

基于最优解析树提取的多功能 雷达状态快速估计方法

代 鹞 鹏, 王 布 宏, 曹 帅, 沈 海 鸥
(空军工程大学信息与导航学院, 陕西西安 710077)

摘 要: 针对基于文法建模的多功能雷达(Multi-Function Radar, MFR)参数估计领域中常规算法具有的高运算复杂度问题,提出一种快速估计算法.该算法利用文法的派生过程仅与文法结构有关,而与文法概率参数无关这一事实,利用库克-杨-卡塞米(Cocke-Younger-Kasami, CYK)算法对截获雷达数据序列进行预处理,构造出可以反映该序列派生过程的解析表,进而从该解析表中提取出序列的最优解析树,然后利用改进的Viterbi-Score算法对雷达文法概率参数进行快速估计.论文仿真分析了该算法的计算复杂度、存储复杂度和估计精度,实验结果表明了该算法相对于常规算法,可以减少60%左右的计算量.

关键词: 多功能雷达;随机上下文无关语法;解析表;解析树

中图分类号: TN918.1 **文献标识码:** A **文章编号:** 0372-2112(2016)03-0514-06

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2016.03.003

Approach to Multi-function Radar Parameters Fast Estimation Based on Best Parse Tree Extract

DAI Li-peng, WANG Bu-hong, CAO Shuai, SHEN Hai-ou

(School of Information and Navigation, Air Force Engineering University, Xi'an, Shaanxi 710077, China)

Abstract: To deal with the huge computing burden of the existing multi-function radar (MFR) syntactic model parameters learning algorithms, a fast learning algorithm is proposed in light of the derivation only relevant to the syntactic architecture but the probabilities. In our method, each training sequence is pre-processed by the Cocke-Younger-Kasami (CYK) parsing algorithm, the parse chart is constructed to accurately describe the sequence derivation. Furthermore, the best parse tree is extracted from the parse chart, and the probabilities are estimated based on the best parse tree with a modified Viterbi-Score algorithm (VS). The time complexity, memory complexity and accuracy are also explored. Simulation results show that compared with the conventional algorithm, more than 60% operation time can be reduced with our proposed algorithm.

Key words: multi-function radar; stochastic context free grammar; parsing chart; parsing tree

1 引言

电子支援系统(Electronic Support, ES)通过分析战场截获的雷达辐射源信号,向我方飞行员提供该辐射源所处的方位、状态和威胁等级等情报^[1].对于传统的机械扫描雷达,ES系统可以通过分析辐射源信号的载频、脉宽等瞬时特性,采用参数类的统计方法便可实现有效告警^[2].但对于多功能雷达(Multi-Function Radars, MFR)而言,由于MFR采用软件算法对波束进行控制,

其波束扫描是无惯性的,且MFR采用了复杂多变的信号模式,使得MFR能够同时对多个目标进行操作.由于MFR的这些特性,基于参数类的威胁告警技术已无法满足MFR的建模需求^[3].为了对MFR的这种多功能性、多工作模式进行有效描述,Nikita与Wang将Markov链与随机上下文无关文法(Stochastic Context-Free Grammars, SCFG)相结合,提出模式类的MFR建模方法^[4],该方法将语言学中的知识引入到雷达信号处理领域,为今后的研究提供了新的思路.

给定能够描述 MFR 雷达动态行为的文法结构,如何利用截获的信号序列对 MFR 的文法概率参数进行估计是模式类 ES 系统面临一个关键问题.通常的方法是将期望最大化算法(Expectation-Maximization, EM)应用于语言学中的 Inside-Outside(IO)算法^[5]和 Viterbi-Score(VS)算法^[6]对概率参数进行迭代求精,但是这两种方法的收敛速度和计算复杂度均无法满足 ES 系统的实时性要求.文献[7]针对上述问题,提出 gEM(IO)和 gEM(VS)算法,但这两种算法只针对单个文法的训练,无法对多个文法进行并行训练.本文将解析表应用于 VS 算法,提出 P(VS)算法对多个雷达文法的参数进行并行快速估计,并通过理论分析与仿真实验证明 P(VS)算法相对于 IO、VS 和基于解析表的 IO 算法(P(IO)),可以有效降低运算复杂度.

2 MFR 雷达的系统结构分析

MFR 通过软件程序对任务进行调度,并将每一种任务映射为基本的脉冲波形进行发射,每个基本脉冲波形称为一个雷达字,雷达字为 MFR 的最小辐射单元^[8].由此可见,MFR 为了满足不同的功能需求而使用了分层的信号结构,因此可利用分层的雷达信号结构模型对 MFR 进行建模.该方法首先将雷达的状态转换为随机离散事件,雷达在每个状态采用特定的随机形式语言交流信息,然后在电子情报基础上利用能够捕获 MFR 复杂信号特征的文法对该语言进行建模,将每一部 MFR 在威胁数据库中对应一个文法语言模型,最后利用机器翻译和解码技术对雷达状态进行解析.

MFR 依据自身所处的战术环境选择雷达需要执行的任务状态,然后将该任务状态通过映射机制转换为雷达字序列,最后将该雷达字序列映射为相应的脉冲序列进行发射.因此,雷达状态选择机制为一个 Markov 过程,即在每个 MFR 周期选择一个合适的任务状态 e_i ,并依据状态 e_i 所定义的调度规则对雷达字进行调度^[9].从 ESM 系统角度来讲,雷达内部工作机制是未知的,因此将每一次调度行为建模为概率性事件,概率值大小代表雷达的资源分配方案.

上下文无关文法(Context-Free Grammars, CFG)具有结构形式简单、表达能力强的特点^[10],因此可利用 CFG 对 MFR 的调度规则进行建模. CFG 由四元组 $G = \{V, N, R, S\}$ 组成: V 为非终结符的非空集合; N 为终止符的非空集合,即雷达字集,且满足 $N \cap V = \emptyset$; R 是产生式集合,其形式为 $A \rightarrow \lambda$,其中 $A \in V, \lambda \in (V \cup N)^+$, R 反应了调度模块与控制模块中的调度准则; $S \in V$ 为初始符. CFG 所产生的语言集定义为 $L_g(G)$. SCFG 为 CFG 的扩展,即赋予产生式概率值 P ,构成 $G_s = \{G, P\}$,我们称 G 为 G_s 的文法结构,其中, P 为赋予文法产生式

的概率,且满足 $\sum_{\lambda} P(A \rightarrow \lambda) = 1$ 的约束条件.

利用 SCFG 对其调度方式建模,可得到 MFR 的信号产生机制如图 1 所示.该建模方法利用马尔科夫链描述雷达状态的动态转移情况,利用 SCFG 描述 MFR 产生雷达信号的调度机制.

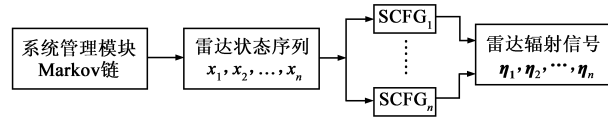


图1 多功能雷达信号产生机制

对于 CFG 而言,其产生的终止符序列 η 可观测得到,但是产生 η 的具体过程无法直观得到.由文献[11]可知,文法产生 η 的派生过程仅与文法结构有关,而与文法的概率值无关,因此可以利用 CYK 算法^[12]或 Early 算法^[13]对每个序列 η 进行预处理,构造可以精确描述该派生过程的解析表,并从中提取出最优解析树,在概率重估时,只有最优解析树参与重估过程,从而减少算法计算复杂度(计算复杂度定义为对一个训练序列进行一次迭代所需要的乘法和除法运算次数).

3 基于解析表的 MFR 参数快速估计算法

3.1 CYK 图表解析算法

对序列 η 进行分析,意味着需要确定一个文法导出 η 的产生式序列. CYK 算法作为一种描述性算法,已应用于语音识别领域. CYK 算法要求所处理的文法产生式为 $A \rightarrow w$ 或者 $A \rightarrow BC$ 两种结构^[14],其中 $A, B, C \in V, w \in N$. 根据不同的产生式定义不同的子树结构来确定 η 的子序列母节点,其中产生式 $A \rightarrow a$ 对应的子树结构为 $A(j, j) \rightarrow a$,表示 $w_j = a$,且 $A \rightarrow w_j \in R$. 产生式 $A \rightarrow BC$ 对应的子树结构为 $A(i, j) \rightarrow B(i, k)C(k+1, j)$,表示文法以 A 为母节点产生子序列 $w_i w_{i+1} \dots w_j$,在该子序列中, $w_i w_{i+1} \dots w_k$ 以 B 为母节点, $w_{k+1} \dots w_j$ 以 C 为母节点,且 $A \rightarrow BC \in R$.

对于给定的长度为 L 的序列,解析表 T 是一个三角形表,每个表中第 i 行第 j 列的参数为 $T(i, j)$,其中 $1 \leq i \leq L, 1 \leq j \leq L+1-i$. 对于 η 的某个子串,若有 $A \Rightarrow w_i w_{i+1} \dots w_j$,则把 $A(i, j) \rightarrow \dots$ 存入 $T(i, j)$ 中. 在构造出解析表后,当且仅当 $S(1, L) \rightarrow \dots$ 存在于 $T(1, L)$ 中时, $\eta \in L_g(G_s)$,且可从该解析表中抽取出该序列的最左导出.

对于序列 $\eta = w_1 w_2 \dots w_L$,可以从左到右、从最高行到最低行顺次构造解析表.由于 G 必须符合 Chomsky 正则形式表示,使得我们能按照下列步骤对 η 构造此解析表:

步骤 1 按照 $i=1$ 到 $i=L$ 的次序求 $T(i, i)$,若 R 中存在产生式 $A \rightarrow w_i$ 时,将 $A(i, i) \rightarrow w_i$ 填入 $T(i, i)$.

步骤 2 假设对 $1 \leq i \leq L$ 已经求出 $T(i, j-1)$,现在

求 $T(i,j)$. 对于 $i=L-1:1$, 有 $j=i+1:L$, 则对于 $1 \leq k \leq j-1$ 中的任何一个 k 值, 当 $A \rightarrow BC \in R$, 且 $B(i,k) \rightarrow \dots \in T(i,k)$, $C(k+1,j) \rightarrow \dots \in T(k+1,j)$, 则将 $A(i,j) \rightarrow B(i,k)C(k+1,j)$ 填入 $T(i,j)$.

步骤 3 重复第二步直至完成此表或表的整行都是空项.

通过对每个序列提前构造解析表可以发现, 解析表中的子树结构参数描述了文法产生 η 的所有派生过程, 我们可在算法每次迭代过程中利用解析表排除不参与派生的参数, 从而可以减少计算量.

3.2 MFR 文法概率快速估计算法

令 $x_{1:m} = (x_1, x_2, \dots, x_m)$ 为 MFR 状态序列, $x_k \in N$. $\eta_{1:m} = (\eta_1, \eta_2, \dots, \eta_m)$ 表示截获得到的 MFR 产生的 m 个终结符序列, 其中 $\eta_i = w_1 w_2 \dots w_{L_i}$, L_i 表示 η_i 的长度. 为了对文法进行解析, 首先引入内部概率 $\alpha_{A(i,j)} = P(w_{ij} | A, G_s)$, 该概率表示 G_s 从非终结符 A 开始, 生成终结符序列 $w_i w_{i+1} \dots w_j$ 的概率.

IO 算法与 P(IO) 算法在基于训练数据集的基础上寻找全局最大似然, 而 VS 算法仅寻找训练序列最优派生过程的最大似然. P(VS) 算法首先利用解析表对每个序列 η 进行预处理, 提取出最优解析树 \check{d}_η , 然后利用 \check{d}_η 对文法参数进行重估, P(VS) 算法的步骤可归纳如下:

步骤 1 自底到顶求解内部概率. 令 $\check{\alpha}_{A(i,j)}^{(n)}$ 表示从非终结符 A 生成终结符序列 $w_i \dots w_j$ 的所有子树的最大内部概率. 若子序列只由一个终止符 $w_j = a$, 且 $A \rightarrow a \in R$, 则依据该序列对应的 CYK 解析表中子树结构 $A(j,j) \rightarrow a$ 计算最大内部概率:

$$\check{\alpha}_{A(j,j)}^{(n)} = P^{(n)}(A \rightarrow w_j) \quad (1)$$

对于每个母节点为 $A(i,j)$ ($i < j$) 的子树结构 $A(i,j) \rightarrow B(i,k)C(k+1,j)$, 其分支概率定义为:

$$\check{\alpha}_{A(i,j) \rightarrow B(i,k)C(k+1,j)}^{(n)} = P^{(n)}(A \rightarrow BC) \alpha_{B(i,k)}^{(n)} \alpha_{C(k+1,j)}^{(n)} \quad (2)$$

因此, 母节点 $A(i,j)$ 对应的最大内部概率为所有以 $A(i,j)$ 为母节点的子树分支概率的最大值:

$$\check{\alpha}_{A(i,j)}^{(n)} = \max_{A(i,j) \rightarrow B(i,k)C(k+1,j)} \check{\alpha}_{A(i,j) \rightarrow B(i,k)C(k+1,j)}^{(n)} \quad (3)$$

对于每个母节点 $A(i,j)$, 使得式(3)最大的子树分支称为最优子树分支, 定义为:

$$A(i,j) \xrightarrow{\Phi} B(i,k)C(k+1,j), (B, C, k) = \Phi_{A(i,j)} \quad (4)$$

令 Φ 保存每个最优子树的路径, 即:

$$\Phi_{A(i,j)}^{(n)} = \arg \max_{(B, C, k)} P^{(n)}(A \rightarrow BC) \alpha_{B(i,k)}^{(n)} \alpha_{C(k+1,j)}^{(n)} \quad (5)$$

在得到 Φ 后, 便可以从 $\Phi_{S(1,L)}^{(n)}$ 开始从上而下迭代构造该序列的最优解析树 \check{d}_η .

步骤 2 依据 \check{d}_η 计算产生式使用次数: 在 VS 算法

中, 由于只有 \check{d}_η 参与重估过程, 因此可依据算法 1 对 Φ 进行访问来提取文法 $G^{(n)}$ 产生序列 η_i 时产生式的期望使用次数 $c^{(n)}(A \rightarrow a | \eta_i)$ 和 $c^{(n)}(A \rightarrow BC | \eta_i)$:

算法 1 计算产生式使用次数算法

```

For each  $r \in R$ 
   $c(A \rightarrow r) = 0$ 
end
 $A = S; i = 1; j = L = \text{length}(\eta)$ 
Visit  $\Phi_{(A, i, j)}$ 
  If  $i = j$ 
     $c(A \rightarrow w_j) = c(A \rightarrow w_j) + 1$ 
  else
     $(B, C, k) = \Phi_{(A, i, j)}$ 
     $c(A \rightarrow BC) = c(A \rightarrow BC) + 1$ 
    Visit  $\Phi_{(B, i, k)}$ 
    Visit  $\Phi_{(C, k+1, j)}$ 
  end
end

```

步骤 3 概率重估. 在得到文法产生每个序列的产生式期望使用次数以后, 文法产生式的概率便可估计为:

$$P^{(n)}(A \rightarrow \lambda) = \frac{\sum_{k=1}^m c^{(n)}(A \rightarrow \lambda | \eta_k) \chi_k(n)}{\sum_{\mu} \sum_{k=1}^m c^{(n)}(A \rightarrow \mu | \eta_k) \chi_k(n)} \quad (6)$$

其中 $\xi_t(i, j) = P(x_t = e_i, x_{t+1} = e_j | \eta_{1:m})$ 表示在 t 时刻 MFR 处于 e_i 状态, $t+1$ 时刻处于 e_j 状态的概率. $\chi_t(i, j) = P(x_t = e_i | \eta_{1:m})$ 表示 MFR 在 t 时刻处于状态 e_i 的概率. 则雷达管理模块对应的 Markov 链状态转移概率可估计为:

$$a_{ij} = \frac{\sum_{t=1}^{m-1} \xi_t(i, j)}{\sum_{t=1}^{m-1} \sum_{j=1}^M \xi_t(i, j)} \quad (7)$$

通过上述步骤进行迭代计算, 直至文法概率参数变化非常小时结束. 每经过一次迭代, 模型对于给定观测数据的似然性都会增加, 这样就保证了每次迭代都会产生更优的模型参数. 通过 P(VS) 算法得到文法的概率参数后, 便可利用 Viterbi 算法对 MFR 所处状态进行估计^[15].

4 算法分析与仿真

为了验证本文的算法, 我们首先对四种算法的计算复杂度进行理论分析, 并利用表 1 所示文法进行仿真验证. 在该文法基础上, 我们设计两个实验: 第一个实验对算法时间复杂度进行分析. 第二个实验对算法的收

敛时间和精度进行分析.

4.1 算法资源分析

为了对算法资源进行理论分析,首先定义以下变量: L 表示训练序列的平均长度; M_t 表示辐射规则的数目; M_{tm} 表示文法终结符的数目; M_{nt} 表示文法非终结符的数目; $|\varphi_t|$ 表示解析表中转移规则的平均数; $|\Delta_e|$ 表示解析表中辐射规则的平均数; $|\Delta l_t|$ 表示解析表中每个转移规则对应的平均子树的个数; $|\Delta l_e|$ 表示左部参数相同的辐射规则平均数.由于算法在产生式概率均匀分布时计算复杂度最大,则在该条件下有:

$$|\varphi_t| = M_{nt} \cdot \sum_{i=1}^L \sum_{j=1}^{i-1} L = M_{nt} \cdot (L^2 - 1) \quad (8)$$

$$|\varphi_e| = M_{nt} \cdot L \quad (9)$$

$$|\Delta l_t| = M_{nt}^2 \cdot L \quad (10)$$

则对于 IO 算法而言,其计算复杂度 T_{IO} 为:

$$\begin{aligned} T_{IO} &= \underbrace{\sum_{j=2}^L \sum_{i=1}^{j-1} \sum_{k=i}^{j-1} 2 \cdot M_{nt}^3}_{\text{IO-Inside}} \\ &+ \underbrace{\sum_{i=1}^L \sum_{j=L}^i \left(\sum_{k=j+1}^L 2 \cdot M_{nt}^3 + \sum_{k=1}^{i-1} 2 \cdot M_{nt}^3 \right)}_{\text{IO-Outside}} \\ &+ \underbrace{M_{nt} \cdot M_t \cdot (L + 1)}_{\text{IO-一元规则}} \\ &+ \underbrace{\sum_{i=1}^{L-1} \sum_{j=i+1}^L \sum_{k=i}^{j-1} 2 \cdot M_{nt}^3 + M_{nt}^3}_{\text{IO-二元规则}} \\ &+ \underbrace{M_{nt}^3 + M_{nt} \cdot M_t}_{\text{IO-概率重估}} \\ &= \frac{4}{3} M_{nt}^3 \cdot L \cdot (L^2 - 1) + M_{nt} \cdot M_t \cdot (L + 2) + 2 \cdot M_{nt}^3 \end{aligned} \quad (11)$$

对于 VS 算法而言,其计算复杂度 T_{VS} 为:

$$\begin{aligned} T_{VS} &= \underbrace{\sum_{j=2}^L \sum_{i=1}^{j-1} \sum_{k=i}^{j-1} 2 \cdot M_{nt}^3}_{\text{VS-Inside}} + \underbrace{M_{nt}^3 + M_{nt} \cdot M_t}_{\text{VS-概率重估}} \\ &= M_{nt}^3 \cdot \left(\frac{L \cdot (L^2 - 1)}{3} + 1 \right) + M_{nt} \cdot M_t \end{aligned} \quad (12)$$

对于 P(IO) 算法,其平均计算复杂度 $T_{P(IO)}$ 为:

$$\begin{aligned} T_{P(IO)} &= \underbrace{2 \cdot |\varphi_t| \cdot |\Delta l_t|}_{\text{P(IO)-Inside}} + \underbrace{4 \cdot |\varphi_t| \cdot |\Delta l_t|}_{\text{P(IO)-Outside}} + \underbrace{2 \cdot |\varphi_t| \cdot |\Delta l_t|}_{\text{P(IO)-二元规则}} \\ &+ \underbrace{2 \cdot |\varphi_e|}_{\text{P(IO)-一元规则}} + \underbrace{M_{nt}^3 + M_{nt} \cdot M_t}_{\text{P(IO)-概率重估}} \\ &= 8 \cdot |\varphi_t| \cdot |\Delta l_t| + 2 \cdot |\varphi_e| + M_{nt}^3 + M_{nt} \cdot M_t \end{aligned} \quad (13)$$

对于 P(VS) 算法,其平均计算复杂度 $T_{P(VS)}$ 为:

$$\begin{aligned} T_{P(VS)} &= \underbrace{2 \cdot |\varphi_t| \cdot |\Delta l_t|}_{\text{P(VS)-Inside}} + \underbrace{M_{nt}^3 + M_{nt} \cdot M_t}_{\text{P(VS)-概率重估}} \\ &= 2 \cdot |\varphi_t| \cdot |\Delta l_t| + M_{nt}^3 + M_{nt} \cdot M_t \end{aligned} \quad (14)$$

定义算法对一个序列进行一次迭代时需要存储的参数个数为算法的存储复杂度.因此,IO 算法的存储复杂度 M_{IO} 为:

$$M_{IO} = \underbrace{M_{nt} \cdot L^2}_{\text{内部变量}} + \underbrace{M_{nt} \cdot L^2}_{\text{外部变量}} + \underbrace{M_{nt}^3}_{\text{转移规则}} + \underbrace{M_{nt} \cdot M_{tm}}_{\text{辐射规则}} \quad (15)$$

VS 算法的存储复杂度 M_{VS} 为:

$$M_{VS} = \underbrace{M_{nt} \cdot L^2}_{\text{内部变量}} + \underbrace{M_{nt}^3}_{\text{转移规则}} + \underbrace{M_{nt} \cdot M_{tm}}_{\text{辐射规则}} \quad (16)$$

由于 P(IO) 与 P(VS) 算法在对训练进行预处理时,需对解析表进行存储,对于长度为 L 的序列,其解析表需要存储的参数个数为:

$$\begin{aligned} M_{CYK} &= \underbrace{\left(\frac{L \cdot (L+1)}{2} - L \right)}_{\text{解析表中由转移规则占据的表格个数}} \cdot \underbrace{(6 \cdot (|\varphi_t| \cdot |\Delta l_t|))}_{\text{每个表格所需存储的参数个数}} \\ &+ \underbrace{\frac{L}{\text{辐射规则占据的表格个数}}}_{\text{辐射规则占据的表格个数}} \cdot |\varphi_e| \end{aligned} \quad (17)$$

因此,于 P(IO) 与 P(VS) 算法的存储复杂度分别为:

$$M_{P(IO)} = M_{IO} + M_{CYK} \quad (18)$$

$$M_{P(VS)} = M_{VS} + M_{CYK} \quad (19)$$

参数 $|\varphi_t|$ 和 $|\Delta l_t|$ 反应了文法的模糊性,因此 P(IO) 和 P(VS) 的平均计算复杂度与文法固有的模糊性具有很大关联.文献[4]中所描述的“水星”MFR 部分文法产生式为 $S \rightarrow ACQ \ RR, ACQ \rightarrow ACQ \ NA \ W_1 \ W_1 \ W_2 \ W_2 \ W_3 \ W_3 \ W_4 \ W_4 \ NA \rightarrow S_1 \ T, RR \rightarrow RR \ NA \ RR_p \ ACQ, T \rightarrow W_2 \ W_3 \ NA_p \rightarrow W_2 \ W_3, RR_p \rightarrow W_4 \ W_5, S_1 \rightarrow W_1 \ W_2 \ W_3, W_1 \rightarrow w_1, W_2 \rightarrow w_2, W_3 \rightarrow w_3, W_4 \rightarrow w_4$ 和 $W_5 \rightarrow w_5$, 其非终止符集 $V = \{S, ACQ, RR, NA, S_1, T, NA_p, RR_p, W_1, W_2, W_3, W_4, W_5\}$, 终止符集 $N = \{w_1, w_2, w_3, w_4, w_5\}$, 初始符为 S . 由该文法产生序列构造的解析表中, $|\Delta l_t|_{\max} = 3$, 将其带入上面各式,可得到不同序列长度与不同算法的计算复杂度、存储复杂度比较结果,如图 2 与图 3 所示.

4.2 算法时间复杂度分析

假设 MFR 有两个状态,分别对应文法 $G^{(1)}$ 和 $G^{(2)}$, 由状态转移矩阵为 $A = (0.7, 0.3; 0.4, 0.6)$ 的 Markov 链随机产生 50 个雷达状态,每个状态依据文法 $G^{(1)}$ 或 $G^{(2)}$ 随机产生一个终结符序列,然后对产生式概率和状态转移概率进行估计.经过 100 次 Monte Carlo 实验,得到算法实验结果如图 4 和图 5 所示.

算法的计算复杂度反应到程序中,便为程序的运行时间,我们通过不同迭代次数、不同序列个数与算法的运行时间进行验证.由图 4 和图 5 可知,随着迭代次数和训练序列的增加,IO 算法的运行时间也急剧增加,这导致了其在实际应用中具有局限性.VS 算法只对最大解析树进行求解,算法对序列进行一次迭代运行时间比 IO 算法减少一半左右.P(IO) 与 P(VS) 算法对训练序列进行了预处理,其中,P(IO) 算法对解析表中的所有路径求解,P(VS) 算法只对解析表定义的最优解析树求解,其算法运行时间比 IO 与 VS 算法降低了一半以上,且实验值可以很好的符合理论分析.

4.3 算法性能分析

该实验通过对比 IO、VS、P(IO) 和 P(VS) 算法收敛

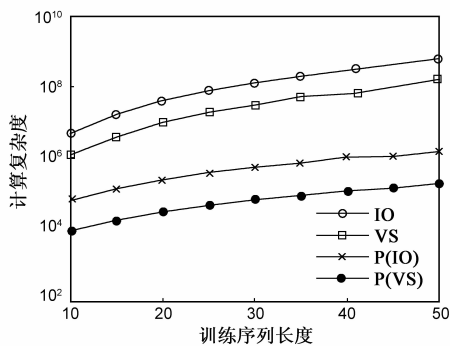


图2 序列长度与计算复杂度比较

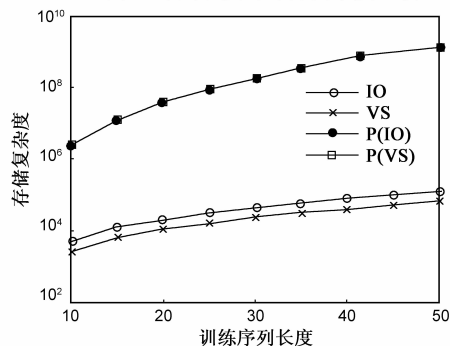


图3 序列长度与存储复杂度比较

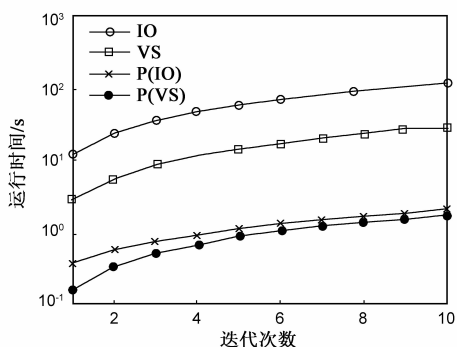


图4 迭代次数与算法运行时间比较

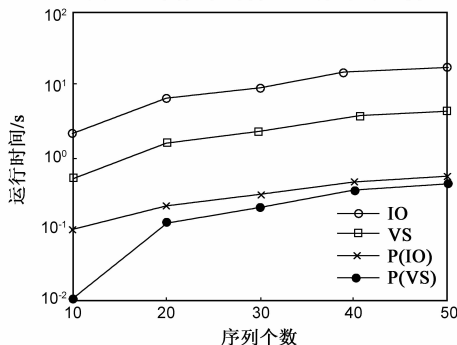


图5 序列个数与算法运行时间比较

后的估计结果与真实值的误差平方和与 MFR 状态估计概率来判别算法性能. 定义第 $n+1$ 次迭代结果和第 n 次迭代结果的误差平方和小于 0.001 时结束. 经过 100 次 Monte Carlo 实验, 结果如图 6、7 所示.

由图 6 可知, 四种算法随着迭代次数的增加, 其估

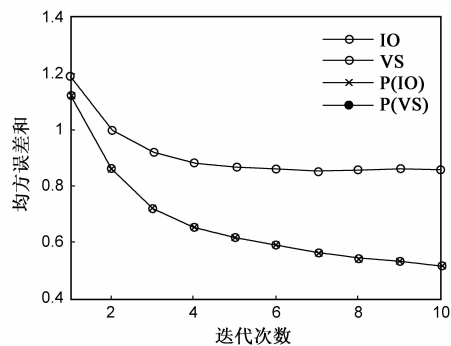


图6 迭代次数与文法估计精度比较

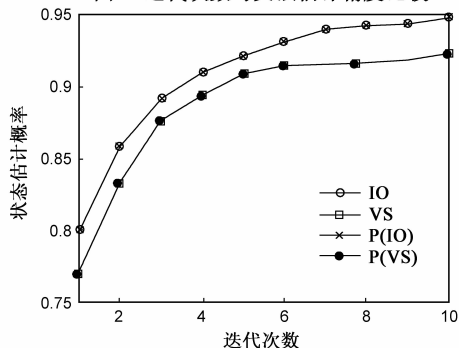


图7 迭代次数与状态估计概率比较

计值与真实值的误差会逐渐降低, 趋于收敛时, $P(\text{IO})$ 和 IO 算法的估计误差相同, 而 VS 与 $P(\text{VS})$ 算法的估计误差值始终比 $P(\text{IO})$ 和 IO 算法大 0.4 左右. 文法概率参数的估计精度会影响到最终 MFR 状态的估计, 不同迭代次数与 MFR 状态估计结果如图 7 所示, 由图可知, IO 算法的与 $P(\text{IO})$ 算法在收敛后的状态估计概率在 0.91 左右, 而 VS 与 $P(\text{VS})$ 算法的估计概率为 0.94 左右.

5 结论

本文针对基于 SCFG 建模的 MFR 参数快速估计算法进行研究, 提出一种基于解析表的参数估计算法. 该算法通过 CYK 解析算法对截获的训练序列集预先构造解析表, 在序列集解析表的基础上排除未参与序列派生过程的产生式, 并寻找出最优解析树, 且只有最优解析树参与重估过程, 大大的减少了算法复杂度. 本文通过理论分析与试验仿真证明, 该方法在减少运算量的同时, 可以确保其精度与 IO 、 $P(\text{IO})$ 算法大致相同, 在 ES 系统的实际应用中具有重要意义, 如何减少该算法的存储需求, 是下一步的研究方向之一.

参考文献

- [1] 刘海军, 李悦, 柳征, 周一宇. 基于随机文法的多功能雷达识别方法[J]. 航空学报, 2010, 31(9): 1809-1817.
LIU Hai-jun, LI Yue, ZHOU Yi-yu. Approach to multi-function radar identification based on stochastic grammars

- [J]. Acta Aeronautica Et Astronautica Sinica, 2010, 31 (9):1809 - 1817. (in Chinese)
- [2] Wang A, Krishnamurthy V. Signal interpretation of multi-function radars; modeling and statistical signal processing with stochastic context free grammar [J]. IEEE Transactions on Signal Processing, 2008, 56(3):1106 - 1119.
- [3] 刘海军. 雷达辐射源识别关键技术研究[D]. 长沙:国防科学技术大学, 2010.
- [4] N Visnevski. Syntactic Modeling of Multi-function Radars [D]. Canada:McMaster University, 2005.
- [5] K Lari, S Young. The estimation of stochastic context-free grammars using the inside-outside algorithm[J]. Computer Speech and Language, 1990, 4(1):35 - 56.
- [6] H Ney. Stochastic Grammars and Pattern Recognition[M]. New York:Springer Verlag, 1992.
- [7] Latombe G, Granger E, Dilkes F A. Fast learning of grammar production probabilities in radar electronic support[J]. IEEE Transactions on Aerospace and Electronic Systems, 2010, 46(3):1262 - 1289.
- [8] 马爽, 柳征, 姜文利. 基于幅度变化点检测的多功能雷达脉冲列解析方法[J]. 电子学报, 2013, 41(7):1436 - 1441.
- MA Shuang, LIU Zheng, JIANG Wen-li. A method for multi-function radar pulse train analysis based on amplitude change point detection[J]. Acta Electronica Sinica, 2013, 41(7):1436 - 1441. (in Chinese)
- [9] Visnevski N, Krishnamurthy V, Wang A, Haykin S. Syntactic modeling and signal processing of multi-function radars: a stochastic context-free grammar approach [J]. Proceedings of the IEEE, 2007, 95(5):1000 - 1025.
- [10] D Manning, H Schutze. 统计自然语言处理基础[M]. 范春法, 等, 译. 北京:电子工业出版社, 2002. 241 - 256.
- [11] R C Gonzalez, M G Thomason. 句法模式识别[M]. 濮群, 译. 北京:清华大学出版社, 1984, 84 - 88.
- [12] Tsuruoka Y, Tsujii J. Iterative CYK parsing for probabilistic context-free grammars [A]. First International Joint Conference on Natural Language Processing[C]. Hainan, China: IJCNLP, 2005. 52 - 60.
- [13] J Early. An efficient context-free parsing algorithm[J]. Communication of the ACM, 1970, 13(2):84 - 102.
- [14] 陆玲, 周书民. 形式语言与自动机及程序设计[M]. 哈尔滨:哈尔滨工业大学出版社, 2014, 11 - 16.
- [15] 代鹏鹏, 王布宏, 蔡斌, 刘军利. 基于 SCFG 建模的多功能雷达状态估计算法[J]. 空军工程大学学报·自然版, 2014, 15(3):24 - 28.
- DAI Li-peng, WANG Bu-hong, CAI Bin, LIU Jun-li. A method for states estimation of multi-function radar based on stochastic context-free grammar [J]. Journal of Air Force Engineering University (Natural Science Edition), 2014, 15(3):24 - 28. (in Chinese)

作者简介



代鹏鹏 男, 1989 年 1 月出生于甘肃省平凉市. 现为空军工程大学信息与导航学院硕士研究生. 主要研究方向为多功能雷达信号处理.
E-mail: dlipeng@163.com



王布宏 男, 1975 年 12 月出生于山西省太原市. 现为空军工程大学教授、博士生导师. 主要从事阵列信号处理、阵列校正等方面的研究工作.
E-mail: wbhyl@aliyun.com



曹帅 男, 1991 年 11 月出生于陕西省西安市. 现为空军工程大学信息与导航学院硕士研究生. 主要研究方向为阵列信号处理.
E-mail: 465782523@qq.com