Author’s response

We got review from two reviewers during discussion interval. The reviewers talked about many factors in the paper. We tried to reflect comments on revised manuscript. We believe that the paper is substantially improved as the result of the revision. We wrote point-by-point response (in blue) to reviewer’s comments (in red). Also, we put revised part of the manuscript in black.

<Major comments>

1) As also the authors point out, the complex network framework is widely used to analyse the spatial dependence of precipitation. Thus, its use does not provide novelty. The authors should better clarify the original contribution of the study. Has the study area been analysed by complex network yet? Are mutual information and multiresolution community detection novel frameworks? Has the study ground-breaking results?

   ➢ Thank you for your good comment. As you mentioned, novelty is fundamental in a research paper. We believe that this study has several novelties. First, while previous complex network studies on precipitation only treated spatial connectivity, we account for both spatial and temporal factors. It made better results in the East Asia relationship, which reflect weather cycle characteristics. Second, we applied new methods (i.e., vital node identification and multiresolution community detection). Those help to analysis subjects in a less complicated and faster manner. Finally, the research framework proposed in this study helps study spatial and temporal connections in large scale regions. From complex network to cross mutual information, the framework contains topological analysis, statistical analysis and also consider temporal factor. This reveals climate connections in regions and reflects weather cycle characteristics.

      - (Introduction) We assessed the effects of each region through centrality analysis and grouped the regions according to clustering analysis. Subsequently, mutual information (MI) was calculated with a time lag (i.e., cross-mutual information) to identify the relationships between each group. In previous complex network studies, they have only considered spatial factors for precipitation analysis. It is available to obtain a good result in a small area, but there is a limit to its applicability in a large area. Because there is a significant time difference between two locations in a large area, therefore, this study proposes an efficient methodology that can evaluate not only spatial but also temporal factors. It also provides a clue to finding the trigger locations of the climate cycle in an area

      - (Discussion) The complex network facilitated a simple analysis of the relationship between East Asian cities. Unlike previous studies, we considered temporal factors in the relationship. Through this, we observed new relationships and characteristic of rainfall in East Asia. Two methods (vital node identification and multiresolution community detection) were very useful for analyzing the network and making reliable results. The research results show that our research framework is helpful for studying relationships in regions. The frame contains not only topological analysis but also statistical analysis and considers temporal factors. Also, in the result, the frame reflects climate cycle factors and reveals its characteristics.

2) I do not understand the significance of the representatives' selection from groups. Why do the authors select these? Are the nodes used in the following analyses? Why are they important?

- We decided to delete the significance of the representatives' selection from groups. Initially, we used the selection of representatives in the groups for future research in this paper. We plan the future study, which will apply complex network analysis to the whole world. In this case, there will be many nodes and links, making it hard to analysis and interpret a network. Therefore, it will be beneficial to use representatives of the groups, instead of all nodes, because representatives are the most influential nodes in the groups and have characteristics of the groups. We needed to check the validity of the method and applied it in this study. The results show that it has a similar result with vital node identification. As a result, we thought that the selection of representatives would be helpful in a future study. However, we did not use the representatives after selecting them because the discussion is minor.

3) Line 178: I am not sure about the reason why Seoul does not belong to any cluster. The authors explain that the distance is great from other nodes, but the distance between Seoul and Shenyang is 560 km and it is less than the distance from this latter city and Beijing (about 630) and they belong to the same cluster. Can the authors better justify the result?

- Your comment was constructive in insight into our research again. After we got your comments, we analysed why Seoul did not make a group with other nodes. We applied event synchronisation, which helps to compare the occurrence pattern of precipitation. Event synchronisation results with Seoul and near cities (Beijing, Tianjin, Shenyang, Osaka, Nagoya, Tokyo) had low value than average event synchronisation reslults of all cities. Through this, it was confirmed numerically that Seoul has a different precipitation pattern from the surrounding cities. The reason for the event synchronisation result is that Seoul is located on the peninsula. The Korean peninsula is influenced by maritime air mass in summer and by continental air mass in winter. Therefore, characteristics of precipitation in the Korean peninsula are affected by continents and oceans' features and has differences from those. This makes Seoul had a low belonging coefficient with the G2 group (affected by the ocean) and G4 groups (affected by continents). We deled the distance reason and added this abovementioned new reason to the paper.

- Because of the location of Seoul, it had low belonging coefficients with nearby nodes. Seoul is in the Korean peninsula. It is influenced by maritime air mass in summer and by continental air mass in winter. Therefore, the precipitation of Seoul is affected by both features and has different characteristics. This feature made Seoul distinguish between G2 and G4.

4) The references should follow the journal's guidelines: the first name initials after the last name. Please, check the references: the authors have often exchanged last and first names.

➢ We rechecked the guidelines of HESS. Then, we checked all references and revised errors in references.

5) Tables and maps are redundant. It would be better if the authors summarise the information in a single figure, rather than duplicate the results in a table and a map. For example, table 3 and figure 4 can be summarised in a map in which a continuous colour scale could represent the adjacency information entropy values and different sizes of points could represent the rank. Even table 4 - figure 5 and table 6 - figure 6 are redundant.

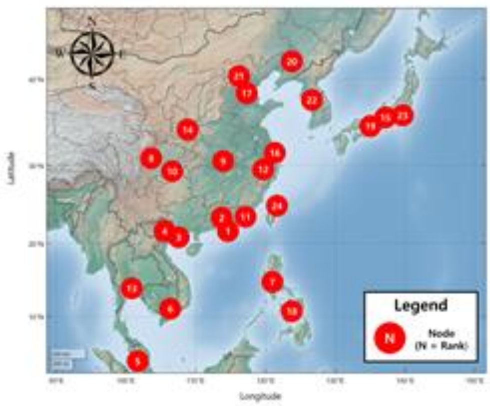
➢ We revised the table and figures accordingly as below.

➢ About Table 3 and Figure 4, we made them as one figure like below.

<Before>

Table 3. Rank of nodes according to adjacency information entropy

Rank	Node	Adjacency information entropy
1	Hong Kong SAR	5.5022
2	Pearl River Delta	5.3833
3	Hai Phong	5.2789
4	Ha Noi	5.2701
5	Kuala Lumpur	4.9206
6	Ho Chi Minh	4.9182
7	Manila	4.7668
8	Chengdu	4.7560
9	Wuhan	4.7441
10	Chongqing	4.7092
11	Shantou	4.6543
12	Hangzhou	4.5983
13	Bangkok	4.5198
14	Xi'an	4.5117
15	Nagoya	4.3093
16	Shanghai	4.3023
17	Tianjin	4.2954
18	Cebu	4.2148
19	Osaka	4.1917
20	Shenyang	4.1507
21	Beijing	4.0442
22	Seoul	3.9890
23	Tokyo	3.9742
24	Taipei City	3.9049



<After>

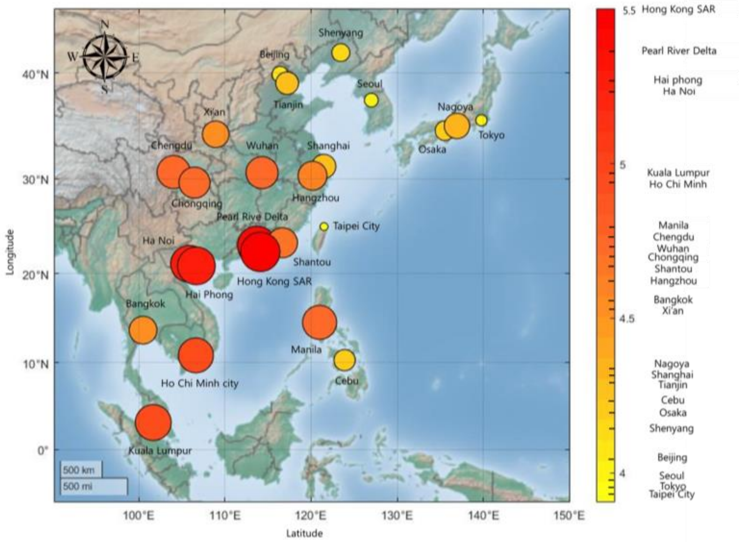


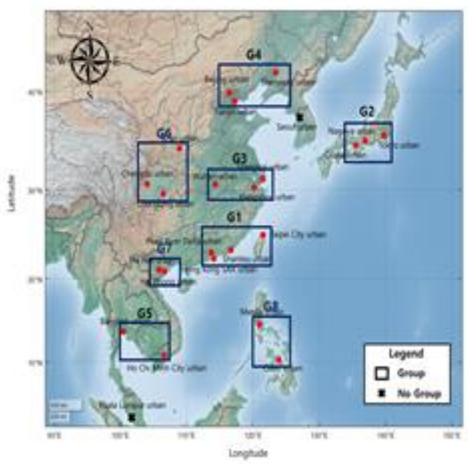
Figure 4. Adjacency information entropy value of cities: Color and size of circle are respectively proportion to the entropy and rank; Ride side of bar shows the adjacency information entropy values of nodes. Except for Taipei city, nodes near South China Sea had higher values.

- About the Table 4 and Figure 5, we deleted Table 4 and put only Figure 5. Readers can find nodes in groups from the figure.

<Before>

Table 1. Groups of nodes and their components

Group	Components
G1	Pearl River Delta, Hong Kong SAR, Shantou, Taipei City
G2	Osaka, Tokyo, Nagoya
G3	Wuhan, Shanghai, Hangzhou
G4	Shenyang, Beijing, Tianjin
G5	Bangkok, Ho Chi Minh City
G6	Chengdu, Xi'an, Chongqing
G7	Ha Noi, Hai Phong
G8	Manila, Cebu
Other	Seoul, Kuala Lumpur



<After>

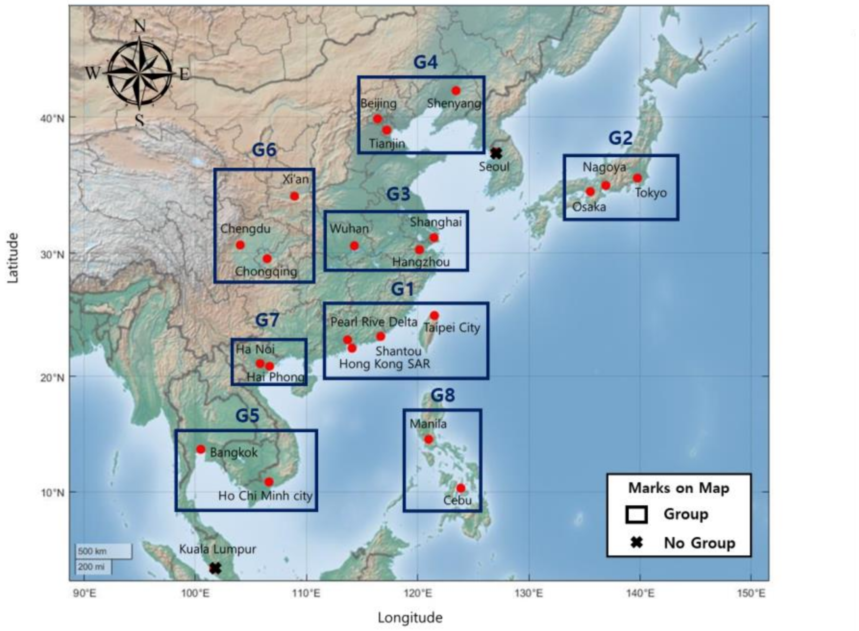


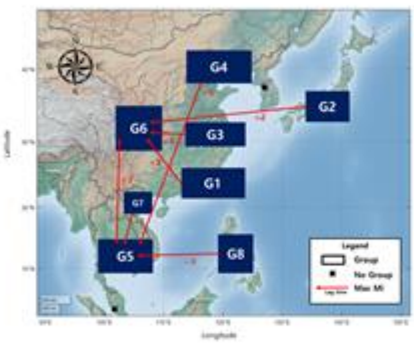
Figure 5. Group of nodes using multiresolution community detection; There are 8 groups in the East Asia; G1(Pearl River Delta, Hong Kong SAR, Shantou, Taipei City), G2(Osaka, Nagoya, Tokyo), G3(Wuhan, Hangzhou, Shanghai), G4(Tianjin, Shenyang, Beijing), G5(Bangkok, Ho Chi Minh City), G6(Xi'an, Chengdu, Chongqing), G7(Hanoi, Haiphong), G8(Manila, Cebu); Seoul and Kuala Lumpur did not make group with other nodes.

- About Table 5 and Figure 6, we erase Table 5 and only put Figure 6. In this part, the most important thing is that what group has a strong relationship with each group. Figure 6 is the suitable than Table 5 for showing this.

<Before>

Table 1. Maximum cross-mutual information and lag time

	G1	G2	G3	G4	G5	G6	G7	G8
G1	-	0.312 (−3)	0.333 (2)	0.312 (2)	0.438 (−9)	0.472 (2)	0.430 (−1)	0.388 (3)
G2	0.312 (3)	-	0.343 (1)	0.273 (2)	0.374 (−8)	0.398 (2)	0.323 (−9)	0.333 (−10)
G3	0.333 (−2)	0.343 (−1)	-	0.271 (0)	0.369 (−9)	0.496 (1)	0.355 (−3)	0.332 (−6)
G4	0.312 (−2)	0.273 (−2)	0.271 (0)	-	0.409 (−5)	0.403 (1)	0.361 (−1)	0.344 (−10)
G5	0.438 (9)	0.374 (8)	0.369 (9)	0.409 (5)	-	0.607 (3)	0.551 (1)	0.564 (3)
G6	0.472 (−2)	0.398 (−2)	0.496 (−1)	0.403 (−1)	0.607 (−3)	-	0.535 (−2)	0.486 (4)
G7	0.430 (1)	0.323 (9)	0.355 (3)	0.361 (1)	0.551 (−1)	0.535 (−2)	-	0.432 (−5)
G8	0.338 (−3)	0.333 (10)	0.332 (6)	0.344 (10)	0.564 (−3)	0.486 (6)	0.432 (5)	-



<After>

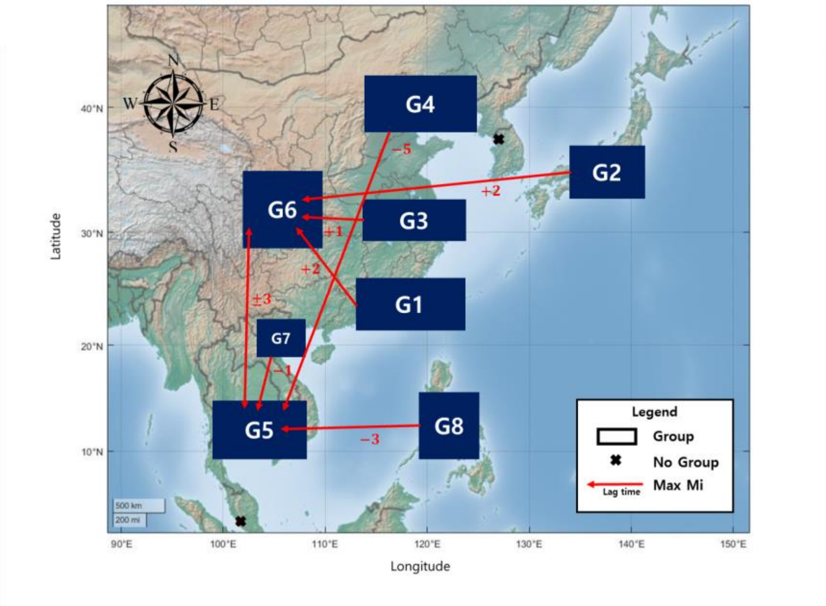


Figure 6. The maximum cross mutual information relationship and its time lag value based on Table 6; Each arrow points out the maximum relationship group and the numbers under the arrows express the lag time(days) of the maximum cross mutual information value; The figure shows relationship of groups and influence time intervals in East Asia;

6) please expand your figure caption to tell us what the figure is trying to show the reader, before searching in the text. This comment applies to all figures and some of the more complex tables

- Thank you for your good comment. We totally agree with your comment. Therefore, we added more detail explanation about Figure and Table in Captions. We hope this will help readers understand Figure and Table.

  - Table 2. Basic statistics values for rainfall data of cities
    ⇒ Table 1. Basic statistics values for rainfall data of cities; Basic statistics contain average, standard deviation, coefficient of variation and skewness;
  - Table 2. Average, maximum, and minimum link weights of each node
    ⇒ Table 2. Average, maximum, and minimum link weights of each node; Parentheses under link weights is nodes that forms a maximum or minimum value for target node;
  - Figure 4. Ranks of nodes using vital node identification
    ⇒ Figure 4. Adjacency information entropy value of cities; Color and size of circle are respectively proportion to the entropy and rank; Ride side of bar shows the

adjacency information entropy values of nodes; Except for Taipei city, nodes near south China sea had higher values;

- Figure 5. Group of nodes using multiresolution community detection

  ⇒ Figure 5. Group of nodes using multiresolution community detection; There are 8 groups in the East Asia; G1(Pearl River Delta, Hong Kong SAR, Shantou, Taipei City), G2(Osaka, Nagoya, Tokyo), G3(Wuhan, Hangzhou, Shanghai), G4(Tianjin, Shenyang, Beijing), G5(Bangkok, Ho Chi Minh City), G6(Xi'an, Chengdu, Chongqing), G7(Hanoi, Haiphong), G8(Manila, Cebu); Seoul and Kuala Lumpur did not make group with other nodes.

- Figure 7. Major water vapor transport routes in East Asia; The routes could explain the reasons why the relationship of groups was made like Figure 6

  ⇒ Figure 7. Major water vapor transport routes in East Asia; The routes could explain the reasons why the relationship of groups was made like Figure 6; Indian monsoon brings vapor from Indian Ocean, East Asian monsoon gets vapor from Pacific Ocean and East China Sea; Anomalous anticyclone provide vapor in East China, Korea and Japan;

7) this last figure '7' is in microns = $1m^{-6}$. This is spurious - try 3 significant figures?

➢ In Figure 7., Max mi means Maximum cross-mutual information value. We want to express relationship between groups based on the maximum cross-mutual information result. However, Max mi in Legend could make confsuion on readers. Therefore, we fixed Legend in the Figure 7. And wrote more explaination in Caption.

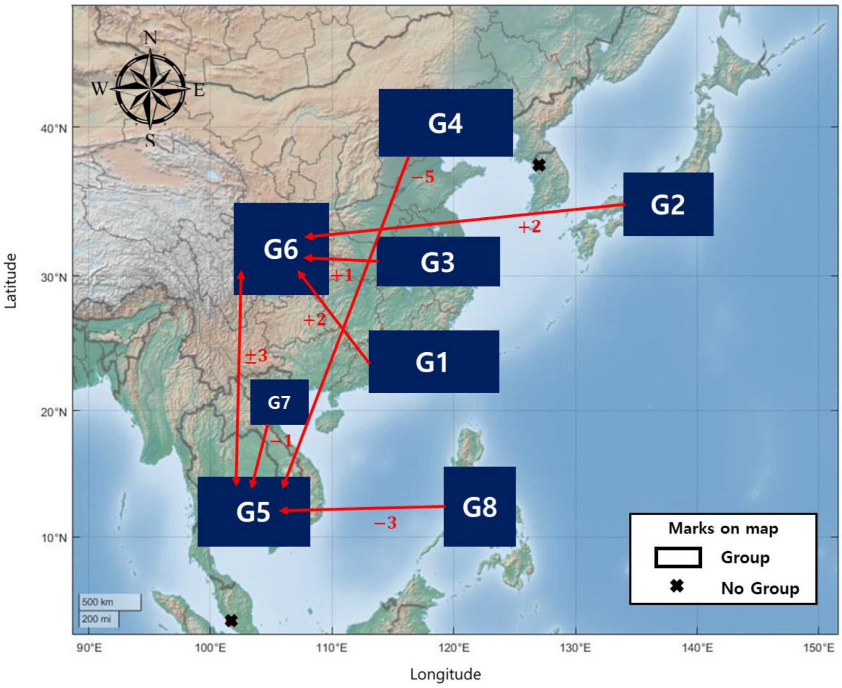


Figure 6. The maximum cross mutual information relationship and its time lag value; Each arrow points out the maximum relationship group and the numbers under the arrows express the lag time(days) of the maximum cross mutual information value; The figure shows relationship of groups and influence time intervals in East Asia;

8) is this k_j an inverse-distance weight, like kriging?

- k_j means 'Degree' which is basic index in complex network analysis. The degree is the number of nodes which have links with target node in unweighted network. About the weighted network, it is sum of weight of links that have connection with target node. Reader need to know about degree for understand vital node identification method. Therefore, we put explanation about degree in equation (2).

- First. Calculate degree(k_j) of each node in the network

$$k_i = \sum_{j\in\Gamma_j} w_{ji} \tag{2}$$

Here, Γ_j is a group of nodes that form links with node j. w_{ji} is weight of link that connect node j and node i. If a network is unweighted, degree is the number of neighbor nodes.

Second, calculate the adjacency degree (A_i) of each node.

$$A_i = \sum_{j\in\Gamma_j} k_j \tag{3}$$

Third, calculate the selection probability (P_{i_j}).

$$P_{i_j} = k_i / A_j \tag{4}$$

Final, calculate the adjacency information entropy (E_i).

$$E_i = \sum_{j \in \Gamma_j} (P_{i_j} log_2 P_{i_j}) \tag{5}$$

After comparing the calculated adjacency information entropy of each node, the importance is determined according to the descending power.

9) Comments on References

➢ We checked guidelines of HESS again. Then, we checked all references and revised errors in all references.

- East Asia accounts for 54% of the global supply chain, providing a wide range of services and products across the world (Ann et al, 2020).
- Wang, Z., Mu, J., Yang, M., and Yu, X.: Reexaming the mechanisms of East Asian summer monsoon changes in response to non-East-Asian anthropogenic aerosol forcing, Journal of Climate, 33(8), 2929-2944, https://doi.org/10.1175/JCLI-D-19-0550.1, 2020.

<Minor comments>

About grammar and sentences, we checked the whole paper by own self and got help from a native speaker.

1) Lines 63-64: "This is because the weights used as input data in each analysis enabled the relationship between regions to reflect in the network and be analysed.";

➢ The sentence was re-written as below

- Because the weights were used as input data in each analysis, network analysis could reflect relationship between regions.

2) Lines 103-104: "Generally, actual systems such as transportation systems or the Internet do not require links to be defined. However, if uncertainty occurs in the connection, researchers must define them.";

➢ The sentence was re-written as below

- Generally, it is easy to define links in systems such as transportation or power grid systems, which have clear physical connections between elements. However, if uncertainty occurs in the connections like social networks, researchers must define them.

3) Lines 154-155: “Various cities had maximum weights for each node, whereas the minimum weights were restricted to a few cities.”.

➢ The sentence was re-written as below

- Each node had a maximum value with several different cities, while the minimum value was for certain cities such as Beijing and Tokyo.

4) Line 43: the first mathematician who formulated complex network theory was Leonhard Euler, in 1735.

➢ The sentence was re-written as below

- Complex network theory, developed by Leonard Eüler in 1735, expresses and analyses a subject or phenomenon as a graph.

5) Figure 1 and Figure 5: delate “urban” from the labels.

➢ We deleted urban in Figure 1 and Figure 5 (Figure 5 is in major comments 5).

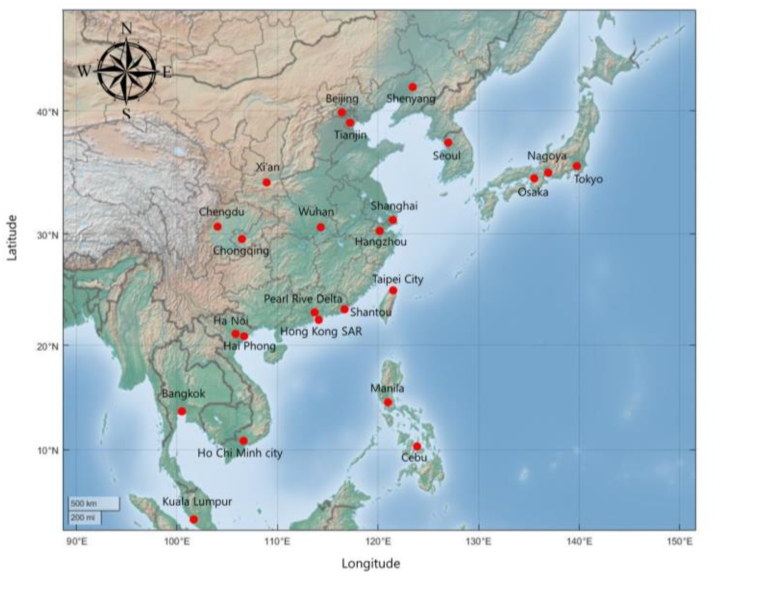


Figure 1. Selected 24 major cities in East Asia

6) Section 3.3: the symbols in the formula are not well defined. The authors should better clarify the meaning of the symbols used. For example, what do v_i, v_j, c_u, and nj mean?

➢ We added explain about symbols like v_i, v_j and c_u.

- First, calculate the link intensity (I_P) of each link.

$$I_P(e_{ij}) = \begin{cases} \sum_{p=1}^{P} \alpha_p \times \dfrac{\sigma(path_p(v_i, v_j))}{\min(w_i, w_j)}, & e_{ij} \in E \\ 0, & otherwise \end{cases} \quad (5)$$

Here, $\sigma(path_p(v_i, v_j))$ is the sum of link weights in the path through p links from node i (v_i) to node j (v_j), P is the parameter of the path, and α_p is a polygonal effect parameter. For edge e_{ij} between node i and node j, w_i and w_j are their respective strengths.

Second, identify the links with link intensity greater than the selected threshold and create a group of nodes with the identified links.

$$\begin{aligned} v_j \in V, \quad & I_P(e_{ij}) > t \\ v_j \in c_u, \quad & I_P(e_{ij}) > t \end{aligned} \tag{6}$$

Here, t $(0 < t \leq 1)$ is the selected threshold, and c_u is a group of nodes.

Third, calculate the belonging coefficient (I_P) of the nodes in node set u (c_u).

$$I_P(c_u, v_j) = \sum_{v_i \in c_u} I_P(e_{ij}) \tag{7}$$