第 34 卷 第 9 期 计 算 机 工 程 2008 年 5 月
Vol.34 No.9 Computer Engineering May 2008
·人工智能及识别技术· 文章编号：1000—3428(2008)09—0225—02 文献标识码：A 中图分类号：TP309

美国“矩阵”情报分析系统的破解研究

高庆德[1]，程 英[2]
(1. 解放军外国语学院军事情报系，洛阳 471003；2. 解放军 63880 部队研究所，洛阳 471003)

摘 要：将认知领域的情报分析和人工智能、计算机技术相结合，实现情报分析的自动化是各国情报系统研究的重要领域，该文对美国“矩阵”情报分析系统的关键环节进行研究，探讨该系统所采用的算法，并对其进行描述，给出该系统的具体分析过程及系统的体系结构。
关键词：军事情报；定量分析；贝叶斯

Dissolve Research on American Matrix Intelligence Analysis System

GAO Qing-de[1], CHENG Ying[2]
(1. Dept. of Military Intelligence, PLA Foreign Language University, Luoyang 471003; 2. PLA 63880 Unite Research Center, Luoyang 471003)

【Abstract】It is a very important field to intelligence analysis system of every nation that psychology recognition combines with artificial intelligence and computer science. The paper studies the key technic of American matrix intelligence analysis system extensively. Arithmetic is discussed in this paper, and it gives the analysis process and the architecture of this system
【Key words】military intelligence; quantities analysis; Bayes

1 概述

“9·11”事件发生后，经过长期的调查，美国司法和情报单位最终从“9·11”事件中 4 架被毁飞机的乘客名单中排查出了劫机者。然而，最先锁定疑凶的却是美国佛罗里达州的一民间人士——汉克·阿舍[1]。

阿舍在 2001 年 9 月 14 日检索出了恐怖分子重点怀疑对象，而此时飞机乘客名单尚未公诸于众。因此，阿舍是在完全独立的情况下找出怀疑对象的。他用于检索的程序被命名为《多条件反恐信息交换系统》，简称“矩阵”。2001 年 9 月 17 日“矩阵”工作室建成。它拥有 20 个工作站，每个都有一台电脑与由数千个因特尔主板构成的“矩阵”超级计算机相连，用于相关的情报分析。

本文将针对“矩阵”系统的体系结构、分析过程和算法等几个方面进行研究。

2 “矩阵”分析系统的框架结构

“矩阵”情报分析处理系统的结构，采用专家系统(ES)的形式。专家系统的特色在于它的知识库和推理机构，其分析系统的结构如图 1 所示。

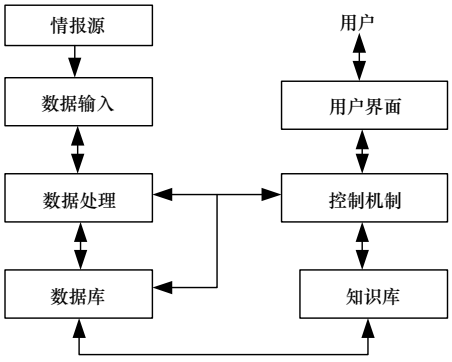

图 1 “矩阵”情报分析系统

在图 1 中，知识库存放着专家对情报分析的知识，确定证据权重、关联程度等；数据库存放着收集到的基本事实和情报结论；控制机构负责控制和推理。

3 “矩阵”情报分析系统的分析过程

在“矩阵”情报分析系统中，采用贝叶斯分析推断的一般模式：先验信息 $\oplus$ 样本信息 $\Rightarrow$ 后验信息，表示为 $\pi(\theta)\oplus p(\chi/\theta)\Rightarrow\pi(\theta/\chi)$。此处的 $\oplus$ 表示为 Bayes 的作用。其分析推理的模式如图 2 所示。

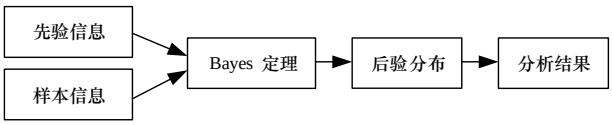

图 2 “矩阵”情报分析的基本模式

先验分布反映了试验前对总体参数分布的认识，在获得样本信息后，人们对这个认识有了改变，其结果就反映在后验分布中，即后验分布综合了参数先验分布和样本信息。由此可以看出矩阵情报分析是一个“从有到有”的过程，且结果清楚自然，符合人们的思维习惯——根据所获得的信息修正以前的看法，不一定从零开始。从本质上说，这种方法概括了一般人的思维、学习过程。

4 “矩阵”分析系统的算法描述

在理论上，统计推断分析是在不掌握完全信息条件下的推断，也就是说，所掌握的信息还不足以决定问题的唯一解，这就提供了贝叶斯(Bayes)方法用于军事情报分析的可能性。基于“矩阵”系统的分析特征，其推断分析符合统计推断分析的特点，笔者认为其基本算法为Bayes方法和更新的Bayes

笔者简介：高庆德(1970－)，男，博士研究生，主研方向：情报分析；程 英，高级工程师、硕士
收稿日期：2007-05-15 **E-mail：**cychh2002@yahoo.com.cn

方法[2]。

Bayes分析方法用于情报分析可以较好解决主观缺陷，特别在解决某一领域情报分析问题时，效果非常明显。Bayes分析完整的表述：假设事件$h_1, h_2, \cdots, h_i, \cdots, h_n$互不相容且构成一个完全事件集合，已知它们的概率$p(h_i)$, i=1, 2, …, n。观察到某种证据d与$h_1, h_2, \cdots, h_i, \cdots, h_n$相伴随而出现，且已知条件概率$p(d/h_i)$，求在新证据$d$出现之后事件$h_i$的概率，即$p(h_i/d)$为

$$p(h_i|d)=\frac{p(h_i)p(d/h_i)}{p(h_1)p(d/h_1)+\cdots+p(h_n)p(d/h_n)}$$

其中，$p(h_i)$, i=1, 2, …, n称为先验概率(基础概率)；$p(h_i/d)$称为后验概率。运用于情报分析，Bayes定理表明对于某一特定时间的情报分析结果不仅依赖于已知的条件概率，也依赖于该事件发生的基础概率。

分析描述如下：设E为与假设相关的事件；NE为事件的否定；A为描述分析问题完备、互斥假设集中的一个元素；$p(A)$为在得到新证据更新前的先验概率；$p(NA)$为A不真实的概率；$p(E)$为独立于假设的事件E的概率；$p(A,E)$为事件E和假设A的联合概率；$p(E|A)$为给定A真实条件下E的条件概率；$p(E|NA)$为给定A非真实条件下E的条件概率；$p(A|E)$为在事件E下假设A的条件概率；$\{E\}_t$为在t时刻以前时间与假设有关的事件集合；E_{t+1}为t+1 时刻新出现的事件。

与假设相关的事件 E 出现后，假设 A 是真实的条件概率为

$$p(A|E)=\frac{p(A,E)}{p(E)} \tag{1}$$

对 $p(A,E)$和 $p(E)$直接计算比较困难，将式(1)变换：

$$p(A|E)=\frac{p(A)p(E/A)}{p(A)p(E|A)+p(NA)p(E/NA)} \tag{2}$$

假设集合有 A, B, C 3 个元素为互斥完备集，则：

$$p(A)+p(B)+p(C)=1,\ p(NA)=p(B)+p(C)$$

则按照 Bayes 原理：

$$p(E_{t+1}|NA,\{E\}_t)=\frac{p(E_{t+1},NA,\{E\}_t)}{p(NA,\{E\}_t)} \tag{3}$$

将 NA 用假设集合中其他元素代替：

$$p(E_{t+1}|NA,\{E\}_t)=\frac{p(E_{t+1},B)+p(E_{t+1},C)}{p(B)+p(C)}=\frac{p(E_{t+1}|B)p(B)+P(E_{t+1}|C)p(C)}{p(B)+P(C)} \tag{4}$$

t 时刻假设 A 的先验概率为 t 时刻前所发生事件的条件概率：

$$p_t(A)=p(A|\{E\}_t) \tag{5}$$

t+1 时刻假设A的概率为t时刻所发生的事件集合$\{E\}_t$和t+1 时刻发生事件E_{t+1}的条件概率：

$$p_{t+1}(A)=p(A|\{E\}_t,E_{t+1}) \tag{6}$$

运用 Bayes 原理，在 t+1 时刻假设 A 发生的概率为

$$p_{t+1}(A)=\frac{p(\{E\}_t)p_t(A|\{E_t\})p(E_{t+1}|A,\{E\}_t)}{p(\{E\}_t,E_{t+1})} \tag{7}$$

为便于计算，将分子、分母变换：

$$p_{t+1}(A)=\frac{p(\{E\}_t)p_t(A)p(E_{t+1}|A,\{E\}_t)}{p(\{E\}_t)p_t(A)p(E_{t+1}|A,\{E\}_t)+p(\{E\}_t)p_t(NA)p(E_{t+1}|NA,\{E\}_t)} \tag{8}$$

即 t+1 时刻后验概率：

$$p_{t+1}(A)=\frac{p_t(A)p(E_{t+1}|A,\{E\}_t)}{p_t(A)p(E_{t+1}|A,\{E\}_t)+p_t(NA)p(E_{t+1}|NA,\{E\}_t)} \tag{9}$$

式(1)和式(2)是在初始时刻(t_0)的第一步，在新证据出现后，这种步骤开始重复，t时刻的后验概率变成t+1 时刻的先验概率，循环往复至融入所有的证据信息，最后取得情报分析结论。

在“9·1l”事件中，用“矩阵” 系统进行分析时，劫机者似应满足以下条件：此人在事件发生前近一两年内才来到美国，有宗教背景，并因此在这个时间段里产生记录——电话账单、水电费单和驾照。相反，如果一位原籍中东的人士在美国定居 10 年，并已获得选举权，则不在怀疑对象之列。将这些证据逐步运用 Bayes 分析方法融入分析系统，逐步得出分析结论。

在阿舍的检索程序中，数据库中凡满足上述某些条件的人都会得到一个“0”以上的分值。最后，在数亿个人名中，有分值的只有大约 12 万人。其中 419 人得分很高，表明嫌疑最大。“9·11”事件发生后的 2 天里，阿舍在位于佛罗里达州博卡拉顿的家中，利用电脑对相关数据进行综合分析，在“9·11”事件发生后的第 2 天晚上 8 点终于分析出可能的劫机者。此后，阿舍将自己的分析方法编写成程序。直到第二天中午，阿舍终于可以用其整晚编写出来的程序在多年积累的海量电子数据库集中运行，在其中收录的约 4.5 亿人的资料中进行检索。通过使用“矩阵”，阿舍检索出了此人所有用过的地址、银行记录、机动车记录、驾照记录、飞机驾驶执照、借贷历史、邻居和房东的名字以及上述所有人的数码照片。

5 结束语

近年发生的几场局部高科技战争、“9·11”恐怖袭击事件以及伊拉克大规模杀伤性武器核查，凸现军事情报分析的地位和作用[3]。美国的情报组织在其机构进行改革的同时，也在寻求更有效的军事情报分析方法，因此产生了“矩阵”系统。本文对“矩阵”系统的算法进行了探索，但由于此系统的保密性很高，“矩阵”工作室高度戒备，包括阿舍在内，西津特公司的任何员工未经许可都不得进入这个工作室。因此本文只是一个初步的探索，其算法、推理过程、体系结构都是笔者基于其特征的推理，因此其正确性、正确程度都不得而知。但可以肯定的是伴随着现代科学技术的快速发展，各种媒体、出版物等更为开放，获取情报素材的手段和能力有了很大提高，这将是拥有高价值的情报素材和低质量的情报分析水平矛盾突出。为解决这个矛盾，自然科学将大量地融合到军事情报分析领域，军事情报量化和自动化辅助分析系统的研究将更加深入。自然科学在认知领域的应用，在及时处理浩如烟海的情报原始素材的同时，最大限度地避免主观错误成为可能。

参考文献

[1] 祁长松. 美国情报首脑全传[M]. 北京, 中国: 中国社会科学出版社, 2006.

[2] 高庆德. 一种军事情报定量分析方法研究[C]//全国博士生论坛: 军事学. 北京: [出版者不详], 2006: 545.

[3] 张晓军. 美国军事情报理论评介[M]. 北京: 时事出版社, 2006.