

BÁO CÁO KẾT QUẢ BÀI TẬP LỚN NHÓM 6

?? tài: “Mô hình hóa mạng lưới 63 tỉnh thành Việt Nam”

Nhóm 6 – KTPM1 – K11

*B? môn Công nghệ phần mềm, Khoa Công nghệ thông tin, ?i h?c Công nghi?p Hà N?i

Email: trantd.vn@gmail.com

TÓM TẮT

Việt Nam có 63 tỉnh thành, là một hệ thống lớn nên việc mô hình hóa mạng lưới các tỉnh thành là một việc cần thiết để có một cái nhìn tổng quan và dễ dàng cho việc quản lý, nghiên cứu và tính toán sau này. Tuy nhiên việc ánh xạ các tỉnh hiện nay chưa theo quy luật rõ ràng và chưa liên tục. **Mô hình hóa 63 tỉnh thành Việt Nam** là một hệ thống mạng lưới các tỉnh thành quan trọng của Việt Nam. Nó thể hiện sự gắn kết giữa các tỉnh thành với nhau cũng như sự liên kết tổng quan về mặt địa lý. Đây là vấn đề mang tính cần thiết. Sau khi nghiên cứu thì nhóm đã thu thập dữ liệu của 63 tỉnh thành Việt Nam và đã mô hình hóa mạng lưới 63 tỉnh thành Gephi.

T? khoá: phân tích; quy trình; mạng phần mềm; 63 tỉnh thành Việt Nam

1. GIỚI THIỆU

Tỉnh là một thuật ngữ thường được sử dụng trong Tiếng Việt để chỉ phần lãnh thổ hành chính cấp 1 trực thuộc Quốc Gia. Mô hình hóa là mô phỏng lại nhóm các đối tượng có liên quan đến nhau dưới dạng mạng lưới gồm các node thường được thể hiện bằng các điểm nối thể hiện mối liên hệ giữa các đối tượng.

1.1: Các khái niệm

Mô hình hóa là phương pháp xây dựng mô hình toán của hệ thống bằng cách phân tích hệ thống thành các khối chức năng, trong đó mô hình toán của các khối chức năng đã biết hoặc có thể rút ra được dựa vào các quy luật vật lý, sau đó các khối chức năng được kết nối toán học để tạo thành mô hình của hệ thống.

Các bước mô hình hóa:

- + Phân tích chức năng
- + Phân tích vật lý
- + Phân tích toán học

1.2: Lý do thực hiện ?? tài

Các tỉnh thành chưa được mô hình hóa và việc ánh xạ chưa hợp lý gây khó khăn cho công tác nghiên cứu tìm hiểu. Vì vậy trước tiên cần thu thập dữ liệu về 63 tỉnh thành, các dữ liệu về bản đồ chính xác nhất. Sau đó nghiên cứu thuật toán 63 tỉnh. Tiếp theo là tìm hiểu về thuật toán mô hình hóa mạng lưới bằng các phần mềm chuyên dụng như Gephi,

CytoScape,... Tiếp theo hành mô phỏng các hệ thống phần mềm. Ta cần tiến hành song song với việc xem lại việc ánh xạ các tỉnh, để nghiên cứu thuật toán ánh xạ mới để phù hợp. Cuối cùng áp dụng vào mạng lưới và hoàn thiện kết quả ?? tài.

2. PHƯƠNG PHÁP VÀ T? LI?U

2.1. D? liệu mạng lưới

Bản đồ hành chính Việt Nam

(Số liệu của **Tổng cục Thống kê** tính đến 0 giờ ngày 1/4/2019)

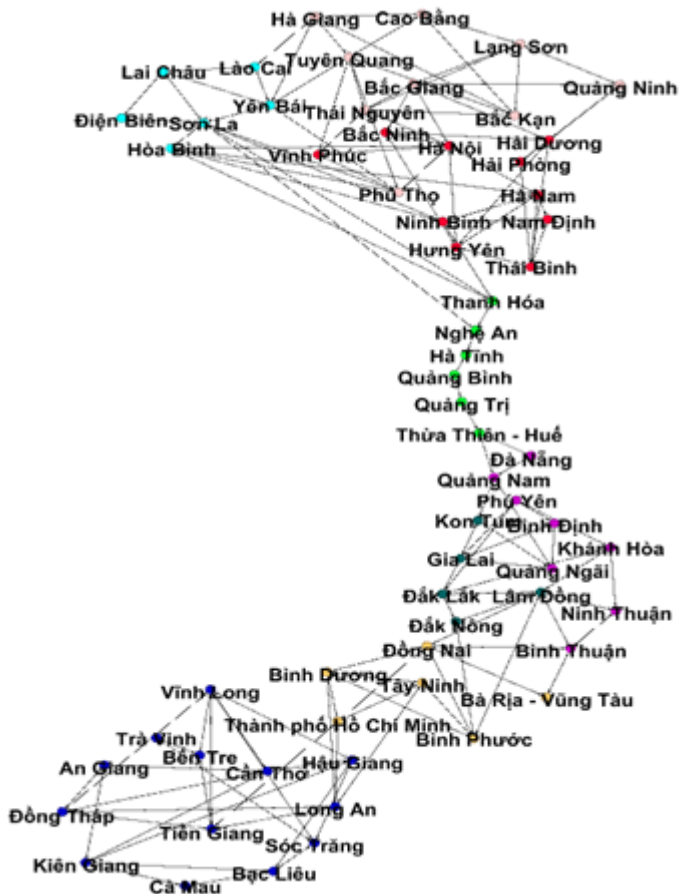
Theo k?t qu? ?i?u tra dân s? ngày 1/4/2019

https://vi.wikipedia.org/wiki/T%E1%BB%89nh_th%C3%A0nh_Vi%E1%BB%87t_Nam#L?ch_s?

Đây là công thông tin d? li?u uy tín, là m?t bách khoa toàn th? t? do.

Nhóm mình tích mô hình hoá mạng lưới các tỉnh thành Việt Nam và phân tích để nêu, đánh giá và tính toán sau này.

Coi m?i t?nh t?ng ?ng v?i 1 Node, ?i?n d? li?u vào datatable theo th? t? trên b?n ??, sau ?ó m? hình hoá l?i b?ng Gephi



2.2 Mô hình mạng là gì?

Kiểm tra xem có thể biểu diễn bằng các mô hình node có dạng $y = a(x^b)$

Nếu: Correlation thu được $[0.5; 1] \Rightarrow$ là mạng Scale-free

Kiểm tra $L = \log(N) \Rightarrow$ là mạng Small-world. Trong đó:

ID	Lable	Modularity	ID	Lable	Modularity	ID	Lable	Modularity
4		0	49	Quảng Ninh	1	11	Bình Thuận	3
13	Cao Bằng	0	52	Sơn La	1	32	Khánh Hòa	3
22	Hà Giang	0	54	Thái Bình	1	38	Lâm Đồng	3
35	Lai Châu	0	55	Thái Nguyên	1	39	Long An	3
36	Lào Cai	0	56	Thanh Hóa	1	43	Ninh Thuận	3
60	Tuyên Quang	0	62	Vĩnh Phúc	1	53	Tây Ninh	3
63	Yên Bái	0	1	An Giang	2	31	Thành Phố HCM	3
20	Điện Biên	0	3	Bạc Liêu	2	16	Đắk Lắk	3
5	Bắc Giang	1	7	Bến Tre	2	17	Đắk Nông	3
6	Bắc Ninh	1	12	Cà Mau	2	18	Đồng Nai	3
26	Hải Dương	1	14	Cần Thơ	2	9	Bình Định	4
27	Hải Phòng	1	29	Hậu Giang	2	21	Gia Lai	4

23	Hà Nam	1	33	Kiên Giang	2	25	Hà T?nh	4
24	Hà N?i	1	51	Sóc Tr?ng	2	34	Kon Tum	4
28	Hòa Bình	1	58	Ti?n Giang	2	45	Phú Yên	4
30	H?ng yên	1	59	Trà Vinh	2	46	Qu?ng Bình	4
37	L?ng S?n	1	61	V?nh Long	2	47	Qu?ng Nam	4
40	Nam ??nh	1	19	??ng Thấp	2	48	Qu?ng Ngãi	4
41	Ngh? An	1	2	Bà R?a –V?ng Tàu	3	50	Qu?ng Tr?	4
42	Ninh Bình	1	8	Bình D??ng	3	57	Th?a Thiên Hu?	4
44	Phú Th?	1	10	Bình Ph??c	3	15	?à N?ng	4

Các modularity Class

$\gamma(G) = \frac{1}{n|S|} \sum_{s \in S} \sum_{i=1}^n I(\langle s \rangle = \langle s_{v_i} \rangle),$
 $\prod_{V_i} \langle s \rangle$??i di?n cho m?t nh? ?i?u hành d? báo ?? trích xu?t các r?t thu hút khách m?t ph?n c?a m?t t?p h?p con cho V_i ?
 V t? $\langle s \rangle$ và $H(\langle s \rangle, \langle s' \rangle)$. ??a ra $\langle s \rangle = s_0 \rightarrow s_1 \rightarrow \dots \rightarrow s_{l-1}$ và $\langle s' \rangle = s'_0 \rightarrow s'_1 \rightarrow \dots \rightarrow s'_{l'-1}$ (1 ? l ? l' ???c gi? s?
không có tính t?ng quát).

$$H(\langle s \rangle, \langle s' \rangle) = \frac{1}{l'} \sum_{j=0}^{l-1} \left(1 - \frac{h(s_j, s'_j)}{K} \right),$$

h là kho?ng cách Hamming và K là kích th??c c?a các tr?ng thái . ?? b?n trong mô-?un th? hi?n m?c ?? mô-?un ch?u
??t bi?n duy trì m?c ?? b?n m? c?c b?. Khi ch? có m?t mô-?un trong m?ng, s? b?n trong mô-?un là t??ng t? nh? m?ng
l??i b?n ? (G).

Mô t? công th?c tính centrality c?a Degree, Closeness, Betweeness, PageRank

Cách làm: S? d?ng Gephi: Sau ?ó

Ch?y các cái nh? hình v? sau chúng ta s?
có k?t qu? tính

File Home View

Gephi 0.9.2

File Workspace View Tools Window Help

Overview Data Laboratory Preview

Appearance x

Nodes Edges

Unique Partition Ranking

#c0c0c0

Graph x Data Table x

Nodes Edges Configuration Add node Add edge Search/Replace Import Spreadsheet

Id	Label
21	Gia Lai
22	Hà Giang
23	Hà Nam
24	Hà Nội
25	Hà Tĩnh
26	Hải Dương
27	Hải Phòng
28	Hòa Bình
29	Hậu Giang
30	Hưng Yên
31	Thành phố Hồ Chí Minh
32	Khánh Hòa
33	Kiên Giang
34	Kon Tum
35	Lai Châu
36	Lào Cai
37	Lạng Sơn
38	Lâm Đồng
39	Long An
40	Nam Định
41	Nghệ An
42	Ninh Bình
43	Ninh Thuận
44	Phú Thọ

842, 270px 1366 x 768px

Type here to search

Ranking	Degree centrality		Betweenness centrality	
	Node name	Characteristics in the real system	Node name	Characteristics in the real system
Highest ranking 1	à N?ng	à N?ng	Bà R?a- V?ng Tàu	Bà R?a- V?ng Tàu
Highest ranking 2	Bà R?a- V?ng Tàu	Bà R?a- V?ng Tàu	Cà Mau	Cà Mau
Highest ranking 3	Cà Mau	Cà Mau	à N?ng	à N?ng
Highest ranking 4	?i?n Biên	?i?n Biên	?i?n Biên	?i?n Biên
Highest ranking 5	Hà T?nh	Hà T?nh	Trà Vinh	Trà Vinh

Lowest ranking 5	Lâm ??ng	Lâm ??ng	Hà T?nh	Hà T?nh
Lowest ranking 4	Tuyên Quang	Tuyên Quang	Qu?ng Bình	Qu?ng Bình
Lowest ranking 3	V?nh Long	V?nh Long	Qu?ng Tr?	Qu?ng Tr?
Lowest ranking 2	Hà N?i	Hà N?i	Qu?ng Nam	Qu?ng Nam
Lowest ranking 1	S?n La	S?n La	Th?a Thiên Hu?	Th?a Thiên Hu?

Ranking	Closeness centrality		PageRank centrality	
	Node name	Characteristics in the real system	Node name	Characteristics in the real system
Highest ranking 1	Cà Mau	Cà Mau	?à N?ng	?à N?ng
Highest ranking 2	B?c Liêu	B?c Liêu	?i?n Biên	?i?n Biên
Highest ranking 3	Cà Mau	Cà Mau	Cà Mau	Cà Mau
Highest ranking 4	Kiên Giang	Kiên Giang	Bà R?a V?ng Tàu	Bà R?a V?ng Tàu
Highest ranking 5	Trà Vinh	Trà Vinh	Lào Cai	Lào Cai
Lowest ranking 5	Kon Tum	Kon Tum	Gia Lai	Gia Lai
Lowest ranking 4	Qu?ng Ngãi	Qu?ng Ngãi	??ng Nai	??ng Nai
Lowest ranking 3	Qu?ng Tr?	Qu?ng Tr?	Hà N?i	Hà N?i
Lowest ranking 2	Qu?ng Nam	Qu?ng Nam	V?nh Long	V?nh Long
Lowest ranking 1	Th?a Thiên Hu?	Th?a Thiên Hu?	Lâm ??ng	Lâm ??ng

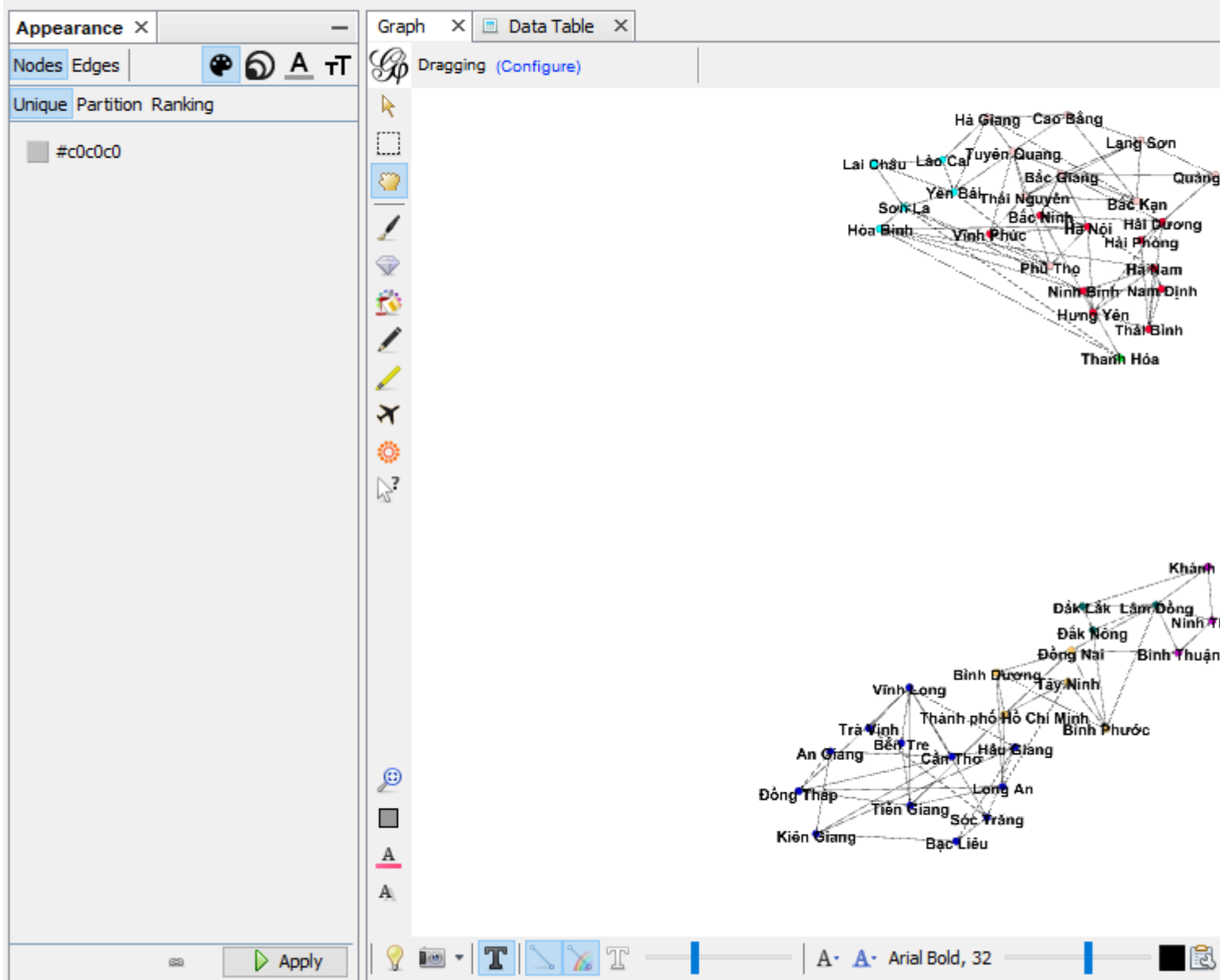
B?ng Centrality of network
Mô t? ph??ng pháp phát hi?n K-core c?a m?ng

S? d?ng Gephi ?? tìm K-Core c?a m?ng:

Ch?n Filters -> Ch?n Topology -> K-Core

Sau ?ó Ch?n Filter. Nâng K-Core Setting lên t?i khi m?ng d?n d?n bi?n m?t thì K-Core cu?i cùng tr??c khi m?ng bi?n m?t chính là K-core c?a m?ng.

T?i: K-Core : 6



3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Phân tích kết quả K-core của mạng

K-core của mạng: 6

Dùng K-core để xác định vùng có số liên kết chặt chẽ nhất trong mạng là 63 tỉnh thành.

Trên hình trên có thể thấy mạng là 63 tỉnh thành được tách thành 2 vùng rõ rệt Bắc – Nam với số liên kết chặt chẽ của các Node trong hai vùng các tỉnh thành là Miền Bắc và Miền Nam.

Vậy phương pháp thực hiện: Sử dụng Gephi để xác định K-Core của mạng là 6 này với công cụ hỗ trợ phân tích mạng là Gephi. Vậy lõi của mạng tập trung và liên kết chặt chẽ ở phía Bắc và phía Nam.

Vậy bộ phận của mạng rất cao vì chúng liên kết khá chặt chẽ ở 2 vùng Bắc – NAM.

3.2. Ý nghĩa của các nút trong mạng

Các node trong lõi của mạng là những node quan trọng nhất có thể liên kết với tất cả các node trong mạng. Các tác nhân trên có thể thấy, trên mạng là những

sánh 10 nút cao nhất và 10 nút thấp nhất.

*Tìm mối quan hệ giữa các centrality và các thông số khác (GDP, Dân số, mật độ dân số).

***Chú thích:** -**GDP:** USD/ng/năm – Số liệu: 2018

- **Dân số:** người – Số liệu :0h 1/4/2019

- **Mật độ:** người/km2 – Số liệu: 0h 1/4/2019

–Nguồn: https://vi.wikipedia.org/wiki/T%E1%BB%89nh_th%C3%A0nh_Vi%E1%BB%87t_Nam

Label	Closness	Betweenness	Degree	PageRank	GDP	Dân số	Mật độ
An Giang	0.0947205	79.21511211	6	0.010698	1491	1.908.352	610
Bà Rịa - Vũng Tàu	0.11296296	0	4	0.009059	5837	1.148.313	580
Bắc Liêu	0.08739255	25.68626984	8	0.015465	1826	907.236	335
Bắc Kạn	0.1101083	28.12284729	10	0.015384	1303	313.905	66
Bắc Giang	0.11151737	168.8090712	12	0.018442	2230	1.803.950	433
Bắc Ninh	0.11030741	18.12761862	8	0.012967	6519	1.368.840	1.467
Bến Tre	0.1014975	50.61089466	6	0.012313	1433	1.288.463	527
Bình Định	0.11641221	121.9966332	9	0.017766	5681	2.426.561	761
Bình Định	0.13436123	46.11904762	6	0.013739	2115	1.486.918	251
Bình Phước	0.1244898	437.3492766	10	0.019185	2520	994.679	140
Bình Thuận	0.12248996	58.03809524	8	0.016116	2215	1.230.808	201
Cà Mau	0.08144192	0	4	0.008788	1880	1.194.476	231
Cao Bằng	0.10912343	21.58205898	8	0.012759	1160	530.341	109
Chợ Mới	0.09546166	110.9290374	10	0.016843	3496	1.235.171	1.007
Đà Nẵng	0.13716814	0	2	0.002381	3612	1.134.310	876
Đắk Lắk	0.138322	1687.773212	10	0.019535	1781	1.869.322	145
Đắk Nông	0.12978723	227.0301508	6	0.0104	1965	622.168	95
Đảng Nai	0.12655602	1040.089202	12	0.023001	4226	3.097.107	529
Đảng Tháp	0.10304054	356.4996659	9	0.013721	1737	1.599.504	499
Định Biên	0.12224449	0	4	0.006773	1186	598.856	59
Gia Lai	0.14186047	1765.377578	10	0.021635	1970	1.513.847	92
Hà Giang	0.11685824	74.77165971	10	0.015442	899	854.679	104
Hà Nam	0.11914063	245.3069015	12	0.01994	2397	852.8	943
Hà Nội	0.12127237	560.2706409	16	0.023489	4080	8.053.663	2.2
Hà Tĩnh	0.1408776	1863	4	0.014703	2150	1.288.866	212
Hải Định	0.10252101	27.83750567	12	0.019338	2445	1.705.059	1.075
Hải Phòng	0.10099338	5.482627517	6	0.010887	4217	1.837.173	1.289

Hòa Bình	0.12951168	696.8570729	12	0.018784	2098	854.131	182
H?u Giang	0.0947205	89.22832778	10	0.017803	1664	733.017	477
H?ng Yên	0.11050725	47.25477141	10	0.016097	2402	1.252.731	1.591
Thành ph? H? Chí Minh	0.11890838	885.4332365	10	0.018681	6725	8.993.082	4.298
Khánh Hòa	0.12815126	88.29821694	8	0.015867	2698	1.231.107	237
Kiên Giang	0.08815029	102.7208231	10	0.018332	2094	1.723.067	281
Kon Tum	0.14352941	855.5238095	6	0.015394	1628	540.438	53
Lai Châu	0.12423625	50.05	8	0.012306	1433	460.196	48
Lào Cai	0.11509434	4.801731602	6	0.010046	2686	730.42	108
L?ng S?n	0.10304054	45.20289486	10	0.015718	1668	781.655	93
Lâm ??ng	0.13347921	1404.1021	14	0.025921	2595	1.296.606	156
Long An	0.11050725	416.9756952	9	0.014666	2980	1.688.547	333
Nam ??nh	0.11509434	57.10277077	6	0.011586	2258	1.780.393	1.111
Ngh? An	0.138322	1846	5	0.013859	1591	3.327.791	189
Ninh Bình	0.12525667	255.7246974	9	0.016372	2106	982.487	692
Ninh Thu?n	0.11914063	3.519047619	6	0.012317	1724	590.467	180
Phú Th?	0.12978723	335.7219738	12	0.016804	1672	1.463.726	393
Phú Yên	0.13203463	106.0271839	8	0.01664	173	872.964	179
Qu?ng Bình	0.14285714	1876	4	0.015116	1628	895.43	110
Qu?ng Nam	0.14489311	1890	7	0.018303	2658	1.495.812	141
Qu?ng Ngãi	0.14386792	965.4761905	8	0.019504	2510	1.231.697	245
Qu?ng Ninh	0.10234899	25.27543244	8	0.013491	5110	1.320.324	200
Qu?ng Tr?	0.14420804	1885	4	0.015211	1894	632.375	132
Sóc Tr?ng	0.09442724	68.39614857	8	0.015055	1629	1.199.653	396
S?n La	0.138322	984.3487163	16	0.018787	1650	1.248.415	86
Tây Ninh	0.11708253	237.4161343	8	0.01508	2727	1.169.165	278
Thái Bình	0.10892857	130.8548532	10	0.017052	1650	1.860.447	1.14
Thái Nguyên	0.11192661	94.26118853	12	0.017846	3375	1.286.751	354
Thanh Hóa	0.13525499	1191.755556	8	0.016193	1785	3.640.128	319
Th?a Thiên- Hu?	0.14489311	1890	5	0.015058	1770	1.128.620	228
Ti?n Giang	0.11050725	694.474327	9	0.016821	2037	1.764.185	696
Trà Vinh	0.09399076	2.066666667	6	0.012021	1911	1.009.168	442
Tuyên Quang	0.11914063	184.587315	14	0.020273	1564	784.811	131
V?nh Long	0.10286678	459.6279176	14	0.024281	1946	1.022.791	688

V?nh Phúc	0.11914063	16.09083875	8	0.012251	3757	1.154.154	871
Yên Bái	0.12761506	303.7992562	2	0.017091	1459	821.03	116
Correlation	-0.0814462	-0.104093206	0.015526	-0.00199	(GDP)		
Correlation	-0.0059703	0.133029882	0.299117	0.509937		(Dân s?)	
Correlation	-0.1474381	-0.159375491	-0.05324	-0.0891			(M?t??)

4. K?T LU?N

Trong nghiên c?u này, chúng em t?p trung phân tích toàn di?n mô hình hóa 63 t?nh thành Vi?t Nam, t? ?ó h? tr? quá trình ?i?u tra c?ng nh? phân tích ??c ?i?m t?ng t?nh thành. V??t qua h?n ch? c?a các ph??ng pháp phân tích khác, ph??ng pháp c?a chúng em ?ã cho phép th?m ??nh mô hình 63 t?nh thành, phân tích ??c tính v? trí ??a lý và phân tích l??ng dân s? c?a t?ng t?nh thành. Th?m ??nh ki?n trúc cho th?y: mô hình 63 t?nh thành Vi?t Nam là m?t h? th?ng ??n nhi?m b?n. K?t qu? phân m?ng l??i 63 t?nh thành là m?t ngu?n tham kh?o tin c?y d? dành cho vi?c qu?n lý, nghiên c?u và tính toán sau này. C? th?, nhóm chúng em ?ã tìm ra s? t??ng quan gi?a các centrality và các thông s? khác, rõ nh?t th? hi?n ? “Dân s?” v?i **Correlation ??t : 0.509937**.

L?i c?m ?n. Nhóm 6 chúng em trân tr?ng c?m ?n th?y Tr?n Ti?n D?ng, Khoa CNTT, ??i h?c Công nghi?p Hà N?i ?ã h??ng d?n th?c hi?n ?? tài này.

5. TÀI LI?U THAM KH?O

1. Tranoris, C. and K. Thramboulidis. *Integrating UML and the function block concept for the development of distributed control applications*. in *Emerging Technologies and Factory Automation, 2003. Proceedings. ETFA '03. IEEE Conference*. 2003.

2. Gilbreth, F.B. and L.M. Gilbreth, *Process Charts*. American Society of Mechanical Engineers, 1921.

3. Dijkman, R., J. Hofstetter, and J. Koehler, *Business Process Model and Notation*. Third International Workshop, BPMN 2011, Lucerne, Switzerland, November 21-22, 2011. Proceedings, 2011.

4. VAN DER AALST, W.M.P., *THE APPLICATION OF PETRI NETS TO WORKFLOW MANAGEMENT*. Journal of Circuits, Systems and Computers, 1998. **08**(01): p. 21-66.

5. Agerwala, T. and M. Flynn, *Comments on capabilities, limitations and “correctness” of Petri nets*. SIGARCH Comput. Archit. News, 1973. **2**(4): p. 81-86.

6. Ganguly, N., A. Deutsch, and A. Mukherjee, *Dynamics On and Of Complex Networks Applications to Biology, Computer Science, and the Social Sciences*. Book, 2009.

7. Tran, T.D. and Y.K. Kwon, *Hierarchical closeness efficiently predicts disease genes in a directed signaling network*. Comput Biol Chem, 2014. **53PB**: p. 191-197.

8. Tran, T.D. and Y.K. Kwon, *The relationship between modularity and robustness in signalling networks*. J R Soc Interface, 2013. **10**(88): p. 20130771.

9. T.D. Tran, C.D. Truong, and Y.K. Kwon, *MORO: a Cytoscape App for Relationship Analysis between Modularity and Robustness in Large-Scale Biological Networks* BMC Systems Biology, 2016.

10. Ibarra, H., *Network Centrality, Power, and Innovation Involvement: Determinants of Technical and Administrative Roles*. Academy of Management Journal, 1993. **Vol. 36**(No. 3): p. 471-501

11. Marsden, P.V., *Egocentric and sociocentric measures of network centrality*. Social Networks, 2002. **24**(4): p. 407-422.
12. Freeman, L.C., *Centrality in social networks conceptual clarification*. Social Networks, 1978. **1**(3): p. 215-239.
13. Kwon, Y.J., et al., *Kinome siRNA screen identifies novel cell-type specific dengue host target genes*. Antiviral Res, 2014. **110**: p. 20-30.
14. Leicht, E.A. and M.E.J. Newman, *Community Structure in Directed Networks*. Physical Review Letters, 2008. **100**(11): p. 118703.
15. Noack, A., *Modularity clustering is force-directed layout*. Phys Rev E Stat Nonlin Soft Matter Phys, 2009. **79**(2 Pt 2): p. 026102.
16. Kwon, Y.K. and K.H. Cho, *Analysis of feedback loops and robustness in network evolution based on Boolean models*. BMC Bioinformatics, 2007. **8**: p. 430.
17. Venkatesh, K.V., S. Bhartiya, and A. Ruhela, *Multiple feedback loops are key to a robust dynamic performance of tryptophan regulation in Escherichia coli*. FEBS Lett, 2004. **563**(1-3): p. 234-40.
18. https://vi.wikipedia.org/wiki/T%E1%BB%89nh_th%C3%A0nh_Vi%E1%BB%87t_Nam