

SAMARITAN RELOADED

RÉALISATION

NOPOZA973

TYPE

ANTI-CHEAT

BUT

LISTER DES PISTES D'AMÉLIORATIONS

RÉALISÉ EN AOÛT 2016



You are being watched. The government has a secret systeme, a machine, that spies on you every hour of every day. I know, because I built it. I designed the machine to detect acts of terror, but it sees everything. Violent crimes involving ordinary people. People like you. Crimes the government considered irrelevant. They wouldn't act, so I decided I would. But I needed a partner, someone with the skills to intervene. Hunted by the authorities, we work in secret. You'll never find us. But, victim or perpetrator, if your number's up, we'll find you.

Samaritan is an anticheat system with a super intelligence to act as a mass surveillance system and try to eradicate cheating.

Indeed, Samaritan can evolve from himself thanks to <u>Deep Learning</u> models. He continuously learn for better identification of new kind of cheat or hack client without any performance impact on your servers.

Goal

The main goal is to provide a sustainable and easy way to protect players against cheater.

The second goal is to give us (you?) a try to learn this kind of concept (MachineLearning, BigData, etc ...) in a big production environment.

Technology

Because of the project size, we have to use simultaneously various tech'!

- <u>Spark</u>: This, and more precisely <u>MLlib</u> seems to be the best for distributed computing applied to Machine Learning.
- <u>Hadoop YARN</u>: We have to manage Spark clustering and he seems to be a good fellow. (Or <u>ApacheMesos</u> or <u>AmazonEC2</u>?).
- <u>Tensorflow</u>: It's an OpenSource library created by Google and that allows us to implement DeepLearning. Far better that his first public release. (Or <u>Theano</u>?).

Java ? Scala ? Python ? Camel ? Cuda ? Quid of implementation ?

Data traitment in Tensorflow is realise in Pyhon or C++ (With <u>Anacona</u>?) We do not have any GPU on our serveurs, so ... No <u>CUDA</u> support

The agent (or sensor) is a <u>EIP</u> Java-based <u>Camel</u> whose function is to redirect specific data flows at Spark driver. It should be in the form of a lambda Minecraft plugin running by a bukkit server.



Why Tensorflow?

While other ML Framework/Lib like Caffe seems to be more powerful, Tensorflow is the best start for our first DeepLearning project thanks to his heavy online documentation and tutorial.

Liste de liens :

http://libccv.org/doc/doc-convnet/

https://code.google.com/p/cuda-convnet/

http://caffe.berkeleyvision.org/

https://github.com/deeplearning4j/deeplearning4j

https://skymind.io/

http://nd4j.org/

http://www.nextplatform.com/2016/02/24/hadoop-spark-deep-learning-

mesh-on-single-gpu-cluster/

http://playground.tensorflow.org/

https://www.college-de-france.fr/site/yann-lecun/seminar-2015-2016.htm

https://www.youtube.com/watch?v=9DZrjBKoxOI

https://www.nuget.org/packages/Hangfire.SqlServer.RabbitMQ/

http://torch.ch/

https://www.youtube.com/watch?v=sphFCJE1Hkl

https://www.youtube.com/watch?v=86qlpX23PLs

https://www.udacity.com/course/deep-learning--ud730

https://aws.amazon.com/fr/lambda/details/

http://deeplearning.net/software_links/

https://hypixel.net/threads/watchdog-roundup.749635/

http://www.datasciencecentral.com/profiles/blogs/top-10-machine-lear-

ning-algorithms

http://jmeter.apache.org/

http://blog.xebia.fr/2016/03/04/apache-kafka-une-plateforme-centra-

lisee-des-echanges-de-donnees/

http://camel.apache.org/

http://blog.soat.fr/2010/01/eip-quest-ce-que-cest/

http://hn.premii.com/#/article/12239096

http://hn.premii.com/#/article/12237998

http://www.iclr.cc/doku.php?id=iclr2016%3Amain#accepted_papers_

conference track

http://deeplearning.net/tutorial/deeplearning.pdf

http://www.cs.nyu.edu/~yann/talks/lecun-ranzato-icml2013.pdf

https://docs.google.com/presentation/d/1UeKXVgRvvxg9OUdh_UiC-

5G71UMscNPlvArsWER41PsU/

http://arxiv.org/pdf/1408.5093.pdf