Motivation

- Tracking a user function over distributed data in online fashion is a fundamental challenge.

![Sensor network](a.png) ![Location based service](b.png)

Background: two-party online tracking

- \( f : \mathbb{Z} \rightarrow \mathbb{Z} \), \( g(t) \in [f(t) - \Delta, f(t) + \Delta] \)
- Naive method: unbounded competitive ratio;
- OptTrack: \( O(\log \Delta) \) competitive ratio.

![Naive method](a.png) ![OptTrack method](b.png)

Chain online tracking

- ChainTrackA: distribute \( \Delta \) averagely.
- ChainTrackO: distribute \( \Delta \) randomly among \( h + 1 \) centralized instances.
- ChainTrackO: assign the whole tracking error \( \Delta \) to the first tracking instance. It achieves \( O(\log \Delta) \) competitive ratio.

Distributed setting

- Each observer \( s_i \) observes an arbitrary function \( f_i \) over time.
- \( T \) wants to keep tracking of \( f(t) = f_1(t), f_2(t), \ldots, f_m(t) \) using \( g(t) \) within an error \( \Delta \) for any time instance \( t \).

Broom online tracking: max

- \( f(t) = \max(f_1(t), \ldots, f_m(t)) \)
- \( g(t) \in [f(t) - \Delta, f(t) + \Delta] \)

![BroomTrack](a.png) ![ChainA](b.png)

Theorem 1: For any algorithm \( A \) in \( A_{broom} \), there exists an input instance \( I \) and another algorithm \( A' \) in \( A_{broom} \), such that cost\((A, I)\) is at least \( h \) times worse than cost\((A', I)\), i.e., for any \( A \in A_{broom} \), ratio\((A)\) = \( \Omega(h) \).

Theorem 2: With respect to online algorithms in \( A_{broom} \), ratio\((BroomTrack) < h \log \Delta \).

Corollary 1: ratio\((m-Chain) = O(h \log \Delta) \).

General tree online tracking: max

- Corollary 2: There is no instance optimal algorithm for \( A_{tree} \).
- Corollary 3: ratio\((TreeTrack) = O(h_{max} \log \Delta) \) with respect to \( A_{tree} \).

Other topologies and functions

- Other functions: min, sum, average and \( \phi \)-quantile

Experiment

- Temperature (TEMP) from MesoWest Project.
- Wind Direction (WD) from SAMOS project.

![Graph topology](a.png)

![BroomTrack vs Chain](a.png)

Conclusion

- Extend the centralized, two party model to the chain model.
- Investigate both the broom model and the tree model, as well as other different tracking functions.