# Article information:

Model poisoning attack in differential privacy-based federated learning  
<https://schlr.cnki.net/en/Detail/index/GARJ2021_4/SJES91E4613ABF7077A65EAF6CDE1CB91F0C>

# Article summary:

1. Federated learning can still reveal user privacy through shared parameters or gradients.

2. Differential privacy is a promising solution to this problem, but it opens up a vulnerability to model poisoning attacks.

3. The proposed Model Shuffle Attack (MSA) can significantly degrade the performance of the global model while remaining stealthy and even causing the model to diverge.

# Article rating:

Appears moderately imbalanced: The article provides some useful information, but is missing several important points or pieces of evidence that would be required to present the discussed topics in a balanced and reliable way. You are encouraged to seek a more balanced perspective on the presented issues by exploring the provided research topics and looking at different information sources.

# Article analysis:

作为一篇关于差分隐私的联邦学习中模型污染攻击的研究论文，该文章提出了一种新颖的模型混洗攻击（MSA），并探讨了在差分隐私保护下，攻击者如何利用噪声来逃避异常检测。然而，在对该文章进行批判性分析时，我们发现以下几个问题：

首先，该文章似乎忽略了联邦学习本身所面临的安全和隐私风险。虽然联邦学习可以保护个体原始数据的隐私，但是共享参数或梯度仍可能泄露用户隐私。因此，在研究差分隐私保护下的模型污染攻击时，应该更加重视联邦学习本身所面临的安全和隐私风险。

其次，该文章提出的MSA攻击似乎只考虑了攻击者如何通过混洗和缩放模型参数来干扰全局模型，并没有考虑其他可能存在的攻击方式。例如，攻击者可以通过篡改训练数据、注入恶意代码等方式来影响全局模型。因此，在研究模型污染攻击时，应该考虑到各种可能存在的攻击方式，并采取相应的防御措施。

第三，该文章似乎没有充分探讨MSA攻击对用户隐私和数据安全带来的潜在风险。尽管差分隐私可以通过添加噪声来保护用户隐私，但是如果攻击者能够成功地进行模型污染攻击，则可能会导致用户数据被暴露或被滥用。因此，在研究差分隐私保护下的模型污染攻击时，应该更加重视对用户隐私和数据安全带来的潜在风险。

最后，该文章似乎缺少对实验结果进行充分验证和解释。尽管作者声称在FedAvg和鲁棒聚合规则下进行实验，并证明MSA能够显著降低全局模型性能并保证悄无声息性（stealthiness），但是具体实验设置、评价指标以及结果解释都需要更加详细地说明和验证。

综上所述，尽管该文章提出了一种新颖的模型混洗攻击，并探讨了在差分隐私保护下如何进行模型污染攻击，但是其存在着忽略联邦学习本身所面临安全和隐私风险、未考虑其他可能存在的攻击方式、未充分探讨潜在风险以及缺少对实验结果进行充分验证等问题。因此，在进一步研究联邦学习中的安全与隐私问题时需要更加谨慎地思考并采取相应措施。

# Topics for further research:

* Security and privacy risks in federated learning
* Other possible attack methods in model poisoning
* Potential risks to user privacy and data security
* Need for more detailed experimental settings and evaluation metrics
* Importance of considering various attack scenarios and defense measures
* Caution in further research on security and privacy issues in federated learning

# Report location:

<https://www.fullpicture.app/item/7557c3c9f83a9c147314bc008c716cc4>