

高负载数据大批量接入后的网关状态预测模型构建

敖知琪, 陈禹旭, 任昊文, 康旖, 代昊琦

(南方电网数字电网集团有限公司, 广东 广州 510700)

摘要: 高负载数据大批量接入后网关信道的容量受数据传输速率影响, 导致网关状态预测不准确。构建高负载数据大批量接入后的网关状态预测模型。利用概率响应函数, 计算各个网关节点的平均传输距离, 构建高负载数据传输模型。利用接入信号矢量与发射矢量的相对熵, 分析网关信道容量。利用时间重复性原理, 计算网关状态的预测概率, 采用基于单周期、多周期综合的多阶段预测模式, 实现网关状态预测模型的构建。实验结果表明, 所构建模型在接入时长为 43 s 时可以迅速提高网关数据的转换率, 并将网关信道容量控制在 85 bit/s 以上。

关键词: 网关状态; 数据接入; 预测模型; 数据传输

中图分类号: TP18; TP393 文献标识码: A 文章编号: 1003-7241(2025)03-0084-04

Construction of Gateway State Prediction Model after Mass Access of High-load Data

AO Zhi-qi, CHEN Yu-xu, REN Hao-wen, KANG Yi, DAI Hao-qi

(China Southern Power Grid Digital Grid Group Co., Ltd., Guangzhou 510700 China)

Abstract: The capacity of the gateway channel is affected by the data transmission rate after the high load data is accessed in large quantities, resulting in inaccurate gateway state prediction. The prediction model of gateway state after high load data mass access is constructed in this paper. The probability response function is used to calculate the average transmission distance of each gateway node and build a high load data transmission model. The relative entropy of the access signal vector and the transmission vector is used to analyze the gateway channel capacity. Based on the principle of time repeatability, the prediction probability of the gateway state is calculated, and a multi-stage prediction model based on single cycle and multi cycle synthesis is adopted to build the gateway state prediction model. The experimental results show that the model in this paper can rapidly improve the conversion rate of gateway data when the access time is 43 s, and control the gateway channel capacity above 85 bit/s.

Keywords: gateway status; data access; prediction model; data transmission

0 引言

在物联网领域, 传感器网络具有广泛的互通性和广域的特性, 需要通过公用网络将传感器的数据与企业以及各个政府部门的数据连接, 因此, 物联网终端的安全性十分重要且不容忽视^[1-3]。为了保证在传统网络和物联网中实现数据的安全传输, 必须预测接入高负载数据后的网关状态。

姚玉坤等人^[4]根据无人机与地面控制台之间的连接情况, 提出了一种新的网关选取算法, 在此基础上, 根据对无人机的活动模式和飞行动向进行预测, 选取与地面控制台相连的长节点为网关的候选节点, 采用功能定时的方法, 选取最大的节点作为网关节点, 并通过仿真模拟

实验, 证明了无人机分布式网关算法可以有效提高网关选取的效率, 使无人机与地面网络传输的时长变短, 并提高传输成功率。赵丙镇等人^[5]在现有认证机制的基础上, 设计了一种分布式动态网关访问验证的体系结构, 该方法在主网关的状态下, 利用区块链协议的一致性, 将访问验证中所包含的装置ID等数字信息保存在一个区块链的数据存储中, 增强了身份验证的安全性。为了验证该方法的可行性, 对其进行理论分析和实例验证, 结果表明, 该方法可以有效提高动态网关访问验证的时效性, 可以有效地防止物联网常见的攻击方式, 该认证机制可以提高物联网设备的安全性。Jrgensen A A 等人^[6]提出基于光纤通信提升数据传输能力的方法, 通过实验验证该方法可产生稳定的波长信道, 实现波长复用数据传输, 保证数据传输的稳定性。

基于以上研究背景, 本文在大批量接入高负载数据的情况下, 构建网关状态预测模型, 保证网络运行的稳定性。

*基金项目: 南方电网数字电网集团有限公司“南网云平台(三期)建设项目”(JY-KF-03-YP-21-005-TQ); 广东省科技项目(202000320001)

收稿日期: 2023-11-24

1 网关状态预测模型构建方法

1.1 高负载数据传输建模

物联网内接入大量的高负载数据造成网关状态拥塞,此时对高负载数据传输建模,保证高负载数据大批量接入后网关状态的稳定性。将物联网 W_i 的响应函数定义为 $f(x)$,利用下式得到高负载数据接入的概率响应函数:

$$f_{x(i)} = g(x) \cdot f(x) [1 - f(x)] \quad (1)$$

式(1)中, $g(x)$ 为物联网 W_i 的概率响应函数。

利用高负载数据接入的概率响应函数,计算高负载数据大批量接入后各个节点的平均传输距离:

$$\bar{d}_k = \bar{x}_{(k)} - \bar{x}_{(1)} \quad (2)$$

式(2)中, $\bar{x}_{(k)}$ 代表物联网中高负载数据大批量接入后各节点第 k 次的传输路径, $\bar{x}_{(1)}$ 代表物联网中高负载数据接入后各节点第 1 次传输路径。

当高负载数据批量接入后,网关节点的速度为 V ,计算物联网中高负载数据节点大批量接入后,完成 k 个大批量接入目标所需要的时间:

$$\bar{t}_k = \frac{\bar{d}_k}{V} \quad (3)$$

从 N 个高负载数据节点中,提取出 k 个参与批量接入的节点函数 $h(k)$ 和高负载数据接入节点分量 $s(k)$,假设,高负载数据有 k 个接入节点时,这个节点为高负载数据大批量接入 a 的目标节点概率^[7],可以通过下式计算:

$$Q(k) = \frac{s(k)}{h(k)} \quad (4)$$

在得到 $s(k)$ 和 $Q(k)$ 后,得到高负载数据大批量接入的等待时间:

$$T = t + \delta_d^2 \quad (5)$$

式(5)中, t 为大批量高负载数据接入的时间, δ_d^2 代表高负载数据大批量接入的时间方差。

为了计算高负载数据大批量接入网关后的边界值,将 T 代入下列不等式中^[8],得到高负载数据传输概率为:

$$p[T|\phi] = 1 - \frac{\sigma_s^2}{(\phi - T)} \quad (6)$$

式(6)中, ϕ 代表高负载数据的接入系数, σ_s^2 代表高负载数据大批量接入网关时单个数据等待的时间方差。

根据高负载数据传输概率,构建物联网中高负载数据安全传输模型,表示为:

$$Y = p[T|\phi] - \frac{\sigma_s^2}{(\phi - T)^2} T \quad (7)$$

根据以上过程,构建了高负载数据传输模型。

1.2 网关信道容量

假设,高负载数据大批量接入网关信道为 φ_k ,得到 K 个批量接入数据库 $\varphi_1, \varphi_2, \dots, \varphi_K$,响应函数为 $p(y|\varphi_k)$,其

中, y 表示大批量接入的高负载数据, i 为接入的高负载数据数量, n 为接入的高负载数据总数量。假如高负载数据的批量接入概率由网关信道的容量决定,则前一个网关信道接入和输出关系相互独立。其表达式为:

$$o = \prod_{i=1}^n p(y_i | \varphi_{1i}, \varphi_{2i}, \dots, \varphi_{Ki}) \quad (8)$$

将网关信道设置成一个高斯信道,在该信道中,每个数据传输矩阵 C 均为一个独立的共同分布点^[9]。如果接入端信道的具体情况不明,则平均分配每个天线上的接入量。得到接入信号的协方差矩阵公式为:

$$R_C = \frac{E_C}{W_N} W_T \quad (9)$$

式(7)中, E_C 代表批量接入高负载数据的总能量, W_T 代表接入矩阵, W_N 为网关信道矩阵。

在实际网关信道传输的条件下,信道衰减因子是一个随机性变量,仅达到平均传输能力^[10]。获得信道能力的平均接入向量,公式为:

$$\chi = R_C^T + \xi \quad (10)$$

式(10)中, ξ 代表信道的衰减因子。

假设信道参数不受其他因素影响,数据传输信道的容量为:

$$C = N(g)_{\max} \Lambda(g, q) \quad (11)$$

式(11)中, $N(g)$ 是接入信号 g 的密度分布, $\Lambda(g, q)$ 是接入信号 g 与发射信号矢量 q 的关联性。通过改变 g 的密度分布状态,使接入矢量最大化,从而实现网关信道容量的最大化^[11-12]。利用信息熵,得出:

$$\Lambda(g, q) = H(q)^T - H(q, g)^T \quad (12)$$

式(12)中, $H(q)$ 代表接入矢量 q 的信息熵, $H(q, g)$ 代表接入矢量 g 为 q 的条件熵。接入矢量 g 和发射矢量 y 二者紧密相连,且相对独立。

对于网关信道的协方差矩阵 γ_q ,其中的接入矢量为 z ,假设 z 作为最大矢量时,熵 $H(q)$ 的值实现最大化。同时,也可以说明矢量 g 是高斯矢量,由 R_C 决定其具体的分布状态。接入信号矢量 q 与发射矢量 y 的相对熵表达式为:

$$\begin{cases} H(q) = \log_2 \det(\pi \gamma_q) \\ H(y) = \log_2 \det(\pi \sigma_s^2 W_T) \end{cases} \quad (13)$$

综合接入信号矢量与发射矢量的相对熵结果,计算网关信道容量,其表达式为:

$$C^* = \beta \frac{\max(H(q)^T + H(y)^T)}{J^*} \quad (14)$$

式(14)中, β 代表高负载数据大批量接入后网关信号的传输比率, J^* 表示数据传输信道的数量。

根据网关信道接入和输出的关系,得到接入信号的协方差矩阵,引入信道衰减因子,得到数据传输信道的容

量,利用接入信号矢量与发射矢量的相对熵,分析网关信道容量。

1.3 网关状态预测

根据网关信道容量的分析结果,在特定时间点定义高负载数据大批量接入后网关状态的概率函数为 $p(e_i, t_i, u_k)$,其中 e_i 表示传感器集合, t_i 表示数据接入的时间序列, u_k 代表接入高负载数据后传感器的状态,在时间窗口 t_c-t_0 中,获取预测网关状态所占的比例:

$$p(e_i, t_p, u_k) = \frac{\{T | u_k\}}{t_c - t_0} \quad (15)$$

式(15)中, t_p 表示网关状态预测时间点。

当循环发生后,在 t_i 时间内发生同样的变化。基于此,利用时间重复性原理,得出网关状态单周期预测模型的预测效果最好。

假设, k_l 代表周期性预测窗口, k_i 表示网关状态的时间窗口,预测概率 $p(e_i, t_p, u_k)$ 可通过下式计算:

$$p(e_i, t_p, u_k) = \frac{\sum_{j=1}^{k_i/k_l} \eta(e_i, t_p - j \times k_l, u_k)}{k_i / k_l} \quad (16)$$

式(16)中, $\eta(\cdot)$ 表示网关状态的统计函数, j 为预测窗口周期。

在一个实体的真实行为中,受许多因素作用,当存在多个周期的影响时,预测准确性降低。为了提高高负载数据大批量接入后网关状态的预测准确性,采用基于单周期、多周期综合的多阶段预测模式,实时预测网关状态。

将 t_p 时刻,网关状态定义为 S_{t_p} ,表达式为:

$$S_{t_p} = (\partial_i, \lambda_i, \zeta_i, \psi_i) \quad (17)$$

式(17)中, ∂_i 表示物联网中传感器的状态, λ_i 表示高负载数据的接入概率, ζ_i 表示预测周期系数, ψ_i 表示周期对应的偏移。

当高负载数据大批量接入后,在 t_p 时刻预测状态 ∂_i 的物联网网关,传感器的接入状态为 ∂_i ,其对应相应的网关状态,由此构建网关状态预测模型:

$$M(e_i, t_i, u_k) = \max \{ \psi_i | (\partial_i, \lambda_i, \zeta_i, \psi_i) \in S_{t_p} \cap \partial_i = u_k \} \quad (18)$$

综上所述,通过定义高负载数据大批量接入后网关状态的概率函数,利用时间重复性原理,计算网关状态的预测概率,采用基于单周期、多周期综合的多阶段预测模式,实现网关状态预测模型的构建。

2 实验分析

2.1 实验硬件环境搭建

为了验证文中方法模型在网关状态预测中的可行性,搭建 ZigBee-Wi-Fi 网络的仿真实验场景,如图 1 所示。

ZigBee 网络和 Wi-Fi 网络是通过协调器中的网关转

换模块连接而成的,网关的硬件框图如图 2 所示。

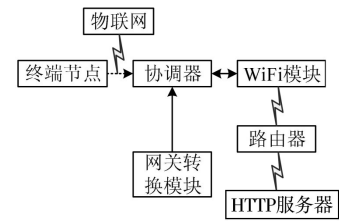


图 1 ZigBee-Wi-Fi 网络的实验场景

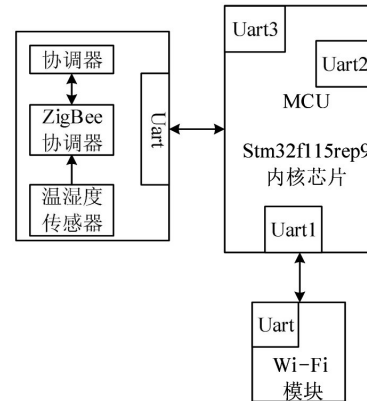


图 2 网关的硬件框图

连接 ZigBee 网络和 Wi-Fi 网络的网关包括温湿度传感器、ZigBee 节点、ZigBee 协调器以及内核芯片和 Wi-Fi 模块,为网关状态的预测提供硬件支持。

2.2 网关状态数据采集

Wi-Fi 网络模块可以实现数据传输,通过 SC 端口与路由器相连,采用 TCP 传输协议将 Wi-Fi 模块与路由器建立连接,形成数据传输通信通道,实现高负载数据大批量接入,由此采集的网关状态数据可以为网关状态预测提供数据支持。数据采集的周期为 0.1 s,在网关状态预测之前,ZigBee-Wi-Fi 网络场景下,采集网关状态数据,如图 3 所示。

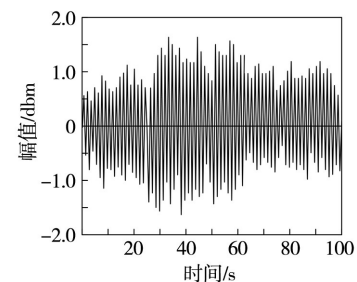


图 3 网关状态数据的采样

当网络接入大批量的高负载数据之后,导致网关运行状态的稳定性变差,根据图 3 采集到的网关状态数据,预测网关状态。

2.3 性能对比

为了突出文中模型的优越性,引入基于链路预测的网关状态预测模型和基于区块链的网关状态预测模型作

对比,当高负载数据大批量接入之后,以数据接入时长为自变量,测试网关数据的转换率和网关信道容量,对比三种方法的优越性。

网关数据的转换率受高负载数据接入系数和时长的影响,测试结果如图4所示。

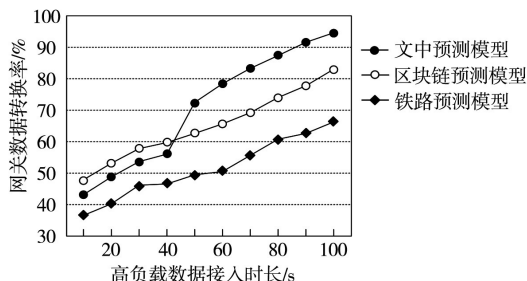


图4 网关数据的转换率

根据图4的结果可知,随着高负载数据接入时长的增加,采用三个模型得到的网关数据转换率逐渐增加,采用基于链路预测的网关状态预测模型时,网关数据的转换率最低,始终在68%以内;当高负载数据接入时长低于43 s时,基于区块链的网关状态预测模型得到的网关数据转换率高于文中模型,但是随着高负载数据接入时长的继续增加,文中方法的网关数据转换率大幅度增加,优化了网关运行的效果。

网关信道容量受网关信号的传输比率和数据传输信道的数量影响。测试网关信道容量,结果如图5所示。

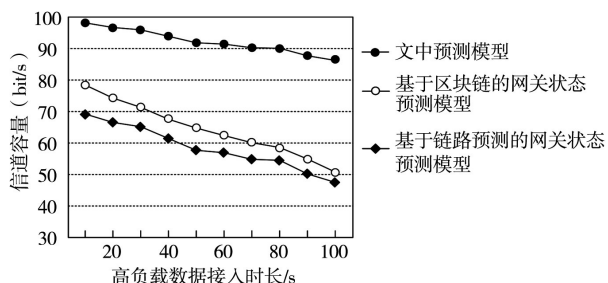


图5 网关信道容量

图5的结果显示,采用基于链路预测的网关状态预测模型和基于区块链的网关状态预测模型时,随着高负载数据接入时长的增加,网关信道容量急速下降,当接入时长为100 s时,网关信道容量已下降到50 bit/s以下;而采用文中模型时,网关信道容量也呈现出下降的趋势,但是由于文中模型通过高负载数据的传输建模,避免了高负载数据大批量接入后,网关信道出现阻塞的情况,将网关信道容量控制在85 bit/s以上。

3 结束语

本文构建高负载数据大批量接入后网关状态预测模型,计算高负载数据大批量接入后各个节点的平均传输距离,根据高负载数据传输概率,构建物联网中高负载数

据安全传输模型;综合接入信号矢量与发射矢量的相对熵结果,计算网关信道容量;计算网关状态的预测概率,采用基于单周期、多周期综合的多阶段预测模式,实现网关状态预测模型的构建。经过实验测试结果分析,该模型通过网关状态的预测,可以提高网关数据的转换率,提升至90%以上;网关信道容量随接入时长变化仍保持在85 bit/s以上,保证网络运行的稳定性。在今后研究中,希望可以在高负载大批量接入之前,对高负载数据进行加密处理,确保网关的安全性。

参考文献:

- [1] 陈福才,何威振,程国振,等.基于DPDK的内网动态网关关键技术设计[J].通信学报,2020,41(6):139-151.
- [2] 陈文艺,高婧,杨辉.物联网网关实时双向通信模块化设计[J].计算机工程与设计,2021,42(5):1215-1221.
- [3] 武超飞,付义豪,孙冲,等.基于模组电能表的居民侧智慧能源网关设计[J].电测与仪表,2020,57(21):140-146.
- [4] 姚玉坤,宋威威,李其超,等.链路预测模型下的无人机网关选择算法[J].小型微型计算机系统,2020,41(2):338-343.
- [5] 赵丙镇,王栋,钱雪,等.基于区块链的电力物联网信任网关设计与实现[J].中国电力,2021,54(7):192-197.
- [6] Jrgensen A A, Kong D, Henriksen M R, et al. Petabit-per-second data transmission using a chip-scale microcomb ring resonator source[J]. Nature Photonics, 2022, 16(11): 798-802.
- [7] Yang H, Li F, Yu D X, et al. Reliable Data Storage in Heterogeneous Wireless Sensor Networks by Jointly Optimizing Routing and Storage Node Deployment[J]. Tsinghua Science and Technology, 2021, 26(2): 230-238.
- [8] 王玉峰,高雅娟.物联网终端安全接入数据网关状态预测仿真[J].计算机仿真,2022,39(4):341-344,356.
- [9] 肖勇,金鑫,王立博,等.基于AMI的智慧用能多模汇聚网关装置研究[J].计算机科学,2020,47(S01):289-293.
- [10] 孙晓川,李志刚,张明辉,等.基于集成学习的广域光骨干网多信道传输质量预测方法[J].通信学报,2020,41(9):7.
- [11] 程玉涛,李宁,李海波.基于加密算法的电力物联网电能质量数据安全传输方法[J].自动化技术与应用,2022,41(7):87-90.
- [12] 禹鑫焱,唐权瑞,施甜峰,等.基于OPC UA协议的工业网关系统设计与实现[J].高技术通讯,2021,31(9):962-968.

作者简介: 敖知琪(1996-),男,硕士,工程师,研究方向:企业架构。