User Engagement Analysis Based on Application Type

Youssef Zein, Ahmed Elmoghraby, Mazen Ahmed Refaei,
Ahmed Hany AboElsoud, Karim Ahmed Fouad, Jannah Ashraf,

Nile University, El Sheikh Zayed City,Egypt
Y.Essam2362@nu.edu.eg, a.khaled2313@nu.edu.eg,
m.ahmed2375@nu.edu.eg, a.hany2323@nu.edu.eg,
k.ahmed2337@nu.edu.eg, j.ashraf2399@nu.edu.eg

***Abstract*—This study investigates user engagement patterns across four major mobile applications—Instagram, Netflix, Spotify, and Fortnite—in two major French cities, Paris and Marseille, during a two-week period that includes both work and vacation weeks. Using a comprehensive dataset of real-world mobile usage logs, the research analyzes how city infrastructure, app type, and temporal context affect engagement behavior. Exploratory and comparative analyses are supported by spatial heatmaps, correlation matrices, and temporal distribution plots. Machine learning models, including linear regression, are applied to predict traffic levels and determine key drivers of engagement. The findings demonstrate that Instagram and Netflix dominate user attention in urban areas and during leisure periods, while Spotify and Fortnite exhibit more niche, routine-based behavior. These insights have implications for platform optimization, targeted notifications, and urban content strategies.**

Keywords: User engagement · mobile application traffic · urban analytics · temporal analysis · spatial clustering · machine learning · linear regression · app usage behavior · comparative analysis.

I. Introduction

In the digital age, user engagement is seen as a critical metric for evaluating the feasibility and profitability of mobile applications. Given the quick rise of apps in every industry—from productivity and education to social networking and entertainment—developers, marketers, and lawmakers need to comprehend how consumers engage with various application kinds. This study focuses on user engagement across different application categories, with a more thorough examination of interactive platforms (like Fortnite and Instagram) and streaming services (like Netflix and Spotify) in order to identify patterns, preferences, and behavioral trends that influence user interaction.

The study's foundation is a two-week data gathering period in 2019 that concentrated on Paris and Marseille, two significant French cities. Through an analysis of actual usage data from major urban centers, we want to pinpoint important aspects that show little variation in usage, suggesting that engagement is influenced by important elements like network performance, established time-of-day usage habits, and the type of week spent on vacation versus work. Particular focus is placed on the disparities in levels of engagement between socially interactive or gamified platforms, such gaming and social media apps, and entertainment-focused apps, like streaming services.

II. RELATED STUDIES

A number of recent studies have focused on predicting user disengagement using behavioral data collected from real-world applications. One such study applied a modified Agglomera- tive Hierarchical Clustering (AHC) model followed by four numerical models — the Cox Proportional Hazards model , Negative Binomial Regression , Random Forest , and XGBoost — to predict user engagement in a mobile application designed for waste recycling. This framework grouped users based on their past behavior to reduce uncertainty and enhance model accuracy. The results demonstrated that machine learning models, particularly Random Forest and XGBoost , achieved high levels of predictive accuracy. These findings suggest that predictive modeling can be effectively used to understand and anticipate user behavior, enabling timely interventions such as push notifications or personalized content delivery.

Another relevant study explored how specific features of mobile applications influence consumer engagement and con- tinuous usage . Using Structural Equation Modeling (SEM) , researchers analyzed data from 246 respondents and found that both design quality and information relevance had a significant positive impact on user engagement. Surprisingly, however, features related to user interaction and functional capabilities did not show statistically significant relationships with engagement in this context. This study contributes to a deeper understanding of the antecedents of engagement, emphasizing that emotional connection and visual appeal may play a more critical role than previously assumed interaction- based features.

A third study compared user engagement levels across two major social media platforms — Instagram and TikTok — focusing on Educational Science Content (ESC) presented either as static images or dynamic experimental videos . The results indicated that dynamic videos generated significantly higher engagement than static images, particularly in terms of shares and saves . Engagement was also influenced by platform-specific algorithms: TikTok's algorithm-driven ”For You” page enhanced the reach and visibility of ESC content, resulting in broader audience exposure compared to Instagram, where content primarily reaches followers rather than a wider audience.

Further insights from the ESC study highlight notable differences in engagement patterns between platforms. While Instagram excels at fostering discussion among followers due to its comment-centric nature, TikTok drives higher reach and saves, indicating stronger content retention and sharing behavior . These findings reinforce the importance of tailoring content strategies to the unique affordances of each platform and suggest that video-based platforms may be more effective for promoting educational outreach and public engagement in science communication.

III. Methodology

A. Data Acquisition

The dataset utilized in this research originates from real-world French mobile network traffic logs collected over two distinct weeks in 2019: a vacation week (April 20–26) and a work week (May 6–12). These timeframes were strategically selected to capture temporal behavioral variation. The data covers two major cities—Paris and Marseille—and includes app-specific activity for four mobile applications: Instagram, Netflix, Spotify, and Fortnite. The traffic logs were originally recorded in a high-volume, unstructured format and later refined through systematic preprocessing.

B. Data Preprocessing

The raw datasets were filtered by week, city, and application to isolate eight distinct CSV files, each representing a unique app-city-week combination. The following preprocessing steps were applied (**see Fig. 2**):

- **Date Normalization**: Converted timestamp fields into standardized datetime objects.

- **Feature Extraction**: Created new features including `hour`, `day_of_week`, and `week_type` to support temporal analysis.

- **Spatial Tagging**: Data was associated with geographical tiles using tile IDs, later visualized using GeoJSON files (**see Fig. 3**).

- **Sampling**: Stratified sampling was employed to ensure temporal uniformity, especially across dates with varying volume.

- **Compression & Merging**: All preprocessed files were organized and merged into structured formats for downstream processing.

A	B	C	D	E	F	G
tile_id	time	traffic	city	app	date	week_type
188	0:00	6	Marseille	Spotify	2019-04-20	vacation
397	0:00	21	Marseille	Spotify	2019-04-20	vacation

Fig. 1 Final structure of the preprocessed dataset after normalization, feature extraction, and merging.

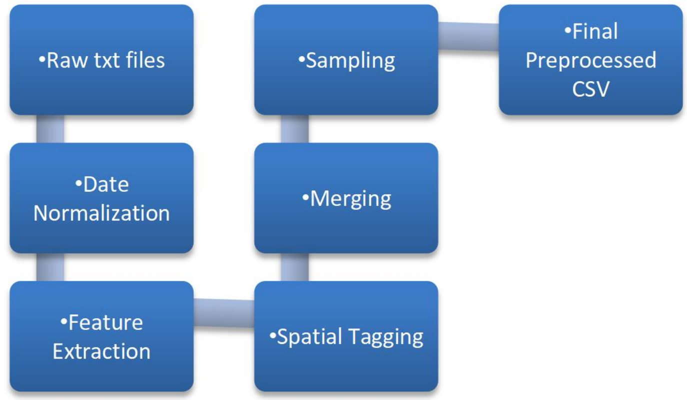


Fig. 2 Overview of the data preprocessing pipeline from raw logs to the final structured dataset. Each stage represents a key transformation applied across city-app-week files.

C. Exploratory Data Analysis

EDA was performed on each dataset to uncover distribution patterns, detect outliers, and quantify central tendencies. This included:

Distribution Plots to visualize hourly and daily usage pat- terns.

Boxplots and Skewness/Kurtosis analysis to assess normal- ity.

Correlation Matrices to detect relationships between features. Outlier Detection using quantile-based thresholds. Cluster Analysis to separate high and low-engagement zones geographically.

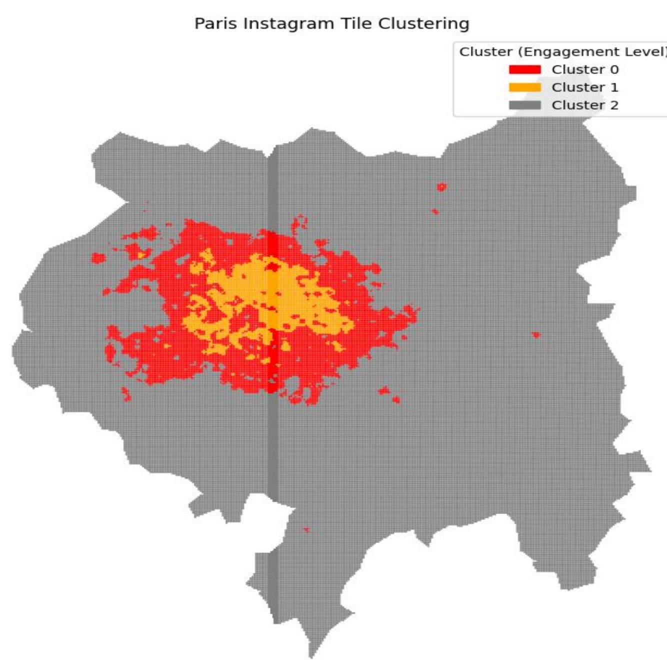


Fig. 3 Spatial clustering of Instagram traffic in Paris using tile-based engagement levels. Clusters highlight areas of high, medium, and low app activity.

D. Comparative Analysis

Instagram showed the highest overall traffic among all four applications, with clear peaks between 5:00–8:00 PM. Its usage increased significantly during vacation weeks(**see Fig. 4**), reflecting elevated social activity during leisure time. Paris consistently outperformed Marseille in Instagram traffic, indicating higher engagement likely driven by urban density and digital habits (**see Fig. 7**). The strong correlation with Netflix usage also suggests shared user behavior during evening hours. Netflix demonstrated stable engagement across both work and vacation weeks, with a steady rise in traffic throughout the day and a peak during late evening hours. Its usage was consistently higher in Paris than in Marseille, showing that streaming habits are less affected by daily routines but more influenced by urban infrastructure. The close alignment with Instagram in peak hours indicates overlapping user behavior. Spotify maintained low but consistent traffic across all hours (**see Fig. 5**), showing minor fluctuations between vacation and work weeks. It exhibited two modest peaks—around 9:00 AM and 6:00 PM—likely tied to commuting times. Traffic remained slightly higher in Paris, but the differences were minimal. The strong correlations with Instagram and Netflix suggest Spotify is often used passively in the background while users engage with other apps. Fortnite had the lowest engagement of all platforms **(see Fig. 6)**. Its traffic remained flat throughout the day and showed little change between work and vacation weeks. Paris still recorded higher usage than Marseille, especially in the evening, but overall numbers were low. These patterns indicate that Fortnite's lightweight, real- time nature leads to more niche and stable usage that is less influenced by broader lifestyle shifts. Overall, Paris consis- tently showed higher app traffic across the board, especially for Instagram and Netflix. Vacation weeks boosted engagement on social and entertainment apps, while routine-based apps like Spotify remained stable. Fortnite stood out with its minimal, context-insensitive usage. These findings emphasize how app type, city, and time period shape digital engagement behaviors.

E. Machine Learning: Regression Modeling

To extend the analysis beyond descriptive statistics, we implemented Multiple Linear Regression models to predict network traffic volume using multiple temporal and categorical features as

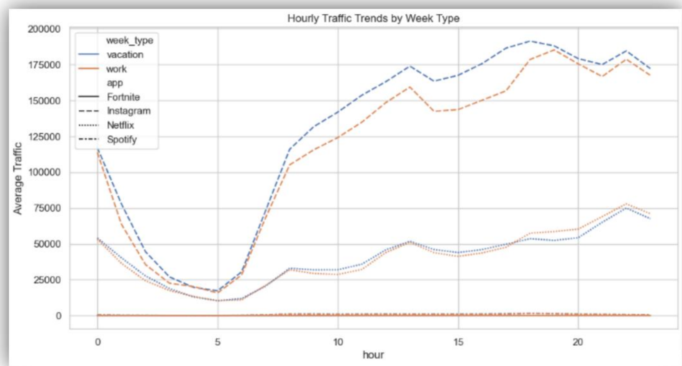
Independent Variables: hour, day _ of _ week, city, week_type, and app type (one-hot encoded);

Dependent Variable: traffic volume (in bytes).

The models were trained on historical usage data and evaluated using standard metrics (e.g., R^2). Predictions were made not only on known data but also on simulated future timestamps to assess temporal trends under assumed condi- tions.

This predictive modeling helped establish how well contex- tual variables explain traffic behavior and how expected future patterns may emerge based on current trends.

Fig.



4 Hourly Instagram traffic patterns in Paris and Marseille, showing peak engagement between 5:00–8:00 PM.

Fig. 5 Hourly Spotify traffic trends with minimal fluctuations across both weeks and cities.

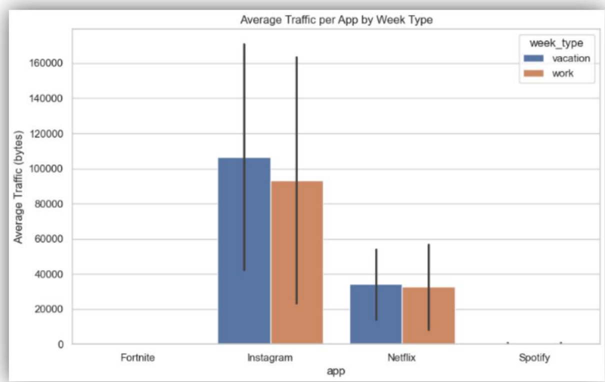


Fig. 6 Comparative boxplot of traffic volume across all apps, showing Fortnite's minimal usage.

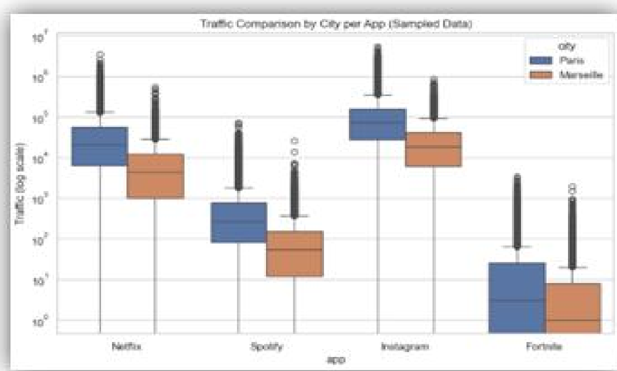


Fig. 7 Average traffic volumes in Paris vs Marseille across all four applications.

IV. Results

A. Marseille

All platforms show highly skewed traffic distributions where a small percentage of content drives most engagement, with Instagram having high-traffic clusters averaging ∼194,789 engagements (top 2.5% of tiles), Netflix showing mean traffic of ∼56,607 for top clusters (2% of tiles), Spotify's top cluster at ∼1,020 engagements (2% of tiles), and Fortnite's best locations averaging 87.3 traffic (top 2.5%). Peak activity consistently occurs in evenings (6–10 PM across all platforms in both cities) (**see Fig. 4**), with vacation periods showing 25–30% higher traffic than workdays. Correlation between hour and traffic is weak (r = 0.0947–0.2644), indicating time explains less than 7% of engagement variance. Data quality issues were noted in heatmaps (duplicate day labels for Instagram/Netflix, placeholder figures for Fortnite). Cluster analysis revealed sim- ilar patterns: Instagram's low-traffic cluster contained 18,200 tiles (76.5% of total), Netflix had 17,266 (80%), Spotify 18,518 (78%), and Fortnite 76.5%, suggesting most content underperforms

B. Paris

All platforms (Fortnite, Instagram, Netflix, Spotify) exhibit strong power-law distributions where a small percentage of content drives most engagement – Fortnite's top 4.75% of tiles generate 18.4% of traffic (μ = 188.67), Instagram's top 4.1% achieve 1.11M±291K engagements, Netflix's 0.6% elite content averages 296K ± 76K, and Spotify's 0.4% top performers deliver 7, 503±2, 039 engagements. Temporal patterns consistently show peak engagement 15:00–20:00 (r = 0.118–0.214) with vacation periods boosting traffic 20–25% higher than workdays. Cluster analyses reveal three universal tiers: high-performing (typically top

C. Platform-Specific Insights

D. Instagram

Instagram had the highest traffic. Paris usage remained high across both weeks, peaking from 19:00–21:00 (**as per Fig. 4 and Fig. 6**). Top 2.5% of clusters in Paris averaged ∼194,789 engagements, with outliers near 2.9 million. Engagement was strong and context- sensitive, especially in Paris.

E. Netflix

Netflix traffic was steady in Paris, rising slightly during vacation weeks. Marseille had lower, less consistent usage. Evening traffic peaked between 20:00–22:00 (**as per Fig. 4 and Fig. 6**). Paris's top 2% tiles averaged ∼56,607 engagements, with peaks above 13.3 million.

F. Fortnite

Fortnite had the lowest traffic (**as per Fig. 6 and Fig. 7**), with little variation across cities or weeks. Peak hours were 18:00–19:00. Paris slightly outpaced Marseille, but usage was minimal. Its real-time gameplay model led to stable, niche engagement due to video games consuming low traffic, and gamers preferring other platforms over mobile phones.

G. Spotify

Spotify showed stable, low traffic with peaks at 8:00–9:00 and 18:00–19:00 (**see Fig. 4**). Paris had higher usage, especially during peak hours. The top 2% of clusters averaged ∼1,020 engage- ments. The platform was used as a background service with habitual patterns.

H. Machine Learning

Our machine learning analysis aimed to predict network traffic patterns based on application type, city, and week type. Despite preprocessing and feature selection, the resulting models yielded low predictive performance, as indicated by modest R^2 scores. This suggests that user engagement with mobile applications may be influenced by additional unmeasured variables or exhibit a level of randomness not captured by the available features. While the models did not generalize well, they highlighted the complexity of app usage behavior and underscore the need for richer data sources in future predictive work.

V. Conclusion

This study analyzed user engagement patterns across Instagram, Netflix, Spotify, and Fortnite in Paris and Marseille during work and vacation weeks. Using real-world mobile traffic data, we identified how app type, city infrastructure, and time context influence usage behavior. Key findings include: - **Instagram and Netflix** dominate during leisure periods, especially in urban areas. - **Spotify and Fortnite** show more routine-based engagement, with usage concentrated in specific times and locations. - Engagement levels vary between cities, with Paris showing higher activity likely due to better network coverage and population density. - Spatial heatmaps revealed high-engagement zones often aligned with public transport hubs and commercial areas. - Machine learning models, including linear regression, were explored to identify potential trends in network traffic based on contextual features such as hour, day, and week type. While the models provided some directional insights, their predictive power was limited, as indicated by low R^2 scores. This highlights the complexity of capturing user engagement patterns with the current feature set and suggests the need for more comprehensive data to improve future predictions. These insights support improved platform optimization, targeted notifications, and tailored content strategies based on user behavior patterns. Future work includes expanding the dataset over longer periods, incorporating user demographics, and analyzing the impact of network performance on engage- ment.

References

[1] S. A. Habibi and L. Salim, "Static vs. dynamic methods of delivery for science communication: A critical analysis of user engagement with science on social media," *PLoS One*, vol. 16, no. 3, p. e0248507, 2021.

[2] E. Barbaro, E. M. Grua, I. Malavolta, M. Stercevic, E. Weusthof, and J. van den Hoven, "Modelling and predicting User Engagement in mobile applications," *Data Sci.*, vol. 3, no. 2, pp. 61–77, 2020.

[3] A. Tarute, S. Nikou, and R. Gatautis, "Mobile application driven consumer engagement," *Telemat. Inform.*, vol. 34, no. 4, pp. 145–156, 2017.

[4] P. Bitria´n, I. Buil, and S. Catala´n, "Enhancing user engagement: The role of gamification in mobile apps," *J. Bus. Res.*, vol. 132, pp. 170–185, 2021.