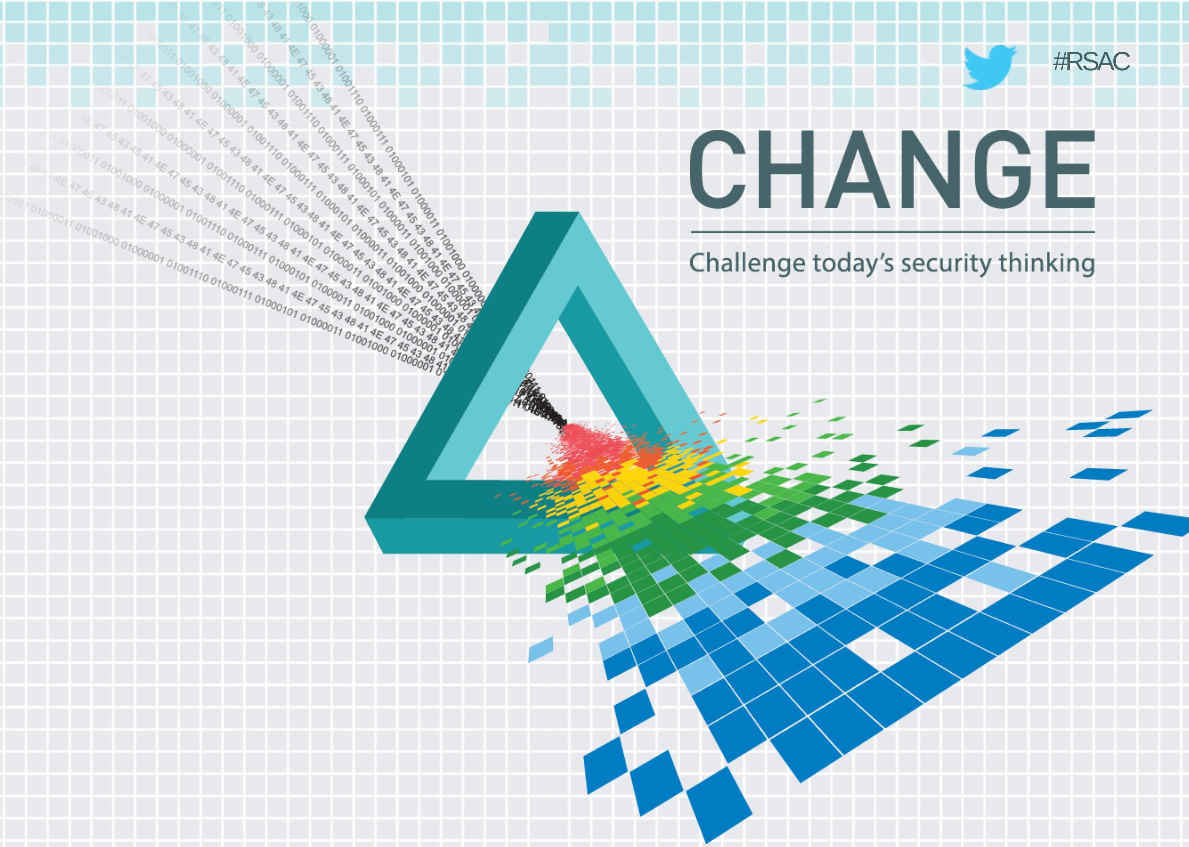
RSAConference2015

San Francisco | April 20-24 | Moscone Center



CHANGE

Challenge today's security thinking

SESSION ID: ANF-F02

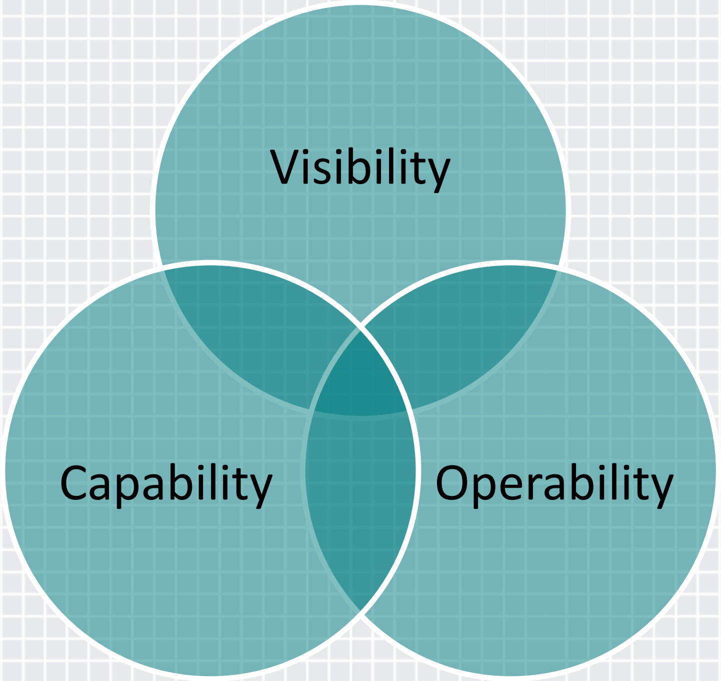
The Physics of Security

Andrew Rutkiewicz

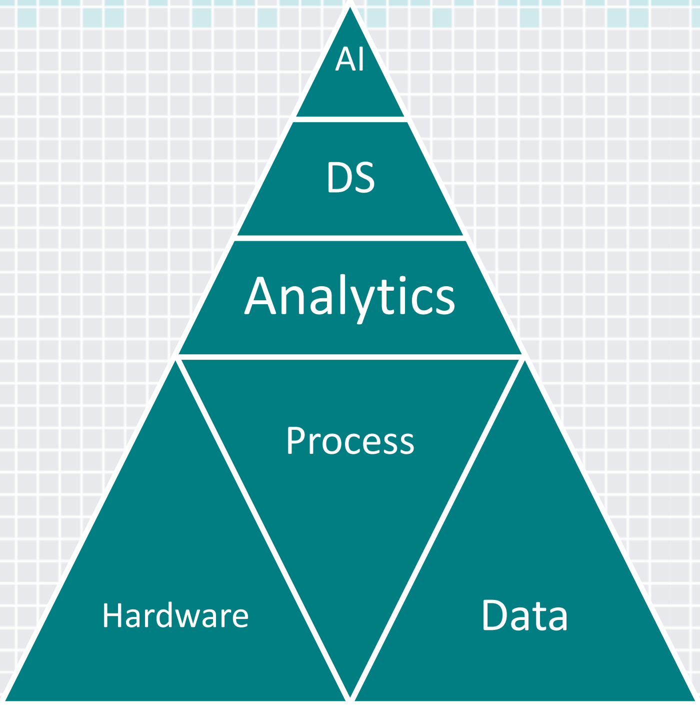
Principal IT Security Analyst
EMC
@packethawk

Challenges For (Analytics Driven) Security

- Visibility
- Normalization of Data
  - Both Packet and Log
  - Transaction Reconstruction
- Traditional Anomaly Analytics Fail
  - Misconfigurations
  - Broken Business Process
  - Can’t Operationalize
- No Standardized Measures or Models



Big Data Pitfalls



- Analytics
  - Apophenia
- Data Science
  - Perception Bias
- Machine Learning
  - Over Fitting
- Traditional analytic methods for network security carry high transaction costs and low yields
- Outcome: Negative ROI – This is changing

Physics and Its Applications

- Physics
  - Knowledge of Nature
- Applied Physics
  - Useful Application of the Knowledge
- Example: Light\Optics
  - Euclid, Alhazen, Newton, Hooke, Kao
  - 300AD First Studies of Light
  - 1973 First Fiber Optic Network

Entropy

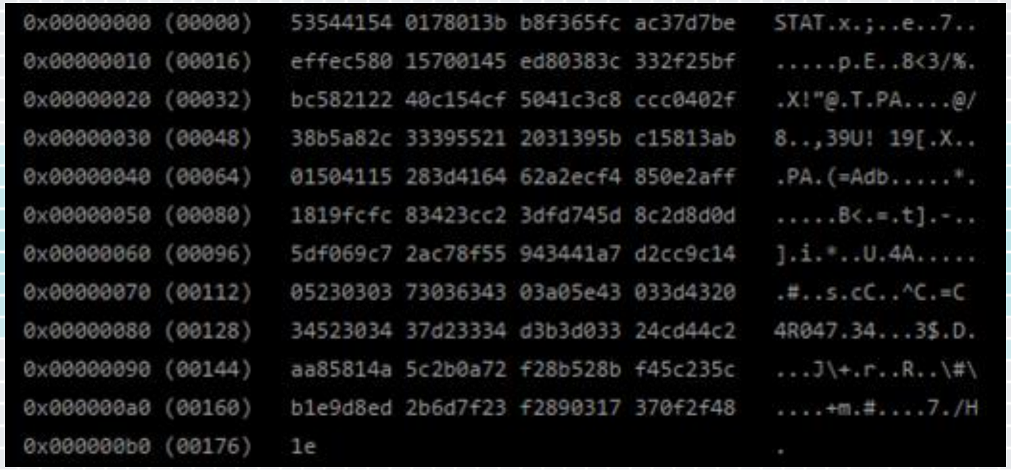
- Thermal Dynamics
  - Boltzmann and Gibbs
    - Extraction of Metals From Oxides
    - Melting/Boiling Point Manipulation
- Information Theory
  - Claude Shannon
    - Communication
    - Compression
    - Cryptanalysis

"surely must be one of the most important master's theses ever written… The paper was a landmark in that it helped to change digital circuit design from an art to a science." - **The Computer from Pascal to von Neumann**

By HH Goldstine

Entropy Hypothesis: RAT Detection

- Detecting Binary C2 communications
  - NON-HTTP Based
- Specifically APT RATs
  - 9002, Pivy, PlugX, Gh0st
  - All use compression and or encryption
- Descriptive Based Detection
  - Non Signature Based
  - Non IoC based



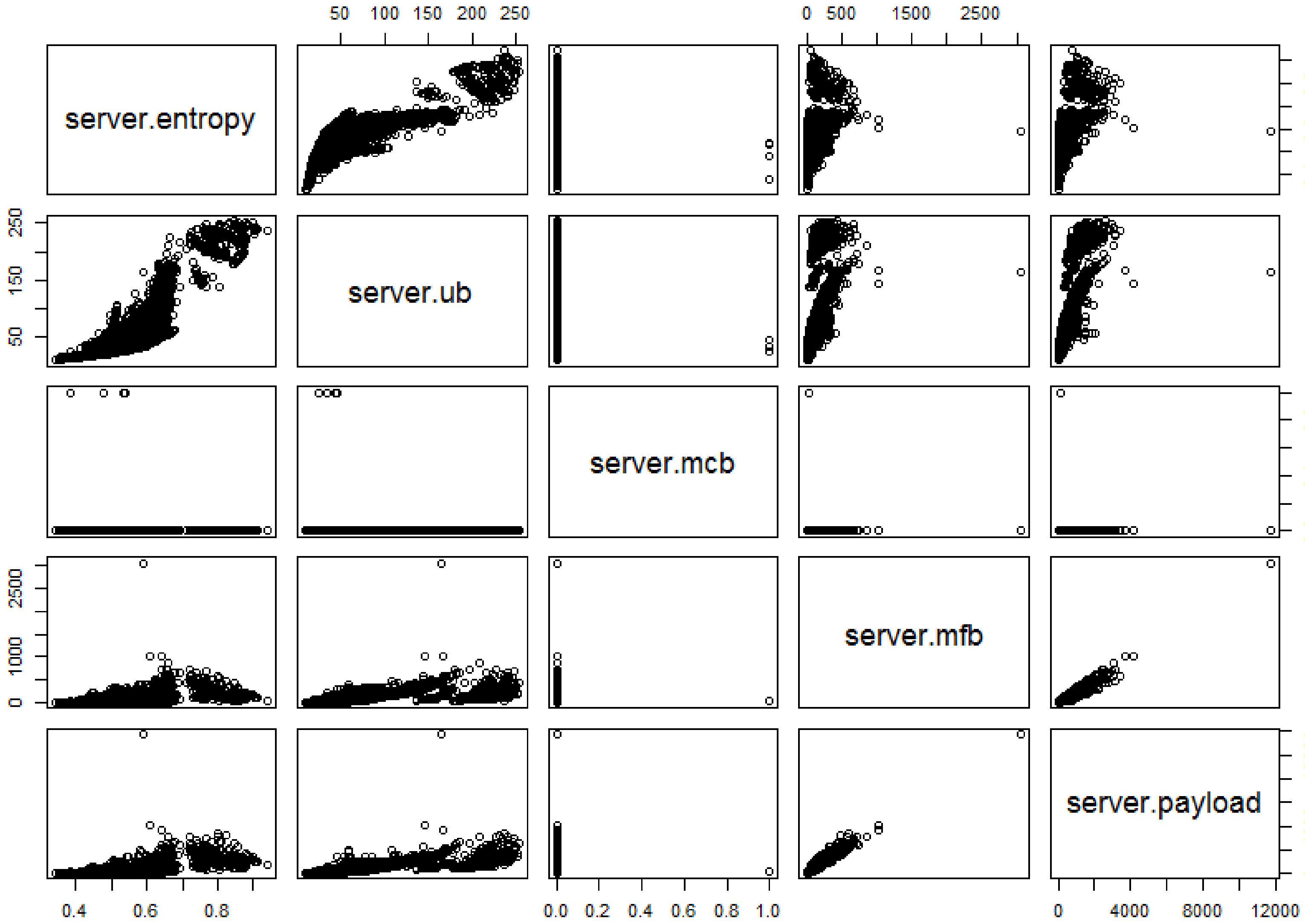
```
0x00000000 (00000)   53544154 0178013b b8f365fc ac37d7be   STAT.x.;..e..7..
0x00000010 (00016)   effec580 15700145 ed80383c 332f25bf   .....p.E..8<3/%.
0x00000020 (00032)   bc582122 40c154cf 5041c3c8 ccc0402f   .X!"@.T.PA....@/
0x00000030 (00048)   38b5a82c 33395521 2031395b c15813ab   8..,39U! 19[.X..
0x00000040 (00064)   01504115 283d4164 62a2ecf4 850e2aff   .PA.(=Adb.....*.
0x00000050 (00080)   1819fcfc 83423cc2 3dfd745d 8c2d8d0d   .....B<.=.t].-..
0x00000060 (00096)   5df069c7 2ac78f55 943441a7 d2cc9c14   ].i.*..U.4A.....
0x00000070 (00112)   05230303 73036343 03a05e43 033d4320   .#..s.cC..^C.=C
0x00000080 (00128)   34523034 37d23334 d3b3d033 24cd44c2   4R047.34...3$.D.
0x00000090 (00144)   aa85814a 5c2b0a72 f28b528b f45c235c   ...J\+.r..R..\#\
0x000000a0 (00160)   b1e9d8ed 2b6d7f23 f2890317 370f2f48   ....+m.#....7./H
0x000000b0 (00176)   1e                                    .
```

The Entropy Experiment

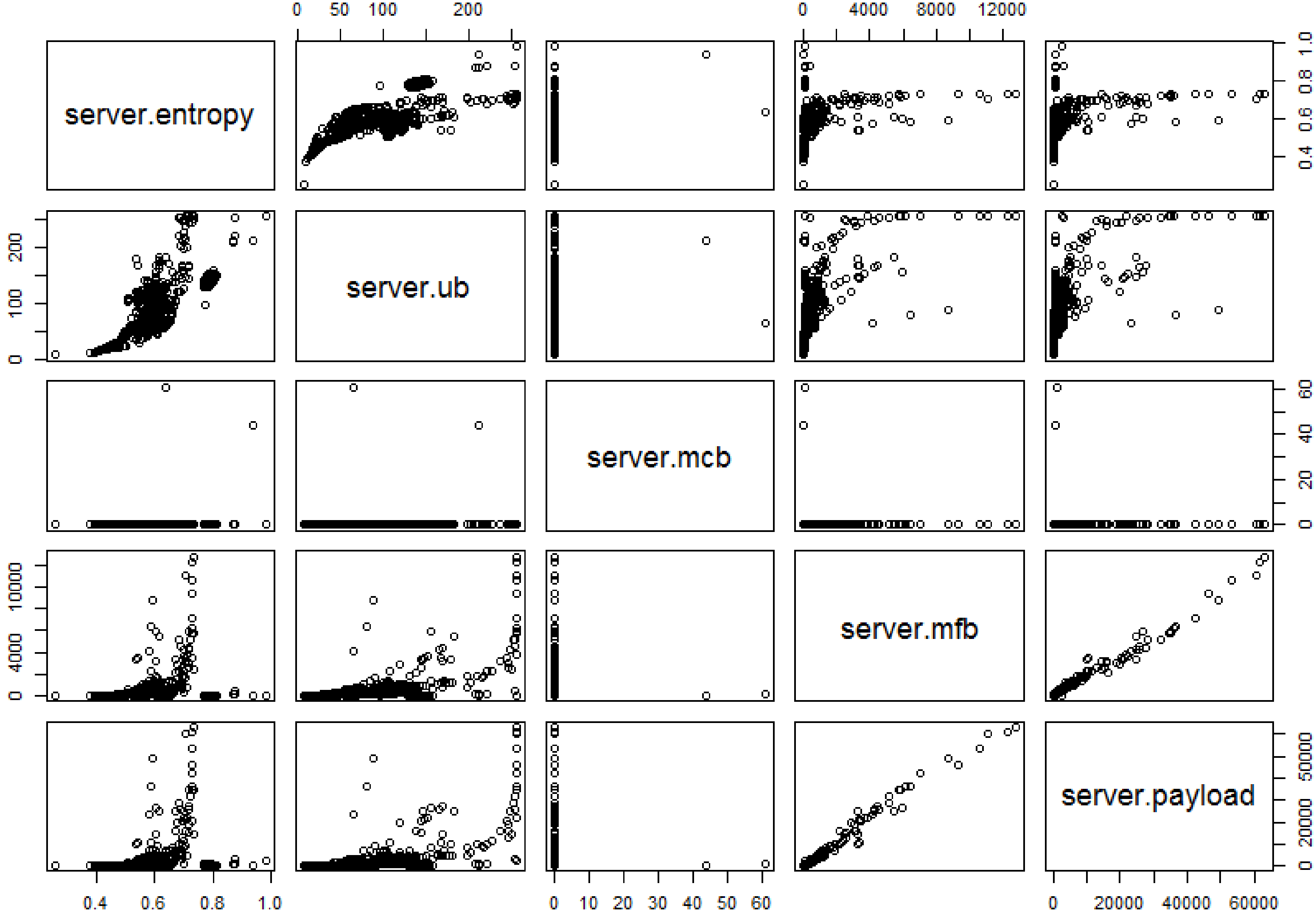
- Calculation of entropy for network traffic
  - Most common C2 channels (20 different TCP/UDP Ports)
- Basic Byte Frequency Measures
  - Most Common Byte (MCB)
  - MCB Frequency (MFB)
  - Unique Bytes (UB)
- Analysis Applications
  - Variance from known protocols
  - Obfuscation, Compression, and Encryption Detection
  - Encoding, Key Space Usage

Results (Still a WIP)

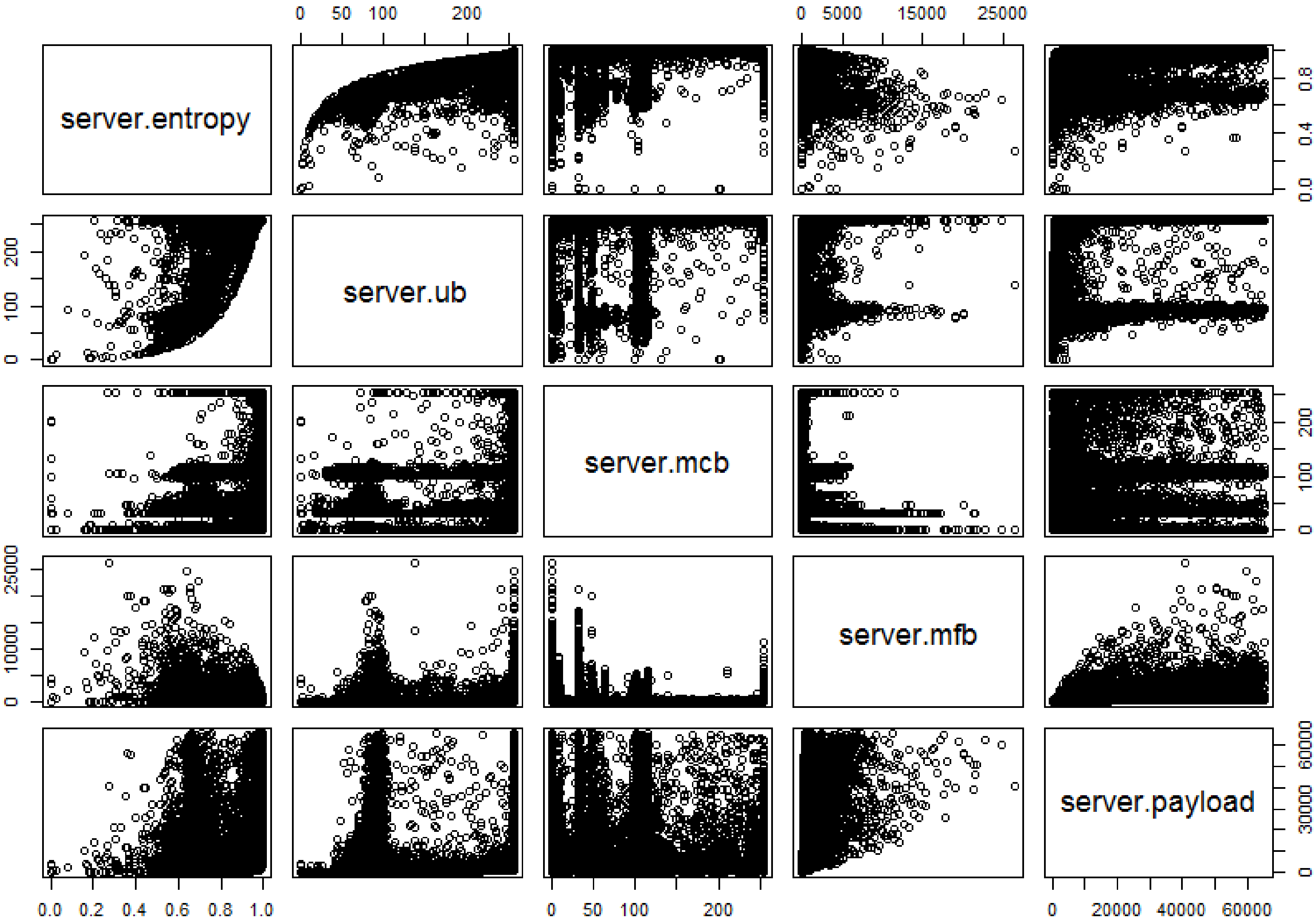
- Encoded and Compressed Data Have Predictable Patterns
  - 39U 19!
  - \x4B63\x6060 → Gh0st
  - LZ Artifacts other than 789C
- Scalability Concerns
  - Entropy calculation at line speeds is difficult
- DNS Anomalies
  - AV Exfil
- Pretty Pictures


server.entropy
server.ub
server.mcb
server.mfb
server.payload
50
100
150
200
250
0
500
1500
2500
0.4
0.6
0.8
0.0
0.2
0.4
0.6
0.8
1.0
1000
4000
8000
12000
10000

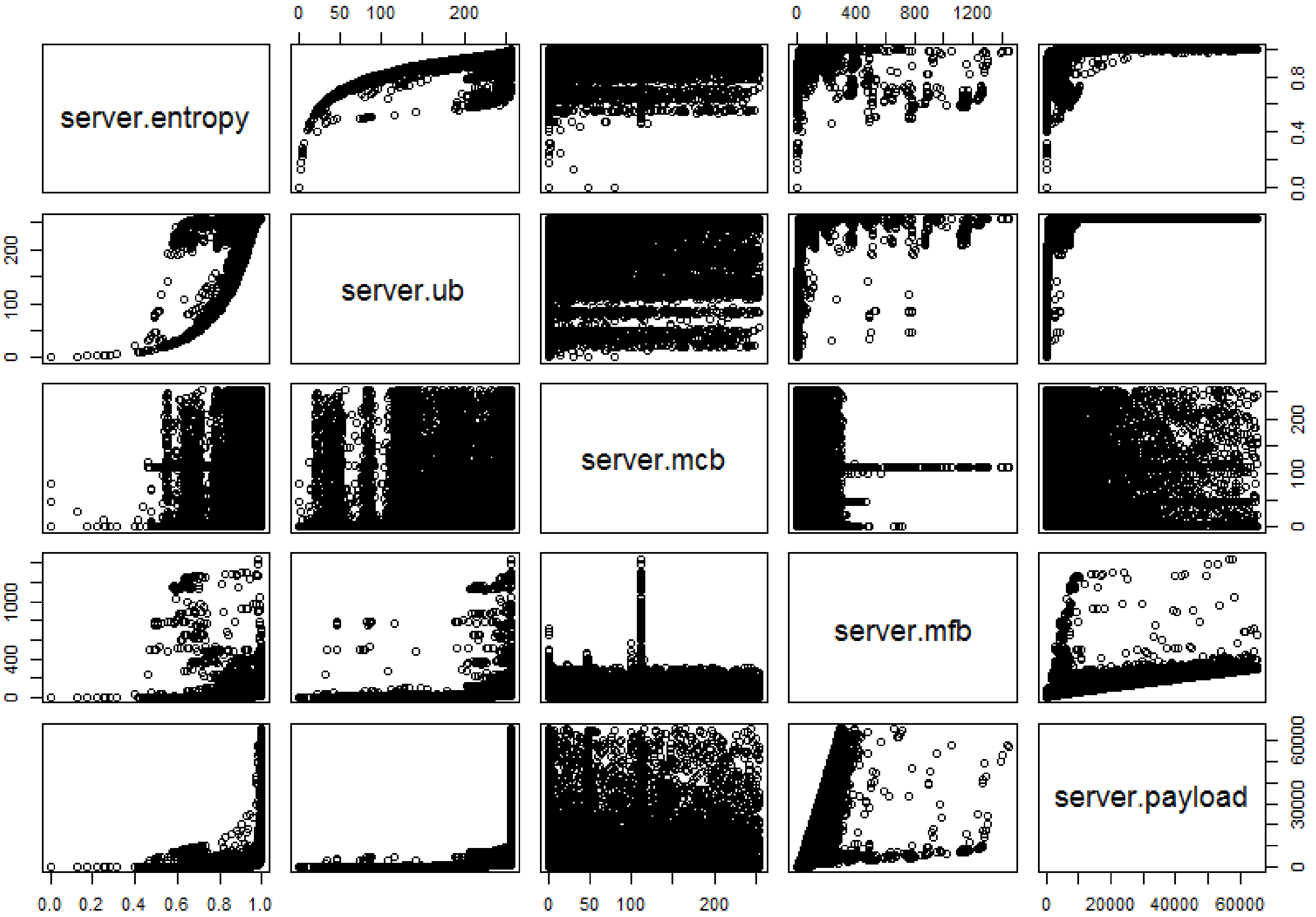
UDP 53


server.entropy
server.ub
server.mcb
server.mfb
server.payload
0
50
100
200
0
4000
8000
12000
0.4
0.6
0.8
1.0
0
100
200
0
20
40
60
0
4000
10000
0
20000
50000
0.4
0.6
0.8
1.0
0
10
20
30
40
50
60
0
20000
40000
60000

TCP 53


server.entropy
server.ub
server.mcb
server.mfb
server.payload
0
50
100
200
0
5000
15000
25000
0.0
0.4
0.8
0
100
200
0
100
200
0
10000
25000
0
30000
60000
0.0
0.2
0.4
0.6
0.8
1.0
0
50
100
200
0
20000
40000
60000

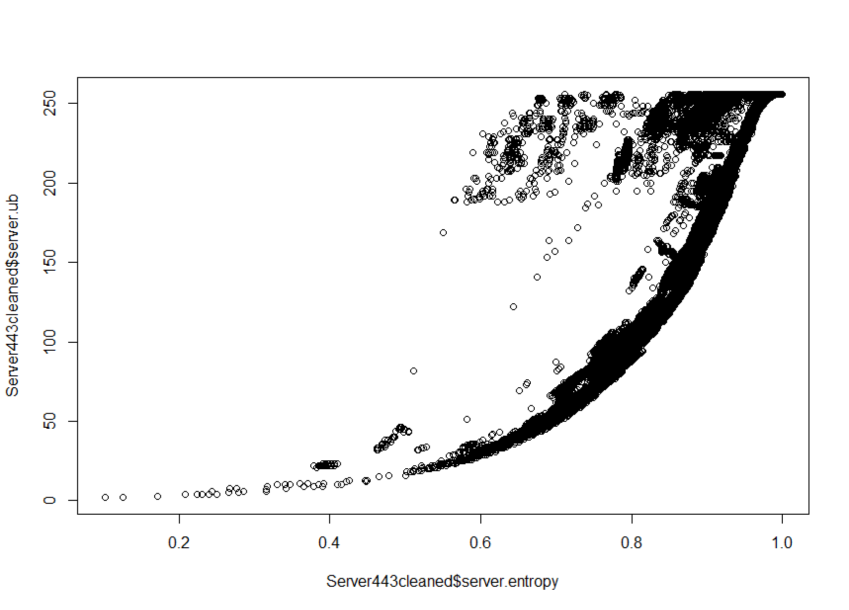
TCP 80


server.entropy
server.ub
server.mcb
server.mfb
server.payload
0
50
100
200
0
400
800
1200
0.0
0.4
0.8
0
100
200
0
100
200
0
400
1000
0
30000
60000
0.0
0.2
0.4
0.6
0.8
1.0
0
50
100
200
0
20000
40000
60000

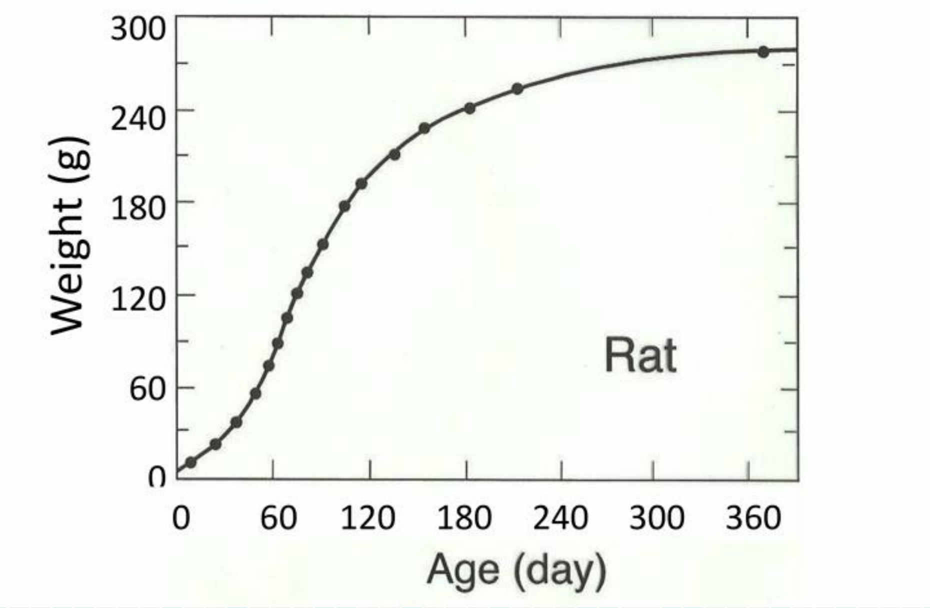
TCP 443

SSL Entropy vs Biological Growth

Unique Bytes Used vs Entropy



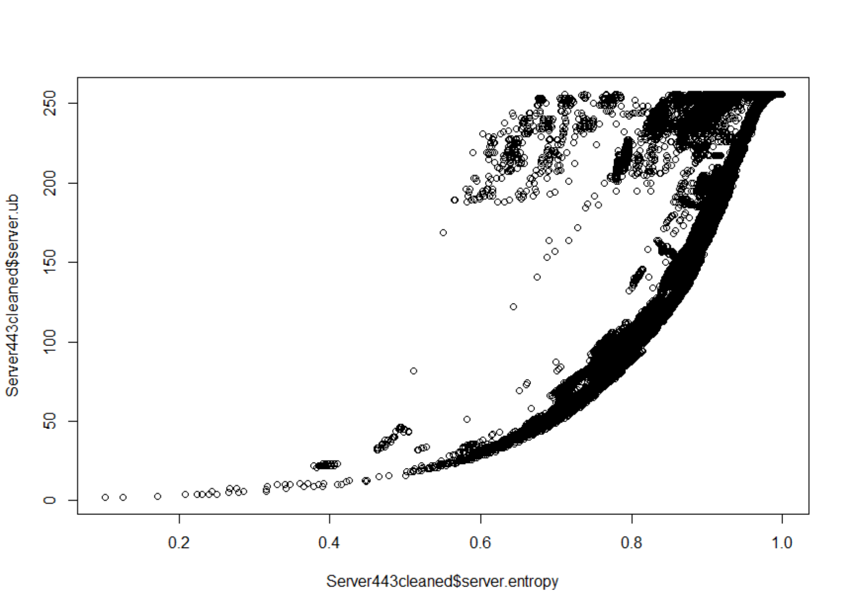
Rat Weight vs Age



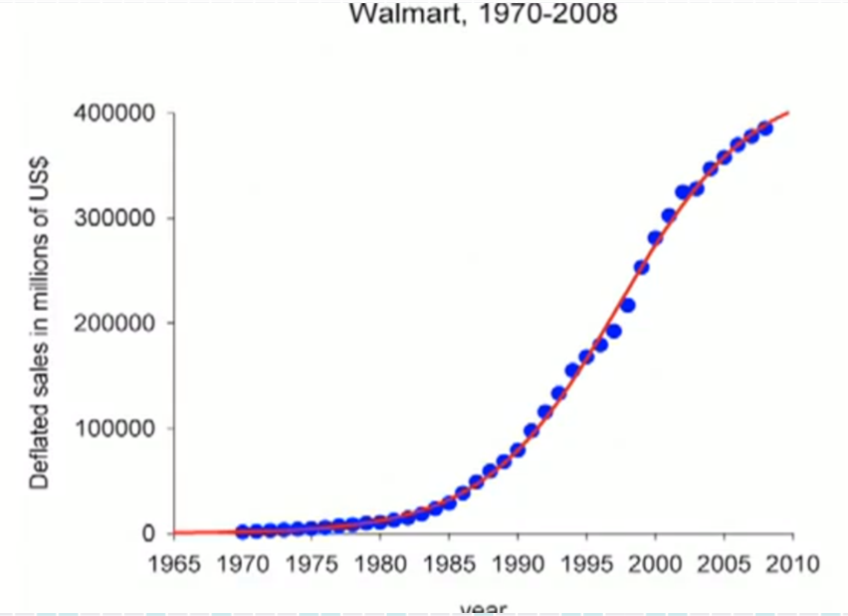
(From Geoffrey West, Ted Talk, July 2011)

SSL Entropy vs Commercial Growth

Unique Bytes Used vs Entropy



Walmart Sales vs Age



(From Geoffrey West, Ted Talk, July 2011)

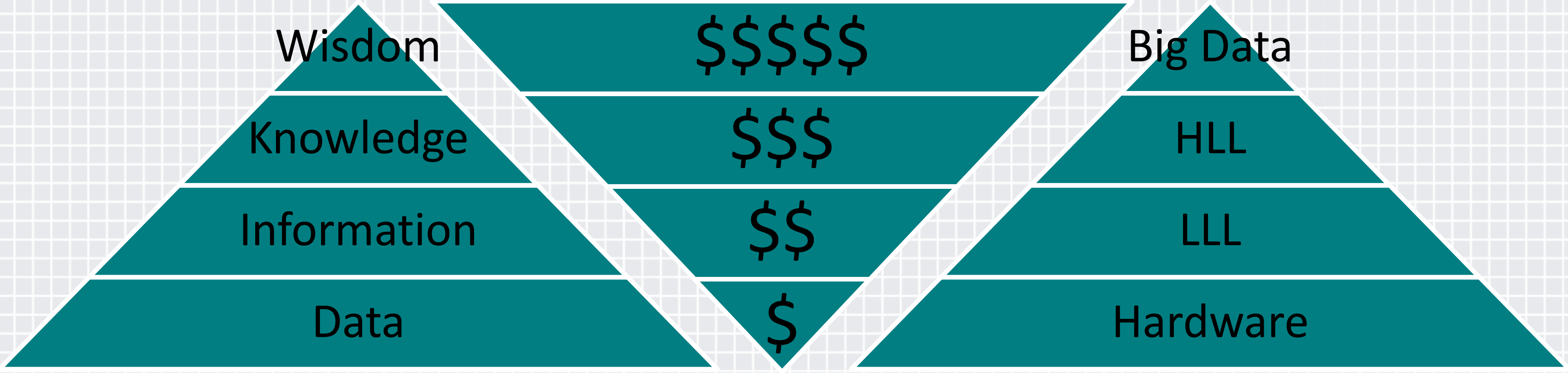
Universal Driving Forces

- Growth
  - Sigmoidal Curve or S Curve
  - Lag, Log, Decel, Plateau

- Economies of Scale
  - Parabolic Curve
  - Advantage, Neutral, Disadvantage

- These forces are as important in the understanding of the data as they are in the system they are built upon.

Cost Benefit Analysis



- Data is cheap
- Data enrichment at collection time is almost as cheap as raw data
- Post processing and enrichment costs grows as you go up levels of abstraction
- “Wisdom is not tactical”

Going Beyond Entropy

- Purpose Built Hardware
  - ASICs
  - DSPs
- Wave Equation
  - Application of frequency, amplitude and wavelength
  - Additional quantitative measures
- Timing Based Analysis
  - Kaminsky BlackOps
- ROWHAMMER
  - Proof physics rule HW and all it is built upon (IMO)

Summary

- Security analytics are still in the lag stage
- Statistics are better than intuition
  - Physics are better than statistics
- Entropy is one of many measures available
  - But an important one
- Growth and Scale
  - Leverage economies of scale
  - S curve as a forecasting tool
- Game Theory Considerations
- **As a community we must move from an art form to a science!**

Apply What You Have Learned Today

- Next week you should:
  - Identify where your organization is on the growth chart
- In the first three months following this presentation you should:
  - Inventory visibility and current data sets
  - Assess operational feasibility of analytics program
- Within six months you should:
  - Evaluate options between DIY or turnkey
  - Establish a plan for partnering with BI teams to conduct a POC for a practical and achievable use case.

Thank you!

Questions?