

基于数据仓库技术的目标威胁估计模型

刘昌云, 宋万德, 贺正洪
(空军工程大学 导弹学院, 陕西 三原 713800)

摘要:介绍了数据仓库的定义、特征以及相关的关键技术,并结合战场威胁估计的实时性、重要性,定义了模型的数据仓库。在此基础上,构造了基于数据仓库的目标威胁估计模型,并进一步建立了目标威胁估计的星型数据组织模型。运用数据仓库技术进行目标威胁估计,能够提高对空中目标的威胁估计快速性、精确性和预测能力。

关键词:数据仓库;威胁估计;模型

中图分类号:TP919 **文献标识码:**A **文章编号:**1009-3516(2001)04-70-73

现代战争,基本上是高技术条件下的局部战争,战争的突然性增加、战场空间扩大、信息量剧增、情况瞬息万变。并且,未来的空袭威胁已呈现如下趋势^[1]:低空和超低空威胁;多批多方向威胁;空袭与电子战相结合;进行饱和和攻击;信息压制。这给现代防空技术提出了更高的要求。当前,利用多传感器技术、多基地雷达、对空域进行全方位、多层面的侦察,实现从不同侧面、不同层次发现目标,但同时也会造成“海量”的数据信息,如何管理这些数据信息并从这些“海量”的数据信息中提取出有效、有用的知识供决策使用,已成为一项迫切需要解决的问题。现今,由于数据仓库技术的发展,为我们能更快、更有效的进行目标威胁估计提供了基础。基于此,本文提出基于数据仓库技术的目标威胁估计模型。

1 数据仓库技术

1.1 数据仓库(DW)的定义

数据仓库技术之父 W. H. Inmon 在其著作《建立数据仓库》中定义:数据仓库是面向主题的、集成的、稳定的、不同时间的数据集合,用于支持管理、决策的制定过程^[2]。数据仓库是一个新型的数据库,通过提取、过滤和集成事务数据,从大量的历史数据、当前的数据中挖掘出有用信息,供决策、分析人员使用。根据文献[3],并结合战场威胁估计的实时性、重要性,定义数据仓库如下:

$$DW = RDBMS + HDB + NDB + PDB + DM + KDD + II + Applic + PI + KDB$$

其中:RDBMS 代表关系型数据库管理系统;HDB 代表历史数据库;NDB 代表当前数据库;PDB 代表预测值数据库;DM 代表数据开采;KDD 代表数据库中的知识发现;II 代表信息界面;Applic 代表应用程序;PI 代表物理基础设施;KDB 代表知识库。

1.2 数据仓库的特点

数据仓库是一个复杂的组合过程,它是建立在数据库系统基础上的高层管理、分析与决策的产物。数据仓库比数据库系统有更好的灵活性、有效性和可靠性,具有以下特征:

1) 面向主题而集成的。数据仓库可以根据最终用户组织和提供数据,它的数据被组织划分成一个个主题域,并将与主题域有关的因素利用多维表体现出来,并且对不同的数据来源要进行数据结构和编码,统一原始数据中的所有矛盾之处。例如:在目标威胁估计中,对空中目标的描绘有空域、高度、速度、航向、飞机类型,这是一个5维问题,可以用一个5维表描绘出来。

2) 完整性。数据仓库管理和组织大量的综合数据、历史数据和当前数据,它的数据一般都存储在多个

介质上。数据仓库需要处理来自不同数据库的信息,需要收集和整理这些应用程序的数据。

3)非易失性。数据仓库是只读的,用户不能回写数据。

1.3 与DW有关的关键技术

1.3.1 数据开采(DM)

DM主要是描述、寻找有价值的模型,专门供决策、分析使用。在DM过程的开始,必须针对样本数据的变化规律,找出逼近该实际数据变化的模型;然后再利用某种方法求出相应的参数,从而找出相应的模式;再按实际值代入该模式,得出预测值,即未来的变化趋势。DM所采用的方法和技术有:归纳学习法、仿生物法、公式发现法、统计分析法等。

1.3.2 知识发现(KDD)

经过DM后,获得有价值的信息,但是并不等于获得了知识,必须将信息再加以规则、推理的合成,才会变成有用的知识,规则采用:IF.....THEN.....的结构。例如:对于目标威胁估计模型中,找到的有用信息有:

- 规则(1):IF 目标速度 $V > v_0$ THEN 战斗机;
- 规则(2):IF 无应答响应信号 THEN 敌机或友机;
- 规则(3):IF 逼近保卫区时间 $< t_0$ THEN 敌机

将规则(1)、(2)、(3)合成为一条有价值的规则:

- 规则(4):IF 目标速度 $V > v_0$,无应答响应信号,逼近保卫区时间 $< t_0$
THEN 目标威胁等级高

2 基于数据仓库的目标威胁估计模型

数据仓库是存储查询和决策分析用的集成化信息仓库,是存储数据的一种数据形式,它从不同的数据源中提取出原始信息,再辅助于数据开采和知识发现功能,从原始信息中提取出有用的知识,供决策分析者使用,该模型如图1所示。

1)多传感器系统。该系统主要用于采集原始数据。为了有效的完成对空域的侦察,并提供给决策者更准确的信息,可以采用多传感器对空域进行侦察:雷达探测器、声纳探测器、红外探测器、磁性探测器等,以完成对空域的多维扫描,获取更多的目标数据。

2)数据预处理系统。该系统通过数据的一致性、相关性、可靠性、紧迫性的检查,过滤无用的数据,减少数据的冗余,进而减少数据仓库的数据量。

3)数据仓库。该系统主要接收从数据预处理系统获得的数据,存储于基本数据库,这是最活跃的数据区。通过对基本数据库的数据的装载、抽取和转换,提取出高度综合的信息存储于综合数据库,供决策使用。历史数据库主要记录过去的数据库,数据仓库根据历史数据可以进行时间趋势分析、预测。元数据库用于组织数据仓库的组织结构,方便用户的查询。从基本数据库、综合数据库和历史数据库数据中提取出原始信息,作为数据开采和联机处理一体化模型系统的输入信息,并存储该系统所发现的知识。

数据仓库主要采用多维数据表存储数据。

4)数据开采和联机处理一体化系统。通过该模型,分析数据仓库中的原始数据,提取出原始数据中隐藏的模式并预测未来的趋势,发现未知的事实。

5)查询系统。该系统主要供分析决策者查询历史知识和当前知识以及威胁估计判断的过程和结果,通过该系统主要实现决策和知识是否一致性的检查。

6)输出系统。该系统输出显示的信息有目标的航迹、目标的类型、目标的飞行速度、目标的威胁等级等;并可打印从当前数据库和历史数据库的信息以及威胁估计判断的过程和结果。

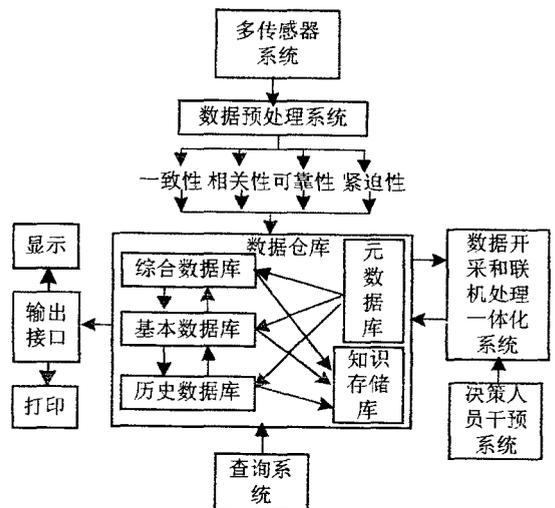


图1 数据仓库的目标威胁估计模型

7) 决策人员干预系统。通过该系统,决策人员可以输入通过其它方式得到的信息,减少数据开采和联机处理一体化系统的运算时间。

3 目标威胁估计的数据组织模型

目标威胁估计数据仓库有两个基本的主题:目标事实主题、我方事实主题。其中目标事实主题是最基本对象,包括:目标类型、飞行速度、飞行时间、飞行高度、航路捷径、所在空域等。我方事实主题是辅助对象,其包括:我方机场信息、友邻部队信息、防空部队信息等事实。数据仓库的主题数据采用星型模型,因为星型模型可以支持多维查询,并具有最佳的性能。星型模型由事实表(大表)以及多个维表(小表)所组成。事实表用于存放大量的事实数据(数量数据);维表用于存放描述性的数据,它是围绕事实表建立的较小的表。本系统的数据模型设计如图2所示。

3.1 数据的归一化

该数据模型是一个多维的事实数据表,存放过程中,要将不同的数据进行一致化、标准化的处理。该模型对数据的一致化处理如下:飞行速度:低速、中速、高速,分别用“0”、“1”、“2”表示;目标模型:民航、侦察机、战斗机、轰炸机、导弹,分别用“0”、“1”、“2”、“3”、“4”表示;飞行高度:低空、中空、高空,分别用“0”、“1”、“2”表示;航路捷径:近、中、远,分别用“0”、“1”、“2”表示;识别信号:无、有,分别用“0”、“1”表示。例如,对这样一张事实表(批次号;空域号;民航;低速飞行;中空;航路捷径:近;识别信号:有),可表示为(批次号;空域号;0;1;0;1)。

3.2 数据的加载

数据加载是数据仓库中重要的操作,从数据源取回的数据要进行合并、筛选、替换、分离、统计处理,并根据数据量的大小、数据有效期阶段进行。数据仓库中进行处理时要以稳定数据为基础,减少重复计算,保证数据质量。加载的方式可以有立即、定时、周期3种,并允许更新某些紧要数据模型。

数据仓库的数据要经过两次映射定义提取规则和数据转换规则,控制数据提取和数据处理的完成。第一次实现远端传递数据的过滤、合并,主要是采用嵌入式模块,尽可能精简的得到所有必须的数据。第二次是完全在本地数据库进行,主要采用存储过程,以期得到尽可能高的效率^[4]。

4 结束语

运用数据仓库技术建立目标威胁估计模型,可以充分利用数据仓库的优点:面向主题、非易失性、基于数据驱动、长时间框架。非战时,可以由我方模拟敌方或其他方式获得空域目标的侦察数据,通过数据仓库的管理工具,获取综合数据和历史数据;并通过数据开采和知识发现系统获取知识,并从知识中得到元知识。在战时,进行目标威胁估计时,可以利用当前数据、历史数据以及元知识,加快空中目标数据的运算速度以及提高目标威胁估计的准确性和预测能力。当然,数据仓库是一项复杂的信息存储库,如何进行数据收集、从信息库中挖掘和发现知识是数据仓库技术应用于实践的关键,当前,数据开采和知识发现的理论和技术的研究正成为信息科学学术界新的研究热点^[5]。

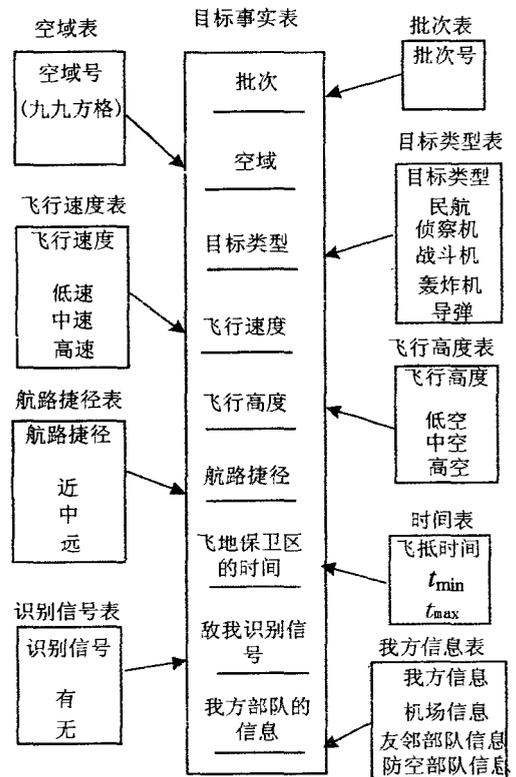


图2 数据仓库的数据组织模型

参考文献:

- [1] 贺正洪,王 睿. 防空指挥自动化信息处理[M]. 西安:空军工程大学,1999.
- [2] 陈文伟. 决策支持系统及其开发(第二版)[M]. 北京:清华大学出版社,2000.
- [3] 陈 燕,杨德礼. 一个数据仓库的设计和实现[J]. 大连理工大学学报,2000,40(2):249-252.
- [4] 张海英,张 滨,王 云. 数据仓库技术在SDH配置管理中的运用[J]. 北京邮电大学学报,2000,23(2):77-80.
- [5] 印 勇,曹长修,林景栋,等. 数据仓库和数据开采研究综述[J]. 重庆大学学报,2000,23(2):116-118.

Model of Object Threat Estimation Based on Data Warehouse Technology

LIU Chang-yun, SONG Wan-de, HE Zheng-hong

(The Missile Institute of the Air Force Engineering University, Sanyuan 713800, China)

Abstract: This article explains the definition, character and the related key skills of data warehouse. Combined with the simultaneity and importance of the war-field threat estimation, the data warehouse of the model is defined. Based on this, a model of object threat estimation is built according to data warehouse. And finally, a star data model of the object threat estimation is produced. Using the data warehouse skills for realizing the object threat estimation can improve speediness, accuracy and prediction ability of the threat estimation of the object in the air.

Key words: data warehouse; threat estimation; model