Response Letter to Manuscript ID TNET-2016-00326: Scalability and Satisfiability of Quality-of-Information in Wireless Networks

Scott T. Rager, Ertugrul N. Ciftcioglu, Ram Ramanathan, Thomas F. La Porta, and Ramesh Govindan

We would like to thank the Reviewers and Editor who have helped us improve the paper significantly. We revised the manuscript addressing all the comments of the reviewers. Below, we provide detailed explanations of the revision points. Additionally, for your convenience, in our revised manuscript, we have marked the text which was either newly added or significantly edited as a result of the reviewer feedback by blue. Because of space limitations, we were able to add only a condensed version of the discussion below in the revised manuscript.

A. Editor's Comments

The reviewers appreciated the proposed contribution, but also recommended a number of changes that can improve the quality of the paper. Their comments can be summarized in three important concerns: (a) the novelty of the paper is questioned by one reviewer, and sufficient argumentation from the authors is needed to address this concern in detail, (b) the related work needs to be updated, extended, and the proposed work needs to be positioned against recent literature improving in this way the presentation, (c) for a new model that aims to replace classical throughput/delay analysis, its wide-applicability to multiple scenarios needs to be demonstrated.

Response to (a): After reading all of the reviewers' comments, we see that the way we structured our manuscript and the amount of space we dedicated to certain sections overshadowed much of our proposed contribution. As we elaborate on in Response 2.3 below, our main contributions of the work lie in Section

V and beyond. We use many techniques from other work in Section IV to introduce an example application that we implemented to motivate the problem sufficiently since we did not wish to assume that all readers would be sufficiently familiar with the concept and importance of using QoI metrics. We have reduced Section IV, moving details to Appendix A to clarify and improve the focus on our main contributions.

Response to (b): Thanks to the reviewers, we realize that we did not elaborate as much as necessary on how this work specifically differs from previous works, especially from our own authors. We have expanded the Related Works section to highlight the differences above and beyond previous work:

"We provide several differences and additional analysis here compared to these works, which mainly stem from our use of QoI requirements instead of static data rates. The first difference is that we are able to evaluate performance and scalability under timeliness constraints, which is not possible under any of these prior works. Second, we illustrate the effects on an application's performance, which is not linear with respect to data rates in most cases. Additionally, we provide a formulation that includes parameters characterized by random variables. This improved modeling allows us to characterize expected delays with a probability distribution, which the previous works do not provide."

Response to (c): We understand that there are questions of the generalizability and applicability of the proposed framework to applications, networks, and protocols beyond what was presented in the submitted manuscript. In order to show more generality of the framework, we have added another network topology from Reviewer 1's suggestion, the NSF Network, including derivation of the factors and equation characterizing its limits. This added example has the benefit of showing analysis for a non-uniform topology and shows that tradeoffs in conflicting QoI requirements can be understood from using this framework. Additionally, we expand on handling other protocols and non-uniform network topologies in Section IX and in Response 3.2 below.

B. Review 1

We would like to thank the Editor who has helped us improve the paper significantly. Below, we provide our responses and indicate how we have revised the paper.

**Comments to the Author:

In this paper authors present a QoI framework to provide mathematical expressions to estimate limitations on network size and QoI requirements. They primarily focus the context dependent measures of timeliness and completeness, but the framework is generic enough to include other parameters as well.

The paper is very well written and organized. Also, the topic is very timely with increasing interest, since the notions of context awareness and information centricity are gaining popularity. In general, authors manages to describe the framework and its various parameters that affect its performance successfully, to the extent that a reader can adapt it at his/her own network topology and application scenario. Below find some comments for the various sections of the paper.

**Review:

- The Related work is very brief, although it contains all the major works regarding QoI. For a 13 page paper and a topic of this importance a reader requires more details for the included related pieces of research. Additionally, a small paragraph with the extra content (i.e., extensions) compared to refs [2][3][4] is also necessary.

Response 1.1: We have expanded the related works section, especially including more clarification on how this work is different from [2][3][4].

- The notion of ad hoc networks is rather old. The paper should probably be aligned with current wireless trends such as D2D wireless networks or wireless social networks. Authors should also search for topologies that might have useful topological attributes and are formed in social networks and can be used in their analysis in Section VI and beyond. For example a star topology is formed in D2D networks enabling the WiFi Direct technology.

Response 1.2: NEED: References to D2D?and should we add a star topology?

- The example of Section IV is very useful and detailed, but rather too extended. It spans almost two pages and disconnects the flow between the network model and the QoI model. Probably authors should reduce its size and keep only that part that is also used later in the paper. For instance the experimental results of section IV.B can be moved to the appendix, whereas section IV.A can be cut down to the very necessary attributes of the image retrieval concept.

Response 1.3: Thank you for this comment. We agree that this section was unfortunately long. However, QoI is a fairly new topic and some readers may not be familiar with it. Moreover, most of the results that follow in the paper use the completeness metrics introduced here, and our goal is to highlight the non-linearity between QoI and data rate that makes using QoI so important. Additionally, Reviewer had positive feedback for including experimental results in this section. For those reasons, we moved as much of the detail as we could from the explanations of the algorithms to the appendix, but we felt keeping the description of the experiment and results in Section IV.

- Section V, which is the main QoI scalability analysis section of the paper, is absolutely nicely presented and organized with all the necessary details.

Response 1.4: Thank you. This is our main contribution of the paper, so we are very pleased that you found it easily understandable.

- Section VI is also nicely presented. The only negative is the extending details regarding the topologies used. Especially, all the details for the grid topology should be moved to the Appendix section, since they also break the flow of the section. The used set of topologies are used for their special characteristics, but authors should include in the paper at least one generic known network topology and derive a scalability equation for part of its nodes. For instance an equation for the NSF network would increase the strength of the analysis (e.g., an equation for the bottleneck node based on its usage). Of course, this equation cannot be generic like those for the used topologies, but will prove the generality of the proposed framework.

Response 1.5: Again, we agree that this section contains details not crucial to the flow of our main contributions, so we have taken you suggestion and moved details on deriving Traffic Factor expressions to appendix B.

The suggestion to add another topology to prove the generality of the framework was also taken. We added the NSF network topology to the paper, including derivation of the factors and its scalability equation, which provides QoI limits in this case, since the network has a fixed size. In Section VIII, we also added a graph to show how one can calculate QoI limits and understand tradeoffs between them as well.

- Section VII is very useful for the adoption of the framework in real wireless opportunistic networks, e.g., sensor networks, especially the subsection for the probability of timeliness satisfiability. [Minor] Please move Eq. 10 at the same page (page 9).

Response 1.6: NEED: move equation 10 to top of appropriate page after everything else is done. I will do this.

- Section VIII is somehow too verbose and can be shortened in an attempt to save some space.

Response 1.7: Thank you for this comment. We realized that this section was repetitive and wordy. To improve it, we removed two of the examples and their explanations, both of which were somewhat redundant. To improve the overall generality and show the utility and flexibility of the work, though, we also added an example using the fixed size NSF network and showing the tradeoffs in completeness and timeliness. So while the overall section is only slightly shorter, we believe it is overall more relevant and

hopefully appears less verbose to the reader.

- Section IX nicely completes the whole framework analysis, since it examines the reverse problem of how large a network can it be assuming a target goal QoI. The only comments here is that it can be positioned earlier in the paper. [Minor 1] Eq 10 is also present in page 11. This should be Eq. 11 since Eq. 10 is on page 8. [Minor 2] Fig. 8 is not readable in a B and W print.

Response 1.8: We realize that Figure 8 (now Figure 7) is difficult to read in black and white. We could not find a way to address this issue directly, and we believe that the graph is very informative when viewed in color, so we decided to leave it and attempt to explain it as well as we could in the supporting text.

C. Review 2

We would like to thank the Reviewer who has helped us improve the paper significantly. Below, we provide our responses and indicate how we have revised the paper as a result immediately following the corresponding comment by the reviewer.

** Comments to the Author:

In this paper, the authors present a QoI-based framework that provide estimates for limitations on network size and achievable QoI requirements. In particular, the authors focus on using completeness and timeliness as QoI attributes, providing an example application and several different ways to measure completeness. The developed framework that can estimate QoI and network size limits and delays for a specific network. Furthermore, they extended this framework to model competing flows and data loads as random variables to capture the stochastic nature of real networks. Finally, this work also present the concept of scalably feasible QoI regions.

I appreciated that the authors presented some "experimental" evaluation of the QoI Model, with set of pictures. In particular, the Example Application: Similarity-based Image Retrieval.

** Review:

I disagree with the authors on the following statement, page 1, column 1: "Often, extensive simulation or experimentation testbeds must be created to test proposed network setups, which is difficult and time-consuming." Since nowadays, there are plenty of open (free-of-charge) and large-scale testbeds where the scientists may evaluate their solutions [1], [2], [3]!

Response 2.1: We included references to available simulation packages and network testbeds and distinguished the difficulties in using them that are alleviated by adopting our modeling approach.

The justification at page 1, column 2: "Experimental techniques, like Response Surface Methodology [7], for example, may be applied to solve the problem we do, but these require complex test beds instead of a compact mathematical framework.", sounds too naive! I do not see why is not that important to have real-world applied solution, since eventually, this is the objective of our research community.

Response 2.2: The goals of our work are to provide a framework to quickly determine a proposed network?s abilities and scalability as well as compare similar network setups with substituted protocols, topologies, equipment, etc. While a real-world solution is desirable for any final proposed network design, creating a real solution for each possible design choice to test limits and compare performance is severely impractical. We have expanded text in the introduction and related works to make this distinction more

clear.

In overall, I found the article superficial, while its contribution remains rather fuzzy. Imho, the authors missed the flow of the paper by concentrating too much in examples and details (such as delay estimation). More specifically, I missed the actual novelty, while its core work is inspired mainly from the literature and extended to certain level:

- "we use similarity-based image collection 21 as an example of an application that is best evaluated using QoI." This application has previously been considered in [14] and [15].

- "We use the same similiary-based image selection algorithm as in [15], but provide new methods of quantifying QoI."

- "Here, as in [16], we specify a vector of minimum values for each QoI metric, and information is evaluated based on whether it satisfies all of the QoI requirements or not".

- "To get a similarity measurement, we use the same choice as was shown to be effective in [15]."

- "A technique called Color and Edge Directivity Descriptor (CEDD) [17] provides a 54-byte vector of qualities inherent to a photograph like lightness, contrast, and color."

- "The similarity between two images can then be given as a scalar by calculating the Tanimoto Similarity [18] between their CEDD vectors. Dissimilarity is simply defined as 1 minus the similarity."

- "For the Spanner algorithm, we employ a greedy algorithm similar to that in [15]."

Response 2.3: The reviewer is correct in pointing out that a number of the techniques utilized in Section IV are not novel. As we point out in the Introduction in the paragraph beginning with "Our main contribution in the paper is...," the main contributions of the paper are in Section V and beyond, not in Section IV. Since QoI is still an emerging field and not all readers may be familiar with it, and since we believe it is beneficial to solidify its importance by showing the difference in value gained from varying amounts of data, we choose to provide an example of a real application that can utilize QoI. This comment and comments from Reviewer 1 have shown us that Section IV is too long and minimizes the main focus of our work, so we have reduced the size of that section.

Furthermore, I do not feel comfortable with the assumption that there is 100% of transmission reliability, in other words considering that by default all data packet transmission will be successful. Since such assumption present certain level of issues in the modeling phase: ?In this network, we assume a simple 3-slot TDMA scheme, which allows each node equal time access to the medium and removes any potential interference or hidden terminal issues.?

Response 2.4: We understand and agree that 100% transmission reliability is a non-realistic assumption. After reexamining our approach and calling on previous work, we realize that our framework can incorporate scenarios with losses by modeling channel rates with an effective rate that relies on packet loss probabilities and retransmission schemes.

Large part of the paper is dedicated on estimating the delay performance, which imho, misleads from the main scope of the study presented here. For example, in Section V. A. QoI Satisfiability Framework, the authors presented detailed, yet non-realistic, delay estimation modeling. In particular, the authors estimate the propagation delay, while not considering the processing and emission delay.

Response 2.5: We realize that we were not precise and clear with our definitions of delays and noting which delays we have included in the model and which we have ignored due to being outweighed. We have rewritten and expanded Section V.A to make this more clear. Specifically we point out that we focus on emission delay and ignore the actual propagation delay since the latter is several magnitudes smaller than the former.

[1] G. Z. Papadopoulos, J. Beaudaux, A. Gallais, T. Noel and G. Schreiner, "Adding value to WSN simulation using the IoT-LAB experimental platform" In Proc. IEEE WiMob 2013. https://www.iot-lab.info/

[2] M. Doddavenkatappa, M. C. Chan, and A. Ananda. Indriya: A Low-Cost, 3D Wireless Sensor Network Testbed. In Proceedings of the Conference on Testbeds and Research Infrastructures for the Development of Networks & Communities (TridentCom), 2011.

[3] http://www.wisebed.eu/

D. Review 3

We would like to thank the Reviewer who has helped us improve the paper significantly. Below, we provide our responses and indicate how we have revised the paper as a result immediately following the corresponding comment by the reviewer.

** Comments to the Author:
The authors present a framework for estimating the scalability and "QoI-satisfiability" limits of the network of interest. Quality of Service is a multi-dimensional metric to measure the value of information with attributes like timeliness and completeness. The authors claim that these metrics should be preferred over the traditional performance metrics (throughput and delay) in the sense that through their framework they

can provide a quick accurate estimation of the network?s abilities without relying on exhaustive testbeds or incredibly hard to derive theoretical bounds for complex networks.

Completeness is a measure for the necessary information you acquired for your query from the network (sum similarity in pictures), while Timeliness is a metric for delay - acquiring the information you need before the deadline for the query ends. The equations they derive are used by applying the known parameters of the network and acquiring the limit of the remaining variable of interest. For example solving the Scalability equation for T, gives back the minimum "timeliness" value (delay notion) for the network.

The results given in section VIII demonstrate the effect of varying network characteristics, like topology, number of nodes or requested QoI to resulting timeliness or completeness (covering a big percentage of the sets) and also the impact that different variables have on the network performance. The figures give an indication of the scalability of the networks to the specific characteristic. The framework can also be used to acquire an estimate of the maximum nodes a network can support based on QoI requirements (results in section IX).

The paper presents, to my best knowledge (which is not great in the area of interest), a novel idea for understanding the scalability and performance of networks through the framework presented using Quality of Information Metrics. The paper is in general well written and is steadily leading the reader to the results. Some modifications in the text are necessary to increase readability, see in the end of the review.

** Review:

The paper is thorough in defining the new aspects of QoI and the significance on selecting them for making scalability/performance observations. The example application of the paper is well presented, but what is not clear for me from reading the paper is if and how this framework is applicable to other networks or different applications. In my understanding for every new network and application this whole analysis needs to start over, from defining the vector of QoI metrics, to the models used for getting the analytic expressions for the metrics.

Response 3.1: Thank you for this comment. The reviewer is correct that factors may need to be derived for new applications and networks, but we believe the significant contribution here is providing the framework and tools to make the approach to analyzing a new network relatively simple. For new

topologies to be analyzed, we show that the network designer need to only determine the bottleneck of the network, which is often intuitive. We also provide a number of examples of deriving factors in this paper in attempt to give readers a thorough understanding of the process one should take to derive a scalability equation for a new network. Based on the reviewers' comments, we have even added the NSF network to show the applicability in finding the bottleneck and analyzing QoI requirements for an arbitrary topology, which we believe further exemplifies the generalizability and ease of extending to new networks and applications.

I think it would be interesting if there was a bigger variety of options in the modeling choices on the latter sections of the text and also have some references to explain the reasoning behind the choices. For example why are line, grid and clique topologies the interesting topologies for wireless networks? Why TDMA? Is round robin like scheduling good to capture the performance of clique ? which appears to be always underperforming to grid?

Response 3.2: We appreciate this comment and have expanded explanation on these points. Again, we point to the addition of a new network topology being analyzed in the paper, providing more variety in application of the framework. We feel that attempting to add even more would lead to some redundancy and take focus away from the main contribution of the framework. Since we are introducing an approach to analysis that readers may not be familiar with, we choose TDMA with round-robin scheduling for ease of exposition, but we address modeling other protocols in Section X.

Sections VII,VIII and IX properly display through plots and text the value of the current work into making network design decisions based on the presented framework for the selected example.

Response 3.4: Thank you for this comment. We very much appreciate that the graphs and explanations were understood and supportive to the focus of the paper.

I would recommend reading thoroughly and correcting the text. Some indicative parts that you should consider clarifying or rewrite are: 1) Section II. Paragraph 3: Dose-¿Does. Also consider rephrasing. 2) Section V. A. Paragraph 1: We use a fixed value the data size? of our applications in section VI. 3) Section V.A. Paragraph 4: Definition of Channel Factor. 4) Section V Last Paragraph: To show this usefulness? 5) Section VI.A. Paragraph 2: Not clear? 6) Section IV.A. : ?into a k clusters...

Response 3.5: We thank the reviewer for their careful reading. We have gone through and fixed each of these typos.