Duke University
Campus Box 98900
Durham, NC 27708

August 15, 2016

Professor John Carter
Associate Editor
SIAM Undergraduate Research Online

To Professor Carter:

We are resubmitting our paper "Rumors with Personality: Differential and Agent-Based Models of Information Spread through Networks" by Devavrat V. Dabke and Eva E. Arroyo for your consideration. We would like to thank the reviewers for their comments and have revised our draft accordingly.

Overall, we continue with the same three models as in the previous draft, but with clarifications on their similarities and differences, as well as on our choice of parameters for each. Because of the confusion in introducing the ISTK model itself—originally just a set of four equations—we added a new flowchart and restructured the section (Section 1.3). We feel these steps greatly clarify the way in which populations move within our model. To better explain our results, we updated several figures, and rewrote more precise descriptions of our data. Finally, the conclusion better describes how individual feature vectors relate to the spread of the rumor throughout the population.

Below, we provide detailed reference to each of the comments left to us by the reviewers. We greatly appreciate the feedback we received, and we feel that our manuscript has improved because of it.

Please let us know if there are any more questions or comments concerning our paper.

Sincerely,

Devavrat V. Dabke Eva E. Arroyo

1 Response to Reviewer 1

1.1 Part 1: Typographical / Grammatical / Stylistic revisions

Comment 1
On page 2, sentence 2 of section 1.2 (ISTK Model) has issues with verb tenses that do not match (We took vs. initialize).

Response 1
This was changed so that tenses agree.

Comment 2
On page 3, near the end of paragraph 2, of section 1.3, there is a sentence that begins "The second term of Equation 1 accounts for the conjugate of 'believing the rumor.' " I don't believe that "conjugate" is the appropriate word here. Perhaps "complement" or "opposite" would be more accurate?

Response 2
The word "conjugate" is now "complement."

Comment 3
The axes on all graphs in this paper (as well as the traces of the data being plotted) were too thin. They barely showed up on the hard copy of the manuscript that I printed for myself. They were also rather faint on the screen rendering of the electronic copy. The vertical and horizontal grid lines on the figures are the hardest of all of these features to see. If the authors could increase the weight of the axes, the traces of the graphs plotted on them, and the grid lines the figures might be easier to read.

Response 3
We darkened and thickened the lines and axes.

Comment 4
The axis label for the independent axis on figure 2 says $\alpha 1$ value (no subscript), when the parameter in question is referred to as α_1 (with a subscript) in the text. These should be consistent.

Response 4
This is fixed to be correctly subscripted (note the figure is now Figure 3).

1.2 Part 2: General Recommendations

Comment
It is my opinion that the manuscript "Rumors with Personality," by Dabke and Arroyo (supervised by Layton) is of sufficient quality and interest for publication in SIURO. While I've found no flaws in the manuscript that warrant major revision, I'd like to make some suggestions for points to make in the conclusion. The authors point out that the richest and least trivial dynamical behavior seems to be present in agent based models with feature vectors that describe both agents and rumors. These models, in many instances, exhibit rumors that die out before all members of the network (even if it is fully connected) observe it. The authors already speculate that it might be possible to engineer a perfect rumor that will fully saturate the network. It's not clear to me if this engineering would be more dependent upon the feature vectors of the agents in the network or upon the topology and architecture of the network itself. Some clarification (or at least mention) of this might be helpful. In any case, other questions for future study that might be of equal interest would be is it possible to design a network topology that is somehow not so restrictive that it still allows agents to interact in ways the people expect of social networks yet it also effectively dampens the spread of as many rumors as possible. Likewise, would it be possible for an agent within a network to engineer their own feature vector (even if it means lying about their personal attributes when creating their account) in a way that inoculates them against the spread of a majority (or at least a plurality) of the prevailing rumors.

Response
The new final paragraph in our discussion addresses this suggestion. Essentially, we consider both how the characteristics of the individuals and the rumor affects the spread, as well as the network topology. Additionally, we would like to note that we assume that our data was representative of a "true" set of features for an individual. The modeling of dishonesty and lying is compelling, and it is something that we would like to consider for the future. Moreover, each feature in the characteristic vector for an individual represents something to which they would positively respond; in that if they had a coding for "English" as a language they spoke, then a rumor with the corresponding feature "English" would appeal to them. However, it is possible that an individual is very interested in "French" and a feature corresponding to "French" in a vector would also cause them to be more likely to spread the rumor. The features we pulled from the dataset provided a reasonable set of data to generate feature vectors. However, we would like to note that the underlying assumption we make is that it *is* possible to find feature vectors which map to how an individual might respond more positively to corresponding features in a "rumor." We have clarified this idea in our introduction and discussion.

2 Response to Reviewer 2

2.1 Introductory Comments

Comment 1
Overall, in the current state of the paper, I do not recommend the paper for publication. There are two main issues:

Response 1
We believe that we have corrected and addressed the issues that were outlined.

Comment 2
It appears that this paper tries to accomplish too much. There seem to be 3 models, none of which are explored much in depth and not enough is said about comparisons between them.

Response 2
Throughout the paper we have tried to make clearer that the three models form a progression, and are based on the same model: the ITSK model. The differential ISTK model provides a baseline for our model by demonstrating rumor spread over a homogenous, generalized population. The simple, agent-based ISTK model introduces more complexity, particularly that of a network. With the agent-based setup, we gain more control over the model and can take into account the rumor's spread over the specific network structure provided to us by our Facebook dataset. Finally, we arrive at the feature-vector, agent-based ISTK model which allows us to track the "personality" of a rumor as it spreads through the network. We can take into account the demographic information over the population (again, by using our Facebook dataset), in addition to the network.

The entire paper demonstrates how including a network and a "personality" changes the way a rumor spreads. Our introduction has been modified to clarify this progression and unify the three versions of the ISTK model. We also have standardized that names we use for these three different versions of the ISTK model throughout the paper, so as to further demonstrate their core connection.

Comment 3
Because of the first issue, there is not enough detail in the models. For example, in the ISTK differential equation model, results are shown, however there is not sufficient detail in all of the choices of parameters to reproduce the results. I believe this is true about all of the models.

Response 3
We have revised the introduction, including major revisions of Section 1.3 that includes a flow chart to better express our model. We made the language in Sections 2 and 3 more precise, so as to better communicate the specific workings on each model. Section 2.2.2 has been updated to clarify the parameter choice for the differential equation. Section 3.1 has been updated to clarify the parameter choice for the simple, agent-based model. Section 3.1 has been updated to clarify the parameter choice for the feature-vector, agent-based model.

2.2 Section 1.2

Comment 1
Often language is used about the personality of the rumor, when I think it the personality of the individual instead.

Response 1
In the feature-vector, agent-based model, we equip rumors with a feature vector. This feature vector parallels the demographics (encoded as vectors) for each agent in the model. Therefore, we define this feature vector as the "personality" of the rumor. The rationale behind this name was two-fold: we believe that it provides a succint link from the feature vector of the rumor to the corresponding characteristic vector of the agents; we also believe that this term incites intrigue and colorfully illustrates the importance and dynamism of the rumor. The "personality" of the rumor is our term for its characteristic feature vector. We kept the terms and clarified its first usage in this paper, which is in this section, in the second paragraph.

Comment 2
For example in the first paragraph: "Instead of assuming that every individual is equally likely to spread any rumor, we assumed that the rumor's personality and the demographic information of each individual affected the likelihood of the rumor to spread." I would argue that the rumor has some characteristics perhaps, but not a personality.

Response 2
(See above comment/response)

Comment 3
A few sentences later you say: "The similarity of the rumor to the individual…" again I would argue that rumors and individuals are not similar. Also, it appears that in your model here that you spend all of your effort on modeling the individuals (either as "ignorant", "spreader", "stiiňĆer" or "knowledgeable") and didn't address the characteristics of the rumor.

Response 3
We clarified that this is the "[t]he similarity of the rumor's personality to that of the individual's" to address this relaxed use of the word "similarity." Overall, when we discuss the rumor, since it has a "personality," i.e. a feature vector that is analogous to those of the agents, we felt comfortable anthropomorphizing the rumor. That being said, we have minimized this effect. Furthermore, we made our language more concise and direct when discussing the rumor, so as to highlight the importance and significance of the rumor's personality.

2.3 Section 1.3

Comment 1
It's a little awkward starting a section with equations. How about: "The equations that describe the ISTK model are:"

Response 1
This section was entirely revamped for clarity, and includes this change.

Comment 2
Overall in this section, I find it difficult to follow. Is any possible way of simplifying the explanation? A possible way to accomplish this is to list all of the terms on the right sides of the equations (without signs) and explain them. For example, the terms all either interaction terms (the products of two individuals) or the term S/d. After explaining the interactions, the equations could then be more easily understood.

Response 2
This section has two major changes to address this concern:

1. We added a flow chart to provide a powerful visual representation of our equations. We believe this greatly clarifies the way in which an individual can move from one population the next.
2. Instead of a block of four equations, we separated them out and provided their corresponding explanations directly below each one.

Comment 3
It would be helpful and quite relevant to explain the differences with these equations and those from the Daley-Kendall model.

Response 3
We have added a paragraph to the end of this section to address this concern.

Comment 4
Equation (1) seems to simplify to $dI/dt = -c(ell)SI$. Perhaps a comment of why you don't do this.

Response 4
We added a comment in the paragraph directly below this equation to address this concern.

2.4 Section 2.1

Comment 1
what does "ran the differential model" mean in the first sentence? Also, is this the ISTK model or something else? You say "differential model" throughout the paper. If this is the ISTK model, use that terminology. I think it's clearer.

Response 1
We changed the inappropriate phrasing "ran the differential model" to "we solved the differential model." We agree that our terminology was loose when referring to the models. Throughout the paper, we enforced consistent terminology for the three equations: differential model; simple, agent-based model; and feature-vector, agent-based model. At times, we also omit the "agent-based" term when discussing the "simple" and "feature-vector" model when we feel the context is sufficiently clear. We have also clarified that all three models are versions of the ISTK model.

Comment 2
You are talking about comparing to a stochastic agent-based model, which is below this in the paper. It's confusing in this order. (After a reread, I understand what you are saying but I'm not sure you need to mention it yet until you do a comparison).

Response 2
This confusing part was removed.

2.5 Section 2.2

Comment
Overall, it seems that you are estimating the parameters on the right side of equations (1)–(4). You mention c and then tau [τ] (which is not one of the parameters in (1)–(4)) and say something about ell, but not specify it. There are also alpha [α] terms and d. What about these? It appears from your results that you vary these, but you should say something about them here.

Response
We have restructured the equations section (Section 1.3) to address this. We have additionally clarified this through Section 2.2.2 on estimating parameters in the differential model.

2.6 Section 3.1

Comment 1
You say "Facebook differential model". It's not clear what you mean by this. Is it the ISTK model or something else?

Response 1
We specifically changed this to simply say "the differential model." As mentioned in our responses to the comments for Section 1.3, we have addressed our sloppy references to the three models by enforcing a consistent naming scheme.

Comment 2
Because I'm not sure what the Facebook differential model is, I'm not sure exactly what you have done. It appears that either (1) you have run the ISTK model with some parameters the you estimated from Facebook or (2) you collected data from Facebook.

Response 2
All three models are versions of the ISTK model. Again, we worked on keeping our names for the three models consistent throughout the paper. In addition, we have clarified the relationship between the differential model and the simple, agent-based one with the first paragraph of this section.

Comment 3
I'm guessing an agent-based model is relatively standard. If so, a reference for this would be quite helpful.

Response 3
We have added a reference for generally using agent-based models for stochastic network dynamics. Specifically, this is the reference to Snijders, et al.

2.7 Section 3.2

Comment 1
Again, as in 3.1, a reference for this would be helpful.

Response 1
We believe that our added reference to Snijders, et al. in Section 3.1 suffices.

Comment 2
It appears that your agents and rumors have dimension 195 and you list some of them. I can understand some of those features being common to both (like language, in which it is interpreted differently: for the agent, can the person speak language X and for the rumor, is it written in language X), but most don't appear to be relevant to one or the other.

Response 2
We assume that a rumor can be targeted towards a particular demographic. The particular features selected, while roughly representative of real demographic information, introduce a mechanism to consider targeting rumors within the context of a (social) network. We do not claim that we are precisely modeling particular rumors and their appeal across certain demographics; instead, more generally, we demosntrate that it is in fact possible to model a targeted rumor. We have added a few sentences to clarify this point.

Comment 3
Again, you mention at the bottom of the section that you ran 300 trials. It's not very clear what this means.

Response 3
We greatly clarified this in the beginning of Section 3.1.

2.8 Results/Discussion

Comment 1
You should list exactly what parameters you used to produce the plot in Figure 1. It would be nice to try to reproduce this, but I can't without the exact parameters.

Response 1
Section 2.2.2 provides parameters for the differential model. This should be enough to reproduce our results. n.b. the addition of another figure has displaced our figure count; this figure is now Figure 2.

Comment 2
Figure 3: you have a disconnect between the legend and the vertical axis. Is it 25% or 50%?

Response 2
This is corrected (now Figure 4). The legend and axis now read 50%.

Comment 3
A bit confused about this also. The steady state ISTK model for each of the 4 values are single numbers (unless you varied the parameters). What exactly does the difference plot show? Is it just the data from the agent-based model minus the steady state and then produced as a plot?

Response 3
Yes, this is clarified in the caption and in the text that follows this figure (now Figure 7).

Comment 4
I think it would be nice to have a plot that directly compares Figure 1 with Figure 5. This seems like a natural comparison to do.

Response 4
This is what Figure 7 attempts to accomplish, and so we have added some sentences to highlight that this is a comparison of Figures 2 and 6.

Comment 5
Again, it appears that Figure 7 shows for the feature-vector model the same plot as figures 1 and 5. This isn't pointed out clearly.

Response 5
A sentence has been added to the description of this figure (now Figure 8) to underscore this fact.

Comment 6
It appears from the discussion that the feature vector may not accurately model real-world (or at least it is not consistent with the other two models). Other than the fact that the authors studied this, is this necessary in the paper? I argue that there is too much in this paper already, perhaps remove this model and stress the similarities of the other two.

Response 6
We have considered this and conclude that our entire paper leads toward the feature-vector model. We would like to stress that the feature-vector, agent-based model is the most interesting one, and it builds directly off of the simple, agent-based model by adding the concept of a rumor's "personality." Furthermore, the simple, agent-based model builds off of the differential model by incorporating a network, instead of using a homogenous population. For these reasons, we believe that the main intrigue of the paper rests with the feature-vector model, and we wish to continue to include it.

2.9 Final Comments

Comment
Another big question that I have is that one item from the introduction that appears important from their point of view is that the authors have included another another term from the Daley-Kendall model. This was explained in section 2.1. This is mentioned, but no comparison to the original (IST) model is explained. Has this been done already? If so, reference it and if not, an explanation is warranted.

Response
We have added more detail on how and why we diverged from the IST model. In conjunction with our references to other explorations of rumor spread, we believe that we now sufficiently address not only how our model differs from the Daley-Kendall (DK) one, but also why we differ from their model.