Spatial data mining is an interesting area and has received a lot of attention (Guha, Rastogi & Shim, 1998; Koperski, Han & Stefanovic, 1998; Ng & Han, 1994; Sander, Ester & Kriegel, 1998). One special class of spatial patterns is the collocation patterns. Collocation patterns describe a set of features that tend to occur together in close spatial proximity. For example, shopping malls and fast food restaurants tend to be located in the same neighborhood. Recently, some researchers have shifted their attention towards mining of topological patterns. Topological patterns are the set of collocated features that satisfy additional pre-defined spatial relationships. Figure 7.1 shows some examples of topological patterns. Mining topological patterns is an interesting research problem with broad applications, such as mining topological patterns in an e-commerce company, a location-based service, an ecology dataset, and so forth. Thus far, existing works primarily focus on the spatial aspect of the pattern while ignoring the temporal aspect. They discover patterns such as: “There is high probability of the occurrence of earthquakes in a region if there is high atmospheric pressure in the nearby region.” However, it is not clear
whether this “high atmospheric pressure in the nearby region” is observed a few days prior to the occurrence of the earthquake, or many months before the occurrence of the earthquake. With the prevalence of spatio-temporal databases, mining of topological patterns with temporal information, such as: “There is a higher incidence of earthquakes in a region, where during the same time, there is a high atmospheric pressure occurs in the nearby region.” This pattern is more useful and helpful to data analysts and decision makers in understanding the underlying process that controls the changes.

Existing techniques for finding topological patterns (Huang, Xiong, & Shekhar, 2003; Koperski & Han, 1995; Morimoto, 2001; Shekhar & Huang, 2001; Zhang, Mamoulis, & Cheung, 2004) cannot be easily extended to mine patterns in spatio-temporal databases. This is because they follow the candidate-generation-and-test methodology (Agrawal & Srikant, 1994). Such approaches do not scale well when the potential number of candidate patterns is large. In spatio-temporal databases, the candidate space is three dimensional rather than two dimensional. In other words, the number of potential candidate patterns is potentially much more than that in spatial databases. Furthermore, topological patterns must satisfy not only the spatial proximity relationships but also the temporal proximity relationships. This translates to higher computational cost for processing the candidate patterns and computing the interestingness of these patterns. New methods are needed to mine the topological patterns efficiently.

Besides the efficiency of mining topological patterns, we have observed that the spatial features in topological patterns are always prompted by the surrounding geographical objects. To enhance the usefulness of the mined topological patterns, we need to include geographical features in the mining of topological patterns. For example, if we can identify a set of spatial features that always happen together when certain geographical features are present, then decision makers or area developers can have the means to issue a warning ahead of a disaster or consider the available alternatives.

In this chapter, we study the problem of mining topological patterns by imposing temporal constraints into the process of mining collocation patterns. We first introduce a summary structure that summarizes the database with the instances’ count information of a feature in a region within a time window. Next, based on the summary structure, we design an algorithm, called TopologyMiner, to find the interesting topological patterns in a depth-first manner. The algorithm follows the pattern growth methodology. We also investigate an efficient way to incorporate geographical features in TopologyMiner.
Related Content

Parallel Data Mining
[www.igi-global.com/chapter/parallel-data-mining/7593?camid=4v1a](www.igi-global.com/chapter/parallel-data-mining/7593?camid=4v1a)

Identify Cross-Selling Opportunities via Hybrid Classifier
[www.igi-global.com/article/identify-cross-selling-opportunities-via/1807?camid=4v1a](www.igi-global.com/article/identify-cross-selling-opportunities-via/1807?camid=4v1a)

Association Rules: An Overview
Paul D. McNicholas and Yanchang Zhao (2009). *Post-Mining of Association Rules: Techniques for Effective Knowledge Extraction* (pp. 1-10).
[www.igi-global.com/chapter/association-rules-overview/8434?camid=4v1a](www.igi-global.com/chapter/association-rules-overview/8434?camid=4v1a)
A Data Mining-Based OLAP Aggregation of Complex Data: Application on XML Documents
*International Journal of Data Warehousing and Mining* (pp. 1-26).
www.igi-global.com/article/data-mining-based-olap-aggregation/1772?camid=4v1a