SaTC: Core: Medium: Protecting Confidentiality and Integrity of Deep Neural Networks against Side-channel and Fault Attacks

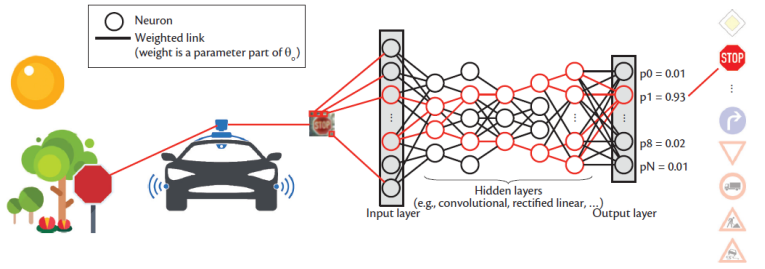
CHEST
CENTER FOR HARDWARE AND EMBEDDED SYSTEMS SECURITY AND TRUST

Challenge:

- Security implication of DNN: IP confidentiality and integrity/availability
- Diverse models, platforms, and applications for DNNs
- Optimization of DNNs and transfer learning

Solution:

- SpyNet: leverage different side-channels for recovering DNN structure and parameters on diverse platforms
- DisruptNet: manipulate DNN operations via practical hardware and software fault injections
- SecureNet: network obfuscation against side-channel attacks, detection of integrity violation of DNNs, and hardening techniques for fault resistance



Overview of SpyNet

	Model Information	HW Implementation	SW Implementation
Structure characteristics	# layers	power SPA	
	type of layer, activation	memory access	μarchitecture (I$, PMC)
	connection/ layers	power SPA, timing	
Hyperparameters	# neurons in FC	power SPA	μarchitecture (I$, PMC)
	# of kernels in CONV	power SPA, memory access	I$, D$, PMC, constraints
	size of kernel in CONV/POOL	memory access	I$, D$, PMC, constraints
Parameters	weights in FC	power DPA, bus snoop	FP
	kernel values in CONV	power DPA, bus snoop	timing μarchitecture, FP

Overview of DisruptNet

		HW implementation		SW implementation	
		Resource	Fault Type	Resource/Stage	Fault Type
Computation		datapath PE	output: stuck-at, random	instruction execution	skip,control/data flow
		control logic	control flow		
Data	reuse	buffer	set/reset, random	DRAM	set/reset, random, flip (rowhammer)
	temporary	registers	set/reset, random (DVFS)	registers	set/reset, random (DVFS)

CNS1929300, Northeastern University,

Yunsi Fei, Shelley Xue Lin, Thomas Wahl
{y.fei,x.lin,t.wahl}@northeastern.edu

Scientific Impact:

- Investigate a new attack surface of DNN inference
- Systematically protect confidentiality and integrity of DNNs
- Deepen understanding of inherent information leakage and fault tolerance of DNN models

Broader Impact:

- Facilitate wide adoption of DNN in security-critical applications
- Advance the state-of-the-art DNN implementations, computer architecture, heterogeneous systems, hardware security, formal methods and verification
- Technology transfer with company partners through a new NSF IUCRC center