doi:10.3969/j.issn.1003-2029.2021.01.002

能量最优下的铱星时序信号分析预测

叶宁祁[1]，官 晟[2,3,4,5]，王凤军[2]，丁军航[1*]，原明亭[1]
（1. 青岛大学 自动化学院，山东 青岛 266071；
2. 自然资源部第一海洋研究所，山东 青岛 266061；
3. 自然资源部海洋环境科学与数值模拟重点实验室，山东 青岛 266061；
4. 山东省海洋环境科学与数值模拟重点实验室，山东 青岛 266061；
5. 青岛海洋科学与技术试点国家实验室 区域海洋动力学与数值模拟功能实验室，山东 青岛 266237）

摘 要：铱星作为低轨道卫星，传输速率快，信息损耗率小，是大范围应用于海洋研究的通信卫星，其信号质量的可靠性对铱星通信终端的数据发送及能量利用率有重要影响，故对铱星信号的监测和分析至关重要。本文在有效监测铱星信号的基础上，使用滑动窗口算法对时序信号质量做总体评估，从而方便对数据进行预处理。并利用极端梯度提升算法（eXtreme Gradient Boosting, XGBoost）的学习预测功能对信号进行分类预测，判断信号是否存在干扰以择优数据发送方案。提高铱星通信终端能量的利用率，减少数据发送的丢失率。研究结果表明：预测模型 AUC（Area Under the Curve of ROC）面积为 0.74 ～ 0.78；预测精确率为 0.76 ～ 0.82；该预测模型可有效实现铱星通信终端设备的能量优化。

关键词：铱星；滑动窗口；极端梯度提升；分类预测；能量优化
中图分类号：TP301.6 **文献标识码**：A **文章编号**：1003-2029（2021）01-0010-07

铱星通信终端由于其全球覆盖的特性，广泛应用于海洋科学研究与野外军事作业[1-3]。由于作业环境的特殊性，铱星通信终端一般需要在自备能源的条件下进行长期不间断工作，故对能量的使用效率需求较高。另一方面，铱星信号的可靠性会直接影响铱星通信终端的通信性能。

铱星是早期发展起来的通信卫星，已有众多学者对铱星的通信信道[4]、信息传输速率[5]、机会信号定位[6]等做了多方面研究，但是由铱星通信终端所处环境造成的信号干扰研究较少。海洋环境下的铱星信号质量干扰因素包括电离层干扰、雨衰、雪衰、日凌和星蚀，所以铱星信号质量数据具有高度的复杂性及随机性[7]。传统铱星信号监测方式是通过测取全天环境信号以判别干扰存在的时期，避免在干扰期向用户端发送数据包。这种方法具有时效慢与能耗大的缺点，而采用分类预测方法可以很好地解决这两类缺点。分类预测方法主要分为传统统计学方法、机器学习和深度学习方法三类。传统统计学方法采用线性拟合预测趋势，此类方法通常无法描述变量的随机性和非线性关系；而深度学习方法用于数据样本量较大的情况，模型收敛速度慢，网络调参复杂，容易陷入局部特征的极小化[8]。机器学习算法的代表XGBoost既可用于铱星时序数据的回归预测，

收稿日期：2020-09-15
基金项目：科技部重大科学仪器专项资助项目（2018YFF01014100）
作者简介：叶宁祁 (1995-)，男，硕士研究生，主要从事智能控制方面研究。E-mail: 885258569@qq.com
通讯作者：丁军航 (1979-)，男，博士，副教授，主要从事智能控制方面研究。E-mail: dingjunhang@163.com

也可用于分类预测，且模型参数少，收敛及预测速度快，对于实时性的铱星信号数据预测具有较高的适用性。

综上所述，文中提出利用滑动窗口算法对信号质量做移动平均整体评估。根据信号质量的评估特性对数据做不同处理，之后采用两种模型评价指标，对模型效果进行评价。0 至 5 共 6 个类别同时看做是连续值，将信号质量大于等于 4 判别为优质信号，信号质量小于 4 判别为干扰，先使用 XGBoost 模型回归预测，计算预测值和真实值之间的平均绝对值误差（Mean Absolute Error, MAE）作为预测偏差的参考。然后使用综合分类模型评价指标接收者操作特征曲线（Receiver Operating Characteristic curve，ROC）和混淆矩阵来衡量模型效果，计算预测模型的准确率、精准率和召回率。基于以上研究结果给出合理的信号采集预测及数据发送控制方案，实现铱星通信终端的高效能量利用率。

1　铱星信号质量采集系统及数据获取

为了能够对铱星信号质量进行分析，首先使用铱星监测平台，通过高增益天线来采集铱星信号质量数据。然后编写铱星通信上位机程序，对铱星信号数据进行接收及存储。铱星监测平台是基于铱星 9602 模块设计的。

1.1　铱星信号质量采集系统

9602 模块具有数据传输、定位跟踪等功能，波特率为 19 200 bps。默认 9 线制串行接口，但只使用传输、接收、接地三线，用户输入 AT 指令以获取模块接收到的数据 [9]。

使用 Labview 编写上位机程序， Labview 作为模块化编程语言，编程效率高，其编程主要分为前控制面板编程和主程序面板编程。主程序面板中主要分为以下几个模块：串口配置模块、指令发送模块、数据采集模块、数据存储模块和图表显示模块。上位机程序框图如图 1 所示。

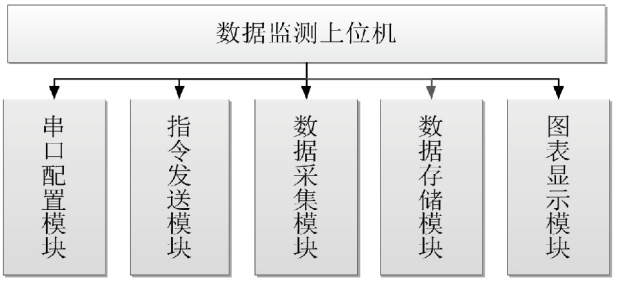


图 1　数据监测上位机框图

1.2　铱星信号质量数据获取实验

在任何开放环境下，周围不同地貌地物形成各自独特的通信环境，对铱星信号收发造成不同形式的干扰，故需要选择不同测量环境。通过实验研究共选样两处实验环境，一处在城市楼顶宽阔地带，周围有高层建筑遮挡，另一处在周围遮挡物较少的海港船舶上。

安装铱星信号监测平台，启动上位机系统，配置串口，波特率设置为 19 200 bit/s。通过指令 A1 启动铱星模块，发送指令 C2 给信号处理板采集数据。由于指令发送及信号采集的延时性，故最短采集间隔为 30 s 采集 1 次，共连续采集 3 d。

按照同样的方式在不同实验环境下分别进行监测。其中采集生成的数据传输给数据存储模块的同时通过图表显示，以便在分析数据前能直观观测出数据整体情况，防止出现大规模数据丢失。

2　铱星信号质量数据分析算法

采集得到的铱星信号质量分为 6 个类别，分别为 0 至 5，0 表示信号质量最弱，5 表示信号最强。其中 4 和 5 可作为数据收发的优良信号质量，数据流形式为时间序列数据流 [10]。目前对时间序列数据分析的算法有很多。经实验数据特性研究对两处环境下采集的信号做移动平均，之后将信

号的类别由六类转换为两类，并通过 XGBoost 的分类预测功能对信号干扰进行预测。

2.1 滑动窗口算法

滑动窗口算法原理如图 2 所示。算法根据指定时间单位长度框住时间序列，在时间序列上做单位平移，并在每次平移前对当前所框住的时间序列对应的值做计算。计算的指标为均值，由于数据采集间隔为 30 s，窗口过大，观察信号质量值容易错过数值分布规律，而窗口过小则样本量较大，不易于直观遍览，故选择窗口大小为 5min 数据，步长为 1。对于大数据而言，遍览整个数据的时间耗费巨大，而滑动窗口算法可以很好地解决这类问题。

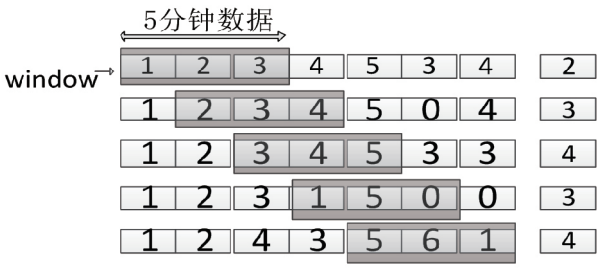


图 2 滑动窗口算法原理图

2.2 XGBoost 分类预测算法

XGBoost 是基于梯度提升决策树（Gradient Boosting Decision Tree，GBDT）改进得出的算法，于 2016 年提出后便受到了广泛关注。其核心思想从决策树中演变而来。相比于其它预测算法，XGBoost 采用多线程并行处理，对于线性以及非线性数据都有很好的处理效果 [11]。

决策树生成的每个分类树是独立的，而 XGBoost 是生成多个弱预测模型，将多个预测结果累积得到强预测模型 [12]。即：

$$Obj^{(t)}=\sum_{i=1}^{n} l\left(y_i,\ \hat{y}_i^{(t-1)}+f_t\left(x_i\right)\right)+\Omega\left(f_t\right) \qquad (1)$$

式中：$\hat{y}_i^{(t-1)}$ 为模型当前的输出结果；y_i 为真实结果；$\Omega(ft)$ 为正则化。通过损失函数揭示训练误差，通过正则化定义复杂度，值越小复杂度越低，泛化能力越强。

2.2.1 高集成度 相对于 GBDT 来说，XGBoost 对损失函数一阶展开的基础上进行二阶展开，并加入代价函数 Ω（ft）的正则化。当树的深度过高时设置 max_depth 停止建设决策树，防止过拟合现象，相比于其它模型具有高集成度 [13]，降低了模型的变化幅度。

2.2.2 内置缺失值处理 XGBoost 内置缺失值处理，可在不同结点遇到缺失值时采用不同的处理方法，不用手工添加弥补缺失值，在模型处理上大大简化了工作量。

2.2.3 最优迭代次数 XGBoost 允许 boosting 迭代中使用交叉验证，因此，可以方便地获取到最优迭代次数。

2.3 预测模型评估指标 *MAE*

预测模型评估指标 MAE 的表达式为：

$$MAE = \frac{1}{N}\sum_{i=1}^{n} \left|\hat{y}_i - y_i\right| \qquad (2)$$

式中：$\hat{y}_i$ 为预测结果；y_i 为真实结果。当预测值越接近真实值，则 *MAE* 越小，预测模型质量越好。反之当误差越大，*MAE* 越大，预测模型质量也就越差。使用平均误差绝对值计算偏差作为预测模型参考。

2.4 预测模型评价指标

如表 1 所示利用混淆矩阵达到预测结果的可视化效果，通过混淆矩阵计算得出 ROC 曲线。并计算综合预测指标，对模型进行综合评估。

表 1 二分类混淆矩阵

	预测值	
肯定类别	TP（True Positive Rate）	FP（False Positive Rate）
否定类别	FN（False Negative Rate）	TN（True Negative Rate）

精确率（Accuracy, *ACC*）=

（*TP*+*TN*）/(*TP*+*TN*+*FP*+*FN*) (3)

精准率（Precision, *PC*）=*TP*/（*TP*+*FP*） (4)

召回率（Recall, *RC*）=*TP*/（*TP*+*FN*） (5)

ROC 曲线的纵轴：真阳性率（True Positive Rate）；ROC 曲线的横轴：假阳性率（False Positive Rate）。ROC 左下角覆盖面积 *AUC* 代表模型分类准确性，当 *AUC*>0.5 的情况下越接近于 1 表示分类准确性越高。

3　实验数据分析及预测模型评估

3.1　信号质量初步评估

使用时间序列滑动窗口算法对采集得到的两组数据进行分析，绘制得到两组数据的信号质量评估图。信号质量评估图如图 3 所示。

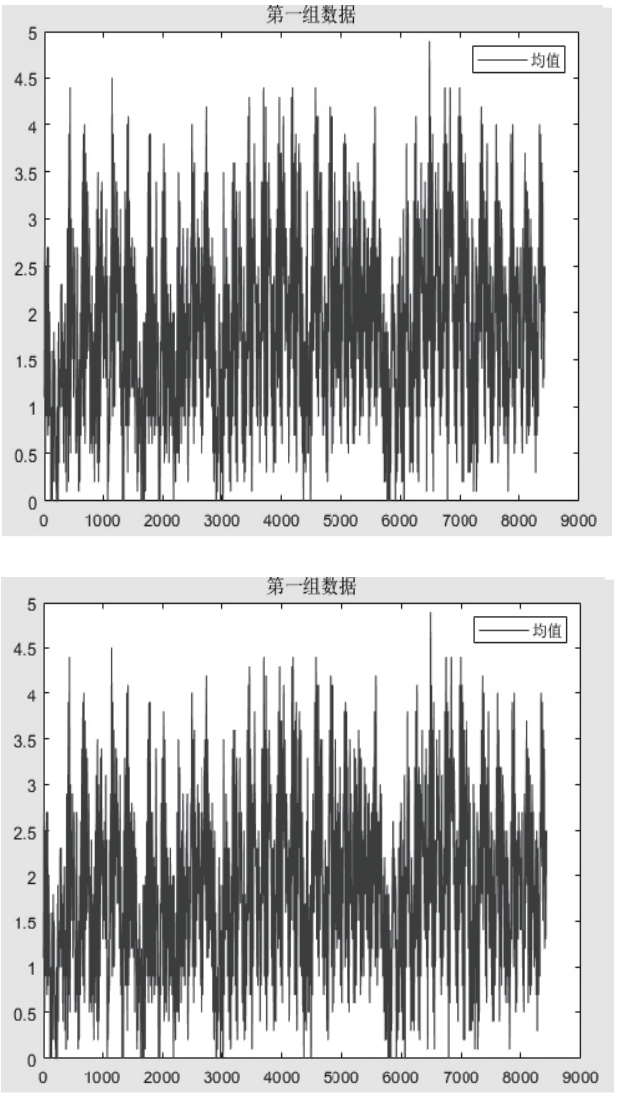


图 3　原始信号质量滑动平均图

图 3 中，第一组数据在城市楼顶环境测量得到，第二组数据在海港环境测量得到。根据第一组信号质量评估图可得，其信号质量长时间处于 4 以下，干扰较多，信号质量恶劣。第二组信号质量评估图信号良好，局部出现干扰。

3.2　数据预处理

通过两处环境信号质量的评估可以发现，不同环境下信号干扰因素不同，由此可得出在不同环境下建立不同预测分类器的重要性。将楼顶环境数据放在工作簿 1 中作为第一组训练样本，取数据列最后 200 个数据样本作为测试样本。海港环境数据放在工作簿 2 中作为第二组训练样本，测试数据样本取样方式同上。再将楼顶数据和海港数据合并放在工作簿 3 中作为第三组训练样本，测试数据样本取样方式同上。分别通过 3 组数据对训练模型进行参数调优。

两地测量数据中含有大于 5 及字母类异常值，为了不影响模型预测的准确率，需将异常值剔除。另外，4 ～ 5 作为优良信号类，4 以下作为劣质信号类，将六类别信号转化为二类别信号，构建二分类预测模型。

3.3　模型参数调优

通过实验验证，max_depth 和 learning_rate 两个参数对模型预测效果影响较大，所以本研究主要针对这两个参数进行调优。

3.3.1　*参数 max_depth 调优*　参数 max_depth 为树的最大深度，系统默认值为 6，典型值为 3 ～ 10。max_depth 的主要功能用于控制过拟合，max_depth 值越大，模型更注重于局部样本，学习更具体。为了选择最优参数，将参数取值范围选为 3 ～ 13，并分别训练及测试观察 *ACC* 和 *AUC* 的变化。

如图 4 所示，max_depth 逐渐增大，模型预测效果逐步提升，当到达一定深度时，预测效果开始下滑。

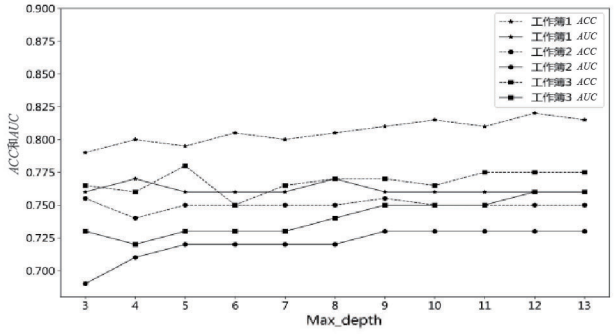


图 4 max_depth 参数调优

3.3.2 *参数 learning_rate 调优* 参数 learning_rate 为学习率，系统默认值为 0.3，典型值为 0.01 ～ 0.2。其功能主要用于减少每一步分类的权重，提高模型预测的鲁棒性。同参数 max_depth 调优方法一样取值范围选为 0.01 ～ 0.2，观察 *ACC* 和 *AUC* 的变化。

如图 5 所示，当 learning_rate 值慢慢增大时，模型预测效果有一定提升，到达峰值后逐步下滑。依据以上研究，得出不同数据模型的最优参数表，如表 2 所示。

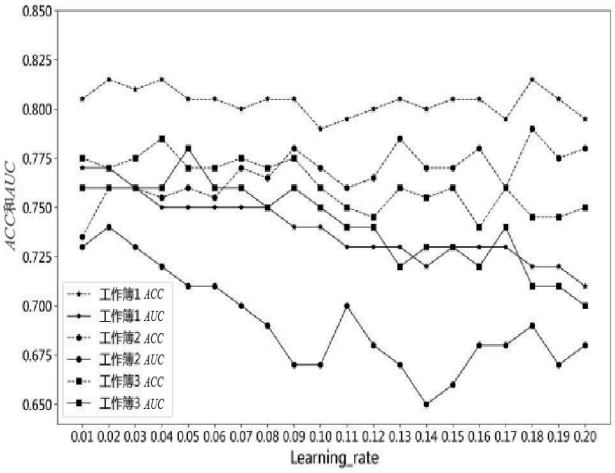


图 5 learning_rate 参数调优

表 2 训练模型最优参数表

最优参数	max_depth	learning_rate
工作簿 1	8	0.02
工作簿 2	9	0.02
工作簿 3	12	0.05

3.4 最优参数下预测结果评估

三组数据的 ROC 曲线图如图 6 所示。工作簿 1 中 *MAE* 为 1.2，混淆矩阵为 [[24 25] [12 139]]，*AUC* 面积为 0.77，预测分类精确度高但是平均误差相对较大。工作簿 2 中 *MAE* 为 0.76，混淆矩阵为 [[140 12] [36 12]]，*AUC* 面积为 0.74，预测分类精确度相对较低，但是平均误差小。工作簿 3 中 *MAE* 为 0.82，混淆矩阵为 [[131 21] [25 23]]，*AUC* 面积为 0.78。由于训练数据类别均衡，综合预测分类效果较为均衡，综合预测评价表如表 3 所示。

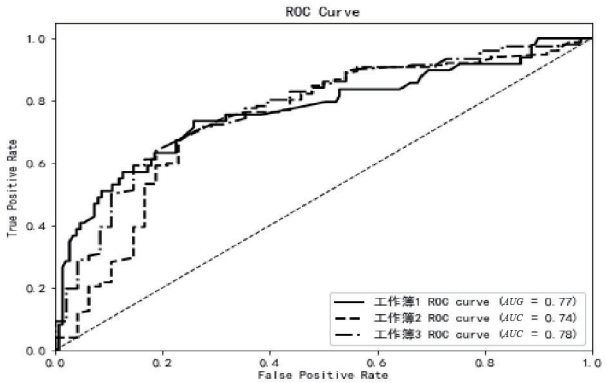


图 6 ROC 曲线图

表 3 调优前后综合预测效果表

		精确率	精准率	召回率
参数调优前	工作簿 1	0.79	0.44	0.59
	工作簿 2	0.74	0.91	0.78
	工作簿 3	0.75	0.84	0.82
参数调优后	工作簿 1	0.82	0.49	0.67
	工作簿 2	0.76	0.92	0.80
	工作簿 3	0.77	0.86	0.83

4 系统方案与能量优化

4.1 系统监测方案

铱星通信终端可基于不同研究方式采取不同监测方案与控制方案，目前铱星通信终端在海洋当中的应用主要分为固定监测和流动监测。当监测方式为流动监测时，系统方案：按日期推移，启动时间修订，同时检测当前环境下铱星信号质量，预测接下来信号质量优劣，选择是否发送数

据包。当监测方式为固定监测时，系统方案：启动系统后检测一天信号质量，预测第二天数据发送时间点的同时进行数据采样。具体方案流程图如图 7 所示。

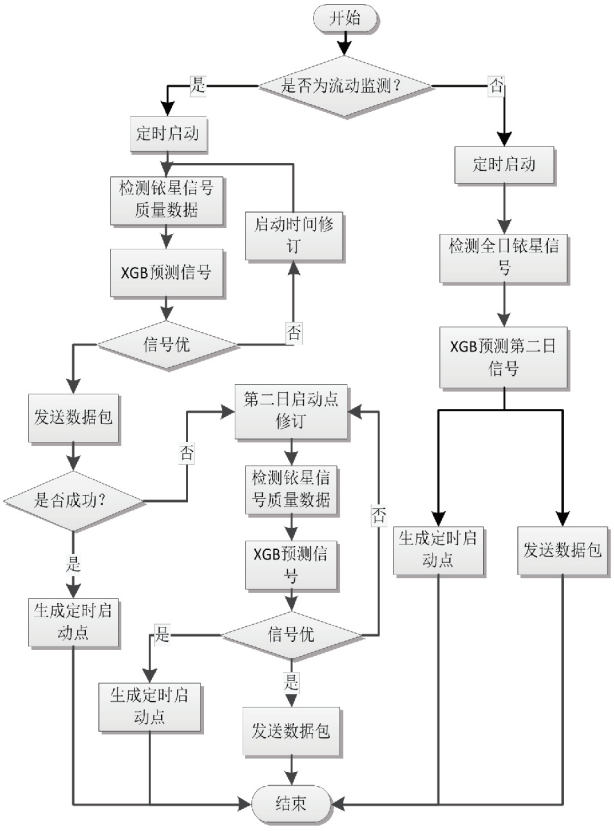


图 7　监测方案流程图

4.2　能量优化计算

铱星通信终端发送数据方式为 2 min 发送一包数据，启动一次等于发送 5 包数据。按照预测模型的召回率，假设在信号质量劣的环境下发送 100 包数据，召回率为 0.67，总共需要发送 149 次能够将数据包发送完。未对信号进行预测，传统数据发送方式的发送次数为 408 次。在信号优的环境下发送 100 包数据，召回率为 0.80，需要发送 125 次。同上，未预测信号的发送次数为 131 次。铱星通信终端功率参数如表 4 所示。

表 4　铱星通信终端耗电参数表

功率参数	均峰值	值
电源输入电压范围		5.0 V DC+/-0.5 V
发射电流	峰值	1 000 mA
发射电流	平均值 *	190 mA

劣信号质量环境下传统数据发送方式所消耗的能量为：

$$W=\frac{U*I}{1000}*T=\frac{5\text{V}*0.19\text{ A}}{1000}*\frac{826\text{ m}}{60}\approx 0.013\text{ kW}\cdot\text{h} \quad (6)$$

进行信号预测能量消耗为：

$$W=\frac{U*I}{1000}*T=\frac{5\text{V}*0.19\text{ A}}{1000}*\frac{308\text{ m}}{60}\approx 0.0048\text{ kW}\cdot\text{h} \quad (7)$$

优信号质量环境下传统发送数据方式所消耗的能量为：

$$W=\frac{U*I}{1000}*T=\frac{5\text{V}*0.19\text{ A}}{1000}*\frac{272\text{ m}}{60}\approx 0.0043\text{ kW}\cdot\text{h} \quad (8)$$

进行信号预测能量消耗为：

$$W=\frac{U*I}{1000}*T=\frac{5\text{V}*0.19\text{ A}}{1000}*\frac{260\text{ m}}{60}\approx 0.0041\text{ kW}\cdot\text{h} \quad (9)$$

根据以上计算可以得出，信号劣的环境下节能 63%，信号优环境下节能 4.6%。

5　结　论

本文针对研究生产中铱星信号出现干扰的情况，提出对铱星信号干扰预测的研究。使用 Labview 语言编写上位机程序对铱星信号进行全天检测，获取铱星信号质量数据。实验证明选择不同环境测量信号的重要性。

根据预测模型的预测指标，计算得出不同环境指标下的能量优化范围为 4.6% ～ 63%，证明该预测模型在具有有效性和可行性的基础上，铱星信号越好的情况下能量优化效率越低，反之能量优化效率越高。预测模型预测效果还有待提高，下一步需对 *ACC* 及 *AUC* 两个指标继续研究，引入更多的分类预测算法进行相关验证，提高分类预测综合效果。

参考文献 :

[1] 王力权 , 胡东伟 . 低轨卫星移动通信系统的星上路由和星间链设计 [J]. 信息通信 ,2016(2):200-201.
[2] 杨宏 . 铱星卫星移动通信系统的关键技术研究 [D]. 重庆 : 重庆大学 ,2013.
[3] 陈质二 , 俞建成 , 张艾群 . 面向海洋观测的长续航力移动自主观测平台发展现状与展望 [J]. 海洋技术学报 ,2016,35(1):122-130.
[4] 汤辉 , 邹钦羊 , 朱立东 , 等 . 卫星通信系统多优先级信道预留分配策略 [J]. 太赫兹科学与电子信息报 ,2019,17(5):765-770.
[5] 余红宴 . 无线携能通信的能量效率优化与物理层安全研究 [D]. 重庆 : 西南大学 ,2017.
[6] 孔佑迪 , 门雅彬 , 齐占辉 , 等 . 基于铱星通信的海洋环境在线监测系统 [J]. 电子技术应用 ,2015,41(8):27-29,33.
[7] 韩晋平 , 许靖 . 一种卫星通信地球站收发通道干扰预测方法 [J]. 数字技术与应用 ,2019,37(11):23-24.
[8] 张永梅 , 陈惠妮 , 张奕 . 基于 XGBoost 的雾霾预测方法 [J]. 计算机工程与设计 ,2019,40(12):3631-3638.
[9] 赖森华 . 基于 LabVIEW 的地铁车载试验数据处理方法研究 [J]. 工业仪表与自动化装置 ,2017(6):104-107.
[10] 闫朝升 , 张承江 , 马英 . 基于滑动窗口的时间序列数据流分析与预测技术研究 [J]. 黑龙江大学自然科学学报 ,2006,23(6):863-867.
[11] Sun, Fei,Wang, Run,Wan, Bo, et al.Efficiency of extreme gradient boosting for imbalanced land cover classification using an extended margin and disagreement performance[J].ISPRS International Journal of Geo-Information,2019,8(7):315.
[12] 陆健健 , 江开忠 . 基于 XGBoost 算法模型的金融客户信用评估研究 [J]. 软件导刊 ,2019,18(4):133-136.
[13] 王国田 , 戴筠一 , 许少钧 , 等 . 基于 XGBoost 算法的实验室安全风险预测模型研究 [J]. 实验技术与管理 ,2019,36(12):245-251.

Analysis and Prediction of Iridium Time Series Signal with Optimal Energy

YE Ning-qi[1], GUAN Sheng[2,3,4,5], WANG Feng-jun[2], DING Jun-hang[1], YUAN Ming-ting[1]

1. *College of Automation, Qingdao University, Qingdao* 266071,*China*;
2. *First Institute of Oceanography, MNR, Qingdao* 266061, *China*;
3. *Key Laboratory of Marine Environmental Science and Numerical Simulation, MNR, Qingdao* 266061, *China*;
4. *Key Laboratory of Marine Environment Science and Numerical Simulation in Shandong Province, Qingdao* 266061, *China*;
5. *Qingdao National Laboratory of Marine Science and Technology, Regional Ocean Dynamics and Numerical Simulation Functional Laboratory, Qingdao* 266237, *China*

Abstract: Iridium satellite, as a low orbit communication satellite with high transmission rate and low information loss rate, has been widely used in marine research. The reliability of its signal quality has an important impact on the data transmission and energy utilization rate of Iridium satellite communication terminal, so the monitoring and analysis of Iridium satellite signal is of top priority. In this paper, based on the effective monitoring of iridium signal, the sliding window algorithm is used to evaluate the quality of timing signal, so as to facilitate data preprocessing. The learning and prediction function of extreme gradient boosting (XGBoost) is used to classify and predict the signals and judge whether there is interference in the signals to select the optimal data transmission scheme. It is suggested to improve the energy utilization of iridium communication terminal and reduce the loss rate of data transmission. The results show that the area of AUC (area under the curve of ROC) is 0.74~0.78; the prediction accuracy is 0.76~0.82; the prediction model can effectively realize the energy optimization of iridium communication terminal equipment.

Key words: iridium star; sliding window; extreme gradient rise; classification prediction; energy optimization