Network Theory for Local Mapview

국 웅

서울대학교

MCC-SNU 공동연구 종료보고

2019.10.10

연구 개요

- 연구명: 신뢰네트워크 구현 및 시각화 (2019.06-2019.08)
- 참여인원: 국웅, 김원세, 진성태, 이강주
- 연구의 내용 및 목표

  MCC-SNU 1차 공동연구 ‘선물의 숲을 통한 신용평가’ (2018.05– 2018.08)의 연속연구

  Good Morning 서비스를 통해 수집한 데이터로부터 회원의 영향력 분석 및 평가방법 제시

  신뢰네트워크 구축 및 mapview개발

  회원들의 네트워크 행동패턴 통계적분석앱 개발

연구 주제 및 방법

- Network visualization:
  TDA (mapper), clustering

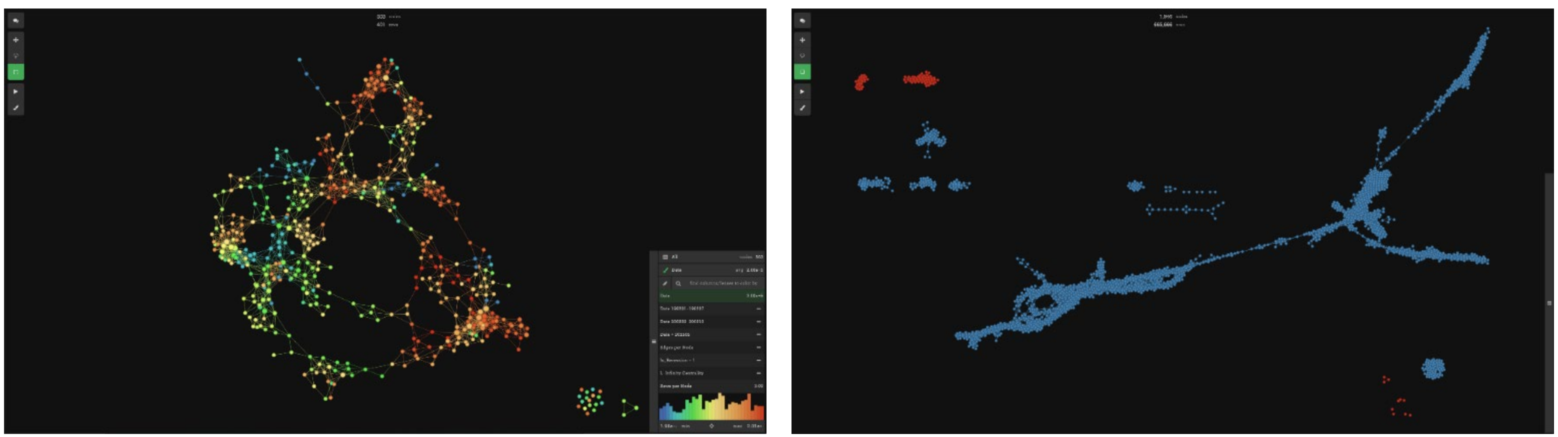
- Network feature and singular user detection:
  graph theory, statistical centralities

- Mapview:
  persistent neighborhood, Page rank

- Future work: network expansion and partition
  harmonic cycles, effective conductance

Mathematics to Industry (M to I)

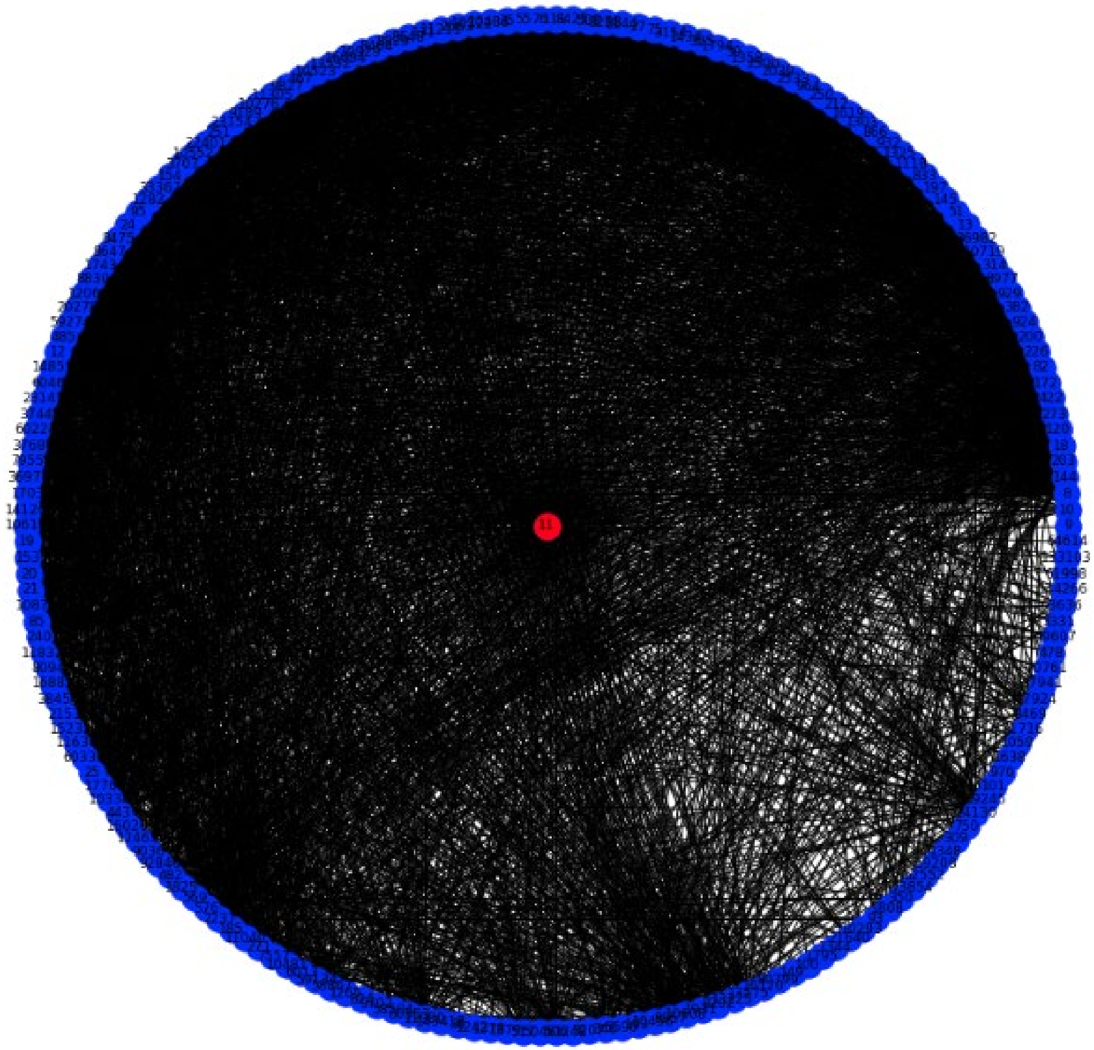
Mathematical Insight → Data Analytics Solution



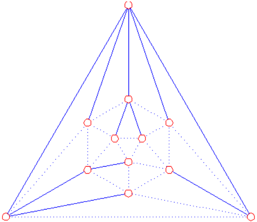
Mapper: Ayasdi Core.

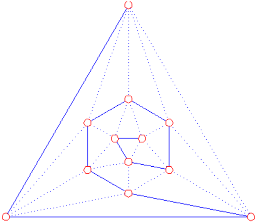
Source: Industrial Mathematics Symposium 2018.05.19

Induced Depth 1 Neighborhood



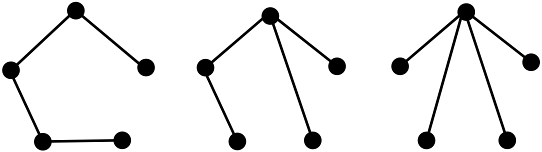
Spanning trees of a graph



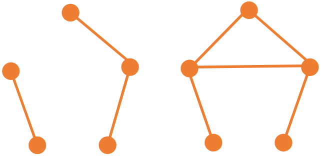


- A spanning tree B is a maximal acyclic subgraph:
  (1) B is connected
  (2) B is acyclic
  (3) B has $n - 1$ edges (n=# vertices)
  (4) $B \cup e$ contains a unique cycle containing e for each $e \notin B$.

Spanning trees and non-spanning trees



spanning trees



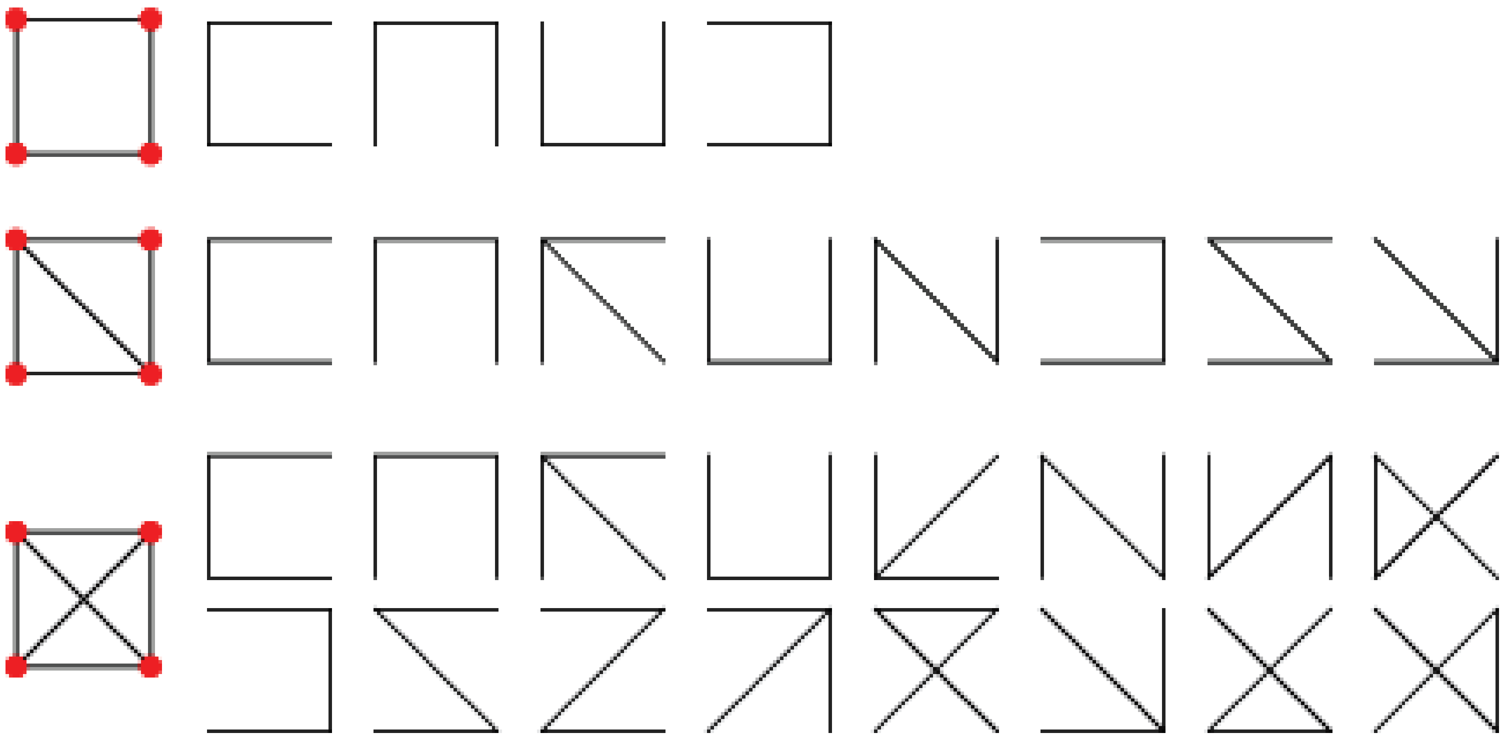
non spanning trees

spanning tree = connected + acyclic

$k(G)$ = # spanning trees in G

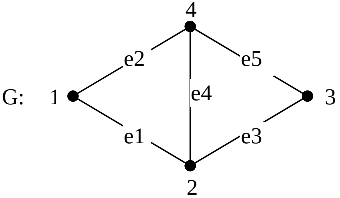
Spanning Trees of Graphs

$k(G)$= # spanning trees in G = *complexity* of G



$k(G_1) = 4$ $k(G_2) = 8$ $k(G_3) = 16$

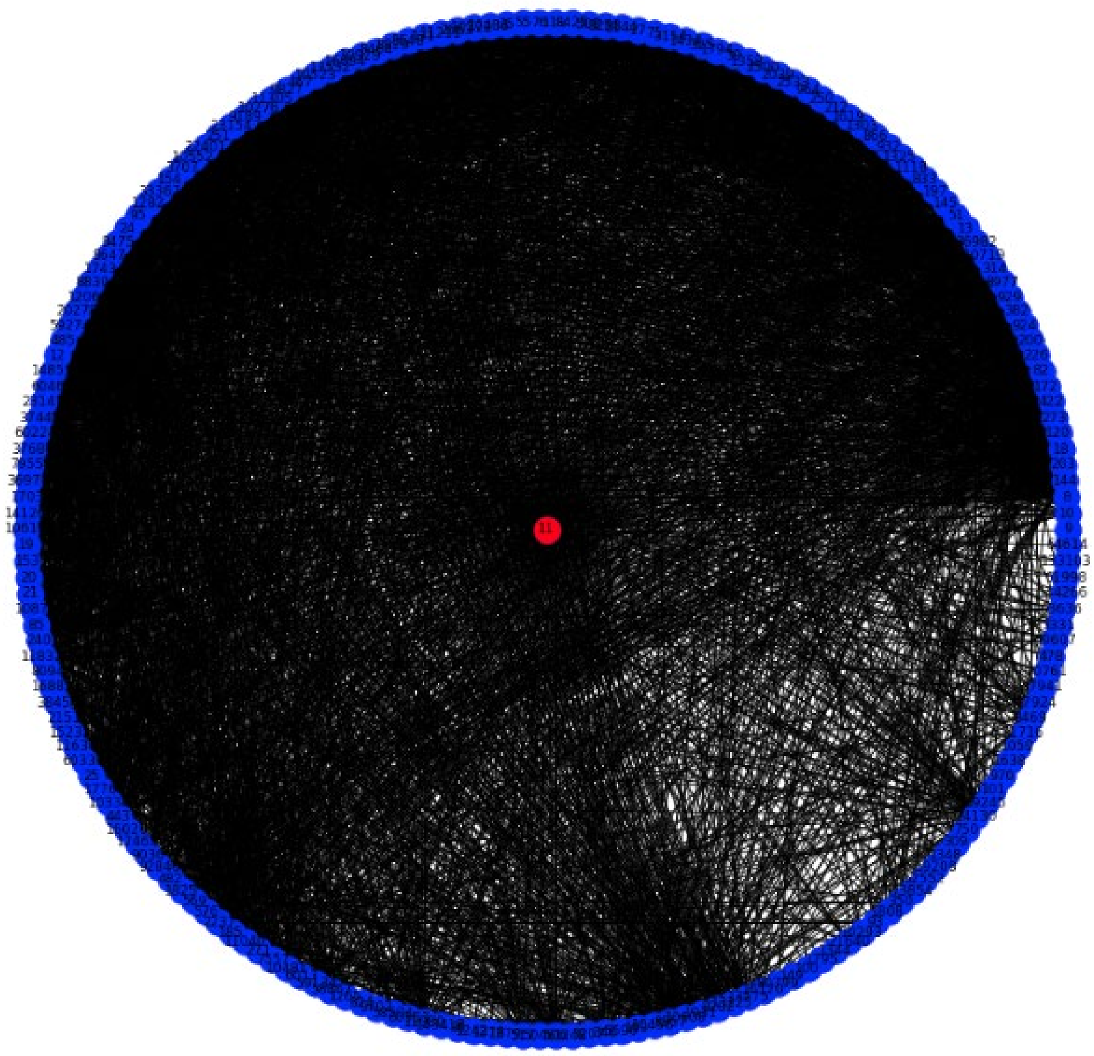
Example: Δ_0



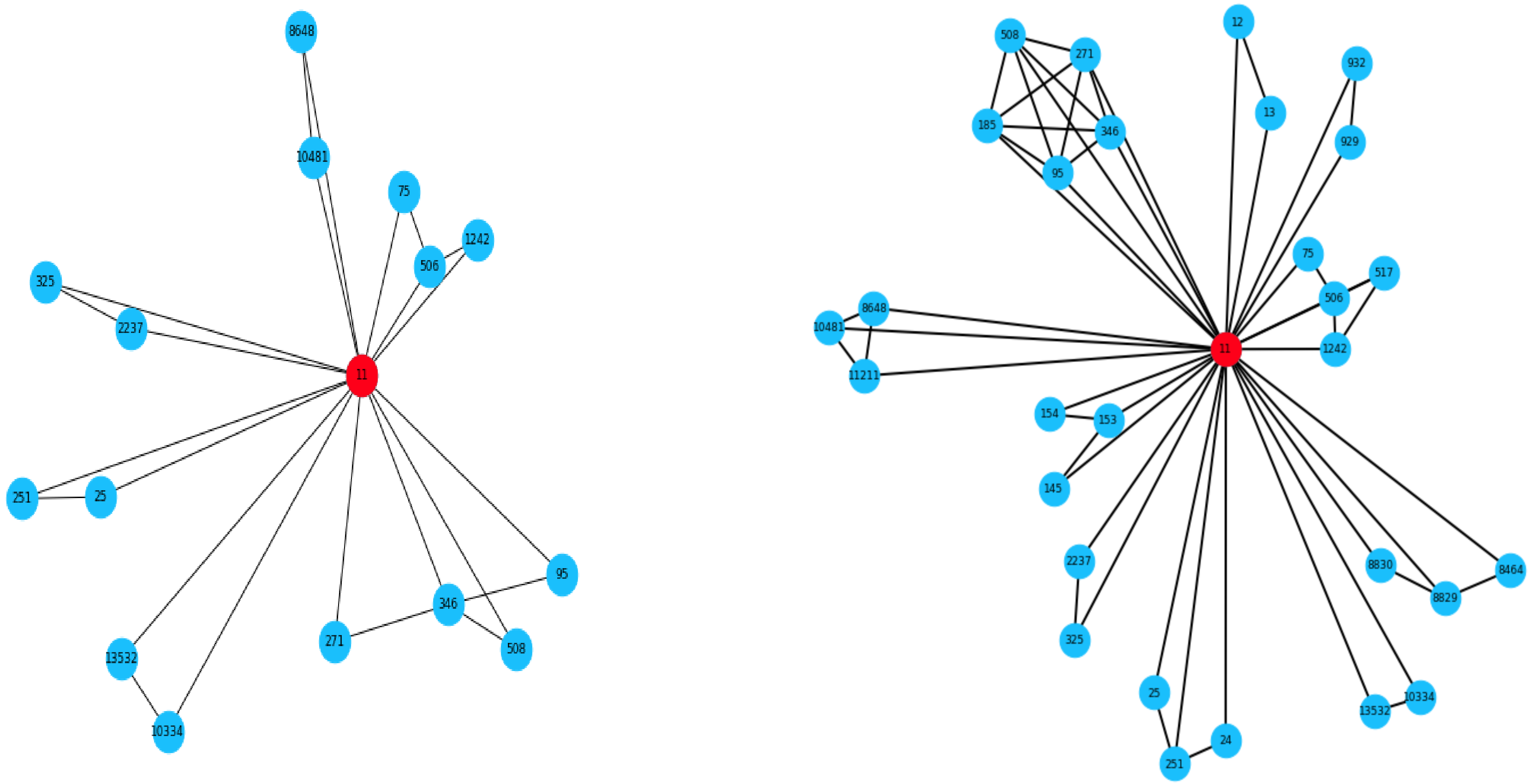
$$\partial_1 = \begin{pmatrix} -1 & -1 & 0 & 0 & 0 \\ 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix} \quad \partial_0 = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}$$

$$\Delta_0 = \partial_1 \partial_1^t + \partial_0^t \partial_0 = \begin{pmatrix} 2 & -1 & 0 & -1 \\ -1 & 3 & -1 & -1 \\ 0 & -1 & 2 & -1 \\ -1 & -1 & -1 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

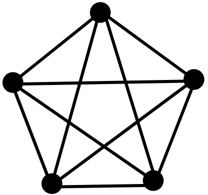
Induced Depth 1 Neighborhood



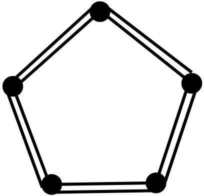
Visualization: Persistent Neighborhood



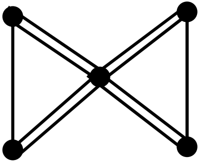
Connectivity vs Complexity



$k(G) = 125$



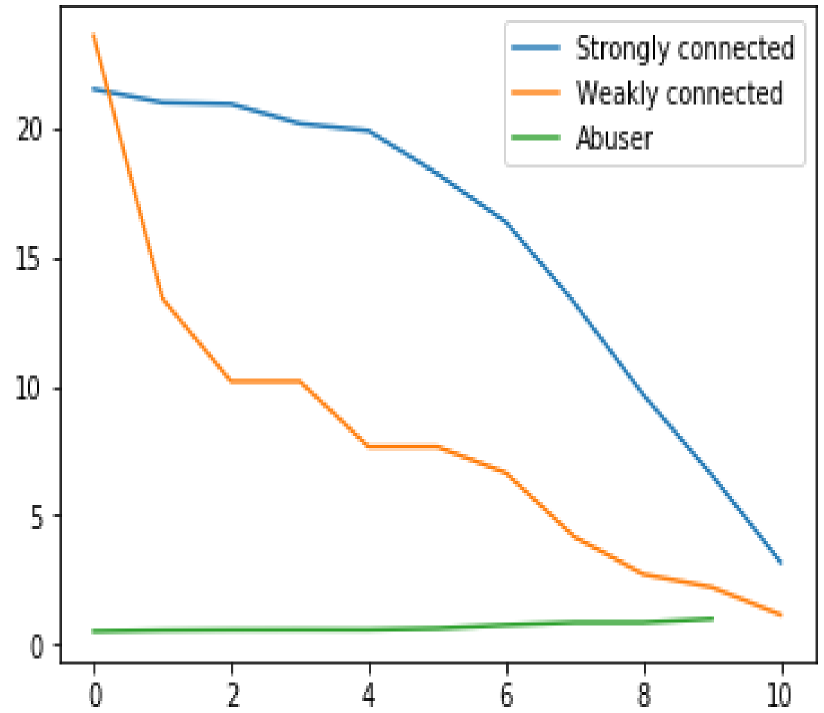
$k(G) = 80$



$k(G) = 64$

Complexity Trace

$$SC = \frac{\log(\#\ \text{of Spanning trees})}{\log(\text{sum of degrees})}$$



Various network centralities

- Betweenness centrality (Importance as a **bridge**)

$$B(v) = \sum_{\substack{s,t \in V \\ s \neq v \neq t}} \frac{\#\ of\ shortest\ paths\ s \to t\ passing\ v}{\#\ of\ shortest\ paths\ s \to t}$$

- Closeness centrality (How close the node is to others)

$$\mathrm{C}(v) = \frac{1}{\sum_u d(u,v)}$$

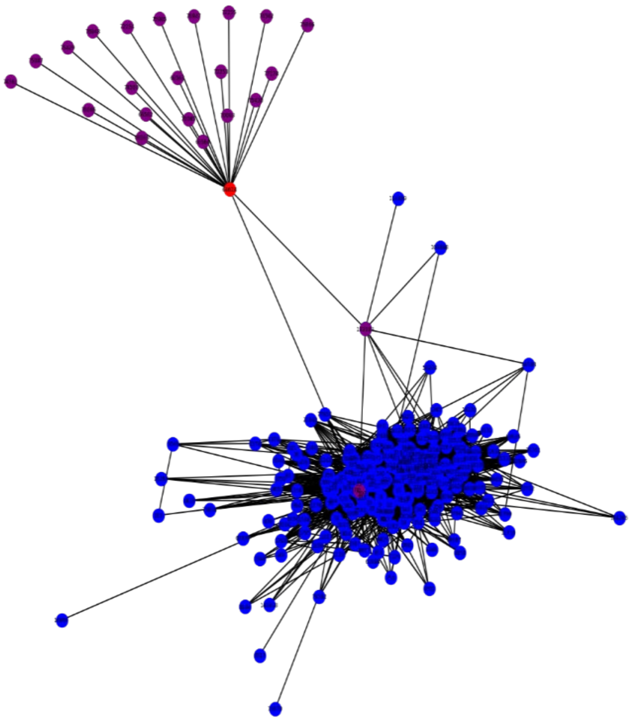
- Eigenvector centrality (Links to “important” nodes are “important)

$$x(v) = \frac{1}{\lambda} \sum_{v \sim t} x(t) \quad \longrightarrow \quad (Matrix\ form)\ Ax = \lambda x$$

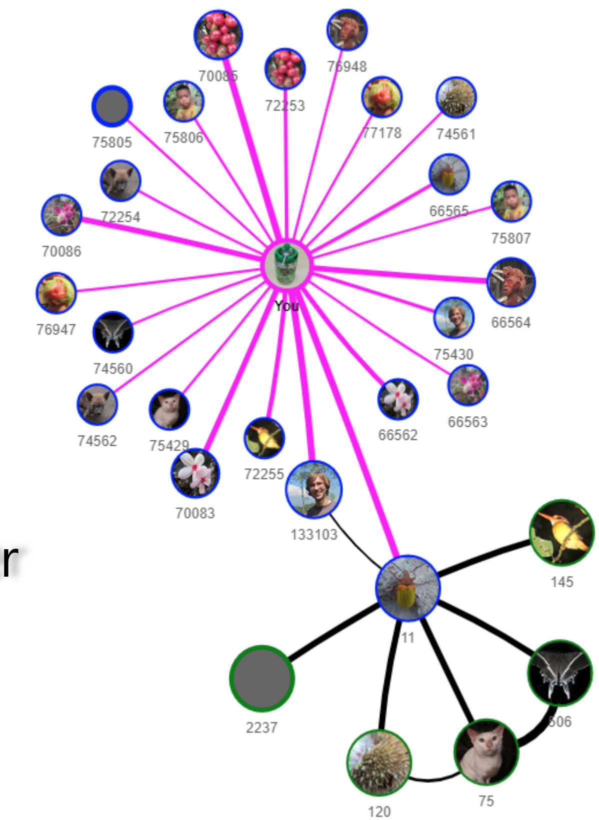
- PageRank (named after **Larry Page** in **Google™**)

$$PR(v) = k \sum_{u \to v} \frac{PR(u)}{outdeg(v)} \quad \longrightarrow \quad (Matrix\ form)\ R = kMR$$

Mapview: Centrality Induced Subnetwork



Centrality Filter



1 Neighborhood

Induced Subnetwork

Combinatorial Laplace Operator

Chain complex

$$\cdots \longrightarrow C_{d+1} \xrightarrow{\partial_{d+1}} C_d \xrightarrow{\partial_d} C_{d-1} \longrightarrow \cdots$$

$$\boxed{\Delta = \partial_d^t \partial_d + \partial_{d+1} \partial_{d+1}^t}$$

Ker Δ	⟹	Shape (TDA)
Det Δ	⟹	Connectivity
Inv Δ	⟹	Centrality

3) **K. Woong**, V. Reiner, and D. Stanton, Combinatorial Laplacians of matroid complexes, *Journal of the American Mathematical Society* (2000).

Harmonic space and combinatorial Hodge Theory

- Harmonic space $\mathscr{H}_i \subset C_i$

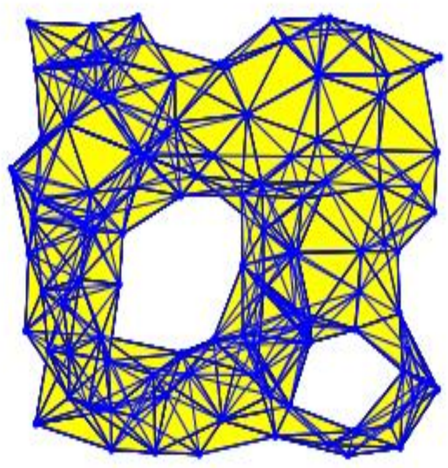
$$\mathscr{H}_i = \ker \Delta_i = \ker \partial_i \cap \ker \partial_{i+1}^t$$

- Energy minimizing property: for $h \in \mathscr{H}_i$ and $x = h + \partial_{i+1} y$

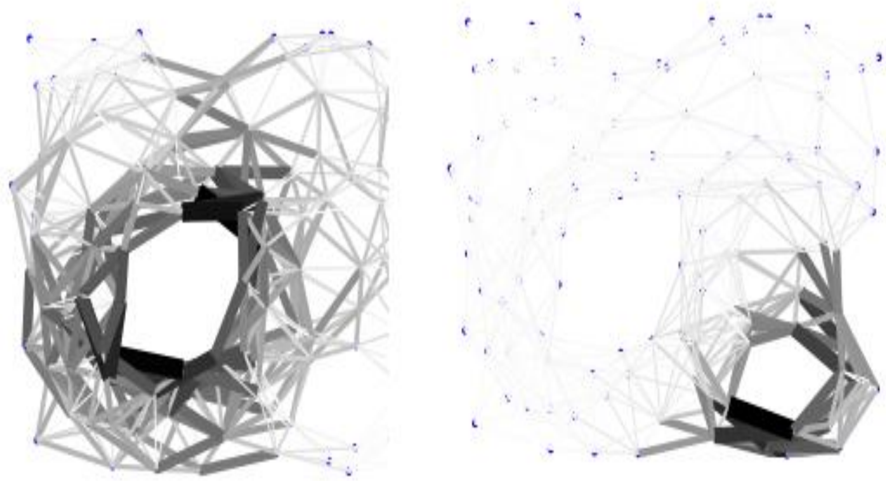
$$\langle h, h \rangle \ \leq \ \langle x, x \rangle$$

Proof: $\langle h, \partial_{i+1} y \rangle = \langle \partial_{i+1}^t h, y \rangle = 0 \quad \square$

Hole Detection via Harmonic Cycles



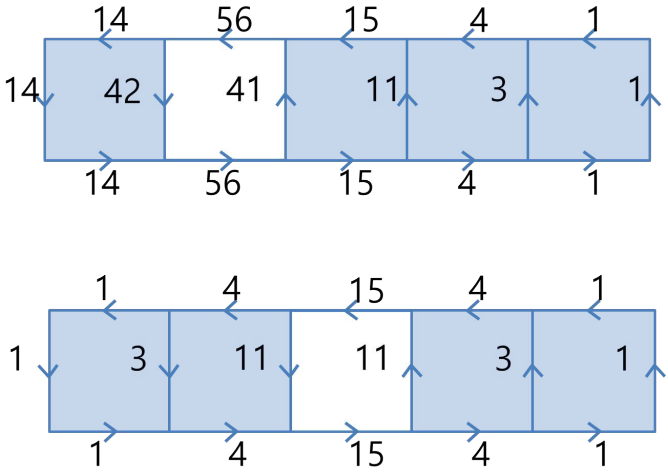
Shape of Data [4)]



Harmonic cycles (조화계)

4) A. Muhammad, and M. Egerstedt, Control using higher order Laplacians in network topologies, In Proc. of 17th International Symposium on Mathematical Theory of Networks and Systems (2006).

Harmonic cycles as shape recognizer



Both shapes have $\beta_1 = 1$ but different harmonic cycles.

Information Centrality

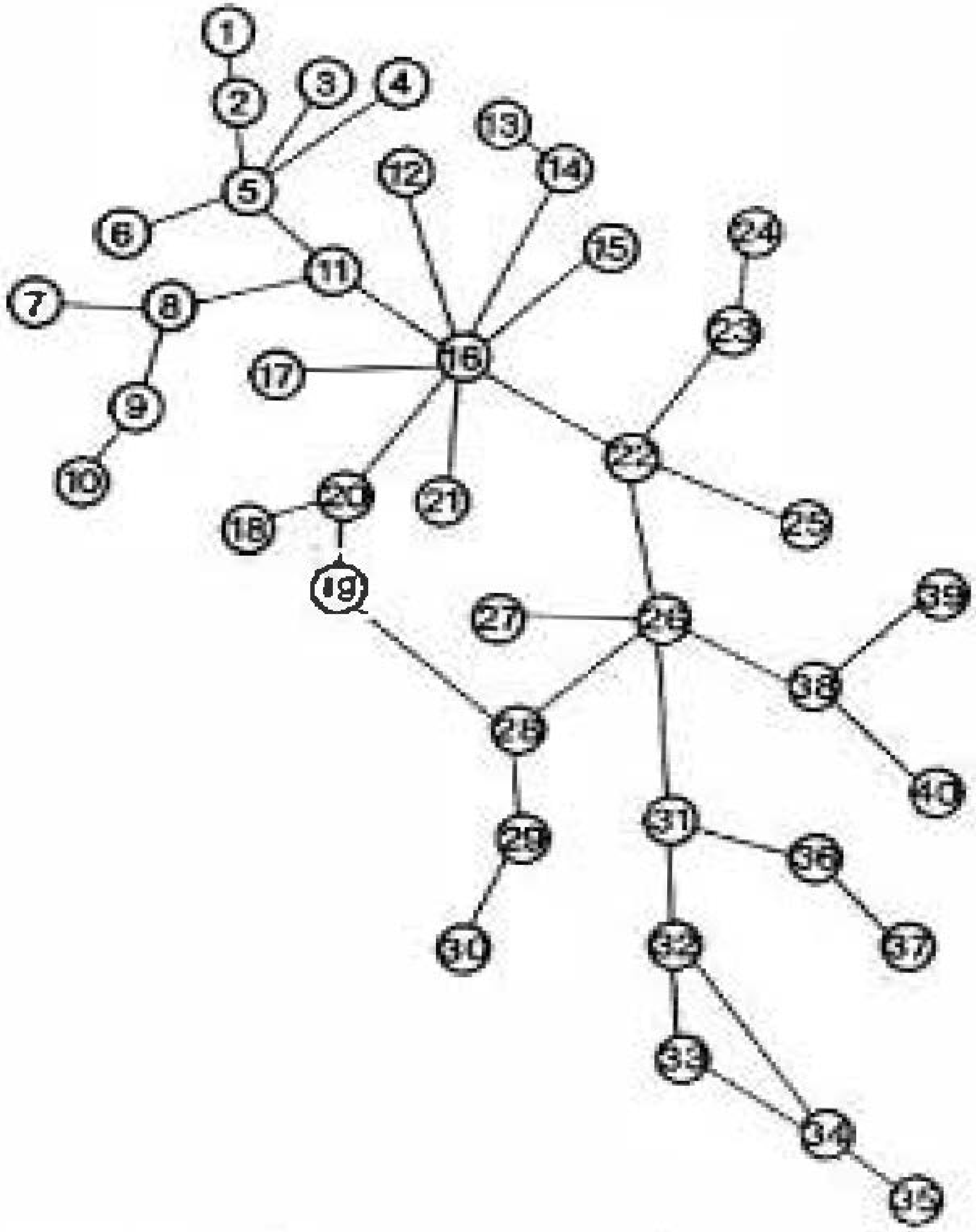
- Combinatorial Green's function $\Delta_0^{-1} = (g_{ij}) \quad$ (K. 2011)

- Information centrality I_{ab} (Stevenson and Zelen 1989)

$$I_{ab} = (g_{aa} + g_{bb} - 2g_{ab})^{-1} = C_{ab}$$

- Information centrality I_a of node a

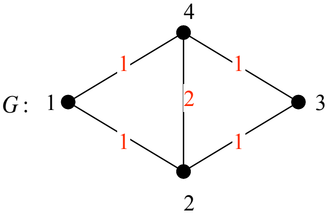
$$I_a = |V(G)| \cdot \Big(\sum_{i \in V(G)} \frac{1}{I_{ai}} \Big)^{-1} \quad = \text{ harmonic mean of } I_{ai}$$



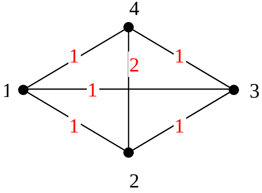
Comparison of centrality measures and rankings

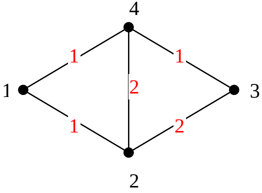
Overall rank	Information	Betweenness	Distance	Degree
1	16 (0.417)	16 (17.7)	16 (0.351)	16 (0.205)
2	22 (0.392)	26 (14.0)	22 (0.345)	26 (0.128)
3	26 (0.388)	22 (13.5)	26 (0.322)	5 (0.128)
4	20 (0.357)	11 (11.6)	11 (0.302)	22 (0.103)
5	11 (0.351)	31 (7.6)	20 (0.281)	8 (0.077)
6	28 (0.348)	5 (6.6)	14 [illegible]	11 (0.077)
7	19 (0.336)	8 (4.1)	19 (0.265)	20 (0.077)
8	31 (0.310)	32 (4.0)	31 (0.265)	28 (0.077)
9	14 (0.303)	28 (3.7)	12 (0.262)	31 (0.077)
10	15 (0.299)	20 (3.1)	15 (0.262)	32 (0.077)

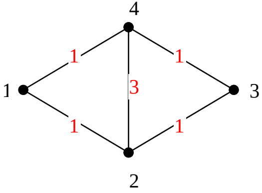
Decentralized network expansion algorithm ($k(G) = 12$)



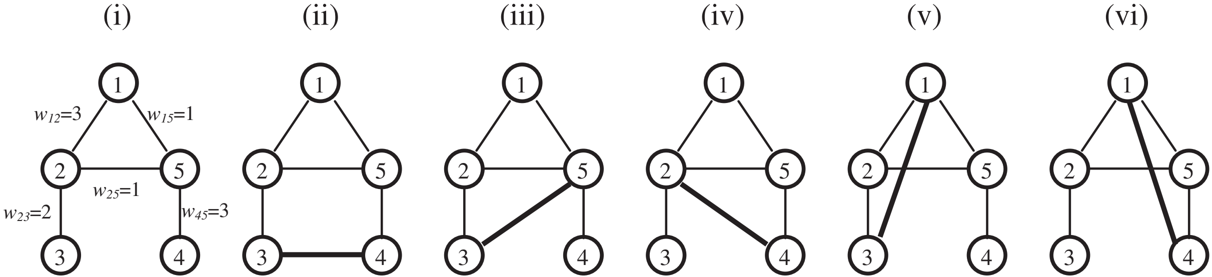
$l_{13} = 1$ $\quad$ $l_{23} = 12/7$ $\quad$ $l_{24} = 3$





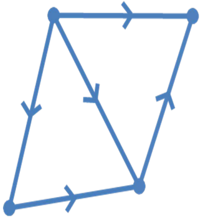


$k(G \cup 13) = 24$ $\quad$ $k(G \cup 23) = 19$ $\quad$ $k(G \cup 24) = 16$

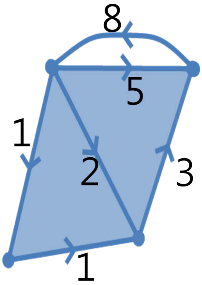


Information Flow	(i)	(ii)	(iii)	(iv)	(v)	(vi)
I_{12}	3.50	3.61	3.63	3.64	4.17	3.64
I_{13}	1.27	1.63	1.71	1.29	2.27	1.29
I_{14}	1.11	1.51	1.23	1.67	1.12	2.11
I_{15}	1.75	2.02	2.07	2.11	1.79	2.50
I_{23}	2.00	2.53	2.64	2.00	2.78	2.00
I_{24}	1.11	1.77	1.34	2.11	1.12	1.67
I_{25}	1.75	2.30	2.42	2.50	1.79	2.11
I_{34}	0.71	1.71	1.18	1.03	0.85	0.91
I_{35}	0.93	1.68	1.93	1.11	1.19	1.03
I_{45}	3.00	3.48	3.00	3.64	3.00	3.64
Mean(I_{ij}) (Std(I_{ij}))	1.71 (0.91)	2.22 (0.76)	2.11 (0.81)	2.11 (0.93)	2.01 (1.05)	2.09 (0.96)
Complexity $c(N)$	42	101	87	80	75	80

Kirchoff's Law



G

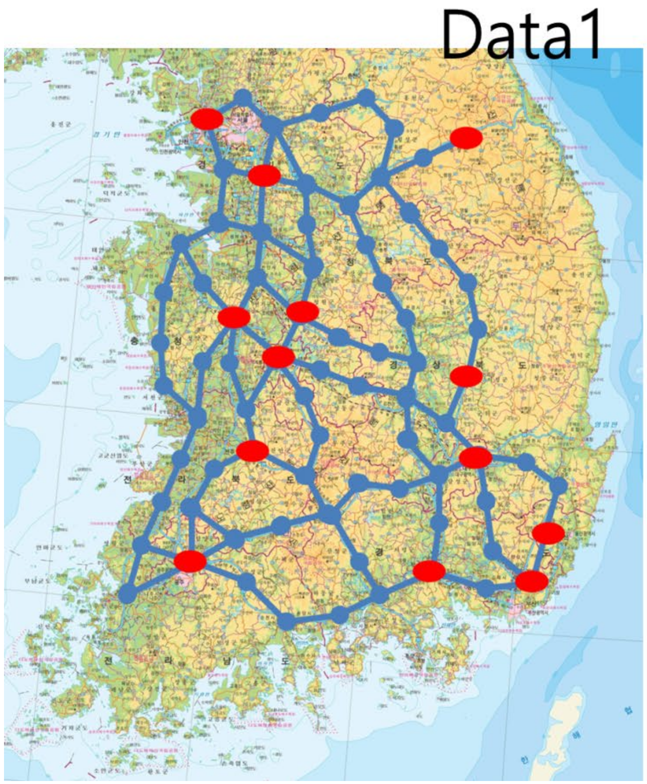


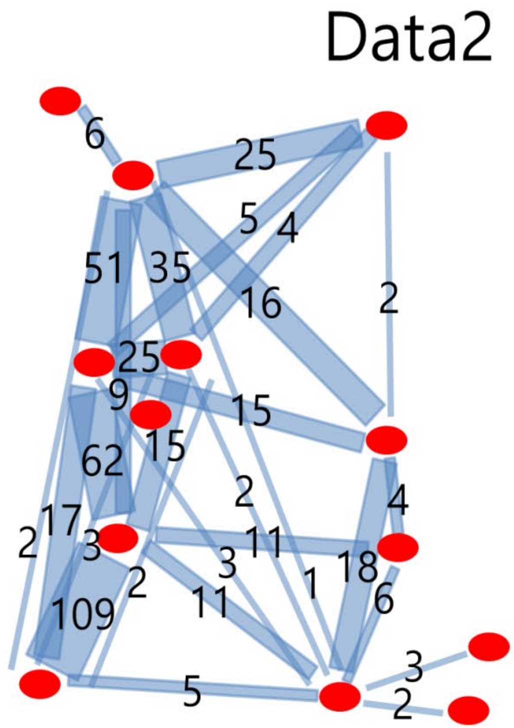
$G^* \cup \omega \cong S^1$

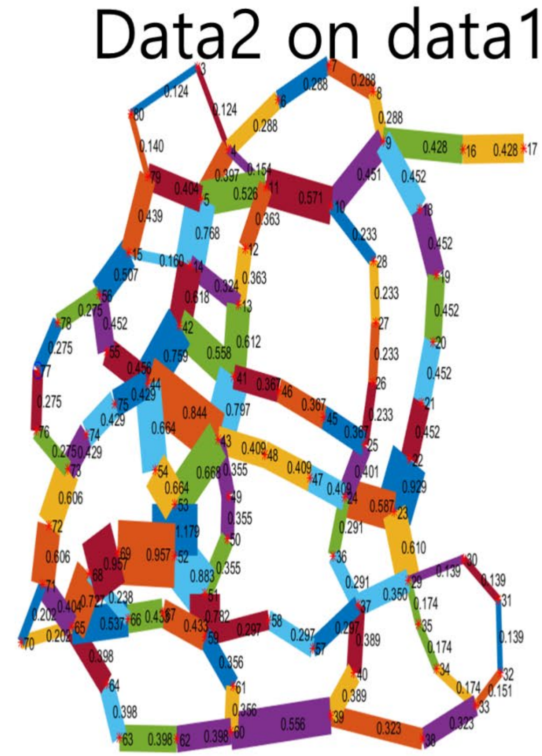
$$I_\omega = I + 8\omega = \partial_1^t \phi + 8\omega = (1,\ -1,\ 3,\ -2,\ -5,\ 8)^t$$

$$I_\omega \in \ker \partial_1 \cap \ker \partial_2^t = \mathscr{H}_1 \cong \mathbb{R}$$

Harmonic Mirroring





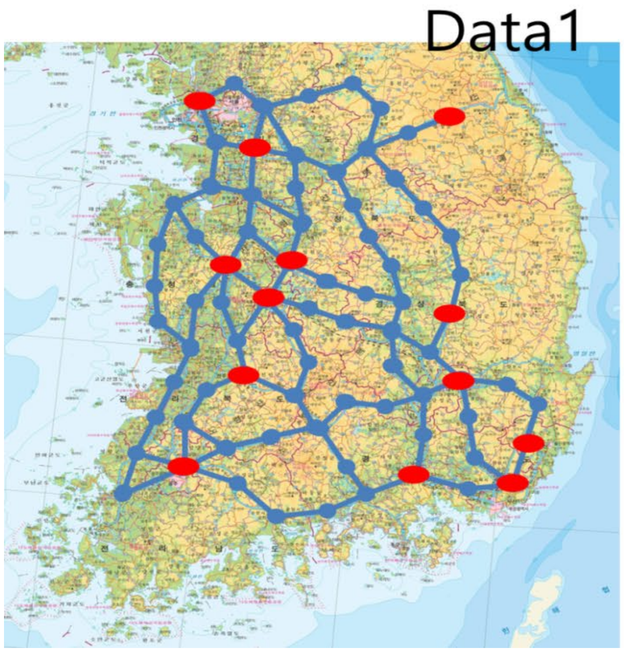


Base Network

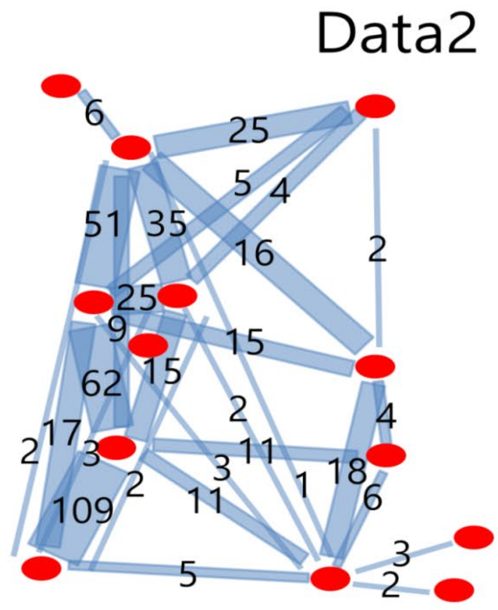
Virtual Network

Implementation
(Harmonic Mirroring)

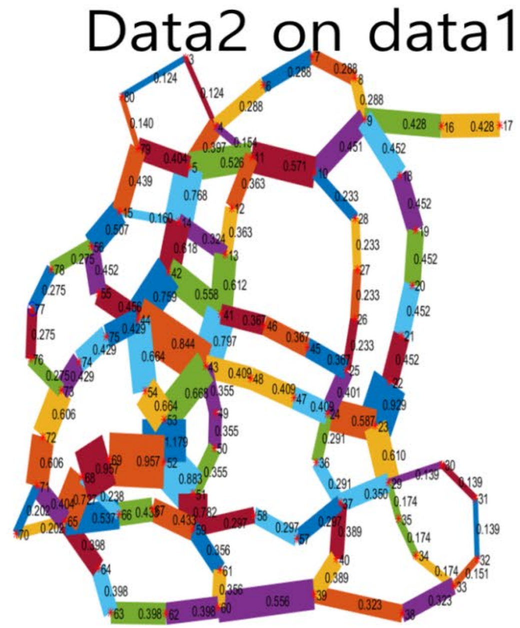
Harmonic Mirroring



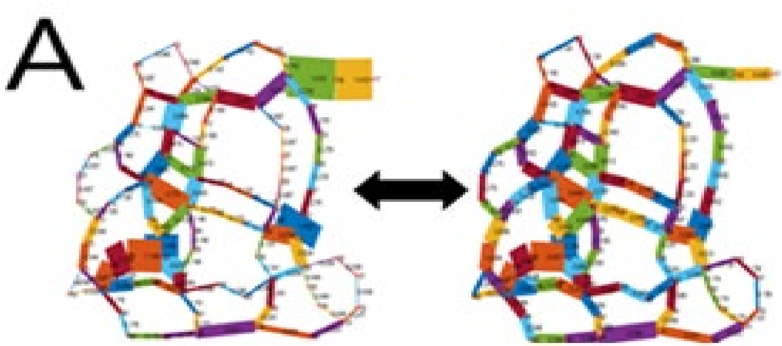
Base Network

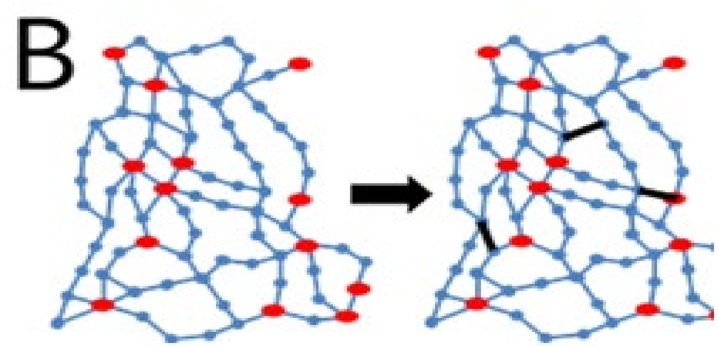


Virtual Network



Implementation
(Harmonic Mirroring)

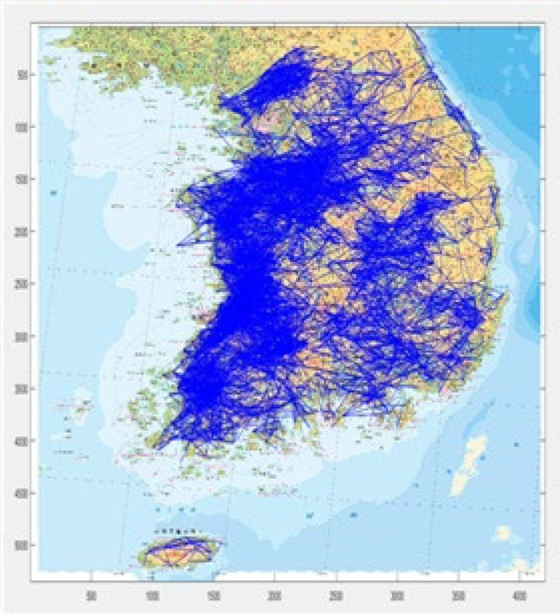




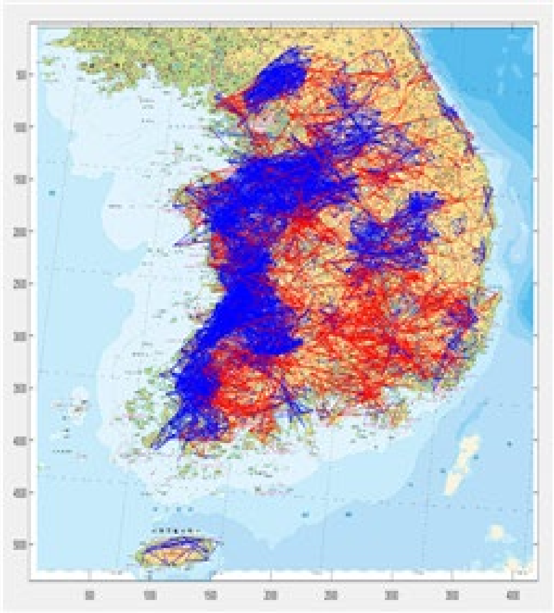
Harmonic Splitting

$$E \equiv \sum_{h_i \in B} h_i^2$$

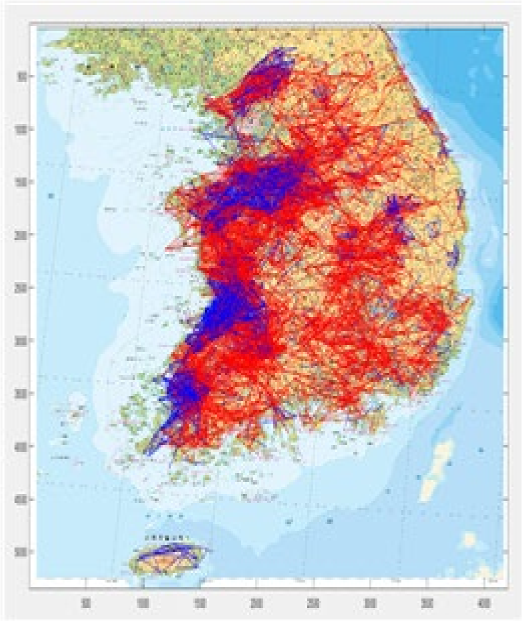
$B =$ orthonomal basis of harmonic cycles



Whole Network



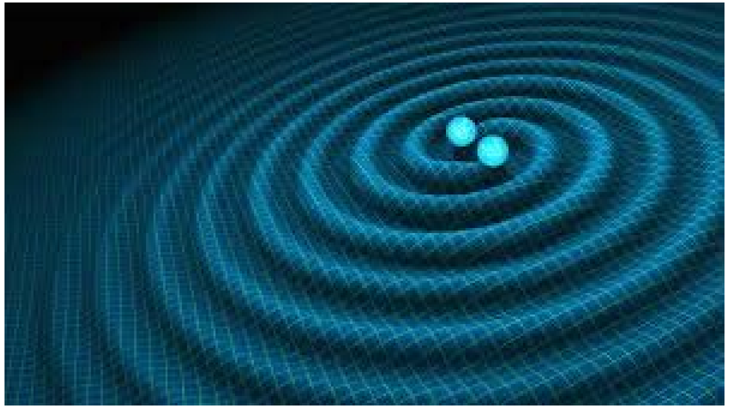
$E > 0.02$



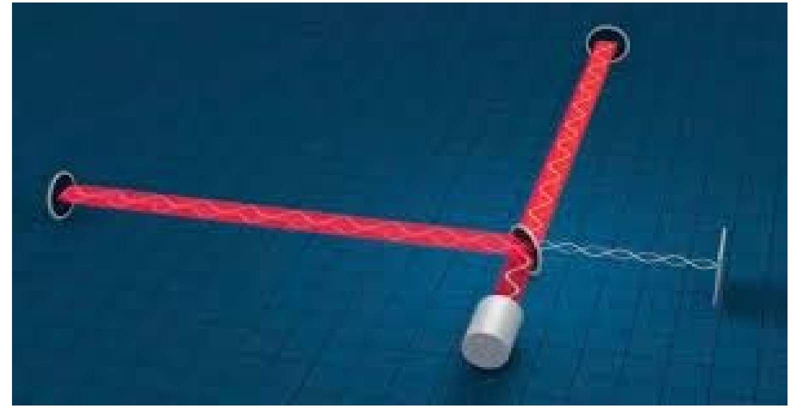
$E > 0.2$

ICIAM 2019 & King of Spain

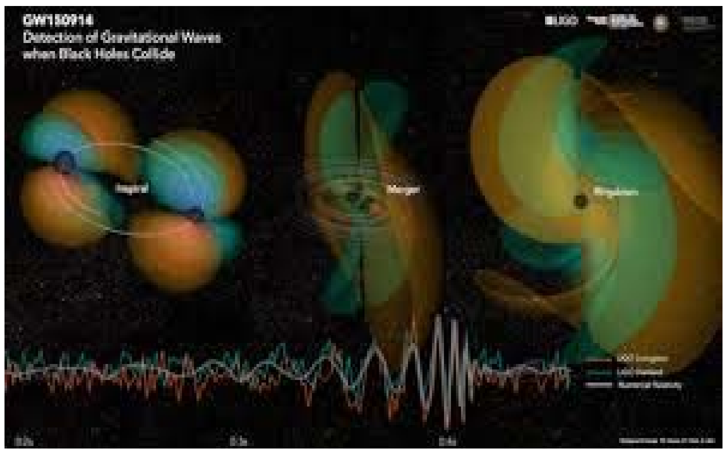
Applications of TDA (Gravitational Waves)



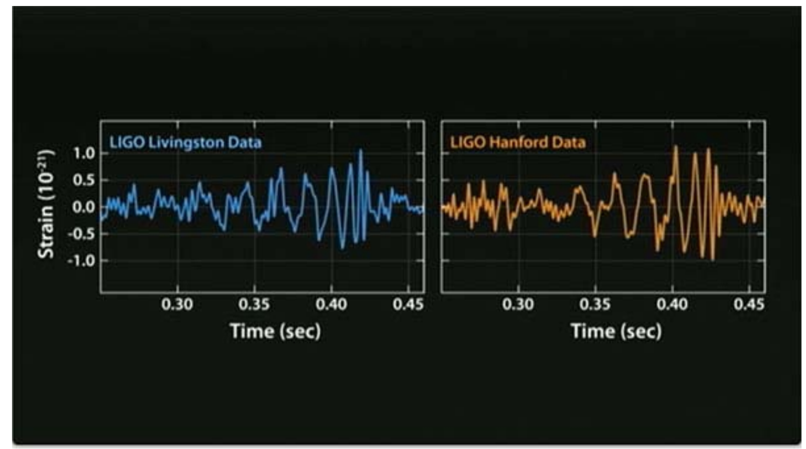
GW generated by binary neutron stars



GW observatory

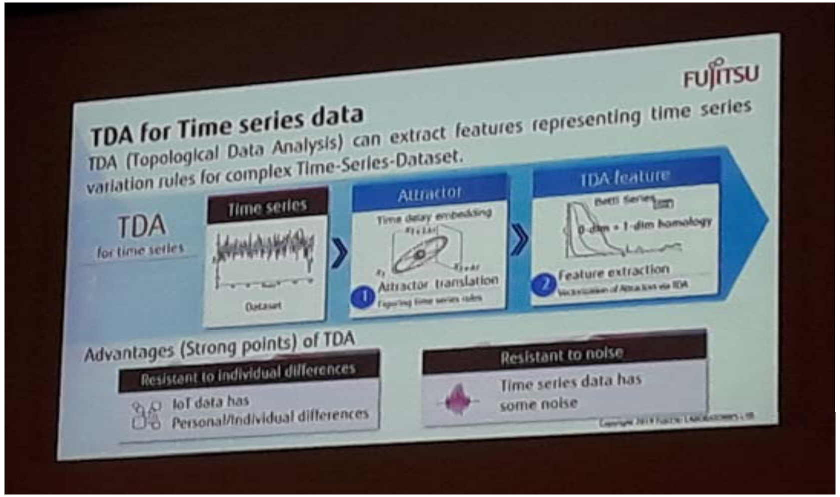


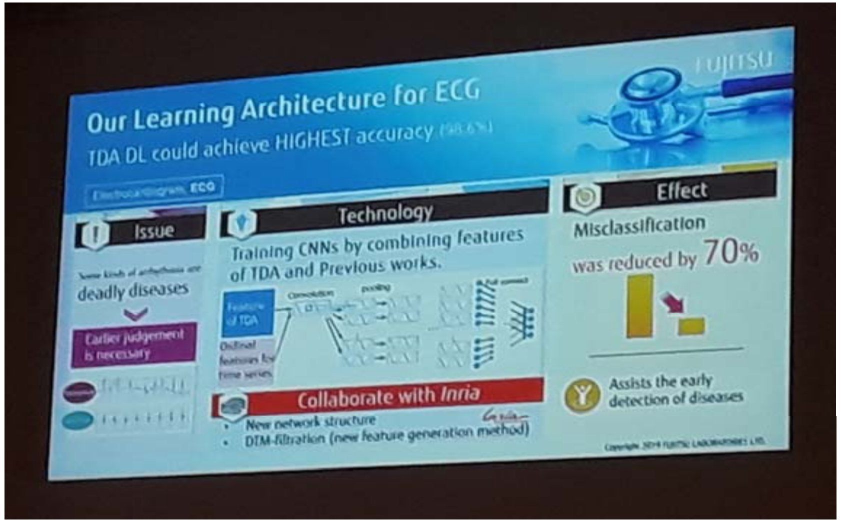
GW detection when black holes collide

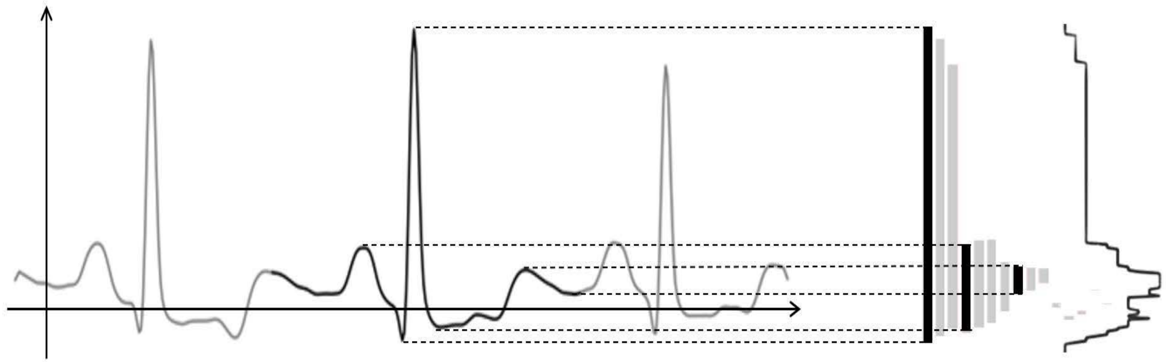


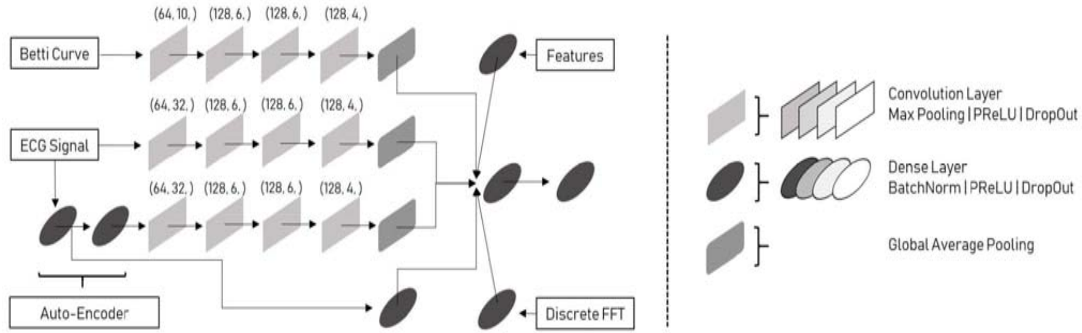
Time series images of GW

Applications of TDA (ECG Data)









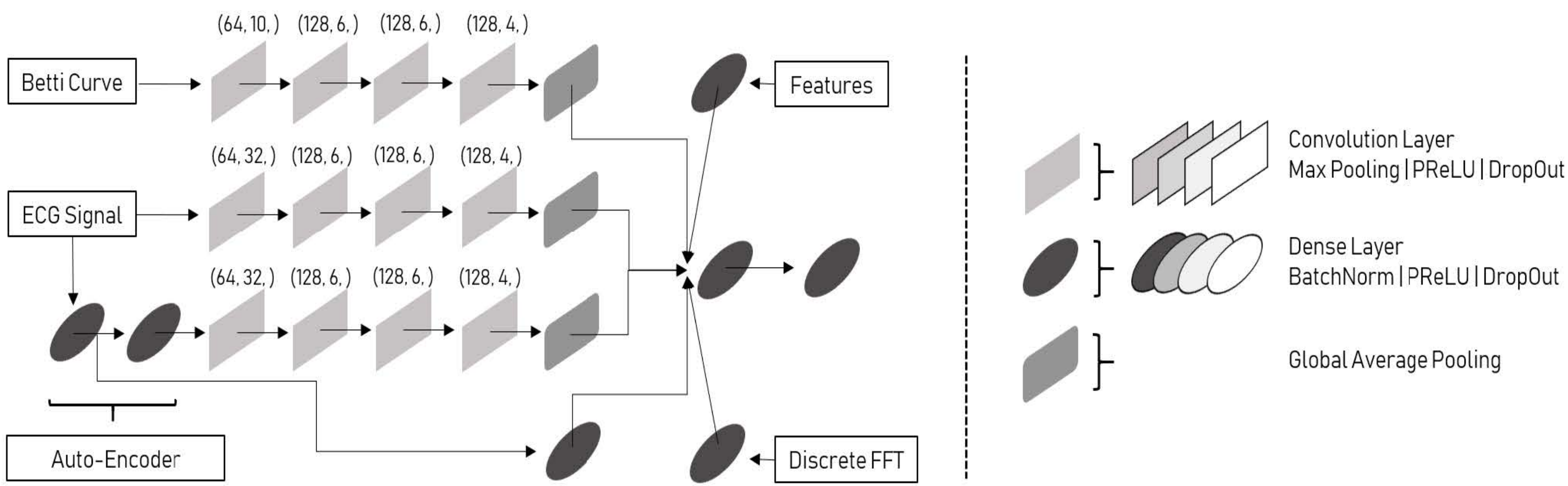
Source: TDA for Arrhythmia Detection through Modular Neural Networks, M. Dindin et al (2019)

Future Applications of Mapview

N mapviews

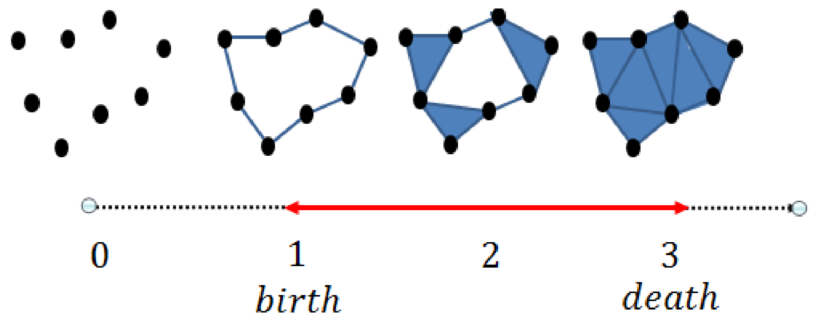
(N=# nodes)

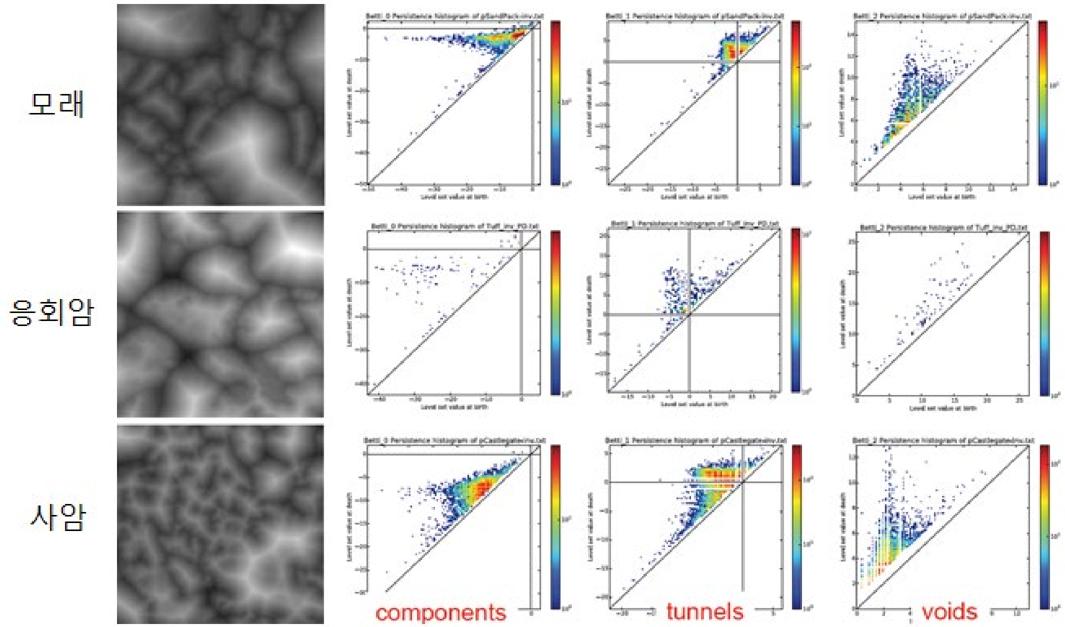
Classification by Machine Learning



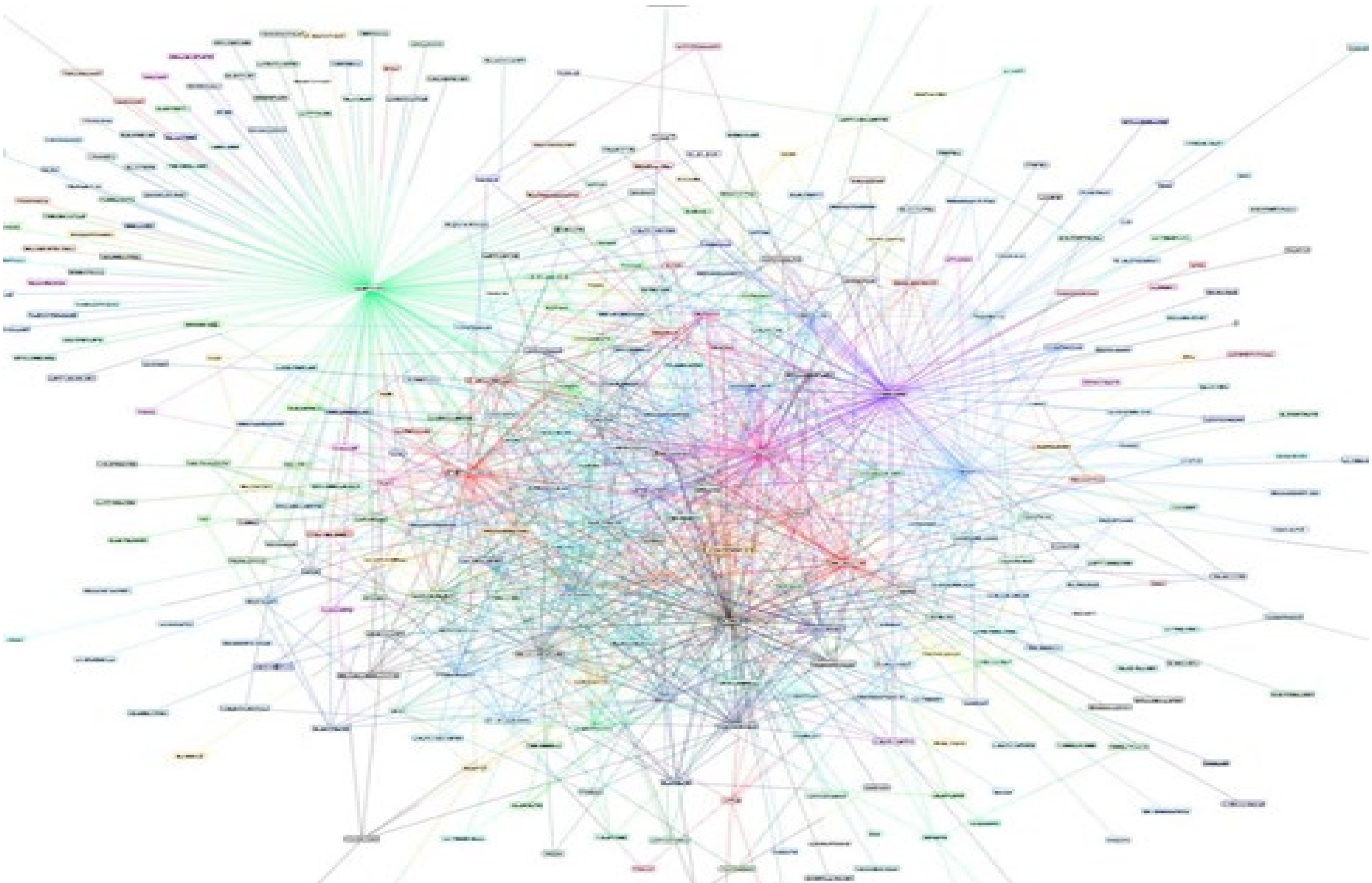
Source: TDA for Arrhythmia Detection, Dindin et al. 2019

Persistent Diagram





Lightening Network



Thank You!