This is the author's accepted version of the article. The final version published by IEEE is M. Pirri, D. Madariaga, Z. Smoreda and M. Fiore, "Disentangling Simultaneous Spatiotemporal Events in Mobile Network Traffic Data: The Case of the 2025 UEFA Champions League Final" Network Traffic Measurement and Analysis Conference, 2026, pp. TBP, doi: TBP.

Disentangling Simultaneous Spatiotemporal Events in Mobile Network Traffic Data: The Case of the 2025 UEFA Champions League Final

Máximo Pirri*§, Diego Madariaga*, Zbigniew Smoreda‡, Marco Fiore*
*IMDEA Networks Institute, Spain, ‡Orange Research, France, §University of Alcalá, Spain
{maximo.pirri, diego.madariaga, marco.fiore}@networks.imdea.org, zbigniew.smoreda@orange.com

***Abstract*—Large-scale social gatherings often generate distinctive patterns of activity in mobile networks due to the concentration of a substantial amount of users. While prior work has shown that specific individual events leave clear identifiable signatures in aggregate network traffic, disentangling the effects generated on mobile data usage by concurrent and co-located happenings entails significant added complexity. In this paper, we present a methodology to detect different but coinciding mass manifestations by unraveling distinct traffic signatures associated with the coexisting behaviors, thus allowing to monitor the spatiotemporal evolution of each event in isolation. We demonstrate the effectiveness of our approach in a practical use case, *i.e.*, the combination of cheerful celebrations and social unrest episodes that accompanied Paris Saint-Germain's 2025 UEFA Champions League victory in the city of Paris, France. Using mobile network traffic measurements collected by a leading network operator, we successfully separate the mobile traffic consumption patterns of peaceful partying crowds from those of rioters that confronted local police forces, and reconstruct the directional flows of the different groups across the city.**

***Index Terms*—Mobile network traffic, application usages, social events, spatiotemporal monitoring, UEFA Champions League.**

I. Introduction

Mobile networks play a key role in our life, enabling a wide range of essential services such as live news, instant messaging, spatial navigation, or expedite financial transactions to name a few. When large groups of people gather at specific locations for a particular event, their collective activity can generate a recognizable footprint in the network usage. Once identified and if strong enough, such a footprint can be leveraged to monitor the events and generate insights, *e.g.*, on the quantity of participants or their mobility over time, which are useful to better understand, organize for or respond to the special circumstances. Recent work has explored the relationship between large social events and the digital world, showing that, *e.g.*, natural disasters [1], emergency situations [2], or organized mass manifestations [3] leave distinguishable and measurable impressions in the mobile network traffic.

While existing studies focus on large social events that occur in isolation, manifestations of different nature may happen at the same or close-by locations and overlap in time. For example, national holidays such as July 4 in the US often feature a combination of parades, public concerts, and gatherings at landmark monuments occurring in the same urban areas; major sport events such as the Olympic Games bring together multi-sport competitions, fan zones, live screenings, and celebrations; major cultural festivals such as the Edinburgh Festival Fringe may offer a wide range of concurrent activities, such as multiple closely located stages oriented to different audiences, interactive art exhibits, and marketplaces; beyond flagship events, everyday city life may also involve an overlay of diverse activities at specific locations where multiple interests coincide due to co-located infrastructures or venues.

Simultaneous social events muddle and hide their respective effects in the digital domain, making the problem of identifying their individual network traffic footprints significantly harder to solve. In this paper, we propose a first methodology to disentangle the signatures of concurrent large events in mobile service consumption, hence allowing monitoring the independent but superposing spatiotemporal dynamics of the happenings from network measurement data. The design and validation of our approach yield the following contributions.

- We propose a methodological framework that isolates and reconstructs the dynamics of multiple overlapping spatiotemporal events using concise traffic consumption signatures, selected through robust feature selection and derived from minimal sets of ground truth observations.
- We demonstrate the viability of our methodology with real-world measurements in a practical use case, *i.e.*, the combination of cheerful celebrations and social unrest episodes that accompanied Paris Saint-Germain's 2025 UEFA Champions League victory in the city of Paris, France.
- Our results show that although signatures and footprints of the mass manifestations are similar and concurrent in space and time, movements from the peaceful crowds can be differentiated from those involved in altercations in scenarios not seen during framework training.

Overall, our study reveals that mobile network measurements can be used by Mobile Network Operators (MNOs) to plan for similar future events by improving network capacity planning and ensuring reliable service during periods of high user density. Additionally, local authorities can leverage insights derived from our proposed methodology to anticipate, prevent, or control large masses from disrupting peaceful celebrations.

978-3-903176-83-6 ©2026 IFIP

We remark that this research and the resulting methods preserve the privacy of the data subjects, since they rely on depersonalized aggregate measurements over large crowds and do not allow identifying or tracking individual or small groups of rioters—see Appendix A.

II. Background and Motivation

Our study falls in the area of human activity analysis with digital data, where previous works have explored the relationships between events occurring in the physical world and the impressions they leave in the digital domain as reviewed in Section II-A. Still, existing approaches are not designed to disentangle concurrent events, such as the peaceful celebrations and violent disorders that characterize our reference use case presented in details in Section II-B.

A. Related work

The analysis of large-scale human activity through mobile network data has shown that collective behaviors leave measurable footprints in digital infrastructures [4]. Prior work has demonstrated the early detection of critical events through anomalies in mobile traffic [2], including natural disasters such as floods [1] and earthquakes [5], as well as epidemics [6], [7]. These studies highlight the strong connection between human activity and network usage dynamics.

Through this analysis, frameworks were developed with the ability to identify and or characterize different types of events. Large gatherings such as sports or leisure activities exhibit distinctive spatiotemporal patterns [8], while mobile data has also been used to infer urban dynamics such as road congestion [9] and to detect public manifestations. Moreover, application-level usage has been shown to provide predictive signals for other phenomena, including elections [10], demonstrations and concerts [11], as well as protests and social unrest [12]–[14].

Beyond mobile sensing, the notion of cascading events has been studied in the context of disaster risk, where sequences of interdependent phenomena amplify overall impact [15]. This perspective highlights the complexity arising from interacting or concurrent events, although it has not been extensively explored in mobile network analysis. However, all these existing works consider events in isolation, where their signatures can be clearly identified. The problem of disentangling overlapping events that are co-located in space and time remains largely unexplored, motivating the approach proposed in this work.

Classical decomposition and clustering techniques commonly used in signal processing and unsupervised learning for latent structure discovery, dimensionality reduction, or data partitioning are not well-suited to our problem setting. These methods are typically designed under assumptions such as statistical independence, low-rank structure, or mutually exclusive partitioning, which are not compatible with our framework. In contrast, our objective is to attribute observed spatiotemporal variations in mobile service consumption to multiple known large-scale events, allowing for overlapping event contributions under a supervised formulation.

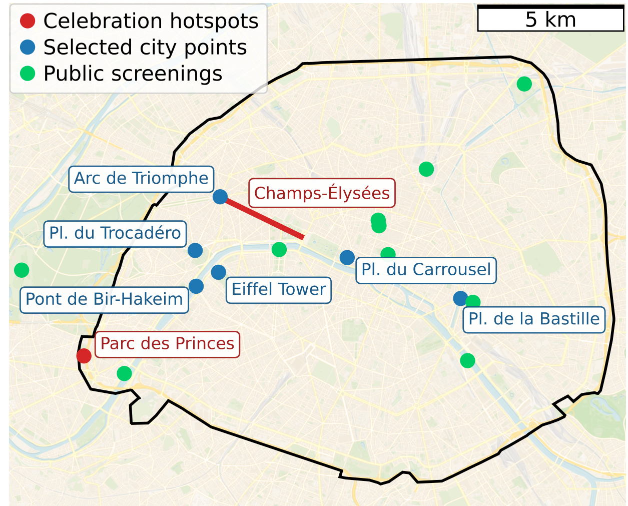


Fig. 1: Key locations in Paris related to our reference use case.

B. Reference use case

A practical scenario that is representative of the concurrent event settings mentioned above occurred in Paris, France, in the aftermaths of the 2025 European Football Associations (UEFA) Champions League semifinal and final—the closing games of the most prestigious football tournament in Europe contested by top-division clubs in the continent—when large crowds engaged in celebration and riotous activities coincided.

Specifically, the second leg of the semifinal was held in Paris at Parc des Princes stadium, home of Paris Saint-Germain (PSG)—the main local football team and one of the semifinalists. The venue was nearly at full capacity, with over 48,000 spectators attending the match to witness PSG's victory and qualification for the final [16]. Following the match, large celebrations erupted around the stadium and along the Champs-Élysées, one of the city's main avenues as shown in Fig. 1. As the evening progressed, parts of the celebrations escalated into unrest, requiring police intervention. Reports indicated incidents including burnt vehicles and property damage: according to media sources, at least three people were injured and 47 were arrested during the disturbances [17], [18].

The final was played as a single-leg match in Munich on the night of May 31, 2025. As tickets for the match quickly sold out, PSG organized a public broadcast at Parc des Princes, where nearly 40,000 fans gathered to watch the game [19]. Additionally, as also shown in Fig. 1, giant screens displaying the match live were installed at multiple locations across Paris [20]–[23]. Following PSG's overwhelming 5–0 victory, large celebrations took place around the stadium and along the Champs-Élysées. Similar to the semifinal night, parts of the celebrations escalated into unrest after midnight, resulting in more than 500 arrests, numerous injuries, and reports of fatalities [24]. In contrast, smaller crowds gathered at Place de la Bastille and other locations across the city to celebrate without reports of violent incidents [25].

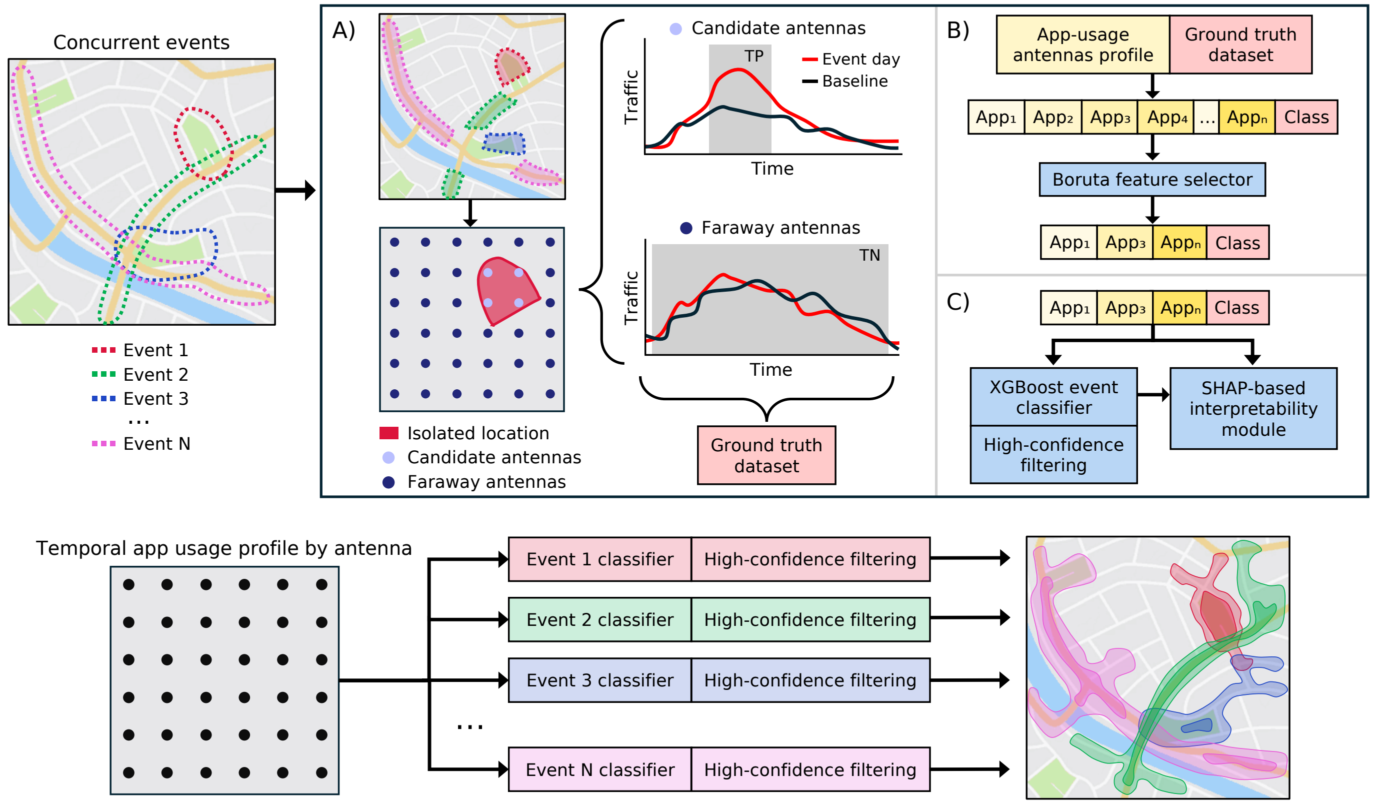


Fig. 2: Methodological framework. Top: Identification of concurrent events, followed by A) ground-truth labeling, B) feature engineering, and C) model training and footprint interpretation applied to each event. Bottom: Inference stage, with the trained models applied to the target data to disambiguate and reconstruct the full sequence of overlapping events.

The day after the final, a victory parade was organized along the Champs-Élysées. The avenue was closed from early morning, with public access allowed from 14:30. At around 17:00, a parade featuring the players was held to present the trophy to fans. Approximately 100,000 people are estimated to have attended the event. Additional celebrations took place at Parc des Princes, where the players continued the festivities with season ticket holders. Concerts were also organized at the stadium as part of the celebrations [26]–[28].

III. Methodology

Our study aims at defining a framework to (i) extract the independent digital signatures of concurrent social events and (ii) employ such signatures to monitor the spatiotemporal evolution of the target co-located happenings. The operation of the framework is illustrated in Figure 2 and presented in Section III-A. The methodology adopted in stages A–C of the inner pipeline is detailed in Sections III-B through III-D.

A. Overall operation

The proposed framework operates on N concurrent events. For each target event, it applies a sequence of steps (A—C in Fig. 2) to extract concise application usage signatures from a conservatively isolated portion of it. These signatures are then leveraged by event-specific classifier models, which are applied jointly to the data to disambiguate and fully reconstruct the overlapping activities. During this inference stage, prediction confidence serves as a guiding criterion, with higher-confidence predictions prioritized to ensure accurate resolution of concurrent events.

B. Ground-truth labeling

In stage A of the inner per-event pipeline of our framework, we first define a set of *ground truth* observations to isolate the traffic consumption patterns specific to each event category. To achieve this, we define well-delimited areas in both space and time, where the crowd behavior can be reliably attributed to a single target event type. Event location boundaries are typically determined with the help of external sources such as local police reports or news media. This process is conducted manually and aims to define highly conservative locations (*i.e.*, narrowly bounded in space and time) that are unambiguously linked to a single event category. Radio access antennas within such boundaries are then *candidates* to be labeled as *true positives* in the ground truth for the characterization of the event category under consideration, in case their mobile traffic demands experience significant changes during the duration of the target event. Similarly, *true negatives* are those antennas located far from the identified spatial boundaries and whose traffic load is completely unaffected during the target event.

Formally, let us denote by $\widetilde{\mathcal{K}}^+ \subset \mathcal{K}$ the set of candidate antennas within the event spatial boundaries, out of all available antennas $\mathcal{K}$ in the overall geographical region where the events occur. We compute for each antenna $k \in \widetilde{\mathcal{K}}^+$ a *baseline* behavior in normal periods not affected by special events, as the mean $\mu_k^{\mathrm{D}}(t) = \frac{1}{|\mathcal{N}|}\sum_{n\in\mathcal{N}} \mathrm{D}_{k,n}(t)$ of the total data traffic demand $\mathrm{D}_{k,n}(t)$ served by k at time t of day n within a set $\mathcal{N}$ of ordinary days. We then identify as true positives antennas that experience a substantial deviation in their traffic $\mathrm{D}_k(t)$ from the baseline $\mu_k^{\mathrm{D}}(t)$ during (portions of) the time period $\mathcal{T}$ influenced the target event category; more precisely, a candidate antenna k is a true positive at $t \in \mathcal{T}$ if

$$\mathrm{D}_k(t) > \mu_k^{\mathrm{D}}(t) + M \cdot \sigma_k^{\mathrm{D}}(t), \tag{1}$$

where $\sigma_k^{\mathrm{D}}(t) = \sqrt{\frac{1}{|\mathcal{N}|}\sum_{n\in\mathcal{N}}(\mathrm{D}_{k,n}(t) - \mu_k^{\mathrm{D}}(t))^2}$ is the standard deviation of the typical mobile traffic demand at time t across usual days in $\mathcal{N}$, and M is a parameter that controls the sensitivity of the true positive labeling process. Days in $\mathcal{N}$ are carefully selected to match the same weekday as the target event day, while excluding periods affected by other large-scale events, holidays, or long weekends.

It is worth stressing that, by targeting concurrent events in space and time, the sets of candidate antennas $\widetilde{\mathcal{K}}^+$ and relevant times $\mathcal{T}$ tend to overlap substantially across different event categories. It is thus critical that the true positives selection is very conservative and diverse events do not mutually confound the downstream characterization of their specific mobile traffic signatures. Therefore, in our implementation, we set M at a high value of 3—meaning that deviations observed at the true positive antennas during the target event must diverge from 99% of the customary traffic volumes recorded at the same time in normal days, under a normal distribution of demands. The final set of true positive antennas is denoted by $\mathcal{K}^+ \subseteq \widetilde{\mathcal{K}}^+$.

For true negatives, an equivalent set of candidate antennas $\widetilde{\mathcal{K}}^- \subset \mathcal{K}$ is defined, in this case encompassing antennas that are outside the boundary areas associated to the event under study. True negative antennas are then labeled from the candidate set above if (1) does not hold at any moment during the event time $\mathcal{T}$. The final set of true negatives is designated as $\mathcal{K}^- \subseteq \widetilde{\mathcal{K}}^-$.

C. Feature engineering

In stage B of Fig. 2, we leverage the ground truth defined in stage A for the target event class and model its digital footprint in terms of associated consumption of specific mobile services. To this end, we augment the ground truth observations for all true positive and negative antennas with information about the traffic demands generated by a set $\mathcal{A}$ of mobile applications. Since an analysis of application usage patterns based on raw traffic data may be dampened by strong imbalances in the traffic volume generated by each service [29], we convert the demands from bytes into a relative traffic consumption metric, *i.e.*, the Revealed Comparative Advantage (RCA) [30]. Originally introduced in international trade analysis, RCA is employed in our context to represent the prevalence of the utilization of each application $a \in \mathcal{A}$ at a ground-truth antenna $k \in \mathcal{K}^+ \cup \mathcal{K}^-$ relative to its typical consumption across the set $\mathcal{K}$ of all antennas in the whole geographical region of interest. Formally, the RCA $\rho_{a,k}(t)$ of mobile service a at antenna k and time t is computed as

$$\rho_{a,k}(t) \quad = \quad \frac{\mathrm{D}_{a,k}(t)/\mathrm{D}_{\mathcal{A},k}(t)}{\mathrm{D}_{a,\mathcal{K}}(t)/\mathrm{D}_{\mathcal{A},\mathcal{K}}(t)}, \tag{2}$$

where $\mathrm{D}_{\mathcal{A},k}(t) = \sum_{a\in\mathcal{A}} \mathrm{D}_{a,k}(t)$ denotes the traffic demand at time t of all services at antenna k, $\mathrm{D}_{a,\mathcal{K}}(t) = \sum_{k\in\mathcal{K}} \mathrm{D}_{a,k}(t)$ is the total traffic demand of application a at all antennas at t, and $\mathrm{D}_{\mathcal{A},\mathcal{K}}(t) = \sum_{a\in\mathcal{A}}\sum_{k\in\mathcal{K}} \mathrm{D}_{a,k}(t)$ is the overall traffic measured in the monitored network at t. To enable a more interpretable comparison between over- and under-consumed applications, we use the symmetric RCA (sRCA), which normalizes the unbounded RCA in (2) to the $[-1, 1]$ range as

$$\bar{\rho}_{a,k}(t) = \frac{\rho_{a,k}(t) - 1}{\rho_{a,k}(t) + 1}. \tag{3}$$

As a final step to the feature engineering stage, we account for the fact that each event category tends to affect different subsets of applications. Therefore, not all applications are relevant to identify every type of event, which paves the way for feature pruning. Specifically, we apply the Boruta feature selection algorithm [31] on the sRCA values $\bar{\rho}_{a,k}(t)$ obtained from (3) to only retain the relevant mobile services for characterizing the ground truth in each event class.

D. Model training and footprint interpretation

In the final stage C, we leverage the per-application sRCA values returned by stage B for the ground truth antennas of the target event category to train an XGBoost classifier (XGBC) for that specific type of event. The classifier aims at predicting which antennas are affected by the considered event based solely on their mobile service demands—hence identifying the digital footprint left by the event on the network traffic. The XGBC comprises 2,000 decision tree estimators and uses logistic regression as the loss function for binary classification, with a learning rate of 0.05. The output of the XGBC is an estimated probability that a given antenna is affected by the event of interest at time $t \in \mathcal{T}$. This probability is computed by aggregating the output of all trees in the ensemble model, and can be interpreted as the model *confidence*.

It is important to note that, in standard binary classification tasks, observations with predicted confidence above 0.5 are marked as positives, *i.e.*, would map in our context to samples $\mathrm{D}_k(t)$ of antenna k being tagged as affected by the event class. However, such a standard parametrization is too lax in presence of co-located, diverse events that can be easily confused with each other. As our problem requires a very high level of confidence in that the event is detected at each data point, we increase the confidence threshold to 0.95, ensuring that only samples for which the model exhibits extreme confidence are predicted to be affected by the event.

Finally, an interpretability module based on SHAP is applied to the trained XGBC, quantifying both the magnitude and direction of each feature's contribution to the model's output. Thus, the module provides insights into the over- or under-consumed applications characterizing each event.

IV. Digital Footprints of Celebrations and Riots

We demonstrate the viability of our framework by applying it to the reference use case presented in Section II-B. To analyze the social activities associated to the 2025 UEFA Champions League games in Paris, we leverage traffic measurements from a large operator, as presented in Section IV-A. By applying our framework to such measurement data, we can identify unique mobile service signatures for both celebration and riotous crowds, as discussed in Section IV-B.

A. Mobile network measurements

Our study builds upon mobile network traffic measurements collected from the production network of Orange, the operator with the leading market position in France, in the days around the semifinal and final games. We monitor the traffic (both downlink and uplink) served by over 24,000 antennas covering the Paris metropolitan area depicted in Fig. 1, where the majority of activities related to celebrations and riots took place. Our measurements cover all available Radio Access Network (RAN) technologies, namely 2G, 3G, 4G, and 5G; although 2G and 3G account for a negligible share of traffic consumption under normal conditions, their contribution increases in high-load scenarios where large crowds gather at specific locations.

At the time of data collection, Orange operated a non-standalone (NSA) 5G deployment, in which mobility management for both 4G gNodeBs and 5G eNodeBs was handled by the 4G Mobility Management Entity (MME). In parallel, the operator had also deployed a standalone (SA) 5G architecture, where mobility control operations are managed through the N2 interface at the 5G Access and Mobility Management Function (AMF). By monitoring the S1-MME and N2 interfaces, it is possible to geo-reference both 4G and 5G traffic sessions by associating each session with its serving antenna and corresponding physical location. For legacy 2G and 3G traffic, geo-referencing relies on the User Location Information (ULI) element embedded in Packet Data Protocol (PDP) contexts, which specifies the antenna to which a device is currently attached. This information is obtained by inspecting GPRS Tunneling Protocol control-plane (GTP-C) signaling over the Gn interface at the Gateway GPRS Support Node (GGSN).

Traffic generated by Orange subscribers is analyzed using proprietary classification techniques based on Deep Packet Inspection (DPI) of IP packets across all user sessions using advanced techniques such as IP-to-domain mapping, protocol identification, TLS fingerprinting, and payload signature analysis. While the internal implementation details and performance metrics of these classifiers cannot be disclosed due to proprietary constraints, they correspond to production-grade solutions designed to operate effectively under practical constraints such as traffic encryption and shared network infrastructure. This is consistent with state-of-the-art approaches reported in the literature, which achieve accuracy levels above 98% [32], [33], even under challenging scenarios similar to those considered in this study. Moreover, these classification outputs are routinely employed in operational settings by the network operator for network monitoring and optimization.

By combining localization data with service-level traffic classification and aggregating across all devices connected to each RAN antenna, the operator derives per-application traffic demand for each antenna. In this study, data are further aggregated into 5-minute intervals, resulting in time series representing traffic volume (in bytes), computed as the sum of downlink and uplink traffic, for more than 400 mobile services across the antennas, sampled every 5 minutes.

As discussed in Appendix A, all data collection and processing procedures comply with applicable regulations. Furthermore, the resulting aggregated dataset preserves user privacy, as it does not enable the identification of individuals or the extraction of personal information.

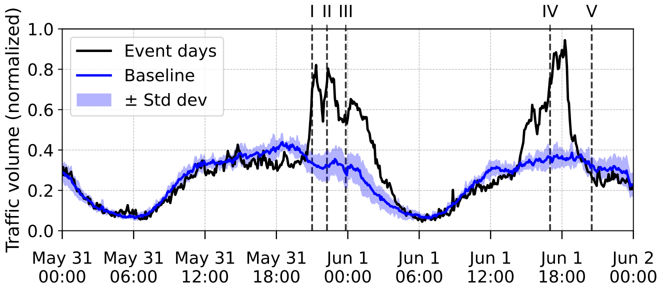
B. Mobile service footprints

We instruct our framework by applying its stages A–C described in Section III to the two separate event categories represented by peaceful celebrations and violent disturbs. Initially we focus on the day of the UEFA Champions League final, which, as shown in Appendix B, yields much stronger effects on the mobile network demand hence is a better scenario to produce reliable ground truth labels and cleaner traffic signatures for the two types of events.

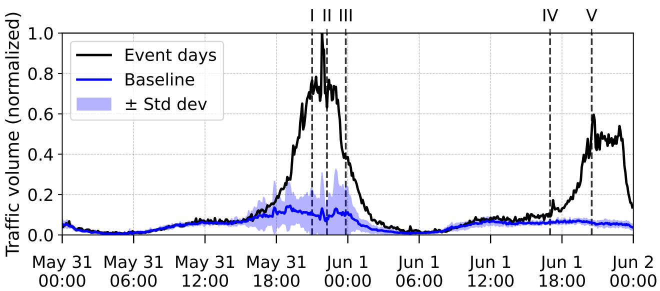
On the evening and night of May 31, we define candidate antennas by drawing boundary areas around locations highlighted in red in Fig. 1, which were reported as major hotspots for both fan celebrations and riots according to local news outlets [25], [28]. Visualizations of the aggregate traffic demand around Champs-Élysées and Parc des Princes during the day of the final and the following day are shown in Fig. 3: they highlight significant deviations from baseline demands $\mu_k^{\mathrm{D}}(t) \pm \sigma_k^{\mathrm{D}}(t)$ computed in the same area from six reference periods that cover the same weekdays in weeks both preceding and following May 31 and June 1. Specifically, the anomalies in the traffic pattern match well key moments linked with the football game and subsequent happenings.

The alterations in the overall demands in Fig. 3 are in fact the outcome of the two major phenomena that characterized the aftermath of the game and largely superposed in time and space: (i) fan celebrations that range from spontaneous gatherings to official organized activities such as the victory parade and (ii) violent riots, characterized by unpredictable movement of people and clashes with the police. As previously described, these events, while substantially co-located, differ significantly in terms of the behavior of their participants—in both the physical and digital domain.

We construct ground truth antenna sets $\mathcal{K}^+ \cup \mathcal{K}^-$ for both types of events using restrictive time frames to ensure, with high confidence, that the different behavioral contexts are effectively isolated. These time frames are primarily defined based on the timeline of events reported in news media sources and corroborated by social media activity. For the modeling of fan celebrations, we consider the area of the Champs-Élysées on June 1 from 16:00 to 19:00, since this covers the scheduled route and timing of the official celebration parade held the day after the final match. Additionally, no violent incidents



(a) Mobile data traffic at Champs-Élysées



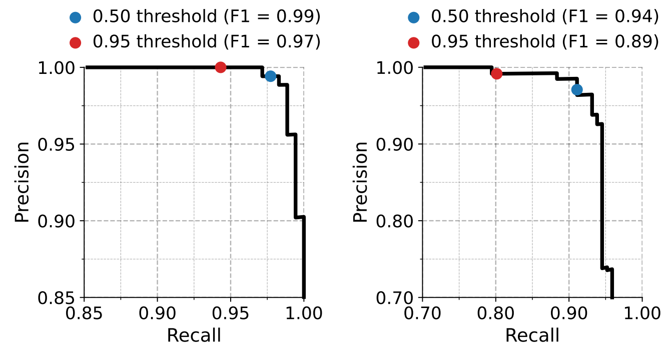
(b) Mobile data traffic at Parc des Princes

Fig. 3: Mobile network traffic volume at chief locations in Paris. Vertical lines indicate key moments linked to the 2025 UEFA Champions League final: (I) start of the match, (II) first reported clashes with the police, (III) end of the match, (IV) start of the parade at Champs-Élysées, and (V) start of organized celebrations at Parc des Princes.

were reported in the proximity of the parade, which makes the selected boundaries reliable for capturing purely celebratory behavior. Isolating the disturbs necessary to model riots and clashes with the police is more challenging, as most of the reported incidents began as celebrations that, at some point, escalated into violence. We opt for very conservative spatial and temporal boundaries, focusing on the Champs-Élysées from 01:00 to 02:00 on the night after the final match. During this period, the highest number of violent events were reported, including multiple arrests, injured individuals, looting, barricades, and cars set on fire; also, peaceful participants were reported to rapidly flee the area, thus weakening the footprint of cheerful celebrations on the mobile traffic at the location.

We run the candidate locations for the two target events through the pipeline presented in Section III, which returns highly accurate models of the social happenings. Fig 4 presents a sensitivity analysis of both models, showing the precision–recall trade-off across different confidence thresholds, including the conventional 0.5 threshold and the more conservative 0.95 threshold adopted in this work as discussed in Section III-D. While the latter does not maximize the F1-score, it enables near-perfect precision at the cost of reduced recall. This behavior is desirable in our setting, as the objective is to identify positive samples with very high confidence, even at the expense of missing weaker event signals.

More importantly, the final interpretation step of our frame-



(a) Celebrations

(b) Riots

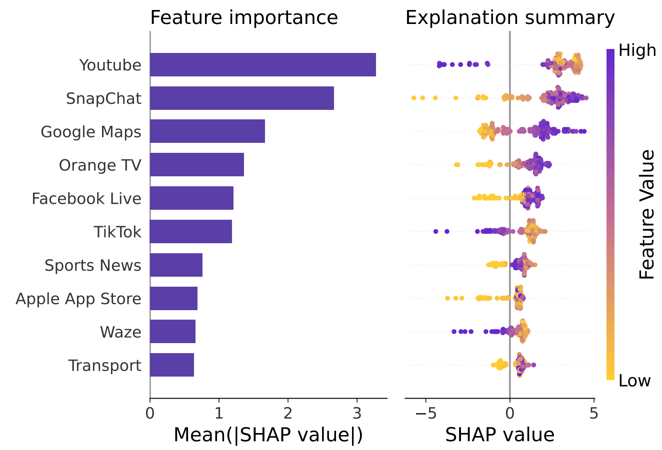
Fig. 4: Sensitivity analysis of the celebration and riot models across different confidence thresholds. Markers indicate the conventional 0.5 threshold and the more conservative 0.95 threshold adopted in this work.

work allows us to quantify the contribution of each feature to the predictions of both models, hence explaining which mobile applications experience a surge of consumption during celebrations and disturbances, respectively. Fig. 5 shows the SHAP value summary for both the celebration and riot models. In each case, features are ranked by their overall impact on the model output when predicting the positive class, *i.e.*, antennas found to be affected by the event. While the Boruta feature selection step reduced the original set of over 60 services to 27 and 25 applications for the celebration and riot models, respectively, the plots display the top 10 most important features only, as those are the most significant for interpreting the mobile traffic signature. Note that we omit the SHAP summary plot for the negative classes, as they aggregate heterogeneous non-event behavior across the target area.

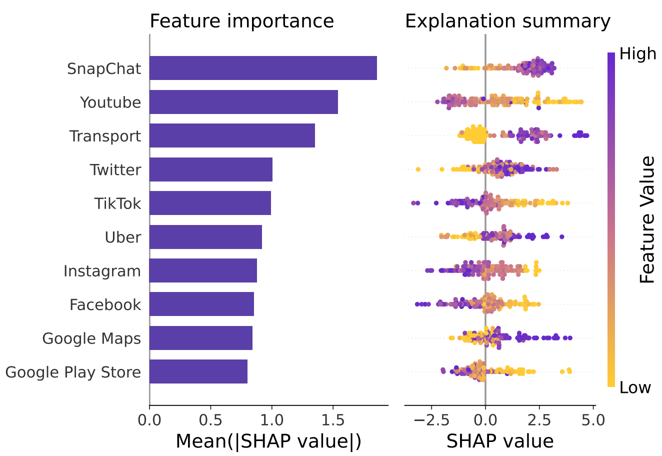
Juxtaposing Fig. 5a and Fig. 5b shows considerable overlap, with five applications appearing in the top 10 most important features for both models. This is expected, given some inherent similarities in the nature of the two events, the presence of causal relationships with some celebrations transitioning into protests, and unavoidable noise that the spatiotemporal co-location leaves in the individual digital footprints.

Although lower relative consumption of YouTube and higher relative consumption of Snapchat are common main patterns in both signatures, their relative importance (with respect to the other features) is substantially higher in the celebration model in Fig 5a. This suggests that user activity during celebrations is more homogeneous and dominated by specific application usage patterns, such as reduced passive content consumption and increased sharing of short, real-time content. In contrast, the riot-related behavior in Fig 5b appears more heterogeneous, which leads to a more distributed feature importance across applications.

Regarding the consumption of mobility-related applications, both types of activities are characterized by higher relative usage of Google Maps and Transport-related applications. However, the riot signature exhibits a stronger influence related



(a) Celebration model, top 10 of 27 selected features



(b) Riot model, top 10 of 25 selected features

Fig. 5: SHAP summary plots showing the application-level traffic signature of the celebration and riot models, based on the top 10 most important features after Boruta selection.

to the over-consumption of Transport-related apps, which, together with the increased use of Uber, may reflect users attempting to leave areas where violent incidents are occurring. Additionally, the under-consumption of Waze, which characterizes the celebration signature, is likely linked to street closures caused by fan celebrations, as drivers would normally rely on Waze in these areas.

The celebration signature also reflects a significant contribution from the increased relative consumption of live broadcasting services, such as Facebook Live and Orange TV. This pattern is consistent with the coverage of fan celebrations by television channels such as TF1 and M6, both of which are available through the Orange TV service. Additionally, the increased consumption of sport-related websites under Sports News appears only in the top 10 most important features for the celebration model. This indicates that celebration behavior is more closely linked to the match outcome than riot-related behavior. In contrast, as reported by local authorities, violent incidents were primarily committed not by PSG supporters, but by individuals exploiting the situation [24].

Finally, the riot model is characterized by over-consumption of Twitter/X, which likely reflects users both sharing and consuming real-time media about the ongoing incidents. This is consistent with previous studies documenting the close relationship between civil unrest and increased Twitter/X activity [34], as well as with evidence that visual records documenting the violent events were shared on that platform.

V. Disentangling Concurrent Events

From the modeling process, we obtain two event-specific classifiers that, although trained on very conservative spatiotemporal boundaries, can be used to identify similar behavioral patterns at other locations and times. In particular, we apply the models across all antennas in the Paris region at 5-minute intervals on the days of the semifinal and final, as well as on the immediately following days. At any given time, a particular antenna may be classified as exhibiting both celebratory and riot-related behavior, since the two models operate independently.

A. Temporal analysis

Fig. 6 shows the number of antennas identified as affected by each model in Paris at 5-minute intervals during the day of the final and the following day. Antennas linked to celebrations start appearing around 15:00 and gradually increase in number, reaching their highest values toward the end of the match, roughly double those identified by the riot model. Before midnight, the number of antennas detected by the riot model remains lower, with a first incremental increase around 22:00, closely aligned with the first reports of incidents near Parc des Princes during the second half of the match [35]–[37]. After midnight, the number of antennas linked to celebrations begins to decrease, contrasting with those associated with riots, which burst and remain at a high level until around 02:30, including roughly one hour during which the riot signal exceeds that of the celebrations. In the afternoon of June 1, the celebration model identifies a significant number of affected antennas throughout the official celebrations. In contrast, the riot model detects only a small number of antennas, in line with reports describing these events as largely peaceful, with only minor incidents between fans and the police occurring outside the parade area designated by the local authorities, as it reached its maximum allowed capacity and was closed off [38].

As shown in Fig. 7, the night of the semifinal match exhibits a pattern similar to that observed during the final: antennas linked to celebrations gradually increase toward the end of the match, while antennas associated with riots appear within a shorter time window after midnight. However, both celebratory and riot-related signals are considerably lower than those on the night of the final (see y-axis scale). Interestingly, between 18:00 and 20:00 on May 8, a reduced group of antennas near the Arc de Triomphe exhibits celebration behavior, although no activity related to fans was reported. Indeed, this minor signal corresponds to the celebration of the 80th anniversary of the Victory of May 8, scheduled between 17:50 and 19:10 on the

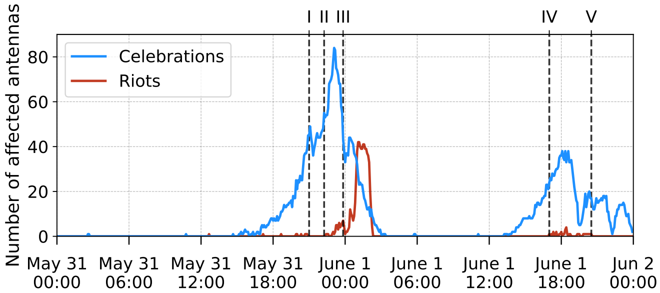


Fig. 6: Number of antennas classified as affected by both models at 5-minute intervals. Vertical lines indicate key moments on the day of the final and the following day: (I) start of the match, (II) first reported clashes with the police, (III) end of the match, (IV) start of the parade at Champs-Élysées, and (V) start of organized celebrations at Parc des Princes.

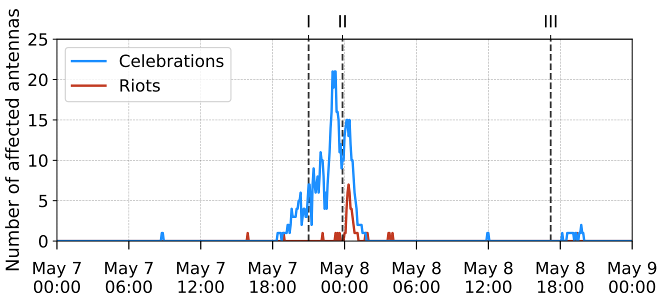


Fig. 7: Number of antennas classified as affected by both models at 5-minute intervals. Vertical lines indicate key moments on the day of the semifinal and the following day: (I) start of the match, (II) end of the match, and (III) commemoration of the Victory of May 8, 1945.

Place Charles de Gaulle (the location of the Arc de Triomphe) and the adjacent top of the Champs-Élysées. The ceremony involved street closures, was broadcast live on television, and included a parade of re-enactors. Thus, roughly matching the fan celebrations used to train the model.

B. *Spatial disentanglement*

As shown in Fig. 6 and Fig. 7, both celebratory and riot behaviors overlap in time during the nights of the semifinal and final. When considering the full periods, their spatial patterns also largely coincide, motivating a temporally disaggregated spatial analysis. To this end, we create maps of celebration and riot activities by aggregating antennas identified by both models over 30-minute intervals. Thus, within a given 30-minute window, each antenna in the Paris region is assigned a value from 0 to 6 for each model, indicating the number of 5-minute intervals during which it is identified as affected. Then, after aggregating the counts at the base station level (*i.e.*, geographic location), we interpolate the values using Ordinary Kriging to generate continuous spatial maps representing the intensity of detected activity across the region.

Fig. 8 presents selected 30-minute maps of detected activities during the night of the final match. During the first half of the match (21:00-21:30), celebratory behavior is detected around the Parc des Princes and other broadcast locations (see Fig. 1), where official watchalongs or giant screens were organized. No riot-related activity is observed during this time. Toward the end of the match (23:00-23:30), celebratory behavior starts converging around key locations, as people gathered to celebrate together. At the same time, riot-related activity appears near Parc des Princes, as the first reports of violent incidents were documented by French media outlets through Twitter/X [35]–[37]. After midnight (00:00-00:30), celebrations accumulate in fewer hotspots, mainly around Parc des Princes and central Paris landmarks (*e.g.*, the Champs-Élysées, and the Eiffel Tower). Meanwhile, riot activity begins to dissipate near Parc des Princes but emerges at the edges of the Champs-Élysées. Between 00:30 and 02:00 (*i.e.*, when most violent incidents were reported in Paris [25]), riot-related behavior intensifies, with the Champs-Élysées becoming increasingly dominated by rioters. Additionally, a resurgence of riot activity is observed near Parc des Princes, as police reports indicated that rioters returned to the area around 01:00. During this time, certain areas with predominantly celebratory behavior persist. Notably, in areas around the Eiffel Tower (Place du Trocadéro and Pont de Bir-Hakeim), lower-intensity riot behavior is observed, although no violent incidents or clashes with the police were reported. According to media reports, mainly fan-recorded videos, these areas were characterized by peaceful yet chaotic celebrations, marked by extensive use of fireworks and flares. In these areas, the predominance of celebratory behavior further increases as the Champs-Élysées transition into a riot zone. Other locations, such as Place du Carrousel and Place de la Bastille, continue to exhibit purely celebratory behavior, though at lower intensities. Between 02:00 and 02:30, the intensity of both activities begins to fade, although a clear geographical separation remains, with distinct clusters of people participating in each type of behavior.

In the case of the night of the semifinal, Fig. 9 reveals that the identified spatial patterns are less dynamic than those on the night of the final, closely aligned with the lower scale of reported activities. Before midnight, only celebration behavior is identified, concentrated around Parc des Princes while the match was ongoing (21:00-21:30), and later near the Arc de Triomphe on the Champs-Élysées after the match (23:30-00:00). After midnight, coinciding with local reports, riots emerge at the points of celebration on the Champs-Élysées but rapidly concentrate in a smaller area by 01:30. After 02:00, both signals have completely dissipated.

Additional results related to the official celebrations on the afternoon of June 1, largely reported as a peaceful happening without disturbs, are shown in Appendix C.

VI. Conclusions

We present a data-driven methodology to characterize and disentangle concurrent urban events by extracting and exploiting their intrinsic digital signatures. After applying our

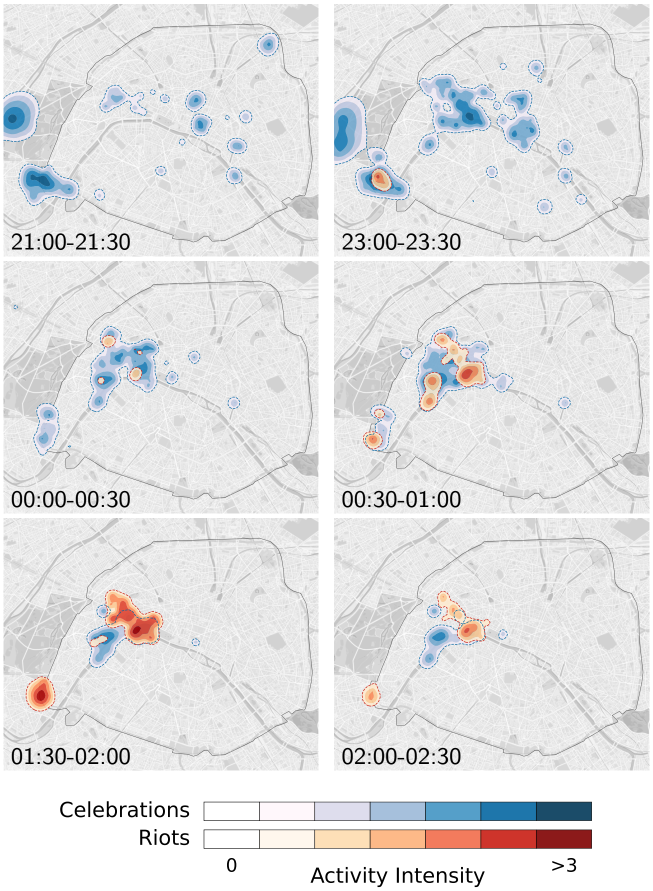


Fig. 8: Spatiotemporal evolution of celebratory and riot activity during the night of the final. The maps show the density of affected antennas in selected 30-minute time windows.

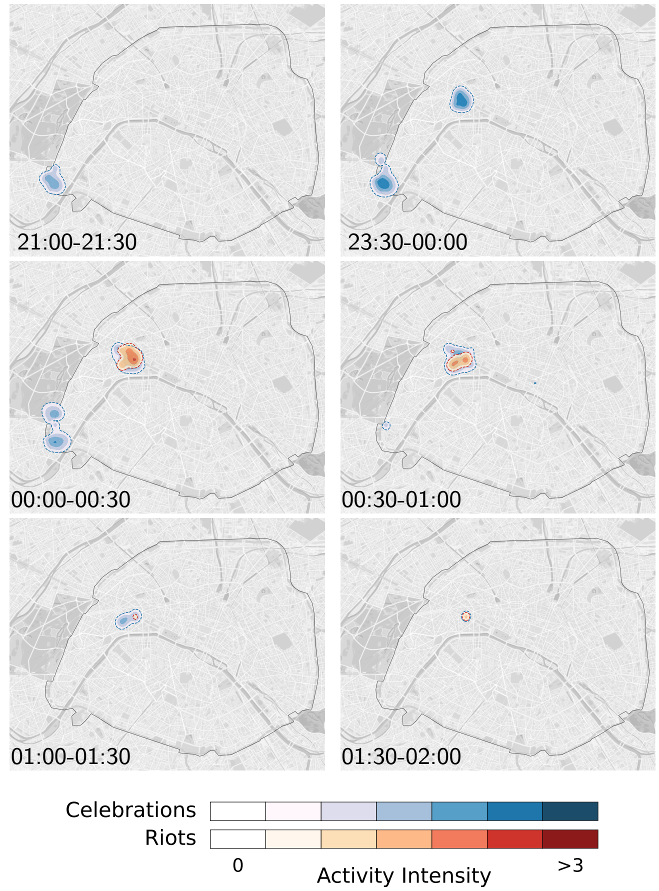


Fig. 9: Spatiotemporal evolution of celebratory and riot activity during the night of the semifinal. The maps show the density of affected antennas in selected 30-minute time windows.

methodological framework to the overlapping celebrations and riots following Paris Saint-Germain's 2025 UEFA Champions League win, we demonstrate the viability of the proposed approach by precisely disentangling these concurrent events and revealing their interactions. Our results provide insights into the movement of people across different areas of the city, the swift oscillation between celebrations and riot behavior, and the identification of predominantly celebratory locations, which received limited media attention due to the absence of violent incidents. These findings highlight the capacity of the proposed framework to uncover subtle behavioral signals, holding significant potential for applications in public safety, urban planning, and event management.

Since the proposed methodology is evaluated on a single case study, its generalizability across different geographical contexts and event types requires further investigation. Nevertheless, given the strong spatiotemporal overlap and behavioral similarity between the two concurrent events analyzed in this work, we expect that the proposed approach can be extended to other scenarios with minor adaptations, provided that reliable information is available to identify ground-truth boundaries for each event type. Future work directions include exploring the applicability of the methodology across a broader range of events, regions, and network deployments. A representative scenario involves multiple events planned to occur simultaneously in Paris on May 30, 2026, including a screening of the UEFA Champions League final at Parc de Princes (potentially followed by celebrations and disturbances), music concerts at three of the largest arenas in Paris (namely Stade de France, Paris La Défense Arena, and Accor Arena), and the 2026 Roland-Garros tennis tournament. The coexistence of these large-scale events, spanning different domains and audiences, is expected to generate complex and potentially overlapping traffic signatures. In this context, unsupervised learning approaches could be explored in future work to identify latent behavioral groups directly from traffic measurements, without the need for manually curated ground truth.

Acknowledgment

This work was funded by CoCo5G (ANR-22-CE25-0016) of ANR. M. Pirri's work is co-funded by Comunidad de Madrid with public funds from predoctoral grant PIPF-2024/COM-34005. D. Madariaga's and M. Fiore's work was supported by the ORIGAMI project (GA 101139270) funded by SNS JU and the European Union.

REFERENCES

[1] L. Hong, M. Lee, A. Mashhadi, and V. Frias-Martinez, "Towards understanding communication behavior changes during floods using cell phone data," in *International Conference on Social Informatics*. Springer, 2018, pp. 97–107.

[2] P. Lemaire, A. Furno, S. Rubrichi, A. Bondu, Z. Smoreda, C. Ziemlicki, N.-E. El Faouzi, and E. Gaume, "Early detection of critical urban events using mobile phone network data," *Plos one*, vol. 19, no. 8, p. e0309093, 2024.

[3] A. F. Zanella, D. Madariaga, S. Mishra, O. E. Martínez-Durive, Z. Smoreda, and M. Fiore, "Characterizing, modeling and exploiting the mobile demand footprint of large public protests," in *Proceedings of the 2024 ACM on Internet Measurement Conference*, 2024, pp. 516–529.

[4] J. Reades, F. Calabrese, A. Sevtsuk, and C. Ratti, "Cellular census: Explorations in urban data collection," *IEEE Pervasive computing*, vol. 6, no. 3, pp. 30–38, 2007.

[5] T. Yabe and S. V. Ukkusuri, "Integrating information from heterogeneous networks on social media to predict post-disaster returning behavior," *Journal of Computational Science*, vol. 32, pp. 12–20, 2019.

[6] N. Oliver, B. Lepri, H. Sterly, R. Lambiotte, S. Deletaille, M. De Nadai, E. Letouzé, A. A. Salah, R. Benjamins, C. Cattuto *et al.*, "Mobile phone data for informing public health actions across the covid-19 pandemic life cycle," p. eabc0764, 2020.

[7] G. Pullano, E. Valdano, N. Scarpa, S. Rubrichi, and V. Colizza, "Evaluating the effect of demographic factors, socioeconomic factors, and risk aversion on mobility during the covid-19 epidemic in france under lockdown: a population-based study," *The Lancet Digital Health*, vol. 2, no. 12, pp. e638–e649, 2020.

[8] F. Calabrese, F. C. Pereira, G. Di Lorenzo, L. Liu, and C. Ratti, "The geography of taste: analyzing cell-phone mobility and social events," in *Pervasive Computing: 8th International Conference, Pervasive 2010, Helsinki, Finland, May 17-20, 2010. Proceedings 8*. Springer, 2010, pp. 22–37.

[9] A. Janecek, D. Valerio, K. A. Hummel, F. Ricciato, and H. Hlavacs, "The cellular network as a sensor: From mobile phone data to real-time road traffic monitoring," *IEEE transactions on intelligent transportation systems*, vol. 16, no. 5, pp. 2551–2572, 2015.

[10] M. K. Chen and R. Rohla, "The effect of partisanship and political advertising on close family ties," *Science*, vol. 360, no. 6392, pp. 1020–1024, 2018.

[11] P. Rozenshtein, A. Anagnostopoulos, A. Gionis, and N. Tatti, "Event detection in activity networks," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2014, pp. 1176–1185.

[12] C. Wu and M. S. Gerber, "Forecasting civil unrest using social media and protest participation theory," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 1, pp. 82–94, 2017.

[13] N. Alsaedi, P. Burnap, and O. Rana, "Can we predict a riot? disruptive event detection using twitter," *ACM Transactions on Internet Technology (TOIT)*, vol. 17, no. 2, pp. 1–26, 2017.

[14] A. M. Ertugrul, Y.-R. Lin, W.-T. Chung, M. Yan, and A. Li, "Activism via attention: interpretable spatiotemporal learning to forecast protest activities," *EPJ Data Science*, vol. 8, no. 1, p. 5, 2019.

[15] G. Zuccaro, D. De Gregorio, and M. F. Leone, "Theoretical model for cascading effects analyses," *International Journal of Disaster Risk Reduction*, vol. 30, pp. 199–215, 2018, understanding and mitigating cascading crises in the global interconnected system. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212420918305041

[16] ESPN, "PSG 2-1 Arsenal (May 7, 2025) game analysis," May 2025. [Online]. Available: https://www.espn.com/soccer/report/_/gameId/733618

[17] Newsflare, "France: Celebrations turn tense in Paris after PSG beat arsenal in Champions League semi-final," May 2025. [Online]. Available: https://www.newsflare.com/video/743572/france-celebrations-turn-tense-in-paris-after-psg-beat-arsenal-in-champions-league-semi-final

[18] Reuters, "PSG victory over Arsenal triggers night of disorder in Paris," May 2025. [Online]. Available: https://www.reuters.com/sports/soccer/psg-victory-over-arsenal-triggers-night-disorder-paris-2025-05-08/

[19] A. Lasry, "Psg – inter: Champions league final to be broadcast at parc des princes in front of 38,000 fans," May 2025. [Online]. Available: https://parissecret.com/en/champions-league-final-parc-broadcast/

[20] Konbini, "PSG – Inter Milan: 10 spots pour voir (sur grand écran) la finale de la ligue des champions," May 2025. [Online]. Available: https://www.konbini.com/lifestyle/psg-inter-milan-10-spots-pour-voir-sur-grand-ecran-la-finale-de-la-ligue-des-champions/

[21] Sortiraparis.com – Paris events guide, "PSG Champions League Final: Giant screens where to watch the Paris-Saint-Germain match," May 2025. [Online]. Available: https://www.sortiraparis.com/en/what-to-do-in-paris/sport-wellness/articles/328207-psg-champions-league-final-giant-screens-where-to-watch-the-paris-saint-germain-match

[22] Soirée à Paris, "Retransmission du match PSG Inter de Milan avec le son a 21h // Jardin Géant // hippodrome de longchamp Paris 16eme Le," May 2025. [Online]. Available: https://www.instagram.com/p/DKEiTMGIxKy/

[23] Sortiraparis.com – Paris events guide, "Bars, pubs, and restaurants where you can watch sports events and matches in Paris: Our top picks," May 2025. [Online]. Available: https://www.sortiraparis.com/en/where-to-eat-in-paris/bars-cafes/guides/175503-bars-pubs-broadcast-matches-paris

[24] T. Mackintosh and A. Lamche, "Two dead and hundreds arrested in France after PSG Champions League win," Jun 2025. [Online]. Available: https://www.bbc.com/news/articles/ckgqyg325gno

[25] Le Monde, "PSG holds victory parade in Paris with Champions League Trophy after overnight violence," Jun 2025. [Online]. Available: https://www.lemonde.fr/en/sports/article/2025/06/01/psg-holds-victory-parade-in-paris-with-champions-league-trophy-after-overnight-violence_6741896_9.html

[26] Sortiraparis.com – Paris events guide, "The PSG parade on the champs-elysées, how to get there and closed subways," Jun 2025. [Online]. Available: https://www.sortiraparis.com/en/news/in-paris/articles/329048-the-parade-on-the-psg-fields-how-to-get-there-and-metros-fermes

[27] Sortiraparis.com – Paris events guide, "PSG parade on the champs-elysées: Traffic restrictions and road closures," Jun. [Online]. Available: https://www.sortiraparis.com/en/news/in-paris/articles/329052-psg-parade-on-the-fields-traffic-restrictions-and-closed-roads

[28] F. Katunga, "Riots erupt in Paris as PSG Champions League Victory Celebrations Turn Violent," Jun 2025. [Online]. Available: https://travelnoire.com/paris-riots

[29] S. Mishra, D. Madariaga, C. Ziemlicki, D. Naboulsi, and M. Fiore, "An urban geography of mobile application usage: Connecting demand dynamics and urban fabrics," in *IEEE INFOCOM 2025 - IEEE Conference on Computer Communications*, 2025, pp. 1–10.

[30] B. Balassa and M. Noland, ""revealed" comparative advantage in japan and the united states," *Journal of International Economic Integration*, vol. 4, no. 2, 1989.

[31] M. B. Kursa and W. R. Rudnicki, "Feature selection with the boruta package," *Journal of statistical software*, vol. 36, pp. 1–13, 2010.

[32] I. Akbari, M. A. Salahuddin, L. Aniva, N. Limam, R. Boutaba, B. Mathieu, S. Moteau, and S. Tuffin, "Traffic classification in an increasingly encrypted web," *Communications of the ACM*, vol. 65, no. 10, pp. 75–83, 2022.

[33] P. Luo, J. Chu, and G. Yang, "Ip packet-level encrypted traffic classification using machine learning with a light weight feature engineering method," *Journal of Information Security and Applications*, vol. 75, p. 103519, 2023.

[34] J. T. Jost, P. Barberá, R. Bonneau, M. Langer, M. Metzger, J. Nagler, J. Sterling, and J. A. Tucker, "How social media facilitates political protest: Information, motivation, and social networks," *Political psychology*, vol. 39, pp. 85–118, 2018.

[35] CLPRESS, "PSG's victory in the Champions League," Jun 2025. [Online]. Available: https://clpress.fr/en/categories/psg-victory-in-the-champions-league

[36] CLPRESS, "Incidents aux abords du Parc des Princes," May 2025. [Online]. Available: https://x.com/CLPRESSFR/status/1928926297075433516

[37] CLPRESS, "Des supporters du Paris Saint Germain envahissent le périphérique," May 2025. [Online]. Available: https://x.com/CLPRESSFR/status/1928906519099191297

[38] P. Allen, "PSG's Champions League Parade turns nasty: Fans clash with riot police in more chaotic scenes as Emmanuel Macron condemns violence that left two people dead and hundreds injured," Jun 2025. [Online]. Available: https://www.dailymail.co.uk/sport/football/article-14770039/PSGs-Champions-League-parade-turns-nasty-Fans-clash-riot-police-chaotic-scenes-Emmanuel-Macron-condemns-violence-left-two-people-dead-hundreds-injured.html

Appendix A
Ethics

Our study relies on mobile network traffic generated by subscribers of a nationwide cellular infrastructure and collected by the MNO for network operation and research activities. Measurements were processed within a secure platform located at the operator's premises and handled exclusively by authorized MNO personnel. This procedure enabled the generation of anonymized aggregates of service-specific traffic volumes at the antenna level. Each aggregate combines the activity of tens to hundreds of users, thereby preventing the re-identification of individual data subjects. In addition, all personal identifiers (*e.g.*, the International Mobile Subscriber Identity, IMSI) as well as sensitive user-level information (*e.g.*, the sequence of visited cells or the specific mobile services used) are removed during the aggregation process.

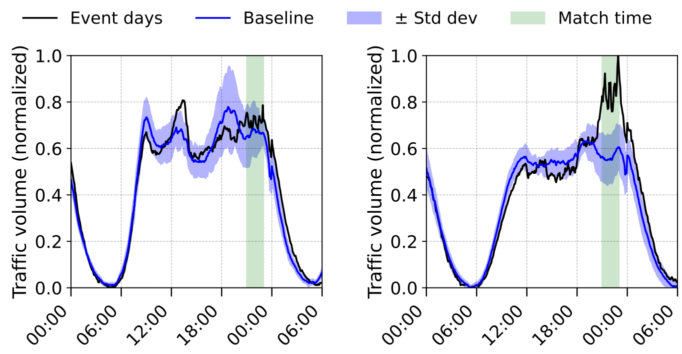
Data collection and processing procedures were reviewed and approved by the Data Protection Officer (DPO) of the MNO within the framework of a collaborative research project. The researchers involved in this work were granted access exclusively to the resulting aggregated and privacy-preserving traffic statistics, which do not qualify as personal data under the definition provided by the GDPR. Our dataset and research do not involve risks for the mobile subscribers, while generating important insights about its potential as an alternate type of data capable of remotely sensing large-scale events.

Appendix B
Impact of the target UEFA Champions League games on the overall Paris mobile traffic

The impact of the 2025 UEFA Champions League semifinal and final games on the citywide mobile network traffic recorded by Orange in Paris is summarized in Fig. 10. The plots compare the total volume of traffic observed across the Parisian region on the days of the semifinal and final (Wednesday, May 7, and Saturday, May 31), as well during the nights following these matches, against the corresponding baseline traffic behavior. Baseline traffic activity is computed over six reference periods (*i.e.*, the same weekdays, excluding holidays) selected from weeks both preceding and following the events. In the case of the semifinal, Fig.10a reveals that, despite the relevance of the event and being held in Paris, the overall network load remains close to baseline behavior. In contrast, on the day of the final, although hosted in Germany, Fig. 10b shows that citywide traffic volume clearly exceeds the baseline, not only during the match but also in the preceding hours and into the early hours of the following day.

Appendix C
Spatial analysis of official celebrations

Fig. 11 shows the spatiotemporal patterns identified by both models during the official celebrations organized on June 1 at 17:00 on the Champs-Élysées and at Parc des Princes at 20:30. The maps reveal celebratory behavior emerging hours before the scheduled events, with a clear movement from the Champs-Élysées to Parc des Princes, marking the transition between the two events. Only very minor and essentially negligible hotspots of riot activity are observed near the parade area, consistent with the minor incidents reported at the periphery of the parade, as noted in Section V-A



(a) Semifinal (b) Final

Fig. 10: Citywide traffic volume during the day of the (a) semifinal (Wed, May 7) and (b) final (Sat, May 31).

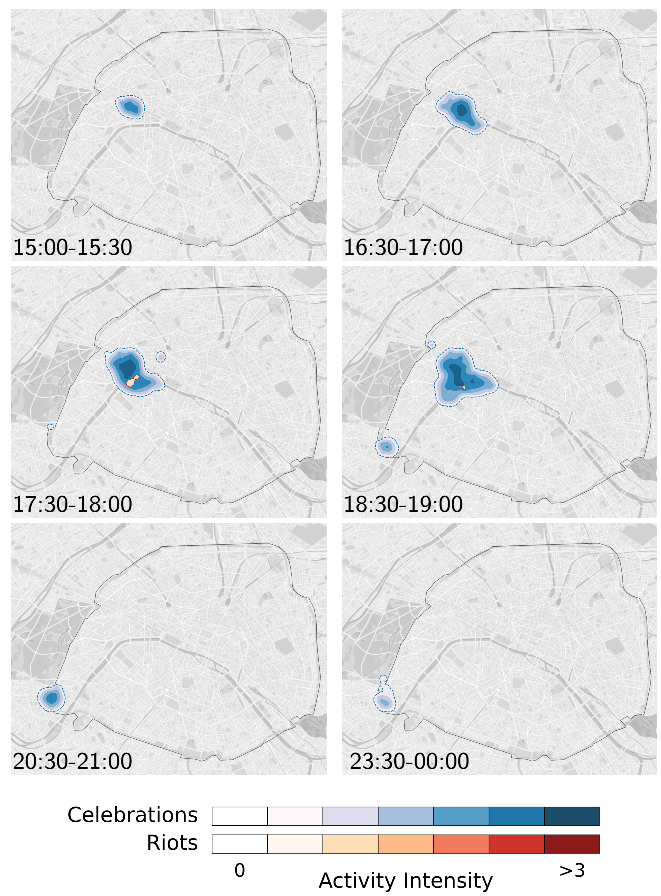


Fig. 11: Spatiotemporal evolution of celebratory and riot activity during the official celebrations. The maps show the density of affected antennas in selected 30-minute time windows.