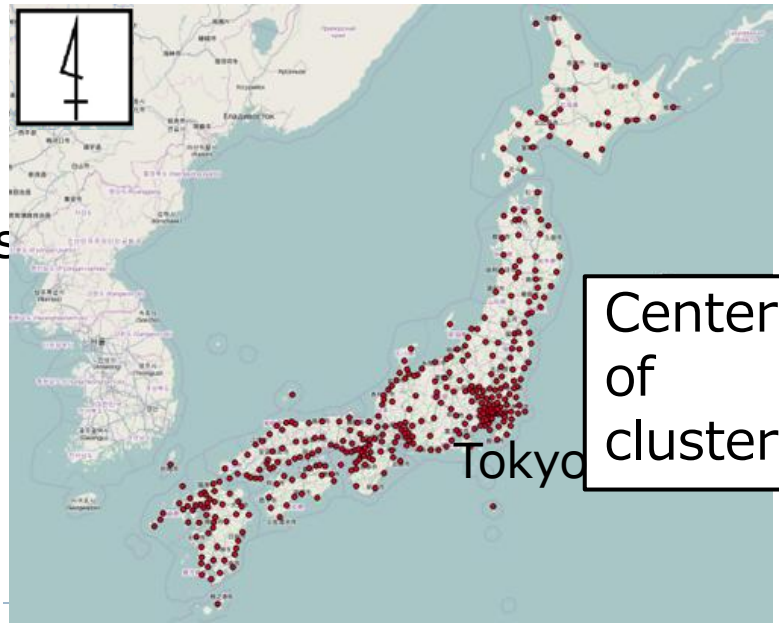
Result-Clustering



Mapping of 1.6 million firms (point data)

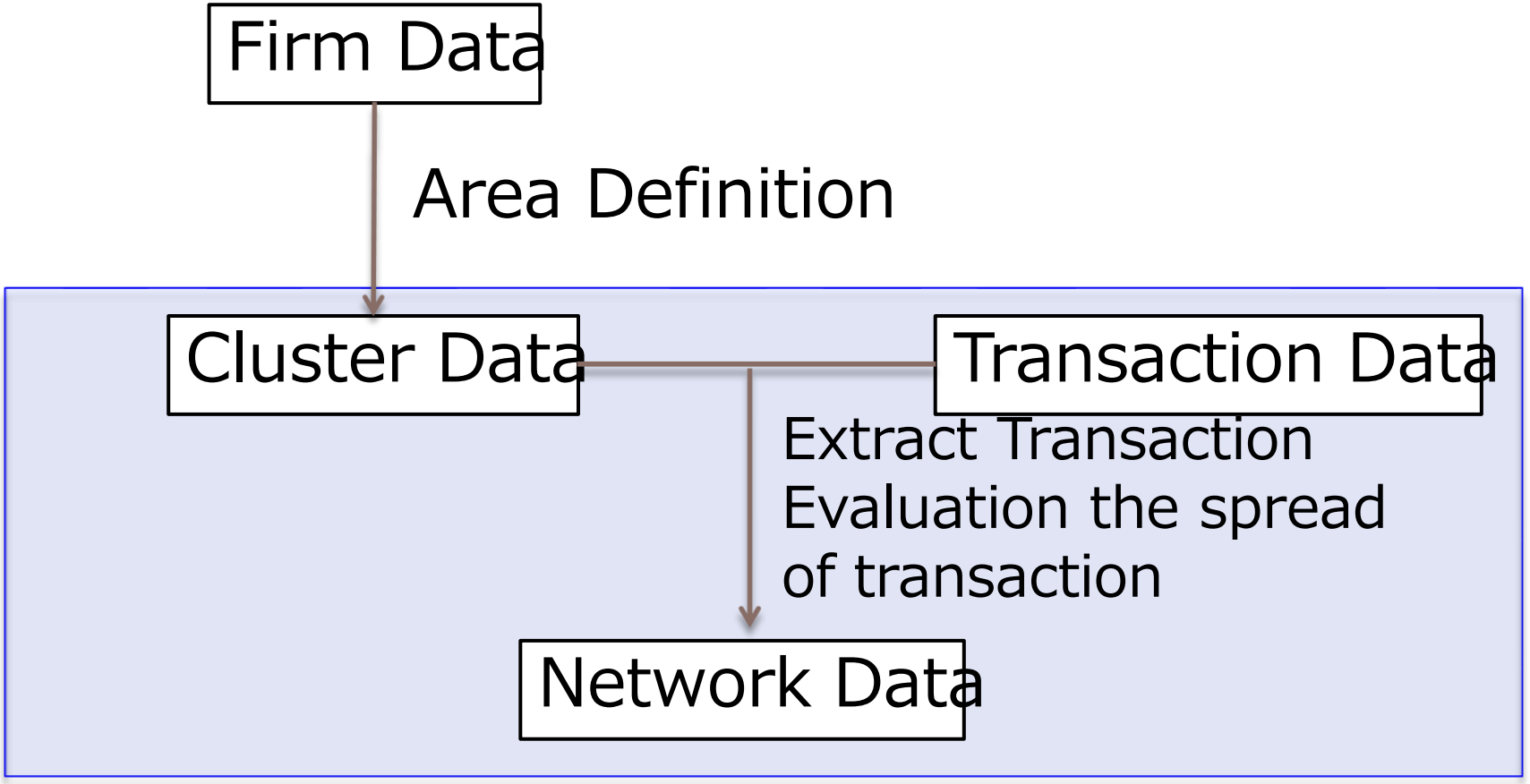
K-means++ clustering (400 clusters





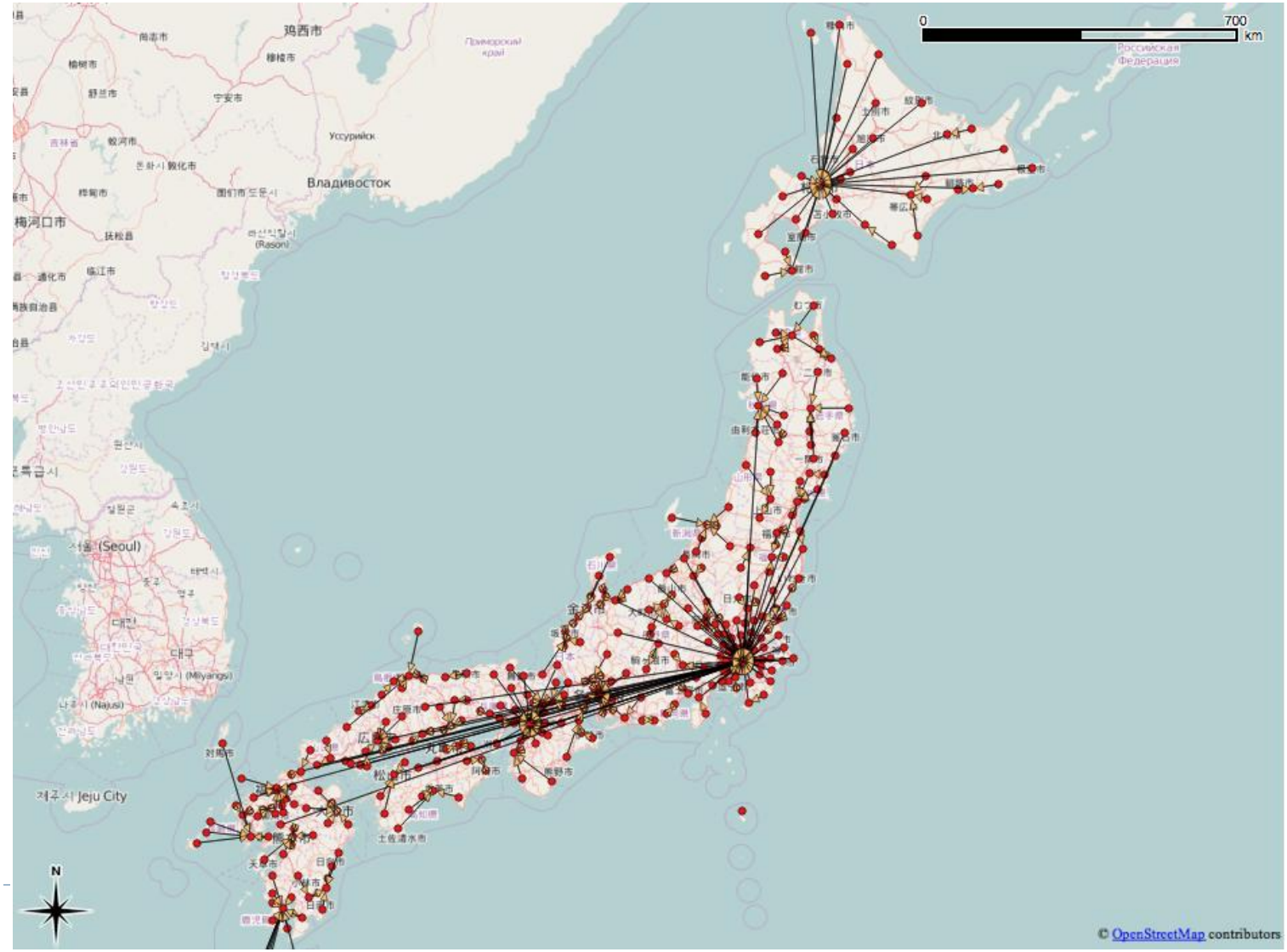
Each color means each cluster.
Many clusters are in urban area.

Processing Flow



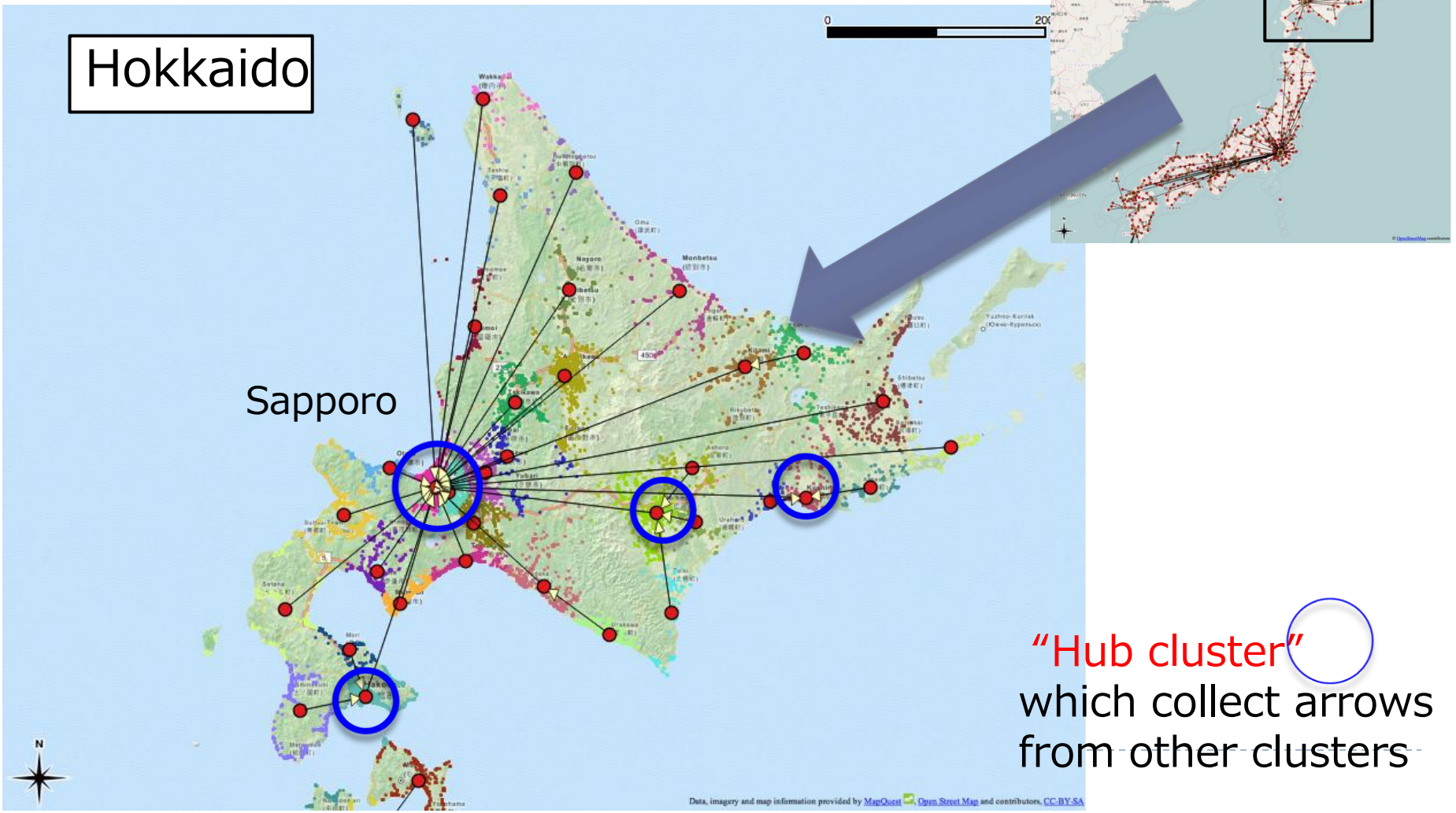
Result & Discussion

Arrows: 1st transaction counterpart of each cluster



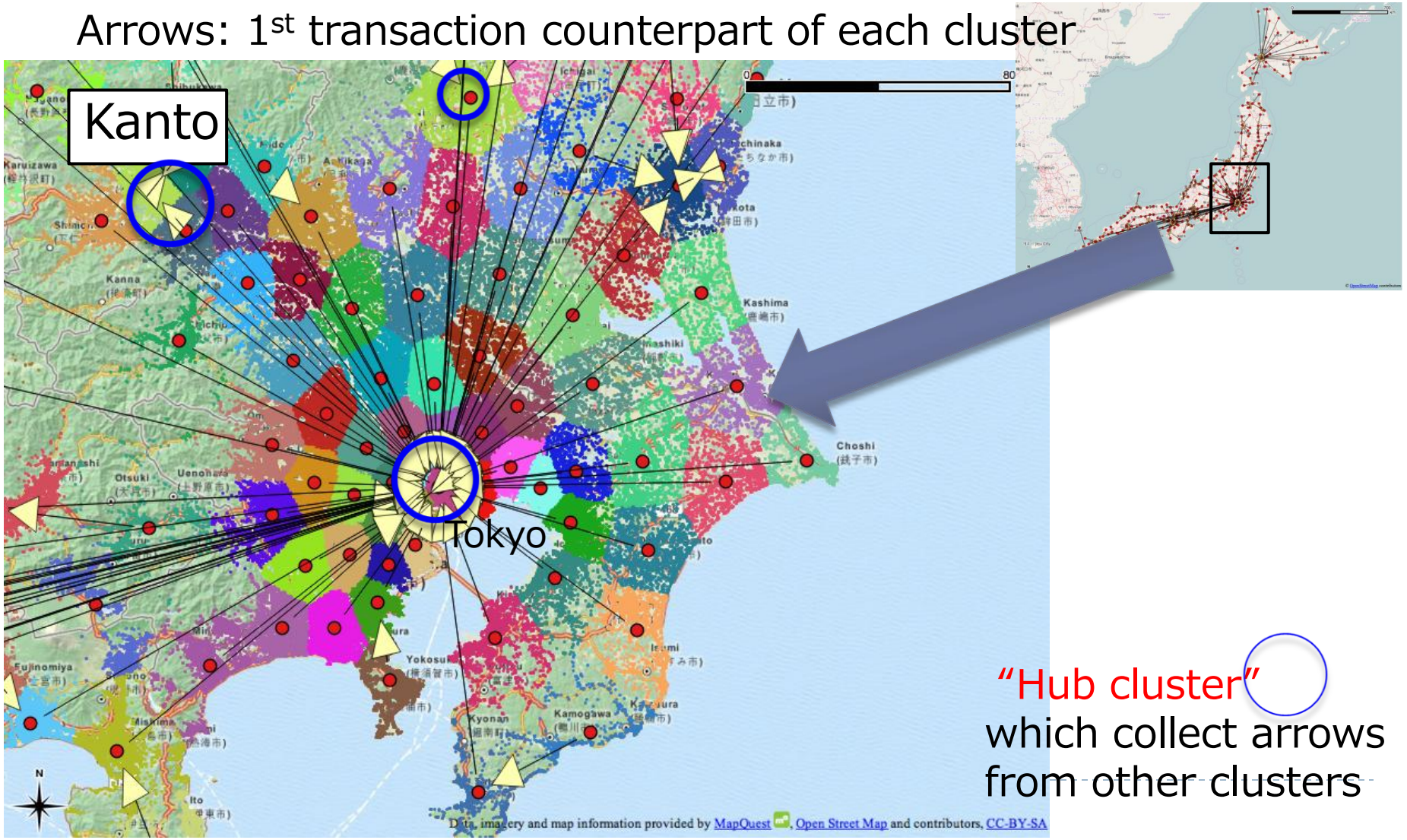
Result & Discussion

Arrows: 1st transaction counterpart of each cluster



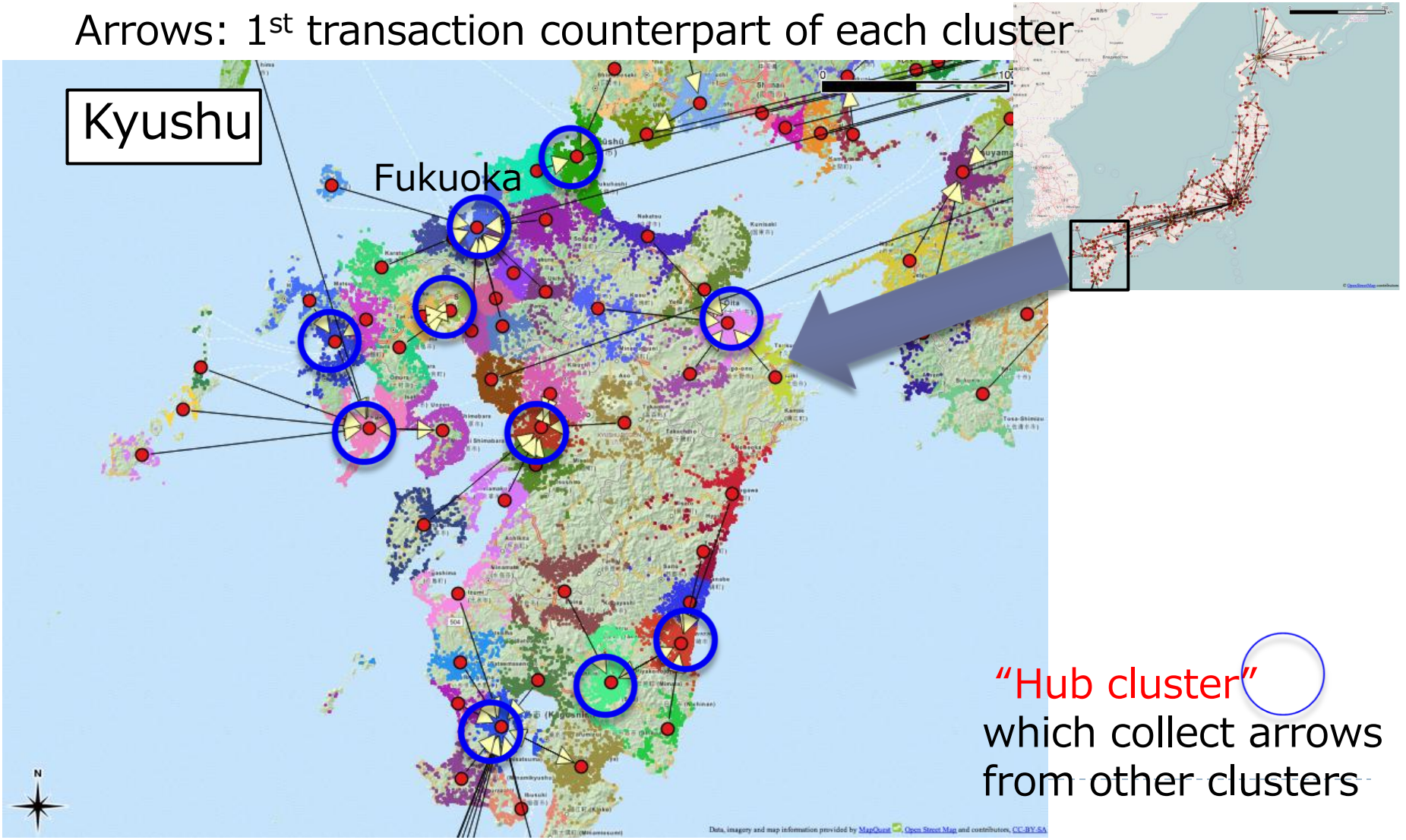
Result & Discussion

Arrows: 1st transaction counterpart of each cluster



Result & Discussion

Arrows: 1st transaction counterpart of each cluster



Conclusion

1. New definition of Local area
   → clustering based on geospatial data of firms
2. Visualizing economic network between local area
   → Dependent cluster ⇔ Hub cluster

Support Strategic Economic Policy
- selection of most important
economic counterparts

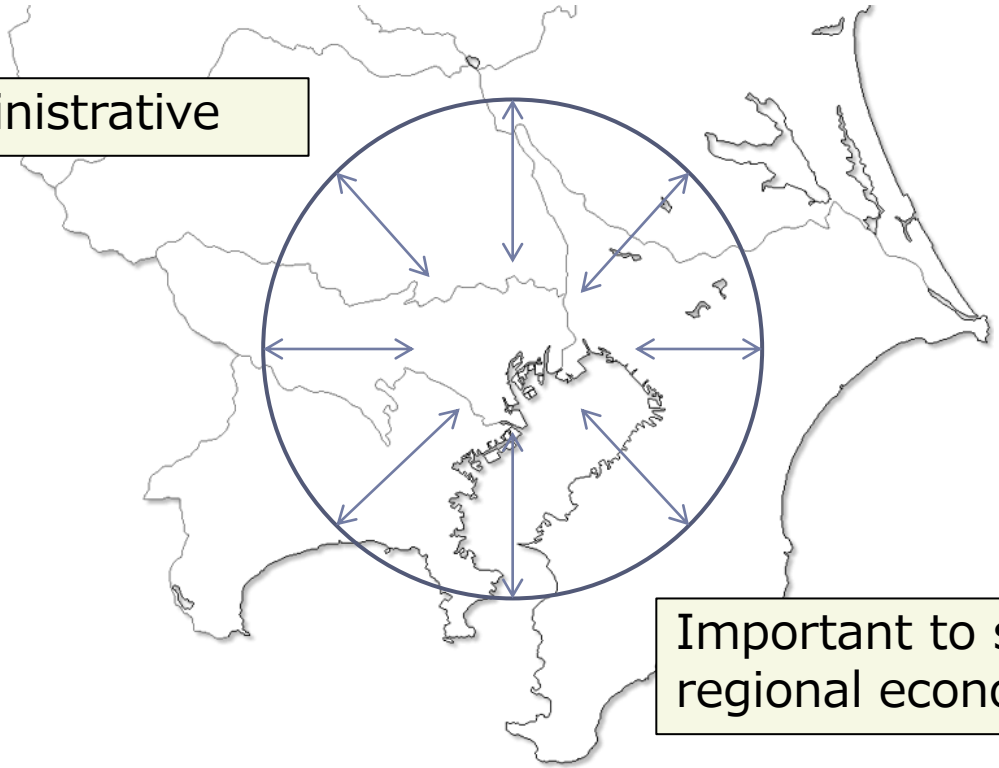
A Network Theory-based Delineation of Metropolitan areas with Mass People Flow Data

Ritsu Sakuramachi 桜町　律
(Shibasaki Lab., The University of Tokyo)
rsakura@csis.u-tokyo.ac.jp

What is a metropolitan area?

- Metropolitan areas are regions defined with focusing on spatial expanse of daily economic activities.

Beyond the administrative divisions



Important to strengthen regional economic power

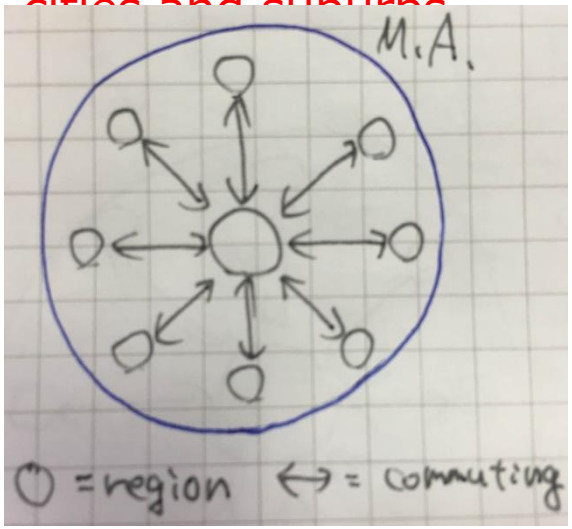
Previous research

表１　中心都市と郊外都市の設定基準

	上段：中心都市の設定基準 下段：郊外都市の設定基準
山田・徳岡（1983）	1. 常住人口が５万人以上。 2. 鉱業を除く非１次産業就業人口比率が75％以上。 3. 昼夜間人口比率が1.0以上。 4. 他の特定の中心都市への流出就業者比率が15％未満。 5. 総流出就業者比率が30％未満。
	1. 中心都市への流出就業者比率が10％以上。 2. 鉱業を除く非１次産業就業人口比率が75％以上。
森川（1990）	1. 卸売・小売業，サービス業従業者数が３千人以上。
	1. 中心都市への通勤者比率が５％以上。
Kawashima et al.（1993）	1. 常住人口が10万人以上。 2. 昼夜間人口比率が1.0以上。
	1. 中心都市への通勤比率が５％以上，または，５百人以上。
総務庁統計局（1999）	1. 東京特別区部および政令指定都市（大都市圏）。大都市圏に属さない人口50万人以上の市（都市圏）。
	1. 中心都市への流出出勤・通学者の常住人口に占める割合が1.5％以上。
日本産業消費研究所（2000）	1. 周辺市町村からの通勤・通学者比率が10％以上。
	1. 中心都市への通勤・通学者比率が10％以上。
金本・徳岡（2002）	1. DID 人口が１万人以上。 2. 郊外市町村の条件を満たすが，従業常住人口比率が１以上で，DID 人口が中心都市の３分の１以上か，10万人以上。
	1. 中心都市への通勤率が10％以上。 2. 郊外市町村への通勤率が10％以上。

（出所）山田・徳岡（1983），森川（1990），Kawashima et al.（1993），総務庁統計局（1999），日本産業消費研究所（2000），金本・徳岡（2002）より作成。

Many previous researches define metropolitan areas based on commuting rates between central cities and suburbs



Hayashi (2014) Journal of economics and sociology, Kagoshima University, 83: 125-137

Previous research

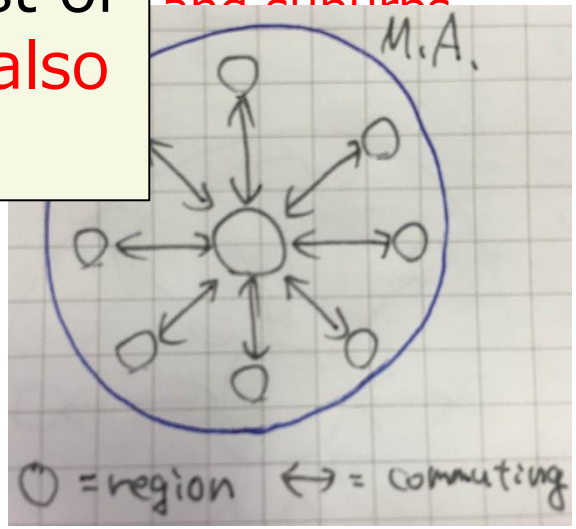
表１　中心都市と郊外都市の設定基準

	上段：中心都市の設定基準 下段：郊外都市の設定基準
山田・徳岡（1983）	1. 常住人口が５万人以上。 2. 鉱業を除く非１次産業就業人口比率が75％以上。 3. 昼夜間人口比率が1.0以上。 4. 他の特定の中心都市への流出就業者比率が15％未満。 5. 総流出就業者比率が30％未満。
	1. 中心都市への流出就業者比率が10％以上。 2
森川（1990）	1 1
Kawashima et al.（1993）	1 2 1
総務庁統計局（1999）	1 1
日本産業消費研究所（2000）	1. 周辺市町村からの通勤・通学者比率が10％以上。 1. 中心都市への通勤・通学者比率が10％以上。
金本・徳岡（2002）	1. DID 人口が１万人以上。 2. 郊外市町村の条件を満たすが，従業常住人口比率が１以上で，DID 人口が中心都市の３分の１以上か，10万人以上。
	1. 中心都市への通勤率が10％以上。 2. 郊外市町村への通勤率が10％以上。

（出所）山田・徳岡（1983），森川（1990），Kawashima et al.（1993），総務庁統計局（1999），日本産業消費研究所（2000），金本・徳岡（2002）より作成。

However, our daily economic activities consist of not only commuting but also other activities (buying behavior, tourism, etc.)

Many previous researches define metropolitan areas
d on commuting
between central
and suburbs



Hayashi (2014) Journal of economics and sociology, Kagoshima University, 83: 125-137

Progress of data

National census (per 5 years)

GPS data (24 hours in 365 days!)

comprehensive and continuous mass people flow data
with high spatiotemporal resolution

Study flow

GPS data (2012) ZENRIN DataCom CO., LTD.

Estimation of stay points by Akiyama (2013)

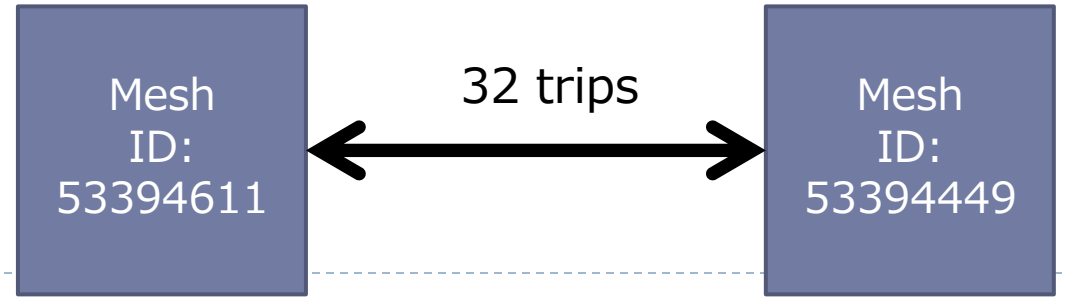
Location data of both origin stay points and destination stay points (643,407,000 trips)

287,320 meshes (1km×1km) in

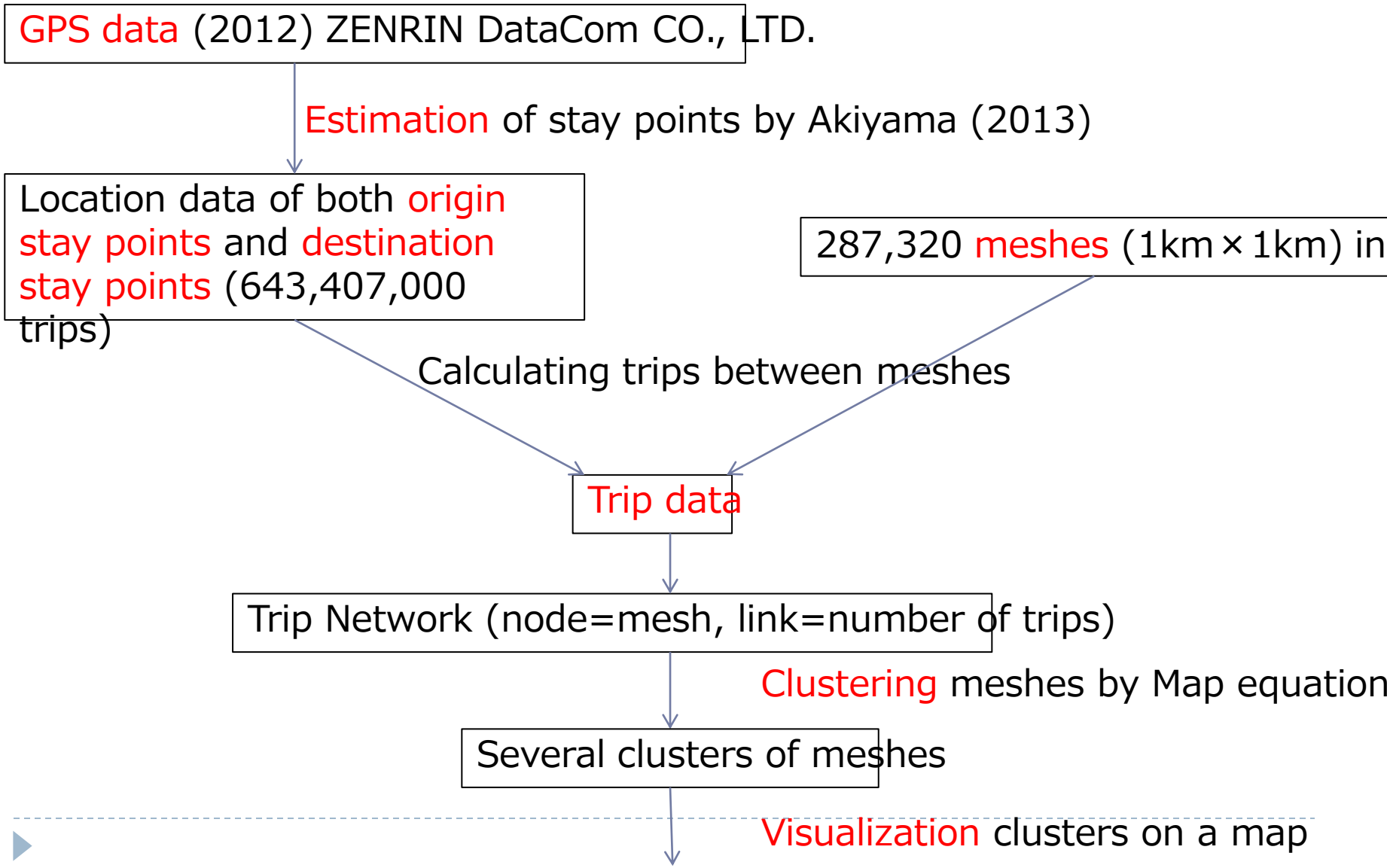
Calculating trips between meshes

Trip data

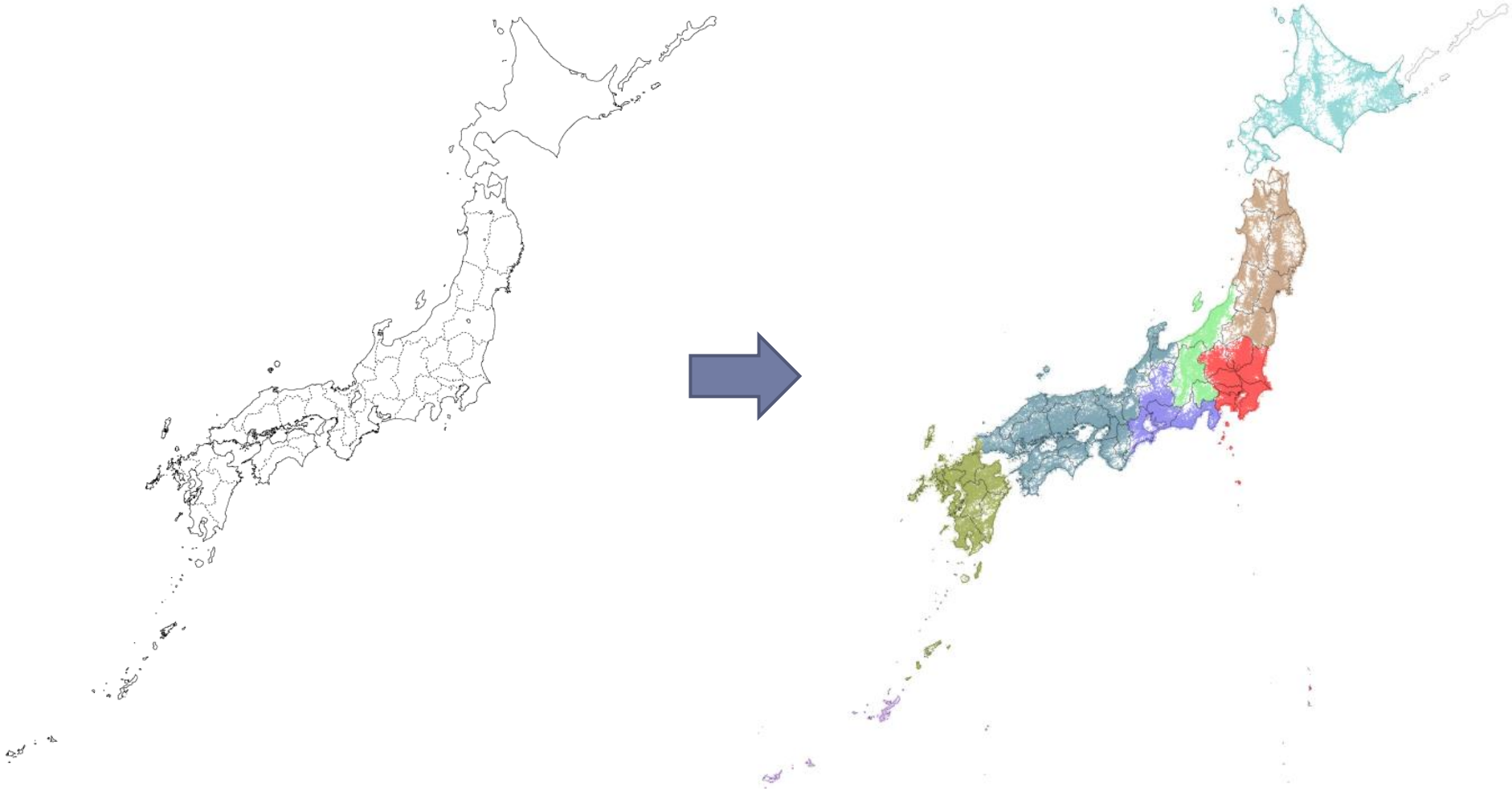
Trip data including information like…



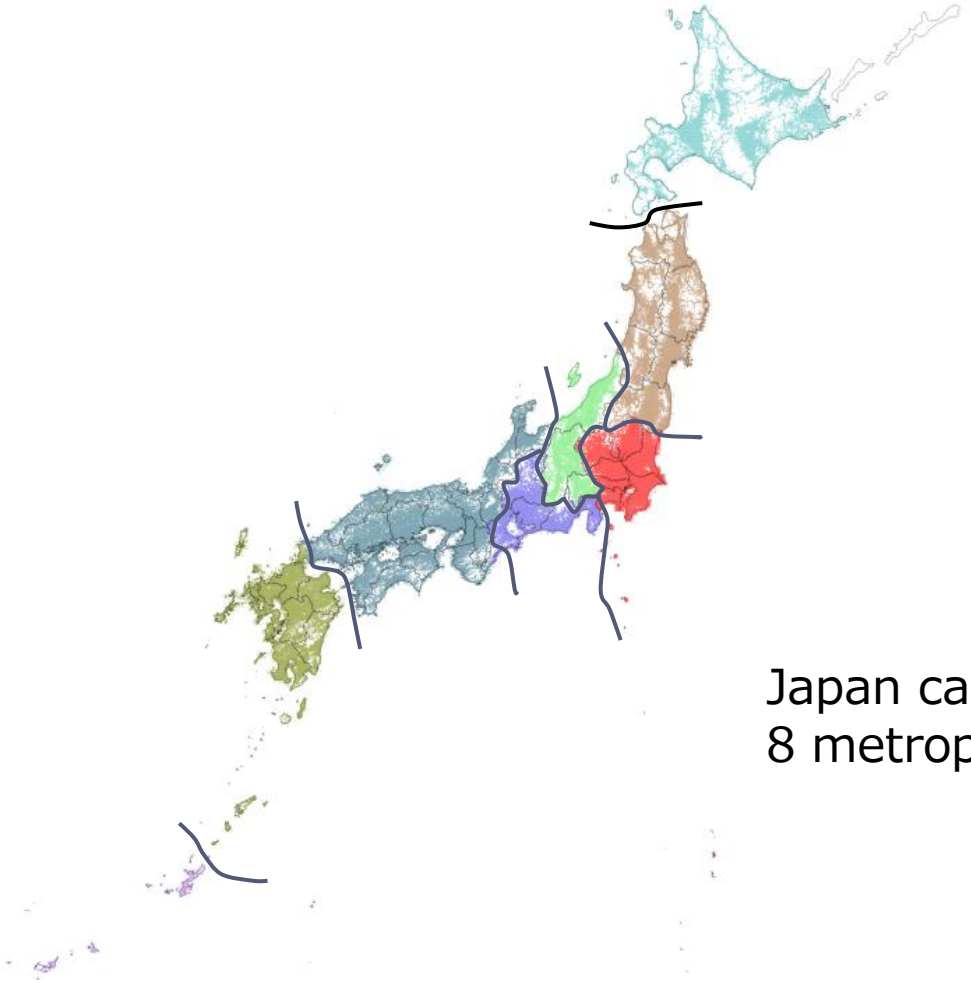
Study flow


GPS data (2012) ZENRIN DataCom CO., LTD.
Estimation of stay points by Akiyama (2013)
Location data of both origin stay points and destination stay points (643,407,000 trips)
287,320 meshes (1km×1km) in
Calculating trips between meshes
Trip data
Trip Network (node=mesh, link=number of trips)
Clustering meshes by Map equation
Several clusters of meshes
Visualization clusters on a map

Result of Clustering

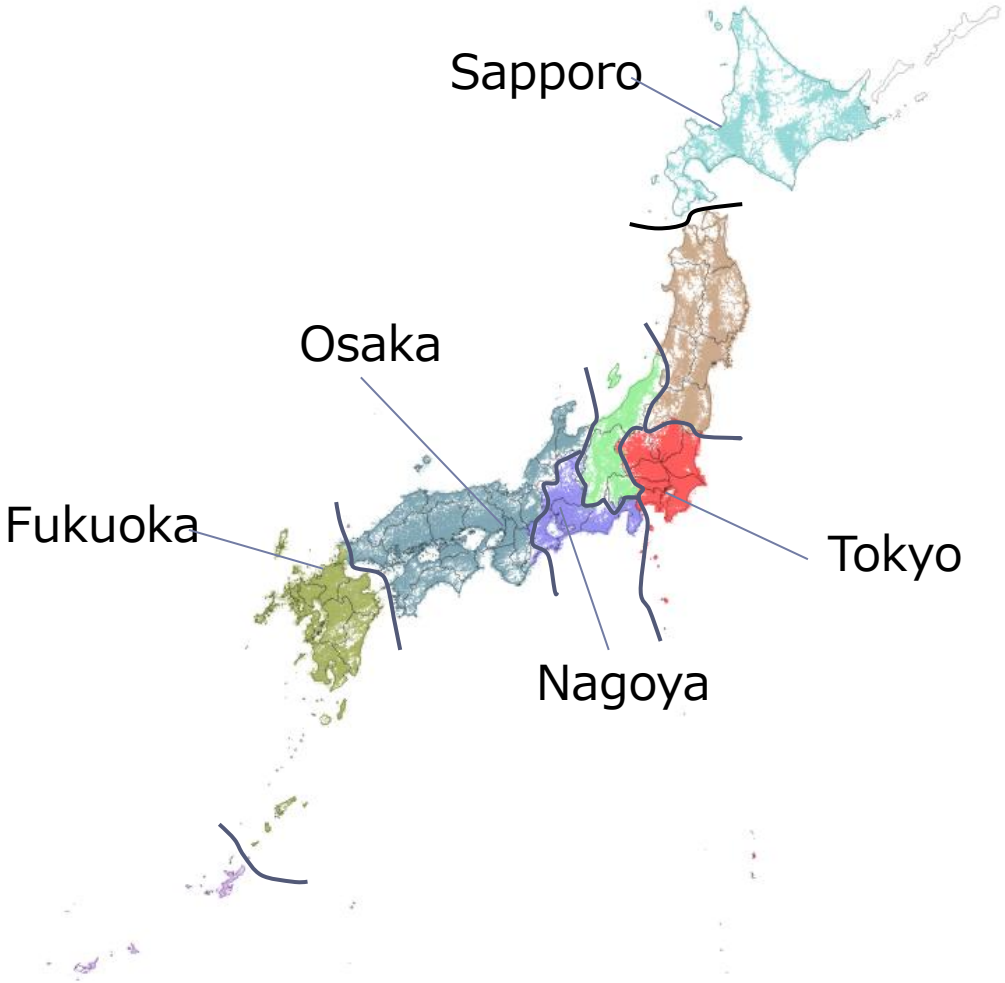


Result of Clustering

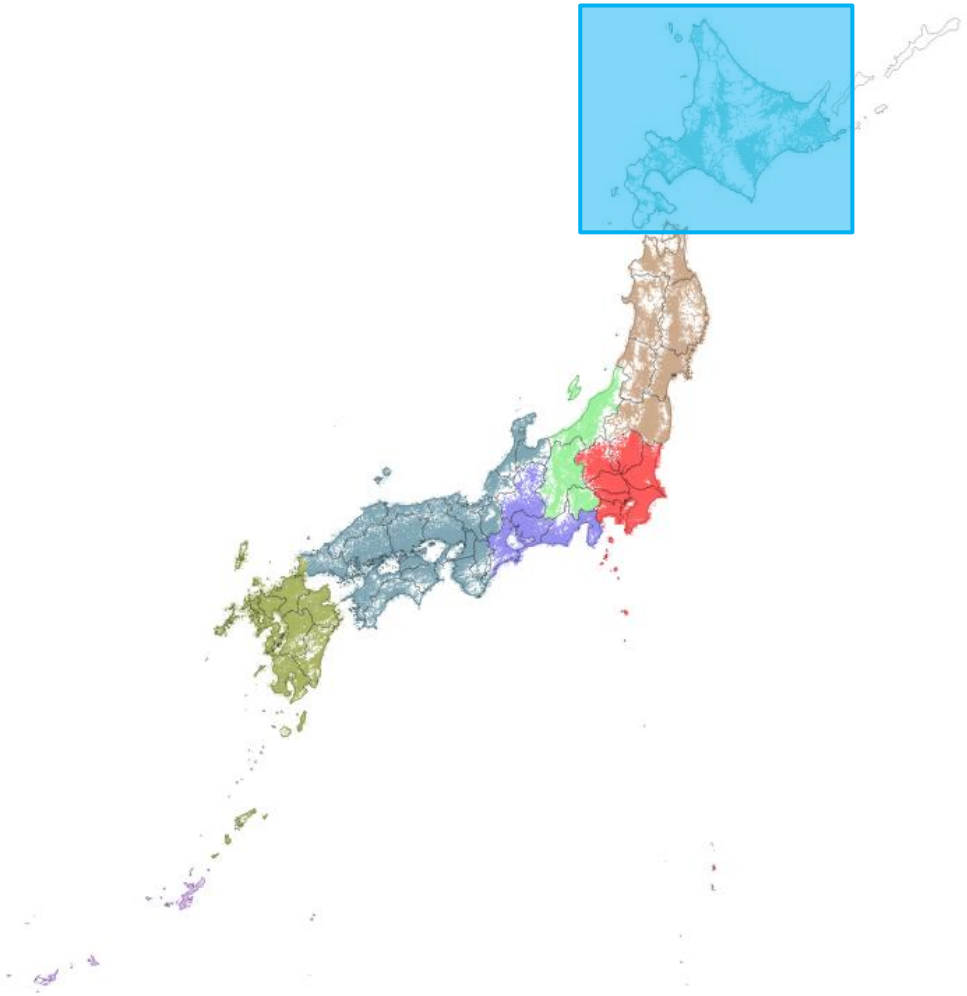


Japan can be devided into
8 metropolitan areas

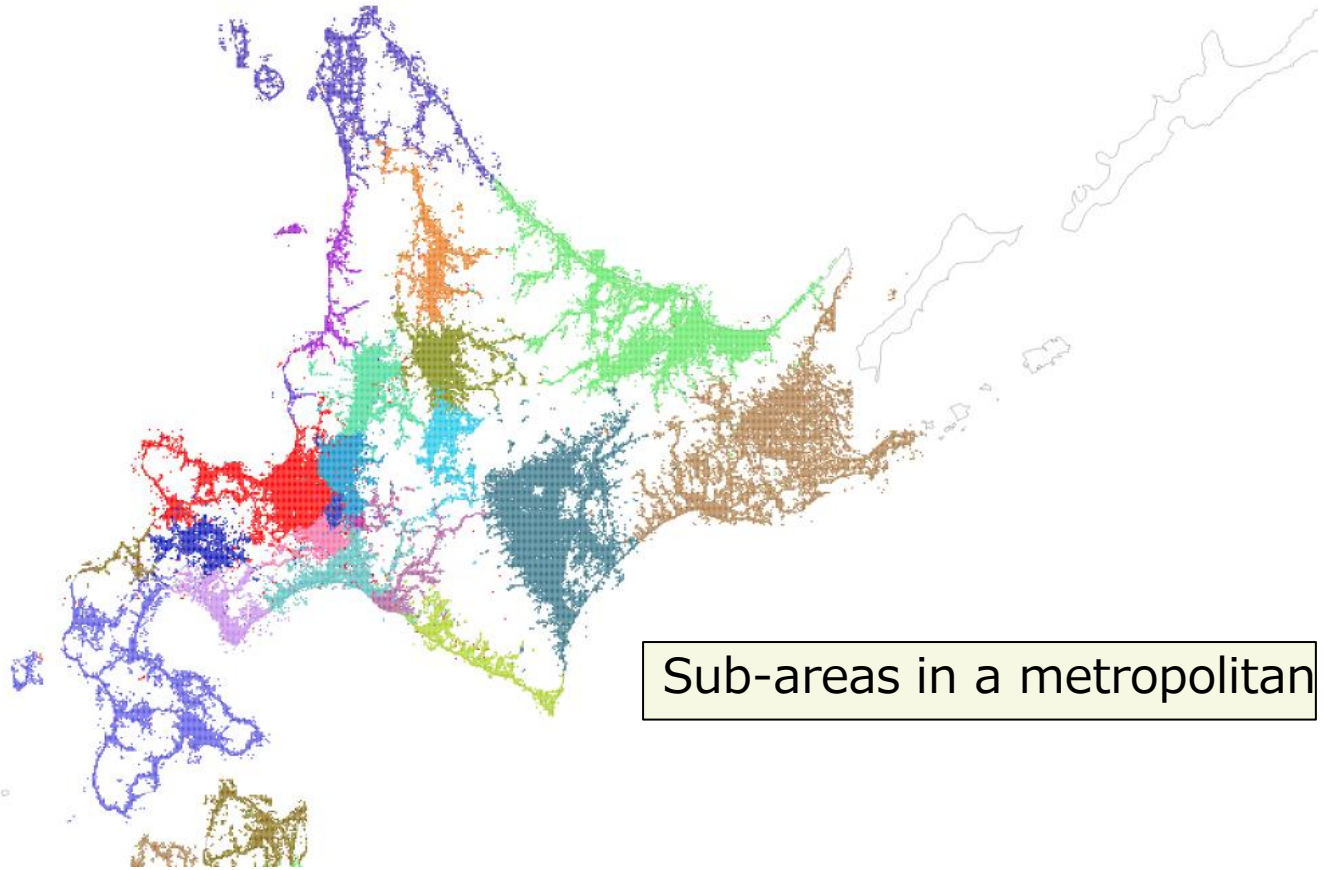
Result of Clustering



Result of Clustering

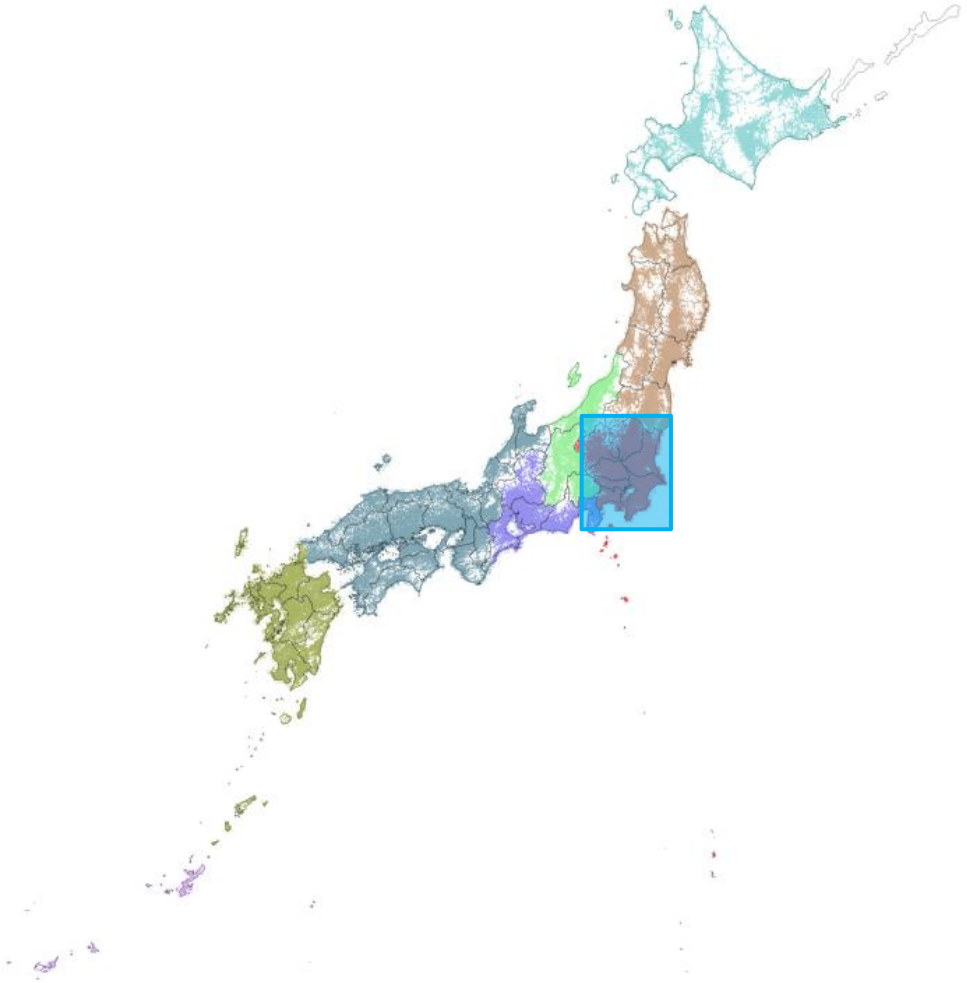


Result of Clustering

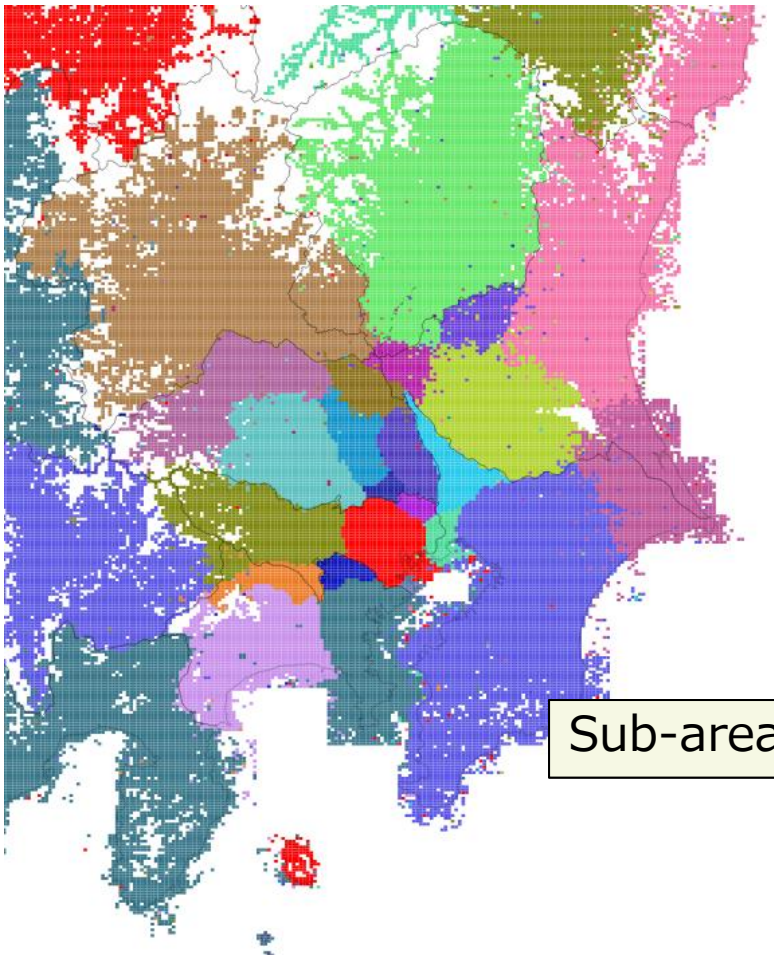


Sub-areas in a metropolitan area

Result of Clustering

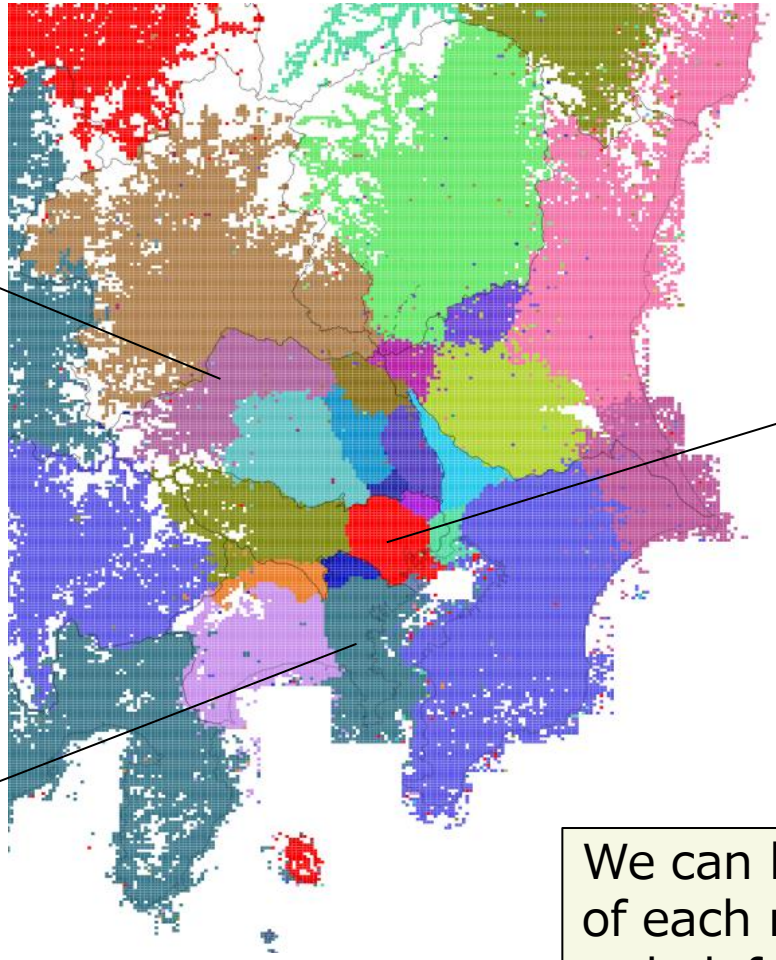


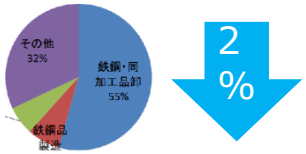
Result of Clustering



Sub-areas in a metropolitan area

Combination of metropolitan area and inter-firm big data

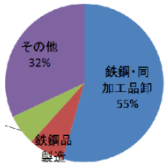




2 %

Business category rates

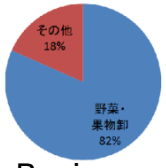
Sale growth



5 %

Business category rates

Sale growth



2 %

Business category rates

Sale growth

We can know the economic situati of each regions
→ helpful for policy making