Commuter Rail Real-Time Information

Commuter Rail Background

MBTA operates Commuter Rail services on 12 lines serving over 120 stations in the Boston metropolitan area. MBTA commuter rail customers are provided with real-time arrival information through LED signs located on station platforms via the MBTA's Passenger Train Information System (PTIS). The purpose of this data feed is to share the real-time information on train locations and schedule adherence with developers, who can build applications that will benefit Commuter Rail riders.

Important Disclaimer

The MBTA could wait until everything about this data was perfect before releasing it, but then no one would be able to benefit from it in the meantime. So the MBTA is releasing this data now, even knowing that it does have some problems. Some trips are not tracked some of the time; under certain condition a trip can be listed twice; vehicles' lateness sometimes fluctuates from one minute to the next; and there may be times when the feed is unavailable. Customers should be advised to use all deliberate caution when using this data, and to do so at their own risk. It's one thing to miss a Red Line train and have to wait five minutes for the next one, but it's quite another to miss a Commuter Rail train when the next one might not be coming for over an hour. However, handled and framed properly, the data in this feed can make Commuter Rail easier and more appealing to ride.

Presentation and Nomenclature

The word "train" can refer to a physical grouping of coaches pulled by a locomotive ("this train has a lot of seats") or to a scheduled trip that runs once a day ("take train 921 from South Station.") To avoid confusion this document uses "vehicle" or "trainset" for the former concept, and "trip" for the latter. Nonetheless when communicating with customers "train 921" is correct and "trip 921" is not.

"Lateness" is included in the data feed measured in seconds, but a normal presentation to customers would convert this value to minutes. Because lateness can fluctuate a little bit from one minute to the next, developers are encouraged to describe anything less than five minutes late as "on time," to prevent a customer from thinking they can arrive two minutes late when that may not turn out to be the case.

Data Organization

There is one source of data for each commuter rail line. The source is available in three formats: CSV, XML, and JSON. URLs can be found on the project web page.

Each "row" of the file represents one trip's scheduled stop at one station, or "trip-stop." A trip-stop will appear in the file 1-2 hours before it is scheduled to occur, and will be removed one or two minutes after it has occurred. If no real-time information is available about the trip-stop then it will not appear at all. Schedule information will appear for all upcoming trip-stops; the current location and lateness of the vehicle are included when known.

There are several trips that operate partly on one line and partly on another. These trips and all their trip-stops are included only in one line's file. Specifically, trips operating on parts of the Franklin Line and the Fairmount Line (via the "Dorchester Branch") are considered Franklin Line trips, and trips operating on parts of the Haverhill Line and the Lowell Line (via the "Wildcat Branch") are considered Haverhill Line trips.

Data Feed

The output files in all formats will contain the following fields of information.

Field	Type	Example	Description
Timestamp	Integer	1294070497	Record timestamp in "epoch time" (seconds since midnight 1/1/1970 GMT)
Trip	String	2718	Identifies the trip, corresponding to the 'Train Number' in the public timetables. In GTFS as a substring. Do not discard leading zeros.
Destination	Integer	Lowell	The station ID of the trip's destination. Matches the MBTA's GTFS feed and is also human-readable.
Stop	String	Wedgemere	The station ID of this trip-stop. Matches the MBTA's GTFS feed and is also human-readable.
Scheduled	Integer	1294070497	The time the trip is scheduled to arrive at or leave the station, in "epoch time"
Flag	String	Pre	Flag characterizing record. See below.
Vehicle	Integer	1505	Unique number identifying the trainset's control cab (physically painted on) if known.
Latitude	Double	42.17951	The vehicle's last known latitude, if known.
Longitude	Double	-71.09089	The vehicle's last known longitude, if known.
Direction	Integer	15	The vehicle's last known compass direction in degrees, if known.
Speed	Integer	47	The vehicle's last known speed (mph) if known.
Lateness	Integer	120	Current lateness of vehicle (seconds), if known.

Flag	Meaning
Sch	Record has scheduled time but not the vehicle lateness. May have vehicle ID and position.
Pre	Record has scheduled time and the vehicle's current lateness and position.
App	Approaching the station now. Record has vehicle location but not lateness.
Arr	Arriving at the station now. Record has vehicle location but not lateness.
Dep	Departing the station now or has departed the station. Record has vehicle location but not lateness. Used at the trip's origin.
Del	"Delayed." Vehicle is late and is not moving. The lateness may or may not be included.