基于预测的自动驾驶 GNSS 欺骗攻击检测

David Pritchard
davidpritchard.org

Gregory S. Macfarlane, Ph.D.

Chieh (Ross) Wang, Ph.D., Corresponding Author
cwang@ornl.gov

Word Count: 10000 words + 1 table(s) × 250 = 10250 words

Submission Date: 2022 年 3 月 14 日

摘要

全球导航卫星系统（GNSS）使用卫星与无线电通信为自动驾驶汽车（AV）提供定位、导航和定时（PNT）服务。然而，由于缺少信号加密，同时 GNSS 粗提取码（C/A）的具有公开可获得性，以及 GNSS 信号较低的信号强度，GNSS 很容易受到欺骗性攻击，从而使得 AV 的导航能力受到损害。欺骗性攻击往往很难被发现，因为欺骗者（实施欺骗性攻击的攻击者）可以模仿 GNSS 信号并向 AV 传输不准确的位置坐标。在这项研究中，我们利用长短期记忆（LSTM）模型（一种循环神经网络模型）开发了一种基于预测的欺骗攻击检测策略。LSTM 模型常被用来预测自主车辆的两个连续位置之间的距离，为了开发 LSTM 预测模型，在本项研究中，我们将使用一个名为 comma2k19 的公开真实世界驾驶数据集。该训练数据集包含不同的特征（即加速度、方向盘角度、速度和两个连续位置之间的行驶距离），这些特征是从汽车的控制区域网络（CAN）、GNSS 模块和惯性测量单元（IMU）传感器中提取的。根据对自主车辆当前位置和下一个位置之间的行驶距离进行预测，利用 GNSS 设备的定位误差和当前位置域下一个位置之间行驶距离的预测误差（即最大绝对误差），建立一个阈值。我们的分析显示，基于预测的欺骗性攻击检测策略可以成功且实时地检测到攻击。

关键词: GNSS，自动驾驶汽车，网络安全，欺骗性攻击，LSTM

引言

自动驾驶汽车需要可靠的定位、导航和计时（PNT）服务，以履行其自主功能 (*1*)。例如，自动驾驶车辆（AVs）使用 PNT 服务来定实现位以及沿道路导航的功能 (*2*)。自动驾驶汽车的安全和运行也会受到不可靠 PNT 服务的严重影响，这是因为自动驾驶汽车依靠 PNT 服务来规划路线从而到达目的地 (*3*)。因此，可靠的 PNT 服务对于自动驾驶汽车的安全以及高效运行是至关重要的。全球定位系统（GPS）/全球导航卫星系统（GNSS）提供的 PNT 服务依赖于卫星和无线电通信，而这些都会受到各种威胁和脆弱性的影响 (*2*)。最常见的影响是环境造成的遮挡，如城市地区的高层建筑、车库和隧道的墙壁以及天花板；甚至是天空中厚厚的云层，也会对信号造成影响 (*4*)。除了这些自然环境的影响外，GPS/GNSS 系统还受到一些其他的威胁，如无线电干扰 (*5*)、通信欺骗 (*6*)、传输信息数据操纵 (*7*)，甚至 GPS/GNSS 基础设施的破坏 (*7*)。所有这些威胁都会影响 PNT 服务的可靠性，从而对 AV 的安全和可靠运行产生严重影响。由于缺乏加密，粗提取码（C/A）的公开可获得性以及 GNSS 信号较低的强度，GNSS 很容易受到自然因素以及伪造攻击的影响，从而使车辆的导航能力受到损害 (*8*)(*9*)(*10*)(*11*)(*12*)(*13*)。其中，欺骗性攻击是最复杂的攻击类型，很难被发现，因为攻击者可以模仿 GNSS 信号并向 AV 传输不准确的坐标。这将使攻击者能够改变 AV 的路线，将 AV 引向一个不安全的地方，甚至造成严重的交通事故。欺骗攻击可以分为三种类型；简单的、中级的和复杂的 (*14*)(*15*)(*16*)。简单型攻击，由于实际的 GNSS 信号和欺骗者的信号之间没有同步，所以很容易被发现，可以使用商业 GNSS 信号模拟器来创建。对于中间人攻击，欺骗者可以一个使用便携式的接收者-欺骗者设备将假信号与实际卫星时间同步，并调整所有卫星信号的多普勒频率、码相。因此，在中间人攻击期间，假信号将会很难被检测到。由于 AV 使用多个 GNSS 接收机天线来连续交叉验证信号，使用多个锁相欺骗器来欺骗所有天线，可以实现复杂的欺骗攻击。因此，开发一个针对复杂欺骗攻击的检测模型是一个挑战。在这项研究中，我们开发了一个基于预测的复杂欺骗攻击检测策略，使用了一个递归神经网络模型，这是一个长短期记忆（LSTM）模型。LSTM 模型被用来预测自主车辆的两个连续位置之间的行驶距离。在不同的递归神经网络（RNN）结构中，LSTM 是最受欢迎的一种，因为它能够说明时间序列数据中的短期和长期依赖关系。

本研究的贡献

我们研究的贡献是利用人工递归神经网络模型，即长短期记忆（LSTM）模型，开发了一个基于预测的欺骗攻击检测策略。LSTM 模型根据来自 AV 传感器的当前数据，预测 AV 在两个连续地点之间行驶的距离。根据预测的自主车辆在当前位置和下一位置之间的行驶距离，利用 GNSS 设备的定位误差和预测

误差（即真实距离和预测行驶距离之间的最大绝对误差）建立一个误差阈值。据我们所知，目前还没有任何研究使用 AV 的预测位置来检测使用即时 AV 传感器数据进行的 GNSS 欺骗攻击。

相关工作

我们回顾了与 GNSS 欺骗攻击模型和欺骗攻击检测模型有关的文献，并将在以下各小节中介绍。

GNSS 欺骗性攻击

欺骗性攻击可分为四步。（i）发射器部署；（ii）接管策略；（iii）控制策略；以及（iv）应用。欺骗者发射器的各种配置被用来实现所需的欺骗行为。单个或多个发射器都可用于实施欺骗 (*17*)。此外，欺骗者可以采取不同的策略，例如：（i）GPS/GNSS 信号极化复制；（ii）多普勒频率范围；（iii）伪随

机噪声（Promannum）序列；（iv）信号带宽；（v）载波频率；（vi）接收功率；以及（vii）调制类型，从而使被攻击的车辆无法检测到任何异常 (*17*)。欺骗者可以进一步匹配数据位的帧结构，使被攻击的车辆确信欺骗信号是真实信号。通过改变信号的时间偏移，欺骗者可以改变卫星的伪距，实现对被攻击车辆的控制。在 (*18*) 中，作者研究了攻击产生的限制以及在道路上破坏 AV 导航系统的可能方式。例如，对 AV 的导航攻击可以通过随机的 GPS 操作来创建，这可能导致不准确的导航指令，如在道路中间右转。(*18*) 中的作者设计了一种攻击算法，通过生成欺骗性的 GPS 数据来操纵自主车辆导航系统。操纵导航系统的目的是产生逐个转弯的欺骗性数据，以便引导自动驾驶汽车到一个错误的目的地。作者在 (*19*) 中评估了惯性导航系统（INS）辅助的 GPS 跟踪和道路导航的安全保障。他们展示了一个综合的、动态变化的 GPS/INS 欺骗攻击，该攻击引导 AV 到一个错误的目的地。(*20*) 中的作者展示了一种使用自适应蒙特卡洛定位（AMCL）算法对位置隐私进行的新的缓存侧信道攻击。

欺骗性攻击检测模型

通常情况下，现有的针对欺骗攻击的对策一般可以分为两类：加密技术以及信号层面的异常检测技术。一般来说，卫星信号是通过使用加密技术的附加信号来验证的，使没有安全密钥的用户无法解密信号 (*19*)。信号电平检测是通过识别物理信号波形或来自信号源的到达角度的差异进行的。另一种检测方法是使用 GNSS 信号的多个接收器来验证信号，以检测任何异常情况 (*20*)(*21*)。然而，欺骗者可以通过生成多颗卫星的同步信号来欺骗这种接收器。GNSS 信号的频率漂移、信号强度以及自由运行的晶体振荡器和 GNSS 信号之间的相关性也可以用来检测欺骗性攻击 (*16*)。Humphreys 等人开发了一个软件定义的民用 GPS 接收机欺骗器，该设备可以用来产生攻击。他们还开发了两种基于软件定义的用户设备的欺骗防御方法：（i）数据位延迟防御和（ii）estigial 信号防御 (*22*)。他们发现，他们的欺骗防御方法能够成功地检测到攻击，而这种攻击是由他们的软件定义的民用 GPS 接收机- 欺骗者产生的。虽然大部分的

探测技术分析了 GNSS 信号，很少有研究关注使用 GPS 坐标数据的欺骗攻击检测。Panice 等人通过比较估计无人驾驶飞行器（UAVs）的当前状态和实际位置，开发了一种基于支持向量机（SVM）的欺骗检测方法 (*23*)。这种方法在检测无人机被欺骗着陆的欺骗攻击时，比检测由捷联惯导系统导致的攻击更有效。在 (*24*) 中，Wang 等人开发了一种重建缺失 GPS 信号的方法，作者使用了实时驾驶数据和 GPS 坐标。他们的方法使用驾驶数据重建驾驶路线，如果输入的 GPS 坐标与预测的 GPS 坐标的偏差超过误差阈值，则标记为攻击 (*24*)。

数据集描述与数据处理

在这项研究中，我们使用了 Comma.ai 的真实世界驾驶数据集，名为 Comma2k19（26），其中包含各种 AV 传感器数据。车辆在加利福尼亚的 280 条高速公路上行驶了 33 个小时，共 2019 个路段（见图 1）。每个路段的行驶时间为 1 分钟。所有的路段都位于圣何塞和旧金山之间，总长度为 20 公里。Comma.ai 使用的 AV 有一个前置摄像头、温度计和 9 轴惯性测量单元（IMU）。这些设备都被用于收集数据。除了这些传感器数据，Comma2k19 数据集还包含来自 GNSS 和 CAN（控制器局域网）的测量数据（见表 1 和表 2）。GNSS 接收模块 u-blox M8 也被用于数据收集，其水平位置精度为 2.5 米。全球定位系统（GPS）和全球轨道导航卫星系统（GLONASS）信号被用于位置测量。此外，一个开放的 GNSS 处理库 Laika 被用来减少定位误差，该处理库可以将定位误差减少了 40%。与本研究有关的 IMU 数据是车辆的加速度（如表 3 所示）。同样，相关的 CAN 数据是车辆的速度和方向盘角度数据（见表 2）。GNSS 数据集包含来自 u-blox 和 Qcom 的实时和原始 GNSS 数据。每个实时部分包括纬

度、经度、速度、utc_timestamp、高度和方位角数据。这个数据集包含密集且多样的驾驶数据，可用于训练递归神经网络模型并预测所需变量。这个用于分析的数据集包含 7200 个 GNSS 观测值，35726 个 CAN 观测值，以及 72148 个 IMU 观测值。由于我们预测了自主车辆当前位置和下一个位置之间的行驶距离，因此在每个时间步骤中，可以利用经纬度坐标和以下哈弗辛大圆公式计算出与前一个时间步骤的行驶距离：

$$d = 2r\sin^{-1}\left(\sqrt{\sin^2\left(\frac{\varphi_2-\varphi_1}{2}\right)+\cos(\varphi_1)\cos(\varphi_2)\sin^2\left(\frac{\psi_1-\psi_2}{2}\right)}\right) \tag{1}$$

其中 d 是地球表面两点之间的距离，r 是地球的半径。φ_1，φ_2 是当前位置的纬度（弧度），ψ_1，ψ_2 是下一时刻位置的经度（弧度）。*haversine* 函数定义如下，

$$hav(\theta) = \sin^2\left(\frac{\theta}{2}\right) = \frac{1-\cos\theta}{2} \tag{2}$$

其中，θ 是角度的测量值，单位是弧度。图 3 显示了从 comma2k19 数据集中的 5987 个观测值中的每个时间戳在当前位置和未来位置之间的移动距离。这些数据与 CAN 和 IMU 一起被用来训练 LSTM 模型

预测模型的开发

LSTM 是 RNN 的一种特殊类型，它可以通过将时间序列数据的时间依赖性存储在递归隐藏层的记忆块中来处理长期依赖性。每个存储块由三个门组成。（i）输入门，（ii）输出门和（iii）遗忘门。块的状态是由这三个门控制的。一个 LSTM 块包括一个 *sigmoid*(σ) 和一个 *tanh* 激活函数。我们使用 LSTM 模型来预测每个时间戳的当前位置和近期位置之间的行驶距离，使用未被攻击的 CAN、IMU 和 GNSS 传感器数据（真实值）。在本文中，我们使用了一个 LSTM 模型，它由一个输入层、一个有 50 个神经元的递归隐藏层和一个输出层组成。输入到 LSTM 模型中用于训练的数据包括：来自 CAN 的速度和转向角数据以及来自 IMU 的前向加速度数据。输出是对于每一个时间戳当前位置和下一时间戳的位置之间行驶的距离。

在训练 LSTM 模型之前，所有的特征都会被归一化到 0 和 1 之间，然后输入到 LSTM 模型。我们将数据集分成由 4500 个样本组成的训练数据集和由 1487 个样本组成的验证数据集。LSTM 模型的超参数，如神经元数量、epoch 数、batch 大小和学习率在提高预测精度方面起着至关重要的作用。我们采用试错法来找出最佳的超参数，并使用 Adam 优化器作为评估指标，以确保我们在训练模型时不存在过拟合和欠拟合问题。平均绝对误差（MAE）被用作损失函数，其定义如下。

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_p - y_g| \tag{3}$$

其中，N 是总的训练样本量，y_p 与 y_g 是真实值和预测距离数据。表 4 中列出了优化后的超参数值。我们在训练和验证数据集上绘制了 MAE，以评估 LSTM 模型的拟合度。图 4 显示了使用最佳超参数的损失曲线，包括训练和验证数据集。这两个数据集的 MAE 的比较表明，使用最优超参数的预测模型具有良好的拟合性。我们使用测试数据集测试了我们所训练的预测模型。RMSE 被用来作为评估指标，以确定模型的有效性，具体如下.

$$RMSE = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(y_p - y_g)^2} \tag{4}$$

其中，N 是总的测试样本量，y_g 和 y_p 分别是地面实况和预测距离数据。图 5 显示了每个时间戳在当前位置和即将到来的位置之间的地面实况和预测行驶距离。我们发现，预测行驶距离的 RMSE 为 0.0242 米，平均绝对误差为 0.0203 米。

CAPTIONS

Figure 1 shows a Gumbel distribution as an example of captioning. As demonstrated, figure captions ought to be sentence capitalized, balded, and can be somewhat longer than in other journals.

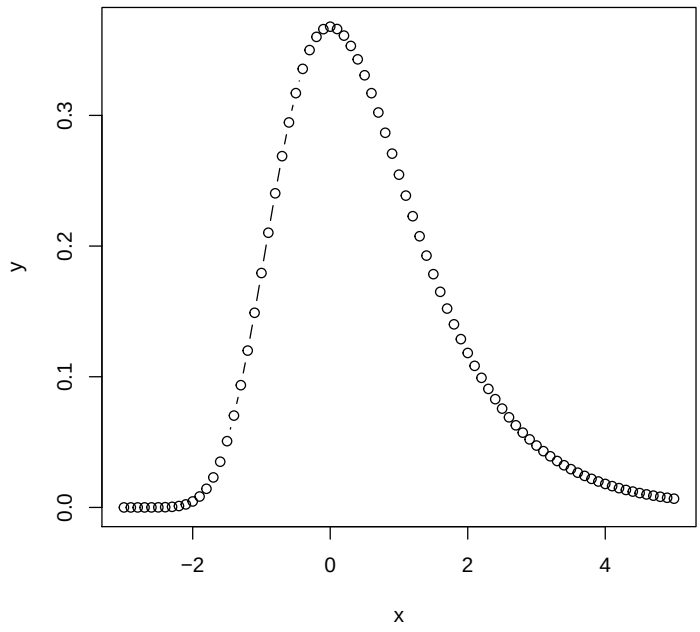

FIGURE 1 This is a random figure to test the counting functionality on the title page. It shows a Gumbel distribution with mode 0 and scale 1. The multinomial logit model assumes that the error terms are distributed identically and independently following this pattern.

Table captions are somewhat different, requiring initial capitals and are more of a title. An example of this is given in Table 1, showing the history of this template.

Bibliography

The TRB bibliography style is defined in the `trb.bst` file which should be in your document folder. A renewed command is specified, `\citep{}` which will print the authors and the number of the reference in the order in which it is supplied. Note that `\citep{}` prints both the author names and the reference number, if you simply need the number of the reference, use command `\cite{}`. The References section will be appended to the end of the document.

TABLE 1 A History of this Template

Version	Date	Author	Contributions
1.0	Sep 2009	Pritchard	Initial work
1.1	Mar 2011	Pritchard	Captions
2.0	Mar 2012	Macfarlane	Automation, documentation
2.1	Jul 2015	Wang	More automation and formatting
2.1.1	Jan 2016	Wang	Minor modifications and uploaded to Github
2.1.1 Lite	Jun 2017	Wang	TEX-only template
3.1	Jun 2017	Wang	Addition of `trbunofficial.cls`
3.1 Lite	Jun 2017	Wang	Addition of `trbunofficial.cls`
4.0 Lite	Jul 2019	Wang	Word count updates for Overleaf/ShareLaTeX compatibility

Equations

Intelligent driver model equations from wikipedia (`https://en.wikipedia.org/wiki/Intelligent_driver_model`) moved to the left using `amsmath` package with `fleqn` options.

$$\dot{x}_\alpha = \frac{\mathrm{d}x_\alpha}{\mathrm{d}t} = v_\alpha \tag{5}$$

$$\dot{v}_\alpha = \frac{\mathrm{d}v_\alpha}{\mathrm{d}t} = a\left(1 - \left(\frac{v_\alpha}{v_0}\right)^\delta - \left(\frac{s^*(v_\alpha, \Delta v_\alpha)}{s_\alpha}\right)^2\right) \tag{6}$$

$$s^*(v_\alpha, \Delta v_\alpha) = s_0 + v_\alpha T + \frac{v_\alpha \Delta v_\alpha}{2\sqrt{ab}} \tag{7}$$

TO DO'S

Two document types, extending from the `[numbered]` option, can be defined to differentiate the initial submission (i.e., with line numbers and in-line figures and tables) and the final manuscript (i.e., without line numbers and all figures and tables are attached to the end).

CONCLUSION

To make the document from source in a Unix-like OS, issue the following commands:

```
latexmk trb_template.tex -pdf -pvc -shell-escape
```

The `--shell-escape` option is required to access the command line for the word count. Normally this feature is disabled because it is a route of entry for malicious software. We promise that there is no such debilitating code in this document, and we encourage you to examine any scripts for suspicious code before permitting `pdflatex` from accessing your system.

Perl is necessary for "texcount" to work and needs a Perl interpreter e.g. [ActivePerl](`http://www.activestate.com/activeperl/downloads`).

ACKNOWLEDGMENTS

The authors would like to thank Aleksandar Trifunovic (`https://github.com/akstrfn`) for creating the `trbunofficial` class document, which has been a very helpful improvement.

REFERENCES

1. Mit, R., Y. Zangvil, and D. Katalan, Analyzing Tesla ‘s Level 2 Autonomous Driving System Under Different GNSS Spoofing Scenarios and Implementing Connected Services for Authentication and Reliability of GNSS Data. In *Proceedings of the 33rd International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2020)*, 2020, pp. 621–646.
2. Yousuf, M., D. J. Dailey, S. Sundararajan, and R. Kandarpa, Enabling Technologies for Vehicle Automation. In *Road Vehicle Automation 3*, Springer, 2016, pp. 225–235.
3. Papadimitratos, P. and A. Jovanovic, Protection and fundamental vulnerability of GNSS. In *2008 IEEE International Workshop on Satellite and Space Communications*, IEEE, 2008, pp. 167–171.
4. Milanés, V., J. E. Naranjo, C. González, J. Alonso, and T. de Pedro, Autonomous vehicle based in cooperative GPS and inertial systems. *Robotica*, Vol. 26, No. 5, 2008, pp. 627–633.
5. Bachrach, A., S. Prentice, R. He, and N. Roy, RANGE–Robust autonomous navigation in GPS-denied environments. *Journal of Field Robotics*, Vol. 28, No. 5, 2011, pp. 644–666.
6. Dovis, F., *GNSS interference threats and countermeasures*. Artech House, 2015.
7. Akos, D. M., Who’s afraid of the spoofer? GPS/GNSS spoofing detection via automatic gain control (AGC). *NAVIGATION, Journal of the Institute of Navigation*, Vol. 59, No. 4, 2012, pp. 281–290.
8. Aleshin, I. M., K. I. Kholodkov, and V. N. Koryagin, Framework for GREIS-formatted GNSS data manipulation. *GPS Solutions*, Vol. 24, No. 2, 2020, pp. 1–7.
9. Schmidt, E., N. Gatsis, and D. Akopian, A GPS spoofing detection and classification correlator-based technique using the LASSO. *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 56, No. 6, 2020, pp. 4224–4237.
10. Amin, M. G., P. Closas, A. Broumandan, and J. L. Volakis, Vulnerabilities, threats, and authentication in satellite-based navigation systems [scanning the issue]. *Proceedings of the IEEE*, Vol. 104, No. 6, 2016, pp. 1169–1173.
11. Volpe, J. A., Vulnerability assessment of the transportation infrastructure relying on the global positioning system. *http://www. navcen. uscg. gov/*, 2001.
12. Jafarnia-Jahromi, A., A. Broumandan, J. Nielsen, and G. Lachapelle, GPS vulnerability to spoofing threats and a review of antispoofing techniques. *International Journal of Navigation and Observation*, Vol. 2012, 2012.
13. Juang, J.-C., Analysis of global navigation satellite system position deviation under spoofing. *IET radar, sonar & navigation*, Vol. 3, No. 1, 2009, pp. 1–7.
14. Huang, J., L. L. Presti, B. Motella, and M. Pini, GNSS spoofing detection: Theoretical analysis and performance of the Ratio Test metric in open sky. *Ict Express*, Vol. 2, No. 1, 2016, pp. 37–40.
15. Khan, A. M., N. Iqbal, A. A. Khan, M. F. Khan, and A. Ahmad, Detection of intermediate

spoofing attack on global navigation satellite system receiver through slope based metrics. *The Journal of Navigation*, Vol. 73, No. 5, 2020, pp. 1052–1068.

16. Psiaki, M. L. and T. E. Humphreys, GNSS spoofing and detection. *Proceedings of the IEEE*, Vol. 104, No. 6, 2016, pp. 1258–1270.
17. Misra, P. and P. Enge, *Global Positioning System: Signals, Measurements and Performance (Lincoln, MA: Ganga*, 2006.
18. van der Merwe, J. R., X. Zubizarreta, I. Lukčin, A. Rügamer, and W. Felber, Classification of spoofing attack types. In *2018 European Navigation Conference (ENC)*, IEEE, 2018, pp. 91–99.
19. Nicola, M., L. Musumeci, M. Pini, M. Fantino, and P. Mulassano, Design of a GNSS spoofing device based on a GPS/Galileo software receiver for the development of robust countermeasures. In *ENC GNSS*, 2010.
20. Scott, L., Anti-spoofing & authenticated signal architectures for civil navigation systems. In *Proceedings of the 16th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS/GNSS 2003)*, 2003, pp. 1543–1552.
21. O' Hanlon, B. W., M. L. Psiaki, J. A. Bhatti, D. P. Shepard, and T. E. Humphreys, Real-time GPS spoofing detection via correlation of encrypted signals. *NAVIGATION, Journal of the Institute of Navigation*, Vol. 60, No. 4, 2013, pp. 267–278.
22. Humphreys, T. E., B. M. Ledvina, M. L. Psiaki, B. W. O'Hanlon, P. M. Kintner, et al., Assessing the spoofing threat: Development of a portable GPS civilian spoofer. In *Proceedings of the 21st International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2008)*, 2008, pp. 2314–2325.
23. Panice, G., S. Luongo, G. Gigante, D. Pascarella, C. Di Benedetto, A. Vozella, and A. Pescapè, A SVM-based detection approach for GPS spoofing attacks to UAV. In *2017 23rd International Conference on Automation and Computing (ICAC)*, IEEE, 2017, pp. 1–11.
24. Wang, Q., Z. Lu, M. Gao, and G. Qu, Edge computing based gps spoofing detection methods. In *2018 IEEE 23rd International Conference on Digital Signal Processing (DSP)*, IEEE, 2018, pp. 1–5.