

Comments from the editors and reviewers:

(In this response, we will use underscore to highlight the major modifications we made in the main text.)

1. Reviewer 1

This paper addresses a relevant topic and offers some insightful results. It does need however in my view to undergo substantial revisions before becoming suitable for publication. In the following, I note 9 comments which I consider major and thereafter list minor comments.

Major comments:

1.1. The paper is currently unnecessarily lengthy. I think the essence of the paper and its substance are such that do not require 16 figures and 29 pages. The paper requires some considerable trimming.

Response: We appreciate the suggestion. To focus on the core conclusions of the paper, we delete many less relevant results, analyses, and especially table and graphs, reducing the paper to a manageable length. There are several major adjustments:

1. We removed the empirical tactic behavioral strategy from the paper. This strategy has some interesting properties, but is not essential to the main message.
2. We removed all non-RTI trip planning strategies from the mapping of spatial patterns. The focus in that section is now on RTI-based trip planning strategies and others are benchmarks (discussed in 1.7).
3. We removed some theoretical space-time graphs from the paper.

We managed to reduce the word count from 29 pages to 24 pages, despite the inclusion of additional analyses requested by the reviewers. We also reduced the number of figures from 16 to 9 and removed unnecessary equations.

1.2. The authors overstate their results. In essence, the obviously naive approach of planning to arrive the same second as the bus is expected to arrive can be expected to often (and even a majority of cases if there is even a slight systematic bias towards over-predicting the remaining time for arrival) result with missing the bus. I find the so-called the insurance buffer to be the more innovative element in this study. Note its affinity with the notion of 'hidden waiting time' which refers to a high (e.g. 95th) percentile of the expected distribution.

Response: We reduced some claims and results about the greedy tactic and added some explanations in sections 3.3.3 and 4.1 (second paragraph). (We also discuss this issue below in responses to comments 2.1, 2.7, and 2.13). However, we believe that the greedy tactic is worth retaining in this study for the following reasons

1. A greedy tactic may be naïve but is the default tactic in many transit planning apps. For example, the Transit apps tells users to “Go in X min.” to catch a specific bus. Users may not act that naïve in reality; they may employ a safer tactic with an insurance buffer. But since a greedy tactical is what is provided at face value by transit planning apps, we believe it is worthwhile considering in the range of behavioral strategies we explore.
2. To improve performance and optimize the insurance buffer, we need to first calculate the performance of greedy tactic, which is the upper bound of the RTI-based trip planning strategies.

Regarding an “affinity with “hidden waiting time” – thanks for this pointer! Furth and Muller (2006) introduced the concept of hidden waiting time or budgeted waiting time, which demonstrates the exact purposes of insurance buffer: to ease the unreliability of the arrival time. Moreover, the concept of offered waiting time in this paper can be connected with insurance buffer using a schedule tactic. We add the reference in the section 3.3.4 (prudent tactic). We use this paper as a reference to justify insurance buffer prudent tactic as a common strategy to avoid missing a bus.

1.3. Section 2.1 is redundant. I suggest the authors instead review the methods of quantifying the impacts of RTI on waiting times and thereafter report their findings. Also the relevance of the paragraph on surveys and their costs in 2.2 can be removed without any loss of information, unless the authors intend to specifically comment on those conducted in the context of RTI and waiting times.

Response: This is a good suggestion. We removed the former section 2.1 and organize the literature review as the reviewer suggested: we first review the methods of quantifying; we delete most discussions about the self-reported surveys to make the discussion more relevant. We also added a new “Findings” section in section 2.2 to report relevant studies’ findings. We first discuss prior studies’ conclusions about the RTI’s positive impact on the perceived waiting time and actual waiting time. Then we discuss other studies that conclude RTI has limited impact on the perceived waiting time and actual waiting time. Finally, we conclude the discussion by demonstrating the academic gap in this area and the motivation of the paper. We believe the literature review is now more complete and relevant to the topic of this paper with the useful supplements proposed by the reviewers. We thank the reviewer again for the great suggestion.

1.4. The literature review misses analytical approaches such as the studies performed by a study published in this journal ("Evaluating the added-value of online bus arrival prediction schemes") and a related study published in the Journal of ITS ("Real-time bus arrival information system: An empirical evaluation") by the same authors. These studies also considered the impact of RTI on waiting times along the line, as well as function of the prediction horizon which is equivalent to the analysis in relation to walking time performed in this paper. Moreover, the comparison with a waiting time resulting from following the schedule has also been performed. The authors should better acknowledge how their work relates to previous work.

Response: Thank you for pointing this out. We added these very useful references to the corresponding part of our analyses:

1. We also add the two mentioned papers in the literature review in section 2.1 and section 2.2.
2. We add *prediction horizon* and relevant conclusions (Cats and Loutos 2016a) to the last paragraph of section 2 and sections 4.3.1 of the risk of missing bus and walking distance influence;
3. We also add the two papers in section 4.3.3 mentioning it is common to compare the performance of RTI scheme/strategy with the schedule.

We believe these adjustments will make the paper more connected with the prior research. We appreciate the values of all the mentioned papers and we thank the reviewer for providing these precious references.

1.5. The presentation of the method can be dramatically shortened. GTFS and APC are by now very standard data sources in transit research. The notion of "reclaimed delay" is also not new. The impact of initial delays on running times further downstream has been extensively studied with conflicting evidence, see for example:

El-Geneidy, A. M., J. G. Strathman, T. J. Kimpel, and D. T. Crout. 2006. "Effects of bus stop consolidation on passenger activity and transit operations." *Transportation Research Record*
El-Geneidy, A. M., J. Horning, and K. Krizek. 2011. "Analyzing transit service reliability using detailed data from automatic vehicular locator systems." *Journal of Advanced Transportation*.
Cats, O. 2018. "Determinants of bus riding time deviations: Relations between driving patterns and transit performance". *Journal of Transportation Engineering*.

Similarly, departure strategies that have been reported elsewhere like AT can be described briefly.

Response: Thank you very much for pointing this out; we found these papers very useful. The notion of "reclaimed delay" and arbitrary tactic are not new. We add the references in the section 3.2 and briefly discuss the conclusions. The prior literature focuses on bus running time (El-Geneidy et al. 2006; El-Geneidy, Horning, and Krizek 2011), system performance (Cats 2019; Park et al. 2019), and ridership (El-Geneidy et al. 2006; El-Geneidy, Horning, and Krizek 2011). Although the papers did not primarily measure the impact based on passengers' perspectives with real-time data, they are good references about the implication of stop consolidation and riding time deviation's impact on the waiting time. In terms of AT, we add corresponding literature to the paper in section 3.3.1 (arbitrary tactic). Again, we appreciate all the mentioned papers' values and hope the added references will improve the paper's connection with the prior research.

1.6. The authors refer to measurement error and schedule recovery efforts as an explanation for an over-estimation of bus arrival time. Even in the absence of those, an over-estimation could

simply result from traffic conditions including short signals and skipping stops (no boarding and alighting passengers).

Response: This is a useful comment. We extend the concept of “reclaimed delay” to be inclusive of driver accelerating and add the scenarios the reviewer mentioned in the fourth paragraph of section 3.2. We also make corresponding adjustments in the fourth paragraph of section 1: we now use schedule recovery efforts as an example of reclaimed delay, instead of the only cause of the over-estimation.

However, this will not impact the conclusions in the paper, because the definition of reclaimed delay in the paper does not depend on the scenario of “driver accelerating”. Meanwhile, our calculation also did not rely on the specific scenarios. Just like the definition, we will calculate the time difference between the actual time of departure and the expected time of departure at the stop. During this process, there is no assumption or dependence of a specific scenario.

1.7. A critical point is that IB is introduced only in PT (prudent tactic) but none of the other TPSs (trip planning strategies) includes an element that is conscious of risk-taking. ST (schedule tactic) can also include an IB (insurance buffer) term, i.e. avoiding just missing the bus. This applies also to ET (empirical tactic).



Response: This is a good question - we also had this same question at the very beginning of this project. However, we originally chose not to include analysis for the schedule tactic and empirical tactic, and we stand by that decision in the revision.

We originally did not include this as part of the ET because realistically, people cannot perform very difficult optimization operations without any help from a computer. The two learning functions – maximum and average – and especially the maximum learning function are simple and practical for a normal passenger to perform independently. However, finding the optimal insurance buffer will take considerable amount of computational power and achieved history data, which cannot be performed realistically by a public transit user. But, this point is now moot since we deleted ET from the paper for length reduction.

The focus of this paper is RTI-based strategies; we use non-RTI strategies as benchmarks, as we discuss in sections 3.3.1 and 3.3.2. Because of this, we chose the simplicity of a straightforward ST. This is similar to the benchmarks of “static information” (equivalent to schedule tactic) and “a commonly deployed scheme” in Cats and Loutos (2016).

1.8. Notations throughout the manuscript are sloppy. For example, introducing if definitions without indicating the value taken otherwise. There are also other matters, needs to be revisited carefully.

Response: We referred to the notations in Cats and Loutos (2016) and Park et al. (2019) and rework all of our notations with simpler and more intuitive expressions. We also change the definition of greedy and prudent tactic from equation to pseudo code, which are mentioned in the comment. We believe the definitions are more rigorous.

1.9. Some of the conclusions may not be transferable, can the authors please reflect on that? For example, the performance of ST obviously depends on the on-time performance of the service under consideration (in particular, the share of early arrivals). The spatial pattern discussed in 4.3.3. is also clearly caused by the deterioration of the on-time performance further downstream.

Response: This is a reasonable comment. The representativeness and transferability issue are always a concern for any studies based on regional data. Specific responses:

Performance of ST and on-time performance of the service.

This is a good point and we added corresponding clarification to section 4.3.3. Schedule tactic's performance does depend on the on-time performance of the service according to its definition. The delay propagation from upstream to downstream is very common for different transit systems (Chen et al. 2009; Huo et al. 2014; G. Liu, Shi, and Qiu 2016). It is noteworthy that schedule tactic usually has a very low share of early arrivals. We also made similar clarification in comment 2.9. The risk of missing bus for schedule tactic is the lowest among other strategies; and 6% is a rather low absolute value. Therefore, this conclusion is transferable for many systems which have same issue of delay propagation and deterioration of the on-time performance.

However, whether the conclusions introduced in this paper are actually applicable for other transit systems still requires the validation by further research based on actual real-time data in the system. This is another contribution of this paper: we provide a robust and reproducible system that can be easily applied to other transit systems with GTFS or GTFS-APC data support. We encourage more studies to expand the concept and analyses to other transit systems to validate the introduced conclusions. We also add to the possible future research directions in the last paragraph.

Spatial pattern of ST-PT optimal waiting time difference and its relevance with delay propagation.

This is a very good point and we add this to the section 4.3.3 to clarify the cause of the phenomenon. Besides delay propagation from upstream to downstream, the prudent tactic's bad performance in the low frequency section and relatively stable pattern in the high frequency section is also another reason. We thank the reviewer for the useful suggestion.

We also comment on other representativeness and transferability issues mentioned by the reviewer 2 in comment 2.4, 2.9, and 2.12 and make corresponding adjustments in the paper according to the suggestions.

Minor comments:

1.10. Suggest to shorten the title. After the question mark can simply have only "An empirical analysis"

We changed the title according to the suggestion.

1.11. TPS (trip planning strategy) is mentioned in the last paragraph of Section 1 but has not been introduced yet

We rephrased the last paragraph of section 1 and removed the new terms that have not been introduced. We also removed all unnecessary acronyms, such as TPS (trip planning strategy), according to the reviewer 2's comment (2.17).

1.12. Add axes titles in figures 1 and 2

We added axes titles for Figure 1. Figure 2 is removed in this draft.

1.13. Broken references in section 3.3

We fixed the broken references. We thank the reviewer again for the effort.

1.14. Frumin and Zhao (2012) is not the original source for Eq. 3. Please refer to the original contributor.

We removed the equation to avoid confusions according to this comment and reviewer 2's comment 2.8.

1.15. The description accompanying Figure 3 is not sufficiently clear, please revisit.

We thank the reviewer for the useful comment. We also see other comment (2.10) from reviewer 2 about this graph. We remove the graph because empirical tactic is removed from the paper as we mention in the comment 1.1.

1.16. Please revisit also the last paragraph in the conclusions, it is not clear to me what is meant by this.

We reorganized the last paragraph. We are suggesting the future research could conduct user group profile surveys, just like reviewer 2 suggested in the comment 2.7. Although it is clear that these trip planning strategies are widely used by passengers or trip planning apps, we still do not know the actual user profile among the passengers, i.e. how many people are using different trip

planning strategies. For example, although we know greedy tactic's performance is very bad, the amount of users who use this is still largely unknown; meanwhile, we calculate the optimal insurance buffer, but the distribution of actual insurance buffer is still largely unknown. These issues cannot be fully discussed and solved in a single paper, therefore we propose several potential topics for future RTI impact studies. We also extensively discuss this problem in the comment 2.7.

2. Reviewer 2

This paper presents an analysis of a single bus route in Columbus, Ohio to explore the impacts of different passenger trip planning strategies, including those using real-time information, on passenger wait times. Overall, I found the manuscript to contain numerous noteworthy flaws. Specifically, the analysis relies on some unusual assumptions that may be driving the results; furthermore, the scope of the analysis is limited to a single bus route in a single city, limiting the generalizability of the findings. Moreover, the authors have not validated their theory and findings with real world behavioral data, such as from surveys or focus groups. In light of these weaknesses, I recommend significant revisions to the paper. My specific comments are detailed below corresponding to the page number (since there were no line numbers in the manuscript).

2.1. Page 2, Paragraph 3

How did the authors come up with the idea that RTI apps can diminish waiting times to zero? Have the authors conducted a survey/focus groups/interviews of riders to demonstrate that this "greedy" strategy is something riders actually do? Numerous prior studies of waiting times have been cited in the literature review section of this paper, and they all included reality high average wait times (e.g., Watkins et al. found 9.23 minutes for RTI users compared to 11.21 minutes for non-users). It seems unrealistic to expect riders to minimize their wait times to zero.

Response: This is a very good question and one of the very reasons we choose to investigate the RTI's impact. We also mentioned the same issue in the comment 1.2. Specific responses:

How did the authors come up with the idea that RTI apps can diminish waiting times to zero? It seems unrealistic to expect riders to minimize their wait times to zero.

We made corresponding changes in the fourth paragraph of the introduction and section 3.3.3 (greedy tactic) to avoid confusions. It is true that RTI apps cannot *always* diminish actual waiting times to zero; we do not suggest that they can. The "zero wait time" refers to its expected waiting time (see its definition in comment 2.11), instead of actual wait time in Watkins et al. "Zero expected waiting time" means users *expect* to have zero wait time; this does not necessarily happen in every situation. If everything does not change during the user's walking and the arrival time keeps the same as the ETA when the user was scheduling the trip, the greedy tactic user's waiting time will be zero. As the reviewer pointed out, this is clearly a very ideal scenario.

As we note in a response to comments by the reviewer 1, this greedy tactic is adopted by many transit apps to calculate their suggested departure times (e.g., “Leave in X minutes.”). For example, Google Map and Transit app and open source trip planning projects (for example, OpenTripPlanner) expect the user to arrive at a stop just as the bus arrives. This can be confirmed by: 1) the suggested arrival time at the stop is always the same as the bus boarding time; 2) there is no “waiting time” or “buffer time” in a trip suggestion between the walking phase and transit phase. Therefore, there trip planning apps follow the greedy tactic.

As we shown in the paper, because the future is unpredictable (reclaimed delay, including bus accelerating and short signals) and the updated data’s high interval (discontinuity delay), greedy tactic has a high risk of missing a bus, and missing a bus is highly penalized by longer waiting time. The empirical analyses also show that the greedy tactic’s actual performance is definitely not zero, just like the reviewer commented.

Have the authors conducted a survey/focus groups/interviews of riders to demonstrate that this “greedy” strategy is something riders actually do?

Our intent is not to use the greedy tactic as an indicator of what people actually do; rather, it is benchmark corresponding to the tactic used by many transit apps to calculate their suggested leaving time. The fact that many apps, especially the most popular ones, adopted this strategy is our primary motivation to measure its performance. If apps and current popular trip planning algorithms are systematically suggesting a trip plan with very poor performance based on real-time information, it is necessary and important to measure its performance and improve it. If the user follows the suggestion of these apps, even if she/he does not realize the risk of missing the bus, the user is following the greedy tactic.

As for the issue of whether people will follow the suggested greedy tactic: it is certainly possible that some transit app users will follow the suggestion because this is the default and primary trip planning strategy provided by the apps. This is especially true for some new app users who are not aware of the high risk of greedy tactic.

For more clarification, please refer to the comment 2.7 (behavioral data) and 2.13 (greedy tactic performance) later.

2.2. Page 3, Section 2.1

The authors do not differentiate between perceived versus actual wait time differences in the literature review, which is an important distinction in prior research on the impacts of real-time information. Please add a brief discussion.

Response: This is very good point and we distinguish the two measures in the literature review. Please find the adjustments in section 2.1 and 2.2.

2.3. Page 6, Paragraph 1

The authors claim that APC data is more accurate in terms of arrival/departure time at each stop compared to GTFS-realttime. I found this surprising. How did the authors come to this finding? What analysis did you conduct to demonstrate the accuracy? Additionally, did you verify a sample of the data with real world observations (e.g., ride the bus and manually record the stop times, then compare them to APC and GTFS)? Please justify.

Response: This is another great question. Our motivation of using APC instead of GTFS is that the actual arrival time derived from GTFS has systematic errors. There are two sources of errors:

1. GTFS real-time (trip update) data is a regularly updated data, which means there is always an interval between the measured time and the actual event time. Similar to the discontinuity delay, this systematic error of measured arrival time will incur considerable error for our calculation.
2. GTFS real-time will use an algorithm to predict the arrival time at each downstream stop. GTFS trip update data will not provide the arrival time of upstream stops. Therefore, all the time provided by GTFS real-time is predicted time. In practice, we found the trip update at the latest timestamp before the actual arrival.

For example, suppose we have two stops (A and B) in a route and the bus arrives at A at 5:00:00 and at B at 5:03:30. The GTFS real-time will update every 60 seconds from 5:00:00. The time at stop A will be the right time since it is captured just in time. The measured arrival time at stop B will be the predicted time at 5:03:00, which is 30 seconds earlier than the actual time. Therefore, depending on how the system predicts the arrival time, which is not always reliable (Cats and Loutos 2016a, 2016b), the measured time could be different from the actual time. However, APC data is produced in an on-demand manner, which means the measured time is always measured when the bus actually stops and moves. Using the same example in the last paragraph, the time at the stop B will be measured and recorded at 5:03:30 instead of 5:03:00. Also, the time in the APC data is all actual time instead of predicted time. There are other factors contributing to the difference between measured and actual time; the most important one could be the precision of the sensors. However, these factors are random errors, which does not affect the unbiasedness of the measured value.

In summary, the recorded time of GTFS data may not be exactly the time when bus arrives at a stop, but recorded time of APC data is exactly the time of the bus arrival. Therefore, we introduce the concept of *temporal accuracy*. Similar to Firmani et al. (2016)'s definition, we define it as: how accurate is the measure's recorded time compared to the actual time of event occurrence (Liu and Miller 2020). It represents the systematic error caused by the temporal delay of measurement. For example, transit passenger surveys are always hours, days, or even weeks later than the time when the user actually took the transit (other than on-board surveys). Therefore, the temporal accuracy of survey methods is much lower than real-time data. Meanwhile, although APC, smart card data, and GTFS can all be categorized as real-time data, smart card and APC data's temporal accuracy can still be higher than GTFS, depending on its update frequency. Therefore, compared with GTFS with systematic errors, APC is a better data source to measure the actual arrival time because of its on-demand nature.

We refined the data introduction part in section 3.1 and added the concept of temporal accuracy.

2.4. Page 6, Paragraph 2

COTA's data is updated once per minute, which seems quite long compared to many other agencies. For example, the MBTA in Boston updates their bus location data every 5 seconds (see <https://medium.com/@sjbarbeau/introducing-the-gtfs-realtime-validator-e1aae3185439>). This is likely an important data limitation from COTA that is driving some of your results (discussed more later). Please add discussion of typical update/refresh times from other transit agencies.

Page 17, Figure 7 (we merged the two comments due to their relevance)

- Figure 7 shows high sensitivity to the 60 second update of real-time data, which, as previously noted, seems to be a reality high value (e.g., the MBTA updates every 5 seconds). Is there a way you can test the sensitivity of this in your model? At a minimum, it should be discussed as a drawback of the case study of COTA, as it appears to be driving the results shown in Figure 7 and may not apply to other agencies with better real-time data.

Response: This is a good question. We address this issue as follows. In the table below, we show the update frequency of all publicly available transit systems in the US that provide GTFS real-time feed from OpenMobilityData.org (OpenMobilityData 2020). We used the GTFS real-time validator (Center for Urban Transportation Research @ USF 2020) to measure the update frequency of each GTFS real-time feed.

Transit system	Update interval (secs)	Transit system	Update interval (secs)
MBTA	~5	Go Metro, Cincinnati	~30
Community transit	~10	DCTA, Denton, Texas	~30
CATA, Lansing, MI	10 – 20	VIA, San Antonio	~30
MST, Monterey, CA	10 – 20	HART, Tampa, FL	~30
RTC, Southern Nevada	10 – 20	LTD, Eugene, OR	~30
Votran, Daytona Beach, FL	10 – 20	Metro Transit, Madison, WI	~30
ART, Arlington, VA	20 – 30	MTA Maryland	~30
Big Blue Bus	20 – 30	RTA, riverside, CA	~30
Calgary Transit	~30	Capital metro	~60
BART	~30	CT Transit, Hartford	>60

Among 20 transit systems we could successfully test the update frequency, 12 of them have update interval larger than 30 seconds as of May 2020; in fact, MBTA is the only known transit authority that releases such high-frequency GTFS real-time trip-update data feeds in the United States. Moreover, in Google's official GTFS playbook webpage, the GTFS real-time data is

expected to have update frequency around 1 minute (Google Developers n.d.). These statistics shows that the majority of the transit systems in the United States still face non-trivial continuity delay in 2020 large than 30 seconds; the discussion of impact of long update interval larger than 30 seconds is still important for most transit systems. Moreover, these statistics were calculated in May 2020; it is very likely that many transit systems had a larger update interval back in 2018. We add the same table and clarifications to section 3.1.

Meanwhile, what users see on their smartphones are not the trip-update data, but results generated by the transit planning apps. The update frequency of the GTFS real-time data is different from transit planning apps update frequency. Several examples can be given by using the Transit app, which is a popular transit planning app with millions of users. We used a regular Android phone and an iOS tablet to randomly select several routes and different time in a day for testing purposes. We observed the information update frequency on the interface is lower than the data per se. For example, MBTA, the transit systems with the highest update frequency, can still have update intervals from 15 seconds to 1 minute for most routes shown on the actual smartphone app interface, despite theoretically the data are served every 5 seconds. We also witnessed similar phenomena in Capital Area Transportation Authority (CATA) in Lansing MI and Monterey–Salinas Transit (MST) in Monterey, California. This suggests that the actual information update can be lower than the ideal data update frequency. Meanwhile, some transit planning apps, such as Google Map, will not update the information automatically; instead, the visualized RTI will only update when the user actively refreshes the interface. This moreover suggests that the actual information update frequency can be even longer if the user does not actively update.

Despite large temporal update intervals being common for most transit authorities at present, the question raised by the referee is still important, because we will witness more transit systems with GTFS trip-update data of higher update frequency in the future. Meanwhile, as the reviewer mentioned, Figure 7 (in last draft, currently Figure 4) surely depends on the frequency of the RTI. For the mainstream update frequency of 30 seconds, the interval in the figure between each trough in the pattern will shift to 30 seconds. However, as for COTA system, we could not get data of different resolution for the same time period, therefore the potential impact of higher frequency still remain largely unknown. It could be a limitation for this paper and could be a very good topic for future studies. We also add corresponding conclusions in section 4.1 and conclusion part of the paper.

2.5. Page 6, Last Paragraph

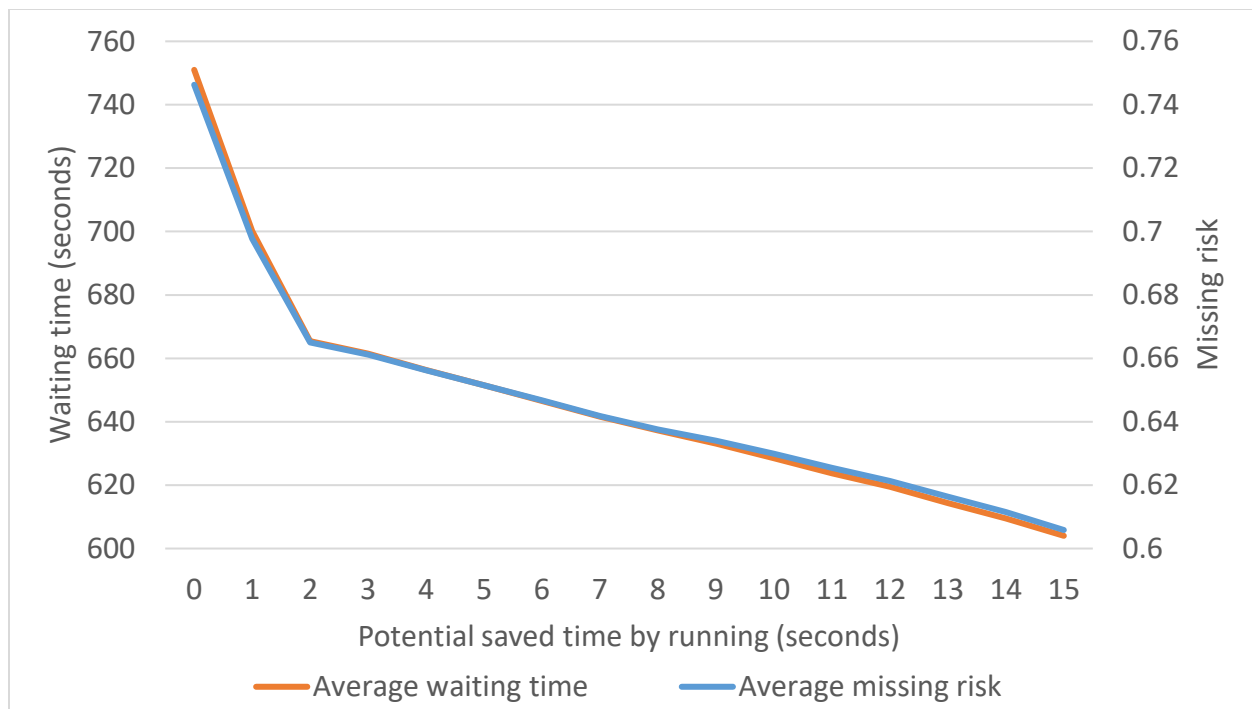
The authors make the assumption that the walking process “is linear with respect to distance.” They later explain that the only way to change the walking time is to depart the home at a different time. However, the assumption of constant walking speed seems highly unrealistic. If a rider sees a bus approaching and they think they might miss it, they are very likely to speed up and potentially even run to meet the bus. Indeed, a study by Dziekan and Kottenhoff (2007) of the subway in Stockholm, Sweden observed passengers entering subway stations and counted the number of passengers running and walking when RTI signage outside the subway station was on,

and the results reveal that significantly more people run when the RTI signage was on rather than when the signage was off. In light of this real world evidence, the assumption of constant walking speed seems very unrealistic, and the authors should test the sensitivity of their results to altering this assumption.

Response: There are several reasons that we chose to use linear walking process and insurance buffer ahead as the primary assumption. One reason is that running is not a viable and desirable option for everyone. Running after seeing the bus approaching can be possible for some younger people, but it is not viable for other passengers such as senior people, disabled people, people with luggage and children. And the obstruction of buildings and trees, the condition of pavements, and the weather all make running a less likely option for most people. Even if a passenger can run to the stop, compared with planning an insurance buffer ahead, passengers may find it exhausting, stressful, and thus undesirable. A linear assumption is more conservative, and inclusive, than assuming passengers can run. In addition, it is not the default assumption for transit planning apps. Because running is not a viable option for many people, the transit planning apps' planning logic do not assume or suggest this behavior.

However, the referee raises a good challenge to our assumption that deserves testing. We conducted a sensitivity analysis and conclude that the linear walking assumption does not affect the results. We first determine a good range of sensitivity buffer to relax the walking time, i.e. how many times the user can save the walking time by running. However, this saved time cannot be too long for two reasons: first, humans cannot run significantly faster than the usual walking speed; and second, the average sprint distance cannot be very long, since the obstruction of buildings and trees usually won't allow the passenger to see a bus at a very long distance, including distinguishing the route number. We use the average speed of 3 m/s as the running speed (Chertoff 2018). We consulted the letter height visibility table (James 2017); for a route number on a regular US bus of 20 cm (7.8 inches), the ideal readable distance is about 21 (70 feet) – 24 (80 feet) meters. We will use 24 meters as the sprint distance. In our study, we use 1.4 m/s the standard walking speed for all the calculation. Therefore, the maximum saved time is $24/(3-1.4) = 15$ seconds. We conducted the sensitivity test based on the extra 0 - 15 seconds during the walking phase. For example, if the saved time by running is 5 seconds, this means that the user will arrive 5 seconds earlier than a normal greedy tactic user at the same stop.

Picture 1 shows the waiting time and missing risk with different potential saved time by running from 0 – 15 seconds. For the maximum potential saved time by running of 15 seconds, the average waiting time is 604 seconds and the missing risk is 60.58%. It shows that even considering the maximum possible saved time by running, the greedy tactic's performance is still bad; in fact, it still has the highest average waiting time and highest average missing risk among all other trip planning strategies in the Table 2 in the paper.



Picture 1: waiting time and missing risk with respect to different potential saved time by running from 0 – 15 seconds.

Therefore, we can conclude that not running is not the cause of the bad performance of the greedy tactic. Instead, its bad performance results from: 1) the optimal scenario (waiting time = 0) is a fragile balance that even a small time shift can make the user miss the bus and wait a very long time; 2) the reclaimed delay and discontinuity delay, as we extensively discussed in our responses and the paper. We added the corresponding conclusions to the first paragraph of section 3.3 in the paper.

2.6. Similar to the previous comment, another recent study by Ferris et al. (2010) found that RTI provided on mobile devices may impact a passenger's decision of where to board the transit vehicle, which would impact the passenger's walking distance to access transit. On a survey of RTI users conducted in Seattle, Washington, 78% of respondents reported they were more likely to walk to a different stop based on RTI (Ferris et al., 2010). Can the authors test this (e.g., changing where to board) in their modelling framework?

Response: The impact of RTI on passenger's path choice is a very important. In this answer, we are going to conceptualize and simulate the decision-making process and investigate if this strategy is rational.

First, we will discuss about the motivation of walking to another stop. A major reason of changing stop is that by walking to another "downstream" stop, the passenger can arrive earlier by taking advantage of the bus running time between the two stops as shown in Picture 1. For

example, a bus will arrive at stop A at 5:00 pm and stop B at 5:05pm, and passenger will leave the home at 4:55am but it will take the user 6 minutes to walk to stop A. Therefore, if the passenger sticks to stop A, she/he will arrive at 5:06pm thus miss the bus. However, the passenger knows that it takes she/he 8 minutes to walk to stop B and when the passenger arrives at B the time will be 5:03pm, thus the bus won't be missed. Therefore, if the user can outperform the bus by going to an alternative stop, then it is possible to save time by catching the bus which should have been missed at the original stop.

However, we need to compare the walking time difference and bus running time difference to validate if this strategy can indeed save time. We first give the definition of the walking time difference:

$$\Delta t = \frac{x_a - x_o}{v} \quad (1)$$

Where x_a is the distance between home and alternative stop, x_o is the distance between home and original stop, v is the walking speed (1.4 m/s). Δt represents the time cost caused by the route changing. Meanwhile, the saved bus running time ΔT is:

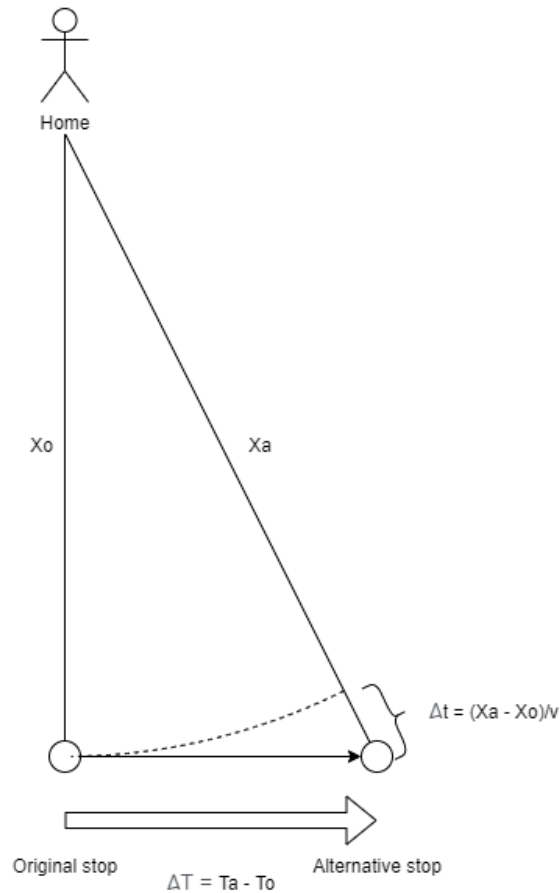
$$\Delta T = T_a - T_o \quad (2)$$

Where T_a is the arrival time at the alternative stop and T_o is the arrival time at the original stop. This will be derived from the history arrival time in the GTFS-APC data. We derive the potential saved time:

$$\delta t = \Delta T - \Delta t = T_a - T_o - \frac{x_a - x_o}{v} \quad (3)$$

If the potential saved time is positive, this says this strategy will save time for the passenger. For the same example above, the saved bus running time ΔT is 5 minutes and walking time difference Δt is 2 minutes, therefore the potential saved time δt is 3 minutes and the user will save time.





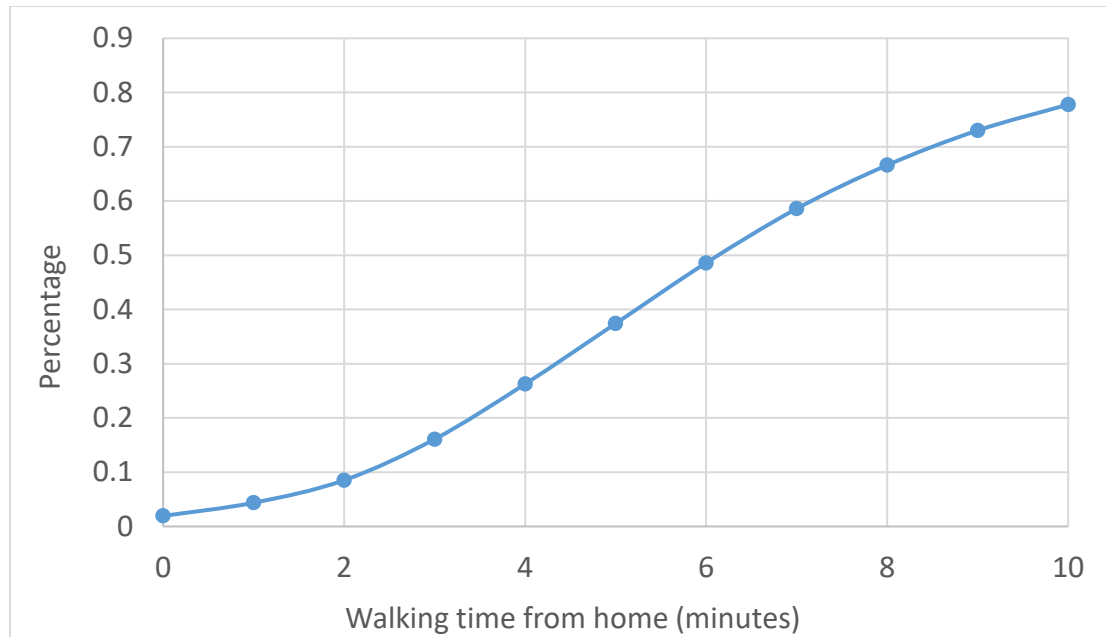
Picture 2: presentation of “walking to another stop” strategy.

Therefore, an important topic about the question raised by the comment is: how often can this strategy save user time in a public system? If the empirical data can prove that this strategy can instead save time for passengers, it can justify and validate the behavioral changes reported by prior research. Therefore, we calculate the potential saved time for each stop and its subsequent downstream stop and summarize the ratio of positive potential saved time. We will test the start locations on the perpendicular line from the original stop with different walking time, from 0 to 10 minutes.

For two subsequent arbitrary stops on a route, we will only consider the positions between the stops. For example, in Picture 2, we will not consider the points to the left of original stop or the points to the right of alternative stop, since passengers who live at these points will have an even closer stop in the upstream or downstream and current two stops will not be the closest two stops.

Picture 3 shows the ratio of positive potential saved time for different walking time from home. It shows the effectiveness of this strategy highly depends on the distance from the home. If the passenger lives very close to the stop, she/he does not want to go to another stop because people’s walking speed is much slower than the bus. Meanwhile, if the user lives further from the stop, the changing stop strategy is more effective therefore more people will be prone to do

that. This can be a very good supportive proof for the conclusion mentioned by the reviewer based on empirical data.



Picture 3: ratio of positive potential saved time for different walking time from home.

2.7. Page 9, Table 1

How did the authors arrive at these 5 trip planning strategies? Similar to my previous comment, have the authors conducted a survey/focus groups/interviews of riders to demonstrate that these are strategies riders actually use?

Response:

Non-real-time trip planning strategy (Schedule, arbitrary, and empirical tactic): Schedule tactic is a default trip planning strategy: users checking a schedule and walking to a stop accordingly. We also added some explanations and reference in section 3.3.2 in the paper. Arbitrarily departure is also a reasonable and common trip planning strategy, if not the most common. Prior studies have extensively proven that arrival time follows a uniform random distribution when the headway is small (Bowman and Turnquist 1981); the referee also brought up the same conclusion in the comment 2.8. We added more references in section 3.3.1 to emphasize this statement. As we mentioned in comment 1.1, we removed empirical tactic from the paper.

But most importantly, as

Greedy tactic: As our responses to the comment 2.1 point out, greedy tactic is the default strategy adopted by different trip planning apps and algorithms; it is not our invention. It is quite plausible that users follow it, particularly since popular trip planning apps such as Transit and Google Maps give specific departure instructions based on this tactic. Perhaps more importantly, the value of investigating the greedy tactic does not lay on whether or how many people use it. The fact that many apps, especially the most popular ones, adopted this strategy is our primary motivation to measure its performance. This is the question we ask in our title: does real-time information (and real-time information apps) reduce waiting time? If apps and current popular trip planning algorithms are systematically suggesting a trip plan with very poor performance based on real-time information, it is valuable to measure its performance, including how bad it is and how bad it can be compared to other traditionally used non-real-time trip planning strategies. Moreover, it would be valuable to find a good way to improve the algorithm. All of these will be based on the investigation of the greedy tactic. We also added these discussions in section 3.3.3.

Prudent tactic: Similar to greedy tactic, there are some facts supporting the actual usage of prudent tactic. We will justify our endeavor to investigate prudent tactic in following aspects:

First, users can actively use prudent tactic. As we mentioned in the section 3.3.4 (prudent tactic), it is a common strategy to avoid risk of missing a bus by leaving earlier than the suggested time (Furth and Muller 2006). This is especially true for two kinds of people: the experienced users and users who are less familiar with the transit system, such as tourists and new residents. Experienced users will notice the unreliability of the suggested time thus leave earlier. New users will feel more insecure thus leave earlier than the expected time to avoid risk. This process can be perceived as a very primitive and coarse prudent tactic; and it is also why we should systematically discuss a more advanced version with well-optimized insurance buffers for specific time and space.

Finally, just like greedy static, the necessity of investigating the performance of the prudent tactic does not lay on whether or how many people use it. We propose prudent tactic as an innovative solution that can be used to improve real-time transit information apps. It is certain that the optimal prudent tactic is not used by anyone right now. However, it could be used if developers adopt our solution or develop better solutions to solve the insurance buffer and apply them to their routing and departure instructions.

In summary, schedule, arbitrary, and greedy tactic are indeed used by at least some passengers. For certain variety of prudent tactic and the prudent tactic optimal, we are sure that no one are using it right now and this is the very reason we are proposing an innovative algorithm: people will benefit from using it as it is adopted by more apps thus more passengers. On the other hand, it can be a very good potential for future studies to expand the topic to the inductive perspective as we discuss in the last paragraph of the conclusion section. This is because the quantitative composition of users' trip planning strategies is still largely unknown, even if we can infer that people are using them. We added this to the last paragraph of the conclusion section as a potential future topic. We thank the reviewer again for providing a useful insight.

2.8. Page 10, Equation 3

- It should be noted that the average waiting time formula for random arrivals is generally only applied to high frequency transit routes (e.g., headways less than 10-15 minutes).

Response: This is very true for formula 6 (last draft). We only provided formula 6 to introduce the traditional calculation method and we did not use headway and its standard variance in our calculation. This is one of motivation to use the high-resolution data without major stochastic assumptions. The actual formula we were using is formula 7 (last draft, formula 2 in current draft): we will find the bus before the targeted bus and produce the average of the two buses' time. Therefore, in this draft, we delete the former formula 6 to avoid further confusions.

2.9. Page 3, Equation 5

- The authors state that the “bus will rarely if ever leave a stop earlier than the scheduled time.” Did the authors verify this statement empirically, such as comparing the GTFS schedule to GTFS-realtime? Drivers occasionally do run “hot.” Please justify the assumption that they don't.

Response: Thanks for pointing this out. We did not assume that buses never run “hot” in our calculation. The best answer to this question must be that we present a risk of missing bus for the schedule tactic. The missing risk of schedule tactic represents the chance for buses to leave earlier than the schedule, which is exactly the scenario of “running hot”. From table 2 in the main text, we can see that the mean of schedule tactic's risk of missing bus is 6.28%. If we assumed no buses would run “hot”, the number will be 0.

As for the authenticity of this claim, the risk of missing bus is the lowest among other strategies; and 6% is a rather low absolute value. Moreover, in COTA's official on-time performance policy, buses are not permitted to run ahead of schedule for normal buses (COTA 2019). Even if a bus leave 1 second earlier than the schedule, the bus is not considered “on-time”. This is also a common policy for many other transit agencies (Bradley 2005; GCRTA 2018). However, we also know that some transit authorities allow running “hot” policy, such as WMATA in Seattle (less than 2 minutes earlier than the schedule).

However, this fact does not affect the authenticity of the proposed methods or conclusions since we did not assume that buses never run hot in our calculation as we explain in the last paragraph. We also discuss the issue in comment 1.9 about the transferability of this conclusion.

2.10. Page 11, Figure 3

- What data was used to create the visualization shown in Figure 3?

Response: We implemented the empirical tactic of different memory and calculated the performance based on the same GTFS-APC data we collected. We did not use an external data source or reference. However, because we remove empirical tactic completely from the paper, the figure 3 is also removed.

2.11. Page 13, Equation 9

- The authors state that “for PT family, insurance buffer should be at least equal to the expected waiting time.” Please explain why this is the case. It wasn’t clear to me.

Response: Per reviewer 1’s suggestion, we delete this part in the main because it is less relevant to the conclusion of this paper.

For prudent and greedy tactic, the users will plan their home departure time according to the bus’s estimated time of departure, instead of the real departure time. Here we define expected waiting time similar to Cats and Loutos (2016a):

$$w^p(\tau) = \pi^p(\tau) - \tau = \pi^p(t + \delta t) - (t + \delta t) \quad (4)$$

where: $\pi^p(\tau)$ is the estimated time of departure, which depends on the expected arrival time, τ is the user’s arrival time at the stop, t is PT strategy’s home departure time, and δt is the walking time. Then, consider the home departure time of prudent tactic (also in Equation (5) in the main text):

$$\begin{aligned} &\textbf{while} \text{ there is a new update } \textbf{do:} \\ &\quad \textbf{if } t' + \delta t + IB \geq \pi^p \\ &\quad \quad \text{return } t = t' \\ &\quad \textbf{else} \\ &\quad \quad \text{wait until next update} \end{aligned} \quad (5)$$

Therefore:

$$t + \delta t + IB \geq \pi^p \quad (6)$$

Combine it with Equation (4) and substitute π^p with expected waiting time w^p plus user’s arrival time at the stop τ :

$$t + \delta t + IB \geq \pi^p = w^p + \tau = w^p + t + \delta t \quad (7)$$

We have: $IB \geq w^p$. Therefore, for PT family, insurance buffer should be at least equal to the expected waiting time.

2.12. Page 14, Last 2 Paragraphs

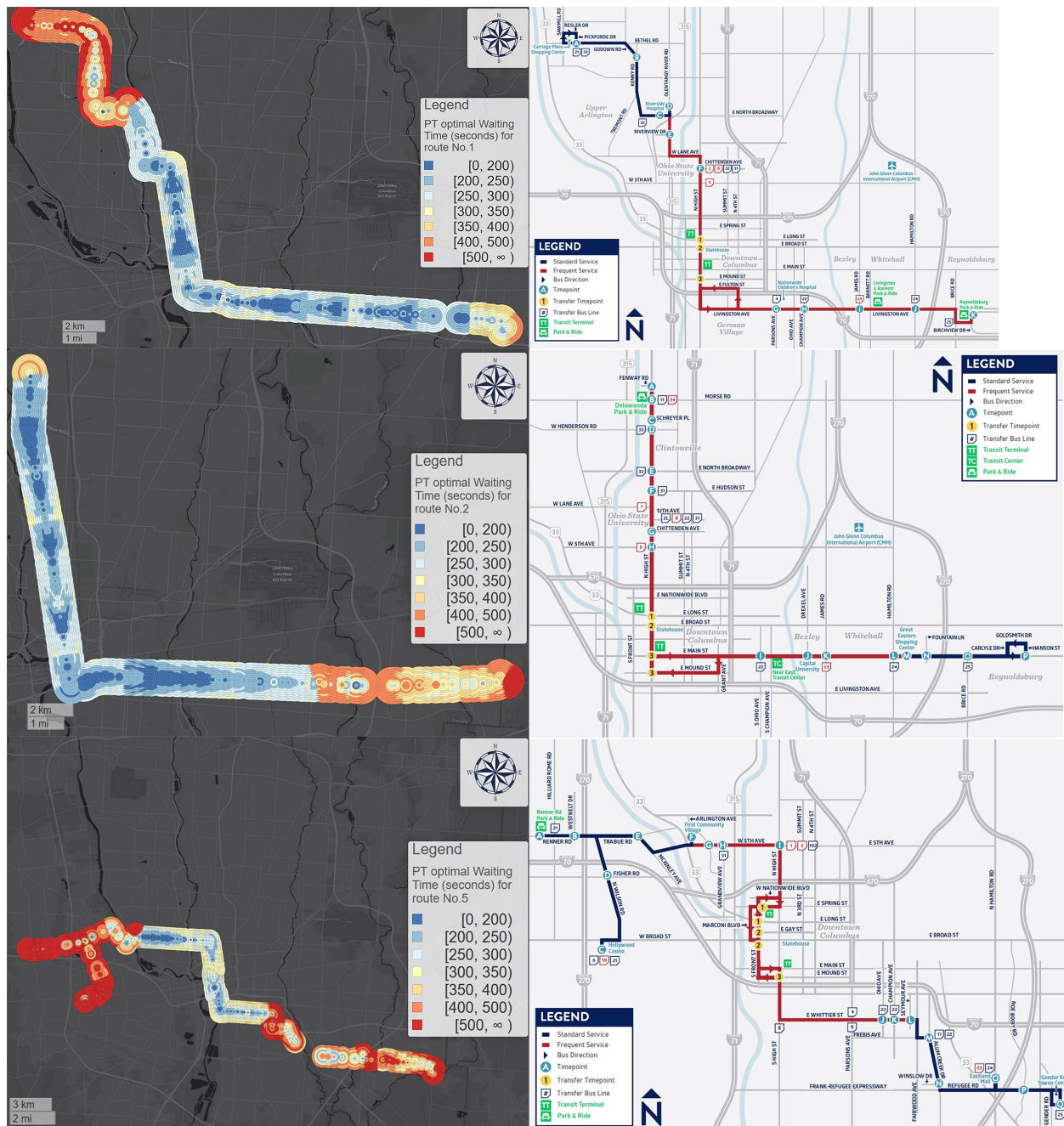
- The authors state that there is a large computational burden to conduct the analysis, so they only selected one bus route for the analysis. This greatly limits the generalizability of the research. Instead of using 1 year of data, why not use 1 week of data and run the analysis for multiple bus routes? I strongly encourage the authors to consider a larger geographic sample.

Response: Thanks for mentioning this. This is a very important question about the representative of route No.2. The core question is: will the selected bus route 2 limit the generalizability of the research?

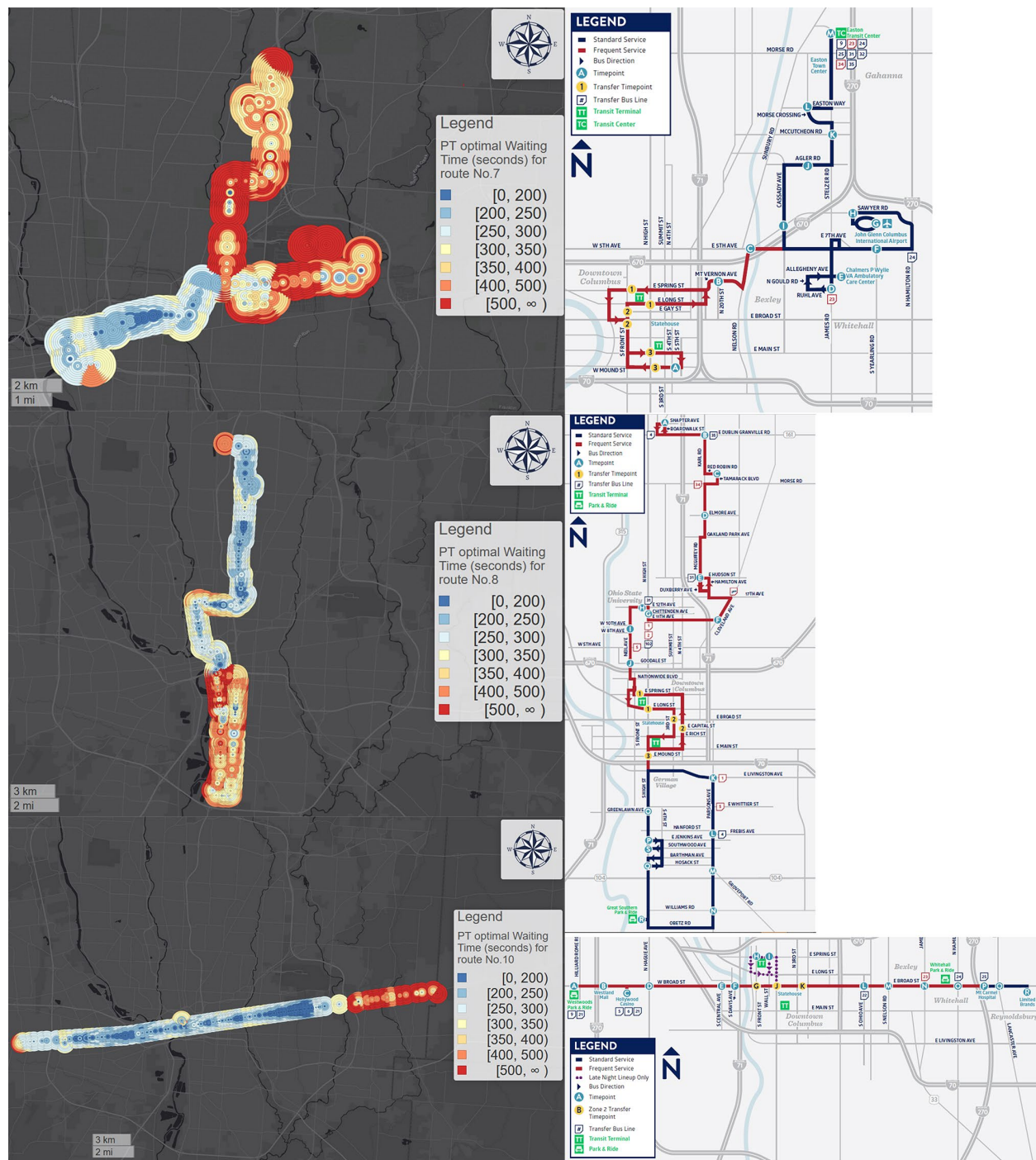
In this response, we select 6 representative major routes: COTA route 1, 2, 5, 7, 8, and 10. All of these routes have different directions, different and wide spatial and temporal coverage, and large ridership share in the City of Columbus. Just as the reviewer suggested, we select the time period from a typical week from 7/15/2018 – 7/21/2018, when there was no major event like football games and extreme weather.

We reproduced the optimization of prudent tactic for each route, which is the core contribution of the paper as reviewer 1 pointed out and have the most heterogeneous spatial pattern. Picture 4 and Picture 5 show the six routes' prudent tactic optimal's waiting time. We can see the spatial distribution of the waiting time is highly similar for different routes: the standard service sections (blue part in the service maps to the right) have higher waiting time while the sections with frequent services (red part in the service maps) have lower waiting time. All the maps have same increasing pattern as the walking time (radius of the concentric circles) increases, as we extensively discussed this in the section of walking time impact. This similarity is not limited to the distribution, but also the specific values of the waiting time. Note the maps have the same color scheme.

Therefore, we can conclude that bus route No.2 is a good sample of the whole COTA bus system and selecting it as the research target will not compromise the generalizability of the research. Moreover, analyses based on longer time period is more generalizable and provide more robust conclusions. Per the reviewer's comment, we add corresponding clarifications in the section of 3.3.4 and 4.3.2. We also add the Picture 4 and Picture 5 to the appendix for readers' reference, since the representativeness of the selected route is an important and common question.



Picture 4: Route 1, 2, and 5's PT optimal's waiting time and route map.



Picture 5: Route 7, 8, and 10's PT optimal's waiting time and route map.

- I found the GT results counterintuitive, and I would have expected them to be similar to the PT results. I suspect this is likely due to the assumptions built into your models (e.g., constant walking speed so riders won't run for an approaching bus and the long 60-sec threshold of real time updates).

Response: Yes, it is surely counterintuitive. However, we show that this counterintuitive phenomenon is justifiable and the conclusion can be one of the major contributions of this paper as we explain in our response to comment 2.1.

As for the two major reasons that the reviewer suspected to be the primary reason of the bad performance of greedy tactic, we make corresponding clarifications in our responses to comment 2.4 and 2.5. First, we explain that 60% of US transit system still have non-trivial update interval larger than 30 seconds, and MTBA is actually the only transit system that provides such as high frequency in US in our responses to the comment 2.4. Second, we also explain why running is not possible for many people and should not be a default assumption for trip planning apps in our responses to comment 2.5. The sensitivity test also confirms that running cannot solve the bad performance issue of greedy tactic. The actual reasons are: 1) the uncertainty of the buses' performance measured by the *reclaimed delay*, and 2), just like the reviewer pointed out, the low update frequency measured by the *discontinuity delay*. We extensively discuss these in the paper. As a naïve trip planning strategy, greedy tactic does not consider the risk of missing bus caused by the two sources of over-estimation, therefore it is actually reasonable why its performance is bad.

On the other hand, the poor performance of greedy tactic does not mean that real-time information is useless. This is another contribution of this paper: we summarize the definition of prudent tactic and insurance buffer and introduce a systematic and sustainable method to optimize the insurance buffer. By using a better and adaptive trip planning strategy which combined the advantages of real-time data and empirical analyses, we can make real-time information more useful for passengers in the future. We also made corresponding adjustments in the paper, such as the last sentence of the third paragraph of the section 1.

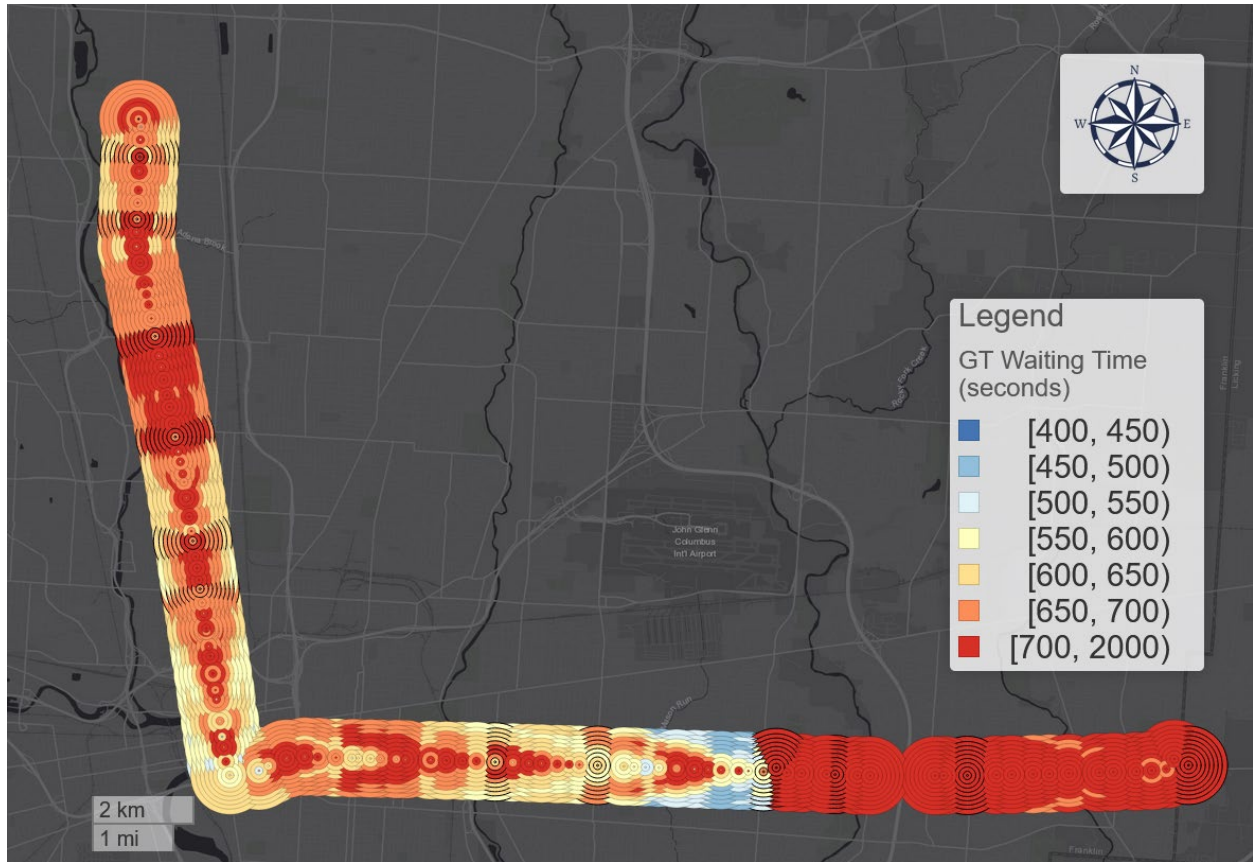
2.14. Page 22, Figure 13

- For the route shown in Figure 13, how many timepoints are there along the route? Where are the timepoints located? How do the timepoints relate to your findings (in this figure and the subsequent maps of the route)? Please add a discussion of timepoints.

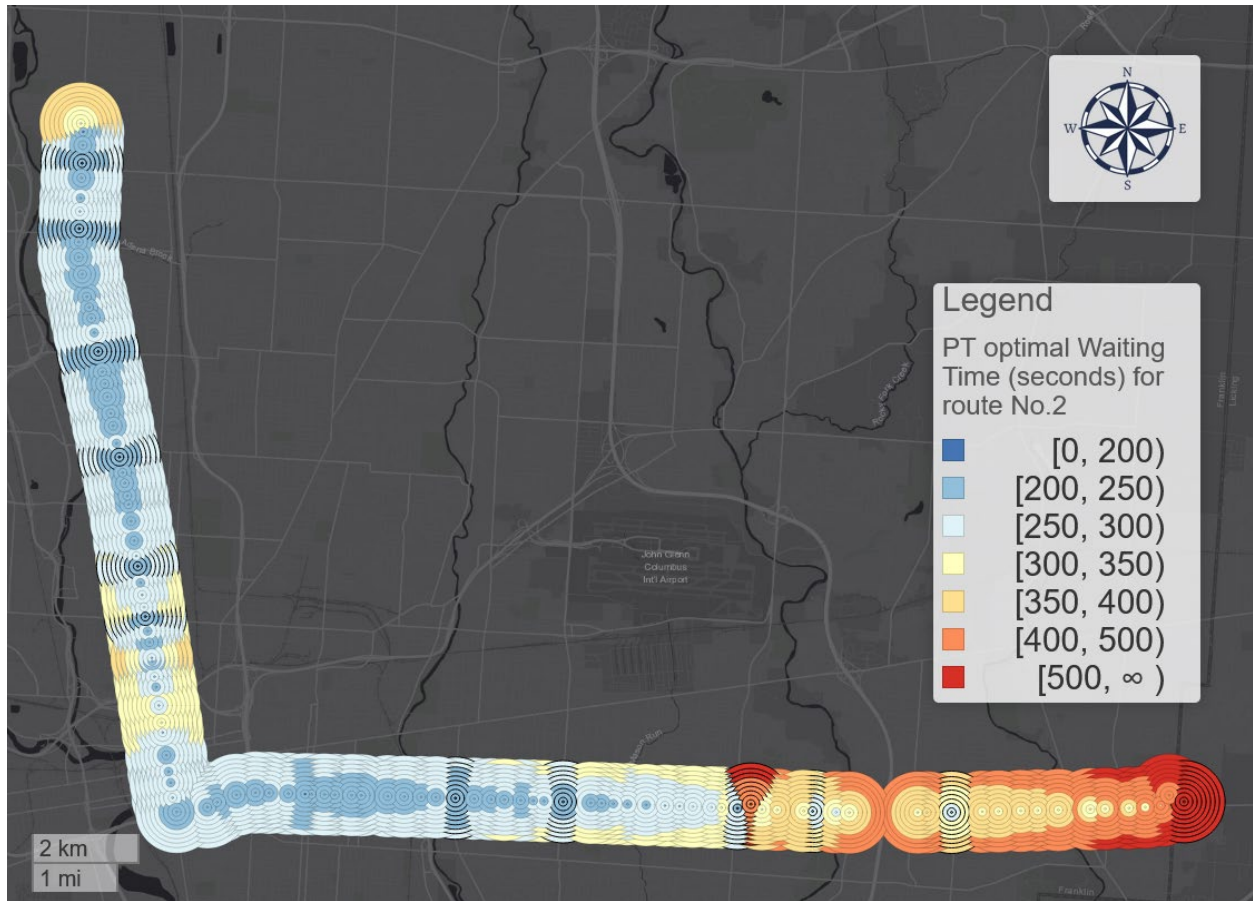
Response: This is a very important issue and we thank the reviewer for pointing it out. We add the analysis and discussion about timepoints in section 4.3.2. Picture 6, Picture 7, and Picture 8 shows the location of timepoints in the corresponding spatial distribution of greedy tactic, prudent tactic optimal, and schedule tactic's waiting time. We can draw three conclusions from the results:

1. Picture 6 shows that the waiting time of greedy tactic at timepoint stops is significantly larger than nearby non-timepoint stops. This can be explained as the result of strict timetable policy: a bus should stick to the timetable strictly at these stops. Therefore, the bus drivers tend to reclaim delay before these stops therefore making the risk of missing bus larger.
2. Picture 7 shows that the waiting time of prudent tactic optimal at timepoint stops is significantly smaller than nearby non-timepoint stops. This proves the effectiveness of the optimal insurance buffer.
3. Picture 8 shows similar spatial pattern for schedule tactic at the timepoints. The primary reason is also because of stricter timetable policy at timepoints. We decide not to put this conclusion in the main paper because it is less relevant.

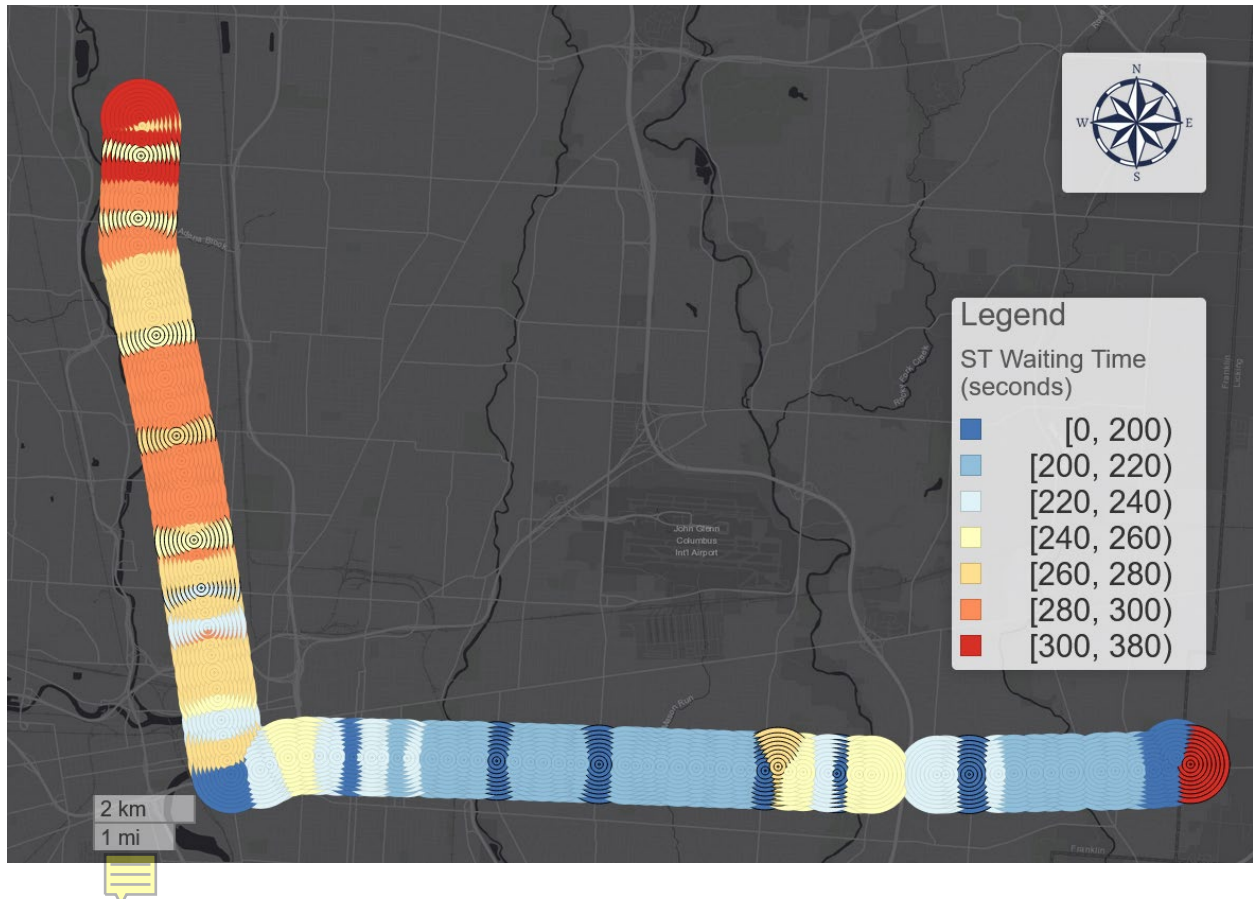
We thank the reviewer again for pointing out the importance of timepoints. It surely adds many valuable conclusions to the paper.



Picture 6: greedy tactic's waiting time for bus route No.2 (circles with black stroke: timepoints).



Picture 7: prudent tactic optimal's waiting time for bus route No.2 (circles with black stroke: timepoints).



Picture 8: schedule tactic optimal's waiting time for bus route No.2 (circles with black stroke: timepoints).

2.15.

- For Figure 13, please make sure the colors in the legends correspond to the same numerical values. It is difficult to compare GT with ST, AT and ET since the ranges for the colors are different. This comment also applies to subsequent figures.

Response: This point is now moot since we deleted the Figure 13 (last draft) from the paper for length reduction as indicated in comment 1.1.

Minor Comments

2.16. On page 9, please fix the reference errors to the figures/tables.

Response: We apologize for the mistake. This is also mentioned by the reviewer 1 in comment 1.13. We make sure all references are correct in this draft.

2.17. The authors introduce many new acronyms throughout the paper, which can be confusing for readers. Please try to limit the use of acronyms that are not commonly found in the prior literature (e.g., consider removing IB, TPS, HDT, etc.).

Response: We remove all the mentioned acronyms in the comments. We also remove some other acronyms that do not appear very frequently in the paper, like ETD, ATD, and RD (reclaimed delay). We keep ST (schedule tactic), AT (arbitrary tactic), ET (empirical tactic), PT (prudent tactic), and GT (greedy tactic) in the paper because they are frequently mentioned and used in different parts of the paper (in total 104 times). To avoid confusions for readers, we use the corresponding full name at the first appearance in each section.

Reference:

- Bowman, Larry A., and Mark A. Turnquist. 1981. "Service Frequency, Schedule Reliability and Passenger Wait Times at Transit Stops." *Transportation Research Part A: General* 15(6): 465–71.
- Bradley, Roosevelt. 2005. "Miami-Dade Transit Service Standards." [https://www.miamidade.gov/citt/library/strategic-financial-studies/2013/cost-other-studies/analysis-activity/2006-1ServiceStandardPresentation \[Compatibility Mode\].pdf](https://www.miamidade.gov/citt/library/strategic-financial-studies/2013/cost-other-studies/analysis-activity/2006-1ServiceStandardPresentation%20Compatibility%20Mode.pdf) (June 11, 2020).
- Cats, Oded. 2019. "Determinants of Bus Riding Time Deviations: Relationship between Driving Patterns and Transit Performance." *Journal of Transportation Engineering, Part A: Systems* 145(1): 4018078.
- Cats, Oded, and Gerasimos Loutos. 2016a. "Evaluating the Added-Value of Online Bus Arrival Prediction Schemes." *Transportation Research Part A: Policy and Practice* 86: 35–55. <http://www.sciencedirect.com/science/article/pii/S0965856415300124>.
- . 2016b. "Real-Time Bus Arrival Information System: An Empirical Evaluation." *Journal of Intelligent Transportation Systems* 20(2): 138–51.
- Center for Urban Transportation Research @ USF. 2020. "Gtfs-Realtime-Validator." <https://github.com/CUTR-at-USF/gtfs-realtime-validator> (May 18, 2020).
- Chen, Xumei, Lei Yu, Yushi Zhang, and Jifu Guo. 2009. "Analyzing Urban Bus Service Reliability at the Stop, Route, and Network Levels." *Transportation research part A: policy and practice* 43(8): 722–34.
- Chertoff, Jane. 2018. "What's the Average Running Speed and Can You Improve Your Pace?" <https://www.healthline.com/health/fitness-exercise/average-running-speed> (May 18, 2020).
- COTA. 2019. "How Does COTA Measure On-Time Performance?" <https://www.cota.com/policies/on-time-performance/> (February 5, 2019).
- El-Geneidy, Ahmed M, James G Strathman, Thomas J Kimpel, and David T Crout. 2006. "Effects of Bus Stop Consolidation on Passenger Activity and Transit Operations." *Transportation Research Record* 1971(1): 32–41.
- El-Geneidy, Ahmed M, Jessica Horning, and Kevin J Krizek. 2011. "Analyzing Transit Service Reliability Using Detailed Data from Automatic Vehicular Locator Systems." *Journal of Advanced Transportation* 45(1): 66–79.
- Firmani, Donatella, Massimo Mecella, Monica Scannapieco, and Carlo Batini. 2016. "On the

- Meaningfulness of ‘Big Data Quality.’” *Data Science and Engineering* 1(1): 6–20.
- Furth, Peter G, and Theo H J Muller. 2006. “Service Reliability and Hidden Waiting Time: Insights from Automatic Vehicle Location Data.” *Transportation Research Record* 1955(1): 79–87.
- GCRTA. 2018. “QUARTERLY MANAGEMENT REPORT.”
<http://www.riderta.com/sites/default/files/qmr/QMR2018Q2.pdf> (June 11, 2020).
- Google Developers. “GTFS Playbook: Using Static with Realtime.”
<https://support.google.com/transitpartners/answer/9047609?hl=en> (May 9, 2020).
- Huo, Yueying, Jinhua Zhao, Wenquan Li, and Xiaojian Hu. 2014. “Measuring Bus Service Reliability: An Example of Bus Rapid Transit in Changzhou.” *Journal of Public Transportation* 17(2): 6.
- James, Nelson. 2017. “Signage 101 – Letter Height Visibility.”
<https://www.signs.com/blog/signage-101-letter-height-visibility/> (May 18, 2020).
- Liu, Gang, Ling Shi, and Tony Z Qiu. 2016. *Evaluation of Factors Affecting Bus On-Time Performance in Edmonton, Canada*.
- Liu, Luyu, and Harvey J Miller. 2020. “Measuring Risk of Missing Transfers in Public Transit Systems Using High-Resolution Schedule and Real-Time Bus Location Data.” *Urban Studies*: 0042098020919323. <https://doi.org/10.1177/0042098020919323>.
- OpenMobilityData. 2020. “OpenMobilityData GTFS-RT.”
<https://openmobilitydata.org/search?q=gtfsrt>.
- Park, Yongha et al. 2019. “Assessing Public Transit Performance Using Real-Time Data: Spatiotemporal Patterns of Bus Operation Delays in Columbus, Ohio, USA.” *International Journal of Geographical Information Science*: 1–26.