Classification Scheme 1: Thematic Classification:

We looked at six papers concerning spatial aspects of association analysis. Of those six papers, three can be placed into one thematic class and three can be placed into a second thematic class. The first thematic class can be thought of as applying mining techniques for the discovery of spatial co-location patterns to new objects and at different scales. For example, Paper 1 proposes a novel buffer-based model for mining co-location patterns concerning extended objects (i.e. line strings and polygons). Paper 2, for example, proposes a novel framework for mining for co-location patterns at different scales. Papers 1, 2, 3 and 6 fall into this first thematic class. The second thematic class concerns the issue of speeding up the co-location pattern detecting and mining process. While all six papers deal with this process on some level, Papers 4 and 5 address this almost exclusively.

Classification Scheme 2: Methodological Classification:

While all of the papers that we looked at use case-studies (using synthetic or real-world datasets, or both), some of the authors test their novel algorithms against alternative algorithms that obtain similar results, while others do not. These are the two classes of the Methodological Classification Scheme: 1. Authors that test their algorithms against alternative approaches in their papers and 2. Authors that do not test their algorithms against alternative approaches in their papers. Papers 1 and 6 fall into the first category while Papers 2, 3, 4 and 5 fall into the second category.

**Paper 1:** A Framework for Discovering Co-location Patterns in Data Sets with Extended Spatial Objects. (Xiong, H.)

In the context of the Thematic Classification Scheme, this paper falls into the first category. One of the major contributions of this paper is that the authors propose a novel buffer-based model for mining co-location patterns. The idea behind the implementation of this model is that objects in space often have some sort of impact on the objects and areas around them. For example, freeways create noise pollution that can be heard blocks away, and factories emit fumes that can affect people for miles around.
Another major contribution of this paper is that it extends the discovery of co-location patterns to line strings and polygons. This is another reason why this paper falls into the first thematic category of applying mining techniques to the discovery of spatial co-location patterns of new objects and/or at different scales, as previous research has dealt with point data.

In the context of the Methodological Classification Scheme, this paper falls into the first category. The algorithm developed by the authors is tested against an alternate, previously applied algorithm using the same spatial data, as outlined below:

**Paper 2**: A Framework for Regional Association Rule Mining in Spatial Datasets. (Wang, J., Eick, C. F., Ding, W., Yuan, X.)

In terms of the Thematic Classification Scheme, this paper falls into the first category. The goal of the research presented in this paper is to come up with a novel framework for mining regional association rules relying on a given class structure, as previous data mining techniques are ill-suited for discovering regional knowledge. This paper fits appropriately into the first thematic category because it deals with applying mining techniques for discovering spatial co-location patterns at different scales, as shown below:
This paper falls into the second methodological category as the authors do not test their algorithm against another algorithm when applying their algorithm to the spatial dataset.

**Paper 3:** Discovering Co-location Patterns from Spatial Datasets: A General Approach. (Huang, Y., Shekhar, S., Xiong, H.)

This paper falls into the first thematic class category. This is the case because, one of the major contributions of this paper is that it introduces a novel algorithm to discover spatial co-location patterns that uses a new multi-resolution pruning technique. As with Paper 2, this paper deals with applying mining techniques for discovering spatial co-location patterns at different scales.

In terms of the Methodological Classification Scheme, this paper falls into the second category. This is because the authors do not test their new algorithm against other algorithms. Rather, the authors test their algorithm against itself, applying it to both a synthetic dataset and to a real-world dataset.

**Paper 4:** Fast Mining of Spatial Collocations. (Zhang, X., Mamoulis, N., Cheung, D. W., Shou, Y.)

In the context of the Thematic Classification Scheme, this paper falls into the second category. This is because this paper deals almost exclusively with speeding up the spatial co-location mining process. One of the major contributions of this paper is that it introduces a novel method that combines the discovery of spatial neighborhoods
with the mining process. The goal of this contribution is solely to speed up the process of discovering co-location patterns.

In terms of the Methodological Classification Scheme, this paper falls into the first category. The algorithm developed for this paper is tested against previous algorithms to depict the amount of time that can be saved by applying the authors’ novel algorithm to the same spatial dataset. This is shown in graphical form, as shown below:

![Graphs showing the effect of different parameters on the time taken to compute](image)

**Figure 6:** Effect of $N$ (star)  
**Figure 7:** Effect of $L$ (clique)  
**Figure 8:** Effect of $m_{\text{noise}}$ (star)

**Figure 9:** Effect of $N$ (clique)  
**Figure 10:** Effect of $L$ (clique)  
**Figure 11:** Effect of $m$ (clique)

**Paper 5:** Mining Frequent Neighboring Class Sets in Spatial Databases. (Morimoto, Y.)

This paper falls into the second thematic classification category. The main contribution of this paper is that it introduces a novel algorithm that efficiently finds sets of service names that are frequently close to each other in a spatial database. In other words, the author develops an algorithm that speeds up the process of mining and discovering spatial co-location patterns.

This paper falls into the second methodological classification category. The algorithm was applied to a telephone directory database. Each facility provided by the database was divided into a class. The author identified the x and y coordinate values for each record, creating a database of point records. The algorithm was then tested using this data to determine the time it took to compute. The author does not test the algorithm introduced against other algorithms on the same dataset.

**Paper 6:** Zonal Co-location Pattern Discovery with Dynamic Parameters. (Celik, M., Kang, J. M., Shekhar, S.)

In the context of the Thematic Classification Scheme, this paper falls into the first thematic category. This is the case because one of the main contributions of this paper is that it introduces an algorithm to discover zonal co-location patterns for dynamic
parameters. This is representative of the first thematic category because it applies mining techniques for the discovery of spatial co-location patterns to new objects.

In terms of the Methodological Classification Scheme, this paper falls into the first category, as the algorithm presented in this paper is compared against another algorithm, as shown below:

Figure 3. (a) Number of Zonal Regions Cost, (b) Varying Zonal Region, (c) Varying Zonal Overlap