Revealing Disparities in Public and Digital Infrastructure of Developing Countries

Abhishek Kumar Mishra, Sachit Mishra

Revealing Disparities in Public and Digital Infrastructure of Developing Countries

Abhishek Kumar Mishra*, Sachit Mishra†
*INSA-Lyon, Inria, University of Lyon, CITI Lab.
†IMDEA Networks Institute, Madrid

***Abstract*—This work, for the first time in the literature, investigates and establishes the disparities between public infrastructure and digital inequality by employing NetMob 2024 datasets. We first define the inference of the state of these infrastructures before quantifying infrastructure inequality using the Gini coefficient in developing countries, i.e., Colombia, India, Indonesia, and Mexico, and unveil the synergy between public and digital infrastructure by leveraging Spearman's rank correlation. The categorization of origin-destination NetMob 2024 data into half-yearly chunks uncovers the varying evolution of the public and digital infrastructure. The results reveal the existence of substantial disparities and temporal variations between the recorded year's halves.**

I. Introduction

In countries worldwide, the governments are putting immense effort into shaping public policy and investment aiming to reduce the disparity and inequality in education level, and urbanization of cities to achieve equality among people living in different regions of the country such as digital equality. In developing countries such as Colombia, India, Indonesia, and Mexico, this particularly holds great importance for development to achieve the development of the state with equality.

Studies of literature examined the level of urbanization based on the availability of public infrastructure [1] and explored the impact of new wireless technology based on the advancement of cities of the country[2].

All of these works separately focus on urbanization and digital imbalance. To the best of our knowledge, this work is the first to unveil the relationship between public and digital infrastructure in developing countries. We leveraged the NetMob 2024 dataset to come up with following key contributions:

- We characterize and quantify the state of the public and the digital infrastructure.
- We reveal the level of disparity of digital infrastructure in developing countries.
- We establish a connection between public infrastructure and digital infrastructure.
- We unveil that the evolution of disparities in the public and the digital infrastructure shows opposite trends.

II. Methodology

A. Characterizing Public and Digital Infrastructure

To analyze infrastructure disparities, we first characterize public and digital infrastructure. Public infrastructure is represented by the average trip speed, while digital infrastructure is represented by the number of digital interactions.

1) Public Infrastructure: The average trip speed S (in km/h) for each trip is calculated as:

$$S = \frac{L}{T} \tag{1}$$

where L is the trip length in kilometers and T is the trip duration in hours. For each geohash g within a country, we compute the average speed $\overline{S_g}$ as:

$$\overline{S_g} = \frac{1}{n_g} \sum_{i=1}^{n_g} S_i \tag{2}$$

where n_g is the number of trips originating from geohash g.

2) Digital Infrastructure: The digital infrastructure is characterized by the number of digital interactions D recorded during each trip. For each geohash g, we compute the average number of digital interactions $\overline{D_g}$ as:

$$\overline{D_g} = \frac{1}{n_g} \sum_{i=1}^{n_g} D_i \tag{3}$$

where D_i is the number of interactions for trip i and n_g is the number of trips originating from geohash g.

B. Metrics for Measuring Disparity

To quantify the disparities in public and digital infrastructure, we utilize the Gini coefficient and Spearman's rank correlation coefficient. The Gini coefficient measures inequality within a distribution, allowing us to assess the disparity in average trip speeds and average digital interactions. This metric is particularly useful for understanding the extent of inequality in infrastructure access and usage across various regions in the country like cities, suburbs, villages, etc.

On the other hand, Spearman's rank correlation coefficient measures the strength and direction of the monotonic relationship between two variables, providing insights into the correlation between average trip speeds and average digital interactions. By employing these two metrics, we can comprehensively evaluate both the level of inequality and the nature of the relationship between public and digital infrastructure metrics.

C. Data Processing

The datasets used in this study include origin-destination data for various countries, comprising metrics such as the number of digital interactions and average trip lengths. We aggregated these metrics by country and by halves of the year to facilitate temporal analysis.

The data for each country were categorized by halves of the year, denoted as $2019H1$ (January to June 2019) and $2019H2$ (July to December 2019). This categorization allowed us to analyze the progress of the disparity metrics over the course of the year.

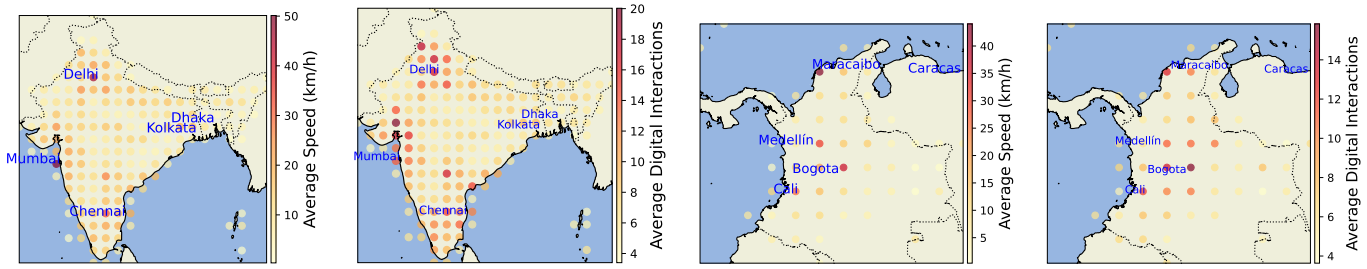
III. Results and Discussion

In this section, we illustrate, reason, and discuss the results of disparity in public and digital infrastructure using the metrics mentioned over the two halves of 2019.

A. Illustrating the State of Infrastructures

Figure 1 illustrates the state of public and digital infrastructure in India and Colombia, for example, when using *geohash3* (availability of *geohash5* in the final version will make locations more precise).

In India, big cities such as Mumbai and Delhi show a significant growth in public infrastructure (higher average speed of trips), as seen in Figure 1a. These cities exhibit high speeds, but much lower speeds are observed in outer regions, indicating disparities in public transportation infrastructure. Similarly, Figure 1b reveals a high concentration of digital interactions in the same big cities along with



(a) Public Infrastructure (b) Digital Infrastructure (c) Public Infrastructure (d) Digital Infrastructure

Fig. 1: Illustrating the state of Public and Digital Infrastructure in India (Figure 1a and 1b) and Colombia (Figure 1c and 1d)

the tech hubs like Bengaluru and Hyderabad, while other regions show much lower levels of digital engagement.

In Colombia, Figure 1c shows that Bogotá and Medellin have high travel speeds, but not in most of the other parts of the country, reflecting the disparities in public infrastructure across the country. The digital infrastructure, depicted in Figure 1d, also shows significant disparities, with higher digital interactions concentrated in major cities compared to rural areas.

B. Disparity in Infrastructures

Figure 2 shows the Gini coefficients for digital interactions and public infrastructure in the four countries. The Gini values for public infrastructure in each country are relatively high, unveiling the differences in infrastructure advancement between small villages and towns compared to big cities. In general, digital infrastructure is progressing much uniformly when compared to the public infrastructure in the developing countries.

Although the coefficients are very high for Indonesia and Colombia, the corresponding Gini index values for the digital infrastructure are lower than those for India and Mexico, demonstrating the presence of a much greater *digital divide*[2] in these countries.

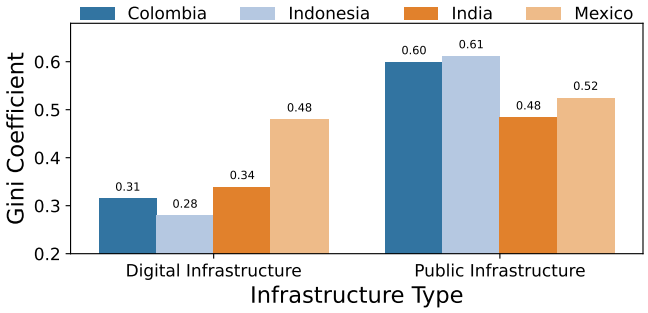


Fig. 2: Gini Coefficients for Digital Interactions and Public Infrastructure (Average Speed) across developing countries.

C. Correlation between Infrastructures

We calculate Spearman correlation coefficients to understand the relationship between public and digital infrastructure. Table I presents the high correlation for all countries indicating that areas with higher digital interactions also tend to have higher trip speeds, suggesting a positive relationship between digital and public infrastructure.

Statistically, India shows a highly correlated existence of good public and digital correlation with a correlation coefficient of 0.80, as visually attested in Figures 1a and 1b also.

Country	Colombia	Indonesia	India	Mexico
Correlation Coefficient	0.76	0.73	0.80	0.73

TABLE I: Spearman Correlation Coefficient between Digital and Public Infrastructure

D. Evolution of disparities

To understand the impact of time on disparity, we analyze the disparity in digital and public infrastructures for India, Indonesia, and Mexico. Colombia is excluded due to incomplete half-yearly data for the current hash 3 (We plan to update this in the final submission upon accessing the complete datasets). Figure 3 illustrates the Gini coefficients for digital interactions and public infrastructure (average speed) across the first and second halves of 2019.

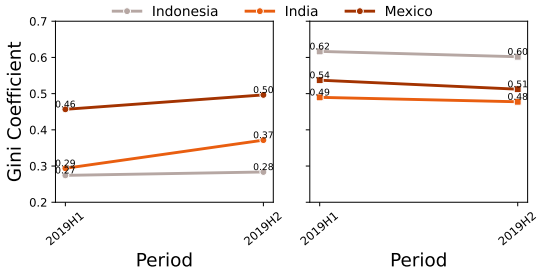


Fig. 3: Half-Yearly Disparity in Digital (left) and Public (right) Infrastructure across India, Indonesia, and Mexico.

The results clearly indicate that, over time, the disparity in public infrastructure decreases for all countries. However, the disparity in digital infrastructure increases for all countries, leading to the conclusion that a digital divide exists and is widening over time, even as the public infrastructure gap closes. India exhibited the highest disparity in digital infrastructure while for Indonesia, this digital divide has increased but is relatively lower than in others, prompting us to further explore with socioeconomic indicators to understand the cause of the smaller digital divide there. The trends indicate a need for targeted interventions to reduce disparities and promote equitable distribution of both types of infrastructure.

IV. Conclusion

The analysis presented in this paper underscores the importance of using robust statistical metrics to quantify and monitor infrastructure disparities. Combining Gini coefficients and Spearman's correlation provides a comprehensive framework for understanding the complex interplay between digital and public infrastructure. Future research should explore the underlying causes of these disparities and develop targeted interventions to address them.

References

[1] C. Cottineau, E. Hatna, E. Arcaute, and M. Batty, "Diverse cities or the systematic paradox of urban scaling laws," *Computers, environment and urban systems*, vol. 63, pp. 80–94, 2017.

[2] S. Mishra, Z. Smoreda, and M. Fiore, "Second-level digital divide: A longitudinal study of mobile traffic consumption imbalance in france," in *Proceedings of the ACM Web Conference 2022*, pp. 2532–2540, 2022.