Storyline of researchers in the Princeton University

Incoming data streams could interfere with representations of previous stimuli and cause us to overwrite or misinterpret important contextual information

Studying brain signals

To represent current and past stimuli simultaneously without mutual interference, the brain essentially "rotates" sensory information to encode it as a memory

"Neural representations of sensory information get rotated 90 degrees to transform them into memories"

The two orthogonal (90 degree) representations can then draw from overlapping neural activity without intruding on each other.

neuroscientist would want to see if people pre-allocating neurons in their neural networks to have stable and switching properties, instead of just random properties, helped their networks in some way,

Our storyline

Previous experience: Success of applying Machine Learning into the signal identification of various drones signals

Can we apply the similar technology to detect & identify ADS-B signals to defend the national airspace against malicious cyber attackers?

We can not let AI to learn every aircraft in the world within a short time, and we need to let AI to learn incrementally by demand.

We are encountering the catastrophic forgetting when AI losses previous memories (old aircraft) uncontrollably when learning to identify new aircraft

We discover that the conflict (interference) between critical memory representations (encoded sensory information) causes the problem. Each memory representation corresponds to an aircraft.

Is there a way to organize those memory representations to minimize those conflicts?

Discovered a theorem and show that: to minimize the interference, the more AI learns, the closer averaged deviation angle between memory representations approximates 90 degree.

Proposed a new incremental learning algorithm, with memories in different learning stages orthogonally separated we do have both stable (by Knowledge Distillation or Elastic Weight Consolidation) and varying neurons.

Now our AI can incrementally learn to identify aircraft from their ADS-B signals and show better performance than existing algorithms

If we force AI to frozen the old memories about learned aircraft, then it refuses to learn new aircraft to a sufficiently good status.

The malicious

The Legitimate

This is aircraft A: abandon landing

This is aircraft A: request landing

Air traffic control centers

Which message is the REAL one?

Our new motivation: the untols risk within ADS-B

Both human & AI are forgetting

Aircraft A

Aircraft B

Nonconflict

Memory representation A

Memory representation B

Aircraft A

Aircraft B

Conflict causes confuse and catastrophic forgetting

Memory representation A

Memory representation B

I don't have mice and neural electrodes but I have mathematics & computers.

Aircraft A

Aircraft C

Aircraft B

I applied my discoveries to real AI before knowing that it's their next step. LoL...