Cyber Security
is an Application of Decision Sciences

or

Why Every Cyber Security Dept Needs a Decision Scientist

Presented to the IEEE Computer Society
10 June 2020
by Brad Morantz PhD

My Focus & Perspective

- I am a Decision Scientist
  - I focus on the problem at hand and how to make the best decision, using whatever I have or can get
  - My background is in electrical engineering, mathematics, CIS, and statistics
- My focus is on intelligence, cognition, information, and intelligent decision making
- I consider the cost of a wrong decision
- I leave the computer to the programmer & network engineer

The Paper

- This presentation is based upon the technical paper of the same name
- The paper is in process of being submitted for peer reviewed publication
- The focus is on the questions and how to arrive at good answers and decisions
- If a standard method was used then the “bad guys” would know how to get around it
- Implementation is left to the code people

Situation Overview

- Cyber Security is:
  - New frontier
  - Very important
    - Military
    - Financial
    - Life
  - Changing rapidly
- We need to protect our assets
- Losing the war
- Shortage of employed qualified people

Components

1) Continuous decision making*
2) Understanding what is happening
3) Technical implementation of decisions

   This presentation will focus on #1

*Life is continuous decision making

What is Cyber Security?

- Protection of Internet connected systems
  - Hardware
  - Software
  - Data
- Gate Keeper
  - Only let in those that are allowed
- Determine if not-allowed are already in
  - What action to take
    - Expel
    - Terminate
    - Changes in response to intrusion

Confidential Computing

- New security model
- Launched by Linux Foundation
- Supported by big tech companies
- Data protected
  - At rest
  - In transit
  - In use
    - Most difficult as it is in open and not encrypted
    - In Trusted Execution Environment (TEE)

Workforce

- Arizona Cyber Talent Organization Estimates
  - 2018: Arizona short 6875
  - 2018: USA short 285,000
  - 2019: USA short 1,500,000
- Hiring methodology
- Accredited degree programs in cyber security

Data Science

- Interdisciplinary field
- Uses scientific methods, processes, algorithms, and systems
- Extracts knowledge and insights from data in various forms
  - Structured
  - Unstructured
- Uses computers, Statistics, & AI
- Finds meaning in “big data”
- Focus is on the data and what it says

Decision Sciences

- Science of making high quality decisions using
  - Computers
  - Mathematics & Statistics
  - Artificial Intelligence (AI)
  - Modeling & Simulation
  - Game Theory
  - Psychology
  - Science & Engineering (domain knowledge)
  - Cost of wrong decision
  - Thinking & problem solving
  - Focus is on the decision

Applications of Decision Sciences

- Best possible decision free of human bias
- Image pattern recognition
- Medical
- Business
- Financial
- Military
- Security including cyber security
- Daily life

Advantages to Decision Sciences

- Free of human bias
- Considers cost of wrong decision
- Multi-disciplinary
- Can create decision support systems (DSS)
- Can work with a data scientist
- Completes the decision process

The Best Cyber Security Team

- Data Scientist
  - To Understand the data
- Decision Scientist
  - To know how to make the decision
  - Info from Data and Computer Scientists
- Computer Scientist
  - To see what is going on in the data feed
  - To implement what the team needs

Cyber Security Decisions

- Do we allow the entity that is trying to enter?
- Do we already have someone in our protected place?
- If we do, what to do?
  - Show the exit?
  - Make changes?
  - Terminate?
- Should we buy insurance?
- What can we learn?

Risk Assessment

- Know where is each risk
- Know what is each risk
- Understand vulnerability
- What are the assumptions
  - Which are questionable
  - Which are solid
- What is the cost of a breach at each place
- Estimate long term effects of a breach

What is Network Security

- Protection of the access to files and directories in a computer network against hacking, misuse and unauthorized changes to the system
- Protecting the system from unwanted intrusions
  - Looking at the data
  - Taking the data
  - Changing the data
  - Changing the program/algorithm
  - Malware
  - Spies or Trojans

Analogy

- We have a building
- Something valuable inside
- Security guard at entrance
- Some allowed access to item; some, not
- Guard decides who to allow in
- Building also has other “openings”
- Maybe:
  - Forgot to close
  - Double agent opened it
  - Trojan horse

What if Adversary Already Inside

- First decision: is agent already inside?
- Next decision: did agent get to “item”?
  - No, then decide what action to take
    - Show exit but first get good image of agent?
    - Terminate agent?
    - Changes to be made?
  - Yes then decide
    - Just show exit?
    - Call police?
    - Terminate agent and how?
    - Changes to system?
    - Changes to “item”?

Set a Trap

- Decide if this is an adversary
- Make the path to the trap realistic
- Lead down hallway to either trap or impostor
- Decide if agent should leave with impostor
  - Would be dis-information
- Allow to leave or arrest?
- What if adversary gets real item?
- More questions/decisions

History of Security

- Encrypted messages via carrier pigeon
- WWII trivia & cultural questions
- Pass phrases & counter pass phrases
- IFF (Identify Friend or Foe, Maybe)
- Computer passwords
- Mother’s maiden name
- Complicated passwords
- Two-stage verification
- USB key, finger print

Malware Detection & Prevention

- Protected sandbox
- Pattern recognition
- Turn off Java
- Disconnect USB socket
- Occam’s razor – law of parsimony
- Keep it simple – look at simple things
  - Time of day
  - Human pattern recognition – feels bad

Simple Pattern Recognition

- Get history and look for patterns
  - Stolen credit card and gas pumps
  - Going to store at closing time
- Learn from the data
- Find patterns:
  - Proper request
  - Invalid request

Transport Layer Security Inspection

- Known as TLSI or TLS
- Replaced Secure Socket Layer or SSL
- Decrypt incoming traffic
- Inspect it
- Decide if legit
  - Legit: re-encrypt and send it on
  - Illegit: decide how to handle but do not allow in
    - Maybe send down rabbit hole to honeypot and mislead
    - Maybe cut off
    - Know that you are under attack

Pattern Recognition

- Cognitive process
- Matching information from stimulus to information in long term memory (LTM)
- People are very good at this
- Human in the loop (HITL)
  - Speed is a limitation in cyber world
  - Can find patterns in R&D phase

Cost of Wrong Decision

- Must consider the cost of a wrong decision
  - Medical example
  - Retail store purchase
  - Automatic Target Recognition (ATR)
- Must bias the system
  - Sometimes would rather be more careful
  - Sometimes costs more to be too careful than loss
    - Credit card losses
  - OK to ask for extra information

Pattern Recognition Methods

- Template Matching
- Prototype Matching
- Feature Analysis
- Statistical Pattern Recognition
- Recognition by Components
- Top Down Processing
- Bottom Up Processing
- Model Based

Template Matching

- Library of templates stored in LTM
- Simplest method compares input to each template until a match is found
- Cross Correlation Analysis
- Problem with generalizability
  - Angle, distance, color, size, etc
- Good for SIMD
  - Multiprocessing

Prototype Matching

- LTM patterns are general or average
- Greatly reduces search space & time
- Matches to a classification
  - To average or class
  - e.g. car but not specific make
- Much faster
- Some problems with generalizability
  - Angle, distance, color, etc

Feature Analysis

- More cognitive process
- Set of features & behaviors compared as a set
- Increasing complexity
- Uses logic, reasoning, & variety of information
- Can use a variety of sensors & sensor types
- Can use a variety of information types

Statistical Pattern Recognition

- Use various statistical methods to find pattern
  - Cluster analysis
  - Self organizing map (SOM)
  - Regression
    - OLS
    - Robust
    - Quadratic
  - Multidimensional Scaling MDS
  - Artificial Neural Networks (ANN)
  - Etc, etc

Cluster Analysis

- Plot each observation in N space
  - A dimension for each variable
  - Most observations will fall into clusters
    - e.g. Big people have big shoes
- Remove variables that do not contribute
  - Or more noise than help
- Identify what each cluster is
- See which cluster new item fits into

Regression

- Use Occam’s razor, start simple
- Ordinary least squares (OLS)
- Robust regression uses median instead
- Quadratic regression uses higher order function
- T statistic can help remove weak variables
- Fisher or F statistic tells how good is the model
- R^2 tells how much the model explains
- Builds a model that can predict classification

Support Vector Machine

- For dichotomous decisions
  - e.g. allow or not allow
- Tries to separate data into two groups
- Not necessarily linear depends on kernel
- Puts fence between them
- Separation distance indicates how strong it is

Artificial Neural Networks

- Based on basic design of biological
- Somewhat like regression
- Many more variables
- Various activation functions
- Can fit non-linear and discontinuous data sets
- Does not make a model, it approximates one
- No knowledge in hand
- Universal function approximator

Recognition by Components

- Visual pattern recognition in homo sapien
- Objects converted into geometric shapes
- Identity in the shapes and relation to each other
  - Table, bicycle, etc examples
- Hough Transform

Top Down Processing

- Starts with previous knowledge in LTM
- Sensors pre-process the data
- Make the best guess based on above
  - Optical illusions

Bottom Up Processing

- Data driven method
- No logical or cognitive processing
- Directly determines identity
- Frog example

Model Based

- Used by Decision Sciences
- Build a model of what it is thought
- Compare item in STM to the model
- Model can incorporate various types of data from various sensors and LTM
- Airplane example
- MBDSS
  - Plug observation into model
  - See if output is correct or closest to desired

Machine Learning

- Start with data
- Learn patterns & gain knowledge
- Uses many of the statistical methods
- Train system on data
- Maybe extract knowledge
- Use trained system to make prediction or classification

Training

- Like people, system must go to school & learn
- Supervised – each observation has answer
- Unsupervised – observation only, no answer
- Semi-supervised – some of the observations have an answer, most do not
  - Can greatly improve performance
- Cyber is new
  - Constantly changing
  - No long history

Clean the Data

- Old saying: “garbage in makes garbage out”
- Need to spend much time cleaning the data
- Missing or impossible values
  - Replace with average
  - Omit observation
- Not all in same metric
- Can take longer than the learning process
- My abalone paper

Time Series Forecasting

- Sometimes not all the variables are known
- Sometimes do not know all of the metrics
- Sometimes can not measure everything
- TSF works on principle that whatever has been going on is reflected in the criterion
- Whatever has been happening in the *near* past should continue into the *near* future
- Uses lagged value to forecast future
- Near is a relative term, depends on the data

Hybrids

- Putting together two or more methods
- Netflix competition
- Ensemble learning
  - Each type has its own area of expertise, just like people
  - Combining them increases power and performance
- Cyber is a time constrained system
- Use Hybrid for learning and R&D or use multiprocessing system

Synthetic Immune Systems

- First developed to study biological immune systems
- System tries to keep out invaders and deal with those that are already inside
- Same goals as cyber security
  - Protect the entity
  - Allow the right things (food, vitamins, etc)
  - Block bad things (virus, bacteria, etc)

Human Component

- Human in the loop (HITL)
- Many breaches caused by humans
  - Plug in infected USB drive
  - Click on link in email
  - Download malware from web page
- Human also excellent at pattern recognition
- Show image to human to see pattern
  - R&D
  - Not for time constrained situation

Effectiveness

- Hit matrix is simple
- Calculate percentage
- Many methods
  - Percent right
  - Percent saved
  - Sigma correct
  - Dollar saved
  - MAPE
  - Etc, etc

	Right	Wrong
Yes	137	12
No	162	15

True positive = 137/149 = 91.95%
False positive =12/149 = 8.05%
True negative = 162/177 = 91.53%
False negative = 15/177 = 8.47%

Cyber Insurance

- If there is a risk, someone will sell insurance
- Know the risk
  - How much is total $
    - Direct loss
    - Indirect loss
    - Long term
- Probability of occurrence X total $ = EMV
- EMV is expected monetary value
- If premium is more than EMV do not do it

Summary

- Network security is a series of decisions
- Stakes are high
- Time is short
- Need R&D
- Need skilled people

Resources

- www.ieee.org
- www.machine-cognition.com
- www.azcybertalent.org