**GTFS2GMNS: Convert Transit Data to Space-Time Multimodal Network for Equity Analysis and Decision-Making**

Xiangyong Luo1,#, Fang Tang1,#, Han Wang2,3, Xuesong (Simon) Zhou 1,\*

1 School of Sustainable Engineering and the Built Environment, Arizona State University, Tempe, AZ 85281, United States

2 Civil and Environmental Engineering, University of California, Berkeley, CA, 94710, United States

3 Electrical Engineering and Computer Science, University of California, Berkeley, CA, 94710, United States

# First Authors; \* Faculty advisor. Email: [xzhou74@asu.edu](mailto:xzhou74@asu.edu)

**ABSTRACT**

GTFS2GMNS is a groundbreaking, open-source tool that acts as a comprehensive platform for transforming the General Transit Feed Specification (GTFS) data into actionable insights. Developed with a class-based architecture, the tool provides an all-in-one solution for reading, converting, and building multi-modal transit networks that encompass various modes such as bus, metro, and rail. This cutting-edge tool distinguishes itself through its expansive capabilities in intricate modeling, rigorous analysis, and dynamic visualization features. With its advanced capabilities, the software caters to a wide range of stakeholders, from policymakers to community advocates. It stands as a pivotal tool in streamlining data-driven decision-making, guiding evidence-based policy formulation, and nurturing interdisciplinary collaboration. Exportable in the General Modeling Network Specification (GMNS) format, the tool promotes seamless data sharing and leverages an open-source ethos for community-driven improvements and long-term sustainability. Architected for inherent scalability and adaptability, the software accommodates the dynamic requirements of expanding and diversifying transit networks. Serving as a catalyst for transformative change, GTFS2GMNS is steadfastly committed to promoting an equitable, accessible, and seamlessly interconnected transit ecosystem, engineered to adapt and evolve in concert with the fluctuating contours of contemporary public transportation systems.

**INTRODUCTION**

As cities around the world confront the challenges of urban sprawl and increasingly congested roadways, there is a heightened focus on optimizing existing infrastructures. Rather than just expanding lanes, many authorities are shifting their focus toward public transit system analysis and planning. They aim to mitigate traffic issues without overburdening the physical framework. Amidst this scenario and the growing adoption of digital twin technology in various sectors, GTFS2GMNS,<https://github.com/asu-trans-ai-lab/GTFS2GMNS>, was born. This tool was conceived out of the urgent need to enhance transit planning and analysis, especially given the complex dynamics of multi-modal transit systems.

The drive to develop GTFS2GMNS stemmed from the challenges posed by multi-modal transit network modeling. The fast-rising surface traffic congestion in urban regions was swiftly overtaking the available infrastructure. A more integrated modeling approach became vital due to the interconnected dynamics of various transportation modes. Our team identified the need to bridge the GMNS format of the generalized highway network and the GTFS data format of the public transit network. The ASU Trans+AI team played a pivotal role, in guiding the design, addressing algorithmic challenges, and refining the tool through iterative processes.

At its essence, GTFS2GMNS provides a solution to the pressing need for a streamlined, unified network framework, facilitating multi-modal transit network modeling. This tool offers a platform that integrates diverse transit data formats, simplifying transit system analysis, optimization, and planning processes. GTFS2GMNS is tailored for a wide audience. While transportation planners might leverage this open-source network structure for multiple intricate tasks, it also serves policymakers, transit service providers, and community advocates. The tool holds immense potential, especially for those keen on enhancing transportation efficiency, safety, and reducing emissions.

**METHODOLOGY**

The GTFS is a universal data format that depicts scheduled transit services, such as buses and trains, which can be downloaded at <https://transitfeeds.com/feeds>. The attributes definition of related files can be found at <https://developers.google.com/transit/gtfs/reference#tripstxt>. Adopted globally, this standardized format promotes interoperability, enabling transit agencies to share and integrate data effortlessly. This not only streamlines multimodal transport planning but also fosters cost-efficient, data-driven decision-making. As many transit agencies offer their GTFS datasets openly, third-party developers can craft diverse applications, enhancing the transit experience for users by providing real-time scheduling, route maps, and more.

**Space-time service network design**

As shown in Figure 1 of APPENDIX, the static network includes a physical layer and a service layer. The physical layer connects with interested zones, while the service layer represents route direction, frequency, and services. Considering the time dimension, the space-time transit network is constructed. Given a network with transit service link sets , this paper introduces a decision variable which define arc is the selected or not by different socio-demographic group for a set of od pairs . Specifically, socio-demographic group is related to car ownership, which represents different access to the nearby transit service, such as park and ride stations. For example, low-income groups with zero car ownership only have the walk-to-transit station option.

A virtual access arc will be established to represent the failure to access. Travel time budget on transit-only network is defined to represent the demand for each OD pair in group . Thus, all the arcs in sets are servable/accessible for each OD pair in group . Otherwise, a virtual service link will be added. In the virtual link, the cost is set as 1, while the cost is 0 in the transit service links.

**Equity and accessibility analysis**

As shown in Table 1, a simple example is utilized to illustrate the gap measurement. The gap is defined as 1 when it is inaccessible and 0 otherwise. For example, the gap of OD pair is 0 (accessible) and is 1 (inaccessible) in a high-income group, while both two OD pairs are 1 (inaccessible) in a low-income group. The total gap over all destinations in the high-income group is 1, while 2 in the low-income group. Therefore, the low-income group with the largest gap has inequitable transit access and has the highest priority to construct new service arcs to improve its accessibility. The mathematical model can be found in **APPENDIX**.

**Table 1 An example of gap measurement**

| **Socio-demographic group** | **Destination** | **Gap** | **The total gap over all destinations in a group** | **Largest gap over all groups** |
| --- | --- | --- | --- | --- |
| High-income | d1 | 0 | 1 | 2 = max(2,1) |
| High-income | d2 | 1 |
| Low-income | d1 | 1 | 2 |
| Low-income | d2 | 1 |

**OUTCOMES**

**Innovation and value proposition**

The GTFS2GMNS framework stands as a testament to innovation in the realm of transit planning, both for our organization and for the broader transit industry. Historically, underserved communities have grappled with a dearth of sufficient transit-related data, impeding their ability to make informed, data-driven decisions. Our tool bridges this gap with an avant-garde approach, reshaping transit planning by leveraging the GTFS data to construct multi-modal transit networks. Such an integrated modeling platform, complemented by an open-source data hub framework, offers a unique blend of versatility and specificity that reshapes how agencies perceive and approach transit modeling. The visualization tools like QGIS and NeXTA help further express GTFS2GMNS’s uniqueness, making transit data not just comprehensible but also visually insightful. For example, as shown in Figure 2 of APPENDIX, (a) represents the multimodal transit network generated by GTFS2GMNS, (b) shows the volume visualization for each metro segment by traffic assignment.

Beyond technical advancements, the behavioral studies incorporated into our framework address the complex social dimensions of transportation. Through innovative data collection methods and equity-centered approaches, we've inspired holistic methodologies that not only discern mobility patterns but also understand the underlying socioeconomic dynamics. The post-COVID-19 mobility pattern analysis, for instance, is a vanguard initiative that gleans insights from the pandemic-induced shifts in human movement, aiding transit planners in policy formulation.

**Achieved outcomes and potential impacts**

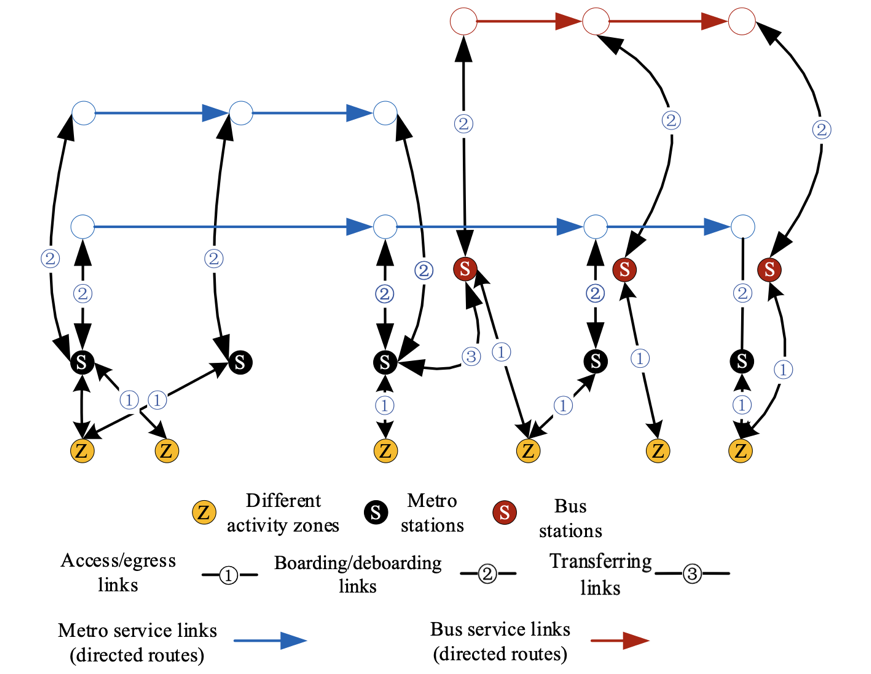
GTFS2GMNS has revolutionized data-driven decision-making by enabling a holistic view of transit multimodal networks with a focus on equity and accessibility. The specific dataset we are working on is the Northern Virginia Transportation Authority (NVTA) transit dataset, which can be found at [https://novagateway.org/Dashboard/Overview](https://urldefense.com/v3/__https://novagateway.org/Dashboard/Overview__;!!IKRxdwAv5BmarQ!Z13nBeg10VZAvZTa8ysOvH-rvQ0ABhwHAJnzcT3z4W1LCbYKC9nnqx17WlIZJZhYLtJEhoJlo-COGg$). GTFS2GMNS is also used in the recent NSF project, POSE: Phase II: CONNECT: Consortium of Open-source plaNNing models for Next-generation Equitable and efficient Communities and Transportation, <https://www.nsf.gov/awardsearch/showAward?AWD_ID=2303748&HistoricalAwards=false>.

Beyond this, its potential to shape a dynamic, interconnected transit ecosystem is profound. Incorporating digital twin technology emphasizes simulating real-world scenarios to enhance policy decisions. As illustrated in Figure 3 of APPENDIX, by integrating census data and the GTFS2GMNS transit data hub, we can explore the correlations between demographic variations and altered travel behaviors during the COVID-19 pandemic. This work could illuminate transportation-related disparities in marginalized communities, offering invaluable insights for city planners and policymakers in the digital age.

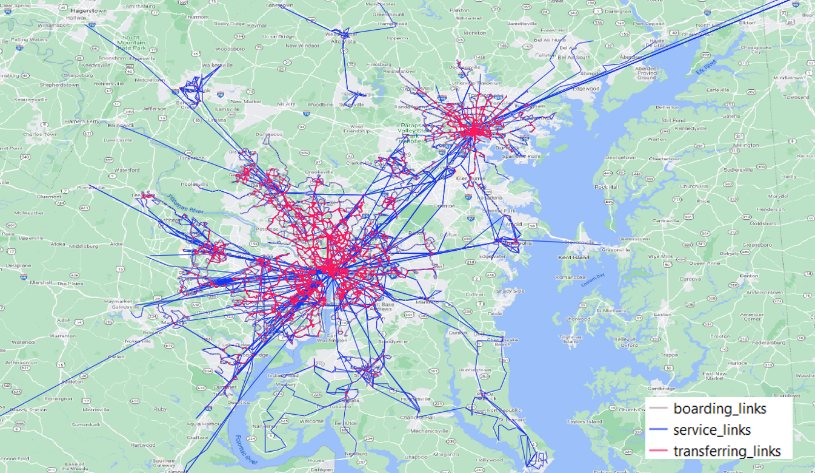
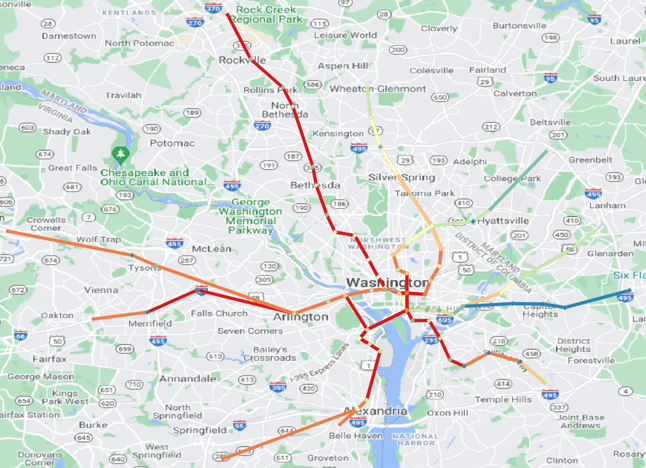
**Adoption by other agencies**

GTFS2GMNS provides a blueprint for transit agencies seeking integrated, equity-centric planning. Using GMNS data in CSV format, agencies can enhance their modeling processes and decision-making. Our open-source approach ensures accessibility for all. Depending on their unique needs, agencies can adopt and tailor the framework, even integrating local data collection techniques for a community-specific solution. With its adaptability, GTFS2GMNS is suitable for various transit settings, from urban centers to rural areas. Agencies of any size can leverage this tool to promote equity and refine their transit operations.

**APPENDIX**



**Figure 1 Physical layer and service layer of multi-modal transit network**

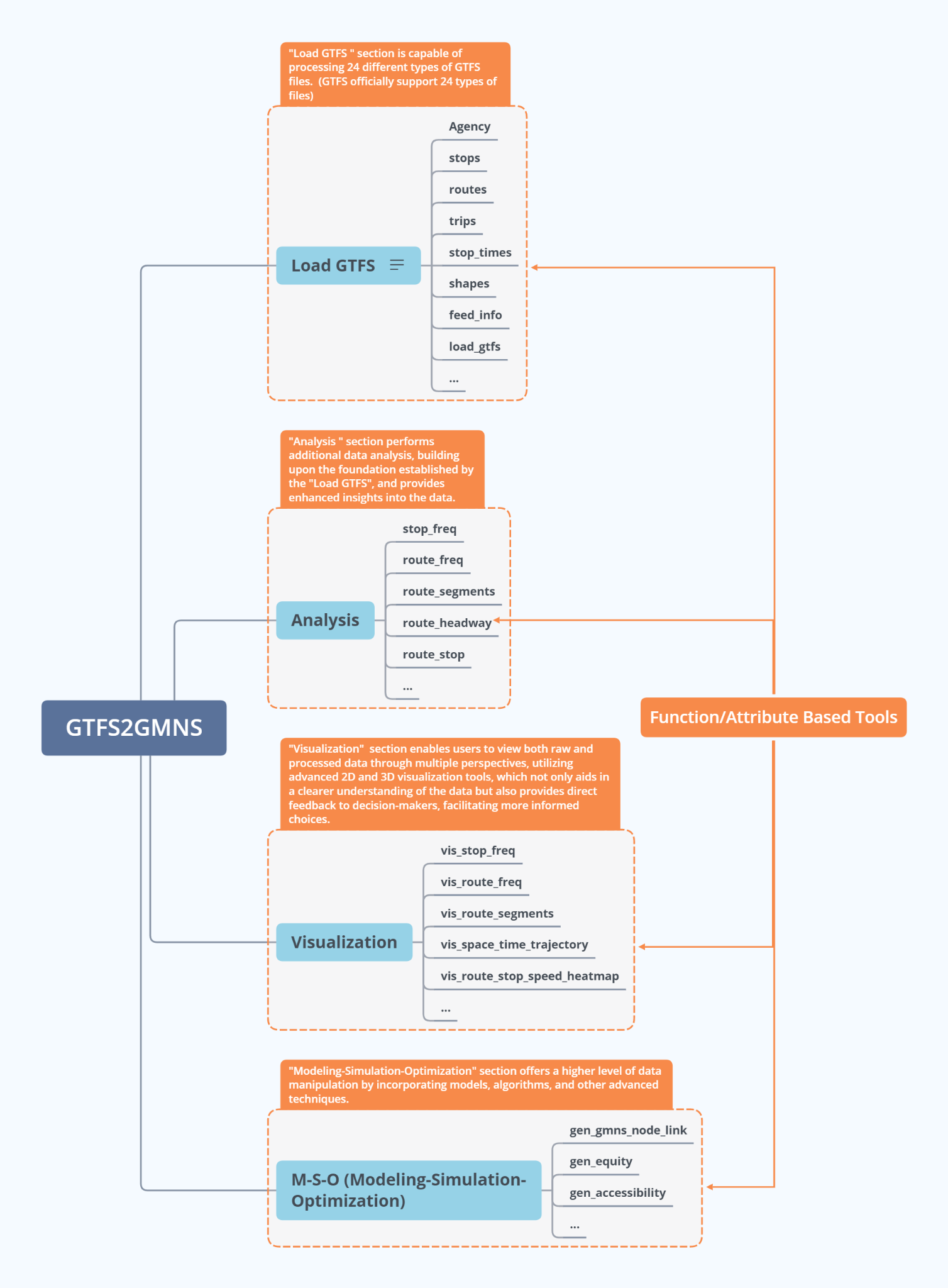
(a) (b)

**Figure 2 A practical example, (a) Transit service network in Washington DC, (b) Transit assignment results of metro in Washington DC at AM period**

A diagram of a diagram

Description automatically generated

**Figure 3 Potential Impacts based on GTFS2GMNS transit data hub and testbeds**

****

**Figure 4 GTFS2GMNS Functional Framework: Interconnected Components of Load GTFS, Analysis, Visualization, and Modeling-Simulation-Optimization Sections**

**Equity-oriented network design problem**

As stated above, the primary concern is the largest gap between transit dependency and accessibility. The gap can be represented as

|  | (1) |
| --- | --- |

Where when a path belonging to OD pair is accessible, otherwise .

Thus, the accessibility Gap for each original zone in group is defined as

|  | (2) |
| --- | --- |

Considering each segment group in the zone , the largest gap across all segment group is formulated as

|  | (3) |
| --- | --- |

Considering each segment group in the entire transit network, the largest gap across all segment group and all zones is defined as an auxiliary variable , as shown in formulation (4).

|  | (4) |
| --- | --- |

The network design model aims to find the best network design and multimodal path solutions to minimize the maximum gaps which can be mathematically interpreted as the following objective function.

|  | (5) |
| --- | --- |

The objective function (5) can be simplified as

|  | (6) |
| --- | --- |

Subject to (a) flow balance constraints for all segment groups

|  | (7) |
| --- | --- |

(b) coupling constraints between space-time arcs and physical links, as design variable,

|  | (8) |
| --- | --- |

and (c) budget constraint

|  | (9) |
| --- | --- |