University of Milan - Bicocca
**School of Sciences**
**Dept. of Computer Science, Systems and Communication**
**Master's Degree in Computer Science**

# Île-de-France Mobilités Metro Transport Network

**Project group:**
*Cristian Piacente* - 866020

**Data Analytics - Academic Year 2023-2024**

# Contents

# Chapter 1

# Introduction

## 1.1 Context

The Île-de-France region has one of the most extensive and complex public transport networks in Europe. Managed by Île-de-France Mobilités, the system includes:

- **Metro:** 16 lines primarily serving central Paris.
- RER (Réseau Express Régional): 5 express train lines connecting the suburbs with the city center.
- Trams and Buses: Complementary services that expand the reach of the network.

Given the scale of this network, our analysis will focus specifically on the **Paris Metro**, a crucial component of the transportation system that carries millions of passengers daily.

## 1.2 Approach

Representing transport networks as graphs allows for a comprehensive evaluation of their **efficiency, resilience and vulnerability**. In particular, we use an undirected weighted graph, since the metro always has both directions and the weight represents the travel time (in minutes). Our goals are:

- **Graph Construction**: Building an undirected graph representation where nodes represent stations and edges represent metro connections, as well as interactive visualizations.
- **Network Metrics Computation**: Calculating key graph measures such as degree distribution, centrality and clustering.

- **Resilience and Failure Simulations**: Simulating random failures and targeted attacks to assess network vulnerability.
- **Load Analysis**: Studying station congestion patterns.

Through this approach, we aim to extract actionable insights into the robustness and performance of the Paris Metro network.

## 1.3 Dataset Source and Selection

This study is based on **General Transit Feed Specification (GTFS)** data, a standardized format for public transport schedules and related geographic information. The dataset used comes from the **GTFS Datahub (Mobility Database)** available at:

`https://mobilitydatabase.org/feeds/mdb-1026`

The choice of this data source was made due to its frequent updates. While other sources like **TransitFeeds** provide GTFS data, their last available dataset for Île-de-France Mobilités was dated August 1, 2019. In contrast, the **Mobility Database stores new datasets twice a week**, ensuring recent and accurate data. The version used in this project is the one released on **February 6, 2025**.

## 1.4 Preprocessing and Data Cleaning

To ensure a robust dataset, the following steps were performed:

- **Metro Filtering:** Extracted only Metro routes (`route_type` = 1).
- **Duplicate Station Removal:** Stations with identical names were merged to avoid artificial fragmentation.
- **Time Correction:** Fixed incorrect timestamps where hour values exceeded 24 by applying modulo 24.
- **Stop-to-Stop Mapping:** Ensured consistency in station identifiers across datasets.

| Dataset | Number of Entries |
|---|---|
| Metro Routes | 16 |
| Metro Trips | 47,258 |
| Metro Stop Times | 1,176,628 |
| Metro Stops (after deduplication) | 321 |
| Metro Transfers | 498 |

Table 1.1: Summary of Processed Paris Metro GTFS Data

The final dataset consists of:

- `metro_stops.csv` - Paris Metro stations (nodes).
- `metro_stop_times.csv` - Travel sequences and times (edges).
- `metro_transfers.csv` - Inter-line transfer connections.

# Chapter 2

# Graph Construction and Visualization

## 2.1 Graph Representation

The Paris Metro network is modeled as an **undirected, weighted graph** $G = (V, E)$, where:

- **Nodes** $(V)$ represent Metro stations.
- **Edges** $(E)$ represent direct travel between stations.
- **Edge Weights** $(w_{ij})$ correspond to the travel time (in minutes) between stations $i$ and $j$.

### 2.1.1 Building the Network

The graph was constructed leveraging:

- Stop sequences from `metro_stop_times.csv`.
- Transfer connections from `metro_transfers.csv`.

Each edge in the network carries a weight $w_{ij}$ corresponding to the travel time, allowing us to analyze travel efficiency.

### 2.1.2 Interactive Visualization

To explore the network structure, an **interactive map** was created:

- **Nodes (Stations)**: Displayed with hover tooltips showing name and coordinates.
- **Edges (Metro Links)**: Colored based on travel time (heatmap gradient).

Figure 2.1 illustrates the network visualization.

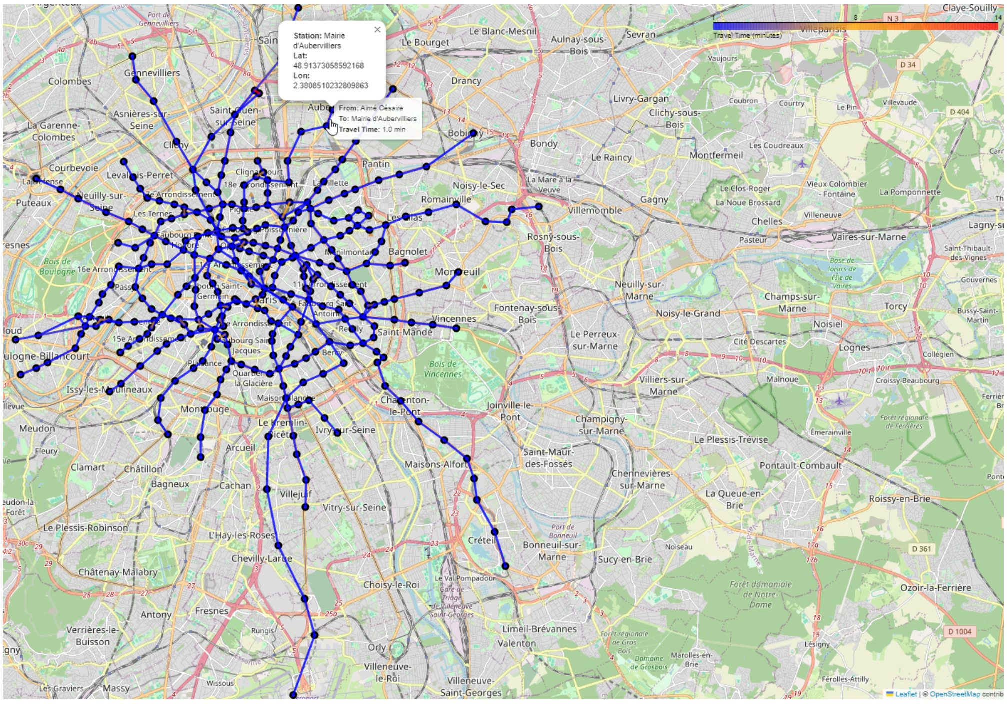


Figure 2.1: Interactive Visualization of the Paris Metro Network

## 2.2 Topological Properties

To assess the structure and efficiency of the Paris Metro network, we computed key **topological properties**. These metrics help understand the connectivity, efficiency and resilience of the system.

### 2.2.1 Degree Distribution

The **degree distribution** indicates the number of connections each station has within the network. A network with a high variance in degree distribution often exhibits a scale-free structure, meaning a few key hubs handle most of the connections.

Figure 2.2 shows the **degree distribution** of the Paris Metro network. Most stations have a degree of **2**, meaning they connect to two other stations, forming linear chains. However, a few major hubs exhibit significantly higher connectivity, reinforcing their critical role in maintaining network cohesion.

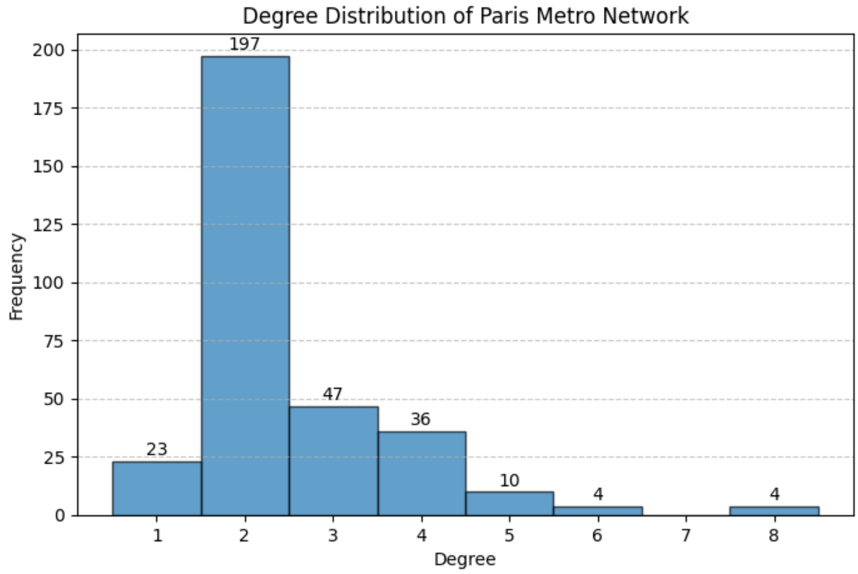


Figure 2.2: Degree Distribution of the Paris Metro Network

The **average degree** of the network is defined as:

$$\langle k \rangle = \frac{2|E|}{|V|} \tag{2.1}$$

where $|E|$ is the number of edges and $|V|$ is the number of nodes in the network. For the Paris Metro, we computed:

$$\langle k \rangle = 2.52 \tag{2.2}$$

This suggests that most stations maintain only a few direct connections. However, the presence of several high-degree hubs indicates that removing these hubs could significantly fragment the network.

### 2.2.2 Network Diameter

The **network diameter** represents the longest shortest path between any two stations in the metro system.

$$D = 36 \tag{2.3}$$

This means that the maximum number of metro stops required to travel between any two stations in the network is **36**. A high diameter suggests that the network is **geographically extensive** but may require multiple transfers for long-distance trips. The diameter value reinforces the importance of **high-centrality stations**.

### 2.2.3 Network Density

**Network density** is the ratio of existing edges to the total possible edges in a fully connected graph:

$$\rho = \frac{2|E|}{|V|(|V|-1)} \tag{2.4}$$

The computed density for the Paris Metro network is:

$$\rho = 0.0079 \tag{2.5}$$

This low value suggests a sparse structure, where most stations connect to only a few neighbors rather than forming a fully connected mesh.

### 2.2.4 Network Efficiency

Network efficiency measures how easily passengers can travel through the network using the shortest paths. The computed efficiency is:

$$E = 0.1124 \tag{2.6}$$

This relatively low value suggests that the network prioritizes geographic coverage over optimal connectivity.

### 2.2.5 Clustering Coefficient

The **clustering coefficient** measures the extent to which stations form interconnected clusters. The computed value is:

$$\langle C \rangle = 0.0088 \tag{2.7}$$

This low clustering suggests that the metro operates as a **mostly linear network**, rather than a highly redundant one.

### 2.2.6 Shortest Path Length

The **average shortest path length** determines the typical journey time across the network. The computed value is:

$$\langle L \rangle = 12.25 \text{ minutes} \tag{2.8}$$

This means that a typical passenger must travel approximately **12 minutes** to reach their destination.

### 2.2.7 Assortativity

**Assortativity** indicates whether high-degree stations tend to connect to other high-degree stations. The computed assortativity is:

$$r = 0.0588 \tag{2.9}$$

Since this value is close to zero, the metro network is **almost neutral**, meaning hubs do not necessarily interconnect.

### 2.2.8 Summary of Topological Properties

| **Metric** | **Value** |
|---|---|
| Average Degree | 2.52 |
| Network Diameter | 36 |
| Network Density | 0.0079 |
| Network Efficiency | 0.1124 |
| Clustering Coefficient | 0.0088 |
| Average Shortest Path Length | 12.25 mins |
| Assortativity | 0.0588 |

Table 2.1: Topological Properties of the Paris Metro Network

These results suggest that the Paris Metro is a **moderately connected system**, with **key hubs** playing a critical role in ensuring connectivity. The **low clustering coefficient** and relatively high shortest path length indicate that **disruptions could significantly impact connectivity**, which we explore in the next chapter.

# Chapter 3

# Network Vulnerability Analysis

## 3.1 Structural Robustness Analysis

To assess the vulnerability of the Paris Metro network, we analyzed key structural properties that indicate its **resilience to failures** and **potential bottlenecks**. The following centrality measures were computed:

- **Degree centrality:** Identifies the most well-connected stations in terms of direct neighbors.
- **Betweenness centrality:** Measures how frequently a station acts as a bridge along the shortest paths in the network.
- **Closeness centrality:** Indicates how efficiently a station can reach all others in the network.
- **Eigenvector centrality:** Identifies stations that are well-connected to other important stations.
- **PageRank centrality:** Highlights stations that are more influential based on the importance of their neighbors.
- **Spectral gap & Algebraic connectivity:** Evaluate the network's overall connectivity and resilience to failures.

### 3.1.1 Critical Stations by Degree Centrality

Degree centrality identifies the most well-connected stations, highlighting transit hubs with **multiple direct connections**.

| Station | Degree Centrality |
|---|---|
| Châtelet | 0.025000 |
| Saint-Lazare | 0.025000 |
| Montparnasse Bienvenue | 0.025000 |
| République | 0.025000 |
| Opéra | 0.018750 |
| Bastille | 0.018750 |
| La Motte-Picquet - Grenelle | 0.018750 |
| Nation | 0.018750 |
| Duroc | 0.015625 |
| Michel-Ange - Molitor | 0.015625 |

Table 3.1: Top 10 Stations by Degree Centrality

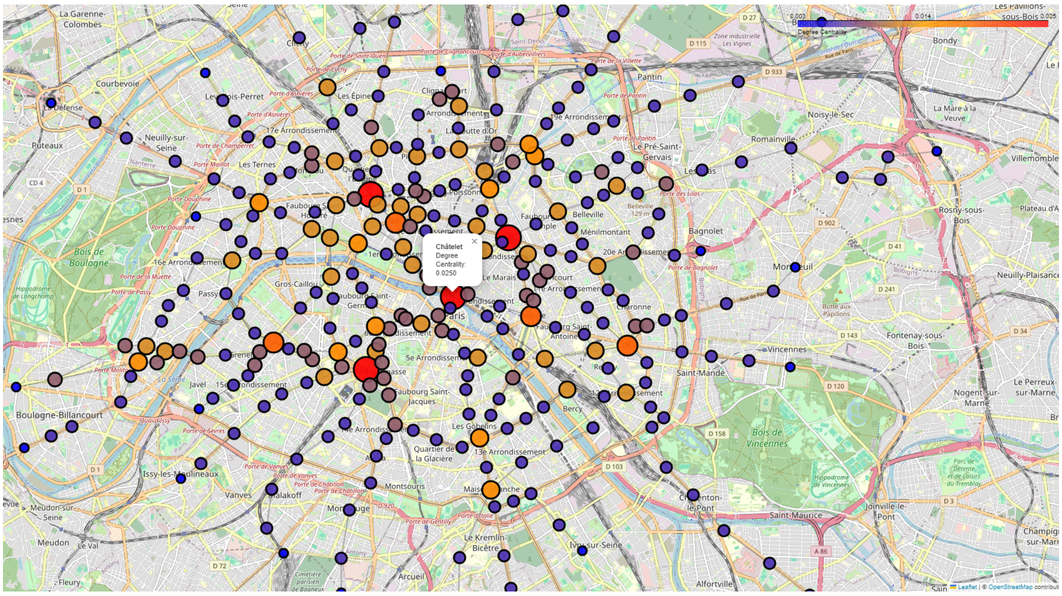

Figure 3.1: Degree Centrality

Stations like **Châtelet, Saint-Lazare and Montparnasse Bienvenue** are highly connected, making them key interchange hubs. Their disruption would affect multiple metro lines simultaneously.

### 3.1.2 Critical Stations by Betweenness Centrality

Betweenness centrality measures how often a station appears on shortest paths between other stations.

| Station | Betweenness Centrality |
|---|---|
| Châtelet | 0.335880 |
| Madeleine | 0.304445 |
| Gare de Lyon | 0.285838 |
| Pyramides | 0.220902 |
| Concorde | 0.216997 |
| Saint-Lazare | 0.206688 |
| Invalides | 0.180556 |
| République | 0.168462 |
| Bercy | 0.154855 |
| Montparnasse Bienvenue | 0.150316 |

Table 3.2: Top 10 Stations by Betweenness Centrality

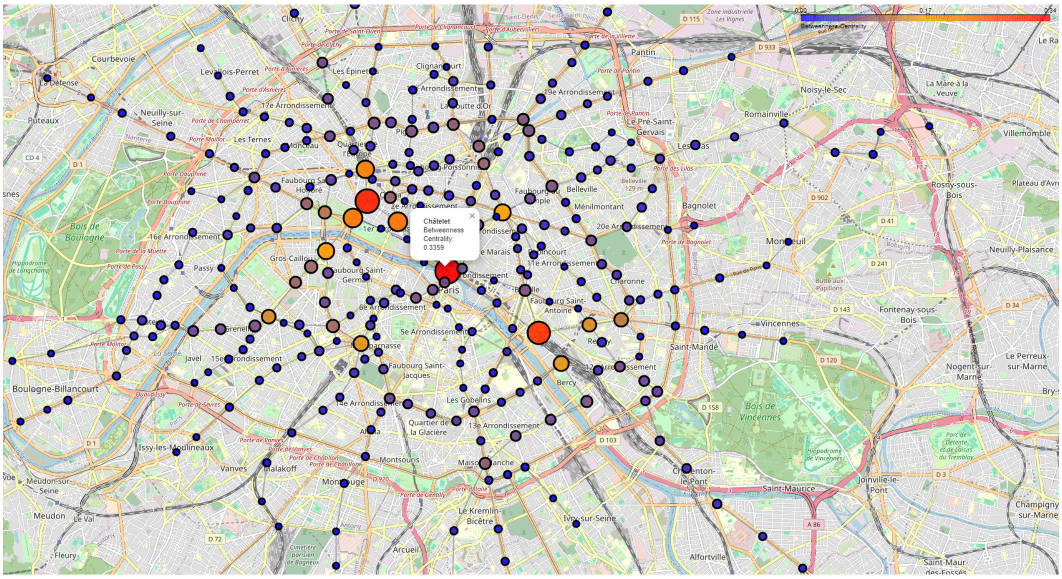


Figure 3.2: Betweenness Centrality

**Châtelet, Madeleine and Gare de Lyon** function as network bridges. Their failure forces rerouting, increasing congestion in alternative paths.

### 3.1.3 Critical Stations by Closeness Centrality

Closeness centrality measures how quickly a station can reach all others.

| Station | Closeness Centrality |
|---|---|
| Châtelet | 0.129450 |
| Madeleine | 0.128411 |
| Pyramides | 0.127847 |
| Opéra | 0.125049 |
| Gare de Lyon | 0.123504 |
| Concorde | 0.122277 |
| Hôtel de Ville | 0.121581 |
| Saint-Lazare | 0.120937 |
| Les Halles | 0.119225 |
| Cité | 0.118212 |

Table 3.3: Top 10 Stations by Closeness Centrality

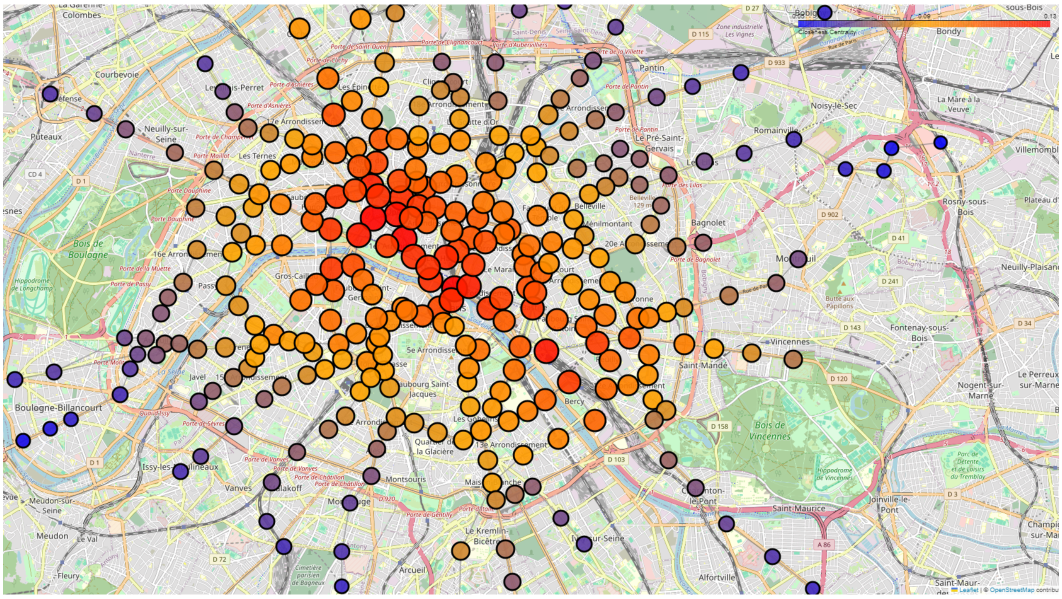

Figure 3.3: Closeness Centrality

Stations with high closeness centrality are **centrally located**, making them well-positioned for fast access across the network. **Châtelet, Madeleine and Pyramides** allow passengers to reach most destinations with minimal travel time.

### 3.1.4 Critical Stations by Eigenvector Centrality

Eigenvector centrality assigns higher importance to stations connected to well-connected stations.

| Station | Eigenvector Centrality |
|---|---|
| Opéra | 0.379321 |
| Saint-Lazare | 0.362677 |
| Havre-Caumartin | 0.307907 |
| Madeleine | 0.295104 |
| Pyramides | 0.261041 |
| Chaussée d'Antin - La Fayette | 0.250421 |
| Saint-Augustin | 0.220115 |
| Richelieu - Drouot | 0.216467 |
| Châtelet | 0.206400 |
| Miromesnil | 0.196227 |

Table 3.4: Top 10 Stations by Eigenvector Centrality

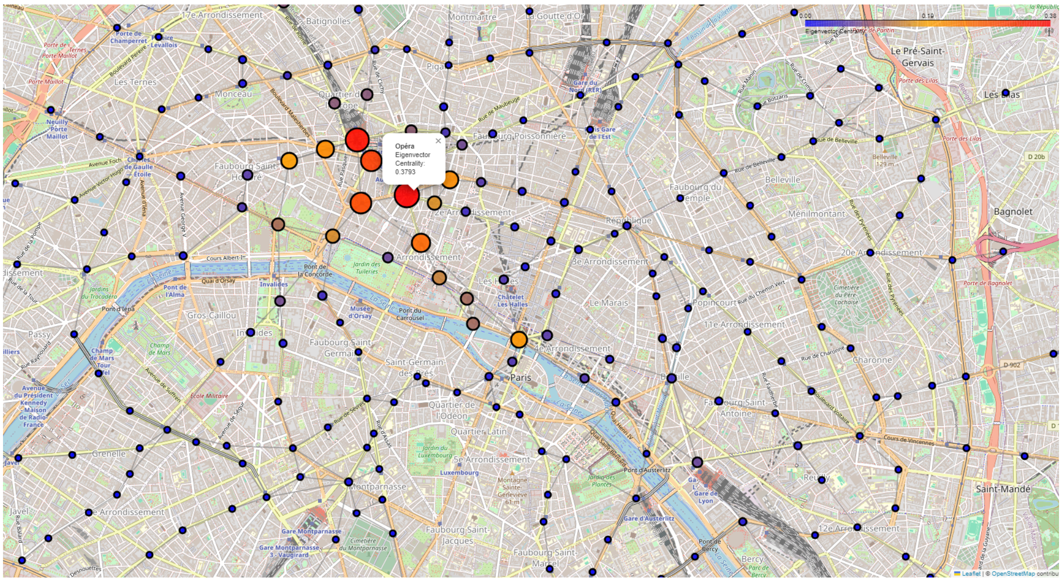

Figure 3.4: Eigenvector Centrality

**Opéra, Saint-Lazare and Havre-Caumartin** are ranked highest, indicating that they connect to other important stations, making them influential despite not having the highest traffic volume.

### 3.1.5 Critical Stations by PageRank Centrality

PageRank centrality ranks stations based on their importance weighted by their neighbors' importance.

| Station | PageRank Centrality |
|---|---|
| Saint-Lazare | 0.007317 |
| Carrefour Pleyel | 0.007145 |
| Maison Blanche | 0.007033 |
| Église d'Auteuil | 0.006848 |
| République | 0.006642 |
| Gare du Nord | 0.006339 |
| Châtelet | 0.006194 |
| Saint-Denis - Pleyel | 0.006085 |
| Charles de Gaulle - Étoile | 0.006034 |
| Place d'Italie | 0.005906 |

Table 3.5: Top 10 Stations by PageRank Centrality

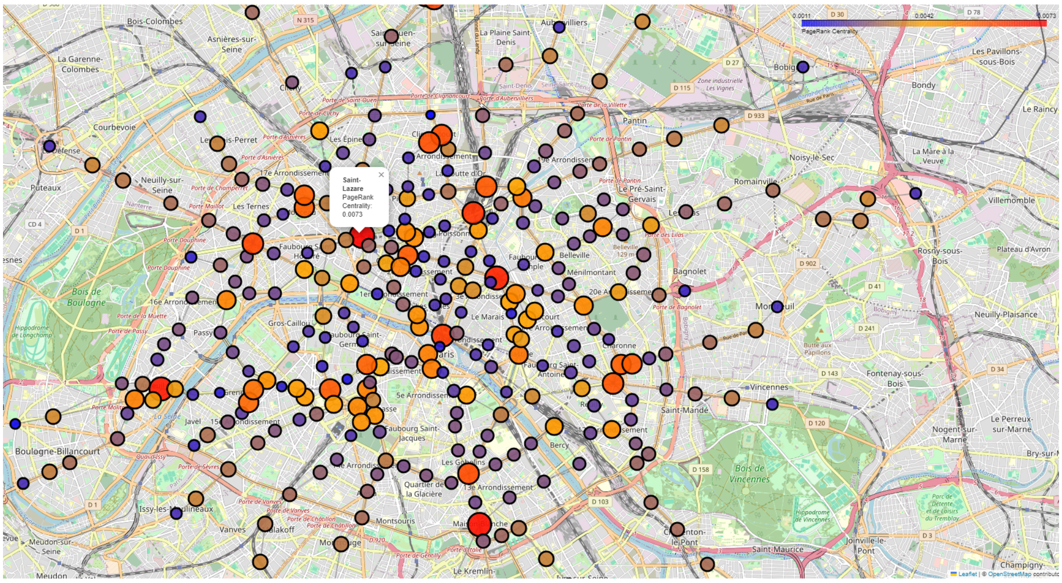

Figure 3.5: PageRank Centrality

**Saint-Lazare ranks highest**, reinforcing its role as a strategic interchange station, serving multiple metro lines (four lines).

### 3.1.6 Spectral Gap and Algebraic Connectivity

Spectral analysis provides insights into how well the metro system remains connected under failures. The two key metrics are:

- **Spectral gap** measures how quickly information spreads across the network.

- **Algebraic connectivity** (the second-smallest eigenvalue of the Laplacian matrix) indicates how well the graph remains connected if stations are removed.

For the Paris Metro, we computed:

$$\text{Spectral Gap} = 0.011314 \tag{3.1}$$

$$\text{Algebraic Connectivity} = 0.011314 \tag{3.2}$$

These values suggest that the network is relatively weakly connected, meaning that removing key stations could significantly disrupt the overall structure. The similarity between the spectral gap and algebraic connectivity suggests that the network has a uniform level of connectivity without extreme structural weaknesses.

### 3.1.7 Identifying Bottleneck Stations

While high-betweenness stations act as major transit hubs, stations with low degree centrality (fewer direct connections) can become **bottlenecks** that restrict passenger flow. The **top 10 bottleneck stations** (stations with the lowest degree centrality) are:

| **Station** | **Degree Centrality** |
|---|---|
| Mairie d'Ivry | 0.003125 |
| Mairie de Montreuil | 0.003125 |
| Pointe du Lac | 0.003125 |
| Pont de Sèvres | 0.003125 |
| Boulogne Pont de Saint-Cloud | 0.003125 |
| Bobigny Pablo Picasso | 0.003125 |
| Pont de Levallois - Bécon | 0.003125 |
| Gallieni | 0.003125 |
| Mairie d'Issy | 0.003125 |
| La Défense (Grande Arche) | 0.003125 |

Table 3.6: Top 10 Bottleneck Stations (Lowest Degree Centrality)

These stations are mostly located at the **ends of metro lines**, meaning they serve fewer connections and rely on a single or limited number of routes for passenger flow. A disruption at these stations would not cause significant network-wide collapse but could isolate neighborhoods or create accessibility issues for passengers in suburban areas.

### 3.1.8 Impact of Removing Key Stations

To measure the effect of disruptions, we computed the **global efficiency loss** when removing the most central stations. The table below lists the percentage drop in network efficiency when each of the top 10 betweenness stations is removed.

| **Station** | **Efficiency Drop (%)** |
|---|---|
| Châtelet | 7.09 |
| Montparnasse Bienvenue | 4.87 |
| Gare de Lyon | 4.48 |
| Saint-Lazare | 4.43 |
| Pyramides | 2.79 |
| République | 2.54 |
| Bercy | 2.27 |
| Invalides | 2.16 |
| Concorde | 2.06 |
| Madeleine | 1.91 |

Table 3.7: Network Efficiency Drop After Removing Key Stations

**Châtelet** is the most critical station: its removal causes a 7.09% drop in network efficiency, making it the most influential hub.
Montparnasse Bienvenue and Gare de Lyon are also highly critical, leading to over 4% efficiency loss each.
Removing Pyramides, République and Bercy causes a more moderate impact, but they still play significant roles in passenger distribution.
Madeleine and Concorde, despite their high betweenness, do not drastically affect efficiency when removed, suggesting some redundancy in their connectivity.

## 3.2 Failure Scenarios

To assess the resilience of the Paris Metro network, we simulated different failure scenarios that could impact its functionality. This analysis allows us to determine how different types of station failures affect the overall efficiency and connectivity of the network. The following types of failures were considered:

- **Random Failures**: The removal of stations at random, simulating minor disruptions such as technical malfunctions, maintenance work, or minor accidents.
- **Targeted Attacks**: The strategic removal of the most critical stations, identified based on degree and betweenness centrality, representing deliberate sabotage or system-wide failures.

- **Cascading Failures**: A dynamic failure scenario where the removal of high-betweenness stations triggers a progressive breakdown.
- **Percolation Threshold**: The minimum fraction of station removals required to fragment the network into isolated components.

For random failures and targeted attacks, only **30%** of the total stations were removed, reflecting a realistic attack threshold rather than a complete breakdown. Cascading failures, however, were modeled dynamically, where station failures evolved iteratively.

### 3.2.1 Impact of Random and Targeted Attacks

To evaluate the impact of station removals, we measured the degradation of:

1. **Network Efficiency**: A measure of how effectively passengers can travel through the network using the shortest paths.
2. **Largest Connected Component**: The fraction of stations that remain interconnected after sequential removals.

Figures 3.6 and 3.7 illustrate how the network behaves under different types of station removals.

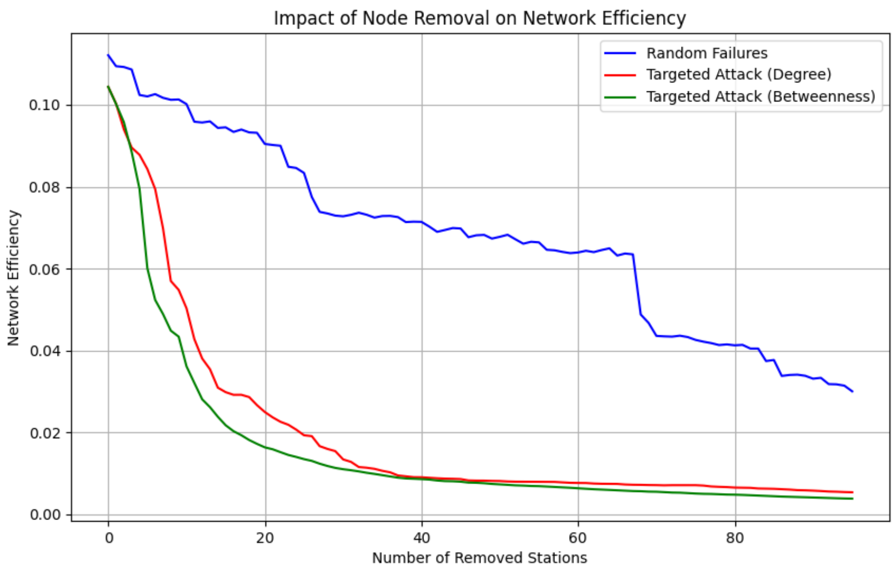


Figure 3.6: Impact of Node Removal on Network Efficiency. The graph compares efficiency degradation under random failures (blue), targeted degree attacks (red) and targeted betweenness attacks (green).

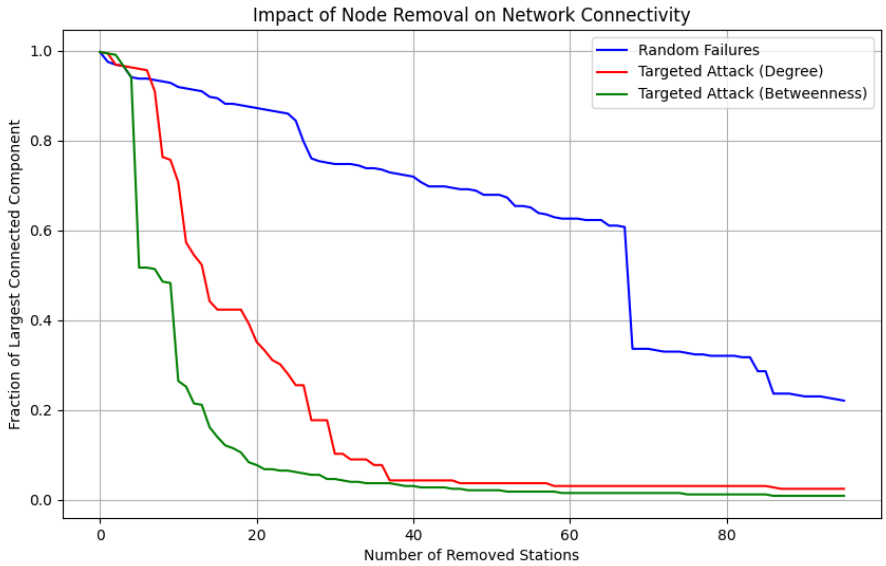


Figure 3.7: Impact of Node Removal on Network Connectivity. The largest connected component shrinks significantly under targeted attacks, particularly when stations with high betweenness centrality are removed.

**Random failures** (blue) result in a slow, gradual decline in network efficiency, indicating the metro has **built-in redundancy** that allows alternative paths to sustain connectivity under minor disruptions.

**Targeted attacks** based on **degree (red)** and **betweenness centrality (green)** lead to a much **faster collapse**. The first 10–20 station removals already disrupt a significant portion of the network.

The **largest connected component** remains largely intact under random failures but **shrinks dramatically when critical hubs are removed**. This suggests that **a few strategically important stations hold the network together**.

**Betweenness-based attacks are the most damaging**, as they remove key transit hubs that connect otherwise distant parts of the network. This shows that the network is highly dependent on major interchange stations.

To better understand the structural impact of failures, we generated geographic visualizations of the Paris Metro network after targeted attacks on high-degree and high-betweenness stations.

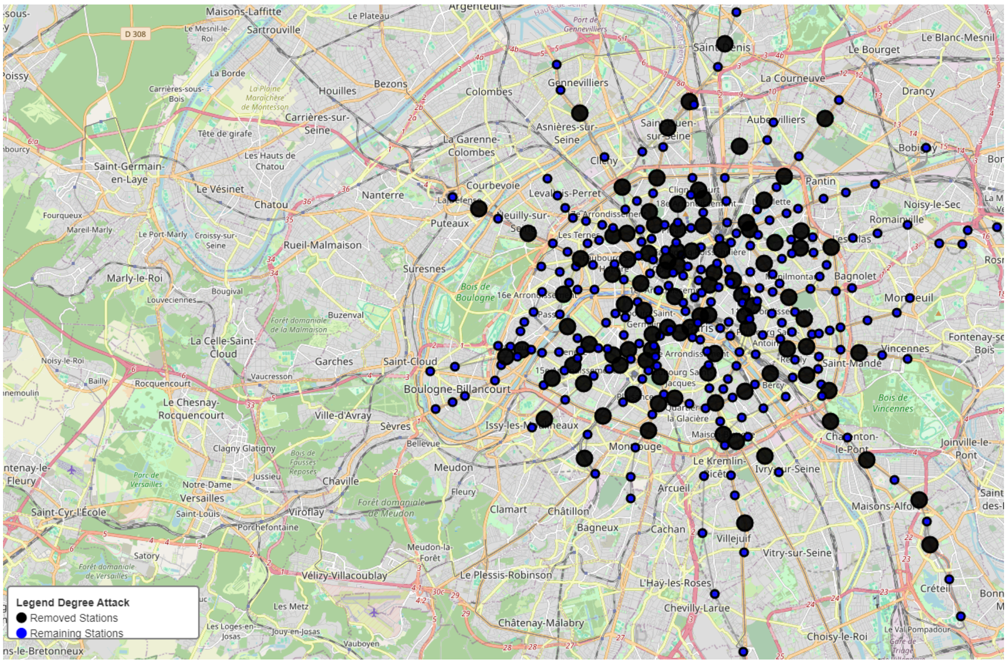


Figure 3.8: Paris Metro after Degree-based Station Failures.

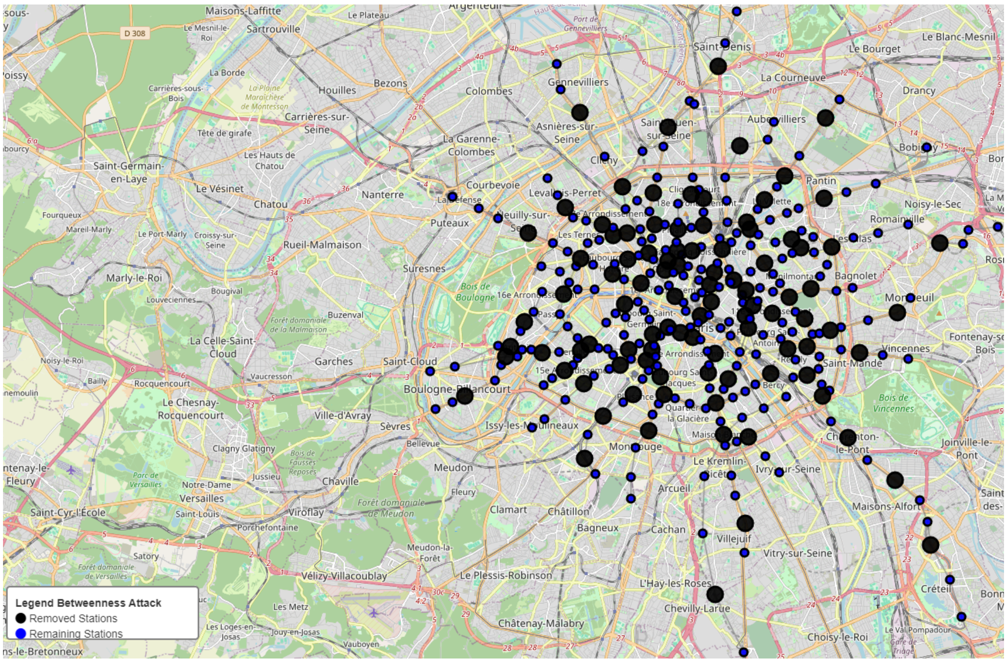


Figure 3.9: Paris Metro after Betweenness-based Station Failures.

### 3.2.2 Cascade Failures and Network Collapse

Unlike the previous scenarios, cascading failures introduce an **iterative breakdown process** where each failure increases pressure on the remaining network, triggering further collapses. Instead of a predefined removal fraction, the system dynamically determines which stations fail based on **betweenness centrality overload**:

- **High-betweenness stations** fail first, as they handle the most passenger transfers.
- At each step, betweenness centrality is recalculated to determine which stations are under the most stress.
- A failure threshold of **10% of maximum betweenness** was used to iteratively remove stations, mimicking congestion-induced collapses.
- The process continues until no more stations exceed the failure threshold.

The results of this simulation are shown in Figures 3.10 and 3.11.

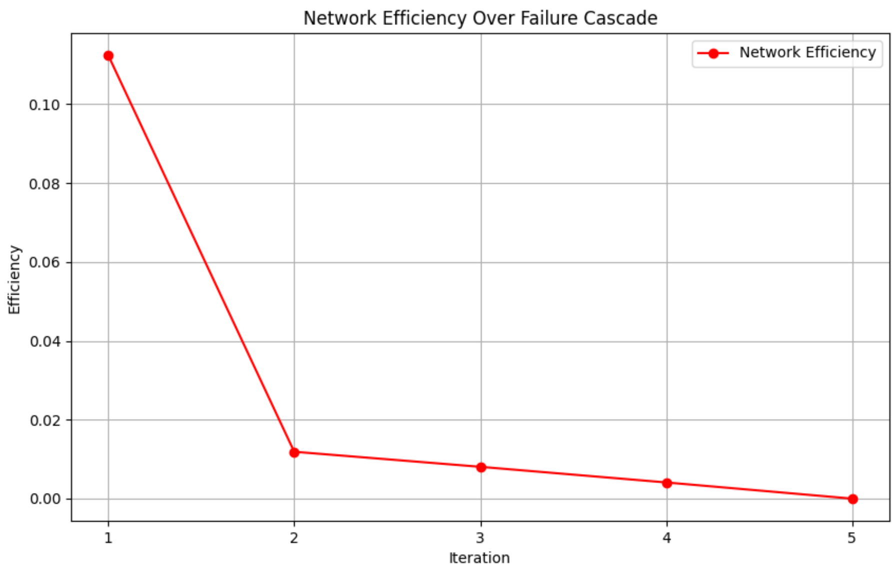


Figure 3.10: Network Efficiency Over Failure Cascade. The removal of high-betweenness stations causes an exponential decline in efficiency.

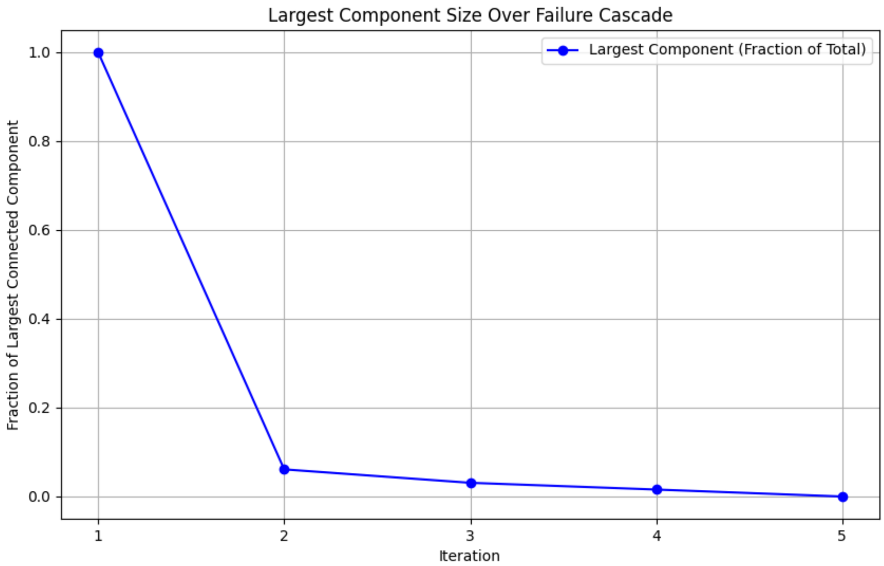


Figure 3.11: Largest Component Size Over Failure Cascade. The network quickly fragments as key transfer stations fail.

The first iteration already removes **108 high-betweenness stations**, causing a sharp drop in efficiency.
The largest connected component shrinks to **less than 20% of its original size** within two iterations.
The network reaches **total collapse by iteration 4**, demonstrating the extreme vulnerability of the system when overloaded.
The number of failed stations at each iteration is summarized in Table 3.8.

| Iteration | Failed Stations |
|---|---|
| 1 | 108 |
| 2 | 51 |
| 3 | 34 |
| 4 | 128 |

Table 3.8: Number of Failed Stations at Each Iteration of the Cascade Failure.

To further illustrate the progression of failures, Figures 3.12 to 3.15 show the metro at each stage of failure.

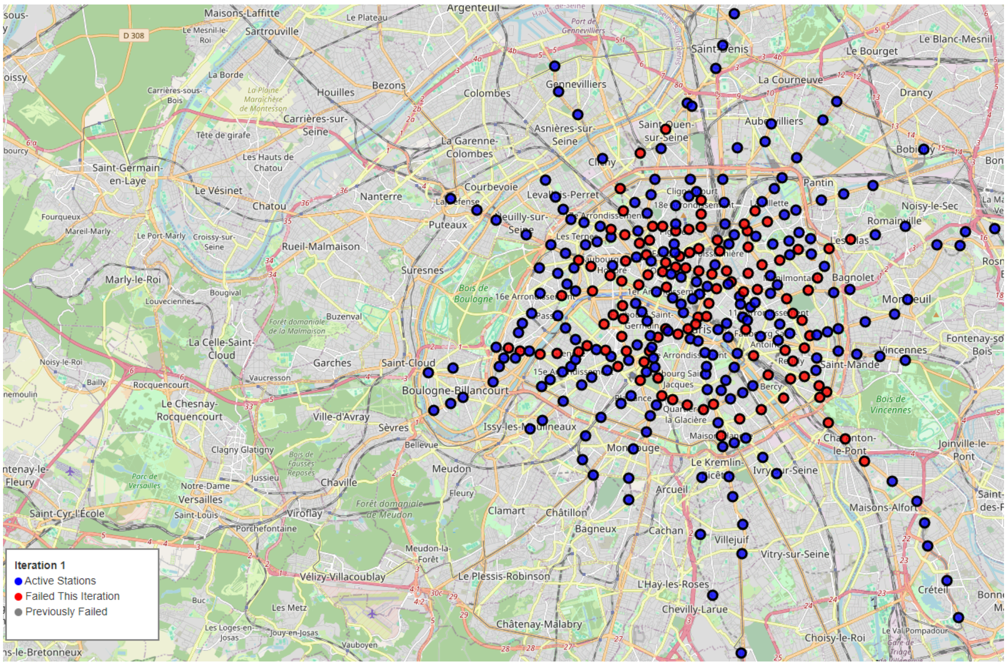


Figure 3.12: Cascade Failure - Iteration 1. Key hubs (red) fail first.

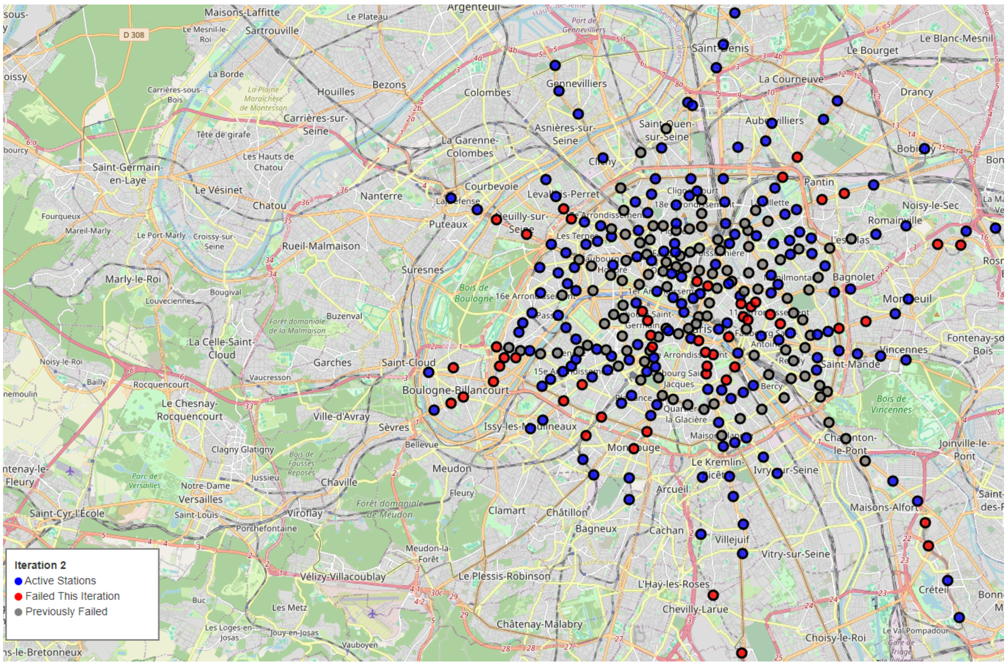


Figure 3.13: Cascade Failure - Iteration 2. Additional stations collapse.

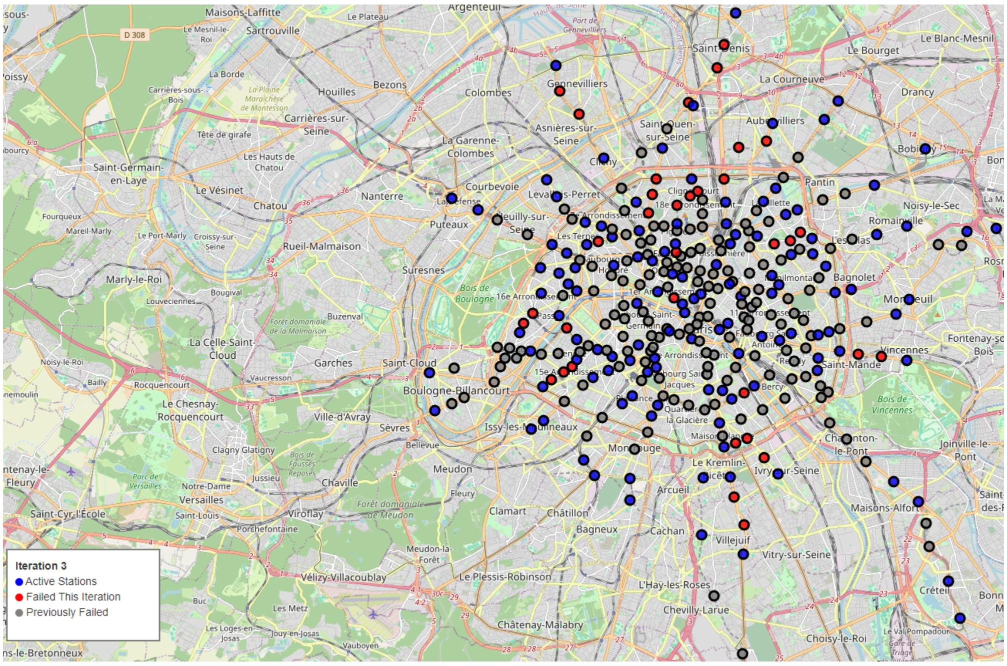


Figure 3.14: Cascade Failure - Iteration 3. Only fragmented sections remain.

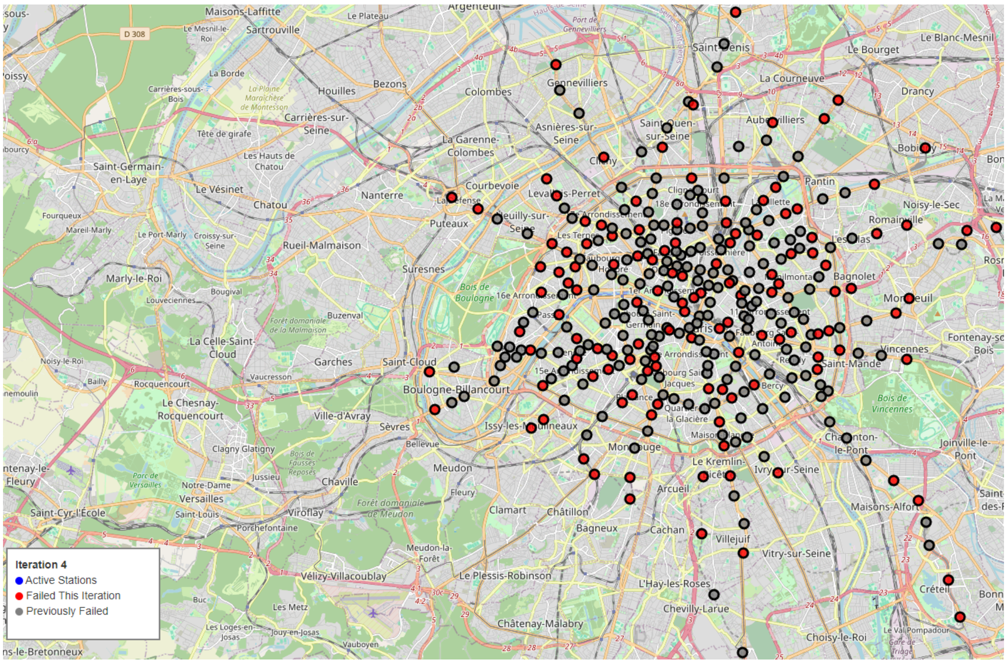


Figure 3.15: Cascade Failure - Iteration 4. The network reaches total collapse.

This analysis confirms that **cascading failures can lead to rapid systemic breakdowns**. For resilience strategies it is important to consider not just direct failures but also their effects throughout the network.

### 3.2.3 Percolation Threshold Analysis

To further analyze the network's robustness, we examined the **percolation threshold**, which determines how the network fragments when stations are progressively removed. This is measured by monitoring:

- The **average degree** of the network (the number of connections per station).
- The **size of the largest connected component** (the fraction of stations that remain reachable in the largest subnetwork).

The results, shown in Figure 3.16, highlight the different rates at which the network collapses under various attack strategies.

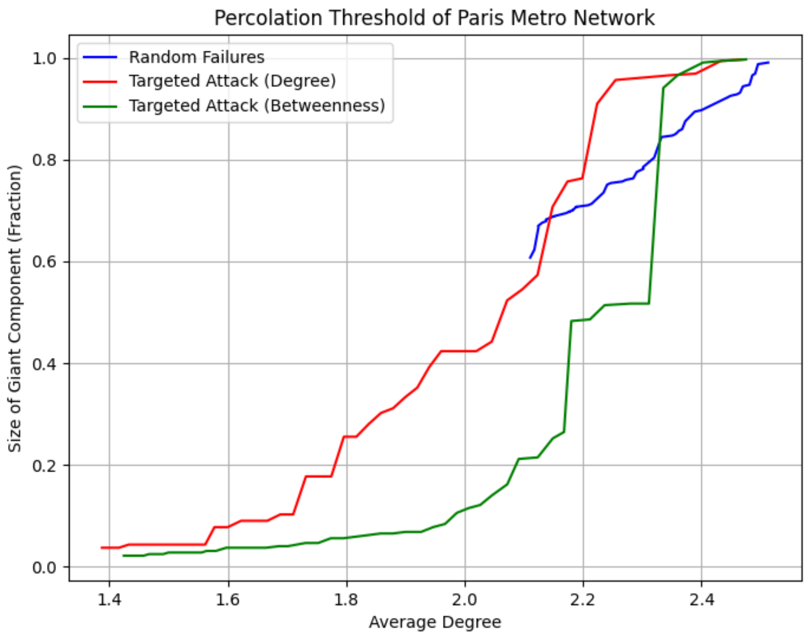


Figure 3.16: Percolation Threshold of the Paris Metro Network. The plot shows how the largest connected component size decreases as the average degree decreases due to station removals. Only 30% of the stations were attacked.

With **Random Failures** (blue), the network degrades slowly and maintains connectivity for a long period.

**Degree-Based Attacks** (Red) show that the network faces a **rapid collapse** as the highest-degree stations (major interchanges) are removed.
The largest component shrinks significantly even when the average degree is still relatively high.

**Betweenness-Based Attacks** (Green) is the most severe scenario, where removing high-betweenness stations (key transit hubs) leads to an **early and sudden** collapse of connectivity.

We can observe:

- The Paris Metro is **highly resilient** to random failures, as expected in a well-designed transportation network.
- **Hubs and transfer stations are critical**, as their removal leads to rapid network fragmentation.
- **Betweenness-based attacks** cause the most severe fragmentation, highlighting the need to reinforce alternative paths for key stations.
- The fact that only 30% of the stations were removed suggests that **even a partial attack can cause severe disruptions**, making it crucial to protect high-betweenness stations.

The concept of percolation is relevant as it helps quantify the network's breaking point. Resilience strategies include adding redundancies and alternative routes.

# Chapter 4

# Station Load Analysis

## 4.1 Station Load And Peak

To evaluate the usage patterns of the Paris Metro network, we analyzed **station load**, defined as the number of train arrivals at each station throughout the day. This measure provides a proxy for congestion levels.

### 4.1.1 Rush Hour and Daily Load Distribution

We examined the network's busiest stations during:

- **Morning Rush (7-10 AM)**: Reflects commuting patterns of workers and students.
- **Evening Rush (5-8 PM)**: Captures return trips and evening transit activity.
- **Total Daily Load**: Represents overall usage throughout the day.

Figure 4.1 presents the top 10 busiest metro stations, ranked by total daily load, with their corresponding morning and evening rush hour activity.

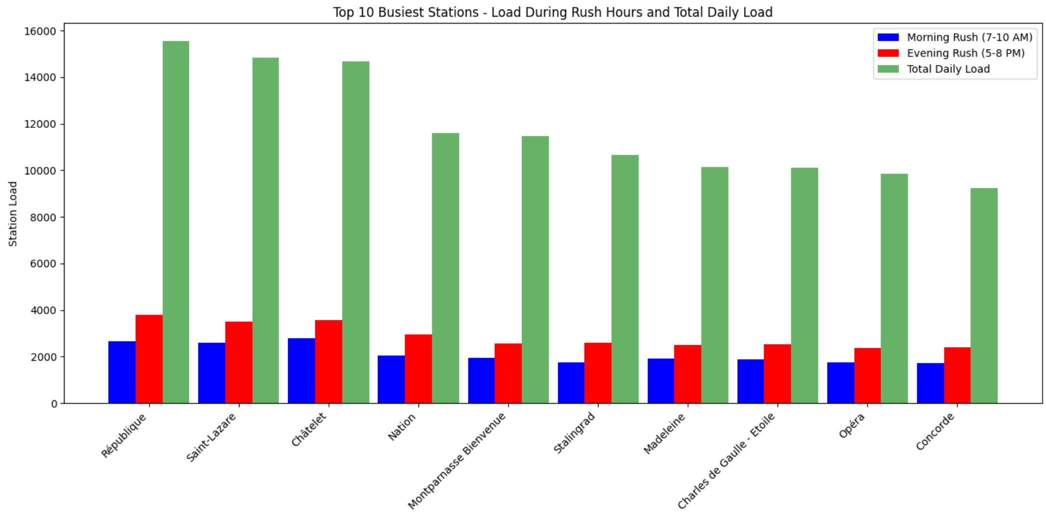


Figure 4.1: Top 10 Busiest Stations

**République, Saint-Lazare and Châtelet** are the most active stations, consistently registering high passenger arrivals in both peak periods and across the entire day.
Some stations like **Nation and Montparnasse Bienvenue** exhibit a significant total load, though their peak-hour contributions are relatively smaller, indicating steady usage throughout the day.
Stations such as **Opéra and Concorde** show more balanced peak-hour activity, with similar levels of morning and evening congestion.

### 4.1.2 Hourly Load Evolution

To gain further insight into passenger dynamics, we examined how station load fluctuates throughout the day.

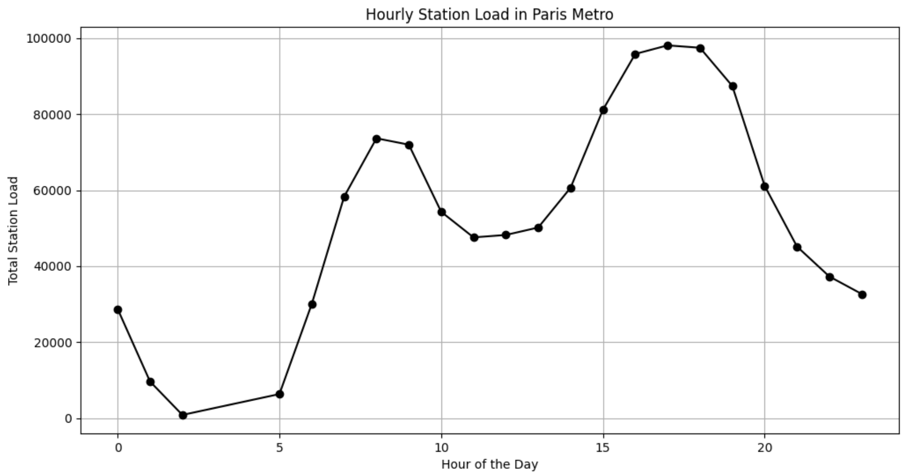


Figure 4.2: Hourly Load in the Paris Metro Network.

Figure 4.2 highlights patterns:

- A **sharp increase** in the load occurs during the morning.
- A secondary **evening peak** emerges following the typical commute.
- Midday usage remains relatively stable, while **late-night activity declines steeply after 10 PM**.

### 4.1.3 Correlation Between Load and Centrality Measures

We investigated whether stations with high **station load** also hold important structural roles in the network. Specifically, we compared station load with **Degree Centrality** and **Betweenness Centrality**.

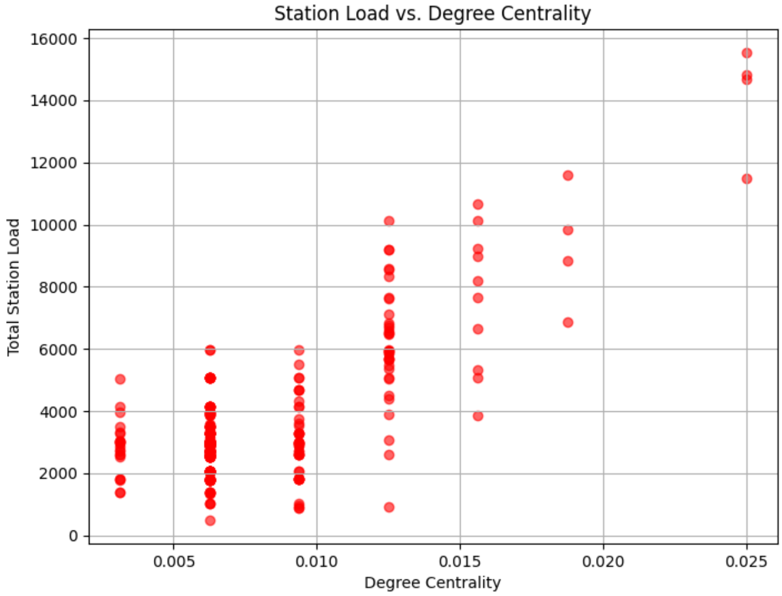


Figure 4.3: Scatter Plot: Station Load vs. Degree Centrality.

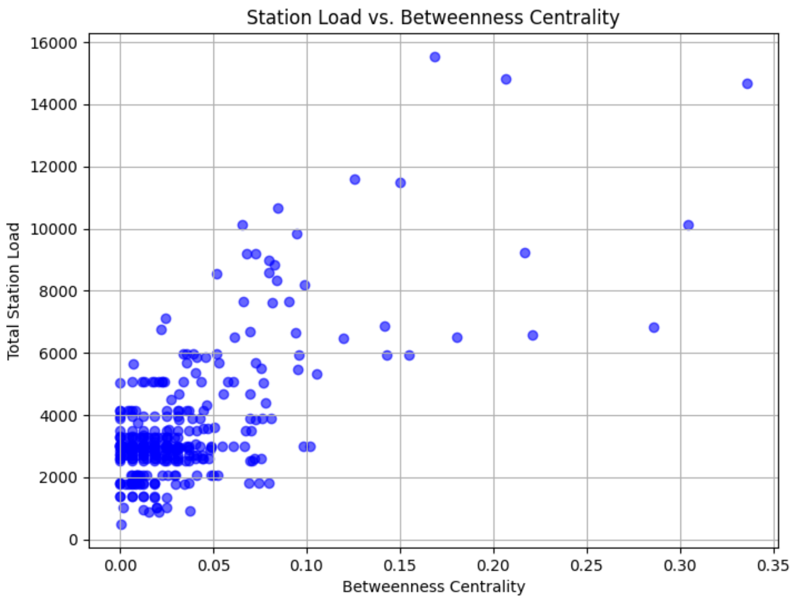


Figure 4.4: Scatter Plot: Station Load vs. Betweenness Centrality.

There is a **positive correlation** between **station load and degree centrality** (Figure 4.3), meaning that stations with many direct connections generally experience higher station load.
The correlation is even stronger with **betweenness centrality** (Figure 4.4), as high-betweenness stations act as critical transfer points.
Some **outliers** exist, certain stations have **high betweenness but relatively lower load**, likely due to their importance in connectivity rather than pure demand.
These results confirm that **highly central stations play a dominant role**, supporting the idea that interchange hubs (e.g., Châtelet) are vital for network efficiency.

## 4.2 Hourly Evolution of Station Load

To visualize how station load dynamically evolves over the course of a day, we animated the metro network where station sizes represent the relative load at each hour. This approach allows us to observe variations in network usage at different times.

### 4.2.1 Visualization of Load Evolution

To analyze the temporal variations in network load, we present snapshots of the Paris Metro network at key times of the day: early morning (2 AM, 5 AM), peak morning hours (8 AM), midday (11 AM) and evening rush hour (6 PM).

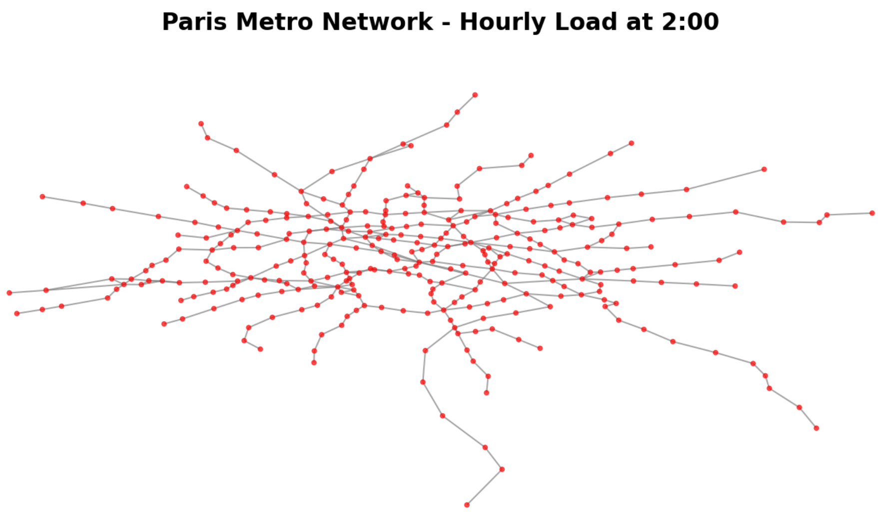


Figure 4.5: Station Load at 2 AM. Minimal network activity is observed.

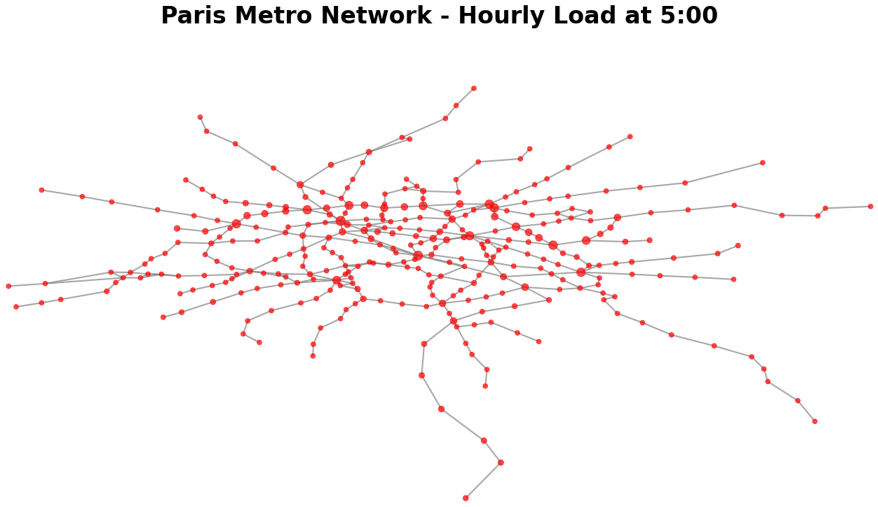


Figure 4.6: Station Load at 5 AM. Activity begins to rise slightly as services begin.

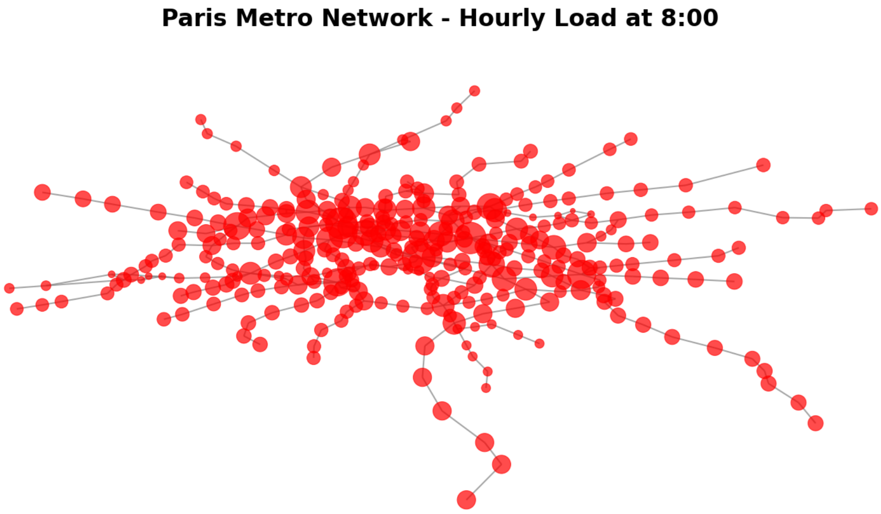


Figure 4.7: Station Load at 8 AM. A strong increase in station load is observed, particularly in central areas and interchange hubs.

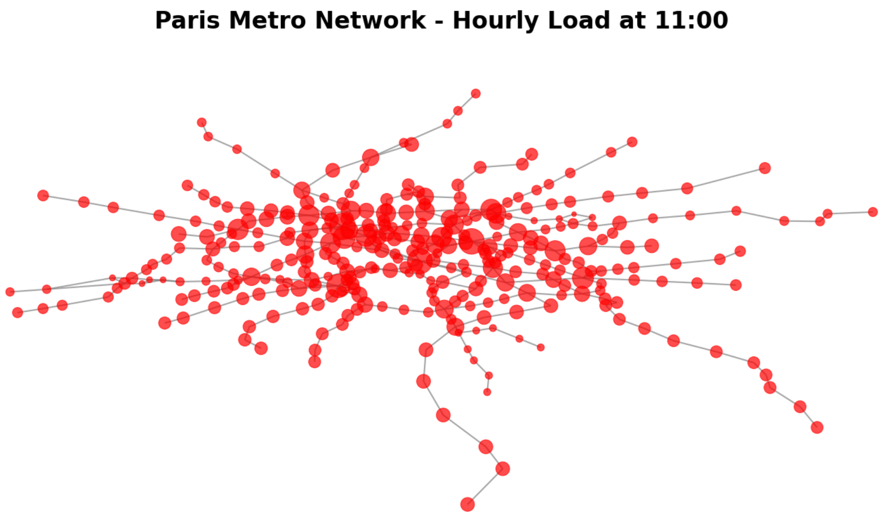


Figure 4.8: Station Load at 11 AM. The load distribution becomes more uniform.

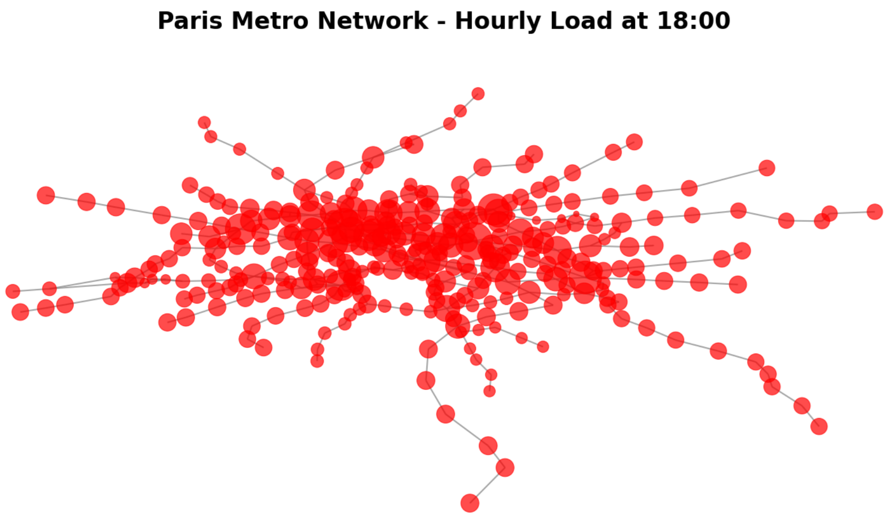


Figure 4.9: Station Load at 6 PM. A renewed increase in load is observed, mirroring the morning peak.

These snapshots reveal distinct trends:

- **Early morning (2 AM and 5 AM)**: Very low metro activity.
- **Morning peak (8 AM)**: A strong increase in station load, indicating peak commuter traffic.
- **Midday (11 AM)**: Congestion levels are reducing in major transfer hubs.

- **Evening peak (6 PM)**: Load becomes similar to the morning, concentrated in transfer hubs and central stations.

### 4.2.2 Insights from Dynamic Load Patterns

The evolution of station load aligns with expected travel patterns in a metro system:

- **Interchange hubs dominate peak hours**: Stations such as Châtelet, République and Saint-Lazare experience significant increases in size during rush hours.
- **Peripheral stations remain stable**: Less central stations do not exhibit major variations in their sizes throughout the day.
- **Off-peak periods show minimal network usage**: Particularly from midnight to early morning, most of the network remains inactive.

These observations emphasize the importance of central transfer points in sustaining network operations during peak periods. The results also suggest that resilience strategies should prioritize high-load hubs to maintain efficiency under stress conditions.

# Chapter 5

# Conclusions & Future Work

## 5.1 Conclusions

This study analyzed the **Paris Metro network** using a **graph-theoretic approach** to assess its **structural properties, vulnerability, and station load distribution**. The key findings are:

- **Network Structure**: The metro network exhibits a **sparse structure** with an **average degree of 2.52** and a **low clustering coefficient** of 0.0088, indicating that it primarily follows a **linear layout** with minimal redundant connections.
- **Critical Stations**: Major interchange hubs such as **Châtelet and Saint-Lazare** play a central role in maintaining network connectivity. Stations with high **betweenness centrality** act as key transit bridges.
- **Network Vulnerability**: The network is **highly sensitive to targeted attacks** on key interchange stations, leading to rapid efficiency loss. The removal of **Châtelet alone** causes a **7.09% drop in network efficiency**.
- **Cascading Failures**: A simulated breakdown of **high-betweenness stations** demonstrated that the metro can collapse within just **four iterations**, with a loss of **321 stations**.
- **Percolation Threshold**: The network remains robust under **random failures** but **collapses quickly under targeted attacks**. Betweenness-based failures are the most damaging, leading to early fragmentation.
- **Station Load Analysis**: High-load stations coincide with high centrality, confirming their role as essential hubs. **Rush hour peaks** concentrate around major transfer points, and load distribution follows expected commuter patterns.

Overall, these insights highlight the Paris Metro's **structural strengths and vulnerabilities**, emphasizing the importance of **reinforcing key hubs and optimizing alternative paths** to enhance network resilience.

## 5.2 Future Work

While this study provides a detailed assessment of the Paris Metro network, the following aspects could be explored in the future:

- **Passenger Flow Modeling**: Integrating ridership data to observe congestion patterns under different scenarios, including disruptions and peak-hour variations.
- **Expanding the Network**: Expanding the study to include **RER, trams, and buses** to understand interconnections between metro lines and other public transport systems.
- **Resilience Strategies**: Testing **mitigation measures**, such as additional station links, to prevent cascading failures.