基于移动通信数据的用户移动轨迹预测方法

刘 震[1] 付俊辉[1] 赵 楠[2]

[1](河南科技学院网络信息中心 河南 新乡 453003)

[2](北京大学信息科学与技术学院 北京 100871)

摘 要 随着无线移动通信设备的发展,获取用户位置的手段更加多样,如何对轨迹进行建模并预测人类行为成为研究热点。现有方法主要针对 GPS 轨迹等连续轨迹进行建模预测,而对移动通信场景中人行为轨迹预测方法尚未研究。针对移动话单数据这种离散程度极大的轨迹数据建模,提出 Match 算法对人类轨迹进行预测。实验证明,85% 的人类轨迹可以利用该算法正确预测。在此基础上,提出轨迹合并的方法,进一步提高了预测的准确率,并发现人类在以天为单位的尺度上,有 30% 的行为是自相似的。

关键词 移动数据 轨迹预测 行为分析

中图分类号 TP311 **文献标识码** A DOI: 10.3969/j.issn.1000-386x.2013.02.003

USERS MOBILE TRACK PREDICTION METHOD BASED ON MOBILE COMMUNICATION DATA

Liu Zhen[1] Fu Junhui[1] Zhao Nan[2]

[1](*Network Information Center, Henan Institute of Science and Technology, Xinxiang 453003, Henan, China*)

[2](*School of Electronic Engineering and Computer Science, Peking University, Beijing 100871, China*)

Abstract With the development of wireless mobile communication devices, there are diverse means to obtain users' location, and the ways to model the track as well as to predict the human behaviours become the focus of the research. Existing means are mainly aiming at continuous trajectory like GPS track to model and predict, but for predicting human behaviour tracks in the scene of mobile communication, it is till the blank yet. In this paper, aiming at modelling the mobile calling list data, which is a kind of track data with very large discrete degree, we propose Match algorithm to predict human tracks. Experiment proves that 85% human tracks can be correctly predicted with this algorithm. On this basis, we then propose a method of tracks merging, which further improves the accuracy of prediction. Moreover, it is found that 30% of the human behaviours are self-similar taking the day as the scale of human beings.

Keywords Mobile data Trajectory predicting Behaviour analysis

0 引 言

随着移动电话的普及以及 3G 网络的部署,移动网络中基于位置信息的服务(LBS)越来越成为热点。如日本的 NTT DoCoMo 推出个人用户定位和导航服务,美国的 Verizon 主推 Family Locator 服务以及欧洲的 Vodafone 等推出了团队位置跟踪服务等;还如 MIT 的 Mobile Landscapes 项目使用移动通话数据得到静态城市人群时空分布,Real Time Rome 项目的实时监测人群密度动态分布等。基于移动位置信息的研究和应用正成为全球学术界和产业界共同关注的焦点问题。

当前,与移动位置相关的行为研究的热点问题之一是对移动用户未来的位置信息进行实时、准确的预测[1-6]。然而,当前对用户未来位置的预测研究主要集中在对连续轨迹(典型代表是 GPS 轨迹)建模的基础上,如文献[1-3]。然而,此类研究难以预测移动通信场景下的用户轨迹。因为在该场景下,用户轨迹十分离散,一天平均只能获得十几个用户轨迹点,现有模型难以根据如此离散的轨迹反演得到用户的历史行为,并建立起有效的预测模型。所以,移动通信场景下的移动轨迹预测成为一个尚未解决的难题。

本文提出了一个新的对离散轨迹进行建模预测的方法。首先,将话单数据所反映的离散轨迹转换为连续轨迹得到用户行为模式,基于该行为模式提出了轨迹预测算法 Match,实验证明,使用该算法有 85% 的人类行为可以预测。通过对历史相似轨迹进行合并,更加准确地刻画了用户的真实轨迹,进一步提高了预测准确率。并得出结论:在以天为单位的尺度上,人类的行为有 30% 是自相似的。

1 相关工作

由于使用的定位手段不同,对人类轨迹的建模方式以及预测方法也有不同。GPS 定位精度最高,通过内置的客户端,可以获得用户的连续轨迹,基于此种数据的研究也最多。从预测方

收稿日期:2012-05-17。国家自然科学基金项目(60703066)。刘震,实验师,主研领域:网络通信与管理。付俊辉,讲师。赵楠,硕士生。

式上,主要有两种方式: 基于 Dynamic Bayesian Network 等网络学习模型对轨迹建模并预测,如文献[1-3]等; 另外,基于历史轨迹频繁项的轨迹预测研究也比较多,其中以文献[4]为代表。作者假设,由于城市中公共交通发达,多数人的出行方式应该相同。将多数人运动的轨迹重合部分,作为对轨迹预测的依据。但是,个人 GPS 轨迹数据难以得到,一般由志愿者提供,样本数量很少,难以大规模开展实验以验证算法可靠性以及研究人类的行为规律。

利用 WiFi 可将用户定位到 WiFi 基站,基站覆盖范围较小,使得定位精度较高。由于 WiFi 网络并不普及,研究大多局限在一个校园或一个社区之内[5,6],代表性并不强,这里不再赘述。

利用移动基站定位用户精度最低,城市中基站的覆盖范围为 0.5km ~5km 不等,其定位精度也在这个范围。然而,由于手机的普及特性,因此研究基于该数据特性的轨迹预测将为大量基于位置信息的应用提供支持。实际中得到的用户话单数据,记录了一个用户在一段时间的通话行为,用户位置通过话单数据中通讯时所联络的基站反映出来。但用户通话行为极其离散,形成了离散的轨迹,如何对其建模成为难题。另外,由于涉及用户隐私,此种研究往往难以开展。所以在这个领域,开展的研究还较少。文献[8]通过在手机安装客户端,记录下用户附近的基站信息,得到连续的轨迹。利用 DBN,对用户的轨迹进行预测。该研究使用了志愿者提供的信息,实际上是低精度类似 GPS 定位数据的连续轨迹预测,有着与 GPS 定位进行轨迹预测相同的缺陷。

Barabasi 等在 Nature[9] 上报到了他们开创性的研究工作。通过与移动公司合作,作者得到了三个月 10,000,000 用户的话单数据。通过系统的分析论证,得到了任何对于人类行为进行预测的准确率不会超过 93% 的结论,但准确率也不会低于 70%(在后文中称该算法为 Most Frequent 算法)。但是,作者并未在文中提出其他具体的轨迹预测方法,仅仅对在移动通信环境下人类行为的可预测性做出上下界的定义。本文是在该文的基础上开展工作的。

2　问题定义

假设 $SetT_{i,j}$ 表示移动用户 i 在第 j 天所有开始通话时刻的集合。$SetTower_i$ 是用户 i 历史通话中出现的全体基站。用户 i 的一次通话记作 $<t_p, Tower_p(x_p, y_p)>$,其中 t_p 是通话开始时刻,$t_p \in SetT_{i,j}$,(x_p, y_p) 是通话时通信基站经纬度,$Tower_p \in SetTower_i$。

用户 i 在第 j 天的通话行为记为话单数据集 $Call_{i,j}$,表示为: $Call_{i,j} = \{<t_1, Tower_1>, <t_2, Tower_2>, \cdots, <t_n, Tower_n>\}$。$Call_{i,j}$ 是一个有序集,$t_1 < t_2 < \cdots < t_n$。

定义 1　轨迹　设 r 是集合 T_i 的一个划分,$r_k = \{Tower_{k1}, Tower_{k2}, \cdots\}$,那么 $Call_{i,j}$ 可重写为 $Trajectory_{i,j} = \{<t_1, r_1>, <t_2, r_2>, \cdots <t_n, r_m>\}$,称作用户 i 在第 j 天的轨迹。

定义 2　行为模式　定义点对 $<t_k, t_{k+1}, r_k>$ 表示用户 i 在时间段 $[t_k, t_{k+1}]$ 处于 r_i 区域,有序集 $Pattern_{i,j} = \{<t_1, t_2, r_1>, <t_3, t_4, r_2>, \cdots, <t_{2n-1}, t_{2n}, r_n>\}$ 称作用户 i 在第 j 天的行为模式。

定义 3　轨迹预测　已知序列 $\{<t_1, r_1>, <t_2, r_2>, \cdots, <t_{n-1}, r_{n-1}>\}$ $(t_1 < t_2 < \cdots < t_{n-1})$,待预测时刻 t_n $(t_{n-1} < t_n)$ 以及用户在 t_n 时刻的真实位置 r_n。使用某种预测算法得到用户在 t_n 所处区域为 r',若 $r' = r_n$,则称预测正确; 若 $r' \neq r_n$,则称预测错误。

定义 4　轨迹预测策略　已知序列 $\{<t_1, r_1>, <t_2, r_2>, \cdots, <t_n r_n>\}$ $(t_1 < t_2 < \cdots < t_n)$ 以及待预测时刻 t',满足 $|t' - t_i| \leqslant \forall k(t' - t_k)$,$t_i \subset SetT_i$,那么预测序列 $\{<t_1, r_1>, <t_2, r_2>, \cdots, <t_{i-1} r_{i-1}>\}$ 在 t_i 的位置,若使用某种预测算法得到的 $r' = r_i$,则称预测正确; 若 $r' \neq r_i$,则称预测错误。

3　轨迹预测方法

手机在进行通讯时,有一定概率发生基站转换,使得话单记录的基站位置并不能很好地反映用户的真实位置。首先对话单数据进行修正,去除通话中发生的基站转换。然后在区域划分的基础上得到用户行为模式,将用户离散的轨迹转换为连续的行为模式。基于该模式,使用 Match 算法对用户轨迹进行预测。在 3.5 节还提出了轨迹合并算法,轨迹合并集较之行为模式更加准确地反映了用户历史真实轨迹,利用轨迹合并集,进一步提高轨迹预测的准确率。

3.1　基站转换

手机在进行通信时,由于信号、基站容量等客观问题,并不总是选择距离最近的基站,使得话单记录的基站位置并不能精确地反映用户的位置,因而有必要去除通话时发生的基站转换。

用户两次通话为 $<t_i, Tower_i(x_i, y_i)>$,$<t_j, Tower_j(x_j, y_j)>$ $(Tower_j \neq Tower_j, t_i < t_j)$,在这两个基站间最小的移动速度为:

$$v_{i,j} = \frac{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}{t_j - t_i}$$

设人类移动速度的上限是 $v_{th} = 120\text{km/h}$,任何基站间最小移动速度高过该上限的都视作发生了基站转换,找出话单数据中发生基站转换的记录进行修正。

3.2　区域划分

用户的一个活动区域往往包括几个基站,将该活动区域中的基站合并就可以得到用户的一个区域。

由于基站的特殊性,我们并没有选择常规的聚类方法。当两个活动基站的信号范围发生重叠时(这里使用基站距离小于 1.5km 作为标准),这两个基站属于同一个区域(算法 1)。

算法 1　Region Division

```
Input: Tower_i
Output: r
1)  for each i,j
2)      if distance(Tower_i, Tower_j) < 1.5km
3)          Tower_i ∈ r_k
4)          r_k = r_k ∪ Tower_j
5)      end
6)  end
```

3.3　用户行为模式

由话单数据得到的用户历史轨迹十分离散,难以通过其对轨迹进行预测。本文假设,如果用户连续的两个通话区域相同,那么用户在这个时间段内并未移动。根据此假设,离散的用户轨迹可以转换成连续的行为模式。

给定轨迹 $Trajectory_{i,j}$,用户 i 在第 j 天的行为模式按照算法 2 得到。

算法 2　Pattern Extraction

```
Input: Trajectory_{i,j}
```

Output: $Pattern_{i,j}$

1) for each k do
2) if $r_k = r_k + 1$
3) if S is NULL
4) $S = <t_k, t_{k+1}, r_k>$
5) else
6) $S = <S(1), t_{k+1}, S(3)>$
7) end
8) else
9) $Pattern_{i,j} = Pattern_{i,j} \cup S$
10) S = NULL
11) end
12) end

3.4 轨迹预测算法

Barabasi 在文献[9]中指出,人类的行为有其潜在的规律性。这个规律体现在人类的后一时刻的行为常常取决于前一时刻或前几个时刻的行为。可以结合用户历史通过前几次行为来预测用户的下一次行为。

例 1 已知 t_3 时刻之前的轨迹序列为{ $<t_1, r_1>$, $<t_2, r_2>$}。用户历史轨迹 $Trajectory_i$ 中包含该序列的最频繁子序列为{ $<t_1, r_1>$, $<t_2, r_2>$, $<t_3, r_3>$},那么用 $<t_3, r_3>$ 作为预测结果,即预测用户在 $t3$ 时刻位于 r_3。

人类行为的规律性还体现在人类的作息以天为单位进行。在一天中的某时刻或时段,人的行为总是相似的。比如,一个人在夜晚常常在家中睡觉,白天在单位上班。多数人的时间还分为工作日和周末,在工作日,人们主要从事劳动;在周末,人们主要休息娱乐。

例 2 对于 $Pattern_i$, $\exists j,k \in$ {工作日}, $\exists p \in$ {周末},使得 $(Pattern_{i,j} \cap Pattern_{i,p}) \subset (Pattern_{i,j} \cap Pattern_{i,k})$

根据上述规律和假设,提出算法 3(称之为 Match 算法)来对用户轨迹进行预测。

算法 3 Match

Input: $<t_1, r_1>, <t_2, r_2>, \cdots, <t_{n-1}, rm-1>, (t_1 < t_2 < \cdots < t_{n-1}), t_n$,
Patterni Output: r'

1) Define t_k in $Pattern_{ij}$ to find a $<t_{2m-1}, t_{2m}, r_m>$ where $t_{2m-1} \leqslant t_k \leqslant t_{2m}$
2) satisfy = { }
3) for each j do
4) if t_n in $Pattern_{i,j}$
5) satisfy = $satisf_y \cup Pattern_{i,j}$
6) end
7) end
8) for each j do
9) find max_i that t_{n-1} in $satisfy_j$, t_{n-2} in $satisfy_j$, $\cdots$, t_{n-i} in $satisfy_j$
10) $match_j = n - i$
11) end
12) p = max(match)
13) if $p \neq 0$
14) $r' = r_q$, that exists q with $<t_{2q-1}, t_{2q}, r_q> \in satisfy_p$ and
15) $t_{2q-1} \leqslant t_n \leqslant t_2$
16) else
17) user Algorithm Most Frequent get r'
18) end

3.5 轨迹合并

用户一天的通话行为十分离散,使得话单数据很难准确反映他的真实轨迹。选取用户历史轨迹中最为相似的两条记录进行合并,则有可能更加真实地反映用户的轨迹。

例 3 用户 i 在 j, k 两天的真实轨迹均为{ $<t_1, r_1>$, $<t_2, r_2>$, $<t_3, r_3>$, $<t_4, r_4>$},表示为: $Trajectory_{i,j}$ = { $<t_1, r_1>$, $<t_2, r_2>$, $<t_3, r_3>$}, $Trajectory_{i,k}$ = { $<t_1, r_1>$, $<t_3, r_3>$, $<t_4, r_4>$},则合并 j, k 这两天的轨迹得到了真实的轨迹: $Trajectory_{i,j} \cup Trajectory_{i,k}$ = { $<t_1, r_1>$, $<t_2, r_2>$, $<t_3, r_3>$, $<t_4, r_4>$ }。

用户两天轨迹的相似程度由轨迹差来度量。

定理 1 三维空间 (x, y, t) 中,由 $Trajectory_{i,j}$ 中序列点对依次连结形成的连续曲线记作 $(Line_j^x(t), Line_j^y(t), t)$。$d1$、$d2$ 两天的轨迹差为:

$$Dis_i(m,n) = \frac{\int_{t_{start}}^{t_{end}} \sqrt{(Line_{d1}^x(t) - Line_{d2}^x(t))^2 + (Line_{d1}^y(t) - Line_{d2}^y(t))^2}}{t_{start} - t_{end}}$$

其中 (t_{start}, t_{end}) 是曲线 $(Line_{d1}^x(t), Line_{d1}^y(t), t)$ 与曲线 $(Line_{d2}^x(t), Line_{d2}^y(t), t)$ 定义域的交集。

证明: 对于用户 i,在同一时刻 t,第 m 天的位置为 (x_m, y_m),第 n 天的位置为 (x_n, y_n),这两天在时刻 t 的距离为:

$$L = \sqrt{(x_m - x_n)^2 + (y_m - y_n)^2}$$

两天平均意义上的距离为:

$$Dis_i(m,n) = \frac{\int_{t_{start}}^{t_{end}} L}{t_{start} - t_{end}}$$

其中 m、n 所对应轨迹曲线为 $(f_m^x(t), f_m^y(t), t)$ 、$(f_n^x(t), f_n^y(t), t)$。则上式可写成:

$$Dis_i(m,n) = \frac{\int_{t_{start}}^{t_{end}} \sqrt{(f_m^x(t) - f_m^x(t))^2 + (f_m^y(t) - f_m^y(t))^2}}{t_{start} - t_{end}}$$

如果 f 取由 $Trajectory_{i,j}$ 中序列点对依次连结形成的连续曲线 $(Line_j^x(t), Line_j^y(t), t)$,那么上式可以写成:

$$Dis_i(m,n) = \frac{\int_{t_{start}}^{t_{end}} \sqrt{(Line_{d1}^x(t) - Line_{d2}^x(t))^2 + (Line_{d1}^y(t) - Line_{d2}^y(t))^2}}{t_{start} - t_{end}}$$

证毕。

在确定两天轨迹的相似程度度量方法后,计算 $Dis_i(p,q)$,选取轨迹差最小的前 $th\%$ 的用户轨迹进行合并,得到用户轨迹合并集,称之为 $Merge_i$。具体轨迹合并算法见算法 4。利用 $Merge_i$ 代替 $Pattern_i$ 进行轨迹预测,实验证明,能够提高预测准确率。

算法 4 Combination

Input: $Trajectory_i$, n, th
Output: $Merge_i$

1) //n 为历史话单数据的总天数,th 为合并记录比例
2) for every p do
3) for every q do
4) if p = q
5) $Dis_{p,q}$ = MAXNUM
6) else
7) $dis_{p,q} = Dis_i(p,q)$
8) end
9) end
10) end
11) for j = 1to th* C_n^2
12) $Dis_{p,q}$ is the j_{th} smallest num in dis
13) $Merge_i = Merge_i \cup \{ Trajectory_{i,p} \cup Trajectory_{i,q} \}$
14) end
15) $Merge_i = Merge_i \cup Trajectory_i$

4　实　验

4.1　实验数据

本文选用中国某地级市 2007 年 9 月 14 天 1240000 用户的话单数据。选择 $f \geqslant 0.5$ 的用户数据进行实验，满足该条件的用户有 15000。

4.2　轨迹预测评估

为了计算预测准确率，我们使用前 13 天的数据作为话单数据集 Call，应用 3.2 节区域划分算法对用户的活动范围 T_i 进行划分，最后使用轨迹预测策略对最后一天每一个整点时刻进行预测，并计算准确率。结果如图 1 所示。

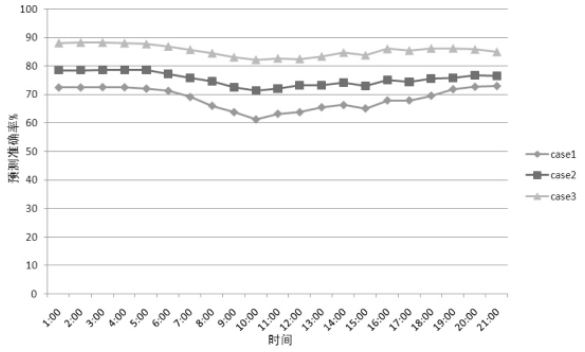

图 1　不同方法的预测准确率

具体解释如下:

1) Case1: 使用 Most Frequent 算法。将最常访问的区域作为轨迹预测结果，对数据不做额外处理。平均预测准确率为 68.5%。

2) Case2: 使用 Most Frequent 算法。使用 3.1 节所述算法对话单数据集 Call 进行修正。平均预测准确率为 75.4%。

3) Case3: 使用本文提出的 Match 算法。使用 3.1 节所述算法对话单数据集 *Call* 进行修正。平均预测准确率为 85.4%。

由图 1 可以得到以下结论:

1) 对数据进行去基站转换处理很有必要。Case1 实验重复了文献 [9] 中所述算法 Most Frequent，验证了其预测准确率在 70% 左右的结论。Case2 实验对 Case1 实验数据额外进行了去基站转换处理，比较 Case1 与 Case2，平均预测准确率提高了 6.9%。在具体实验中，发现检测出发生基站转换的通话记录只占总记录数的 1.9%。说明虽然发生基站转换的比例并不高，但基站转换错误地暴露了用户位置，去除基站转换，能够很大程度上提高预测准确率。

2) 比较 Case2 与 Case3，发现本文提出的 Match 方法较之前 Most Frequent 的方法预测准确率提高了 10%。本文提出的方法不仅考虑了用户前几个时刻表现出来的行为对之后的影响，还考虑了以天为单位，用户行为的规律性。使得在轨迹预测时，能够找到最为相似的历史中一天，给出合理的预测结果。

3) 人类在休息时间(一天的较早时间和较晚时间)的规律性要大于工作时间的规律性。对于休息时间的轨迹预测准确率要高于工作时间的轨迹预测准确率。尤其是在正午时间，三条预测曲线都达到了最小值，说明在正午时间人类行为的规律性最差，最为难以预测。

对 Match 算法错误预测用户轨迹的来源进行分析，发现主要有三个因素影响了预测的结果:

1) 人类行为不可知性。在文献 [9] 中，作者详细论述了该问题，人类的行为虽然有着很强的规律性，但大约有 7% 的人类行为是突发的、不可预知的。

2) 基站转换。每次通话时都会有一定概率发生基站转换，而不是与最近的基站进行通讯，使得部分验证数据错误反映用户位置，影响预测准确率。

3) 行为模式并不能完整的反映用户的历史行为。用户的通话极为离散，即使是选择通话行为较频繁的用户，其一个小时内所打出的电话也不过 1、2 个，也就是说，一个小时中，只能够得到 1、2 次用户的位置。用户的轨迹很可能并未完全暴露，得到的行为模式只反映了用户部分历史行为。

4.3　轨迹合并预测评估

针对 4.2 提出的第 3 个影响预测结果的因素，我们进行了轨迹合并实验，希望使用用户轨迹合并集提高预测的准确率。

首先，使用前 13 天的数据作为话单数据集 *Call*，并对数据做去基站转换修正，应用 3.2 节区域划分算法对用户的活动范围进行划分。

接下来的步骤与 3.2 节轨迹预测实验方法略有不同。使用算法 3 产生用户轨迹合并集 *Merge*，用 $Merge_i$ 代替 $Trajectory_i$ 产生行为模式 $Pattern_i$，并计算预测准确率。结果见图 2。

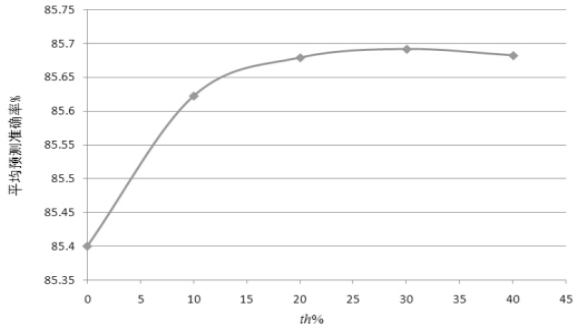

图 2　经过轨迹合并后预测准确率

由图 2 可知，随着轨迹合并比例 *th* 的增大，平均预测准确率不断上升，在 $th = 30\%$ 左右时，预测准确率达到峰值，*th* 继续增大，预测准确率开始下降。

这是因为当 *th* 很小时，只合并轨迹差异最小、最相关的记录，合并后的轨迹能够正确反映用户的真实轨迹，能够增大正确预测轨迹的比率。当 *th* 过大时，一些不相关的轨迹记录也被合并，合并后的轨迹不能反映用户的真实轨迹，减小了正确预测轨迹的比率。

图 2 还可以得出结论: 人类在以天为单位的时间尺度上，有 30% 的行为是自相似的。

5　结　语

已有研究大多针对于 GPS 等连续轨迹数据，无法预测离散轨迹。本文对离散轨迹(话单数据)建模得到用户行为模式，使用本文提出的 Match 算法，约有 85% 的人类轨迹可以预测，接着使用轨迹合并算法，预测准确率得到进一步地提高。另外，在对轨迹合并进行研究的基础上，我们得到了人类在以天为单位的时间尺度上，有 30% 的行为是自相似的结论。

基于移动数据的轨迹预测，还应有提高预测准确率的空间，结合社会网络的轨迹预测将是我们进一步研究的方向。

(下转第 17 页)

② 将时间 T 的分钟值与 10 做取模运算，记为 $t = T_{分} \bmod 10$ ($t = 1,2,\cdots,9$)，并将此时的时间记为 T [i]。截短口令初值每次选取三组 M [i]，分别为 M [i]、M [i + 10]、M [i + 20]。

③ 将三个数组顺序组合为 3 个字节的十六进制数，即 OTP = M [i] M [i + 10] M [i + 20]。

④ 将 OTP 转化为十进制数 OTP′。

⑤ 将 OTP′与 10^6 取模，得到截短的六位数字口令 P。

举例说明，如果当前时间是 14 点 53 分，则 t = 53mod10 = 3，此时时间记为 T [3]，明文经过加密，输出值为 M = debe9ff9 2275b8a1 38604889 c18e5a4d 6fdb70e5 387e5765 293dcba3 9c0c5732，取 M [3] = f9、M [13] = 8e、M [23] = 65 顺序组合生成 OTP = f98e65，转换为十进制数 OTP′ = 16354917。则最终生成的动态口令 $P = 16354917 \bmod 10^6 = 354917$。此种方法保证了生成口令的二次动态性，使得非法用户通过猜测的方法获得密钥几乎是不可能的，进一步提高了动态口令算法生成的安全性。

3　实验验证

本文设计和实现的基于时间的动态口令卡，以 SM3 算法作为加密算法，在 DEMO 板上实现了每分钟生成一次动态口令，以 6 位数字显示在液晶屏上。在 keil C 软件模拟环境下，口令每分钟更新一次，测得单次口令的平均计算时间和时间刷新时间均为 200ms 左右。动态口令系统在不同工作模式下的功耗如表 1 所示。

表 1　系统不同状态功耗值

系统状态	MCU 状态	LCD 状态	功耗
待机状态	打开	打开	1.1uA
		关闭	0.9uA
时间刷新	在 32.768Khz 下工作	打开	4.5uA
		关闭	4uA
口令更新	在 2Mhz 下工作	打开	240uA

系统整体功耗 = 待机时间功耗 + 时间刷新时的功耗 + 口令更新时的功耗。计算每分钟内的系统平均功耗: 设定时间计时 500ms 处理一次，则每分钟可以处理 120 次。口令每分钟更新一次，对照表 1 系统功耗在不同工作模式下的数据，可计算出动态口令卡系统的平均功耗为:

$$(1.1 \times 60 + 4.5 \times 120 \times 0.2 + 240 \times 0.2) / 60 = 3.7\text{uA}$$

口令卡使用时间 = 电池容量 × 放电系数/平均功耗。放电系数是考虑到动态口令卡在实际应用中，由于外界因素和自身性能的影响，有一定的电量损耗。通常放电系数按照为 0.8 计算，即电池能够有效使用的电量为总电量的 80%。钮扣电池的容量为 210mAh，则理论使用年限为:

$$210 \times 0.8 \times 1000 / 3.7 = 45405 \text{ 小时} \approx 5.18 \text{ 年}$$

国内市场的现有的动态口令卡，常显式产品的使用寿命大都为 3 年左右，本文设计的口令卡将使用时间延长了近 70%，达到了降低产品功耗，延长使用寿命的设计要求。

4　结　语

本文采用基于时间同步的认证机制，设计和实现了一种基于 SM3 算法的动态口令卡，是国有密码算法在身份认证领域的具体应用。本文提出了基于时间的截短口令算法，实现了截短口令的二次动态性，进一步提高了动态口令算法的安全性。此外，口令卡采用低功耗芯片设计，通过测试和仿真实验表明，达到了降低系统功耗、延长了使用寿命的设计目标。本文设计实现的动态口令卡，为 SM3 算法的现实应用提供了一个可行的选择。

参　考　文　献

[1] 张文. 动态口令系统的设计与实现 [J]. 微计算机信息，2005，21(3): 236-237.

[2] 顾韵华，刘素英. 动态口令身份认证机制及其安全性研究 [J]. 微计算机信息，2007，23(33): 57-59.

[3] 叶晰，叶依如. 基于 MD5 算法的动态口令技术的软件实现 [J]. 计算机应用与软件，2009，29(11): 287-288.

[4] 王小云，冯登国，来学嘉，等. 散列函数中的碰撞 [EB/OL]. http://www.cstc.net.cn/docs/docs.php? id=323.

[5] 中华人民共和国国务院令第 273 号. 商用密码管理条例 [S/OL]. http://topic.csdn.net/t/20011229/11/448986.html.

[6] 李晓东，贾慧斌. 基于时间同步的动态口令认证系统 [J]. 信息网络安全，2010，31(5): 71，77.

[7] 国家密码管理局. SM3 密码杂凑算法 [EB/OL]. 2010-12-22. http://www.oscca.gov.cn/UpFile/2010122214185778866.pdf.

[8] 九段. 评价微处理器低功耗的五个标准 [EB/OL]. 2008-09-24. http://www.eetrend.com/forum/100023574.

[9] 高小霞. 基于哈希算法的动态口令令牌的分析设计与改进 [EB/OL]. http://wenku.baidu.com/view/23d30943a8956bec0975e32e.html.

(上接第 13 页)

参　考　文　献

[1] Han Sangjun, Cho Sungbae. Predicting user's movement with a combination of self-organizing map and markov model [C]//ICANN 2006, 883-893.

[2] Daniel Ashbrook, Thad Starner. Using GPS to learn significant locations and predict movement across multiple users [J]. Personal and Ubiquitous Computing, 2003, 7: 275-286.

[3] Jan Petzold, Faruk Bagci, Wolfgang Trumler, et al. Global and local state context prediction [C]//Artificial Intelligence in Mobile Systems 2003 (AIMS 2003) in Conjunction with the Fifth International Conference on Ubiquitous Computing, 2003.

[4] Anna Monreale, Fabio Pinelli, Roberto Trasarti, et al. WhereNext: a location predictor on trajectory pattern mining [C]//KDD'09, 2009: 637-645.

[5] Hsu Weijen, Ahmed Helmy. IMPACT: investigation of mobile-user patterns across university campus using WLAN trace analysis [R]. Technical report, University of Souch California, 2005.

[6] Faruk Bagci, Florian Kluge, Theo Ungerer, et al. Optimisations for LocSens - an indoor location tracking system using wireless sensors [J]. International Jounal of Sensor Network, 2009, 6: 157-166.

[7] http://www.miit.gov.cn/.

[8] Nathan Eagle, Aaron Clauset, John A Quinn. Location segmentation, inference and prediction for anticipatory computing [C]//AAAI 2009.

[9] Song Chaoming, Qu Zehui, Nicholas Blumn. Limits of predictability in human mobility [J]. Science, 2010, 327: 1018-1021.