# Article information:

SOTER: Guarding Black-box Inference for General Neural Networks at the Edge | USENIX
<https://www.usenix.org/conference/atc22/presentation/shen>

# Article summary:

1. SOTER is a system that can protect model confidentiality, integrity, and accuracy while providing low inference latency on third-party edge devices.

2. SOTER transforms associative operators into parameter-morphed, confidentiality-preserved operators to execute on untrusted GPUs.

3. An oblivious fingerprinting technique is used to detect integrity breaches of morphed operators outside TEE to ensure correct executions of inferences.

# Article rating:

May be slightly imbalanced: The article presents the information in a generally reliable way, but there are minor points of consideration that could be explored further or claims that are not fully backed by appropriate evidence. Some perspectives may also be omitted, and you are encouraged to use the research topics section to explore the topic further.

# Article analysis:

The article provides an overview of the SOTER system which is designed to protect model confidentiality, integrity, and accuracy while providing low inference latency on third-party edge devices. The authors provide evidence for their claims by citing experiments conducted on six prevalent models in the three most popular categories which show that SOTER achieves comparable performance with partition-based baselines while retaining the same high accuracy as insecure inference.

The article appears to be trustworthy and reliable as it provides evidence for its claims and does not appear to be biased or one-sided in its reporting. However, there are some points of consideration that are missing from the article such as potential risks associated with using the system and how it could be improved upon in the future. Additionally, there is no mention of any counterarguments or alternative solutions that could be used instead of SOTER which could have been explored further in the article. Furthermore, there is no promotional content present in the article which suggests that it is unbiased and impartial in its reporting.

# Topics for further research:

* Alternative solutions for model confidentiality
* Risks associated with SOTER system
* Improvements to SOTER system
* Counterarguments to SOTER system
* Partition-based baselines for model inference
* Low inference latency on third-party edge devices

# Report location:

<https://www.fullpicture.app/item/796462c03105b94c5490c5f50f8dee68>