Temporal Hierarchical Visualization of COVID-19 Epidemic Data

Sicong Wu* Shenzhen University
Min Lu† Shenzhen University
Hui Huang‡ Shenzhen University

ABSTRACT

In this work, we present two novel visualizations for COVID-19 visual analysis, to deal with the hierarchical relationship in temporal data. By designing a novel tree visualization, we are able to give an overview of the temporal traits as well as spreading relationship over the epidemic transmission. To compensate for a detailed analysis, This work designs a visualization using the metaphor of musical notation, for an easy temporal analysis within a hierarchy. We collected the daily-reported COVID-19 cases in Shenzhen and show the usage of these two visualizations.

Keywords: Tree visualization, COVID-19, temporal data, hierarchical relationship

1 INTRODUCTION

With the spread of the novel coronavirus, named COVID-19, contaminated more than 3.4 million people and killed more than 240 thousand people as of 2 May 2020. Visualization and Visual Analysis can be used to analyze the spread of coronavirus and the pattern of infection. Existing works have created a variety of visualization systems for exploring epidemic data. For instance, Jonh developed a visualization system based on ArcGis dashboard to track worldwide epidemic situations in time [2]. Nguyen et al. [4] presented a new system for rapid modeling of mesoscale biological models and demonstrated the utility on the modeling of the SARS-CoV-2 virus.

But it poses significant challenges for a comprehensive exploration over the epidemic data, such as the spread trend and the pattern of infection for its temporal hierarchical traits. Tree graph is designed for dealing with hierarchical data. Li et al. [3] devised a novel tree grammar to create different types of tree graphs to handle hierarchical tree data. Throughout this work, we combine tree and staff visualization to give an overall trend and show family cluster cases that reflect different infectious patterns. We report our methods applying on Shenzhen COVID-19 patient data set in detail.

2 DATA DESCRIPTION

Our data set was collected from the daily epidemic news pressed by Shenzhen Health Commission. We extracted the temporal hierarchical data from every available patient text information and defined a unique ID for each one. The temporal item consists of the time of arriving in Shenzhen, having symptoms, being diagnosed, and hospitalized. The hierarchical item main comprises the domestic relationship with other patients.

3 VISUAL DESIGN

Generally, in time-series data visualization, using line graph with the time axis is an effective way to exhibit the temporal traits [5]. However, if the data contains a hierarchical structure, orthodox oriented-time visualization can be limited [1].

*e-mail: wuusicong@gmail.com
†e-mail: lumin.vis@gmail.com
‡Corresponding author: e-mail: hhzhiyan@gmail.com

Tree Visualization. Considering the tree graph as a common visualization to deal with hierarchical data and Motivated by the straight shape of the pine tree, we designed timeline tree visualization for exploring, tracking, and comparing time-series epidemic data. Timeline visualization is used for reference to replace tree trunk. Every time node consists of a single or multilevel tree to give an overview and reflect the hierarchical structure which is a relationship of patients.

Staff Visualization. The staff visualization was designed to compensate the deficiency of showing a detail of hierarchical relationship in timeline tree visualization. The musical notation consists of a set of five horizontal lines and four spaces that each represent a different musical pitch. Inspired by its approach to stratify pitch, the space between two lines is defined as the time interval in the staff graph. Fig. 1 shows several examples of infected family structures. Infected people (solid point) at lines show a different interval of infection after contacting the first patient (hollow point).

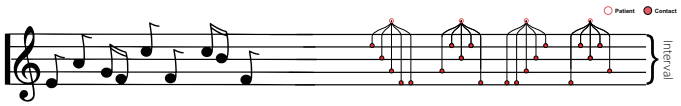

Figure 1: Hierarchy structure visualization using the metaphor of musical notations.

4 CASE STUDY

We apply two visualizations in the task of exploring COVID-19 epidemic data. First of all, we conducted two visualization tasks. (1) Give an overview of the COVID-19 situation and the trait of daily confirmed cases during one period. (2) Demonstrate the mode of transmission relationship among cases. In the first task, daily confirmed cases were encoded as nodes and hence demonstrate the epidemic situation and the temporal transmission characteristics in real-time. For the second task, we conducted a multilevel tree graph to give an overview of transmission events and staff visualization to demonstrate the detail of different transmission modes.

COVID-19 Infection Trend. As shown in Fig. 2.1, we can notice that the trend of the epidemic in Shenzhen showing a Gaussian distribution during the period from 1.19 to 2.14 which indicates the situation of the epidemic has gradually leveled off. Around the beginning of February, it was the outbreak and along with massive clusters which indicates a highly contagious trait of COVID-19. From early February to mid-February, the epidemic gradually returned to its flat period. It can be guessed that the virus had been effectively controlled, and there were only sporadic cases of relative transmission.

Cluster Infection Exploration. Fig. 2.1 gives an overview of all clusters and the ID is corresponding to the cluster in Fig. 2.2. A tree consists of the first confirmed case, infected cases, and their relationship. The distance between infected and confirmed cases represents the interval after close contact with the first one. Fig. 2.2 reflects the massive cluster infection event and shows a fast transmission speed and uncertain incubation period after close contact.

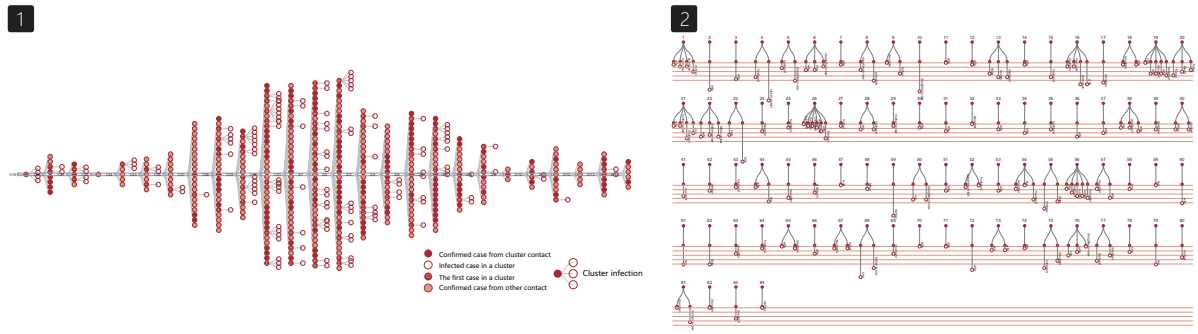

Figure 2: **Timeline Tree and Staff Visualization**. Utilizing subtree (1) represents cluster transmission event. Line and Space define the COVID-19 diagnosis interval from close contact (2). The corresponding cluster shows a detail including the contact interval and relationship with the first case.

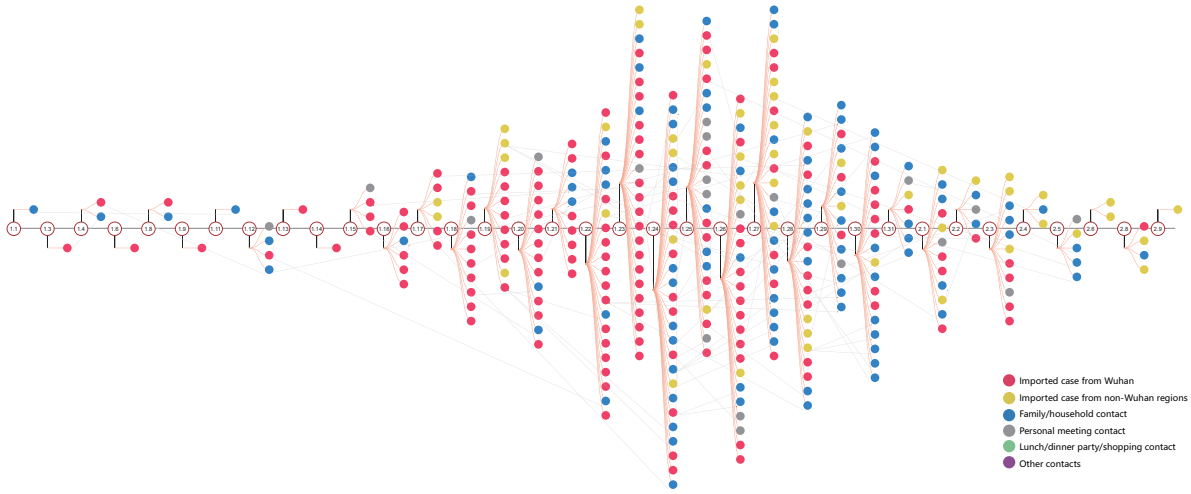

Figure 3: **Comparison of Transmission Route**. The timeline tree supports display multiple transmission routes of COVID-19 in Shenzhen. Using color to encode different transmission routes and lines to point to close contact.

Transmission Route Comparison. We used another period of Shenzhen COVID-19 data set but contains the transmission route and the type of cases. See Fig. 3.2, arrows were employed to display infection routes between cases, and colors were encoded for displaying different types of infection. The result shows that the patient imported from Wuhan was the most in the daily number of cases before 1.27. After that, the number of non-Wuhan cases had dramatically increased which indicates the epidemic has spread locally.

5 Conclusion

This work proposes two novel visualization methods for COVID-19 Epidemic data. The timeline tree and staff graph are based on the metaphor of pine tree and musical notation. The two visualizations were applied to the COVID-19 data in Shenzhen. We combine the two methods to give an overview of the COVID-19 situation and detail about the mode of transmission relationship. The timeline tree visualization provides a new idea of using a tree diagram to process temporal hierarchical data. The approach of combining tree and staff graph can present different aggregation characteristics in detail and may become a new type of visualization method for processing large-scale hierarchical data.

Acknowledgments

The authors wish to thank reviewers and IEEE PacificVis committee. The authors wish to thank Min Lu, Jin Shen, and Hui Huang for contributing to this work.

References

[1] W. Aigner, S. Miksch, W. Müller, H. Schumann, and C. Tominski. Visualizing time-oriented data—a systematic view. *Computers & Graphics*, 31(3):401–409, 2007.

[2] E. Dong, H. Du, and L. Gardner. An interactive web-based dashboard to track covid-19 in real time. *The Lancet infectious diseases*, 20(5):533–534, 2020.

[3] G. Li, M. Tian, Q. Xu, M. J. McGuffin, and X. Yuan. Gotree: A grammar of tree visualizations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, p. 1–13. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3313831.3376297

[4] N. Nguyen, O. Strnad, T. Klein, D. Luo, R. Alharbi, P. Wonka, M. Maritan, P. Mindek, L. Autin, D. S. Goodsell, et al. Modeling in the time of covid-19: Statistical and rule-based mesoscale models. *IEEE transactions on visualization and computer graphics*.

[5] Y. Wang, F. Han, L. Zhu, O. Deussen, and B. Chen. Line graph or scatter plot? automatic selection of methods for visualizing trends in time series. *IEEE transactions on visualization and computer graphics*, 24(2):1141–1154, 2017.