Minutes of Monthly Webinar No. 5:
INFORMS RAS · 2020 Problem-Solving Competition
June 24, 2020 · 9:00 AM EDT

RAS Attendance:

Krishna Jha (Host / Problem Owner) · Jay Baillargeon (Problem-Solving Competition Chair)

Agenda:

- Welcome & Important Reminders
- Open Forum for Participants' Q&A
- Adjourn (Next Monthly Webinar: **Wednesday, July 1, 2020**)

Conference Call / Zoom Webinar:

- Jay Baillargeon and Krishna Jha welcomed those who attended the fifth webinar for the 2020 INFORMS RAS Problem-Solving Competition and re-iterated that the goal of the webinar is to provide an open forum to gather feedback for the problem at hand from the RAS community as well as address questions from participants.
- Mr. Baillargeon then proceeded to discuss some important reminders related to the 2020 INFORMS RAS Problem-Solving Competition:
    - Given the affect this unfortunate circumstance has had on the schedule, the Committee agreed to change the date that the Quiet Period will begin from Friday, June 26, to Friday, July 3, starting at 11:59 PM CDT.
    - In addition, an additional webinar will be hosted before the Quiet Period next Wednesday, July 1, to ensure all participants are comfortable with the validation dataset and have has a chance to have their questions addressed.
- Following his update on recent news and an overview of the validation dataset, Mr. Baillargeon opened the forum for questions, comments, and any discussion related to the problem statement and datasets. This brought about the following questions:
    - **Is the "Train Mvmt Data" provided in the validation dataset the actual arrival/departure information, or the timetable information?**
        - Mr. Jha responded that the arrival/departure times in the "Train Mvmt Data" sheet are timetable information; the times reflect the planned arrivals/departures of each of the trains represented in the dataset.
        - A follow-up question asked, **"The timing for the same train is different between the two days provided in the dataset. Why is that?"** In response, Mr. Jha noted that the dataset provided was modified to assist with the optimization of the entire system. As such, it was advised not to assumed that the schedule will be the same for both days and, in turn, each train provided can be considered separate. In this way, the problem is capturing some realistic scenarios of freight railroad operations.

- o **Is the problem aimed at predicting the arrival time of the trains to the station and minimizing the deviation from the planned arrival times provided? If so, wouldn't we need the actual arrival/departure times?**
  - Mr. Jha confirmed that this is the correct way of interpreting the problem. Although, the data provided is considered to be the timetable for the trains provided and the operations should follow that as closely as possible. Also, he re-iterated that it is important to consider each train provided (even if they have the same train number but operate on a different day) as a completed independent train.
  - Mr. Jha also noted that the goal of the problem is to run the trains as closely to the provided schedule as possible.
- o **Will there be an additional dataset provided with the actual arrival/departure times and, if so, how large is the dataset?**
  - Mr. Jha noted that there are no plans to release additional data, so what has been provided will be all that is available for this year's competition.
  - In addition, he also noted that the way the judges will be assessing each of the submitted solutions will be (1) determining how effective the model is at reducing the overall delay within the system, and (2) the number of trains of delayed. This is in addition to the criteria provided on the 2020 INFORMS RAS Problem-Solving Competition website. As such, an additional dataset with actual times is not necessary.
  - Mr. Baillargeon added further that a rubric is being developed for the judges of the competition and, as long as there are no issues with doing so, that will be shared with participants so they are aware of how each submission will be judged exactly.
- o **Which is better with regard to the problem: meeting all the constraints and conditions of the problem, or minimizing the deviation from the planned arrival times provided?**
  - Mr. Jha stated that the latter is more important; that is, reducing the difference between the actual and planned arrival times. Although, this is still an optimization problem in that parameters (e.g., crew availability) will result in some amount of delay and it is the goal of the participants to effectively manage the delay and adhere to the schedule as much as possible.
- o **Can we delay the departure time of a train leaving from the origin?**
  - Mr. Jha noted that this is indeed possible.
- o **Given there are two routes in the network provided, do participants need to optimize separately or it will be integrated optimization?**
  - Mr. Jha noted that we are doing an integrated optimization and, as such, both routes should be considered simultaneously.
  - A follow-up question asked, **"In that case, how should this second route be approached as there are no trains operating on this route?"** In response, Mr. Jha and Mr. Baillargeon will look into why no trains operating on this alternate route were provided in the validation dataset.
- o **Since each train is assumed to be different on each day, the planning horizon is assumed to be two days, correct?**

- ▪ Mr. Jha stated that this is correct.
- o **In "Distances" sheet, from row 107 to 131, the "From" and "To" stations are same and distance is also given for them. What do these mean?**
  - ▪ Mr. Baillargeon noted that this was an oversight; this data was included in the datasets from the 2018 competition. The intention was to delete them when the modifications were completed but, in the rush to get the validation dataset out to the participants, this was missed. Please ignore them for now and, in the meantime, they will be deleted other discrepancies will be address and a corrected dataset will be re-posted.
- o **In the "Train Mvmt Data" sheet, for train number 815, the Work_Ordr flag = Y at station Wt. Although, in the "Stn Order & Details" sheet, station Wt does not have a yard. Is the data incorrect?**
  - ▪ Mr. Baillargeon noted that, even though a Work Order flag is given, it doesn't necessary mean it will be at a yard. The yards given in "Stn Order & Details" are listed as addition places for temporary storage. So, for the Work Order flags, participants should assume that an industrial spur will be available at those locations to pick up or drop off cars. These spurs will be separate from any sidings or yards present and, as such, will not hinder other train movement into the sidings/yards.
  - ▪ For example, if a train is at Wt dropping off cars in the industrial spur, a following train can still use one of the sidings at Wt as well (assuming the sidings are not occupied in the first place). Likewise, if a train is using one of the sidings at Wt, a train scheduled to pick up or drop off cars can still do so at Wt as the industrial spur is separate. This is established in Assumption #7. Although, only one train and pick up or drop off cars at Wt at any time.
- o **What specific information needs to be provided for the evaluation of the solution participants submit (i.e., the expected output)?**
  - ▪ Mr. Jha noted that, for each train, the actual arrival/departures times for each station and how much delay was entered into the system from each of the random variables provided.
  - ▪ Mr. Jha also added that the judges will look at how the submitted solutions handles the delays and how effectively the model minimizes the delay in the system.
- o **Which is more important for this competition: How quickly the model can arrive at a solution, or how effective is the modeled solution?**
  - ▪ Mr. Jha stated that each of these aspects will be considered by the judges for the submissions.
- o **For the current competition, should the amount of delay be incrementally reported at multiple times over the two-day timeframe, or only once?**
  - ▪ Mr. Jha stated that the delays can be reported only once at the end of the two-day planning horizon.

- Although, in a real-world environment, freight railroad operations will be running models such as these very frequently to determine the amount of delay at any given point in time in their network.