

基于隐马尔可夫模型的行为轨迹还原算法

冯 涛, 郭云飞, 黄开枝, 吉 江

(国家数字交换系统工程技术研究中心, 郑州 450002)

摘 要: 针对行为轨迹还原过程中观察序列状态缺失、无法对终端轨迹进行精确还原的问题, 提出一种基于隐马尔可夫模型的行为轨迹还原算法。利用基站布局的空间相关性, 在不考虑缺失观察状态的情况下, 对隐马尔可夫模型求解过程中的局部概率进行修订, 还原出轨迹序列。性能分析和仿真结果表明, 状态倾向度越大, 轨迹还原成功率越高, 当状态倾向度取 0.8 时, 轨迹还原成功率在 90% 左右。

关键词: 行为轨迹; 状态倾向度; 轨迹还原; 状态缺失; 局部概率; 隐马尔可夫模型

Behavior Trajectory Restoration Algorithm Based on Hidden Markov Models

FENG Tao, GUO Yun-fei, HUANG Kai-zhi, JI Jiang

(National Digital Switching System Engineering & Technological R&D Center, Zhengzhou 450002, China)

【Abstract】 This paper proposes a behavior trajectory restoration algorithm for observation sequence state missing problem, which leads to terminal trajectory restoration inaccurately. The algorithm utilizes base station layout's spatial correlation and revises the partial probability of the solution process of the Hidden Markov Models(HMM) to restore the track sequence without considering the missing observation states. Performance analysis and simulation results show that the greater the degree of state propensity is, the higher the success rate of trajectory restoration is. When the degree of state propensity is 0.8, the success rate of trajectory restoration is about 90 percent.

【Key words】 behavior trajectory; state propensity degree; trajectory restoration; state missing; partial probability; Hidden Markov Models(HMM)

DOI: 10.3969/j.issn.1000-3428.2012.18.001

1 概述

近年来, 随着无线通信技术的快速发展, 全球范围内移动终端的规模和数量日益增大, 针对终端的定位及轨迹跟踪技术也得到快速发展。当终端在特定区域内运动时, 会在基站范围内留下行为痕迹, 如越区注册、短信、呼叫等。因此, 可以在不依赖于其他辅助定位手段而只通过基站获取终端信息的条件下, 获知终端大范围内的移动轨迹。然而, 由于阴影衰落、漏判以及移动终端关闭等原因, 基站长时间未发现终端或终端两相邻观测点之间的时间间隔较长, 从而无法获知终端较为精确的行为轨迹。

文献[1-3]引入隐马尔可夫模型算法, 在不同场景中根据运动目标轨迹复杂程度的不同, 建立相应的隐马尔可夫模型进行轨迹识别。然而, 运动目标轨迹识别算法是基于计算机视觉对所监视场景中运动目标轨迹的行为模式进行分析与识别, 无法适用于大范围内的终端行为轨迹跟踪。文献[4-6]介绍了缺失特征方法, 该方法在语音方面的应用较为广泛, 它根据语音局部信噪比的不同将语音分为“可靠”及“缺失”状态, 首先对“缺失矢量”进行还原,

得到完整矢量后再进行语音识别, 缺失特征方法为观察序列缺失状态的估计提供了有力的参考。文献[7-8]提出了一种基于隐马尔可夫模型的局部最优状态路径的数据重建(LOPDI)算法, 该算法按最大后验概率准则(MAP)重建出“缺失状态”。然而, LOPDI 算法只适用于序列各个状态点由多个状态参量构成的情况, 它利用多个状态参量内部之间的关系对缺失状态点进行还原, 无法适用于序列状态点只由单个状态参量构成的情况。因此, 针对观察状态缺失的终端行为轨迹还原, 仍存在以下问题: (1)如何判断观察序列中两相邻状态之间的信息丢失; (2)如何在只获知终端部分观察序列的前提下, 有效地对终端行为轨迹进行还原。

本文提出了一种基于隐马尔可夫模型的行为轨迹还原算法。设计一种缺失状态估计方法, 判断出观察序列中两相邻状态之间的信息丢失, 在不考虑缺失观察状态的情况下, 对维特比算法中的局部概率计算方法进行修订, 令 $d_i(i) = \max_j(d_{i-1}(j)a_{ji})$, 还原出轨迹序列。同时, 对算法性能的分析给出了影响算法性能的 2 个主要因素: 状态倾向

基金项目: 国家自然科学基金资助项目(61171108)

作者简介: 冯 涛(1988—), 男, 硕士研究生, 主研方向: 无线移动通信; 郭云飞, 教授、博士生导师; 黄开枝, 副教授; 吉 江, 博士研究生

收稿日期: 2011-11-28 **修回日期:** 2012-01-19 **E-mail:** aifengtao@163.com

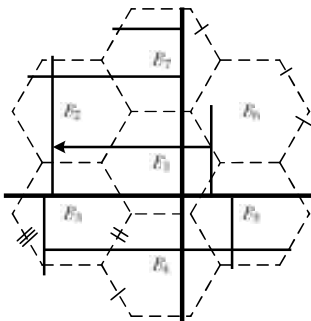
度和缺失状态比重, 状态倾向度越大, 轨迹还原成功率越高。

2 算法背景

假设某一地区有 n 个基站, 用 $B = (B_1, B_2, \dots, B_i, \dots, B_n)$ ($i=1, 2, \dots, n$) 表示, 基站覆盖情况如图 1(a)所示, 其中, $n=7$ 。终端由一个基站进入另一个基站的概率(状态转移概率)用 $a_{ij} = P(B_j/B_i)$ 表示, 所有基站之间的转移概率构成转移概率矩阵 $A = (a_{ij})$, ($i, j=1, 2, \dots, n$)。基于基站布局及道路信息得到该地区的拓扑图如图 1(b)所示, 利用宽度不同的实线依次代表宽度不同的道路。根据相邻基站之间连接的道路类型及数量, 归一化后可以确定相邻基站之间的转移概率, 进而得到该区域基站间的状态转移矩阵。



(a) 基站布局



(b) 拓扑图

图 1 基站布局及拓扑图

用 $\pi = [\pi_1, \pi_2, \dots, \pi_i, \dots, \pi_n]$, ($i=1, 2, \dots, n$) 表示 $t=1$ 时刻终端所处状态的概率分布, 它满足 $\sum_{i=1}^n \pi_i = 1$ 。当终端在特定区域内运动时, 基站可以获知终端的越区注册、短信、呼叫等通信行为。

$\bar{B} = [\bar{B}_1, \bar{B}_2, \dots, \bar{B}_i, \dots, \bar{B}_N]$, ($i=1, 2, \dots, N$) 表示 N 个观测

到移动终端的基站, 可以通过基站的观测记录对基站覆盖范围内移动终端的行为轨迹进行还原。假设基站侧记录的移动终端行为序列为:

$$COR = \begin{bmatrix} t_1 & t_2 & t_3 & \dots & t_i & \dots & t_N \\ S_1 & S_2 & S_3 & \dots & S_i & \dots & S_N \end{bmatrix} \quad (1)$$

其中, t_i 为基站 S_i 检测到终端的时间。终端在基站 B_i 中的持续时间为 τ^{B_i} , 即终端从区域 B_i 转移到区域 B_j 的时间间隔为:

$$\tau^{B_i} = t_i - t_j \quad (2)$$

利用终端行为序列进行轨迹还原主要包含 2 个随机过程, 如图 2 所示: (1)终端从一个基站 B_i 的覆盖范围转移到另一个基站 B_j 的覆盖范围的概率 a_{ji} ; (2)由于高楼阻挡等原因, 实际接收到终端信号的基站与终端的位置关系存在一定的随机性; 假设基站 B_j 观测到终端的概率为 $P(B_j)$, 但实际上终端在基站 B_k 的物理范围内的概率为 $w_{kj} = P(B_k/B_j)$, 其中, $\sum_{k=1}^n w_{kj} = 1$, ($k=1, 2, \dots, n$), 所有基站之间的误判概率构成概率矩阵 $W = (w_{kj})$ 。同时, 由于高楼阻挡及终端关闭等原因, 导致基站侧没有记录到终端轨迹点, 即基站侧记录的终端观察状态有缺失, 如图 2 中状态 \bar{B}_{t_0} 所示。因此, 如何在观察序列轨迹点(状态)缺失的情况下还原出终端真实的行为轨迹是本文要解决的问题, 即由观察序列 $\bar{B}_1, \bar{B}_2, \dots, \bar{B}_N$ 还原出隐藏状态 \hat{B}_{t_0} 和隐藏序列 $\hat{B}_1, \hat{B}_2, \dots, \hat{B}_N$ 。

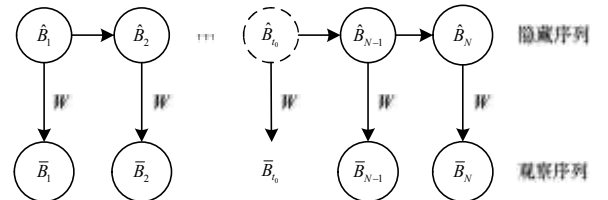


图 2 轨迹还原过程

3 行为轨迹还原算法

针对以上问题, 本文设计了一种基于隐马尔可夫模型的行为轨迹还原算法。首先, 判断观察序列中两相邻状态之间是否有信息丢失。如果存在终端轨迹点缺失, 那么两相邻观测点之间的时间间隔会较长。因此, 可以根据缺失特征方法^[4-6], 利用持续时间 τ^{B_i} 作为度量参数来估计缺失状态, 即:

$$\bar{B}_{i,j+1} = \begin{cases} 0 & \tau^{B_i} > \tau \\ 1 & \tau^{B_i} \leq \tau \end{cases} \quad (3)$$

其中, $\bar{B}_{i,j+1}=1$ 表示状态信息“可信”, 即状态没有缺失; 反之, $\bar{B}_{i,j+1}=0$ 表示状态信息缺失。 τ 为时间度量标准, 可以对基站覆盖范围内移动终端的持续时间进行统计得到该值。

如图 3 所示, 大部分终端在基站范围内的运动时间在 t_m 上下波动, 令:

$$\tau = t_m \times (\kappa + 1) \quad (4)$$

其中, $\kappa \in [0, 1]$ 为时间容忍因子。

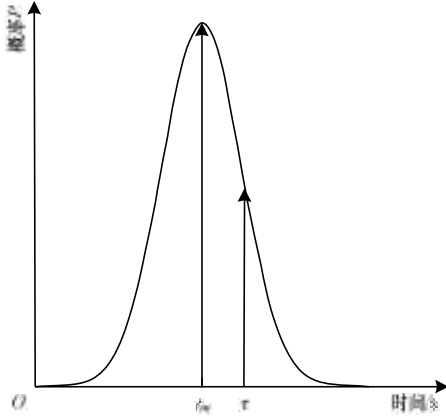


图3 持续时间概率分布

当确知有状态点 \bar{B}_{t_0} 缺失后, 可以利用基站布局的空间相关性、最大关联概率还原出中间缺失状态。如图4所示, 隐藏状态之间的转移概率即为状态转移概率 A , 可以通过观察状态 \bar{B}_i 感知到隐藏状态 \hat{B}_i 的存在。对于网格中的每一个中间及终止状态 i , 都有一个到达该状态的最可能路径, 定义为局部概率 $\delta_i(i)$, 表示 t 时刻到达状态 i 的最可能路径的概率^[9]。因而, $\delta_i(i)$ 是 t 时刻到达状态 i 的所有序列中概率最大的。

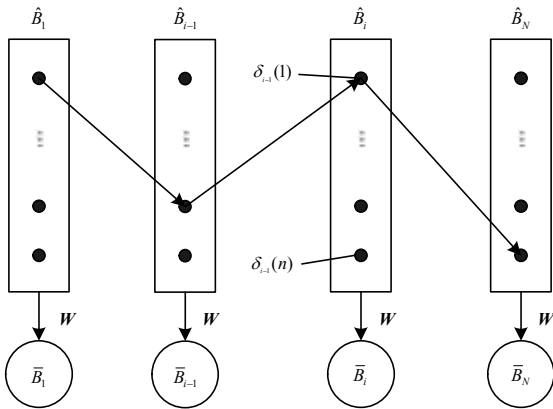


图4 轨迹还原模型

首先, 计算初始时刻到达某一状态的最大概率, 记为 $\delta_1(i)$; 然后, 将观察状态序列分为缺失状态与非缺失状态。若为非缺失状态, 则直接对局部概率 δ 进行更新。一阶隐马尔可夫模型的当前状态只与其前一个状态有关, 因此可以取 t 时刻所有基站的状态转移概率与 $t-1$ 时刻获得的最大局部概率之积的最大值, 对状态 i 的局部概率 $\delta_t(i)$ 进行更新, 同时乘以隐藏状态与观察状态的统计对应关系 $w_{\bar{B}_t, B_t}$, 即:

$$\delta_t(i) = \max_j (\delta_{t-1}(j) a_{jy} w_{\bar{B}_t, B_t}) \quad (5)$$

若为缺失状态 \bar{B}_{t_0} , 则无法知道 $w_{\bar{B}_{t_0}, B_t}$, 只能根据基站状态转移的相关性对局部概率 $\delta_{t_0}(i)$ 进行修订, 即:

$$\delta_{t_0}(i) = \max_j (\delta_{t_0-1}(j) a_{ij}) \quad (6)$$

同时, 对于当前时刻, 需要记录下上一时刻触发此概率的状态。当状态遍历完成后, 选取概率值最大的序列, 从而确定系统完成时最可能的隐藏状态序列。最后, 按最可能的状态路径在整个网络回溯, 所得到的序列即所求结果。

此外, 当需要确知实际的缺失观察状态时, 对于缺失观察状态 \bar{B}_{t_0} , 因为特定区域内基站数目一定, 所以其值为 B 中一个, 令:

$$\bar{B} = (\bar{B}_1, \bar{B}_2, \dots, B_m, \dots, \bar{B}_N) \quad (7)$$

其中, m 依次取 $1, 2, \dots, n$, 对于新生成的每一个观察序列, 都可由维特比算法得到与之对应的隐藏状态序列。当所有观察状态遍历完成后, 得到 n 个序列 $l_1, l_2, \dots, l_i, \dots, l_n$ 及其相应的概率值 $p_1, p_2, \dots, p_i, \dots, p_n$ 。取概率值最大的隐藏状态序列作为所求结果, 即:

$$L_{\text{Res}} = l_z$$

$$z = \arg \max_i (p_i) \quad i = 1, 2, \dots, n \quad (8)$$

此时 B_m 的取值即缺失的观察状态。

综上所述, 针对观察序列状态缺失的轨迹还原算法流程如图5所示。

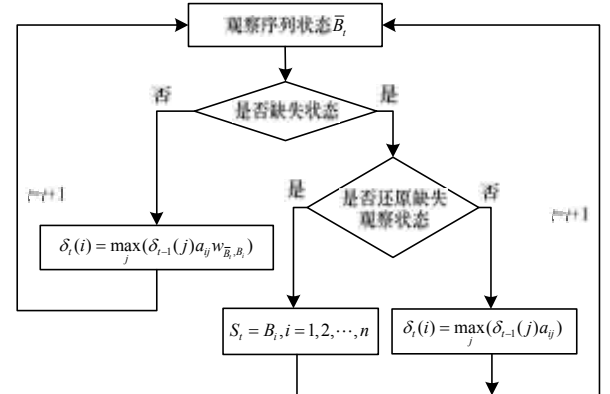


图5 轨迹还原算法流程

算法流程的具体步骤如下:

(1) 初始化: 确定初始概率矩阵 $\pi = [\pi_1, \pi_2, \dots, \pi_i, \dots, \pi_n]$ ($i = 1, 2, \dots, n$)。

(2) 局部概率更新: 当 $t=1$ 时, $\delta_1(i) = \pi_i w_{\bar{B}_1, B_1}$ 。当 $t > 1$ 时, 对于观察序列状态 \bar{B}_t , 若为非缺失状态, 那么 $\delta_t(i) = \max_j (\delta_{t-1}(j) a_{jy} w_{\bar{B}_t, B_t})$ 。若为缺失状态 \bar{B}_{t_0} , 那么 $\delta_{t_0}(i) = \max_j (\delta_{t_0-1}(j) a_{ij})$ 。

(3) 记录状态: $\Psi_t(i) = \arg \max_j (\delta_{t-1}(j) a_{jy})$ 。

(4) 选取最优序列: $\hat{B}_N = \arg \max_i (\delta_N(i))$ 。

(5) 路径回溯: 令 $\hat{B}_t = \Psi_{t+1}(\hat{B}_{t+1})$, 所得到的序列 $\hat{B}_1, \hat{B}_2, \dots, \hat{B}_N$ 即所求结果。

4 性能分析和仿真验证

4.1 性能分析

如图6(a)所示, \bar{B}_{t_0} 、 \bar{B}_{t_0-1} 为 t_0 、 t_0-1 时刻隐藏状态的

估计。一阶马氏链当前状态只与其之前状态有关，因此，将 t_0-1 时刻的局部概率 $\delta_{t_0-1}(i)$, ($i=1,2,\dots,n$) 看作初始概率分布, $P_{e1}, P_{e2}, \dots, P_{ei}$ 分别表示 $1, 2, \dots, i (i < N)$ 个缺失状态还原错误的概率, 则:

$$P_{e1} = p(\tilde{B}_{t_0} = \hat{B}_{t_0} / B) = \sum_{B_i \neq \hat{B}_{t_0}} p(\tilde{B}_{t_0} = B_i / B) = \sum_{B_i \neq \hat{B}_{t_0}} \sum_j p(\tilde{B}_j / \hat{B}) p(\tilde{B}_{t_0} = B_i / \tilde{B}_j, B) = \sum_{B_i \neq \hat{B}_{t_0}} \sum_j \sum_k p(\tilde{B}_j / \hat{B}) p(\tilde{B}_{t_0-1} = B_k / \tilde{B}_j, B) p(\tilde{B}_{t_0} = B_i / \tilde{B}_j, \hat{B}, \tilde{B}_{t_0-1} = B_k) \quad (9)$$

其中, \tilde{B}_j 表示在已知隐藏序列条件下, 获知的观察状态序列。对于式(9), $p(\tilde{B}_j / \hat{B})$ 主要由客观物理条件决定, 即误判概率矩阵 W 。 $p(\tilde{B}_{t_0-1} = B_k / \tilde{B}_j, \hat{B})$ 表示观察状态确定时, 在 t_0-1 时刻处于状态 B_k 的概率, 由维特比算法决定。 $p(\tilde{B}_{t_0} = B_i / \tilde{B}_j, \hat{B}, \tilde{B}_{t_0-1} = B_k)$ 表示 t_0-1 时刻处于状态 B_k 、 t_0 时刻估计为状态 B_i 的概率, 与转移概率 a_{ik} 有关。由式(9)可得:

$$P_{e1} = \sum_j p(\tilde{B}_j / \hat{B}) \sum_{B_i \neq \hat{B}_{t_0}} \sum_k \delta_{t_0-1}(k) \times \xi_{i,k} \quad (10)$$

其中, 当 $a_{ik} = \max_m(a_{mk})$ 时, $\xi_{i,k} = 1$, 否则 $\xi_{i,k} = 0$ 。 $\delta_{t_0-1}(k)$ 主要包括 2 个部分, \tilde{B}_{t_0-1} 估计正确部分 ($\tilde{B}_{t_0-1} = \hat{B}_{t_0-1}$) 以及估计错误部分。

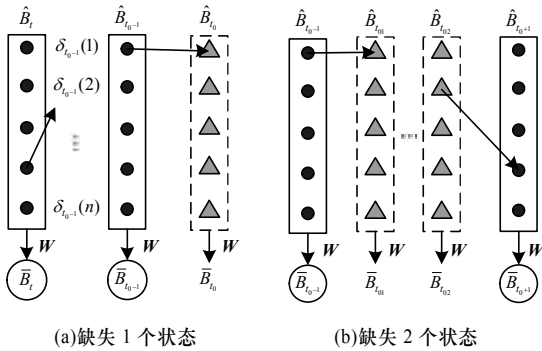


图 6 算法分析

因此, 缺失状态还原正确与否与 t_0-1 时刻终端所处的状态以及该状态的转移概率大小有关。

定义(状态倾向度) 将状态转移概率的大小定义为状态倾向度, 用 $\zeta, 0 \leq \zeta \leq 1$ 表示。 ζ 越大, 转移概率状态倾向性越强, 反之转移概率越均衡, 倾向性越差。

当缺失状态数增多时, 如图 6(b)所示。 t_{01} 、 t_{02} 为 2 个状态缺失时刻, 令 P_{e2}^* 表示 t_{02} 时刻缺失状态还原错误的概率, 状态还原错误率主要包括 2 种情况:

(1) 当两缺失状态相邻时, 由式(10)可得:

$$P_{e2} = P_{e1} + (1 - P_{e1}) P_{e2}^* = P_{e1} + (1 - P_{e1}) \sum_j p(\tilde{B}_j / \hat{B}) \sum_{B_i \neq \hat{B}_{t_{01}}} \sum_k \delta_{t_{01}}(k) \xi_{i,k} \quad (11)$$

(2) 当两缺失状态不相邻时:

$$P_{e2} = P_{e1} + (1 - P_{e1}) P_{e2}^* = P_{e1} + (1 - P_{e1}) \sum_j p(\tilde{B}_j / \hat{B}) \sum_{B_i \neq \hat{B}_{t_{02-1}}} \sum_k \delta_{t_{02-1}}(k) \xi_{i,k} \quad (12)$$

由式(11)和式(12)可知, $P_{e2} \geq P_{e1}$, 即随着缺失点的增多, 缺失状态还原错误率增大。且当两缺失状态相邻时, t_{01} 时刻状态缺失点的还原正确率会对 t_{02} 时刻状态还原正确率直接产生影响。

4.2 仿真验证

基站数量 $n=37$, 首先随机生成若干组状态倾向度不同的转移概率矩阵, 构建概率矩阵 W , 其中, 概率 $w_{B_i, B_j} = \max_j(w_{B_i, B_j}), B_j \in B$; 然后依转移概率分布确定下一时刻状态, 确定隐藏序列, 根据误判概率矩阵 W , 生成与隐藏状态序列相对应的观察状态序列; 采用随机抠点方法, 从观察序列中抠除相应数量的状态点; 最后利用算法对带有缺失状态观察序列进行还原。

在仿真过程中, 当缺失状态数为 1 时, 将序列中除初始状态以外的所有状态依次作为缺失状态。定义还原正确率为:

$$P_r = \frac{\sum_{i=1}^K R_i}{K \times (T-1)} \times 100\% \quad (13)$$

其中, K 为总循环次数; R_i 为第 i 次还原正确的次数; T 为观察序列长度。

取状态倾向度 $\zeta=0.8$, 仿真结果如图 7 所示, $x=K \times (T-1)$, 轨迹还原成功率较为稳定, 约为 90%。实验结果表明, 轨迹还原算法能够很好地还原出终端行为轨迹。

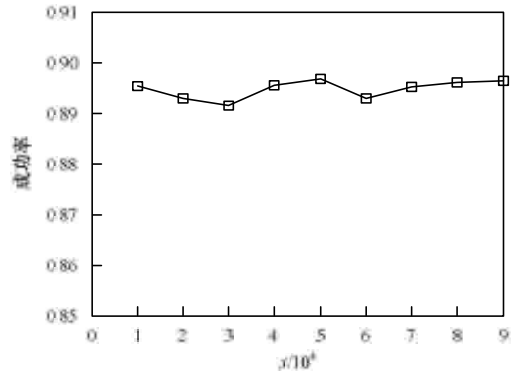


图 7 轨迹还原成功率的变化趋势

图 8 给出了直接利用维特比算法进行轨迹还原与考虑缺失状态后进行轨迹还原的算法(本文算法)性能对比。

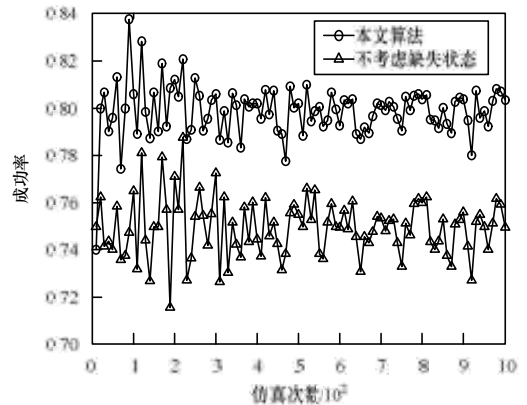


图 8 还原成功率随仿真次数的变化趋势

由图 8 可知, 本文算法的性能要优于不考虑缺失状态时的算法性能, 并且随着仿真次数的增加, 2 种算法的成功率都逐渐趋于稳定。图 9 给出了 2 种算法的成功率随序列长度变化的趋势图, 因为随着序列长度的增加, 缺失状态(此处取缺失状态数为 1)在序列中所占的比重逐渐减小, 所以缺失状态的影响力逐渐减小。当序列长度为 15 时, 2 种算法成功率基本相同。

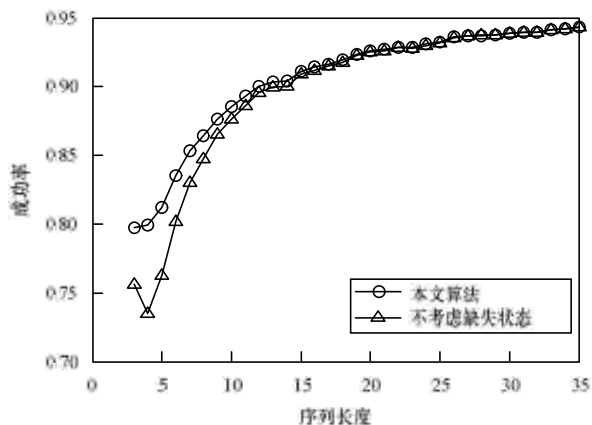


图 9 还原成功率随序列长度的变化趋势

轨迹还原算法有效地利用隐马尔可夫模型, 对终端行为轨迹进行还原, 减小了终端状态的时空不合理性。此外, 由于状态转移概率及观察状态序列由程序依概率随机生成, 观察序列可能会陷入局部循环, 导致最终结果也可能陷入一种局部循环。

图 10 给出了缺失状态数为 1 时轨迹还原成功率随 ζ 的变化趋势。

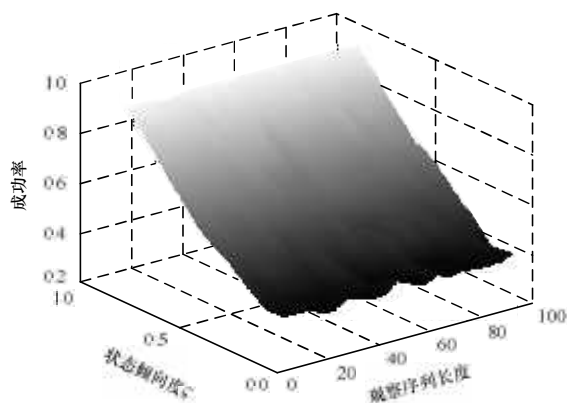


图 10 轨迹还原成功率随状态倾向度的变化趋势

由图 10 可以看出, 随着转移概率矩阵状态倾向度的减小, 行为轨迹还原成功率逐渐降低, 验证了算法性能分析的正确性。

图 11 给出了观察序列长度一定 ($T=20$) 时算法成功率随序列缺失状态数的变化趋势。由图可知, 当观察序列长度一定时, 随着缺失状态数的增多, 行为轨迹还原成功率

逐渐降低。

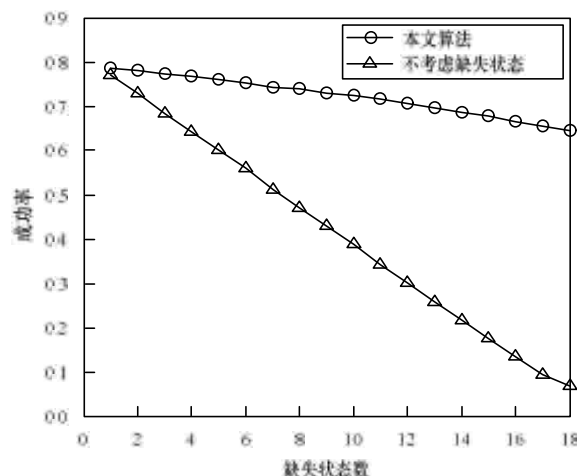


图 11 算法成功率随缺失状态数的变化趋势

5 结束语

本文针对观察序列状态可能存在缺失的情况, 给出一种缺失状态估计方法, 然后提出一种基于 HMM 的行为轨迹还原算法, 通过对仿真图的比较分析表明, 该算法能够较好地适用于观察状态存在缺失的情况, 确保了终端行为轨迹的准确性。

参考文献

- [1] 潘奇明, 程咏梅. 基于隐马尔可夫模型的运动目标轨迹识别[J]. 计算机应用研究, 2008, 25(7): 1989-1995.
- [2] 沈波, 刘云. 基于 HMM 的移动预测模型与仿真[J]. 系统仿真学报, 2007, 19(18): 4118-4121.
- [3] 何彦斌, 杨志义, 马荟, 等. 一种基于 HMM 的场景识别方法[J]. 计算机科学, 2011, 38(4): 254-256.
- [4] Morris A C, Cooke M, Green P. Some Solutions to the Missing Feature Problem in Data Classification[C]//Proc. of ICASSP'98. Seattle, USA: [s. n.], 1998: 737-740.
- [5] Cooke M, Green P, Josifovski L, et al. Robust Automatic Speech Recognition with Missing and Unreliable Acoustic Data[J]. Speech Communication, 2001, 34(3): 267-285.
- [6] Raj B. Reconstruction of Incomplete Spectrograms for Robust Speech Recognition[D]. Pittsburgh, USA: Carnegie Mellon University, 2000.
- [7] 罗宇, 杜利民. 基于隐马尔可夫模型局部最优状态路径的数据重建算法[J]. 电子与信息学报, 2004, 26(5): 722-726.
- [8] 苏腾荣, 吴及, 王作英, 等. 利用空间相关性的改进 HMM 模型[J]. 计算机工程与设计, 2010, 31(5): 1023-1026.
- [9] Rabiner L R, Juang B H. An Introduction to Hidden Markov Models[J]. IEEE ASSP Magazine, 1986, 16(3): 4-16.

编辑 任吉慧