Elucidating Transition State Behaviour from Mobility Data by Cascades of Markov Chains

Michiel Dhont†, Elena Tsiporkova†
† EluciDATA Lab, Sirris, Bd A. Reyerslaan 80, Brussels, Belgium
‡ ETRO Department, Vrije Universiteit Brussel, Brussels, Belgium

Abstract
With the ongoing trend towards digitisation, vast amounts of, often very fine-grained, data are being collected. The ultimate goal is to capture and understand the behaviour of a system, such as the traffic in a city. However, making sense of such data is not straightforward due to its high level of detail and complex dependencies in time and space. Exploring heuristic approaches is essential to arrive at data representations that enable better insights into the underlying system dynamics by zooming out from the detail. In this paper, a novel approach for representing and reasoning about traffic state transition behaviour via a multitude of parameterised Markov chains models, cleverly designed to fit in a cascade, is proposed. The benefits of working with a multitude of individual Markov chains are outlined and subsequently, it is illustrated how to combine them into daily transition graphs such that their graph representation can be exploited to extract insights about daily traffic behaviour. In addition, targeting context-specific studies, an alternative approach is introduced combining in a dynamic fashion a cascade of Markov chains covering longer and overlapping time windows. A recursive algorithm is conceived and validated allowing to exploit this cascade structure for computing state transition probabilities over time. The potential of the proposed approach for mining traffic state transitions is demonstrated on a use case derived from real-world data.

Keywords: mobility data, traffic states, Markov chain, Markov clustering, reasoning

1. Motivation

Today, people and companies store sensor data from a wide range of systems with the goal to monitor its behaviour over time [1]. However, in practice, it is often hard to interpret the behaviour of a system based on complex multi-variate time series originating from dynamic operating contexts. Data collected from traffic monitoring systems are even more difficult to understand and extract meaningful insights from due to the additional complexity caused by the spatial dimension [2]. Furthermore, a variety of external factors, including public events, road construction, and weather conditions, significantly impact traffic behaviour, making it difficult to identify their individual effects.

One way to grasp the big/overall picture is to try zooming out from the detail, e.g. by extracting a limited set of discrete states, each representing a unique context or temporal state of the (traffic) system for a certain time period. The identified temporal states can be labelled (annotated) by further subjecting them to in-depth contextual analysis allowing to reveal the underlying dynamics of the temporal behaviour observed. Semantically annotated temporal states can be very powerful instrument for studying system behaviour in terms of state transitions in time. Markov chains are a very suitable mathematical formalism to study state transitions. However, it is not trivial to model traffic dynamics using Markov chains. Traffic, and in particular urban traffic, is typically a periodic phenomenon dictated by people’s daily/weekly/monthly routines. This implies that the probability of moving from one state to another can be very different during the day, and may also vary between weeks and months depending on the season. This requires a modelling framework that adequately accounts for traffic periodicity.

© 2023 by author(s) as listed above. The article is licensed under a Creative Commons Attribution (CC BY 4.0) International license (https://creativecommons.org/licenses/by/4.0/legalcode), except where otherwise indicated with respect to particular material included in the article. The article should be attributed to the author(s) identified above.
We propose here a novel approach for representing and reasoning about traffic state transition behaviour via a cascade of Markov chains. The conceived cascade of Markov chains is carefully designed to capture the transition state behaviour with high granularity, while guaranteeing the robust estimation of the transition probabilities assigned to the individual chains in the cascade. Working with a multitude of smaller Markov chain models opens several different opportunities for their employment in traffic state transition analysis. For instance, a daily (24-hour period) transition matrix (graph) can be composed by chaining one after the other a cascade of suitably constructed Markov chains over equidistant and consecutive time windows covering 24 hours. Such a daily transition matrix can be subjected to various transformations, e.g. using Markov clustering, which allow to study daily traffic behaviour across different locations and seasons. Alternatively, daily traffic behaviour can be studied with a cascade of Markov chains constructed over overlapping time windows, which allows for more accurate representation of traffic state transitions.

The potential of the proposed modelling and reasoning framework is demonstrated on a real-world use case. Namely, experiments are conducted on traffic states extracted and annotated in our previous work (for more detail see [3]) using a real-world traffic dataset of 16 locations along a busy road. Our experimental validation demonstrates how the proposed approach can be exploited to obtain better understanding of traffic behaviour and gain deeper insights into the underlying traffic state transition dynamics. We also convincingly exemplify how relevant questions for traffic operators, such as "What is the chance that the current traffic congestion will be resolved in a given amount of time?", can be answered with our cascade of Markov chains.

2. Related Work

In the field of mobility, geo-referenced time series are commonly used. The fine-grained spatio-temporal character makes it very complex for traffic operators to extract insights from such data, e.g. in real time traffic monitoring systems. In order to facilitate the analysis, one can try to convert (discretise) the time series data into a set of labels representing distinct traffic states, e.g. free-flow, congestion and agitated traffic.

The approaches in the literature for extracting traffic states from multi-variate time series can be roughly split in two streams. The first option is to extract interpretative features which characterise the traffic condition for the moments of interest. Based on those features, traffic states can be extracted by well-known clustering algorithms such as k-means clustering [4], agglomerative clustering [5], spectral clustering [6, 7] and fuzzy clustering [8, 9]. Constantinos et al. [10] propose an approach to define clusters based on differentiated time series (flow, density, and speed in their use case). In this way, clusters identify states with similar increase or decrease conditions, allowing to characterise expected changes in traffic and identify anomalies. A drawback of this strategy is that no direct link is made between the transitions and traffic conditions as understood by humans, e.g. free-flow or congestion. The second option is to acquire traffic states using a deep black-box method. Asadi and Regan [11] propose an approach for extracting temporal and spatial clusters with a deep embedded clustering (DEC) model. A DEC model learns to map the data into a lower-dimensional latent feature space using a reconstruction loss function as done in an encoder-decoder neural network (auto-encoder), but also a clustering loss function which aims to arrive at dense cluster distributions by pulling samples near cluster margins [12].

The extraction of well characterised and semantically interpreted states allows for further investigation of the state dynamics of the system, e.g. the transitions in between two different states might elucidate new insights since they indicate a change in behaviour. In the literature, traffic is typically modelled directly by traffic flow or occupancy rates (e.g. [13, 14]). In this approach, physical characteristics of the infrastructure define the model
constraints, while the dynamics are modelled using, for example, a hidden Markov model. Alternatively, Wang et al. [15] illustrate how to model one-step transition probabilities between 6 daily quasi-stationary daily traffic states. Wang et al. mention that traffic does not follow Markovian behavior, i.e. transition probabilities change over time. As a result, they can only extract transition probabilities over short time windows. Our paper solves this problem by dynamically chaining a selection of time-specific Markov chain models.

3. Materials and Methods

3.1. Data and Use Case

This paper considers data coming from 16 automatic number plate recognition (ANPR) cameras on the small ring of Brussels, the capital of Belgium. Their geographical locations are visualised in Figure 1. The dataset contains the average velocity, vehicle flow and road occupancy per minute for approximately two years and three months (16/01/2020 - 04/05/2022). This dataset was mined in real-time from an open API of Brussels mobility.

3.1.1. Traffic State Extraction

In a previous work [3], we proposed a multi-stage approach to extract (temporal) traffic states. For this approach, 8 relevant features (conceived in consultation with traffic domain experts) were first extracted per 15-minute time segment. The k-means clustering algorithm was then employed to categorise the dataset (see Section 3.1) into 6 distinct traffic states. Subsequently, kernel density estimations of the derived features and the fundamental traffic diagrams were utilised to derive semantic interpretations per cluster as summarised below:

● A) Traffic build-up: Traffic density is gradually increasing within the time window of 15 minutes.
● B) Free-flow: Only few vehicles are detected on the streets. People can drive their desired speed (are not blocked by other drivers), resulting in a wide range of average velocities.
● C) Stable non-saturated traffic: In this state, many vehicles are on the road. However, the roads are not yet reaching saturation, such that the mean velocity is only slightly influenced by the flow rate. Within the 15-minute time window, density is rather stable.
● D) Traffic intensity reduction: This state has very similar characteristics as state A, but with the exception of a clear decrease in traffic density within each time window.
● E) Congested traffic: Traffic is slow due to too many vehicles on the road, leading to inefficient road usage.
● F) Variable non-saturated traffic: Traffic occupation is similar as in traffic state C, although, large fluctuations of density are observed within each time window.

A potential shortcoming of the above clustering strategy is the risk of missing some latent features, i.e., creating a blind spot, since the features are derived manually based on some prior (potentially limited) domain knowledge. As an alternative to the manual feature extraction, we experimented with a deep learning approach to extract features in a black-box manner. For this purpose, a convolutional neural network auto-encoder was trained to map each of the 15-minute time segments composed of the 3 raw (normalised) parameters into a latent space vector of 8 features. Subsequently, those 8 features were clustered following the same approach used for the manually extracted features. Despite resulting in the same

1Brussels mobility API: https://data.mobility.brussels/traffic/api/counts/
number of clusters, the clustering approach using deep features resulted in differently located
temporal clusters compared to those obtained with the manually extracted features. The
deep features are probably capturing different (hidden) characteristics of the data. However,
the latter makes it much harder to interpret the new clusters in terms of meaningful traffic
states. Moreover, further analysis indicated a somewhat arbitrary division between the new
(deep) clusters, inconsistent with how traffic engineers describe traffic situations. For this
reason, the experiments in the remainder of this study are illustrated on the traffic states
extracted via the manually selected features.

3.2. Markov Chains

A Markov chain is a stochastic model which captures the probability of the transition
between any two states in the state space \( \{S_1, S_2, \ldots, S_M\} \). The Markov chain considers
discrete time steps (e.g. per minute, hour, day...), and allows for self-loops, i.e. transition
to the same state. It further assumes that the transition to the next state does only depend
on the present state, i.e. it has no memory about the past states.

A Markov chain model can be represented by a transition matrix \( P \) of order \( M \), with \( p_{ij} \)
the transition probability from state \( i \) to \( j \). If the number of states is not too high, one can
visualise the Markov chain model as a directed weighted graph as shown in Figure 2a. Each
of the \( M \) states is visualised as a node, and the weights on the directed edges denote
the transition probability. Based on the initial state, one knows the probability of each state
at the next time step. Moreover, the chance of arriving at a specific state \( j \) in \( n \) time steps,
given the current state \( i \) can be calculated with the recursive function as follows [16]:

\[
p^n_{ij} = \sum_{k=1}^{M} p^{n-1}_{ik} \cdot p_{kj}.
\]  \hspace{1cm} (3.1)

An interesting characteristic of the transition matrix of a Markov chain is that the proba-
bilities in Equation (3.1) can also be obtained by raising \( P \) to the \( n \)th power. For instance,
the transition probabilities of arriving at each state from each state in 5 time steps, can be
computed by taking the 5th power of \( P \). For an increasing forecast horizon, this approach
will be less informative as it reaches the so-called steady state, representing the overall
probability of each state (i.e. neglecting the origin state). Note that only regular Markov
chains converge over time. The non-regular Markov chains will loop periodically through a
finite set of probability matrices. Following the theorem of Wielandt [17], a Markov chain
is considered regular if there exists a power \( n \) of the transition matrix which has only strict
positive values, with \( n \in [1, 2, \ldots, (M-1)^2 + 1] \).

3.3. Markov Clustering Algorithm

If the nature of a dataset allows that it can be represented as a graph, it may be advan-
tageous to exploit this structure when extracting clusters. The Markov clustering algorithm
(MCL) is such an elegant and effective graph-based clustering approach, based on the ran-
dom walk principle. To apply MCL, the weight matrix of the graph is first normalised in a
way that the columns sum up to 1 (similar as the transition matrix from a Markov chain).
From here the algorithm alternates between expansion and inflation until convergence:

• **Expansion**: helps making farther nodes reachable by taking the \( e \)th power of the matrix
• **Inflation**: helps to strengthen strong neighbours and demoting less close neighbours by
taking the \( r \)th power of each individual value in the matrix, followed by a normalisation.

In the above, \( e \) and \( r \) are hyperparameters which need to be optimised. After conver-
gence, low probability transitions of the matrix are pruned away, resulting in multiple sub-
graphs [18].
One approach to find optimal values for the hyperparameters $e$ and $r$ is by calculating the modularity $Q$ after applying the cluster algorithm for different combinations of $e$ and $r$. The hyperparameters for which the clustering results produce the highest $Q \in [0, 1]$ are considered to be the most appropriate. $Q$ can be calculated as follows:

$$Q = \sum_{i=1}^{K} (c_{ii} - a_{i}^2),$$

(3.2)

with $K$ the number of clusters, $c_{xy}$ is the fraction of all edges in the network that link nodes from cluster $x$ to cluster $y$. More explicitly, $Q$ is the sum of the differences between the edge distribution observed within each cluster ($c_{ii}$) and the expected one if the edges were shuffled randomly ($a_{i}^2$) [19].

### 3.4. Markov Chains of Traffic States

The ability to convert the mobility data stream into a temporally ordered sequence of state labels enables a more systematic examination of the state transition dynamics of the targeted mobility system through the application of Markov chains.

Let us assume that $M$ distinct temporal traffic states have been identified following the approach described in Section 3.1.1. Subsequently, the time series traffic data can be converted into a temporally ordered sequence of state labels by replacing for all time windows per location the features vectors with the corresponding temporal traffic states. Each neighbouring pair of labels can be seen as denoting a transition in time between states. Thus, the transition probability $p_{ij}$ of moving from a given state $i$ to state $j$ can be reliably estimated for different time periods and locations from historical data as follows:

$$p_{ij} = \frac{t_{ij}}{M \sum_{k=1}^{K} t_{ik}},$$

(3.3)

with $t_{ij}$ the the total number of transitions from state $i$ to state $j$. In this manner, a transition probability matrix $P$ (as explained in Section 3.2) can be estimated per specific location or time period of interest, allowing to construct Markov chains as shown in Figure 2a. This allows mining and comparing state transition behaviour across different locations and time periods. The comparison is facilitated by the fact that our temporal states have been derived in such way that they have the same semantic meaning across different locations and different time periods.

Although Markov chains are quite powerful, they have certain limitations when it comes to realistically reflecting the dynamics of the mobility system. Traffic, and in particular urban traffic, is typically a periodic phenomenon dictated by our daily/weekly/monthly routines. This implies that the probability of moving from one state to another can be very different during the day, and also may vary between weeks and months depending on the season. Therefore, it does not make sense to try to construct an overall, ignoring the time of the day, Markov chain connecting the different traffic states. Such a representation does not discriminate between the different moments in a day when estimating the state transition probabilities and thus will not account for the periodic behaviour during the day. There is also an additional negative side effect to not discriminating between the different moments in a day, namely this will lead to extremely high probabilities of staying in the same state, which in its turn will result into negligibly low probabilities associated with transiting to any other state. We therefore introduce a parameterised Markov chain model $M_{s,t_{start},s}$, defined by 3 key parameters:

- timeframe granularity $s$ expressed in minutes;
- start time $t_{start}$ taking values between $T_{min}$ and $T_{max}$ (e.g. 0:00 and 24:00);
• timeframe coverage \( \kappa \in \mathbb{N}_{>0} \) defining the time window \([t_{\text{start}}, t_{\text{stop}}]\) over which the transition probabilities are estimated, with \( t_{\text{stop}} = t_{\text{start}} + s \cdot \kappa \).

In this way, a family of Markov chains \( \mathcal{M} \) is being constructed, as formally defined below:

\[
\mathcal{M} = \left\{ M_{s,t_{\text{start}},\kappa} \mid s \in \mathbb{R}_{>0}; \ t_{\text{start}} \in [T_{\text{min}}; T_{\text{max}}); \ \kappa \in \mathbb{N}_{>0} \right\}. \quad (3.4)
\]

\( \mathcal{M} \) can be considered as the set of building blocks available to study the traffic state transitions across various temporal horizons. Note that for small values of \( \kappa \) (e.g. \( \kappa = 1 \)), the transition probabilities of the Markov chains are calculated over relatively small time windows. To find the best trade-off between fine-grained daily coverage vs. reliable transition probabilities estimates, it is important to evaluate the corresponding confidence intervals (CI’s). As proposed by Goodman [20], the confidence interval \( CI_i \) of state transition \( i \) can be found as follows:

\[
CI_i = [p_i - \delta_i, p_i + \delta_i], \quad \delta_i = \sqrt{\frac{\chi^2(\alpha, 1) p_i (1 - p_i)}{N}},
\]

with \( N \) the size of the considered population of state transitions, \( p_i \) the estimated probability of state transition \( i \), \( \delta_i \) the maximum deviation of state transition \( i \) for the desired confidence level, \( \chi^2(x, y) \) the upper \( 1 - x \) quantile of the \( \chi^2 \) distribution with \( y \) degrees of freedom, \( \alpha \) equals 1 minus the desired confidence level (often \( \alpha = 0.05 \)) and \( M \) the total number of possible state transitions.

Using the Markov chain paradigm several aspects of a mobility system can be examined. It is feasible, for instance, to compute the probability of an exact temporal sequence of traffic states at a given moment in time. Additionally, it is possible to derive the likelihood of arriving at a specific state following a predetermined number of time steps, predicated on the current state. Another interesting option is to determine the system equilibrium, i.e. determine whether and how fast the steady state is reached.

3.5. Modelling and Reasoning Frameworks for Mining State Transitions

The family of parameterised Markov chain models defined in the foregoing section gives rise to the development of modelling and reasoning frameworks. The rationale behind such frameworks is to empower and facilitate flexible exploration of traffic state transitions, e.g. by creatively connecting individual Markov chains in cascades. Such cascade approach allows to realise different use cases by carefully selecting a multitude of Markov chain models, which can capture the behaviour of state transitions with relatively high granularity, while guaranteeing the robustness of the transition probabilities assigned to the individual chains in the cascade.

In this section we develop two different modelling and reasoning frameworks: 1) zooming in the daily state transition dynamics by rolling out individual Markov chains over time into daily transition graphs; 2) estimating the probability of arriving at any state during the day given the present state by chaining individual Markov chains in a dynamic fashion.

3.5.1. Zooming in Daily State Transition Dynamics

A daily transition graph (matrix) is composed by chaining one after the other (in a cascade) a set of suitably constructed Markov chains over equidistant and consecutive (non-overlapping) time windows covering any period (i.e., of at most 24 hours, with \( T_{\text{min}}=0:00 \) and \( T_{\text{max}}=24:00 \)) during the day. Formally, for any timeframe granularity \( s \) and coverage \( \kappa \), a family of Markov chains can be constructed as shown below:

\[
\mathcal{M}_{s,\kappa} = \left\{ M_{s,t,\kappa} \mid t = t_0 + i \cdot s \cdot \kappa; \ t_0 \geq T_{\text{min}}; \ t \leq T_{\text{max}} - s \cdot \kappa; \ i = 0, 1, 2, \ldots \right\}. \quad (3.6)
\]
The individual Markov chain models can be stacked one after the other chronologically in time. This practically means that the original transitions graph (Figure 2a) is rolled out in time as illustrated in Figure 2b. Such a daily transition matrix can be subjected to various transformations, e.g. using Markov clustering (see Section 3.3), which allow to study daily traffic behaviour across different locations and seasons. The resulting clusters provide insights into which combinations of traffic states and time windows frequently occur in a similar context. Additionally, it is compelling to examine if and how those clusters change when considering only a subset of the data to calculate the state probabilities, such as only considering a specific location, vacation period, corona lockdown, weekend, etc.

One drawback of the daily transition graph approach for small values of $\kappa$ (e.g. $\kappa = 1$) is that a fine-grained cascade of Markov chains is required. This makes it challenging to arrive at small enough CIs for the transition probabilities when the available historical data is not large enough. In such contexts, it might be wise to trade the fine-grained daily coverage granularity for a larger $\kappa$. Alternatively, as illustrated in the next subsection, daily traffic behaviour can be studied with a cascade of Markov chains constructed over overlapping time windows, which allows for more accurate and robust representation of traffic state transitions, while preserving the high coverage granularity.

It is worth noting that a transition graph can also be rolled out along the spatial dimension, e.g. along successive locations of a trajectory, rather than the temporal dimension.

### 3.5.2. Estimating the State Transition Probabilities Across the Day

Suppose that we want to construct and explore a daily transition graph for only a portion of the data, such as a particular location or period of interest. There is a risk that the population of state transitions is too small, which will result in too wide CIs. For this reason, we propose an alternative approach, which allows to dynamically cascade larger (i.e. for larger values of $\kappa$) individual Markov models and covers potentially overlapping time windows. In this way, every state transition has a larger population, resulting in smaller CIs (see Equation (3.5)). If we assume that all the different transitions are uniformly distributed across all time windows of width $\kappa$, than the population of those transitions will increase with factor $\kappa$, resulting in a CI that is narrowed down with a factor of $\frac{1}{\kappa^3}$.

For any given timeframe granularity $s$, let us consider a family of Markov chains, all having a coverage $\kappa > 1$ and each one starting at a consecutive timeframe, i.e. the individual Markov chains are covering overlapping windows:

$$
\mathcal{M}_{s,\kappa} = \left\{ M_{s,t,\kappa} \mid t = t_0 + i \cdot s; \ t_0 \geq T_{min}; \ t \leq T_{max} - s; \ i = 0, 1, 2, \ldots \right\}. \quad (3.7)
$$

The aim is to conceive a computational workflow (algorithm) allowing to connect a selection of Markov chains in a cascade, which can be used for estimating the probability of arriving at any state during the day, given the present state. The following recursive algorithm is
proposed, where \( p_0 \) is a vector of length \( M \), giving the starting state probability distribution at time \( t_0 \) and \( P(s, t_i, \kappa) \) is the probability matrix associated with Markov chain \( M_{s,t_i,\kappa} \), for \( i = 0, 1, 2, \ldots \):

**Algorithm 1 Estimating Transition Probabilities via a Cascade of Markov Chains**

**Precondition:** \( p_0 \) is a vector of length \( M \), giving the starting state probability distribution at time \( t_0 \), \( n_{steps} \) is the number of time steps in the future for which we want to know the state probabilities, \( \kappa \) is the frame size and \( s \) the frame granularity.

1: function \( \text{CALCULATE\_TRANSITION\_PROBABILITIES}(p_0, t_0, n_{steps}, \kappa, s) \)
2: if \( n_{steps} = 0 \) then
3:     return \( p_0 \)
4: else
5:     \( i \leftarrow \min(n_{steps}, \kappa) \)
6:     \( p_{current} \leftarrow p_0 \ast \text{power}(P(s, t_0, \kappa), i) \)
7:     \( t_{current} \leftarrow t_0 + i \ast s \) \hspace{1cm} \triangleright \text{Update starting point of frame}
8:     \( n_{steps} \leftarrow n_{steps} - i \) \hspace{1cm} \triangleright \text{Update remaining time steps}
9:     return \( \text{CALCULATE\_TRANSITION\_PROBABILITIES}(p_{current}, t_{current}, n_{steps}, \kappa, s) \)
10: end if
11: end function

4. Real-world Experiment

In this section, the traffic states from the use case as described in Section 3.1.1 are used to demonstrate the potential of the proposed modelling and reasoning methodology. The traffic states have been derived for timeframe granularity of 15 minutes, i.e., \( s = 15 \) in everything that follows below.

4.1. Markov Chains of Traffic States

As described in Section 3.4, our traffic states can be used to construct Markov chains by computing transition probabilities over a period of interest. Calculating the transition probabilities over the whole historical dataset at our disposal, we obtain the Markov chain depicted in Figure 3a. It can be observed that any traffic state has a probability of at least 0.4 to stay in the same state, which is in particular high for the free-flow (B) and congested traffic (E) states, reaching 0.77 and 0.89, respectively. Although the overall Markov chain highlights some dependencies between the traffic states, they are rather general because they are summarising the behaviour over different contexts during the day. More concrete, traffic characteristics are expected to differ depending on the moment of the day, e.g., whether it is weekend and which specific location we are investigating. In Figures 3b to 3f, a selection of Markov chains is depicted for different starting times and \( \kappa = 4 \), i.e., an hour. A close inspection of those graphs confirms our assumption that considering Markov chains over shorter time periods is effective since the new graphs display different state transition characteristics. For example, at night (Figures 3b and 3c) the edges to arrive to traffic state B (free-flow) from any state suddenly appear, while during the day (Figures 3d to 3f) some edges arise in the opposite direction, i.e., the chance of moving from free-flow to another state increases.

4.2. Zooming in Daily State Transition Dynamics

As proposed in Section 3.5.1, a subset of Markov chains over equidistant and consecutive time windows can be linked chronologically into a daily transition graph allowing to capture the daily characteristics of traffic state transitions. For instance, we have constructed such
a daily transition graph when using all the available data and linking consecutive Markov chains (i.e. starting at consecutive time stamps) with coverage $\kappa = 1$. Subsequently, the resulting transition matrix was subjected to Markov clustering as outlined in Section 3.3, resulting in splitting the temporal traffic states into 10 different clusters across the day. In Figure 4a, the resulting graph after the clustering is depicted. The thickness of the edges in this graph are weighted based on the values of the transition probabilities and edges in-between clusters are pruned away to facilitate interpretation. It is interesting to observe that the clustering elucidates a certain daily interaction pattern among the different traffic states. For instance, traffic states A (build-up) and D (intensity reduction) are continuously clustered together along the day, meaning that these traffic states frequently alternate, but also that there is no systematic daily moment when this alliance ends or evolves into a third traffic state. The remaining 9 clusters highlight clear temporal bounds mostly where one could expect them: 1) at night, having clusters 2 and 3, indicating predominantly alternations between free-flow and non-saturated traffic; 2) early morning rush hours with many active clusters 4, 5, 6 and 7, indicating rapid changes in traffic conditions; 3) morning and afternoon traffic, dominated by clusters 8 and 9, respectively, which denote intensive alterations between stable and variable non-saturated traffic states; 4) and finally the evening rush hours with cluster 10, which denotes very different transition pattern in comparison to the early morning rush hours and also lasts longer than the morning one.

Besides the strong daily pattern, traffic is also expected to exhibit a weekly seasonality. For instance, Figure 5 depicts the occurrences of each cluster in function of the time of the week. One interesting insight that can be derived from this figure is that cluster 3 (free-flow and non-saturated traffic) is often shortly interrupted before midnight by cluster 2 (free-flow and congested traffic), while this appears less frequently on Friday and Saturday night.

To take into consideration the different aspects of a particular location such as the quality of the road infrastructure or characteristics of the surrounding area, it is interesting to zoom in on each individual location. Unfortunately, the latter would not end up in reliable transition probabilities due to relatively low counts of state transitions per location. Section 4.2 and Figure 4b depict the $CI$’s for one particular location for the frame starting at 4 AM with $\kappa = 1$ and $\kappa = 8$, respectively. It can be observed that for $\kappa = 1$, most of the $CI$’s are
too large i.e., being hardly informative, while for $\kappa = 8$, the CI’s become manageable again. Subsequently, the daily transition graph for the same location and $\kappa = 8$ was constructed and subjected to Markov clustering. Figure 4b depicts the resulting graph, which (although less granular than the one over all locations) still elucidates insightful transition patterns.

4.3. Estimating the State Transition Probabilities Across the Day

In addition to providing a powerful mining environment allowing to reason about state transition behaviour and generate insightful visual interpretations, the representation of traffic state transitions via a multitude of Markov chain also allows to address, as proposed in Section 3.5.2, relevant questions related to expected traffic states given the present situation. For instance, a traffic operator might want to know how long it will take for congested traffic to become free-flowing, or even more specific, what the chances are of still being in congested traffic after a certain amount of time. For this purpose, it is important to carefully select days with the right context when constructing the models. The context might be temporal (such as weekend, workday or holiday) and/or spatial (only consider a specific location). However, as already explained above, the larger the CI’s for the estimated transition probabilities become (see Figure 6).
Subsequently, focusing on the location from Figure 6 and following the procedure of Algorithm 1 for timeframe coverage $\kappa = 8$, the probabilities to arrive at any of the traffic states starting from state E (congestion) for different forecast horizons within a day have been calculated and depicted in Figure 7. The chance of moving from congestion to free-flow increases gradually with the forecast horizon. This is expected since these are the two most extreme and opposite to one another situations. Further, it can be observed that during the day, traffic state E has high probabilities to evolve into states C and F (i.e., either stable or variable non-saturated traffic), while in the evening it is most likely to arrive at traffic state B, which means the traffic intensity gets reduced drastically. Additionally, we investigated the chance of remaining in the same state for an increasing forecast horizon as depicted in Figure 8. The probabilities to remain in traffic states B (free-flow) and E (congestion) seem to be the ones which are most negatively affected (decreased) by the forecast horizon, i.e., they appear to be the most temporally unstable states. During the day, the other 4 states show almost no decline for an increasing forecast horizon.

5. Conclusion

Traffic is a complex and dynamic phenomenon that exhibits periodicity across time periods of the day, the week, the month, and the year. This makes it challenging to study and understand temporal state transitions using traditional probabilistic frameworks. In this article, we propose a heuristic approach that creatively and flexibly uses a multitude of Markov chain models to account for traffic periodicity. Our approach enables the modelling and interpretation of traffic state transition behaviour over different time periods and locations. We demonstrated the potential of our approach to elicit and reason about state transition behaviour in two different scenarios: 1) zooming in on daily state transition dynamics; 2) estimating state transition probabilities across diverse forecasting horizons. These scenarios by far do not exhaust the different opportunities offered by our methodology, which we intend to explore further. For instance, we are interested in modelling traffic state transition behaviour over different trajectories.
Acknowledgements

This research was subsidised through the project MISTic by the Brussels-Capital Region – Innoviris and received funding from the Flemish Government (AI Research Program).

References