Visualizing Time in Temporal Event Sequences

Jessica Magallanes-Castaneda[1], Lindsey van Gemeren[2], Steven Wood[2] and Maria-Cruz Villa-Uriol[1]*
[1]INSIGNEO Institute for in silico Medicine, Department of Computer Science, University of Sheffield, United Kingdom
[2]Sheffield Teaching Hospitals NHS Trust, Sheffield, United Kingdom

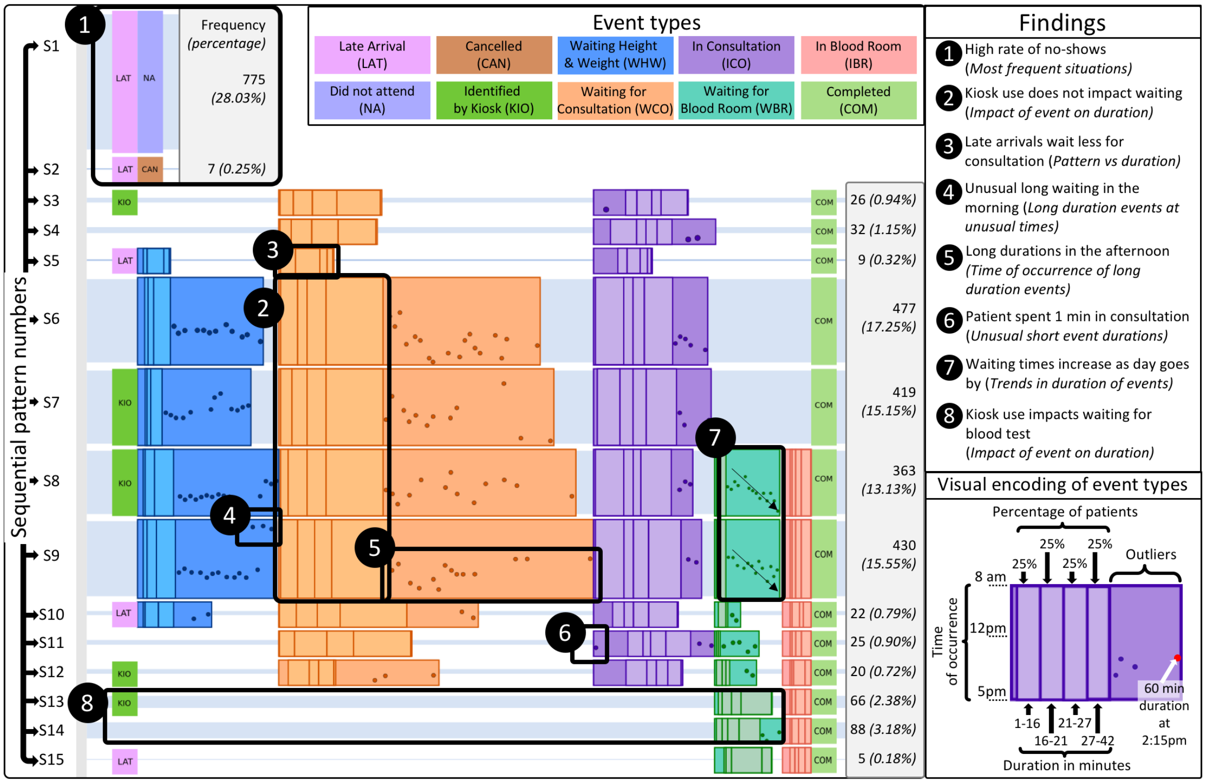

Figure 1: Sequential and time patterns for a dataset provided by Sheffield Teaching Hospitals (STH) [1] (*main view*). The dataset consists of simulated patient flow data, containing 11,899 records, 2,764 sequences and 15 sequential patterns. In the *main view*, event types are color-coded (*top*) and visually encoded (*bottom right*). Width of a color-coded event indicates duration and height represents frequency. Duration outliers are ordered by time of occurrence (vertical axis). Examples of findings derived from the proposed visualization (*top right*).

Abstract

Current methods to visually explore event-data focus on discovering sequential patterns but present limitations when studying time attributes in event sequences. Time attributes are especially important when studying waiting times in patient flow analysis. We propose a visual methodology that allows the instant identification of trends and outliers with respect to duration and time of occurrence in event sequences. Moreover, we show how using Multiple Sequence Alignment (MSA) helps deriving conclusions that otherwise could not be easily reached using single event alignment. The proposed visualization has been applied to a dataset provided by Sheffield Teaching Hospitals (STH), for which four classes of conclusions were derived.

* m.villa-uriol@sheffield.ac.uk

Keywords: Temporal event sequences, outliers, time patterns, patient flow analysis.

Index Terms: Human-centered computing; Visualization techniques—Human-centered computing; Visual analytics—Applied computing; Health care information systems

1 Introduction

There is a pressing need to go beyond basic statistics to analyze the performance of clinical environments. Analysis of patient flow could reveal anomalies in healthcare delivery and provide an opportunity for service improvement. Patient flow data, collected by a workflow management application as event logs, typically consist of a set of temporal events representing the changes in status of patients over time.
The visual exploration of temporal event data is a popular topic; methods to visualize event data often aggregate event sequences in a tree view and encode them according to frequency and duration [2]–[4]. In the context of patient flow data, existing approaches allow the analysis of the most visited pathways (frequent event sequences), but they present limitations in the study of waiting times and lengths of stay, which are key performance indicators in healthcare. The visual encoding of

IEEE VIS 2018 Posters
21-26 October 2018, Berlin, Germany
Author(s) retain copyright.

current techniques complicates the identification of duration outliers and/or time patterns. This limitation is mainly because: (i) duration is encoded using the average value; and (ii) the time of occurrence is not encoded.

Exclusively representing average duration means that outliers are missed, in particular those with an unusual duration. The exact time each event occurs might not be of interest alone, but when combined with the rest of the attributes, time patterns relevant to the understanding of temporal event sequences can be established.

2 VISUALIZING TEMPORAL PROPERTIES OF EVENTS

In temporal event data, an event can be expressed as what happened (event type), when (time of occurrence), for how long (duration), and who was involved (identifier). For each identifier, an ordered list of events can be obtained (event sequence); sequential patterns are obtained by grouping event sequences sharing the same order of events.

In the context of patient flow analysis, event types could represent the rooms in a clinic (e.g. in consultation) or the patient's status (e.g. waiting for consultation, late arrival). Event sequences represent the visit of a patient to the clinic (e.g. late arrival-> waiting for consultation->in consultation->completed). For multiple patients, sequential patterns and their frequency indicate the number of patients visiting each pathway.

2.1 Visualizing duration

Typically, event sequences are grouped according to their sequential order, then the average duration per group is obtained, meaning that events with outlier durations are blended with the rest having a "normal" duration.

Previous literature [5] indicates the importance of identifying infrequent outlier sequences, but no emphasis has been made regarding outliers in terms of their duration. Techniques, such as Eventflow [2], allow temporal querying or even offer a histogram of the different variables in the data [6], however, this information is not encoded in the aggregated view.

In this work, the duration of an event is represented using a box-plot-like graph, where expert users in patient flow analysis can instantly identify unusual waiting times and lengths of stay.

2.2 Visualizing time of occurrence

Apart from detecting the outlier durations, it is critical to understand when they usually occur. In this work, duration outliers are vertically ordered by time of occurrence within the block representing each event type. If an outlier is close to the upper side of the event block, the record happened early in the morning – and as it is closer to the bottom it indicates the event happened towards the end of the business day.

2.3 Multiple sequence alignment

Current visualization tools align event sequences by a single event, either by the first one or the one specified by the user. Eventpad [6] uses Multiple Sequence Alignment (MSA) to show overlap and differences in event sequences. In this work, we also use MSA as this facilitates comparing time attributes across event types and produces a less spread visualization for our test dataset.

3 EXAMPLE APPLICATION

Fig. 1 shows the proposed visualization tested with simulated patient flow data provided by STH [1]. Eight examples of derived findings are numbered and highlighted accordingly in the figure (top right). They can be grouped into the 4 classes discussed next.

Class I - Frequent sequences (finding 1): The most frequent sequence is S1 "Late Arrival" -> "Did not attend" (775). This means that the rate of 'no-shows' in this clinic is very high. Moreover, only few people call in to notify they will not attend, as shown in the S2 pattern "Late Arrival" -> "Cancel" (7).

Class II - Impact of an event in subsequent events duration (findings 2,3,8): Late arrivals (S5) spend significantly less time in the clinic when compared to other pathways (S6 and S7). Sequence pattern S5 has the smallest duration range of "Waiting height & weight", "Waiting Consultation" and "In consultation", than any of the other sequential patterns including a consultation event.

Class III - Trends in time of occurrence vs duration (findings 4,5,7): Note that for sequential patterns S6-S9, and in particular for S9, the longest waiting times happen during the afternoon shift.

Class IV - Unusually short or long event durations (finding 6): For sequential pattern S11, the "In consultation" event (in purple) had an unusually short duration (~1 minute) shown as an outlier.

4 CONCLUSION

The visual encoding of duration and time of occurrence facilitates deriving conclusions II, III, and IV – these conclusions could not be derived, as easily, using existing approaches. The four classes of conclusions presented open the possibility of identifying patterns of interest to support hospital service improvement in their delivery of healthcare to patients.

REFERENCES

[1] L. van Gemeren, "Modelling Outpatient Visits as Agent Driven Workflows," The University of Liverpool.

[2] M. Monroe, R. Lan, H. Lee, C. Plaisant, and B. Shneiderman, "Temporal event sequence simplification," IEEE Trans. Vis. Comput. Graph., vol. 19, no. 12, pp. 2227–2236, 2013.

[3] Z. Liu, B. Kerr, M. Dontcheva, J. Grover, M. Hoffman, and A. Wilson, "CoreFlow: Extracting and Visualizing Branching Patterns from Event Sequences," Comput. Graph. Forum, vol. 36, no. 3, pp. 527–538, 2017.

[4] K. Wongsuphasawat and D. Gotz, "Exploring flow, factors, and outcomes of temporal event sequences with the outflow visualization," IEEE Trans. Vis. Comput. Graph., vol. 18, no. 12, pp. 2659–2668, 2012.

[5] K. Vrotsou, J. Johansson, and M. Cooper, "ActiviTree: Interactive visual exploration of sequences in event-based data using graph similarity," IEEE Trans. Vis. Comput. Graph., vol. 15, no. 6, pp. 945–952, 2009.

[6] B. C. M. Cappers and J. J. van Wijk, "Exploring Multivariate Event Sequences Using Rules, Aggregations, and Selections," IEEE Trans. Vis. Comput. Graph., vol. 24, no. 1, pp. 532–541, Jan. 2018.