論 文 の 内 容 の 要 旨

論文題目　Variant Topological Data Analysis
（バリアント位相データ解析）

氏　　名　チャン　クオック　ホアン Tran Quoc Hoan

Real-world data exhibit multi-scale and hierarchical structures having patterns, such as in information spreading and epidemic diseases, appearing at various scales. Revealing these patterns to understand the underlying dynamics of data is fundamentally difficult. The major part of this thesis aims to develop a novel set of features to provide insights into the dynamics and reconstruct underlying patterns in the real complex world that the data are supposed to represent. This research lies in the field of algebraic topology, and more specifically, topological data analysis, which is a set of computational topology methods to indicate the shape of data.

The most powerful method in topological data analysis is *persistent homology*, which considers data as a finite set of points in a high dimensional space and constructs a geometrical model that is primarily a collection of geometrical shapes depending on the scale to group nearby data. From this geometrical model, data can be modeled as a sequence of shapes along with the change of the scale. The transformation of the topology structure in this sequence is encoded into quantitative features to reveal the hidden structures which are crucial to understanding the underlying structure of data but are sufficiently complicated to be constructed in other standard approaches. Because topology is a qualitative property that is stable under the influence of noise, persistent homology has been applied successfully to detect the qualitative changes in material science. However, it remains questions in the existing methods to represent variant-scale and dynamics in other types of data, such as time series and complex network data.

Such research questions are addressed in two separate parts of this thesis. In the first part, we propose *variant topological data analysis*, which is a general framework grounded on topological data analysis to encode complex patterns in data as quantitative features over variable scales. It presents a novel set of features, i.e., variant topological features that directly depict topological information as well as provide insights into the variant-scale and dynamics in data. We show how to integrate these features into the kernel method to apply in statistical-learning tasks such as classification and change point detection.

The second part of this thesis gathers our contributions in applications of using the variant topological features with time-series data and complex network data. In the statistical-tasks such as classification and transition points detection of time-series and complex networks, the numerical calculations for both synthetic and real-world data demonstrate that the proposed framework is highly effective and presents a unified approach for characterizing the dynamics. Furthermore, we show the robustness of these variant topological features under perturbation applied to the time-series and complex network data.

The *variant topological data analysis* framework presented in this thesis is evaluated and analyzed on multiple data sets of different dynamical systems to show the robustness of the framework. As such, this thesis suggests that *variant topological data analysis* is a powerful and generic framework for efficient discovery and analysis that can be applied to many types of data in dynamic systems.