Neurocomputing 203 (2016) 22–33

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Event detection from traffic tensors: A hybrid model

Hadi Fanaee-T [a,*], João Gama [b]

[a] Laboratory of Artificial Intelligence and Decision Support, FCUP/University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal
[b] Laboratory of Artificial Intelligence and Decision Support, FEP/University of Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

ARTICLE INFO

Article history:
Received 31 December 2014
Received in revised form
24 February 2016
Accepted 17 April 2016
Communicated by Jianbin Qiu
Available online 6 May 2016

Keywords:
Traffic data
Origin/destination matrix
Tensor decomposition
Tucker
Core size

ABSTRACT

A traffic tensor or simply *origin* × *destination* × *time* is a new data model for conventional origin/destination (O/D) matrices. Tensor models are traffic data analysis techniques which use this new data model to improve performance. Tensors outperform other models because both temporal and spatial fluctuations of traffic patterns are simultaneously taken into account, obtaining results that follow a more natural pattern. Three major types of fluctuations can occur in traffic tensors: mutations to the overall traffic flows, alterations to the network topology and chaotic behaviors. How can we detect events in a system that is faced with all types of fluctuations during its life cycle? Our initial studies reveal that the current design of tensor models face some difficulties in dealing with such a realistic scenario. We propose a new hybrid tensor model called HTM that enhances the detection ability of tensor models by using a parallel tracking technique on the traffic's topology. However, tensor decomposition techniques such as Tucker, a key step for tensor models, require a complicated parameter that not only is difficult to choose but also affects the model's quality. We address this problem examining a recent technique called adjustable core size Tucker decomposition (ACS-Tucker). Experiments on simulated and real-world data sets from different domains versus several techniques indicate that the proposed model is effective and robust, therefore it constitutes a viable alternative for analysis of the traffic tensors.

1. Introduction

The understanding and characterization of traffic tensors has many applications in network information systems, transportation systems and many other areas. In particular, event detection in these systems enables operators to make better decisions about emerging problems and perform some prevention tasks. Some intuitive examples include the identification of attacks and malicious activities inside networks and traffic jams in transportation systems. However, one of the most serious problems in event detection from traffic tensors is the complexity and diversity of event types. Fig. 1 illustrates a simplified scenario of a hypothetical bike sharing system during the operational period of 100 days. The stations are specified by letters A to F and the connections between them are represented by directed lines. For extra simplicity let us assume that the system shows a normal behavior with stable traffic among four stations of A, B, C and D and no traffic from nodes E and F until day 95. Let us also presume that events take place in the system exclusively during days 95–100.

The first event type occurs on days 95 and 96 when the system experiences an increase in traffic flow throughout the network. As we can see, a constant amount is added to the volume of traffic between all stations. This is similar to what happens in a large impact citywide event such as a big festival which affects the city's whole population. For this example, $t=95$ relates to a severe event and $t=96$ to a low-scale event. Note that the alteration in traffic volume is not necessarily additive. Weather-related events such as rain lead to the same patterns, but in a subtractive form. For instance, in a normal working day, heavy rain may remarkably reduce the requests for bike rental.

The next event type occurs during days 97 and 98 when some new connections are established between stations in a part of the network. As we see, traffic between main stations remains unchanged as the system behaves normally, but moderate traffic shows up from stations E and F to C. This kind of event can appear due to operational changes or occurrence of some regional events. For instance, let us imagine a scenario in which stations E and F are out of service for a long period and suddenly become available. Or in another possible scenario, if users do not find available bikes in station B they may refer to the closer stations E and F. Some local events such as a sports rivalry can also account for this type of events. We know that during a football rivalry, people come from different regions to the event's location. Hence, many rare links with zero traffic might be connected by these users. For instance, a user who lives far away from the stadium may establish a new

* Corresponding author.
E-mail address: hadi.fanaee@fe.up.pt (H. Fanaee-T).

http://dx.doi.org/10.1016/j.neucom.2016.04.006
0925-2312/

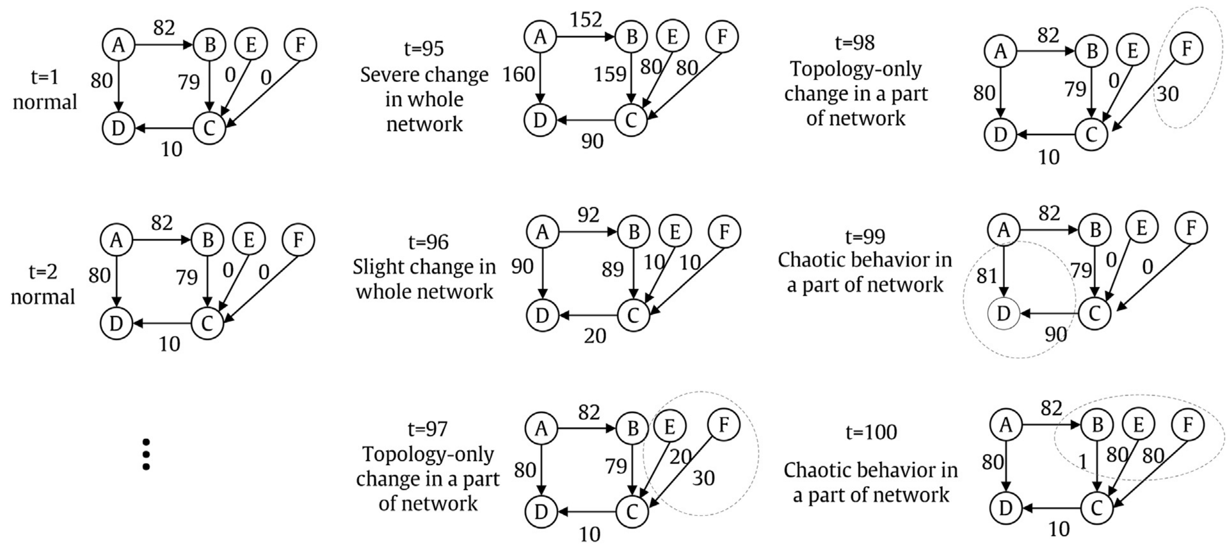

Fig. 1. Motivational example: a simplified hypothetical scenario in bike-sharing network. Between $t=1$ and 94 the system has a stable behavior. At $t=95$–96 a mutation occurs through whole network. In $t=97$ and 98 topology of traffic changes in a part of the network. During $t=99$–100 a chaotic behavior shows up in a part of the network. How can we accurately detect these events via a unified model?.

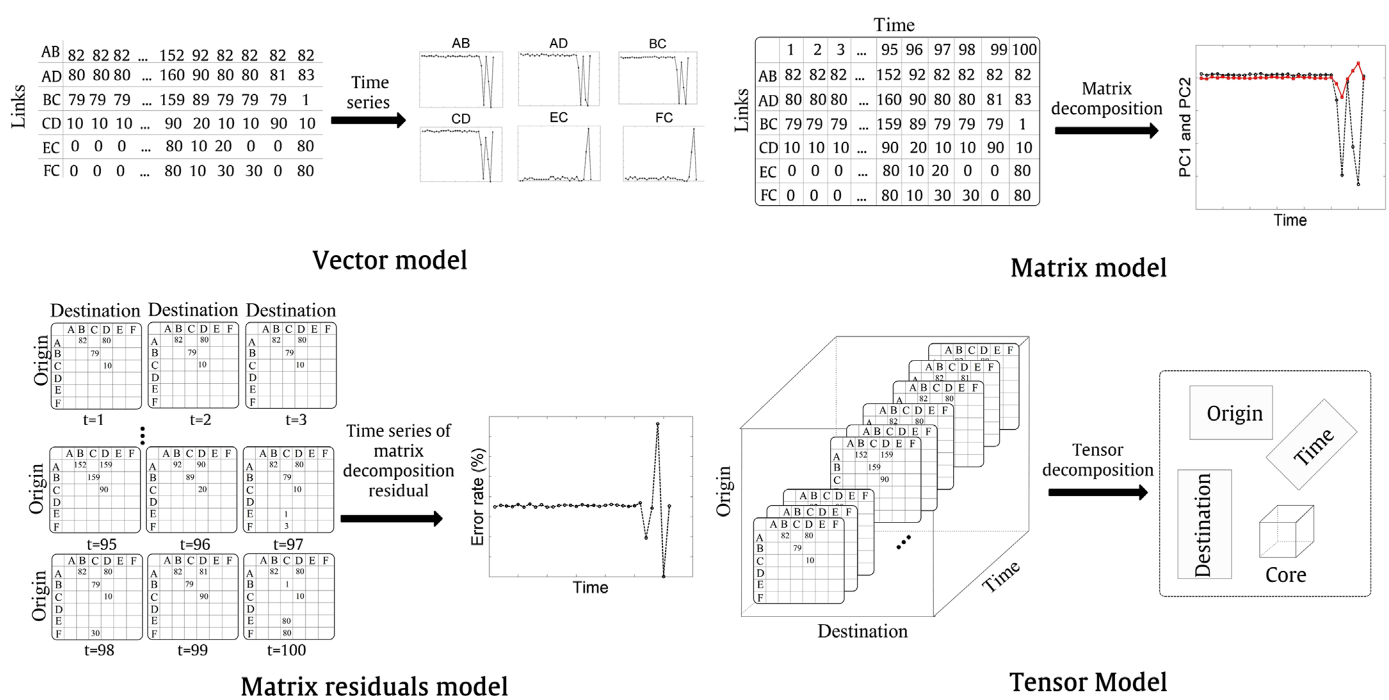

Fig. 2. Four major traffic data analysis techniques. Numbers in the figure are derived from the example scenario in Fig. 1.

connection from a station close to his neighborhood to the stations close to the stadium's location.

Finally, during the last two days, i.e. $t=99$, 100 we observe a chaotic behavior in a part of the network where events do not follow any regular pattern. For instance, even though the traffic pattern at $t=99$ seems similar to the expected for stable conditions, links A to D and C to D exhibit an odd behavior. In one of them we observe a very slight change while in the other one the fluctuation is very intense. Likewise, in $t=100$ we perceive a mutation in the network topology as well as an irregular inconstancy in the flow volume. The reason for this type of events is not evident, because multiple factors are usually involved in their occurrence. An intuitive example of such a chaotic behavior can be an occurrence of various events at the same time such as a football rivalry, alongside a blackout event in some stations added to a weather-based event like rain.

As we can notice in the above cases, some events are associated with fluctuations in flow (e.g. days 95 and 96), while some are linked with alterations to the network topology (e.g. days 97 and 98) and others, such as 99 and 100 to a chaotic behavior in part of the network. This kind of patterns may be repeated several times and are not necessarily limited to a specific type. In a practical manner, most of these patterns take place in the system during its life cycle. The question is how to construct a model that simultaneously detects all these types of events. The answer to this question is the matter of this research and will be discussed in the following, but before that let us briefly review the existing solutions for this problem.