Responses to Reviewers’ Comments on the Manuscript TPDS-2010-06-0358.R1

Stability-Optimal Grouping Strategy of Peer-to-Peer Systems

by

Zhenhua Li, Jie Wu, Junfeng Xie, Tieying Zhang, Guihai Chen, and Yafei Dai

We are grateful for the detailed comments and constructive suggestions made by the reviewers and editors for our original submission (TPDS-2010-06-0358) and major revision submission (TPDS-2010-06-0358.R1). In this minor revision submission (TPDS-2010-06-0358.R2), we have carefully studied and responded to the reviewers’ and editors’ concerns in the major revision version. We sincerely hope the reviewers and editors will find this revision satisfactory.

In this file, we listed our responses to the comments item by item. In particular, the fundamental changes in this minor revision version are briefly summarized as follows:

1. ***Solution and performance in an open P2P system***. In an *open P2P system*, the session time of a newly joining node is hard to get although the session time distribution of all nodes can be easily got. Moreover, the total user number may be quite unstable (fluctuate sharply). Then the problem is *how to deal with such an open system*. Our solution is to estimate the session time of a new node as the average session time of existing nodes. As time goes, the information of this new node would be learnt, and then we can allocate it into a more proper group. As to the sharp fluctuation in user number, our solution is to recollect all nodes' information and then regroup them at intervals. We evaluate the corresponding performance through trace-driven simulations. Reviewer 2’s valuable comments have stimulated us to carefully consider the applicability of our grouping strategy in an open P2P system. They have greatly improved the completeness of our paper.

Response to the Editor’s Suggestions

Editor’s suggestions

Please carefully address all the issues raised by reviewers.
* Define the definition of stability more clearly.
* Justify how your algorithm will work without knowing some parameters in advance. If certain perdition is used, evaluate how accurate the prediction is and the impact of inaccurate prediction.
* Address the question of the simulation time and group period.

Our response

We appreciate the editor’s positive and useful suggestions, which have improved our paper step by step. In our paper of the revised version, we have carefully read and considered all the review comments point by point. We have provided a detailed response to every point, no matter positive or negative. Additionally, we have further proofread the paper and have made additional revisions.

As to Issue 1, Reviewer 2's comments have let us realize that our former explanation in Section 2.1 was brief and abstract, lacking concrete description. As a result, we use a metaphor which compares our idea to another more familiar scenario, in hopes of better and easier understanding of our idea. The corresponding paragraph has been added to Section 6.4 of the supplementary file in this minor revision version.

As to Issue 2, Reviewer 2's comments have stimulated us to carefully consider the applicability of our grouping strategy in an open P2P system. They have improved the completeness of our paper. In an open P2P system, the session time of a newly joining node is hard to get although the session time distribution of all nodes can be easily got. Then the problem is: how to estimate the session time of a node when it joins the system? Our solution is to estimate the session time of a new node as the average session time of existing nodes. As time goes, the information of a new node would be learnt, and then we can allocate it into a more proper group. Refer to Section 2.4.1 of the TPDS manuscript for the solution, and Section 10 of the supplementary file for the performance evaluation.

As to Issue 3, Reviewer 3's comments have indicated that our careless presentation might lead to seemingly confusing setting and explanation of group period. Our simulation data set is composed of three data sets. Two data sets are extracted from real-world systems, so the period is set to 24 hours. Another data set is the artificially generated data set, where "second" is the *simulation second* rather than real time. Our simulation tool for the generated data set is implemented as a discrete time event generator, and an event is generated or triggered per simulation second. We have added the corresponding explanation into Section 7.1 of the supplementary file. Besides, the reason why we usually set the period to be 24 hours lies in that the two relevant real-world systems exhibit obvious diurnal user access pattern. For a system with a shorter period, our proposed grouping strategy would be still applicable, since it does not require the period to be some specific value.

Response to Reviewer 1's Comments

Reviewer's comments:

Accept With No Changes

I find the paper of excellent quality, sound reasoning, well experimented, it could be accepted as it is. BTW, we appreciate the novel format (8-page paper + supplementary file).

Our response

We deeply appreciate the encouraging comments. The current quality, reasoning, and experiment are achieved step by step, greatly owing to the valuable reviews in two rounds. Besides, we have made further improvements in this version.

Additional Questions

1. Which category describes this manuscript: Research/Technology
2. How relevant is this manuscript to the readers of this periodical? Please explain your rating under Public Comments below: Very Relevant

Our response

Thanks for the positive comments.

Additional Questions

1. Please explain how this manuscript advances this field of research and/or contributes something new to the literature. : To construct a DHT with relative high level of stability and scalability in high-churn environments is an interesting and fundamental research issue. This paper proposes a natural way to address this issue by utilizing the benefits of hierarchical topology. However, the current version of this paper suffers from some essential problems, which should be well addressed before being accepted.
2. Is the manuscript technically sound? Please explain your answer under Public Comments below: Yes

Our response

Thanks for the summary of our paper and the problems pointed out. In fact, the clear and sound exposition is (or will be) achieved after several rounds of modifications, extensions, and improvements. We have taken (and will take) each comment item seriously, and will not stop our steps until the manuscript is truly sound.

Additional Questions

1. Are the title, abstract, and keywords appropriate? Please explain under Public Comments below: Yes
2. Does the manuscript contain sufficient and appropriate references? Please explain under Public Comments below: References are sufficient and appropriate
3. Does the introduction state the objectives of the manuscript in terms that encourage the reader to read on? Please explain your answer under Public Comments below: Yes
4. How would you rate the organization of the manuscript? Is it focused? Is the length appropriate for the topic? Please explain under Public Comments below: Satisfactory
5. Please rate the readability of the manuscript. Explain your rating under Public Comments below: Easy to read

Please rate the manuscript. Please explain your choice: Good

Our response

Thanks for the encouraging comments.

Response to Reviewer 2’s Comments

Reviewer's recommendation:

Author Should Prepare A Minor Revision

This is a revised version. Authors made many improvements based on the reviewers' feedback, thus this revision has addressed most of the issues in the original submission.

Our response

We appreciate Reviewer 2 for his (her) recommendation and detailed review comments, carefulness, and patience with our manuscript. We have tried best to carefully read these comments and make corresponding thoughts and revisions according to them. More importantly, we feel these comments do help a lot in improving the quality of our paper.

Reviewer's comments

However, I am still thinking that this paper should be further improved through following aspects:

1. The definition of stability is still hard to understand. The illustration of game theory zero-sum can show that grouping some nodes can increase the session time of the group if the living time of these nodes is disjointed. So we can say grouping more nodes can make the group stable. For the stability of the whole system, I do not understand why less variance of the stability of the groups can make the system more stable. What is the system stability, and what is the big issue that affects system stability? The definition of scalability and stability plays important roles in the grouping algorithms presented in this paper. If it is not explained clearly, readers may still have questions why nodes should be grouped in this way, why such a grouping policy can make system stable.

Our response

These are useful comments that drive us to express our definition of stability more clearly. Our former explanation in Section 2.1 was brief and abstract, lacking concrete description. Here, we use a metaphor which compares our idea to another more familiar scenario, in hopes of better and easier understanding of our idea. The following paragraph has been added to Section 6.4 of the supplementary file in this minor revision version.

Suppose we want to operate 100 web sites, and we have 500 servers to support these web sites. One server can only support one web site. Suppose each server can only run for 6 hours per day. Then the node stability is 6/24 = 0.25. We do not know when they fail, and they fail independently. Surely, it is impossible to build a perfect system where every web site has high stability. A natural idea (egalitarianism) is to uniformly allocate 500 servers to 100 web sites, with each web site supported by 5 servers. Then each web site is expected to have the same stability: Ψ = 1-(1-0.25)^5 = 0.763. For a certain web site (says A), you may feel this group stability (0.763) is not satisfactory, and then the only way to improve Ψ(A) is grabbing a server from another web site (says B). Then, Ψ(A)’ = 1-(1-0.25)^6 = 0.822, but Ψ(B)’ = 1-(1-0.25)^4 = 0.684. The gain inΨ(A) is 0.822-0.763 = 0.059, while the loss inΨ(B) is 0.763-0.684 = 0.079. Obviously, as to the whole system, the loss is larger than the gain. If A grabs more servers from B, the loss will be much larger than the gain. This simple example indicates that the natural idea (egalitarianism) is in fact the best idea. The symbol of egalitarianism is the minimum variance (zero in this example). This is why we define the system stability as inverse proportional to the system variance.

Reviewer’s comments

2. The proposed design needs to know a node’s session time when it joins the system. This information is used to choose the right groups with a similar session time. In the simulations, the session time of nodes can be easily got from the trace. However, such information is hard to get in real application. Sure you can say the session time of a node can be predicated by monitoring node session history, however, in an open p2p environment, such information for a single node is hard to get although the session time distribution of all nodes in the system can be easily got. In this revised version, authors only provided more detail to show that the session time of nodes does follow some distribution. My question is not about the distribution of session time. How to estimate the session time of a node when it joins the system? Will the algorithm still work if the session time cannot be predicated?

Our response

These are valuable comments that have stimulated us to carefully consider the applicability of our grouping strategy in an open P2P system. In an open P2P system, the session time of a newly joining node is hard to get although the session time distribution of all nodes can be easily got. Then the problem is: how to estimate the session time of a node when it joins the system? Our solution is to estimate the session time of a new node as the average session time of existing nodes. As time goes, the information of a new node would be learnt, and then we can allocate it into a more proper group. Refer to Section 2.4.1 of the TPDS manuscript for the solution, and Section 10 of the supplementary file for the performance evaluation.

Reviewer’s comments

3. The proposed design does not work for the open p2p environment. The total number of peers in the system must be known beforehand, and the user group is relatively stable. This is not applicable for many popular peer-to-peer systems.

Our response

These are also good comments that have enhanced the completeness of our paper. In an open P2P system, the total user number may be quite unstable (fluctuate sharply). Then the problem is how to deal with such an open system. Our solution is to recollect all nodes' information and then regroup them at intervals. We evaluate the corresponding performance through trace-driven simulations in Section 10 of the supplementary file.

Additional Questions

1. Which category describes this manuscript: Research/Technology
2. How relevant is this manuscript to the readers of this periodical? Please explain your rating under Public Comments below: Relevant

Our response

Thanks for the positive comments.

Additional Questions

1. Please explain how this manuscript advances this field of research and/or contributes something new to the literature. : This paper presented an analytical model to investigate the trade-off between stability and scalability of the peer grouping policy in peer-to-peer systems. By formalizing the maximum stability grouping problem, and analyzing the intractability and infeasibility of such a problem, it proposed a homogeneous grouping strategy to achieve optimal stability with guaranteed scalability. Compared with previous peer grouping policy that choosing stable peers as the group leader, the homogeneous grouping strategy tries to assign more nodes to a dwarf group and fewer to a giant group, so the dwarf group can survive for a time period equal to that of the giant counterpart.
This work is well-motivated. The idea is novel.
2. Is the manuscript technically sound? Please explain your answer under Public Comments below: Yes

Our response

Thanks for a proper summary of our paper. This summary has well caught our idea and is quite complete.

Additional Questions

1. Are the title, abstract, and keywords appropriate? Please explain under Public Comments below: Yes
2. Does the manuscript contain sufficient and appropriate references? Please explain under Public Comments below: References are sufficient and appropriate
3. Does the introduction state the objectives of the manuscript in terms that encourage the reader to read on? Please explain your answer under Public Comments below: Could be improved
4. How would you rate the organization of the manuscript? Is it focused? Is the length appropriate for the topic? Please explain under Public Comments below: Could be improved

5. Please rate the readability of the manuscript. Explain your rating under Public Comments below: Readable - but requires some effort to understand
Please rate the manuscript. Please explain your choice: Good

Our response

Thanks for the encouraging comments.

Response to Reviewer 3's Comments

Reviewer's recommendation:

Author Should Prepare A Minor Revision

Our response

We are grateful for Reviewer 3's recommendation and review comments. We have carefully addressed the problem about the group period.

Reviewer's comments

The reviewers are very satisfied to find out that the group definition in this paper is much clearer than previous version.

Our response

Thanks for the positive comment.

Reviewer's comments

The basic property of a group is that several nodes in this group can provide continuous and stable service for a period. This paper said that the period is usually set to be 24 hours for practical system.
Why the period is usually set to be 24 hours? We can guess a group may include a lot of nodes (PDA or other mobile nodes) when the period is setting to be 24 hours. Do there exist some other systems that set the period to be 24hours?
If the authors really set the group period to be 24 hours, why authors do simulation for only 10000 sec which is much less than a period???

Our response

These are very careful comments! In Section 2.1 of the major revision version, we wrote that "*Period* is usually set to 24 hours for a practical system." And in Table 2, we also wrote that "For a practical P2P system, Period is usually 24 hours." However, in Section 7.1 of the supplementary file, we wrote "We set *Period* = 1000 seconds. The system is simulated to run long enough (10,000 seconds) to collect sufficient data." Our careless presentation leads to seemingly confusing settings. Our simulation data set is composed of three data sets: 1) generated data set, 2) AmazingStore trace, 3) CoolFish trace. For data sets 2) and 3), they are extracted from real-world systems, so the period is set to 24 hours. Data set 1) is the artificially generated data set, and here “second” is the *simulation second* rather than real time. Our simulation tool for the generated data set is implemented as a discrete time event generator, and an event is generated or triggered per simulation second. For convenience, we can simply take 1000 simulation seconds as 24 hours in real life. We have added the corresponding explanations into Section 7.1 of the supplementary file.

Besides, the reason why we usually set the period to be 24 hours lies in that the two relevant real-world systems (AmazingStore and CoolFish) illustrate obvious diurnal user access pattern, as shown in Fig. 5-6 and Fig. 15-17 of the major revision version, and Fig. 3-5 of the supplementary file. Although AmazingStore and CoolFish are both mid-scale P2P systems, their diurnal user access pattern may reappear in some large-scale regional systems like PPStream, PPLive, and so on. However, for some worldwide systems, such as Skype and KaZaa, the time difference makes the user access pattern more complicated: as to users in a certain region, the period is still 24 hours; but as a whole, the system has a shorter period, maybe several hours. We have not got such worldwide trace till now, which may be an interesting future work. For a system with a shorter period, our proposed grouping strategy would be still applicable, since it does not require the period to be some specific value.

Additional Questions

1. Which category describes this manuscript: Research/Technology
2. How relevant is this manuscript to the readers of this periodical? Please explain your rating under Public Comments below: Relevant

Our response

Thanks for the positive comments.

Additional Questions

1. Please explain how this manuscript advances this field of research and/or contributes something new to the literature. : The paper studies the problem of how to group unstable nodes together to form an adequate stable service group. The objective of this paper is to maximize the stability of whole system through grouping nodes into different service groups. In solving the grouping problem, this paper gives the mathematical problem formulation of Maximum Stability Grouping, and proves the problem to be NP-hard. After restricts the problem to a homogeneous MSG problem (H-MSG), this paper proposes a homogeneous grouping strategy to fulfill the optimal solution to the H-MSG problem, and performs simulations on generated data sets and real-world traces to check the effectiveness of the proposed grouping strategy.
2. Is the manuscript technically sound? Please explain your answer under Public Comments below. : Yes

Our response

This is an appropriate summary of our paper. It has well caught our idea and is quite complete.

Additional Questions

1. Are the title, abstract, and keywords appropriate? Please explain under Public Comments below: Yes
2. Does the manuscript contain sufficient and appropriate references? Please explain under Public Comments below: References are sufficient and appropriate
3. Does the introduction state the objectives of the manuscript in terms that encourage the reader to read on? Please explain your answer under Public Comments below: Yes
4. How would you rate the organization of the manuscript? Is it focused? Is the length appropriate for the topic? Please explain under Public Comments below: Could be improved
5. Please rate the readability of the manuscript. Explain your rating under Public Comments below: Readable - but requires some effort to understand

Please rate the manuscript. Please explain your choice: Good

Our response

Thanks for the encouraging comments.

Additional Revision

1) We have proofread the paper and found some undiscovered minor presentation errors.
2) Since we have added contents into the manuscript, for space limitation, Section 3.6 is moved to the supplementary file as Section 7.5.