We thank all the reviewers and editors for their feedback and suggestions. In the following we will address the each questions and comment from the viewers.

1 Reviewer 1:

1. First of all, clustering technique is not new and it is one of the well-researched areas in ad hoc network, cognitive radio, HetNet etc. The article does not review the latest publications in this area. The references are not adequate due to the fact that no recent works have been investigated. Therefore, it makes it insufficient (or unsatisfactory) work in present context.

   Answer: As mentioned in the introduction section, there has been a lot of works on clustering in ad hoc networks or sensor networks. They address problems such ~~like~~ as routing, power consumption and communication quality. Compared with the wireless networks where spectrum availability is not opportunistic, the clustering in CRN facilitated the better usage of the available spectrum, meanwhile clustering itself is restricted by the spectrum availability. This fact makes clustering in CRN a difficult topic. Up to now, only [1–3] have tried to address the robustness issue of the formed clusters in CRN. After searching, we find one ~~relavent~~ additional relevant paper [4] from 2015 IEEE Crowncom. It proposes a clustering algorithm which aims to speed up the process of re-clustering in case a cluster doesn't uphold in from of active primary users. But this work doesn't consider to improve the ability of the formed clusters to ~~maintain~~be maintained, besides, the messaging cost which is involved in this process is not considered. We have included this citation in the related work section in the revised version. We will appreciate if the reviewer could provide more specific references that we have not mentioned yet.

2. The presentation of the paper is not good enough and needs a lot of improvements. For instance, ROSS scheme has been mentioned in the paper in many places, but it is not properly defined (though reference has been given). I had to check the reference paper to know it is Robust Spectrum Sharing. The journal article must be self-contained.

   Answer: ~~The explanation of the scheme name is provided in the introduction section. The paper has been thoroughly proofread.~~ In the revised version, we explain the name ROSS in the fifth paragraph when we introduce our scheme.

3. The paragraphs are too long in the Introduction. As a result, the paragraph contains too many information which are not related each other. This makes readers difficult to understand the concept.

   Answer: We ~~polish the paper and now the section of~~ polished the paper according to the reviewers request. Now the introduction consists of ~~4~~5 paragraphs: introduction of CRN with mentioning of spectrum sensing → clustering in CRN → robustness of clustering → the role of cluster size in robust clustering in CRN → introduction of our work. The part of related works becomes one separate section.

4. In page 3, it is mentioned that "A cluster can be formed only by the cluster head". As a matter of fact, the cluster head formation is itself a very challenging task in any cluster based approach. There are many previous works which describe the cluster head selection techniques. Please justify your statement with proper assumption or previous works that how the cluster heads are selected. I can see the description is in the few pages later but some information must be provided where it is described first time. By the way, the CH formation technique described in 4.1.1 is not novel.

   Answer: We agree with the reviewer that this statement is not well explained. At first, we put this line in the first submission and wanted to say that after the clusters are decided, it is the cluster head's job to form a cluster out of the neighboring nodes. In the revised version, we have changed the statement as "Our scheme enables that the cluster heads are selected in a distributed manner, and a cluster can only be formed only by these selected cluster head. "

5. In my understanding, the values of connectivity degree in Fig. 1 came from the assumption of K_A, K_B $\ldots..K_H$. Each node has number of available channels and I assume that they have been randomly chosen as mentioned in Fig. 1 Caption. It would be good to mention how to obtain such information in ad hoc cognitive radio network?

   Answer: The available channels on the nodes in Fig. 1 is randomly chosen, which is only used to explain our scheme. But in the simulation, the available channels on the secondary users are decided by the primary users. As stated in the system model, both primary and second users work on the same set of channels, i.e., the former use the channel whenever it is needed, and the later can only use it opportunistically. ~~We assume the primary users are randomly located and work on each channel with a certain probability. With spectrum sensing, the secondary users can detect the presence of active primary users, and thus obtain the available channels.~~

   In Section 5.1 in the revised version, we describe the generation of connectivity vectors, and reformulate the paragraph as follows,

   "Individual connectivity degree d_i and neighborhood connectivity degree g_i together form the *connectivity vector*. Connectivity vector is calculated by every secondary user after channel availability is obtained, and is then broadcasted. Figure 2 illustrates an CRN where every node's channel availability and connectivity vector is shown. This CRN is used as an example in the rest of this section to illustrate the proposed distributed schemes."

6. Theorem 4.1 is too obvious, authors need to discuss in terms of convergence and complexity of the proposed method. In fact, Theorem 4.1 and description immediately are so confusing and misleading. If step is defined as executing Algorithm 4.1 for one time which again takes at most N steps. A further clarification is needed.

   Answer: We agree with the reviewer that 'step' is not fully explained. We have represented this theorem thoroughly in Section 5.1.1 in the revised

version~~, which will be made after three lemmas~~. Besides, we use 'time' in stead of 'step' to quantify the process of the phase of cluster head selection and initial cluster formation.

In the phase I of ROSS, the biggest concern is when will this process end. Considering the fact that the secondary users keep receiving the latest information from neighbors, it is not easy to say whether a secondary user will change its role between cluster head and remember for multiple times or even infinitely. Thus we represented the theorem in the revised version, which will be made after three lemmas.

Lemma 5.1 shows all the secondary users will be involved in the process, and decide on their roles decisively in the end.

Lemma 5.2 looks obvious or even trivial but it is critical to decide the time needed for phase I of ROSS to complete. According to our regulation on the *Individual connectivity degree* of a secondary user (it becomes a large positive integer M after being included in a cluster, and being the same when it is included in a second cluster, besides it never decreases), the cluster head will remain to be cluster head once it becomes a cluster head. This excludes the possibility for a secondary user to frequently or even infinitely switch between the two different roles in clusters.

Lemma 3 is the natural outcome from Lemma 5.2.

Theorem 5.1 is the takeaway which gives the decisive maximum time to complete this process.

7. In 4.1.3, parameter t depends on density of the network. Please clarify what is the impact of selecting value of t in system design, such as selecting t to be higher than 1.3.

   Answer: The impact of t in the system design is elaborated in Section 5.1.3 in the revised version. The added text are as follows,

   "t is between 1 and the ratio between the average size of ~~the every~~ each secondary user's neighborhood and the desired size. When t is smaller, for example ~~,~~ $t = 1$, the formed cluster in the phase I will be δ. ~~Then when the nodes which belong to multiple clustersdecide on their clusters, they will be excluded from the other cluster whose size~~ For the cluster whose members are included by other clusters, the size of this cluster will be smaller than δ ~~. A large~~ after the membership clarification phase. If t is chosen large, for example, ~~the t which makes~~ $t \cdot \delta$ equals to the size of neighborhood, then the mechanism of adjusting the cluster size will not work any more."

8. In page 7, what does "the toy network in Figure 2" mean? I hope it's a typo. Otherwise, the Sections 4.2.1, 4.2.2, 4.2.3 are nicely written.

   Answer: It is a typo and we have changed the sentence ~~to~~ as "Now we apply both ROSS-DGA and ROSS-DFA to the network in Figure 2 ~~which has been applied the~~ after phase I of ROSS is applied."

9. In simulation, have authors considered the spectrum sensing parameters on each secondary nodes? Are the channels randomly made available to users? Please clarify. What value of A is chosen in simulation?

Answer:

The way the available channels are decided complies with the our system model in Section 3. We would like to answer the question here again.

In the most part of the simulation, we assume the secondary users have ~~accurate spectrum sensing ability~~ the ability to sense the spectrum accurately. They are able to detect the active primary users in their vicinity. We can say the channel availability on the secondary users is decided by the primary users. As ~~to~~ for a secondary user, if ~~none of the primary users whose transmission rangeinclude the secondary user works on a~~ it is not in any primary user's transmission range, which are working a certain channel, then the channel ~~, that channel~~ will be available for that secondary user. ~~We also consider~~ Based on a comment of reviewer 3, we also considered the existence of sensing error ~~,~~ in Section 6.2.4 ~~in~~ of the new version, where we investigate the influence caused by false negative sensing error.

As mentioned in Section 6.1, 'A' represents the length of one side of the simulation square. ~~And~~ Note that the radius of transmission of both primary and secondary users is a fraction of A. By doing this, we ~~separate our scheme with concrete physical layer communication technologies~~try to abstract from the influence of any given physical layer technology, and A can be given a concrete value in practice when the physical layer technology is decided. We have added the motivation to use 'A' in Section 6.1 in the revised version.

10. It is mentioned in the article that "The number of licensed channels is 10, each PU is operating on each channel with probability of 50%". It means there will be only 5 channels available (maximum) to each secondary users. In section 5.1, 7 channels are available for each CR node. It means there is chance of interference in two nodes. Please clarify those conflicts.

    Answer: The simulation consists of two parts, the comparison between the centralized and distributed schemes, and the comparison between the distributed schemes. Some parameters are identical for both parts (stated in the second paragraph in Section 6), but there are some parameters which are different (as stated in the beginning of Section 6.1 and 6.2 respectively). The difference lies in the number of ~~the CRN and the~~ secondary and primary users, and the transmission ranges in CRN. The reason is because there doesn't exist an efficient method to solve the centralized scheme, thus we have to do simulation in a smaller network.

    The parameter setting questioned by the reviewer is for the simulation involving centralized schemes in Section 6.1. Please note that the transmission radius of primary user is A/3, which means some nodes are not affected by all the primary users. Other than that, from the simulation we know ~~averagely~~ that on average 7.1 channels are available for each secondary user.

11. Further comparison is needed against the previous works to justify the robustness of the proposed method. **Answer:** We add the definition of robustness in both the introduction section (3rd paragraph) and system model section (Section 3.1). The robustness is illustrated by the number of the unclustered CR nodes in the CRN, which is investigated both in

Section 6.1 and 6.2. We have modified the analysis of the simulation result in Section 6.1.3 and 6.2.2. We have also added additional simulations about the effects of sensing errors on the distributed schemes in Section 6.2.4.

2 Reviewer 2:

Comments to the Author In this paper, authors investigate the optimization algorithms for node clustering strategy in ad hoc based cognitive radio networks, where both the centralized solution and distributed solution are considered.

1. In Section 2, that the definition of channel $k \in K$, i.e. fading channel or noisy channel? Does the channel contain pathloss and shadowing and/or fast fading? Does the value of k indicate the channel gain in either time or frequency domain?

   Answer: As shown in the section of system model, "We adopt the unit disk model for both primary and CR users' transmission." In other words, we don't make any assumption on the concrete propagation model, but just assume ~~the communication is successful when the two ends locate within each other's transmission range~~that nodes are generally in communication range if their physical distance is lower than some threshold value. One can argue that such a model resembles the impact of path loss and the average impact from shadowing and fading. We adopt this simple communication model ~~so as~~ in order to focus on secondary users' behavior to pursue the robust cluster structure.

2. Does the channel invariant or variant (how fast) during the clustering for how long windows range? In practical, the channel changes along with the transmission, the performance clustering may be impacted by the varying channel. What's the impact of insufficient channel knowledge?

   Answer: We assume the channel availability doesn't change during the clustering process. We agree with the reviewer that the channel variance affects the clustering by disrupting the communication. In reality, if the communication during the clustering process doesn't succeed ~~as~~ while the channel quality deteriorates, the involved nodes can still make decisions based on the available knowledge of neighbors ~~according to our algorithm.~~ which is already obtained. In this case, many smaller clusters or even singleton clusters may be generated. In the other hand, the chosen model would fit a slowly varying scenario, where for instance most primary and secondary users are stationary. Thus, single messages might be lost, but statistically the proposed schemes could still determine a stable clustering (by retransmissions etc.).

3. The work emphasize on the robustness of clustering. However, such an evaluation in terms of robustness metric is not clear. The definition of robustness and relative discussions as well as the results should be further elaborated. The definition appears in Section 5 but my suggestion is to clearly state it in the system model and describe why choose this robustness metric according its relation with system design.

Answer: We ~~take~~ took the advice and have done the modifications in two places. ~~1.~~

(a) We mention the definition of robustness of clusters in the 3rd paragraph of the introduction section. ~~2.~~

(b) In Section 3.1, the robustness is associated with the number of secondary users which are not included in any cluster. ~~with~~ With these modifications, we hope that the readers will understand the robustness metric in Section 6 (Section 5 in the previous version) easily.

4. In Section 4, algorithm flow charts should be included.

    Answer: According to the comments of the reviewer, we added a flow chart as Figure 1 ~~is added~~ in Section 5 (which is Section 4 before this major revision)

5. In Section 5, Fig. 10 is missing. Fig. 14-16 can be subfigures, similar to Fig. 17-19.

    Answer: ~~Advice is taken and change is made~~ We took the advice and made changes accordingly. Please refer Fig 8 - 9.

3   Reviewer 3:

Comments to the Author In this paper, the authors consider the robust clustering problem in the ad-hoc Cognitive Radio Networks (CRNs). The paper is well formalized. The research background and motivation are well described. The research problem is well addressed. The simulation based experiment is conducted to show the performance of the suggested solution approach to the addressed problem.

Overall, the considered topic is very interesting. The authors well presented the considerations and assumptions on several key research issues, which inherently exist in the ad-hoc networks, for instance, decentralized spectrum sensing, synchronization on the sensing results and congestion control. To solve the addressed problem, the authors jointly use different research methods such as graphic theory and game theory, which have been widely reported in recent studies on CRNs. The experiment is well conducted with the thorough performance analysis and comparison on different experimental scenarios.

1. However, it seems that the authors totally miss a very important research issue, which is referred to as spectrum sensing errors in terms of overlook and misidentification on the channel availability by Secondary Users (SUs). Because the sensing errors may highly affect the effectiveness and performance of the suggested solution approach, the authors need to complement the paper by accordingly either advancing the feasible solutions to alleviate sensing errors or analyzing the effect of sensing errors on the suggested solution approach.

    Answer: It is true that the sensing error was not considered in the first submission. We agree with the reviewer that this is an important issue, as sensing ~~error is~~ errors are inevitable in practical applications. We have done additional simulations about the effects of sensing errors on the distributed schemes in Section 6.2.4. With the false negative rate increases

from 0 to 50%, the percentage of nodes in singleton clusters generated by the variants of ROSS is increased from 0 to 10%, as comparison, the percentage by SOC increases from 9% to 25%. Besides, the variance of sizes of the formed clusters by the variants of ROSS is much smaller than that by SOC.

References

[1] D. Li and J. Gross. Robust clustering of ad-hoc cognitive radio networks under opportunistic spectrum access. In *Proc. of IEEE ICC '11*.

[2] Sisi Liu, Loukas Lazos, and Marwan Krunz. Cluster-based control channel allocation in opportunistic cognitive radio networks. *IEEE Trans. Mob. Comput.*, 11(10):1436–1449, 2012.

[3] Nafees Mansoor, A. Islam, Mahdi Zareei, Sabariah Baharun, and Shozo Komaki. Construction of a robust clustering algorithm for cognitive radio ad-hoc network. In *Proc. of CROWNCOM 2015*.

[4] Nafees Mansoor, A. K. M. Muzahidul Islam, Mahdi Zareei, Sabariah Baharun, and Shozo Komaki. Construction of a robust clustering algorithm for cognitive radio ad-hoc network. In *Proc. of CROWNCOM 2015*.