Tracking the Evolution of Temporal Patterns of Usage in Bicycle-Sharing Systems Using Nonnegative Matrix Factorization on Multiple Sliding Windows

Remy Cazabet*, Pablo Jensen†, Pierre Borgnat †
*LIP6, CNRS, Universite Pierre et Marie Curie,
Paris, France
†Univ Lyon, Ens de Lyon, Univ Claude Bernard, CNRS, Laboratoire de Physique,
F-69342 Lyon, France

***Abstract*—Bicycle-Sharing Systems (BSS) are growing quickly in popularity all over the world. In this article, we propose a method based on Nonnegative Matrix Factorization to study the typical temporal patterns of usage of the BSS of Lyon, France, by studying logs of rentals. First, we show how this approach allows us to understand the spatial and temporal usage of the system. Second, we show how we can track the evolution of these temporal patterns over several years, and how this information can be used to better understand the BSS, but also changes in the city itself, by considering the stations as social sensors.**

I. Introduction

Bike Sharing Systems (BSS) are now ubiquitous in many cities all over the world. They offer a new type of public transportation system, often considered complementary with more traditional Public Transport. Because most of BSS systems are composed of fully automated electronic docking stations, large datasets describing the usage of the system are usually accessible.

Velo'v, the BSS system installed in Lyon in 2005, is one of the oldest. We obtained the usage dataset for the last 5 years (2011-2015).

In this article, we propose to use Nonnegative Matrix Factorization (NMF) to find the stations' typical spatio-temporal patterns of usage. We propose a method to identify automatically the stations whose patterns of usage have changed, revealing modifications in their neighborhood.

In section 2, we show how to use NMF to uncover weekly temporal patterns of usage (TPU) and the associated spatial usage. In section 3, we present a method for uncovering the evolution of the prevalence of these TPU. In section 4, we apply the proposed method on Lyon's BSS, and show how it can be used to better understand both the system and the changes in the city.

II. Usage pattern detection

A. Nonnegative Matrix Factorization

Nonnegative Matrix Factorization (NMF) has been used in a wide variety of applications since the seminal paper of Lee et al. [1]. Famous applications include facial recognition [2], document clustering [3] and overlapping community detection [4].

NMF can also be used to uncover temporal features in networks [5], such as transportation networks [6].

NMF has already been used to study mobility patterns, for example on taxi trips [7]. In that case, three temporal features are enough to explain the traffic during week days. In [8], the authors use a locality preservation constraints based NMF (LPNMF) to obtain a low-dimension representation of network-level traffic states. They show through experiments on realistic simulated traffic data how this information can be used to predict long-term spatial patterns of traffic flow. In these two articles, the NMF is computed a single time for the whole studied dataset. Here, we use sliding windows to study the evolution of temporal patterns found by the NMF.

Nonnegative Matrix Factorization (NMF) [1] can be written as the following problem: given a nonnegative matrix $\boldsymbol{V}$ of dimension $E \times T$, find a factorization $\boldsymbol{V} \approx \boldsymbol{W}\boldsymbol{H}$ where $\boldsymbol{W}$ and $\boldsymbol{H}$ are two nonnegative matrices of dimensions respectively $E \times Q$ and $Q \times T$, with Q the number of features (usually small). $\boldsymbol{W}$ is described as the features matrix and $\boldsymbol{H}$ the corresponding levels of activation over time. Formally the NMF problem comes down to solve:

$$(\hat{\boldsymbol{W}}, \hat{\boldsymbol{H}}) = \arg\min_{W,H} \sum_{e=0}^{E-1} \sum_{t=0}^{T-1} d(\boldsymbol{V}_{et} | [\boldsymbol{W}\boldsymbol{H}]_{et}) \qquad (1)$$

with $\boldsymbol{V}_{et}$ the value at row e and column t of the original matrix $\boldsymbol{V}$, $[\boldsymbol{W}\boldsymbol{H}]_{et}$ the same element of the reconstructed matrix, and $d(x|y)$ a scalar cost function (several cost functions have been proposed in the literature), that is often a parameter of the method. To simplify the notation, in the rest of the paper, we will note $\hat{\boldsymbol{W}}$ as $\boldsymbol{W}$ and $\hat{\boldsymbol{H}}$ as $\boldsymbol{H}$.

Several methods have been proposed to solve this problem, using different cost functions and heuristics [9]. Here we use a public implementation in the scikit-learn library [10], using the method described in [11] and the cost function $d(x|y) = |x - y|$. Note that the solution is not necessarily unique (the problem to solve is not convex) and that the solution found is an approximation. It is therefore important to pay attention to the stability of the result. In this article, we show that on our BSS dataset, results are stable not only for multiple runs on a same dataset, but also on several runs on different sliding windows of the same dataset.

B. *Constitution of the action matrix*

Let's define an action a as a couple associating an entity e and an instant in time t. Entities, in this study, will be BSS stations, but we could use the same method to find Temporal Patterns of Usage of users, or TPUs of particular trips between pairs of stations. Actions, in this study, will be arrivals of bicycles at stations.

Let's represent each action by a couple (e, t), with $e \in E$, the set of all stations in the system.

A is the log of all occurring actions:

$$A = [(e_1, t_1), ...(e_n, t_n)] \tag{2}$$

It corresponds to the list of all arrivals of bicycle in all stations.

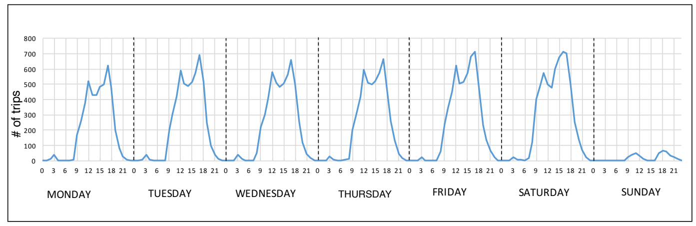
We first need to create a matrix of actions $\boldsymbol{V}$ of size $E \times T$, in which T depends on the chosen temporal granularity. The temporal granularity is defined by two factors: the *temporal unit* and a *typical pattern period*.

Temporal Unit: because we need to discretize time to create a matrix, we need to choose a minimal temporal unit to be considered. In this article, we use a Temporal Unit of 1h, meaning that all events occurring during the same hour will be binned together. A 1h period seems to be a good trade-off between the need of having many events in each bin, and having a fine temporal granularity. Previous works have shown that similar Temporal units such as 30 minutes or 2 hours do not affect qualitatively the results [6].

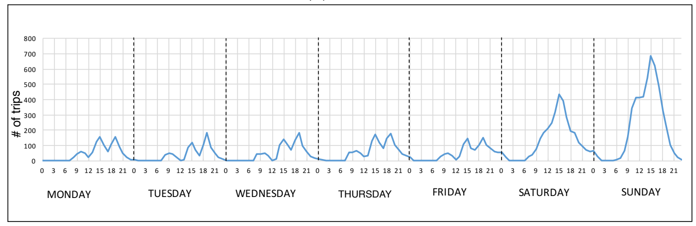
Typical pattern period: In order to get interpretable results, we must also choose a typical pattern period, which corresponds to the duration of the typical TPU we are searching for. In this article, we use a period of one week, typically used when searching for patters of usage [12], [13]. One could also consider only two different types of days, week-days and week-end days. However, one advantage of the NMF is precisely not to have to make such *a priori* decision. As a matter of fact, we do observe relevant differences between days of the same category for some patterns, such as a strong difference between Saturdays and Sundays in activity in commercial areas, or differences between week-days nights in late-evening activities.

Now we can define T as the number of temporal units contained in the typical pattern period, i.e with Temporal Unit = 1h and Typical pattern period = 1 week, $T = 24 \times 7 = 168$.

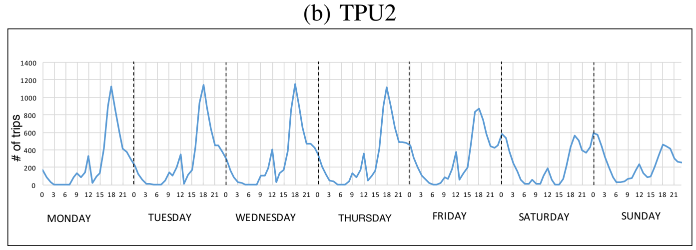
As a consequence, $\boldsymbol{V}_{et}$ corresponds to the number of actions accomplished by entity e during the t^{th} hour of a typical week.



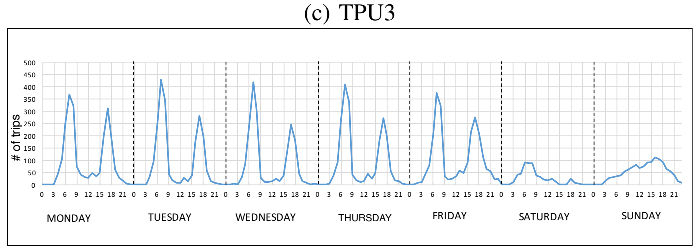
(a) TPU1



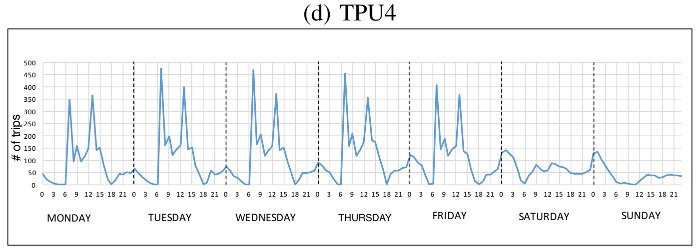
(b) TPU2



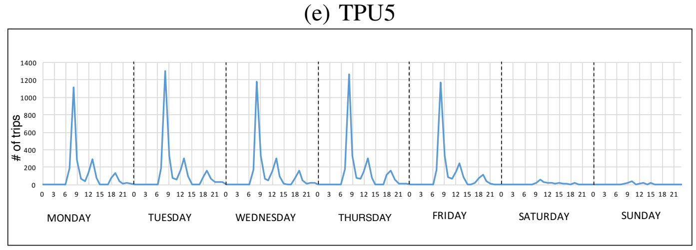
(c) TPU3



(d) TPU4



(e) TPU5



(f) TPU6

Fig. 1: Temporal Pattern of Usage (TPU) found for Lyon's BSS

C. *Computation of Pattern Impact Matrices*

To interpret the results, we generate three complementary matrices from the decomposition matrices $\boldsymbol{W}$ of dimension

$E \times Q$ and $\boldsymbol{H}$ of dimension $Q \times T$, as explained in Formula (1). As a reminder, E is the number of entities, Q the number of desired features and T the number of temporal units.

$\tilde{\boldsymbol{W}}$ represents the profiles of each entity over the features, i.e. its number of actions for each feature :

$$\tilde{\boldsymbol{W}}_{eq} = \sum_t \boldsymbol{W}_{eq}\boldsymbol{H}_{qt} \tag{3}$$

$\tilde{\boldsymbol{H}}$ represents the entities' profiles over the time steps :

$$\tilde{\boldsymbol{H}}_{qt} = \sum_e \boldsymbol{W}_{eq}\boldsymbol{H}_{qt} \tag{4}$$

$\boldsymbol{F}$ gives, for each feature, the total number of actions:

$$\boldsymbol{F}_q = \sum_e \sum_t \boldsymbol{W}_{eq}\boldsymbol{H}_{qt} \tag{5}$$

III. Temporal evolution of patterns of usage

Modern cities are changing at a fast pace, and we can expect usages of BSS installed in these cities to change accordingly. The questions we want to answer are, first, if there are changes at the system scale, i.e if the patterns of usage of the system change along time, while users get used to the system and learn how it can be used, and, second, how we can track punctual changes, such as differences in the pattern of usages of stations affected by city modifications.

A. Computing the evolution of usage patterns

NMF is typically used to uncover Temporal Patterns of Usage (TPU) in an time-aggregated dataset. For instance, to study the TPUs of stations in a BSS, one will count the number of arrivals per hour for each station in the whole dataset. The rows of the matrix to be decomposed therefore correspond to the stations of the BSS.

In the method we propose, we use overlapping sliding windows to create smaller instances of the dataset, and, for each of these windows, we compute, for each station, the number of arrivals per hour. We later aggregate all these profiles in a single matrix to decompose. A row of this matrix corresponds to the activity of one station in one time-window.

More formally, we consider that in our original dataset, we have a list of timestamped actions, where an action is, for instance, a user taking a bicycle from a given station. There are E stations, the entities whose activity we want to study. The method we propose can be described as follows:

1) Extract part of the data using a sliding window (SW) of aggregation. We will obtain a list of i SW, each SW being a list of actions, a subset of the original dataset
2) Compute the action matrices corresponding to each window as explained before. We obtain i matrices $\boldsymbol{V^i}$ of size $E \times T$
3) Stack all matrices column-wise into a single matrix $\boldsymbol{V} = [V^{i_1}, ..., V^{i_n}]$
4) Decompose this matrix using the NMF as described in sections II-B and II-C, and Formula (1).

The result of this procedure is:

- The TPU matrix $\tilde{\boldsymbol{H}}$, corresponding to patterns that are relevant over all sliding windows
- The pattern impact matrix $\tilde{\boldsymbol{W}}$, which gives the importance of each pattern for each entity for each time window.

We can use this information to study how usage patterns have evolved for each entity, and for the system as a whole.

At this point, it might be necessary to clarify the different periods of time we use :

- The temporal unit (1 hour), is used to count the number of actions per temporal unit
- Typical pattern period (Week), the length of our TPUs, which means that we are searching for weekly temporal patters
- Length of the sliding window (1 Year)
- Shift between windows, or size of the sliding window (1 month)

B. Detection of changes in usage patterns

BSS stations are located at fixed points in a city. Previous works have considered that the activity of a station is mostly affected by a buffer zone of around 300 meters around each station [14]. We can therefore expect that changes occurring in this buffer zone affect the activity of a station. Detecting these changes might be interesting at several levels:

- It can alert the BSS operator on the necessity of upgrading the station capacity (for instance if there is a rise in rush-hour usage)
- It can help the BSS operator to update its balancing strategy for the affected stations
- Through the "social sensor" perspective [15], it can allow city planners to detect either progressive, unplanned changes (loss in commercial attractiveness, increase in late-night activities, etc.), or to observe the effect of an urban development project (opening of a Public Transport station, a park, etc.)

To compare the TPU of a same station in different windows, we first compute a normalized temporal profile for each station in each window, as:

$$NTP(e, w) = f(e, q, w), q \in [1..k] \tag{6}$$

with e a station, k the number of temporal patterns, q the index of a temporal pattern, w the index of a time window, and $f(e, q, w)$ is computed as:

$$f(e, q, w) = \frac{\tilde{\boldsymbol{W}}^w_{eq}}{\sum_{r\in[1..k]} \tilde{\boldsymbol{W}}^w_{er}} \tag{7}$$

where $\tilde{\boldsymbol{W}}^w_{eq}$ is the number of actions occurring for entity e, for temporal pattern q during window w.

For each station, we compute its mean NTP, mNTP, such as:

$$mNTP(e) = \{\frac{1}{\max(w)} \sum_w f(e, q, w), q \in [1..k]\} \tag{8}$$

where $\max(w)$ corresponds to the number of windows.

We then define the difference between two NTP by using the Kullback-Leibler divergence [16], also called Information divergence, which computes the difference between an observed probability distribution P and a reference probability distribution Q such as:

$$D_{KL}(P||Q) = \sum_{i} \log \frac{P(i)}{Q(i)} \tag{9}$$

The average KL divergence of a station informs us on how much the temporal patterns are unstable along time. High values correspond to stations that are continually changing, for which there is no typical distribution conserved during the studied period. The average KL divergence of a station e is defined as:

$$mD_{KL}(e) = \{\frac{1}{\max(w)} \sum_{w} D_{KL}(NTP(e,w)||mNTP(e))\} \tag{10}$$

The highest this value, the more the temporal patterns of this station have changed along time.

IV. Application on Lyon's BSS

A. Description of the dataset

Lyon's bicycle-sharing system has been already studied in previous papers. A more in-depth presentation of the system can be found in [17]. The dataset we are using in this article spans 5 years, from 2011 to 2015, and contains all trips done during this period, including station of origin, station of destination, time of arrivals and time of departure. Lyon's BSS stations have stayed mostly identical during the studied period, with 345 stations being active at least a year. The number of trips slowly increased, with an average of around 6 Million per year. The parameters we use to compute the TPUs and their evolution are the following:

- Entity: Bicycle sharing station
- Action: a bicycle arrives at a station
- Temporal Unit: hour
- Typical pattern period: Week
- Windows size: 1 year
- Shift between windows: 1 month

We adopt a 1-year window size to smooth out seasonal patterns. We label each time window by the last month included in this time-window. For instance, the time window labeled as July 2013 corresponds to trips done between August 2012 and July 2013. For each time window, we keep only stations for which we observe at least one activity for every month spanned by the time window, therefore removing stations that were temporally inactive.

B. Usage Patterns description

In this section, we present the Temporal Patterns of Usage (TPU) obtained by NMF. The NMF takes an argument, the number of features (TPU) wanted, k. There is no obvious method to choose k, as increasing k will shrink the reconstruction error, but at the cost of a higher complexity, making results less interpretable. We choose arbitrarily $k = 6$, meaning that we are searching for the 6 most relevant temporal patterns. We experimented with slightly different values, and observed that picking a lower k leads to merging most similar temporal patterns, while a greater value leads to some temporal patterns being split, but the main features remain consistent.

These patterns are presented in Fig. 1, and Fig. 2 show the corresponding strengths of each station (through their $mNTP$ values). We can note that despite the NMF process having no prior information about the location of stations, stations that are spatially close often have similar NMF profiles. This method can be seen as an alternative to methods already proposed to discover clusters of stations with similar patterns of activity (as the Poisson mixture model [18] or the Bayesian network approach [19]). While these clustering approaches associate only one cluster to each station, all stations belonging to a same cluster therefore being considered identical, the NMF approach considers that the activity at each station is a combination of several possible typical activities. For instance, in previous clustering approaches, a station has either a usage pattern typical of train stations, or a usage pattern typical of a commercial district, but not both. As can be seen in Fig. 2, in Lyon, the main shopping mall is located next to the west side of the main train station. As a consequence, BSS stations on the west of the train station have strong values both for TPU1 (commercial activity) and TPU4 (train station commuters activity), while BSS stations on the east of the train station have high values for TPU4 but not for TPU1.

Below, we propose to associate to each temporal profile a name, based on the shape of the temporal profile itself, and the stations associated with it. We must point at that these names have been chosen manually, and their accuracy could be discussed. They can, however, help us to interpret the kind of activity associated with each station.

- TPU1: commercial activity. Activity during typical opening hours of shops, museum and public places, but not on Sunday, during which most shops are closed in France. On the map, largest dots surround the main city mall and the main shopping street.
- TPU2: Leisure activity. Activity mostly on week-ends during afternoons, and after working hours during the week. This correlates spatially with public parks and banks of rivers.
- TPU3: Residential, restaurants, bars. Activity mostly during the evenings (people leaving their office and either coming back home, or going out), and around midnight on the days people tend to go out in France (Thursday, Friday, Saturday).
- TPU4: train stations. This activity is mostly present around the main train station of the city. Temporally, it is consistent with a commuting activity of people leaving the city before and after working hours.
- TPU5: Universities. Activity mostly during the morning rush hour and noon, but also around midnight on Thursday, Friday and Saturday, which corresponds to students coming back to dormitories in campuses. These patterns of activity spatially correlate with the main

universities and with other locations that correspond to student activity (neighborhoods with restaurants and bars ...). Unlike TP3, there is no after work commuting activity, as students usually do not have imperative to come back to their dormitories just after class.

- TPU6: Business district. Temporally, the activity mostly corresponds to the morning rush hours. Spatially, it correlates with the city center, where most of the economic activity concentrates.

C. Global evolution of the prevalence of Temporal Usage Patterns. We can observe that their proportions are mostly stable.

A first aspect that we can study is the global evolution of the temporal patterns, all stations taken together. In Fig. 3, we represent the proportion of trips corresponding to each TPU, for each time window. Because activities observed for stations in different time windows are merged together in a single matrix without explicit temporal information, and the NMF procedure has only this matrix as input, there is no technical reason for different time windows to have similar profiles. However the TPU prevalence stays remarkably stable across time windows, revealing that the way people used the BSS system did not change in the course of the five years spanned by the dataset. The only observable variable that could be considered significant is the fluctuations of TPU2 (activity correlated with leisure activities): decreasing during 2012, then increasing from Mid-2013 to early 2015. Several explanations could be proposed to this temporary change, such as a meteorological origin, an effect of city planning, or a change in pricing policy. For instance, new bicycle paths were created along the banks of Lyon's rivers in 2013, while the price of the yearly subscription raised from 15 to 25 euros in May 2012. Although this global analysis is not sufficient to understand in detail the evolution of the BSS, it can give us some insights, and reveal some broad behavioral changes that could later be studied more in-depth.

D. Analysis of mean-KL-Divergence

Fig. 4 is the distribution of the mD_{KL} values. The mode is reached at 0.005, the mean value is 0.001 and the standard deviation 0.0011. A few stations appear clearly as outliers, in particular values above 0.032, at more than 2 standard deviations from the mean. In Fig. 5, we present the evolution of TPU for 5 stations: the 3 stations of highest mD_{KL}, (7052, 9042, 7013), the stations of lowest mD_{KL} (7056), and a station taken among the ones with $mD_{KL} = 0.005$ (7033), i.e with the most common mean divergence.

To explain the observed changes of the highest mD_{KL} stations, one has to look at the history of their surroundings. For stations 7052 and 9042, we observe a gap in the data, meaning that the station was inactive during a period. After being active anew, the patterns have completely changed. By searching information regarding these stations, we finally found that both stations had been moved from one block to another, while staying in the same *arrondissement* (the city of Lyon is split in 9 administrative divisions called *arrondissement*). The dates of these changes perfectly match the observe discontinuity in the data. As we had no information about these changes in our dataset, we can see the method proposed as a good means to detect such irregularities in a collected dataset.

The case of station 7013 is different. For this station, there is no gap in the data, but we can observe large variations, that are correlated with changes occurring near the station: In January 2014, a new bridge was built just in front of the station and a streetcar line started to operate. In December 2014, a major new museum was opened to the public at a walkable distance. The important variations of TPU1 and TPU5 can reasonably be linked to these changes in the station surroundings.

E. Observing impacts of the city's change on BSS activity

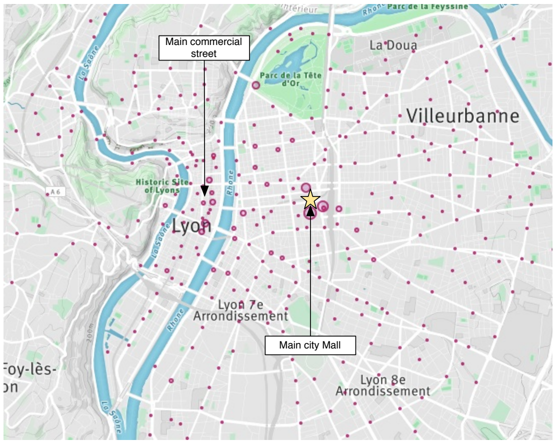
In the previous section, we have used an indicator to automatically detect the most important modifications in BSS usage in the city. We can also reverse the process, and check the impact of some known changes in the city or the BSS system on the way people use it.

A first case study corresponds to a small change in the position of a BSS station. We found that in May 2013, a station (2004) was moved for technical reasons, without interruption of service, from the west side of a public square to its east side, some 200 meters away. The same public square has another station (2022). In Fig. 6, we show how this modification affected the activity of these stations. Surprisingly, moving the station by less than a hundred meters on the same public space had a strong effect on the TPUs of the moved station, in particular a strong decrease in TPU4 (this public square is located in front of the second most important train station of the city). But this change also had a noticeable effect on the neighboring station, in particular an increase in TPU4, starting at the same period. We can assume that this change is due to a report of users arriving at the train station.

A second case study is the impact of the opening a large city mall, the second most important in the city, in April 2012. A new station was created in front of the mall, and opened simultaneously. We observe the evolution of the two closest stations, located respectively at 300 meters (station 2007) and 600 meters (station 2028), together with the station created in front of the mall (2005), in Fig. 7 Because the 2005 station was created simultaneously with the mall, we do not observe a rise in TPU typical of commercial activity (TPU1), but instead a rise in TPU3. This can be linked to the development of bars and restaurants in the surroundings of the mall. On the neighbouring stations however, we can clearly observe a rise in TPU1, which is likely to be a consequence of the opening of the mall. The station closest to the Mall is the most affected.

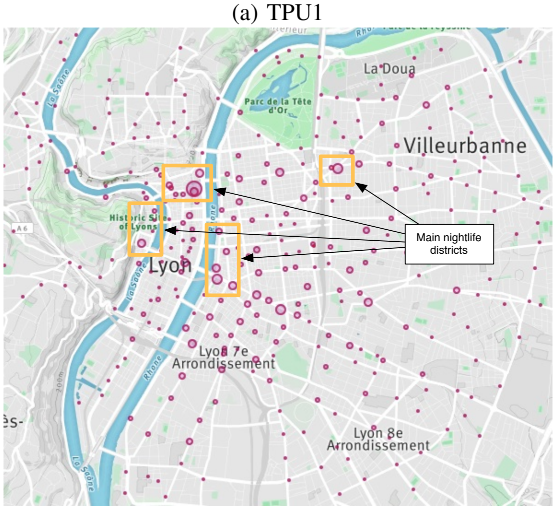
V. CONCLUSION

In this article, we have presented a new method for studying the evolution of the prevalence of Temporal Patterns of Usage in Bicycle Sharing Systems, using Nonnegative Matrix Factorization. We have shown how we can follow the usage evolution of the system as a whole. While we observed a remarkable



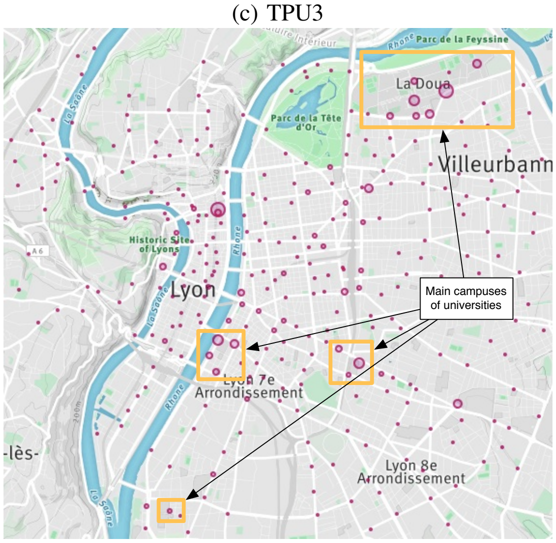
(a) TPU1

(b) TPU2

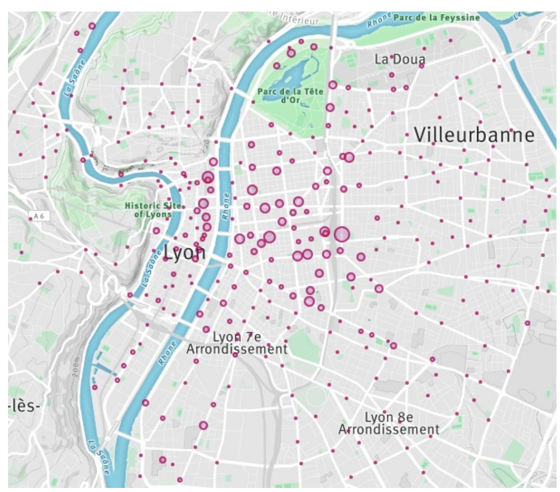


(c) TPU3

(d) TPU4



(e) TPU5



(f) TPU6

Fig. 2: Maps of the mean Normalized Temporal Profiles (mNTP) of each station for each TPU.

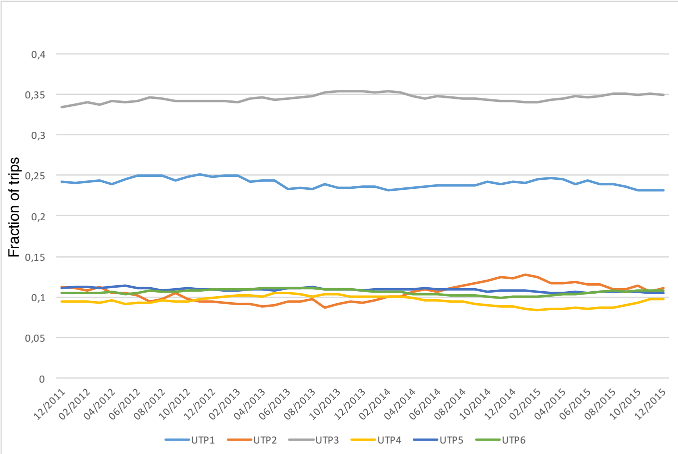


Fig. 3: Evolution of the prevalence of TPUs, all stations taken together

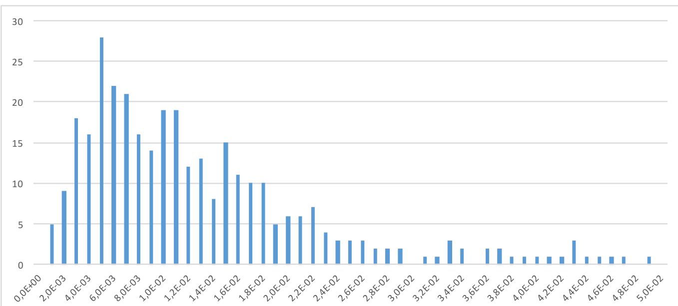


Fig. 4: Overall distribution of mean-KL-Divergence

stability for the city of Lyon, it would be interesting to apply the same process to different cities, and in particular to the first years following the opening of a new system, to see if users take some time to adapt to the system. It is also interesting to observe that an important increase in price did not deeply impact usage. This observation is complementary to a previous observation that the overall number of trips was not notably impacted by this change.

We have also shown how to detect important changes for some stations, by studying variations of their NTP . As we discussed, we could imagine to use this information for two usages: monitoring the BSS system, or monitoring changes in the city itself, following the principle of "social sensors".

One could for instance use such a tool to monitor, nearly in real time, if some activity becomes more or less important in a given neighborhood, using the station activity as a proxy.

This work could be extended in several directions. In this article, we have only studied the possibilities of analysis at the level of entities.. It is also possible to study how many users are active in each TPU, how the motions done in a given TPU are distributed among stations, or the frequency of entities being active in different TPUs, to give a few examples. In datasets for which metadata are available for entities, it would be possible to correlate these metadata with the profiles of entities. For instance, one could correlate socio-economical information of buffer zones, such as employment, resident population, presence of universities, park visitors, presence of bike lines, etc. with the TPU profiles of stations. We could later imagine using these correlations to predict the activity that a station would have if it were located in a given neighborhood.

Finally, improvements of this method could be investigated. Currently, we only investigate stable TPUs, that are relevant when considering all windows together. In the future, it could be relevant to develop a method able to identify TPUs relevant for some sliding windows only. In this dataset, we have only considered annual aggregation, because this allows us to ignore seasonal variations. But if we were doing the same type of analysis using monthly aggregation windows, it is likely that usages would change during holiday periods, compared with working weeks. However, because we are searching for TPUs that are relevant globally, and that holidays represent only short periods, such TPU would currently not be discovered by the method.

ACKNOWLEDGMENT

This work is funded in part by the ANR VelInnov: VelInnov ANR-12-SOIN-0001-02

The authors would like to thank JCDecaux and Cyclocity for sharing the data.

REFERENCES

[1] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[2] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis," *Neural Networks, IEEE Transactions on*, vol. 13, no. 6, pp. 1450–1464, 2002.

[3] W. Xu, X. Liu, and Y. Gong, "Document clustering based on non-negative matrix factorization," in *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*. ACM, 2003, pp. 267–273.

[4] J. Yang and J. Leskovec, "Overlapping community detection at scale: a nonnegative matrix factorization approach," in *Proceedings of the sixth ACM international conference on Web search and data mining*. ACM, 2013, pp. 587–596.

[5] R. Hamon, P. Borgnat, P. Flandrin, and C. Robardet, "Nonnegative matrix factorization to find features in temporal networks," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 1065–1069.

[6] R. Hamon, P. Borgnat, C. Févotte, P. Flandrin, and C. Robardet, "Factorisation de réseaux temporels: étude des rythmes hebdomadaires du système vélo'v," in *Colloque GRETSI 2015*, 2015.

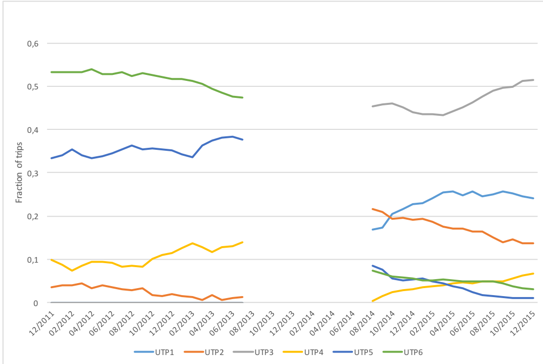
[7] C. Peng, X. Jin, K.-C. Wong, M. Shi, and P. Liò, "Collective human mobility pattern from taxi trips in urban area," *PloS one*, vol. 7, no. 4, p. e34487, 2012.

[8] Y. Han and F. Moutarde, "Analysis of network-level traffic states using locality preservative non-negative matrix factorization," in *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2011, pp. 501–506.

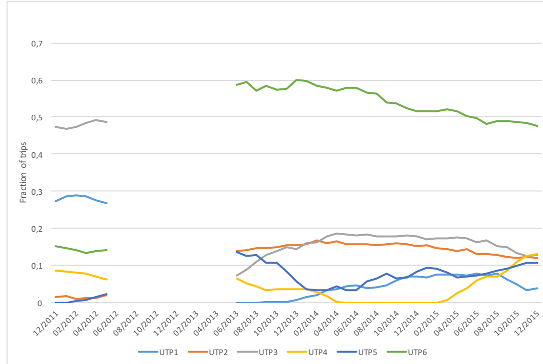
[9] C. Févotte and J. Idier, "Algorithms for nonnegative matrix factorization with the β-divergence," *Neural Computation*, vol. 23, no. 9, pp. 2421–2456, 2011.

[10] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

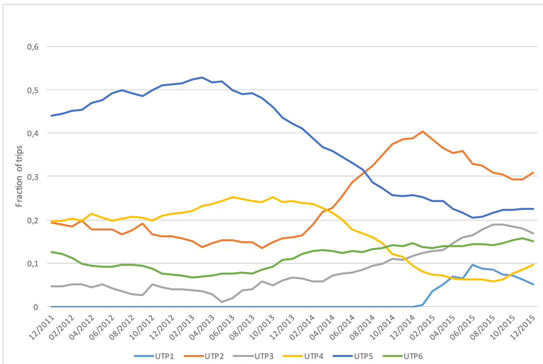
[11] C.-J. Lin, "Projected gradient methods for nonnegative matrix factorization," *Neural computation*, vol. 19, no. 10, pp. 2756–2779, 2007.



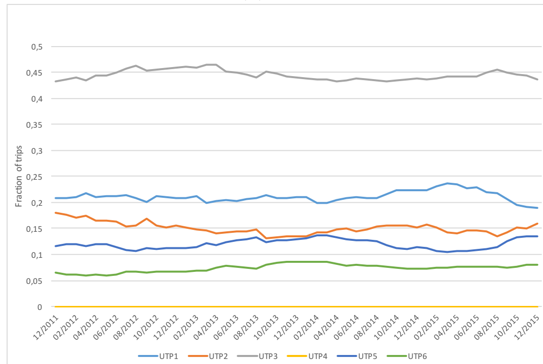
(a) 7052



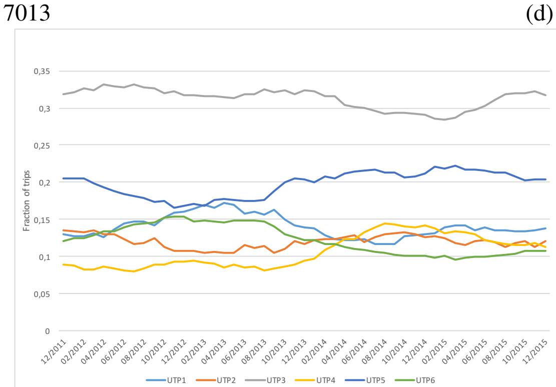
(b) 9042



(c) 7013



(d) 7056



(e) 7033

Fig. 5: Evolution of TPUs for 5 characteristic stations

[12] L. Guo, E. Tan, S. Chen, X. Zhang, and Y. E. Zhao, "Analyzing patterns of user content generation in online social networks," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2009, pp. 369–378.

[13] S. A. Golder, D. M. Wilkinson, and B. A. Huberman, "Rhythms of social interaction: Messaging within a massive online network," in *Communities and technologies 2007*. Springer, 2007, pp. 41–66.

[14] T. D. Tran, N. Ovtracht, and B. F. dArcier, "Modeling bike sharing system using built environment factors," *Procedia CIRP*, vol. 30, pp. 293–298, 2015.

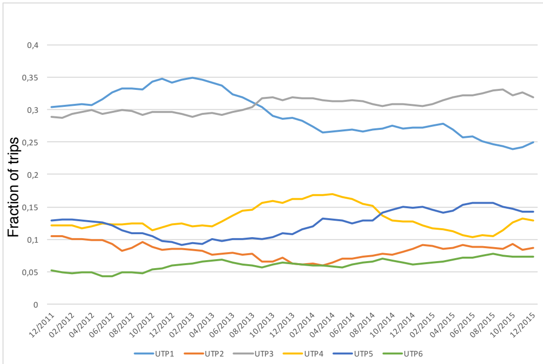
[15] G. Sagl, B. Resch, B. Hawelka, and E. Beinat, "From social sensor data to collective human behaviour patterns: Analysing and visualising spatio-temporal dynamics in urban environments," in *Proceedings of the GI-Forum*, 2012, pp. 54–63.

[16] S. Kullback and R. A. Leibler, "On information and sufficiency," *The annals of mathematical statistics*, vol. 22, no. 1, pp. 79–86, 1951.

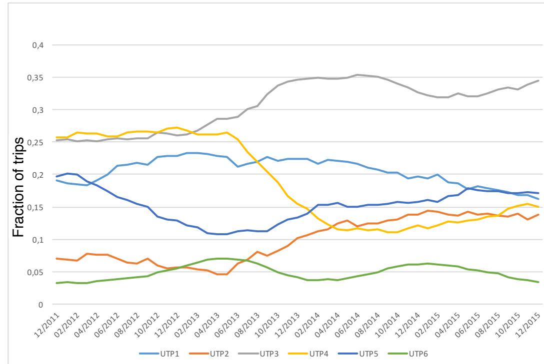
[17] M. Vogel, R. Hamon, G. Lozenguez, L. Merchez, P. Abry, J. Barnier, P. Borgnat, P. Flandrin, I. Mallon, and C. Robardet, "From bicycle sharing system movements to users: a typology of vélo'v cyclists in lyon based on large-scale behavioural dataset," *Journal of Transport Geography*, vol. 41, pp. 280–291, 2014.

[18] C. Etienne and O. Latifa, "Model-based count series clustering for bike sharing system usage mining: A case study with the vélibsystem of paris," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 5, no. 3, p. 39, 2014.

[19] J. Froehlich, J. Neumann, and N. Oliver, "Sensing and predicting the pulse of the city through shared bicycling." in *IJCAI*, vol. 9, 2009, pp. 1420–1426.

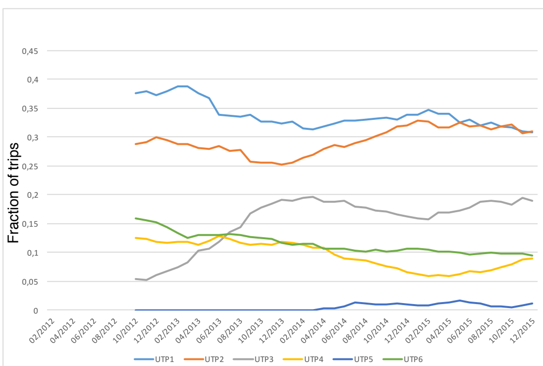


(a) 2022

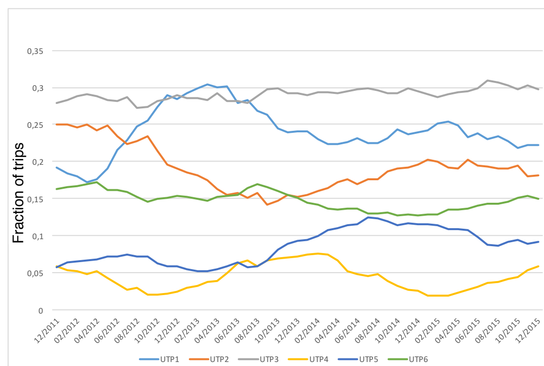


(b) 2004

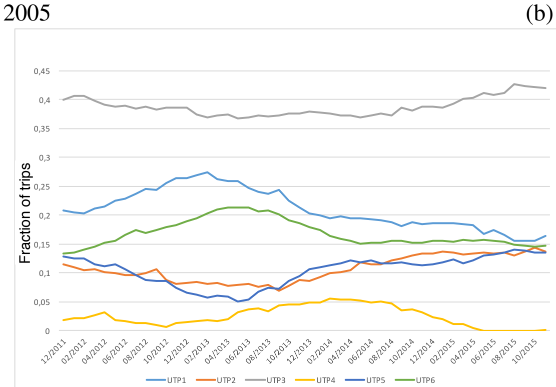
Fig. 6: Evolution of TPUs for neighboring stations 2022 and 2004. We can observe the impact of the relocation of station 2004 in May 2003



(a) 2005



(b) 2007



(c) 2028

Fig. 7: Evolution of TPUs for stations affected by the opening of a large shopping mall in April 2012