

## 物联网中适用于内容搜索的实体状态匹配预测方法

张普宁\* 刘元安 吴帆 唐碧华 吴超  
(北京邮电大学智慧无线移动信息技术中心 北京 100876)

**摘要:** 对实体匹配用户内容搜索的状态进行预测可显著提高物联网搜索的效率,降低搜索过程的通信开销。该文提出等时距与周期内实体状态预测方法,估计实体在用户查询时刻的状态;设计了适用于内容搜索的有序验证方法,依据实体匹配用户查询内容的概率对实体进行排序验证,以保证用户搜索结果的可靠性。结果表明,所提实体状态预测方法具有较高的精度,结合所提预测方法与匹配验证方法的搜索机制具有较低的通信开销。

**关键词:** 物联网; 内容搜索; 状态预测; 匹配验证

中图分类号: TP393

文献标识码: A

文章编号: 1009-5896(2015)12-2815-06

DOI: 10.11999/JEIT150191

## An Entity State Matching Prediction Method for Content-based Search in the Internet of Things

Zhang Pu-ning Liu Yuan-an Wu Fan Tang Bi-hua Wu Chao  
(Institute of Smart Wireless Mobile Information Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China)

**Abstract:** Matching prediction with high accuracy of entity state can significantly improve the efficiency of content-based search in the Internet of Things and reduce communication overhead while searching. The equal-interval and during the period entity state prediction methods are proposed, which are applied to the estimation of the entity state at the moment of querying. Moreover, the ordered verification approach is designed to verify the entities in sequence based on the degree of compliance with the searching content, for the sake of enhancing the reliability of searching results. Numerical results show that the proposed entity state prediction approaches can achieve high accuracy, which combines with the ordered verification approach to dramatically improve the performance of communication overhead during the searching process.

**Key words:** Internet of Things; Content-based search; State prediction; Matching verification

### 1 引言

面对物联网中海量、多样化的物理实体,如何快速、准确地搜索用户需要的物理实体是物联网搜索服务需解决的根本问题<sup>[1]</sup>。传统的搜索引擎主要用于搜索网页及pdf, doc文件等静态或缓变的非结构化内容,而物联网中数据具有较强的时空特性,其建立的索引为爬取时刻的内容<sup>[2]</sup>,无法即时呈现物理实体的实时状态。因此,如何设计高效的物联网实体搜索机制是物联网中的关键性问题。

目前,国内外针对物联网实体搜索的研究刚刚起步。文献[3, 4]对物联网搜索服务的内容、特点和典型系统以及物联网搜索技术面临的挑战和发展趋势进行了概括。为避免物联网搜索过程中信息资源泛滥,降低搜索过程中的通信开销,文献[5]设计了Dyser搜索引擎,将传感器与物理实体的状态信息均抽象为Web网页,并提出可估计实体周期性状态的预测模型。通过计算实体页面与搜索内容匹配的概率,获得可能匹配搜索请求的物理实体的集合。然而,文中基于传感器节点可获取实体的周期性高级状态的假设。因此,上述搜索机制对采用低成本传感器且实体状态转换周期性不显著的应用场景适用性较差。

针对上述问题,文献[6]提出支持基于内容的搜索机制(CSS),利用传感器的原始输出时间序列历史数据构建实体状态预测模型,估计实体与查询内容匹配的概率,以搜索与用户需求匹配概率较高的传感器。然而,CSS的实体状态估计方法具有时间独立性,忽略了物联网中实体状态随时间动态变化的

收稿日期: 2015-02-02; 改回日期: 2015-08-19; 网络出版: 2015-11-01

\*通信作者: 张普宁 zhangpuningpaper@gmail.com

基金项目: 国家自然科学基金(61170275), 国家 863 计划项目, 民用航天十二五预研科技项目, 北京市高等学校青年英才计划项目和安全生产智能监控北京市重点实验室基金

Foundation Items: The National Natural Science Foundation of China (61170275); The National 863 Program of China; Civil Space "12th Five-year" Plan Pre-Research Project of Science and Technology; Beijing Higher Education Young Elite Teacher Project; Beijing Key Laboratory of Work Safety Intelligent Monitoring

特性, 未准确感知实体状态的动态变化趋势, 造成对实体状态的估计精度较差, 从而给物联网的实体搜索过程带来较多的通信开销。

为解决上述问题, 本文提出适用于物联网内容搜索的实体状态匹配预测(Entity State Matching Prediction, ESMP)方法, 通过传感器采集的实体状态历史记录构建实体状态预测模型, 预测实体在用户查询时刻的状态, 并结合用户提交的查询内容, 提出实体状态匹配估计方法, 估计实体与用户查询内容的匹配概率, 进而, 依据匹配概率进行实体状态的有序匹配验证, 以降低搜索过程中的通信开销。

## 2 物联网内容搜索服务架构

本文所提 ESMP 搜索服务架构如图 1 所示。用户通过客户端发出内容搜索请求; 物联网网关负责搜索请求的下发及传感器上报数据的收集; 预测模型数据库用于存储传感器上报的历史数据与预测模型; 传感器用于采集实体状态数据, 实体为最终的搜索对象。

物联网搜索系统中有多个物联网网关分布在 Internet 中, 每个网关负责一个或多个传感器网络, 并与该传感器网络的预测模型数据库关联。物联网网关通过周期性地爬取传感器发布的 web 页面来获取传感器当前的读数, 并将读数存入预测模型数据库中用来构建实体状态的预测模型。

客户端通过 GUI 界面发起搜索请求, 用户请求首先由上层网关进行处理, 按照地理位置等具有区分度的信息进行分类, 并下发到下层网关进行内容查询的搜索。由于物联网中实体的状态具有时变性,

用户查询具体的状态值的实用价值较小, 因此, 物联网中搜索的形式如:  $T = (a, b, t_q, R_q)$ , 表示搜索在  $t_q$  时刻温度在  $a^\circ\text{C} \sim b^\circ\text{C}$  之间的实体,  $R_q$  为用户定义的需返回的匹配实体个数。

下层网关在接收到上层网关下发的搜索命令后, 将访问本地的预测模型数据库, 搜索匹配概率较高的传感器, 并将实体匹配排序列表及其匹配概率发送给上层网关, 上层网关将接收到的排序列表进行汇聚, 并依据匹配概率进行排名, 进而, 对排名靠前的实体进行逐一验证, 以确定实体是否真正满足用户的查询需求。最后, 网关将经过验证的搜索结果返回给用户。

## 3 实体状态预测方法

为保证搜索结果的时效性与可靠性, 使得搜索过程快速收敛, 本节通过利用传感器周期性上报的实体状态信息, 建立可预测实体在下一个周期内任意时刻状态的预测模型, 在用户发起基于内容的搜索请求后快速决策用户所需实体。

### 3.1 实体等时距状态预测方法

假设传感器周期性采集并上报实体的状态至汇聚网关。定义  $S = \{x(t_1), x(t_2), \dots, x(t_n)\}$  为实体状态的有序时间序列集合, 其中  $x(t_i)$  为第  $t_i$  个周期实体的状态值,  $t_1, t_2, \dots, t_n$  为传感器采集实体状态数据的周期。则  $S$  为等时间距离的序列, 时间距离  $\Delta t = t_{i+1} - t_i = \omega$  为传感器上报数据的周期。可见  $S$  中隐含了实体状态的时间相关性信息。为充分利用实体状态历史信息, 感知等时间距离下实体状态的变化趋势, 需首先对  $S$  进行矩阵变换。

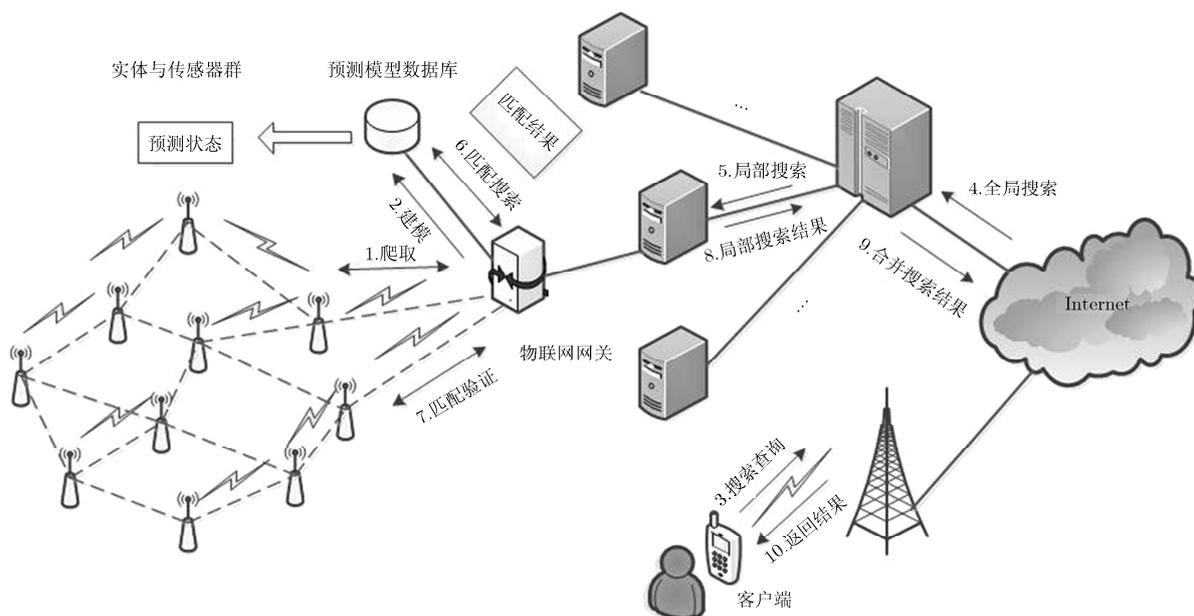


图 1 ESMP 搜索架构

(1) 实体状态时间序列矩阵变换：为构建预测模型，需对传感器采集的时间序列进行重构，将实体状态时间序列转化为可最大化时间序列相关性的矩阵形式。通过滑动时间窗口  $\chi_t = [x(t), \dots, x(t+m-1)]$ ,  $t = 1, 2, \dots, n_t - m$  控制训练样本的输入。其中  $x(t+i)$  为第  $(t+i)$  个周期传感器采集的实体状态值， $n_t$  为训练集的大小。 $y_t = x(t+m)$  表示实体在第  $t$  个周期的实际状态值，也为预测模型的期望输出值。则  $y_t$  与  $\chi_t$  满足映射关系  $f: R^m \rightarrow R$ 。变换后的训练集矩阵如式(1)，式(2)所示：

$$\mathbf{X} = \begin{bmatrix} \chi_1 \\ \chi_2 \\ \vdots \\ \chi_{n_t-m} \end{bmatrix} = \begin{bmatrix} x(1) & x(2) & \cdots & x(m) \\ x(2) & x(3) & \cdots & x(m+1) \\ \vdots & \vdots & \ddots & \vdots \\ x(n_t-m) & x(n_t-m+1) & \cdots & x(n_t-1) \end{bmatrix} \quad (1)$$

$$\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n_t-m} \end{bmatrix} = \begin{bmatrix} x(m+1) \\ x(m+2) \\ \vdots \\ x(n_t) \end{bmatrix} \quad (2)$$

(2) 实体等时距状态预测模型构建：物联网中传感器所采集的实体状态数据具有非线性、随机性等特点，为减少传感器采集数据的存储开销，降低实体状态预测的计算复杂度，本文基于最小二乘支持向量机 (LSSVM)<sup>[7]</sup> 提出实体等时距状态估计 (Equal-interval Entity State Prediction, E-ESP) 方法，进行等时间距离下实体状态的预测。

为实现传感器采集的实体状态数据线性可分，需构造最优线性回归函数  $f(x)$ ：

$$f(x) = \mathbf{w}^T \phi(x) + b \quad (3)$$

式中， $\phi(x): R^n \rightarrow H$  为非线性映射函数； $\mathbf{w} \in H$  为特征空间中的权值向量； $b$  为偏移量。

为提高所提实体状态预测方法的泛化能力，根据结构风险最小化 (Structural Risk Minimization, SRM) 准则，将最小二乘支持向量机表述为

$$\left. \begin{aligned} \min_{\mathbf{w}, b, e} J(\mathbf{w}, e) &= \frac{1}{2} \|\mathbf{w}\|^2 + \frac{1}{2} \gamma \sum_{i=1}^l e_i^2 \\ \text{s.t. } y_i &= \mathbf{w}^T \phi(x_i) + b + e_i, \quad i = 1, 2, \dots, l \end{aligned} \right\} \quad (4)$$

其中， $J$  为耗费函数； $\gamma > 0$  表示惩罚系数； $e_i$  为偏差值。通过拉格朗日函数将其转化为非约束优化问题，从而得到该问题的对偶形式：

$$L(\mathbf{w}, b, e, \boldsymbol{\alpha}) = J(\mathbf{w}, e) - \sum_{i=1}^l \alpha_i [\mathbf{w}^T \phi(x_i) + b + e_i - y_i] \quad (5)$$

式(5)中  $\alpha_i$  为拉格朗日乘子。根据 KKT 条件并消去权值  $\mathbf{w}$  和偏差值  $e_i$  后，则可得到对应的矩阵方程

$$\begin{bmatrix} 0 & \mathbf{1}^T \\ \mathbf{1} & \boldsymbol{\Omega} + \mathbf{I}/\gamma \end{bmatrix}_{(l+1) \times (l+1)} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix} \quad (6)$$

式中， $\mathbf{y} = [y_1, y_2, \dots, y_l]^T$ ； $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_l]^T$ ； $\mathbf{I}$  为  $l \times l$  阶单位矩阵； $\boldsymbol{\Omega}$  为  $l \times l$  阶矩阵，且  $\Omega_{ij}$  具有如下形式：

$$\Omega_{ij} = \phi(x_i)^T \phi(x_j) \quad (7)$$

采用最小二乘法求解式(6)中参数  $b$  与  $\boldsymbol{\alpha}$ ，可推得实体状态预测函数  $y(x)$  为

$$y(x) = \sum_{i=1}^l \alpha_i K(x, x_i) + b \quad (8)$$

为降低所提实体状态预测方法的计算复杂度，本文选用广泛采用的高斯径向基 (RBF) 核函数：

$$K(x, x_i) = e^{-\|x-x_i\|^2/(2\sigma^2)} \quad (9)$$

则实体状态预测函数  $y(x)$  可进一步转化为

$$y(x) = \sum_{i=1}^l \alpha_i e^{-\|x-x_i\|^2/(2\sigma^2)} + b \quad (10)$$

### 3.2 实体周期内状态估计方法

前述预测模型可预测实体等时间距离的状态值。然而，用户发起内容搜索请求的时间具有较强的随机性。因此，本文进一步提出周期内实体状态的估计 (During the Period Entity State Prediction, DPESP) 方法，对周期内的实体状态进行估计，以满足用户在周期内的搜索需求。

为提高实体状态预测模型的精度，本文采用牛顿插值多项式<sup>[8]</sup>对实体周期内的状态进行预测。假设给定的实体状态样本点集为  $S_{NT} = \{(t_0, x(t_0)), (t_1, x(t_1)), \dots, (t_n, x(t_n))\}$ ， $t_i$  表示用户查询的时刻， $x(t_i)$  为用户查询时刻的实体状态值。定义  $x(t)$  关于  $t_0$  与  $t_1$  时刻实体状态的一阶均差  $\sigma[t_0, t_1]$  为

$$\sigma[t_0, t_1] = \frac{x(t_1) - x(t_0)}{t_1 - t_0} \quad (11)$$

同理可得  $x(t)$  关于  $t_0$ 、 $t_1$  与  $t_2$  时刻的二阶均差  $\sigma[t_0, t_1, t_2] = \frac{\sigma[t_1, t_2] - \sigma[t_0, t_1]}{t_2 - t_0}$ ，则  $x(t)$  关于  $(t_0, t_1, \dots, t_n)$  时刻的实体状态  $n$  阶均差为

$$\sigma[t_0, t_1, \dots, t_n] = \frac{\sigma[t_1, t_2, \dots, t_n] - \sigma[t_0, t_1, \dots, t_{n-1}]}{t_n - t_0} \quad (12)$$

由上文可知， $\sigma[t, t_0] = \frac{x(t_0) - x(t)}{t_0 - t}$ ，因此可得

$t$  时刻实体的状态  $x(t)$  与实体在  $t_0$  时刻的状态  $x(t_0)$  及其一阶均差  $\sigma[t, t_0]$  的关系为

$$x(t) = x(t_0) + (t - t_0) \sigma[t, t_0] \quad (13)$$

由于实体状态  $x(t)$  关于  $t_0$  与  $t_1$  时刻的二阶均差

$\sigma[t, t_0, t_1] = \frac{\sigma[t_0, t_1] - \sigma[t, t_0]}{t_1 - t}$ , 因此, 可将  $\sigma[t, t_0]$  进一步转化为

$$\sigma[t, t_0] = \sigma[t_0, t_1] + (t - t_1)\sigma[t, t_0, t_1] \quad (14)$$

将式(14)代入式(13)中可得到实体在  $t$  时刻的状态  $x(t)$  为

$$x(t) = x(t_0) + (t - t_0)\sigma[t_0, t_1] + (t - t_0)(t - t_1)\sigma[t, t_0, t_1] \quad (15)$$

由此, 可进一步获得预测多项式  $f_n(t)$  为

$$\begin{aligned} f_n(t) = & x(t_0) + (t - t_0)\sigma[t_0, t_1] \\ & + (t - t_0)(t - t_1)\sigma[t_0, t_1, t_2] + \dots \\ & + (t - t_0)(t - t_1)\dots(t - t_{n-1}) \\ & \cdot \sigma[t_0, t_1, \dots, t_n] \end{aligned} \quad (16)$$

周期内实体状态预测多项式  $f_n(t)$  构建完成后将开始对实体在当前周期内的状态进行预测。为保证用户得到搜索结果的可靠性, 网关将进行选定实体状态的验证。

为充分利用验证过程的实体状态信息, 提高实体状态预测方法的精度, 本文进一步提出利用验证数据的周期内实体状态估计(During the Period Entity State Prediction with Verification Data, DPESP-VD)方法, 将用户的查询时间  $t_q$  及被选定验证的实体状态值  $x(t_q)$  作为样本点代入预测多项式中, 对  $S_{NT}$  进行更新:

$$\begin{aligned} S'_{NT} = & \{(t_0, x(t_0)), (t_1, x(t_1)), \dots, (t_n, x(t_n)), \\ & \dots, (t_q^1, x(t_q^1)), \dots, (t_q^m, x(t_q^m))\} \end{aligned} \quad (17)$$

其中,  $t_q^i$  为用户请求搜索的时刻,  $x(t_q^i)$  为匹配用户内容搜索的实体实际状态值。反之, 若实体的当前预测状态不匹配用户的搜索内容, 即该实体不为选定验证实体, 则不更新该实体的状态多项式样本点集  $S_{NT}$ 。

## 4 实体有序验证方法

通过前述建立的预测模型, 网关可估计实体的状态与用户搜索的内容是否匹配, 然而, 实体状态的匹配程度尚缺乏统一的度量指标。并且, 为保证搜索结果的可靠性, 网关需对实体的实际状态作进一步验证, 而逐一验证的方式将极大地增加系统的通信开销, 影响搜索结果的时效性。因此, 本文提出实体有序验证方法, 由网关按照匹配排序列表依次对实体的状态进行验证。

### 4.1 实体匹配排序方法

为便于比较实体之间的匹配概率, 估计实体状态与用户搜索内容的匹配程度, 本文提出实体状态

匹配排序方法, 将实体状态预测值映射到向量空间中进行度量。假定实体在  $t_i$  时刻的状态预测值为  $\hat{x}(t_i)$ , 则将  $\hat{x}(t_i)$  映射入向量空间中并转化为如式(18)形式:

$$\hat{\mathbf{x}}(t_i) = \left[ |\hat{x}(t_i)_a|, |\hat{x}(t_i)_b| \right] \quad (18)$$

其中  $|\hat{x}(t_i)_a| = |\hat{x}(t_i) - a|$  为预测值  $\hat{x}(t_i)$  与查询范围下界  $a$  的距离,  $|\hat{x}(t_i)_b| = |\hat{x}(t_i) - b|$  为预测值  $\hat{x}(t_i)$  与查询范围上界  $b$  的距离。定义  $\mathbf{N}_a^b = [(a+b)/2, (a+b)/2]$  为查询范围  $[a, b]$  在向量空间中的基准向量。显然, 由于预测值存在随机误差, 在查询区间内  $((a+b)/2, (a+b)/2)$  点处的匹配稳定性最高, 相应的匹配概率也最大, 而与该点的距离越远则匹配概率越小。由此可得实体状态预测值  $\hat{x}(t_i)$  与查询内容的匹配概率  $M$  定义为

$$M = \frac{\hat{\mathbf{x}}^T(t_i) \mathbf{N}_a^b}{|\hat{\mathbf{x}}(t_i)| |\mathbf{N}_a^b|} \quad (19)$$

### 4.2 实体匹配验证方法

为保证搜索结果的可靠性, 需选定匹配实体并对其进行验证。本部分提出实体匹配验证方法, 依据实体匹配排序列表  $ML$  对排名靠前的实体的实际状态进行验证。

如前所述, 下层网关在获取本地实体匹配排序列表  $ML$  后, 将该列表发布给上层网关, 由上层网关对下层网关的多个排序列表进行融合处理, 从而得到验证列表  $VL_q = \bigcup_{k=1}^{\delta} VL_q^k$ , 融合决策规则为

$$\left. \begin{aligned} VL_q^1 &= \text{MAX} \{ML_1, ML_2, \dots, ML_n\} \\ VL_q^2 &= \text{MAX} \{ \{ML_1, ML_2, \dots, ML_n\} - VL_q^1 \} \\ &\vdots \\ VL_q^\delta &= \text{MAX} \left\{ \{ML_1, ML_2, \dots, ML_n\} - \sum_{k=1}^{\delta-1} VL_q^k \right\} \end{aligned} \right\} \quad (20)$$

其中  $ML_i$  为下层网关  $i$  上报的排序列表,  $\delta$  为上层网关发布的验证次数, 函数  $\text{MAX} \{\bullet\}$  为获取列表中具有最大匹配概率的节点。定义  $R_q$  为用户指定的需返回的搜索结果数量, 则  $\delta$  与  $R_q$  满足式(21), 式(22):

$$R_q = \sum_{k=1}^{\delta} e(VL_q^k) \quad (21)$$

$$e(VL_q^k) = \begin{cases} 1, & VL_q^k \in [a, b] \\ 0, & VL_q^k \notin [a, b] \end{cases} \quad (22)$$

其中,  $e(VL_q^k)$  为示性函数, 当被验证节点  $VL_q^k$  满足搜索范围时,  $e(VL_q^k)$  为1; 反之,  $e(VL_q^k)$  为0。

## 5 仿真验证与结果分析

本文采用 Intel Berkeley 数据集<sup>[9]</sup>对所提方法进行

行仿真验证。查询范围  $[a, b]$  为区间  $[x_{\text{Min}}, x_{\text{Max}}]$  内的随机值。 $x_{\text{Min}}$  与  $x_{\text{Max}}$  分别为所有传感器读数的最小值与最大值，查询时间  $t_q$  为  $[0, \omega]$  内的随机值。计算机仿真软件环境为 Matlab R2014a，仿真结果均为运行 20 次后的平均值。

5.1 实体等时距状态预测方法性能验证

本节对 E-ESP 与典型的等时距实体状态预测算法 GM-ESP<sup>[10]</sup> 进行了对比，仿真结果如图 2 所示。

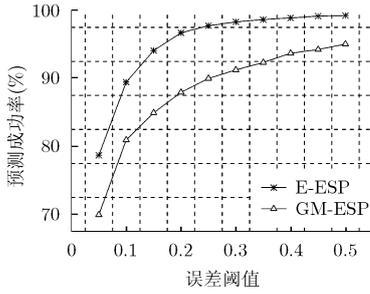


图 2 不同误差容忍阈值下的成功率对比

由图 2 可知，E-ESP 及 GM-ESP 的预测成功率随阈值的增大都呈现出显著上升趋势。E-ESP 较 GM-ESP 在预测成功率方面平均提高了约 7.6%。其中在  $\gamma = 0.2$  时，E-ESP 的预测成功率可达到 96.7%，在误差阈值较小时即达到了较高的精度。所提 E-ESP 算法通过将数据转换至矩阵空间中，从而增强实体状态数据之间的时间相关性，并采用 LSSVM 模型进行实体等时距状态的单步预测。因此，相比 GM-ESP，E-ESP 在预测成功率方面有较大幅度的性能提升。

5.2 周期内实体状态预测方法性能验证

本部分验证了所提周期内实体状态预测方法 (DPESP) 的性能，并与利用验证数据的周期内实体状态预测方法 (DPESP-VD) 进行了对比，仿真结果如图 3 所示。

由图 3 可知，随着查询次数的增多 DPESP-VD 方法的预测精度呈现上升趋势，而 DPESP 并未呈现出明显的变化趋势。两种周期内实体状态预测方

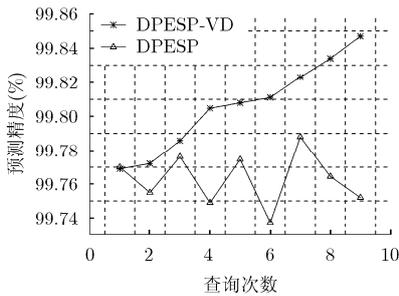


图 3 单周期内不同查询次数下的预测精度对比

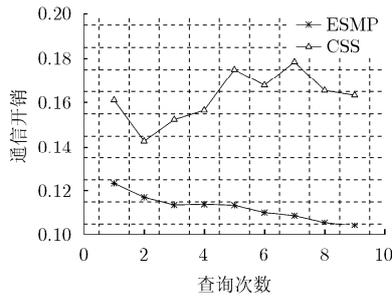


图 4 单周期内不同查询次数下的通信开销对比

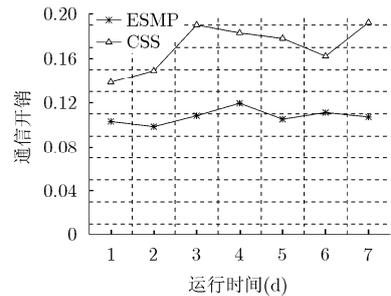


图 5 运行时间的通信开销对比

法均具有较高的精度，其中 DPESP-VD 较 DPESP 在预测精度方面有一定提升。DPESP-VD 方法合理利用了匹配验证过程中的历史数据，适用于所需预测精度较高的应用场景中，而 DPESP 方法未利用历史数据，适用于网关的存储空间受限的场景中。

5.3 实体匹配预测的搜索方法有效性验证

本节验证了融合所提实体匹配预测方法的搜索机制的有效性。由于所设计的搜索机制包括实体有序验证环节，因此用户总是得到符合搜索要求的实体列表，则搜索过程中的通信开销就只与实体的验证列表相关。因而，定义搜索过程的归一化通信开销  $co(VL_q)$  为

$$co(VL_q) = \frac{\sum_{i=1}^{|VL_q|} e_{VL_q}(i)}{m(VL_q)[|VL_q| - m(VL_q)]} \quad (23)$$

其中  $|VL_q|$  为验证列表  $VL_q$  的长度， $m(VL_q)$  为在  $VL_q$  中匹配的实体数量。 $e_{VL_q}(i)$  为实体的排名误差，若  $i$  是一个不匹配的实体，则  $e_{VL_q}(i) = 0$ 。由上可知，通信开销  $co(VL_q)$  为与排名误差成比例的归一化值。本节分别在不同查询次数与运行时间条件下验证 ESMP 与前述 CSS 算法的通信开销，仿真验证结果如图 4 所示。

由图 4 可知，随着单个周期内查询次数的增多，所提 ESMP 算法的通信开销呈现出显著的下降趋势，而 CSS 算法的趋势则并不明显。ESMP 的通信开销相比 CSS 平均降低了约 30.6%，较 CSS 算法有较大幅度的提升。

如图 5 所示，ESMP 算法的通信开销随运行时间的延长保持了相对稳定的趋势，充分说明了 ESMP 方法具有较好的泛化能力，对实体状态的预测精度保持了较为稳定的态势。而 CSS 对实体状态的估计存在较大的随机误差，因而，CSS 机制的通信开销随运行时间的延长表现出较大幅度的波动。并且，ESMP 较 CSS 在通信开销方面平均降低了约 36.2%，在通信开销性能方面 ESMP 相比 CSS 有较大幅度的性能增益。

## 6 总结

本文提出了适用于物联网内容搜索的实体状态匹配估计方法, 设计了等时距与周期内实体状态估计方法对未来实体状态进行估计, 并依据预测所得实体状态提出实体有序验证方法, 验证选定实体与用户搜索需求的符合状态, 以提高物联网内容搜索的效率。仿真结果表明, 相比现有的典型方法, 本文所提实体状态预测方法具有较高的预测精度, 基于所提匹配预测方法的搜索机制具有较低的通信开销。

## 参 考 文 献

- [1] 于海宁, 余翔湛. 物联网中基于编码的多路径并发实体通用搜索算法[J]. 智能计算机与应用, 2013, 3(2): 25-31.  
Yu Hai-ning and Yu Xiang-zhan. A multi-path concurrency object universal search (MCOUS) algorithms based on code in internet of things[J]. *Intelligent Computer and Applications*, 2013, 3(2): 25-31.
- [2] 茹立云, 李智超, 马少平. 搜索引擎索引网页集合选取方法研究[J]. 计算机研究与发展, 2014, 51(10): 2239-2247.  
Ru Li-yun, Li Zhi-chao, and Ma Shao-ping. Indexing page collection selection method for search engine[J]. *Journal of Computer Research and Development*, 2014, 51(10): 2239-2247.
- [3] 李养群, 沈苏彬, 许斌. 物品万维网技术综述[J]. 南京邮电大学学报(自然科学版), 2014, 34(2): 32-42.  
Li Yang-qun, Shen Su-bin, and Xu Bin. Technology of web of things: a survey[J]. *Journal of Nanjing University of Posts and Telecommunications (Natural Science)*, 2014, 34(2): 32-42.
- [4] 于海宁, 张宏莉, 方滨兴, 等. 物联网中物理实体搜索服务的研究[J]. 电信科学, 2012, 28(10): 111-119.  
Yu Hai-ning, Zhang Hong-li, Fang Bin-xing, et al. Research on search service for physical entities in the internet of things[J]. *Telecommunications Science*, 2012, 28(10): 111-119.
- [5] Benedikt O, Kay R, Friedemann M, et al. A real-time search engine for the web of things[C]. *Internet of Things*, Tokyo, 2010: 1-8.
- [6] Cuong T and Kay R. Content-based sensor search for the web of things[C]. *Global Communications Conference*, Atlanta, 2013: 2654-2660.
- [7] 唐舟进, 彭涛, 王文博. 一种基于相关分析的局域最小二乘支持向量机小尺度网络流量预测算法[J]. 物理学报, 2014, 63(13): 57-66.  
Tang Zhou-jin, Peng Tao, and Wang Wen-bo. A local least square support vector machine prediction algorithm of small scale network traffic based on correlation analysis[J]. *Acta Physica Sinica*, 2014, 63(13): 57-66.
- [8] 李增科, 高井祥, 王坚, 等. 利用牛顿插值的GPS/INS组合导航惯性动力学模型[J]. 武汉大学学报(信息科学版), 2014, 39(5): 591-595.  
Li Zeng-ke, Gao Jing-xiang, Wang Jian, et al. Inertial dynamic model of GPS/INS integrated navigation based on newton interpolation[J]. *Geomatics and Information Science of Wuhan University*, 2014, 39(5): 591-595.
- [9] <http://db.csail.mit.edu/labdata/labdata.html>[OL]. 2014.10.
- [10] 尚军亮, 方敏. 一种优化的高精度灰色 GM(1,1)预测模型[J]. 电子与信息学报, 2010, 32(6): 1301-1305.  
Shang Jun-liang and Fang Min. New optimized method of high-precision grey GM(1,1) forecasting model[J]. *Journal of Electronics & Information Technology*, 2010, 32(6): 1301-1305.

张普宁: 男, 1988年生, 博士生, 研究方向为物联网搜索、无线传感器网络。

刘元安: 男, 1963年生, 教授, 研究方向为电磁兼容、泛在无线网络。

吴帆: 女, 1981年生, 副教授, 研究方向为物联网搜索、泛在无线网络。