

第一章 绪论

1.1 选题背景

1.1.1 虚拟仿真技术

1.1.2 缺陷检测技术

1.2 研究意义和内容

1.2.1 研究意义

1.2.2 研究内容

1.3 论文的内容安排

1.4 本章小结

第二章 软件缺陷检测技术

1) 软件缺陷分类

2) 缺陷检测技术

第三章 虚拟仿真：解释翻译和功能模拟

第四章 嵌入式二进制软件缺陷检测的仿真支持

第五章 实验及数据

第六章 总结及展望

- 1) 论文总结
- 2) 后续工作展望

1.5 需求是科技发展的源动力

从20世纪80年代中期以来,各个中、大型企业都逐步建立了功能强大、面向事务型的信息系统(OLTP—Online Transaction Processing)。采用OLTP能充分满足关键任务的数据请求,高效、精确地处理企业的日常事务,完成收集和管理数据的功能。由于关系型数据库具有参照完整性、良好的故障恢复能力、能并行处理大量小事务等优点,因此几乎所有OLTP系统都是基于关系型数据库的。数据库开发商也开发了大量关系型数据库系统,并且提供最新型的产品使得这些数据库的使用更容易、更有效。

经过多年的积累,目前每个应用事务型信息系统的企业,都保存了大量关于产品、用户和市场的数据。随着竞争的加剧,越来越多的企业意识到有必要对累积在OLTP中的数据加以提炼、利用,帮助他们更好地进行下一步计划和决策。

1993年,E. F. Codd,关系型数据库的发明者,在《Providing OLAP to User Analysis: An IT Mandate》白皮书中,首次提出了OLAP(Online Analysis Processing)概念。OLAP通过对数据的切片、切块、钻取和旋转等操作,从不同角度提取有关数据,提供多层面、多角度的逻辑视图,以查询、分析为特征,根据决策者对信息的需求,给出即时、准确的信息,使分析人员、管理人员或执行人员能够对从原始数据中转化出来的、能够真正为用户所理解的、并真实反映企业特性的信息进行快速、一致、交互地存取,获得对数据的更深入了解,回答诸如“谁的”或者“干什么的”等等问题,并可以通过类似“如果……,怎么样”的假设,观察到不同的情况。利用OLAP分析数据,可以帮助公司管理某个特定领域的销售目标、项目进展,甚至预测发展趋势,所有这一切都将有助于公司在竞争中取胜。

由于OLAP的功能是用户参与的汇总和比较,是基于管理人员对问题求解的需求,对数据进行有目的处理、分析,因此跟管理人员的水平直接相关。随着数据的大量积累,以及市场竞争对信息与知识的迫切需要,人们希望通过更进一步的技术,回答难以用确定的方式来提问的问题,找到尚未发现的、隐藏在数据中的、人的先验知识和经验无法确定的信息与知识。随着统计学、人工智能和数据库技术的成熟,数据挖掘作为解决上述问题的技术手段,自1989年被正式提出以后,迅速发展起来。

数据挖掘(Data Mining)就是从大量的、不完全的、有噪声的、模糊的、随机的数据中,提取隐含在其中的、人们事先不知道的、但又是潜在的有用信息和知识的过程^[1]。沃尔玛的老板再精明,如果没有数据挖掘的帮助,也绝对

想不到“啤酒与尿布”这两个风马牛不相及的东西还有着千丝万缕的联系，而将它们放在一起，竟然增加了啤酒的销量。数据挖掘正是利用一定的知识发现算法，对原始数据进行模式总结，获取一些难以用确定的方式来提问的问题，做出前摄的、基于知识的决策，然后以一定的置信度对未来趋势及行为进行预测，从而为企业的决策提供强有力的支持。

1.6 数据挖掘应用的发展现状

1.2.1 数据挖掘产品研究现状

经过十多年的发展，大量有效的数据挖掘算法已经被研制出来，这些算法都已经比较成熟，可以帮助人们发现隐藏在海量数据后的“秘密”。目前，数据挖掘研究的重点逐渐从发现方法转向提高算法实际应用效率和完整数据挖掘系统的研发，并注重多种发现策略和技术的集成，以在一个系统内完成数据采集、预处理、数据挖掘和结果呈现这一系列任务，并可以将挖掘得到的结果嵌入到应用系统。Giga Research所做的研究报告指出，虽然目前数据挖掘在商业智能（包括报表系统、OLAP、数据挖掘等等）中只占15%的份额，市场大约在10亿美元左右，但它是商业智能中增长最快的领域，越来越多的传统商业应用，如ERP、CRM等等都逐渐开始包含数据挖掘功能[2]。目前，市面上大约有几百家数据挖掘厂商和咨询公司，在KDNuggets (www.kdnuggets.com)中有关于这些公司及其产品的详细信息，主要的数据挖掘厂商有以下几个。