基于智能交通的隐私保护道路状态实时监测方案

李家印[1,2]，郭文忠[1,3]，李小燕[1]，刘西蒙[1,2]

（1. 福州大学数学与计算机科学学院，福建 福州 350108；2. 福州大学网络安全福建省高校重点实验室，福建 福州 350108；
3. 福州大学福建省网络计算与智能信息处理重点实验室，福建 福州 350108）

摘 要：为缓解道路的交通压力，减少道路拥堵现象的出现及避免交通事故的发生，结合安全、K 最近邻（KNN）算法，提出了一种基于智能交通的隐私保护道路拥堵状态的实时监测（PPIM）算法。为了确保交通数据的安全，采用安全多方计算策略将数据内容随机分成独立的部分，通过不共谋的多服务器对数据分量进行存储和加密。为了提升道路状态监测的精度，提出了一种改进型的 KNN 交通监测算法，借助数据的相似度计算，获取衡量道路之间交通状态关系程度的相关值，并将其作为权重系数与传统的 KNN 算法进行整合。为加快密态数据的处理速度，设计了一系列的数据安全计算协议，实现了数据的安全处理。另外，利用真实的交通数据对该算法进行验证，实验结果表明改进型 KNN 算法有助于提高道路监测的准确度。实验分析表明，所提算法在保证数据的安全同时可以提高交通监测的精度。

关键词：智能交通；隐私保护；空间距离；K 最近邻

中图分类号：TP302

文献标识码：A

doi: 10.11959/j.issn.1000−436x.2020110

Privacy-preserving real-time road conditions monitoring scheme based on intelligent traffic

LI Jiayin[1,2], GUO Wenzhong[1,3], LI Xiaoyan[1], LIU Ximeng[1,2]

1. College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350108, China
2. Key Lab of Information Security of Network Systems, Fuzhou University, Fuzhou 350108, China
3. Fujian Provincial Key Laboratory of Networking Computing and Intelligent Information Processing, Fuzhou University, Fuzhou 350108, China

Abstract: To alleviate the traffic pressure on roads, reduce the appearance of road congestion, and avoid the occurrence of traffic accidents, a privacy-preserving intelligent monitoring (PPIM) scheme based on intelligent traffic was proposed in combination with the safe and k-nearest neighbor (KNN) algorithm. To ensure the security of traffic data, the data content was randomly divided into independent parts via the secure multi-party computing strategy, and the data components were stored and encrypted separately by non-colluding multi-servers. To improve the accuracy of road condition monitoring, an improved KNN traffic monitoring algorithm was proposed. By virtue of the similarity calculation of data, the correlation value to measure the degree of traffic condition relationship between roads was obtained. And it was integrated with the KNN as the weight coefficient. To speed up the processing of dense data, a series of data security computing protocols were designed, and the data security processing was realized. In addition, real traffic data were used to verify the algorithm. The results show that the improved KNN algorithm is helpful to improve the accuracy of traffic monitoring. The analysis shows that the algorithm can not only guarantee the safety of data but improve the accuracy of traffic monitoring.

Key words: intelligent traffic,privacy-preserving, space distance, KNN

收稿日期：2020–03–24；**修回日期**：2020–05–14

通信作者：刘西蒙，snbnix@gmail.com

基金项目：国家自然科学基金资助项目（No.61702105, No.U1804263, No.61672159, No.U1705262）

Foundation Item: The National Natural Science Foundation of China (No.61702105, No.U1804263, No.61672159, No.U1705262)

1 引言

随着经济的飞速发展、社会城市化进程的加快，全球各大城市均面临交通拥堵带来的压力。近些年，我国机动车保有量持续快速增长，城市交通拥堵已成为大众出行讨论的焦点。交通的拥堵不但延长了人们出行的时间，增大了机动车的油耗，而且加剧了环境的污染[1-2]。更甚者，交通拥堵导致交通事故频发，给人们的生命和财产带来极大威胁。为缓解城市交通压力，构建智慧城市，建立完善的智能交通系统成为城市建设的关键。智能交通系统（ITS, intelligent traffic system）指对道路交通状况实施监测的系统，该系统借助数据挖掘方法获取道路的拥堵状态，然后根据系统分析的结果对机动车实施有效的调度[3-4]。事实上，许多监测手段已经应用到了智能交通系统，并实现了对交通拥堵状况的判断。例如，通过道路两侧安装静态传感器获取道路的实时数据，并将其传输到数据中心加以处理与分析，得到反映道路拥堵状态的数值。另外，借助高清摄像头采集交通图像数据，根据对图像的处理分析结果做出对道路拥堵情况的准确判定[5-7]。然而，静态传感器、高清摄像头的安装成本和维护成本十分高昂，尽管可以有效地实现交通状况的监控，但不适合大规模地部署[8]。此外，固定的静态传感器的感知区域有很大局限性，所获取的数据也过于稀疏，不利于数据的处理与分析。随着传感器功能的增加，越来越多的机动车安装了智能车载传感器设备[9]。研究学者利用车载传感器不断地收集机动车交通数据（如实时位置、行驶速度、方向、数据上报时间等信息），实现低成本、高精度的交通监测。此外，机动车数量庞大且行驶具有随机性，能够有效解决交通数据过于稀疏的问题[10]。然而，利用机动车交通数据实现道路状态监测，忽略了因数据泄露引发的一系列危害。由于交通数据包含诸多敏感信息，攻击者很容易通过数据提供者的历史数据推断出未来某个时间段数据提供者所处的位置，甚至还可以通过频繁出入的地方推测出数据提供者的个人信息，诸如兴趣爱好、健康状况、收入等敏感信息[11-12]。研究学者提出利用传统的加密算法对交通数据进行加密，以确保数据传输及存储过程中的安全性。处理中心在密态数据的基础上实现交通状态的监测。借助传统的加密算法尽管能够满足数据提供者对数据安全性的要求（如高级加密标准（AES, advanced encryption standard）[13-14]等），但传统的加密算法存在明显的弊端，即无法对密态数据直接进行处理与分析。为达到监测的目的，需要先对密文进行解密，这不但将明文数据暴露给了数据处理中心，而且还延长了交通监测所需的时间。Paillier 算法尽管可以避免加密数据的解密过程，但其完成数据处理与分析的过程非常耗时，违背了交通监测对时延的要求。更重要的是，借助加密算法能够保证在交通数据中的敏感信息不被泄露的同时，更需要实现机动车驾驶员对交通状态的安全查询，因此，实现对加密数据的查询功能是完成智能交通隐私保护的关键[15]。其不仅需要满足任意时刻机动车驾驶员对智能交通系统分析出的某一具体道路交通拥堵状态结果的查询，而且能够保证查询值的正确性及反馈结果的唯一性。通过对加密数据的查询，能够保证在整个智能交通系统的实现过程，机动车所产生的交通数据始终受到保护而不被泄露。因此，在保证交通数据隐私安全的前提下，既能快速地挖掘出数据的内在信息、又能安全且准确地实现驾驶员对交通道路拥堵状态结果的查询是研究隐私保护智能交通系统的核心难题。

安全多方计算（MPC, secure multiparty computation）为隐私保护智能交通道路监测的实现提供了可行的方案[16-18]，该方案在确保数据安全的前提下，完成对明文数据的处理、分析与查询。但是，如何设计适用于处理与分析机动车交通数据的安全计算协议，降低处理数据耗费的时间仍是实现交通监测所面临的难题。

针对上述问题，本文对传统 K 最近邻（KNN, K-nearest neighbor）距离算法进行改进，提出了一种基于智能交通的隐私保护道路状态实时监测（PPIM, privacy preserving intelligent monitoring）算法。借助基础安全计算协议，并设计了一系列的安全计算协议，在保证数据隐私安全的前提下，完成了交通数据的处理与分析，通过部分已知道路的拥堵状态，达到对整个路网中其余未知道路拥堵状态的监测。此外，还提出一种改进型 KNN 算法，通过引入皮尔逊相关系数，得到已知道路与未知道路之间拥堵状态的权重系数，提高了对道路拥堵状况监测的准确度。

2　预备知识

2.1　路网的拓扑结构与拥堵状态

交通路网（简称“路网”）指供机动车行驶的庞大道路，其最基础的结构是路段与路口，如图 1 所示。通常，路网的拓扑结构用图 $G(L,C,V)$表示，其中，L 表示路网中的道路；C 表示道路 L 的中间位置；V 表示道路 L 某一时间间隔内道路的平均速度，即此时道路允许的机动车通行能力。假设图 $G(L,C,V)$包含 n 条道路，定义 $i\in[1,n]$（i 表示路网中道路的编号）。倘若道路 i 在时间间隔 T 内采集了 m 辆机动车的数据，并且上报 k 条的数据，v_τ（$\tau\in[1,k]$）表示每条数据中车辆的瞬时速度，单位为 km/h。基于上述定义，可以计算出道路 i 在时间间隔 T 内的平均速度 $v_i(T)=\frac{1}{k}\sum_{\tau=1}^{k}v_\tau$。结合真实的驾车体验，道路的拥堵状态可分为 4 种不同程度。1) 当道路的平均速度在 $0<v_i\leqslant 10$ 范围时，表示该道路处于十分拥堵的状态；2) 当 $10<v_i\leqslant 30$ 时，表示该道路处于较拥堵的状态；3) 当 $30<v_i\leqslant 45$ 时，表示明机动车可以在该道路上畅通地行驶；4) 当 $v_i>45$ 时，表示该道路此时的通行状态十分畅通。

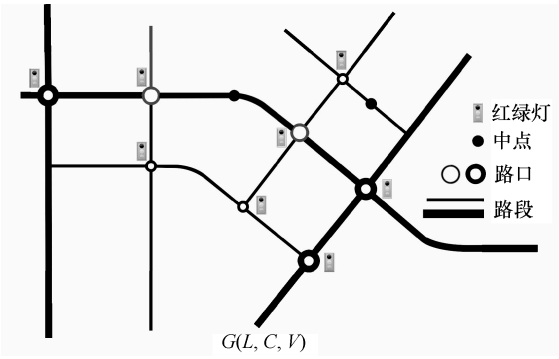

图 1　路网的拓扑结构

2.2　传统的 KNN 算法监测框架

KNN 算法常被用于监测道路交通的拥堵状态。若 $v_\tau(T)$ 表示时间间隔 T 内道路 τ 待求解的预测值，$\boldsymbol{v}(T)$ 表示时间间隔 T 预测 $v_\tau(T)$ 所对应的特征向量。对于特征向量 $\boldsymbol{v}(T)$，KNN 算法在历史数据集中搜索与其距离最近的 K 个历史特征 $x_i,x_{i+1},\cdots,x_{i+K}$，利用式(1)将 K 个特征向量通过加权估计的方法，求解 $v_\tau(T)$ 的值。

$$v_\tau=\sum_{i=1}^{K}\frac{d_i^{-1}}{\sum_{i=1}^{K}d_i^{-1}}v_i(T) \tag{1}$$

由式(1)可知，KNN 算法中的唯一参数是确定 K，其值直接关系到判断未知道路拥堵情况的准确性。通常采用交叉验证法确定最优 K 值，具体步骤如下。

步骤 1　定义选取 K 的最大值 $K_{\max}$ 和最小值 $K_{\min}$。根据路网中道路总数可知 $K_{\max}=n$，$K_{\min}=1$。

步骤 2　将 n 条道路的历史数据以时间间隔 T 为单位计算 n 条道路的 n 个特征值，得到道路历史数据集的集合 $\{v_1,v_2,\cdots,v_n\}$，然后依次将数据集 $\{v_1,v_2,\cdots,v_n\}$ 作为测试数据集，其他 $n-1$ 个数据集作为历史数据集。

步骤 3　将 K 从 $K_{\min}$ 以步长为 1 进行取值，直至取到 $K_{\max}$ 后结束。

步骤 4　优化 $\min\|v_\tau(T)-v'_\tau(T)\|_2^2$ 二范式，使待求解道路 τ 的预测值 $v_\tau(T)$ 与真实值 $v'_\tau(T)$ 相差最小。

步骤 5　计算 n 个数据集对应选取 K 值的平均误差百分比，并对 n 个值求均值，即 $K=\frac{1}{n}\sum_{i=1}^{n}\mathrm{num}_{(i,K)}$。其中，$\mathrm{num}_{(i,K)}$ 表示计算道路 i 时 K 的具体数值。

本文基于 KNN 算法实现交通状态的实时监测，主要是因为其更能满足交通监测对时延的要求。首先，相比于其他数据预测算法，KNN 算法不仅具有简单易用的特点，而且能达到较好的预测效果；其次，KNN 算法模型的训练时间快，其模型训练主要是确定 K 的取值，对于特定的某个区域，一旦 K 值被确定，在相当长的时间内可以对未知数据进行预测；最后，KNN 算法对异常值不敏感，换而言之，交通数据内经常出现的异常值对 KNN 算法的预测精度不会造成很大的影响。

2.3　安全两方计算

安全两方计算是指双方在互不泄露各自数据的前提下完成数据的处理与分析。假设 2 个独立的参与实体 Alice 和 Bob，两者分别存储各自的私有数据，Alice 的私有数据是随机数 a，Bob 的私有数据是随机数 b。安全两方计算的目的是获取 $a\times b$ 的结果，但是 Alice 和 Bob 都不想泄露其私有信息的数值。安全两方计算过程如下。首先定义函数 Γ 是期望实现的乘法计算，经过执行 $z-1$ 次函数 Γ 得到最终结果 S_z，即存在 $S_z\xleftarrow{z-1}\Gamma(a,b)$。然后 S_z 被随机分成 a_z 和 b_z，并满足 $S_z=a_z+b_z$ 的关系。最后 Alice 获取 a_z 的值，Bob 接收 b_z 的值。具体过程如图 2 所示。

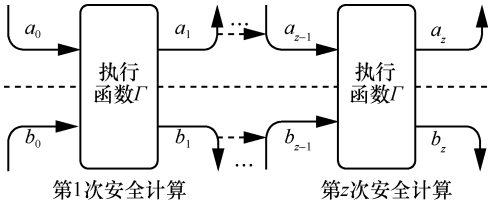

图 2 安全两方计算过程

通过安全两方计算来维护交通数据的隐私安全，在不泄露数据敏感信息的前提下，实现对数据的安全处理。安全两方计算能够加快处理数据的速度，降低获取城市路网道路交通状态的时延，达到道路实时监测的目的。

2.4 基础安全计算协议

本文根据已有研究成果及基础安全计算协议，来实现所提 PPIM 算法。基础安全计算协议基于 2 个云服务器 SP_1 和 SP_2，通过两者之间的安全交互数据来完成计算，具体如下。

基础安全乘法协议（BSMP, basic secure multiplication protocol）。假设云服务器 SP_1 和 SP_2 分别存储私有数据 a 和 b，将满足 $k_a+k_b=r_ar_b$ 条件所生成的 4 个随机数 k_a、k_b、r_a 和 r_b 作为密钥。SP_1 计算 $a'=a+r_a$，SP_2 计算 $b'=b+r_b$，并将 a' 和 b' 进行互换。然后，SP_2 产生随机数 v_b 并计算 $t=a'b+k_b-v_b$，SP_1 计算 $v_a=t+k_a-r_ab'$。经上述过程即可实现 $v_a+v_b=ab$ 的计算，并将 v_a 和 v_b 替代 SP_1 和 SP_2 原有的数据 a 和 b。因此，基础安全乘法协议可表示为映射 $v_a+v_b\leftarrow \mathrm{BSMP}(a:b)$。

基础安全除法协议（BSDP, basic secure division protocol）。SP_1 和 SP_2 分别存储私有数据 a 和 b，将满足 $k_a+k_b=r_ar_b$ 条件所生成的 4 个随机数 k_a、k_b、r_a 和 r_b 作为密钥。SP_1 计算 $a'=a+r_a$，SP_2 计算 $b'=\dfrac{1}{b+r_b}$，并将 a' 和 b' 进行互换。然后，SP_2 产生随机数 v_b 并计算 $t=a'\left(\dfrac{1}{b}\right)+k_b-v_b$，$SP_1$ 计算 $v_a=t+k_a-r_ab'$。经上述过程即可实现 $v_a+v_b=\dfrac{a}{b}$ 的计算，并将 v_a 和 v_b 替代 SP_1 和 SP_2 原有的数据 a 和 b。因此，基础安全除法协议可表示为映射 $v_a+v_b\leftarrow \mathrm{BSDP}(a:b)$，其中，$v_a$ 和 v_b 分别表示原始数据 a 和 b 的密文值。

3 交通监测方案

为实现智能交通隐私保护道路状态的监测，本节给出了系统监测的模型，并对模型内实体的功能逐一介绍；结合系统模型，阐明系统存在的安全威胁；为防止数据泄露带来的危险，设计了适于隐私保护改进型 KNN（IKNN, improved KNN）算法的高效安全计算协议。

3.1 系统模型

道路状态监测的系统模型如图 3 所示，该模型包含 4 个实体的参与，具体如下。

可信任实体（TA, trusted authority）。一个独立可信任的第三方，它的主要任务是生成随机数并将随机数通过安全信道发送给其他参与方。

数据提供车辆（VD, vehicle data）。负责机动车交通数据的收集，并将收集的数据实时地传输到云服务器，即云服务提供商。

云服务提供商（SP, server provider）。接收来自数据提供车辆发送的数据并对其存储和处理与分析。

导航提供商（NAV, navigation）。负责接收 2 个云服务提供商计算出的结果，并将其发布给行驶在道路上的机动车，用于提升司机的驾驶体验。

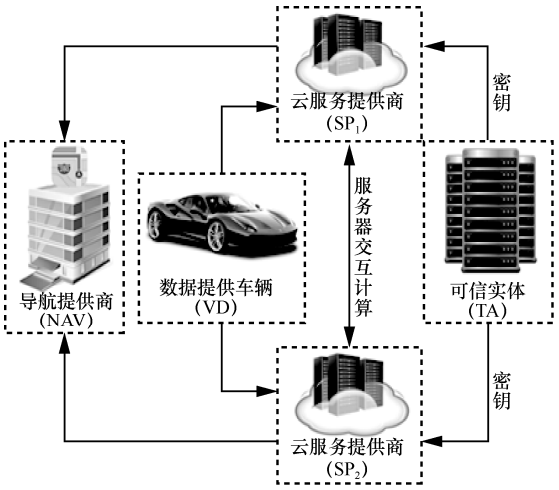

图 3 系统模型

3.2 安全威胁

安全威胁主要来自恶意攻击者对提供数据存储服务的第三方云服务提供商进行的恶意攻击，该攻击容易造成数据提供者信息的泄露。假设存在一个模拟器 Υ，其任意地产生随机值并获取安全计算协议真实的视图 $\mathcal{H}_1$。基于 $\mathcal{H}_1$，模拟器 Υ 试图在多项式的时间内生成一个模拟的视图 $\mathcal{H}_2$。假设存在 2 个不共谋的恶意攻击者 $\mathcal{A}_1$ 和 $\mathcal{A}_2$。攻击者 $\mathcal{A}_1$ 可以成功地攻击并获取存储在服务器 SP_1 内的交通数据，

一次成功的攻击意味着 $\mathcal{A}_1$ 可以找到一个概率二项式算法成功地区分 $\mathcal{H}_1$ 和 $\mathcal{H}_2$，并结合获取的数据推测出真实的数据值。攻击者 $\mathcal{A}_2$ 也可以成功地攻击并获取存储在服务器 SP_2 中数据，$\mathcal{A}_2$ 完成一次成功的攻击同样意味着其可以找到一个概率二项式算法对 $\mathcal{H}_1$ 和 $\mathcal{H}_2$ 进行区分，结合获取的数据推测出其真实值。此外，云服务提供商是好奇并诚实的实体，在执行数据处理时会利用已知的部分数据来猜测真实的数据值。

3.3　设计目标

本文主要目的是实现对智能交通隐私保护道路交通状态的监测，因此，系统既要满足实现交通监测的目标，又要符合数据隐私保护的要求，具体要求如下。

数据的安全性。任何收集的机动车行驶数据都应该受到保护，比如车辆 ID、车辆行驶的速度、车辆的经纬度等信息。

数据计算的低复杂性。用于数据加密的加密算法应当是轻量级的、快速的。

数据的轻量化传输。为了降低交通监测的时延，应尽可能地减少交通数据的传输量。

预测的实时性。为满足道路交通监测的实时性，利用少量的数据实现对整个路网道路状况的快速监测。

3.4　IKNN 监测算法

为提高传统 KNN 算法的监测精度，本文提出了 IKNN 算法，其主要思想是利用已知道路的部分平均速度对未知道路的平均速度进行预测。通过增加权重调节因子 ω，借助机动车的历史数据，整合距离预测道路最近的 K 条道路与其之间的相似度值，调节不同道路在预测未知道路时的影响程度。对于权重调节因子 ω 的计算，本文首先采用归一化的欧几里得公式将获取的路段速度值归一化到 0～1。然后利用式(2)所示的皮尔逊相关系数进行计算，获取所需的权重值。

$$\omega(\boldsymbol{x},\boldsymbol{y})=\frac{\mathrm{E}(\boldsymbol{xy})-\mathrm{E}(\boldsymbol{x})\mathrm{E}(\boldsymbol{y})}{\sqrt{\mathrm{E}(\boldsymbol{x}^2)-\mathrm{E}^2(\boldsymbol{x})}\sqrt{\mathrm{E}(\boldsymbol{y}^2)-\mathrm{E}^2(\boldsymbol{y})}} \tag{2}$$

其中，$\boldsymbol{x}$ 表示待求解道路的历史平均速度的向量，$\boldsymbol{y}$ 表示距离其最近的 K 条道路中某一条路段的历史平均速度向量。由皮尔逊相关系数的性质可知，ω 值越大，说明向量 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 两者之间的相似度越大，反之亦然。基于 ω 给出 IKNN 交通监测计算式，如式(3)所示。

$$v_\tau(T)=\sum_{i=1}^{K}\frac{\omega_i d_i^{-1}}{\sum_{i=1}^{K}\omega_i d_i^{-1}}v_i(T) \tag{3}$$

ω 的引入更加凸显出已知道路中每一条道路对未知路段具有不同程度的影响。根据选取的已知道路与待求解道路历史数据之间的皮尔逊相关系数，降低预测值与真实值之间的误差。

由式(3)可知，IKNN 算法的前提是确定 K 值及计算 ω，具体如下。

步骤 1　利用收集的整个路网的历史数据，根据式(1)计算出时间间隔 T 内每一条道路的平均速度及经纬度信息。采用道路中间位置的经纬度 $C_i(x,y)$ 表示道路 i 的空间地理信息。$C_i(x,y)$ 利用两点之间的中点求解公式 $x=\frac{x_1+x_2}{2}$ 和 $y=\frac{y_1+y_2}{2}$ 进行求解，其中 (x_1,y_1) 表示道路 i 最靠近路口一端的经纬度信息，(x_2,y_2) 表示道路 i 最远离该路口另一端的经纬度位置信息。

步骤 2　通过训练历史的交通数据，确定用于估计未知道路平均速度时的 K 值。

步骤 3　依次计算选取的 K 条道路与其余 $n-K$ 条道路之间的权重向量 $\boldsymbol{\Theta}=(\omega_1,\omega_2,\cdots,\omega_{n-K})$。

步骤 4　根据步骤 3 得到的 K 值和式(3)计算出未知道路时间间隔 T 内待求解道路的平均速度 $v_\tau(T)$。

4　隐私保护道路监测

本节阐述了智能交通隐私保护道路监测的实现过程。为说明算法执行的具体步骤，首先定义了机动车交通数据的数据格式及数据到云服务提供商的上传格式。然后阐述云服务提供商在保护数据隐私安全的基础上，采用 IKNN 算法，实现道路交通状态的监测。

4.1　数据格式和数据传输

实现监测的关键是收集必要的数据。定义 $R(\mathrm{lat},\mathrm{lon},v,t)$ 表示数据的上传格式，时间间隔 T 内道路 i 上行驶的车辆共上报 k 条数据，$D_\tau=\{(\mathrm{lat}_\tau,\mathrm{lon}_\tau,v_\tau,t_\tau)\,|\,\tau\in[1,k]\}$ 表示数据的集合。然后，将每一条数据内的每一个元素随机地分为两部分，即 $D'_\tau=\{(\mathrm{lat}_\tau,v'_\tau,t'_\tau)\}$ 和 $D''_\tau=\{(\mathrm{lon}_\tau,v''_\tau,t''_\tau)\}$，其中，$v_\tau=v'_\tau+v''_\tau$，$t_\tau=t'_\tau+t''_\tau$。最后，每个元素被随机地分为 2 个分量分别传输给云服务提供商 SP_1 和 SP_2，即 D'_τ 作为一个整体发送给 SP_1，D''_τ 传输给 SP_2。

4.2 经纬度距离安全计算协议

将数据安全地外包给云服务提供商后，云服务提供商利用存储的数据计算相应道路的平均速度。对于道路 i，SP_1 根据存储的速度数据分量 $v_i'=\{v'_{(i,\tau)}\mid\tau\in[1,k]\}$，对集合 v_i' 计算 $v_a=\frac{1}{k}\sum_{\tau=1}^{k}v'_{(i,\tau)}$；$\mathrm{SP}_2$ 计算 $v_b=\frac{1}{k}\sum_{\tau=1}^{k}v''_{(i,\tau)}$。根据机动车的历史经纬度计算能够体现道路空间位置的信息。路段 i 的 GPS 纬度信息的集合 $\mathrm{lat}_i=\{\mathrm{lat}_{(i,\tau)}\mid\tau\in[1,k]\}$ 存储在 SP_1，经度信息集合 $\mathrm{lon}_i=\{\mathrm{lon}_{(i,\tau)}\mid\tau\in[1,k]\}$ 存储在 SP_2。根据 3.4 节中步骤 1 所述，近似地计算出道路 i 的空间经纬度位置信息 $C_i(\mathrm{lat}_C,\mathrm{lon}_C)$。为计算空间两点经纬度之间的距离，本文将两点间距离计算方法推广到求解空间两点间经纬度的欧氏距离，如式(4)所示。

$$d=R\arccos(l_il_\tau p_{(i,\tau)})+b_ib_\tau \tag{4}$$

其中，R 表示地球的半径，p 表示两点之间经度差的余弦值，$l_i=\cos(\mathrm{lat}_i)$，$l_\tau=\cos(\mathrm{lat}_\tau)$，$b_i=\sin(\mathrm{lat}_i)$，$b_\tau=\sin(\mathrm{lat}_\tau)$，$p_{(i,\tau)}=\cos(\mathrm{lon}_i-\mathrm{lon}_\tau)$。根据式(4)计算基于密文状态下 2 个不同经纬度之间距离，具体如下。

步骤 1 SP_1 执行 $l=l_il_\tau=\cos(\mathrm{lat}_i)\cos(\mathrm{lat}_\tau)$ 和 $b=b_ib_\tau=\sin(\mathrm{lat}_i)\sin(\mathrm{lat}_\tau)$ 的计算。

步骤 2 SP_2 对两点之间经度进行差值运算，即 $p_{(i,\tau)}=\cos(\mathrm{lon}_i-\mathrm{lon}_\tau)$。

步骤 3 根据泰勒公式的展开近似等式，即 $\arccos(x)=x+\frac{1}{6}x^3+\frac{3}{40}x^5+\frac{5}{112}x^7+\frac{35}{1152}x^9+o(x^9)$ 以及二项式展开 $(a+b)^n=C_n^0a^nb^0+\cdots+C_n^na^0b^n$，计算对应数值反余弦的近似值。当 $f\in[9,7,5,3]$、$\rho=\left[\frac{35}{1152},\frac{5}{112},\frac{3}{40},\frac{1}{6}\right]$（$f$ 和 ρ 每次选取的元素位置相同）时，根据每次选取的 f 和 ρ 值，执行 BSMP，结合 f 和 ρ 的取值范围，协议执行可以表示为 $v_{a(f,g)}+v_{b(f,g)}\leftarrow\mathrm{BSMP}(C_f^gl^{\max(g)-g}:\rho p^g)$，其中，$g\in[1,f]$。然后，将计算得到的 $v_{a(f,g)}$ 发送给 SP_1，将获得的 $v_{b(f,g)}$ 发送给 SP_2。

步骤 4 SP_1 计算 $v_a'=\sum v_{a(f,g)}$，SP_2 计算 $v_b'=\sum v_{b(f,g)}$。

步骤 5 SP_1 执行 $v_a=Rv_a'+b$ 的计算，SP_2 执行 $v_b=Rv_b'$ 的计算。

经上述步骤，得到路网中任意 2 条道路之间的空间距离 d。并且满足 $d=v_a+v_b$ 的映射关系。其中，v_a 存储在 SP_1，v_b 存储在 SP_2。

假设时间间隔 T 内，智能交通系统共收集到来自 h 条道路上的机动车的交通数据，根据前文所述的平均速度计算方法，系统计算出 h 条道路的平均速度，组成向量 $\boldsymbol{V}=[v_1,v_2,\cdots,v_h]$。另外，向量 $\boldsymbol{V}$ 随机地被分为两部分 $\boldsymbol{V}'$ 和 $\boldsymbol{V}''$，然后分别存储在 SP_1 和 SP_2。结合存储在云服务器中 n 条道路的历史数据，提取出任意 2 条道路之间的空间相关性，得到 n 条道路的空间相关性矩阵 $\boldsymbol{\Omega}$，如式(5)所示。

$$\boldsymbol{\Omega}=\begin{bmatrix}d_{11}&d_{12}&\cdots&d_{1n}\\d_{21}&d_{22}&\cdots&d_{2n}\\\vdots&\vdots&\ddots&\vdots\\d_{n1}&d_{n2}&\cdots&d_{nn}\end{bmatrix} \tag{5}$$

同样地，空间相关性矩阵 $\boldsymbol{\Omega}$ 被随机地分为 $\boldsymbol{\Omega}'$ 和 $\boldsymbol{\Omega}''$ 两部分分量，被依次存储在 SP_1 和 SP_2，并且满足 $\boldsymbol{\Omega}=\boldsymbol{\Omega}'+\boldsymbol{\Omega}''$ 的等式关系。其中，$\boldsymbol{\Omega}'$ 和 $\boldsymbol{\Omega}''$ 分别如式(16)和式(17)所示。

$$\boldsymbol{\Omega}'=\begin{bmatrix}d'_{11}&d'_{12}&\cdots&d'_{1n}\\d'_{21}&d'_{22}&\cdots&d'_{2n}\\\vdots&\vdots&\ddots&\vdots\\d'_{n1}&d'_{n2}&\cdots&d'_{nn}\end{bmatrix} \tag{6}$$

$$\boldsymbol{\Omega}''=\begin{bmatrix}d''_{11}&d''_{12}&\cdots&d''_{1n}\\d''_{21}&d''_{22}&\cdots&d''_{2n}\\\vdots&\vdots&\ddots&\vdots\\d''_{n1}&d''_{n2}&\cdots&d''_{nn}\end{bmatrix} \tag{7}$$

4.3 权重系数安全计算协议

为获取已知 h 条道路与未知的 $n-h$ 条道路之间的权重系数，通过已知 h 条道路的历史平均速度向量 $\boldsymbol{V}_i=\{v_i\mid i\in[1,h]\}$，完成对未知道路的历史平均速度向量 $\boldsymbol{V}_\tau=\{v_\tau\mid\tau\in[h+1,n]\}$ 的求解。具体步骤如下。

步骤 1 SP_1 计算向量 $\boldsymbol{V}_i$ 的期望 $\mathrm{E}(\boldsymbol{V}_i)$、$\mathrm{E}(\boldsymbol{V}_i^2)$ 和 $\mathrm{E}^2(\boldsymbol{V}_i)$，同理，$\mathrm{SP}_2$ 计算得到 $\boldsymbol{V}_\tau$ 的期望 $\mathrm{E}(\boldsymbol{V}_\tau)$、$\mathrm{E}(\boldsymbol{V}_\tau^2)$ 和 $\mathrm{E}^2(\boldsymbol{V}_\tau)$。

步骤 2 SP_1 求解 $\boldsymbol{V}_a'=\sqrt{\mathrm{E}(\boldsymbol{V}_i^2)-\mathrm{E}^2(\boldsymbol{V}_i)}$，$\mathrm{SP}_2$ 计

算 $\boldsymbol{V}_b'=\sqrt{\mathrm{E}(\boldsymbol{V}_\tau^2)-\mathrm{E}^2(\boldsymbol{V}_\tau)}$。

步骤 3　执行基础安全乘法 BSMP，即完成 $\boldsymbol{V}_{a1}+\boldsymbol{V}_{b1}\leftarrow\mathrm{BSMP}(\boldsymbol{V}_i:\boldsymbol{V}_\tau)$ 的映射关系。另外，计算 $\boldsymbol{V}_{a2}+\boldsymbol{V}_{b2}\leftarrow\mathrm{BSMP}(\boldsymbol{V}_a':\boldsymbol{V}_b')$。然后，$\mathrm{SP}_1$ 计算 $\boldsymbol{V}_{a1}$ 的期望值 $\mathrm{E}(\boldsymbol{V}_{a1})$，$\mathrm{SP}_2$ 计算 $\boldsymbol{V}_{b1}$ 的期望值 $\mathrm{E}(\boldsymbol{V}_{b1})$。

步骤 4　执行 $\boldsymbol{V}_{a3}+\boldsymbol{V}_{b3}\leftarrow\mathrm{BSMP}(E(\boldsymbol{V}_i):E(\boldsymbol{V}_\tau))$ 的计算后，SP_1 计算 $\boldsymbol{V}_{a4}=\boldsymbol{V}_{a1}-\boldsymbol{V}_{a3}$，$\mathrm{SP}_2$ 计算 $\boldsymbol{V}_{b4}=\boldsymbol{V}_{b1}-\boldsymbol{V}_{b3}$。

步骤 5　系统继续执行 BSDP，完成 $\boldsymbol{\Theta}'+\boldsymbol{\Theta}''\leftarrow\mathrm{BSDP}\big((\boldsymbol{V}_{a2},\boldsymbol{V}_{a4}):(\boldsymbol{V}_{b2},\boldsymbol{V}_{b4})\big)$ 的计算。

通过上述步骤可以获得已知的 h 个道路与未知道路之间的相关性强度向量 $\boldsymbol{\Theta}=\boldsymbol{\Theta}'+\boldsymbol{\Theta}''$，其中，$\boldsymbol{\Theta}'=\{\omega'_{(\tau,u)}\}$，$\boldsymbol{\Theta}''=\{\omega''_{(\tau,u)}\}$，其中，$\tau\in[1,n]$，$u\in[1,h]$，分别存储在云服务器 SP_1 和 SP_2。

4.4　道路监测安全计算协议

本文基于服务器 SP_1 存储的 $\boldsymbol{\Omega}'$、$\boldsymbol{\Theta}'$ 和 SP_2 存储的 $\boldsymbol{\Omega}''$、$\boldsymbol{\Theta}''$ 数据，实现式(3)所示的安全计算。假设预测道路 τ 的平均速度，选取用于预测道路 τ 的已知道路为 $i\in[1,K]$，通过如下步骤实现预测。

步骤 1　在确定 K 取值的前提下，从 $\boldsymbol{\Omega}'$、$\boldsymbol{\Theta}'$、$\boldsymbol{\Omega}''$ 和 $\boldsymbol{\Theta}''$ 中选取参与相关计算的数值。

步骤 2　通过选取的不同权重系数 $\omega'_{(\tau,i)}$ 和空间距离分量 $\boldsymbol{\Omega}''$ 中的元素 $d''_{(\tau,i)}$，智能交通系统执行 BSMP，有 $x_{(1,\tau,i)}+y_{(1,\tau,i)}\leftarrow\mathrm{BSMP}(\omega'_{(\tau,i)}:\frac{1}{d''_{(\tau,i)}})$ 的映射关系。SP_1 计算 $x'_{(1,\tau,i)}=\omega'_{(\tau,i)}d'_{(\tau,i)}$。然后选取权重系数 $\omega''_{(\tau,i)}$ 和距离分量 $\boldsymbol{\Omega}'$ 中的元素 $d'_{(\tau,i)}$，执行 $x_{(2,\tau,i)}+y_{(2,\tau,i)}\leftarrow\mathrm{BSMP}(\omega''_{(\tau,i)}:\frac{1}{d'_{(\tau,i)}})$，$\mathrm{SP}_2$ 计算 $y'_{(1,\tau,i)}=\omega''_{(\tau,i)}d''_{(\tau,i)}$。

步骤 3　SP_1 计算 $v_{(x,\tau,i)}=x_{(1,\tau,i)}+x'_{(1,\tau,i)}+x_{(2,\tau,i)}$，$\mathrm{SP}_2$ 计算 $v_{(y,\tau,i)}=y_{(1,\tau,i)}+y'_{(1,\tau,i)}+y_{(2,\tau,i)}$。

步骤 4　基于步骤 3 获取的 $v_{(x,\tau,i)}$ 和 $v_{(y,\tau,i)}$，智能交通系统继续执行 BSMP，得到 $\beta_{(1,\tau,i)}+\gamma_{(1,\tau,i)}\leftarrow\mathrm{BSMP}((v_{(x,\tau,i)},v_i'):(v_{(y,\tau,i)},v_i''))$ 的映射。

步骤 5　SP_1 计算 $v_a=\sum\beta_{(1,\tau,i)}$，$\mathrm{SP}_2$ 执行计算 $v_b=\sum\gamma_{(1,\tau,i)}$。

步骤 6　系统执行 BSDP，完成 $v'_\tau+v''_\tau\leftarrow\mathrm{BSDP}((\sum v_{(x,\tau,i)},v_a):(\sum v_{(y,\tau,i)},v_b))$。

经上述计算，最终未知道路 τ 在时间间隔 T 内道路的平均速度 v_τ 被随机地分为 v'_τ 和 v''_τ，并分别存储在 SP_1 和 SP_2 中。

通过执行设计的经纬度距离安全计算（SLD, secure length distance）协议、权重系数安全计算（SWF, secure weight factor）协议和道路监测安全计算（SMS, secure monitoring scheme）协议这 3 个安全计算协议，可以完成智能交通系统隐私保护道路状态实时监测的任务。在 IKNN 算法的基础上，本文首先利用 SLD 协议获取体现道路之间空间位置的空间相关性矩阵 $\boldsymbol{\Omega}$；然后根据 SWF 协议计算出路网中任意 2 条道路之间的 ω，并得到权重向量 $\boldsymbol{\Theta}$；最后在获取的空间相关性矩阵 $\boldsymbol{\Omega}$ 和权重向量 $\boldsymbol{\Theta}$ 的基础上，依据式(3)并按照 SMS 协议的执行过程即可实现隐私保护交通状态的实时监测。总而言之，依次执行 SLD、SWF 和 SMS 这 3 个安全计算协议，借助存储在云服务器中的交通数据即可实现整个智能交通监测的算法。因此，基于上述过程，可根据已知 h 条路段的平均速度，在保证数据隐私安全的前提下，实现对其余 $n-h$ 条未知道路的平均速度的预测。然后，将系统最终的处理分析结果提供给导航提供商。导航提供商将获取的数据直接进行相加，获取时间间隔 T 内整个路网所有道路的平均速度，并根据速度的相加结果，判断出道路交通真实的拥堵情况。最后，将分析出的真实拥堵状况实时地发布给机动车驾驶员。

5　安全性分析

5.1　数据存储安全性分析

根据本文的数据传输策略，SP_1 存储数据分量 x_1，SP_2 存储数据分量 x_2。假设 2 个攻击者 $\mathcal{A}_1$ 和 $\mathcal{A}_2$，$\mathcal{A}_1$ 对 SP_1 发起攻击并获取 SP_1 的内部数据，$\mathcal{A}_2$ 对 SP_2 发起攻击并获取 SP_2 的内部数据。由于云服务提供商 SP_1 和 SP_2 存储的数据仅仅是原始数据的分量，当成功攻击 SP_1 获取存储在其中的数据 x_1 后，由于 x_1 本身的随机性，$\mathcal{A}_1$ 并不能通过数据分量 x_1 对原始数据 x 做出正确的推断；同样地，当成功攻击 SP_2 获取存储在其中的数据 x_2 后，由于 x_2 同样具备很大的随机性，$\mathcal{A}_2$ 也不能由数据分量 x_2 对原始数据 x 做出正确的推断。

5.2　数据计算安全性分析

假设 $\mathcal{A}_1$ 对 SP_1 实施攻击并能够获取其内部存储的数据，并拦截云服务提供商 SP_1 和 SP_2 之间交互运算时所产生的中间值；$\mathcal{A}_2$ 同样对 SP_2 进行攻击，

并拦截两服务器之间数据交互计算所生成的中间值。当 SP_1 安全地接收可以作为密钥的随机数 r_a 时，在 r_a 的基础上将获取的原始数据分量 x_1 与之相加，即计算 $x_3 = x_1 + r_a$。当 x_3 被获取后，攻击者 $\mathcal{A}_1$ 很难根据 x_3 的值推算出 x_1 的值，因此 $\mathcal{A}_1$ 更无法对完整的原始数据 x 实现正确地推断。同理，$\mathcal{A}_2$ 也无法对执行上述操作的数据进行正确的推断。当云服务提供商 SP_1 和 SP_2 执行安全计算协议后得到最终的路段平均速度，该速度又被随机地分成了两部分并分别存储在 SP_1 和 SP_2。对于发送给导航提供商的最终数据也仅仅只是体现道路的拥堵情况，并不能推测出机动车的原始数据。本文所设计的 SLD 协议、SWF 协议和 SMS 协议均依照上述分析的安全交互过程进行计算，因此，本文所提 PPIM 算法可以保证数据的隐私安全不被泄露，进而完成对数据的一系列操作。

6 性能分析及实验验证

6.1 性能分析

计算复杂度。为评估本文所提 PPIM 算法的计算复杂度，先对算法中每个安全的子协议进行计算复杂度分析。假设 T_{BSMP}、T_{BSDP} 和 T_{Paillier} 分别表示 BSMP、BSDP 及 Paillier 算法执行一次所需的时间，具体数值通过仿真实验获得，如表 1 所示。由表 1 可知，$T_{\text{BSMP}} \approx T_{\text{BSDP}} \ll T_{\text{Paillier}}$。对于 SLD 协议，因其需要借助泰勒公式对反余弦函数近似求解，完成一次迭代的计算需要执行 9 次 BSMP，因此，执行输入数据大小为 n，计算复杂度为 $9O(n)T_{\text{BSMP}}$。对于 SWF 协议，系统完成一次完整协议计算需要进行 3 次 BSMP 计算及一次 BSDP 计算，执行输入数据大小为 n，SWF 协议所需时间为 $3O(n)T_{\text{BSMP}} + O(n)T_{\text{BSDP}}$。同样，对于完成道路监测协议 SMS 协议，其执行 BSMP 的次数与 K 的取值有关，并且完成一次 SMS 协议也需执行一次 BSDP。对于执行输入数据大小为 n，SMS 共需的时间为 $3KO(n)T_{\text{BSMP}} + O(n)T_{\text{BSDP}}$。因此，将 3 个过程的计算复杂度进行相加即可得到 PPIM 算法的计算复杂度，即 $(11+3K)O(n)T_{\text{BSMP}}$。假设数据的大小仍然为 n，采用 Paillier 算法完成对道路状态的监测，则需要的时间复杂度是 $nO(n^2)T_{\text{Paillier}}$。通过表 2，可以更加清晰地比较 SLD 协议、SWF 协议、SMS 协议、PPIM 算法及 Paillier 算法之间的计算复杂度。

表 1 协议执行一次所需时间

数据量/个	T_{BSMP} /s	T_{BSDP} /s	T_{Paillier} /s
500	0.012 2	0.012 0	126.950 0
1 000	0.023 9	0.023 8	254.696 0
1 500	0.047 8	0.041 9	380.596 0
2 000	0.056 3	0.055 9	507.666 0

表 2 协议/算法计算复杂度

协议/算法	计算复杂度
SLD 协议	$6O(n)T_{\text{BSMP}}$
SWF 协议	$3O(n)T_{\text{BSMP}} + O(n)T_{\text{BSDP}}$
SMS 协议	$3KO(n)T_{\text{BSMP}} + O(n)T_{\text{BSDP}}$
PPIM 算法	$(11+3K)O(n)T_{\text{BSMP}}$
Paillier 算法	$nO(n^2)T_{\text{Paillier}}$

通信开销。借助对每个子协议的分析，获取整个算法执行过程中的通信开销。表 3 给出了 SLD 协议、SWF 协议、SMS 协议及 PPIM 算法各自完成一次计算所需的数据通信轮数。通过分析基础协议 BSMP 和 BSDP，完成 BSMP 或者 BSDP 都需要 2 轮的通信开销。根据计算复杂度分析，完成一次 SLD 协议需要运行 6 次 BSMP，因此其需要进行数据交互的轮数是 12；对于 SWF 协议，其需要运行 3 次 BSMP 和 1 次 BSDP，所以共需 8 轮的数据交互；然而，执行 SMS 协议同样与 K 的取值有关，需要数据交互的轮数是 $6K$+2。因此，根据 3 个协议计算时交互的轮数，得出 PPIM 算法需要的通信开销，即交互 22+6K 轮。

表 3 执行协议/算法通信轮数

协议/算法	通信轮数/轮
SLD 协议	12
SWF 协议	8
SMS 协议	$6K$+2
PPIM 算法	22+6K

6.2 实验验证

基于真实的交通数据，本节通过实验对 PPIM 算法的性能进行验证。首先，通过未知数据预测误差衡量 IKNN 算法的有效性；然后，给出每一个安全计算协议及实现 PPIM 算法运行时间结果；最后，给出完成 PPIM 算法各阶段所需通信成本的仿真。为验证隐私保护下 IKNN 算法的有效性，本文用 e

表示对道路平均速度预测的误差率[19]，其计算如式(8)所示。

$$e=\frac{\|v_i-v\|_2^2}{\|v\|_2^2} \tag{8}$$

其中，$\|\cdot\|$ 表示计算范数，v_i 表示预测道路的平均速度，v 表示预测道路的真实平均速度。

实验的仿真数据为福州市 200 辆出租车的相关数据，包含出租车行驶的瞬时速度、GPS 经纬度位置信息、数据上报的时间及机动车的 ID 信息。仿真实验平台是在 RAM 为 16 GB，CPU 为 1.80 GHz 的 Intel Core(TM) i7-8550U 的笔记本电脑。本文将 K 的值设为 4，时间间隔 T 设为 5 min[20]。

在保证数据隐私安全的前提下，分别采用 KNN 与 IKNN 这 2 种算法估计未知数据的误差，曲线如图 4 所示。从 1 000 个数据中，依次随机在集合 $M=[200,300,\cdots,800]$ 选取不同数量的数据 M_l（l 表示集合 M 中元素位置的编号）作为已知数据。利用 M_l 个已知数据，估算其余没有被选取的 $1\,000-M_l$ 个数据。由图 4 可以得出，随着已知数据与数据总量比值的增大，无论是 KNN 算法还是 IKNN 算法，对未知数据估计的准确度降低了 6%～8%的误差。更重要的是，本文所提 IKNN 算法的曲线一直处于 KNN 算法曲线的下方，说明 IKNN 算法在实现道路状态监测上具有更好的准确度。

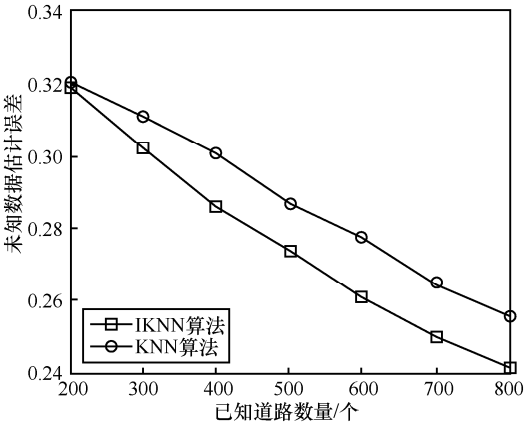

图 4　2 种算法估计未知数据估计曲线

本文所提 IKNN 算法与贝叶斯（Bayes）算法和支持向量机（SVM, support vector machine）算法这 2 种算法进行的交通监测预测准确度比较如图 5 所示。数据总量同样为 1 000，选择 200～800 范围，数据步长为 100。需要说明的是，SVM 算法采用的核函数是高斯核函数（RBF, radial basis function），正则化参数 δ=5，核函数参数 ε=0.01。由图 5 可知，IKNN、Bayes 和 SVM 这 3 种算法的未知数据预测误差都随已知数据量的增加出现了不同程度的下降。从曲线变化过程可知，IKNN 算法和 SVM 算法的预测精度均优于 Bayes 算法，IKNN 算法与 SVM 算法曲线尽管有部分相交，但随着数据量的增加，IKNN 算法的准确性更高。

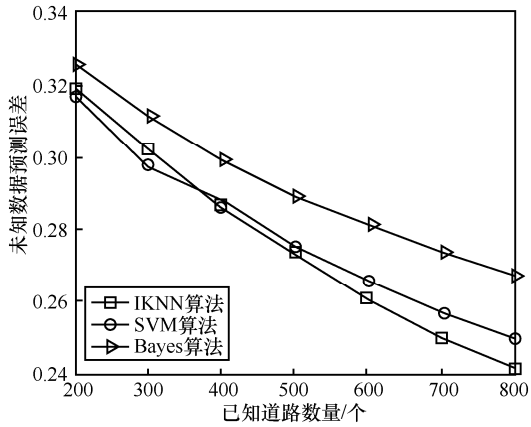

图 5　3 种算法交通监测预测准确度曲线

SWF、SLD 和 SMS 这 3 个协议在不同的数据量下的运行时间如图 6 所示。从图 6 可以看出，每个协议在加密不同量的数据量时，其效率都达到了毫秒级，并且 3 个协议之间运行时间存在的差异与 6.1 节的理论分析结果一致。

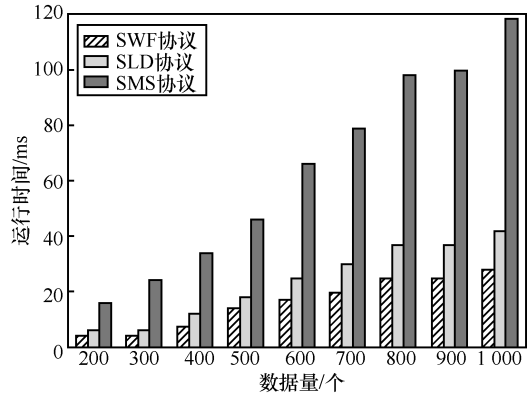

图 6　3 个协议的运行时间

为说明本文所提 PPIM 算法不仅完成了对道路状态的监测，而且具备安全、高效处理数据的优点，本文所提 PPIM 算法与 Paillier 算法进行了对比。Paillier 算法是加法同态加密算法，可以在不泄露数据隐私安全的前提下，实现对数据的安全处理。通过设置变量 N 及 2 个大素数 p 和 q，产生用于保护数据隐私的密钥，其整个过程通常分为密钥生成、

加密和解密 3 个阶段。

1) 密钥生成

首先，选取 2 个大素数 p 和 q，并使其满足 $\gcd(pq,(p-1)(q-1))=1$ 的要求。

其次，计算 $N=pq$，$\lambda=\mathrm{lcm}(p-1,q-1)$，并定义函数 $L(x)=\dfrac{x-1}{N}$。

最后，随机选取 g 且满足 $g<N^2$，计算 $u=((L(g^{\lambda}\bmod N^2))^{-1}\bmod)N$，可生成所需的公钥 (n,g) 和私钥 (λ,u)。

2) 加密

对需要加密的数据 m（$0<m<N$），并计算其密文 $c=g^m r^N \bmod N^2$。

3) 解密

基于密钥生成阶段的私钥 (λ,u)，计算出明文 $m=(L(c^{\lambda}\bmod N^2)u)\bmod N$。

需要说明的是，Paillier 算法满足系统加密的 2 个消息相乘的结果解密后恰好是 2 个消息明文状态相加的结果，这种方式与本文数据传输与处理策略相同。因此，将本文所提 PPIM 算法与 Paillier 算法进行对比，得到图 7 所示的仿真结果。需要说明的是，采用 Paillier 算法时，设置变量 N=2 048，并且生成 2 个 1 024 位的大素数 p 和 q。由图 7 可知，本文所提 PPIM 算法比 Paillier 算法具备更加快速地实现对道路状态监测的功能，更好地满足智能交通系统对低时延的要求。

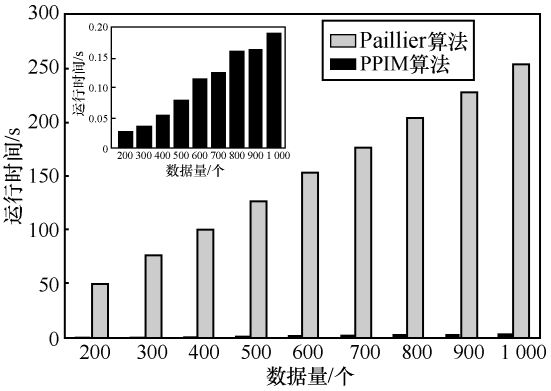

图 7 2 种算法的交通监测实现效率对比

本文用于衡量通信开销的单位是 bit，其测量依据是基于 IEEE 754 标准[21]。由该标准可知，任何以单精度浮点形式存储的值都需要 32 bit，双精度浮点则需要 64 bit，而真实的交通数据属于单精度浮点。各协议和算法所对应的通信量如图 8 所示。

由图 8 可知，实现整个 PPIM 算法，即使数据大小为 1 000 时，其数据的通信量还不足 0.2 MB，现有的网络带宽很容易满足其要求。

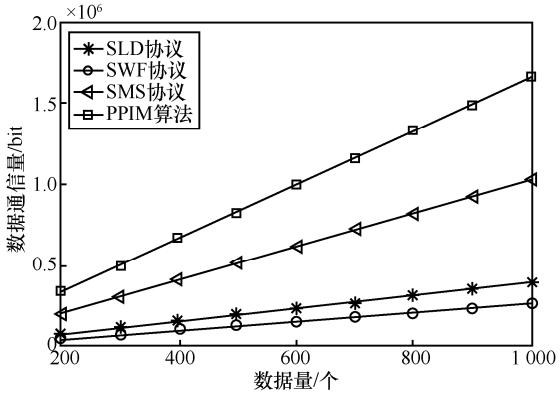

图 8 协议/算法的数据通信量

7 结束语

本文实现了机动车交通数据隐私保护的道路交通监测。为了优化 KNN 算法，提出了 IKNN 算法，通过引入道路之间的相似程度，对预测值进行有目的的调整，提高了交通监测的精度。为了提升数据处理的效率，还设计了适用于 IKNN 算法的安全计算协议。最后，通过真实数据的仿真实验证明，所提算法不但有效地实现了隐私保护下的智能交通道路监测，而且满足交通监测对实时性的需求。

参考文献：

[1] ARNOTT R, KENNETH S. The economics of traffic congestion[J]. American Scientist, 1994, 82(5): 446-455.

[2] KIM J, MEI-PO K. Beyond commuting: ignoring individuals activity-travel patterns may lead to inaccurate assessments of their exposure to traffic congestion[J]. International Journal of Environment Research and Public Health, 2019, 16(1): 89-109.

[3] MU S D, XIONG Z X, TIAN Y X. Intelligent traffic control system based on cloud computing and big data mining[J]. IEEE Transactions on Industrial Informatics, 2019, 15(12): 6583-6592.

[4] LATIF S, AFZAAL H, ZAFAR N A. Intelligent traffic monitoring and guidance system for smart city[C]//Mathematics and Engineering Technologies.[S.n.:s.l.], 2018: 1-6.

[5] WANG X B, LIU C, ZHU M L. Instant traveling companion discovery based on traffic monitoring streaming data[C]//IEEE Web Information Systems and Applications Conference. Piscataway: IEEE Press, 2016: 89-94.

[6] CELESTI A, GALLETTA A, CARNEVALE L, et al. An IoT cloud system for traffic monitoring and vehicular accidents prevention based on mobile sensor data processing[J]. IEEE Sensors Journal, 2017, 18(12): 4795-4802.

[7] WANG Y, ZHANG Y, PIAO X, et al. Traffic data reconstruction via

adaptive spatial temporal correlations[J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 20(4): 1531-1543.

[8] DATONDJI S R E, DUPUIS Y, SUBIRATS P, et al. A survey of vision-based traffic monitoring of road intersections[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(10): 2681-2698.

[9] JAIN N K, SAINI R K, MITTAL P. A review on traffic monitoring system techniques[C]//Soft Computing: Theories and Applications. Berlin: Springer, 2019: 569-577.

[10] PANG C C C, LAM W W L, YUNG N H C. A novel method for resolving vehicle occlusion in a monocular traffic-image sequence[J]. IEEE Transactions on Intelligent Transportation Systems, 2004, 5(3): 129-141.

[11] BROWN J W, OHRIMENKO O, TAMASSIA R. Privacy-preserving real-time traffic statistics[C]// Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. New York: ACM Press, 2013: 540-543.

[12] GISDAKIS S, MANOLOPOULOS V, TAO S, et al. Secure and privacy preserving smartphone-based traffic information systems[J]. IEEE Transactions on Intelligent Transportation Systems, 2014, 16(3): 1428-1438.

[13] HAHANOV V, ZHALILO A. Cloud driven traffic control: formal modeling and technical realization[C]//IEEE Mediterranean Conference on Embedded Computing. Piscataway: IEEE Press, 2015: 21-24.

[14] ZEGHID M, MACHHOUT M, KHRIJI L, BAGANNE A, et al. A modified AES based algorithm for image encryption[J]. International Journal of Computer Science and Engineering, 2007, 1(1): 70-75.

[15] BRISSAUD P O, FRANCCIS J, CHRISMENT I, et al, Transparent and service-agnostic monitoring of encrypted Web traffic[J]. IEEE Transactions on Networks and Service Management, 2019, 16(3): 842-856.

[16] LINDELL Y. Secure multiparty computation for privacy preserving data mining[C]//IEEE in Encyclopedia of Data Warehousing and Mining. Piscataway: IEEE Press, 2005: 1005-1009.

[17] BEN-DAVID A, NISAN N, PINKAS B. FairplayMP: a system for secure multi-party computation[C]//Proceedings of the 15th ACM Conference on Computer and Communications Security. New York: ACM Press, 2008: 257-266.

[18] DOU Z, CHEN X B, XU G, et al. An attempt at universal quantum secure multi-party computation with graph state[J]. Physica Scripta, 2020, 95(5): 55-106.

[19] LI J Y, GUO W Z, MA Z, et al. Privacy-preserving compressive sensing for traffic estimation[C]//IEEE Global Communications Conference. Piscataway: IEEE Press, 2019: 1-6.

[20] LI J Y, ZHENG H F, FENG X X, et al. Traffic estimation in road network via compressive sensing[C]//IEEE International Conference on Wireless Communications and Signal Processing. Piscataway: IEEE Press, 2017:1-6.

[21] KAHAN W. IEEE standard 754 for binary floating-point arithmetic[J]. Lecture Notes on the Status of IEEE 754, 1996, 5(11): 1-30.

[作者简介]

李家印（1990– ），男，山东济宁人，福州大学博士生，主要研究方向为移动数据采集、智慧城市、云外包数据存储、隐私保护、密文计算等。

郭文忠（1979– ），男，福建惠安人，博士，福州大学教授、博士生导师，主要研究方向为计算机智能及其在计算机网络中的应用等。

李小燕（1988– ），女，福建福州人，博士，福州大学讲师，主要研究方向为数据中心网络、网络安全、算法的设计与分析等。

刘西蒙（1988– ），男，陕西西安人，博士，福州大学研究员，主要研究方向为隐私计算、密文数据挖掘、大数据隐私保护、可搜索加密等。