**Revision Report of Submission TOSN-2016-0008**

We thank the comments with cares and insights made by the reviewers, which are helpful for improving the quality and readability of our paper. In our revised paper, we have made detailed explanations and changes in response to all the reviewers’ comments. Here, we explain our revisions based on each of the comments and suggestions.

**Response to Reviewer 1’s Comments**

1. **Comment:** *For better understanding of this paper, please explain the processes of allocation and scheduling are decoupled and optimal separately. Is it equal to the joint optimization of allocation and scheduling?*

**Response:** Thanks for your careful reading. We have reorganized **Section 1** and presented the allocation of tasks and scheduling of sampling intervals in the fifth paragraph. Besides, we have added the definitions of task allocation and scheduling of sampling intervals in **Section 3.1**.

When an application produces multiple tasks, they should be allocated to and performed by nodes. Consider the potential benefits of data sharing amongst overlapping tasks, the strategy of task allocation and the scheduling method of sampling intervals should be carefully designed, so as to maximize data sharing. Specifically, the optimal strategy of task allocation should make sure that those overlapping tasks with maximal amount of data sharing should be allocated to a same node. The optimal scheduling method of sampling intervals should fully exploit the potential benefits of allocated tasks on the same node.

However, optimizing the two challenging issues separately is not equivalent to the joint optimization, and may loss some extent of gain of data sharing for a wireless sensor network. Since the strategy of task allocation has a great impact on the performance of the scheduling method of sampling intervals, it is reasonable to consider the joint optimization problem and find a near optimal solution.

1. **Comment:** *"overlapping tasks" is an important term in this paper, so it is better to give a clear definition of "overlapping task windows" and "overlapping tasks".*

**Response:** Thanks for your careful reading. Inspired by the suggestion, we have added the definitions of the *“Overlap”* in **Section 3.1***.* If two sampling tasks are overlapping, their time windows share a common time region. Their sampling intervals are thus moved to the common time region for achieving the gain of data sharing.

1. **Comment:** *In def.3.4, it is better adding v1>u2 when define overlapping intervals.*

**Response:** Thanks for your careful reading. We have revised the definition of *“Overlap”* in **Definition 3.3**, and adopted this suggestion by adding the explanation in the following paragraph.

1. **Comment:** *Eq.(2) seems not so complete. Ii is changeable with j since the intervals of task i on different node j maybe different because of the interval movement.*

**Response:** Thanks for your careful reading. Inspired by your suggestion, we have revised the definition of the joint optimization problem. The notation *Ii* in Equation (2) in the previous version has been replaced by *Iij* in the Equation (4) in the new submission.

1. **Comment:** *Eq. (6)-(8) might be more simple if introducing the variable of overlapping window(e-b).*

**Response:** Thanks for your careful reading. We have tried to introduce the variable of overlapping window into such equations. However, we found that this will not bring more benefits to improve the readability of this paper. According to your suggestion, we have revised Equations (7)-(9) in the new version carefully to make them be understandable and simple.

1. **Comment:** *give the line numbers in all the algorithms and be consistent with the explanation in the text.*

**Response:** Thanks for your careful reading. Inspired by the suggestion, the line numbers have been added in all algorithms in the new version.

**Response to Reviewer 2’s Comments**

1. **Comment:** *The authors should further improve the presentation about the major difference between the existing point sampling tasks and the interval data sampling tasks addressed in this paper. It will be better if such an issue can be discussed in the problem statement and formulation sections besides the introduction section.*

**Response:** Thanks for your careful reading. We have reorganized **Section 1** and added the third paragraph to discuss the difference between the point sampling tasks and the interval data sampling tasks. The point sampling tasks are performed at a certain time; while the interval data sampling tasks should be performed for a time interval. Therefore, the energy consumption of the interval data sampling tasks is much more than that of the point sampling tasks.

We have revised the definition of “Interval sampling task” in **Section 3.1** and the last paragraph in **Section 3.1** to discuss the difference between the point sampling tasks and the interval data sampling tasks. The task model of interval data sampling tasks contains three elements: the begin time, the end time and the sampling interval. The point sampling tasks are the special case of the model because they can be performed by sampling data only once; while the interval sampling tasks have to be performed by sampling data for a time interval. Furthermore, the data sharing of overlapping interval sampling tasks is difficult to compute due to the variety of the length of sampling intervals.

1. **Comment:** *Another major contribution of this paper is to minimize the interval data sampling from the aspect of the whole WSN instead of a single sensor node. The authors should emphasize such contribution so as to ease the presentation of this paper.*

**Response:** Thanks for your careful reading. Inspired by the suggestion, we have revised the fifth and sixth paragraphs in **Section 1** to highlight the contribution. Consider that multiple tasks will be produced by an application across all sensor nodes in a WSN. Since data sharing exists amongst overlapping tasks, those tasks with the maximal gain of data sharing should be allocated to a same. However, existing work about the interval sampling tasks just focuses on the task scheduling problem on a single node. Although the involved scheduling method is effective, it does not work well in WSNs due to lack of task allocation procedure. The main reason is that the performance of the scheduling method is dominated by the strategy of task allocation at the scope of entire network. We consider and formulate the two problems as a whole, and provide a general and practical solution for a deployed system.

1. **Comment:** *It seems that “task scheduling” means to schedule the sampling interval in a time window for a sampling task. It is better to use “scheduling of sampling interval” instead of “task scheduling”.*

**Response:** Thanks for your careful reading. Inspired by the suggestion, “task scheduling” in the previous version has been replaced by “scheduling of sampling interval” in the new submission.

1. **Comment:** *It is not very clear why the strategy of task allocation has a great impact on the data sharing during task scheduling? It is necessary to formally define the scheduling and allocation problems in the new version. The definition of MAX-DHP problem should also be given before its first use.*

**Response:** Thanks for your careful reading. We have revised **Section 1** and added the fifth paragraph to highlight the importance of a good strategy of task allocation to gain data sharing for the performance of the scheduling of sampling intervals. If multiple sampling tasks are allocated to a same sensor node and their time windows are not overlapping, the scheduling methods of sampling intervals cannot exploit data sharing to reduce redundant data sampling. Instead, if the overlapping tasks adopt an optimal strategy of task allocation, then the sampling intervals of overlapping tasks which have been allocated to a node can be moved in their time windows to exploit the potential data sharing. Consider that a data-intensive application may produce a large set of sampling tasks. It is vitally important to adopt a good strategy of task allocation for gaining potential data sharing.

Moreover, inspired by your suggestion, we have added the definitions of the allocation of tasks and the scheduling of sampling intervals in **Section 3.1**in the new version. Additionally, the definition of MAX-DHP has been added in the S**ection 3.3**.

**Response to Reviewer 3’s Comments**

1. **Comment:** *The readability of this paper needs to be improved. The current structure is not easy for reviewer to understand the major contribution. Authors need to highlight the contribution the demonstrate the innovative clearly.*

**Response:** Thanks for your careful reading. Inspired by your suggestion, we have reorganized and polished the paper carefully in the new version. In **Section 1**. The major motivation of this paper has been carefully polished. We have highlighted the vital importance of minimizing the volume of sampled data by using data sharing. In the second paragraph, we have demonstrated the application scenarios to improve the readability of the motivation and the focused problem in this paper. Furthermore, we have revised the fourth paragraph to present the vital importance of maximizing the data sharing amongst tasks. The fifth paragraph has been revised to present the challenges when maximizing the data sharing amongst tasks in a wireless sensor network. The major difference between the previous work and this paper is presented in the sixth paragraph. Additionally, the contributions are outlined in the seventh paragraph.

We have reorganized **Section 3**. To further to improve the readability of our work, we have first presented the task model and the network model, and then formulated our problem. Finally, we have analyzed the complexity of our problem theoretically. We have polished **Section 4** and **Section 5**. In **Section 4**, we have demonstrated a basic but effective method: *COMBINE* to compute the volume of sampled data. In **Section 5**, we have proposed a solution: *CATS* to allocate tasks and to schedule the sampling intervals of tasks in WSNs.

1. **Comment:** *Authors proposed four algorithms. However, it is very necessary to explain the difference and connection among four of them.*

**Response:** Thanks for your careful reading. We have reorganized the paper to improve the presentation, and revised all algorithms in the new version. The algorithms in the previous version have been revised in the new version.

As discussed in **Section 1** and **Section 3**, task allocation and schedule of sampling intervals are extremely difficult. The main reason is that the data sharing amongst multiple tasks is hard to be computed (See in **Lemma 3.13**). We thus propose the first algorithm, i.e., *COMBINE* in **Section 4** which schedules the sampling intervals of overlapping tasks to achieve the maximal data sharing in the current state. It is noting that the memory complexity of *COMBINE* is *O(n2)*. Since the memory of a wireless sensor node is scarce, we then propose the second algorithm, i.e., *COMBINE\_2*, which adopts the same strategy with *COMBINE*, but its memory complexity is *O(n)*. Both *COMBINE* and *COMBINE\_2* are effective methods to compute the data sharing amongst overlapping tasks (See in **Theorem 4.5**).

Using *COMBINE* or *COMBINE\_2* to schedule sampling intervals and compute the data sharing amongst overlapping tasks, we propose the third and fourth algorithms, i.e., *PRUNE* and *CATS* to solve the joint optimization problem. Both *PRUNE* and *CATS* are heuristic methods to allocate tasks across wireless sensor nodes in a WSN. In a *k*-coverage and *r*-redundant WSN, *PRUNE* allocates each a task to its all *k* candidate sensor nodes, and then removes such a task from *k-r* nodes due to the amount of data sharing; while *CATS* schedules the sampling intervals of tasks to achieve as much data sharing as possible, and then allocates such tasks to a same node together. *CATS* ensures that data sharing achieved is locally maximal for every task allocation. When the task allocation is completed, both PRUNE and CATS will use *COMBINE* or *COMBINE\_2* to compute the total amount of sampled data for the wireless sensor network.

1. **Comment:** *Although authors provide an experimental test, it is also important that authors need to evaluate whether the developed experiment is valuable in practice. It will be better to configure a wireless sensor network from a practical scenario.*

**Response:** Thanks for your careful reading. Inspired by your suggestion, we have deployed a wireless sensor network in a practical scenario by using 22 nodes. It is noting that this wireless sensor network is built in a general topology, instead of the grid topology in the testbed. We have revised **Section 6.3**, and added the comparison of the amount of sampled data in the fourth paragraph in **Section 6.3** when using different strategies of data sampling.

In this experiment, we adopt the settings of sampling tasks in the **Section 6.1** and compare the amount of sampled data when varying the number of sampling tasks in Figure 13(b), and varying the value of *r* in Figure 13(c). It is obvious that our strategies of task allocation and the scheduling of sampling intervals, i.e., *PRUNE* and *CATS,* outperform *RANDOM*. The main reason is that *PRUNE* and *CATS* reduce the volume of unnecessary sampled data by exploiting data sharing effectively.

Besides, we have revised the last paragraph of **Section 6.3**, and presented the details of our simulation settings in Figure 15 which are adopted widely.