1. Maintaining Data Privacy in Association Rule Mining

链接：https://www.vldb.org/conf/2002/S19P03.pdf

作者：Shariq J. Rizvi and Jayant R. Haritsa

期刊或会议名称：VLDB '02: Proceedings of the 28th international conference on Very Large Data Bases

引用次数：733

主要内容：在关联规则挖掘中，保护用户隐私是一个重要的问题。本文提出了一种新的保护用户隐私的关联规则挖掘算法，称为MDL(Maintaining Data Privacy)算法。MDL算法基于数据扰动技术，通过给原始数据添加噪声的方式来保护敏感数据。具体地，MDL算法中，在保证数据准确性的前提下，通过改变数据分布来提高数据隐私性。本文还在公开数据集上进行了实验，实验结果表明，MDL算法能够有效地提高数据隐私性，同时保持较高的准确性。因此，MDL算法具有在关联规则挖掘领域应用的潜力。

关键词：关联规则挖掘、数据隐私、差分隐私、隐私保护、数据保护

1. Privacy-Preserving Distributed Mining of Association Rules on Horizontally Partitioned Data.

链接：http://www.bell-labs.com/user/minos/DMKD02/Papers/kantarcioglu.pdf

作者：Murat Kantarcioglu and Chris Clifton

期刊或会议名称：2002 {ACM} {SIGMOD} Workshop on Research Issues in Data Mining and Knowledge Discovery, Madison, Wisconsin, USA, June 2, 2002

引用次数：456

主要内容：数据挖掘可以从大型数据集中提取重要知识，但有时这些数据集被分散在不同的方。出于隐私考虑，各方可能无法直接共享数据和某些类型的关于数据的信息。我们解决了横向划分数据的安全关联规则挖掘问题。该方法采用加密技术最小化共享的信息，同时为挖掘任务增加很少的开销。

关键词：隐私保护 、分布式计算、关联规则挖掘、水平分区数据、加密技术、安全多方计算

1. Privacy preserving association rule mining in vertically partitioned data.

链接：https://doi.org/10.1145/775047.775142

作者：Jaideep Vaidya and Chris Clifton

期刊或会议名称：Proceedings of the Eighth {ACM} {SIGKDD} International Conference on Knowledge Discovery and Data Mining, July 23-26, 2002, Edmonton, Alberta, Canada

引用次数：443

主要内容：隐私方面的考虑通常会限制数据挖掘项目。本文解决了事务跨源分布的关联规则挖掘问题。每个站点都持有每笔交易的一些属性，并且这些站点希望协作来识别全局有效的关联规则。但是，网站不得透露个人交易数据。我们提出了一种两方算法，用于有效地发现具有最低支持水平的频繁项集，而没有任何一个站点透露单独的交易价值。

关键词：隐私保护、关联规则挖掘、垂直分区数据、加密技术、Apriori算法、数据拆分

1. Discovering calendar-based temporal association rules.

链接：https://doi.org/10.1109/TIME.2001.930706

作者：Yingjiu Li and Peng Ning and Xiaoyang Sean Wang and Sushil Jajodia

期刊或会议名称：Eigth International Symposium on Temporal Representation and Reasoning, TIME-01, Civdale del Friuli, Italy, June 14-16, 2001

引用次数：387

主要内容：一个时间关联规则是在特定时间间隔内成立的关联规则。例如，早上鸡蛋和咖啡的销售经常同时发生。本论文研究了用户给定的日历模式下的时间关联规则。通常使用日历模式可以使发现的时间关联规则更易理解。一个日历模式的例子是 (年，月，日)，它可以产生形如 (d/sub 3/, d/sub 2/, d/sub 1/) 的一组基于日历的模式，其中每个d/sub i/ 可以是一个整数或符号 \*。例如，(2000, \*, 16) 就是这样一个模式，对应于时间间隔，每个间隔都由 2000 年某个月的第 16 天组成。本文定义了两种类型的时间关联规则：精确匹配关联规则要求关联规则在每个时间段内都成立，模糊匹配要求关联规则在大多数时间段内成立。本文扩展了著名的Apriori算法，并开发了两种优化技术来利用基于日历的模式的特殊属性。实验表明，算法和优化技术是有效的。

关键词：关联规则、日历、交易数据库、信息系统、数据挖掘、数据分析、合同

将这10篇论文按照研究主题进行分类，可以分为以下5类：

数据隐私保护和关联规则挖掘

Maintaining Data Privacy in Association Rule Mining

Privacy-Preserving Distributed Mining of Association Rules on Horizontally Partitioned Data

Privacy preserving association rule mining in vertically partitioned data

这三篇论文的研究都围绕着如何在保护数据隐私的前提下进行关联规则挖掘。它们介绍了不同的方法来实现数据安全挖掘，包括对数据进行加密、分区和转换等。这些方法能够保护敏感信息，并增强数据安全性。

时间关联规则挖掘

Discovering calendar-based temporal association rules

该论文介绍了一种基于日历的时间关联规则挖掘方法，它使用日历模式和基于日历的模式，对给定数据中存在的时间间隔进行过滤和筛选，以提高挖掘效率并减少结果中的噪声。该方法可以应用于推荐系统或其他相关领域中时间序列分析和事件流分析的场景。

基于关联规则的分类

CPAR: Classification based on Predictive Association Rules

该论文介绍了一种基于预测性关联规则的分类方法，即使用概率映射函数来将交易映射到预测的类别。该方法可以应用于许多具有高维离散属性的数据集，并且可以大大减少分类错误率。

关联规则挖掘和算法

Association Rule Mining, Models and Algorithms

Mining gene expression databases for association rules

Efficient Adaptive-Support Association Rule Mining for Recommender Systems

Mining Both Positive and Negative Association Rules

这四篇论文介绍了关联规则挖掘和算法及其在不同领域的应用，包括基因表达数据挖掘、推荐系统和消费市场等。它们分别讨论了如何选择合适的数据集、特征提取和建模等方面的问题，并提出了相应的解决方案。其中的算法和模型也能够帮助研究者更好地理解并应用关联规则挖掘。

模糊关联规则挖掘

Fuzzy association rules: general model and applications

该论文讨论了关联规则挖掘中的模糊性质，并提出了一种通用模糊关联规则模型。该模型可以应用于各种数据集和领域，并且可以找到不同的方式来定义支持和置信度等关联规则参数。