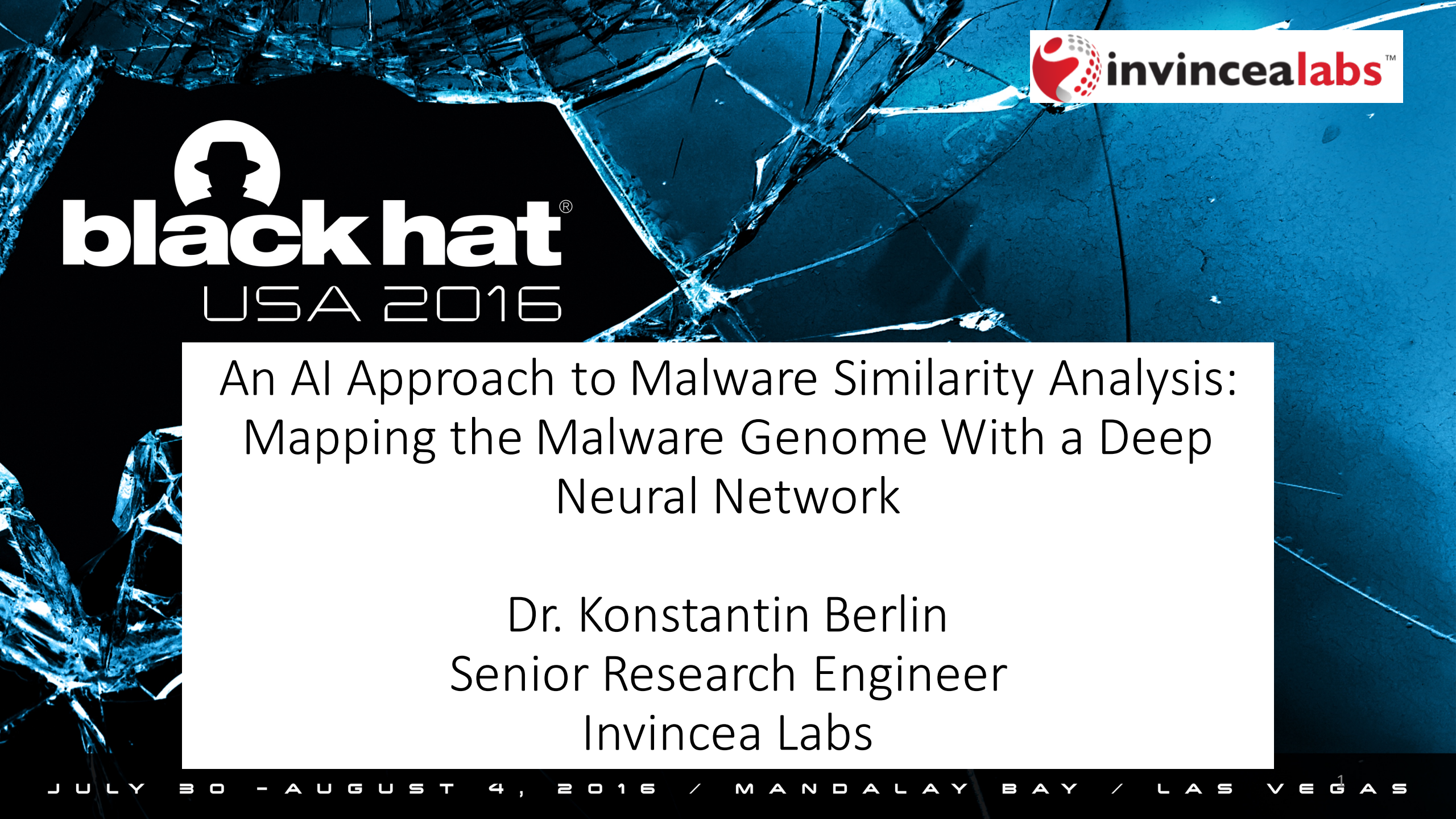


An AI Approach to Malware Similarity Analysis: Mapping the Malware Genome With a Deep Neural Network

Dr. Konstantin Berlin
Senior Research Engineer
Invincea Labs

Enterprise Networks are Under Constant Attack

- Intelligence is critical for prevention
- Cyber defenders are overwhelmed
- AI can help find important relations

Number of Network Breaches Per Year
(Verizon's 2016 Data Breach Investigations Report)

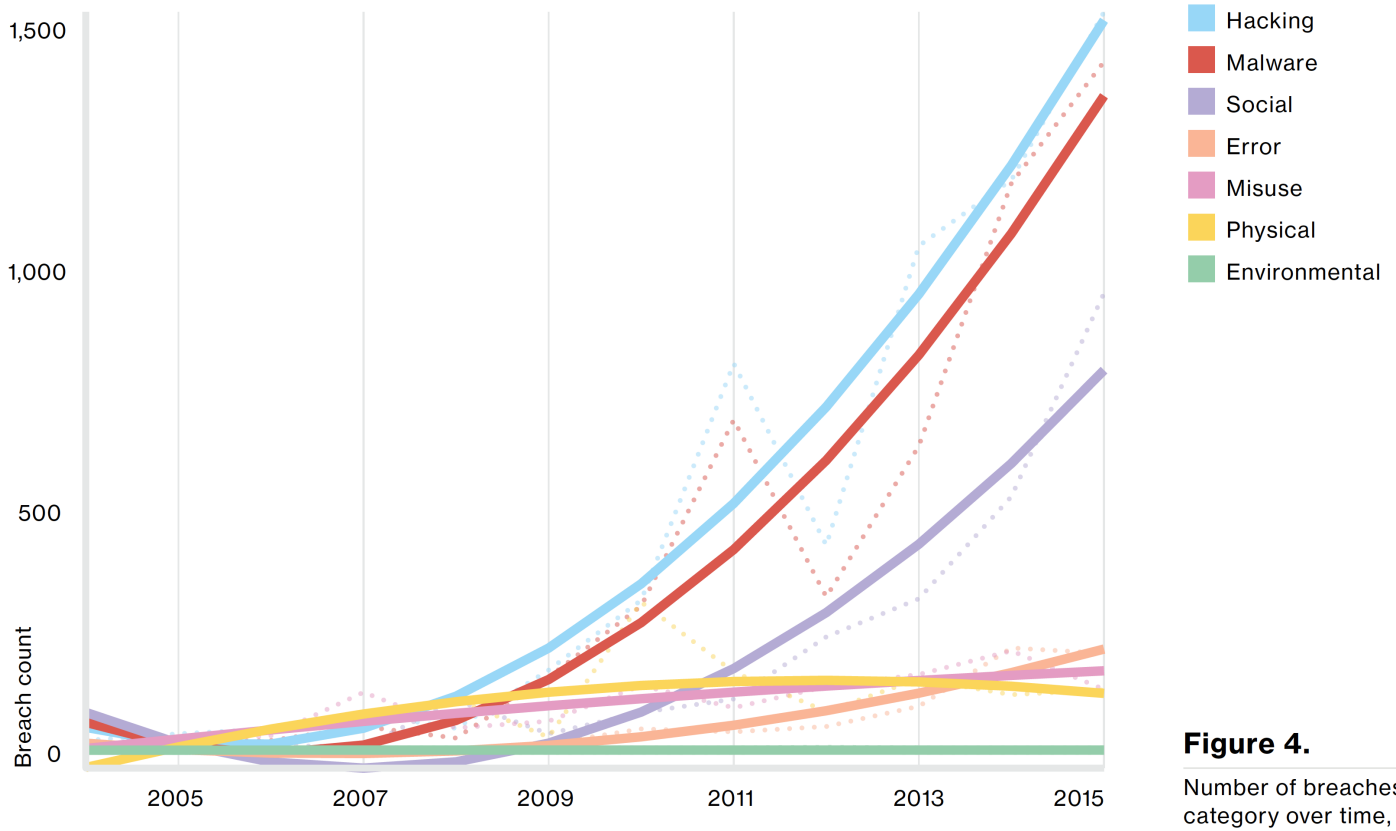
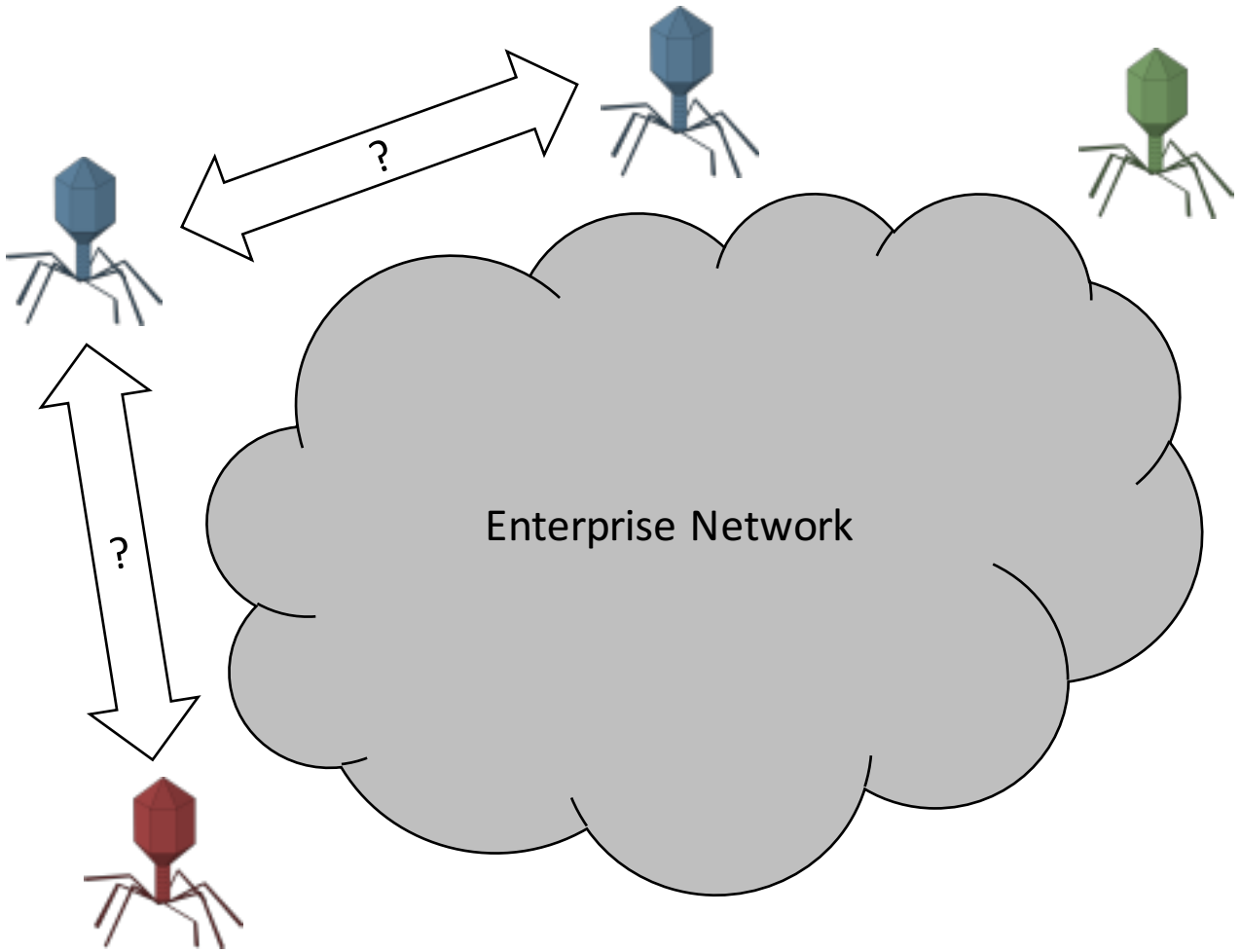


Figure 4.

Number of breaches per threat action category over time, (n=9,009)

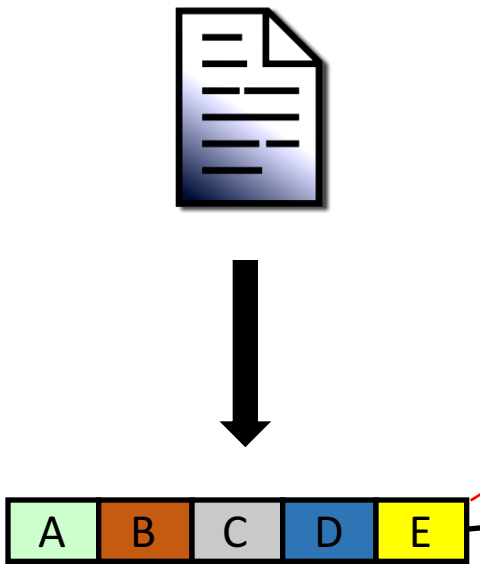
Intelligence through Malware Triage



- Benefits
  - Identify threat actors
    - Link various attacks to a single actor
  - Quickly understand functionality
    - Speed up reverse engineering
  - Mitigation
    - Signatures
    - Network Rules

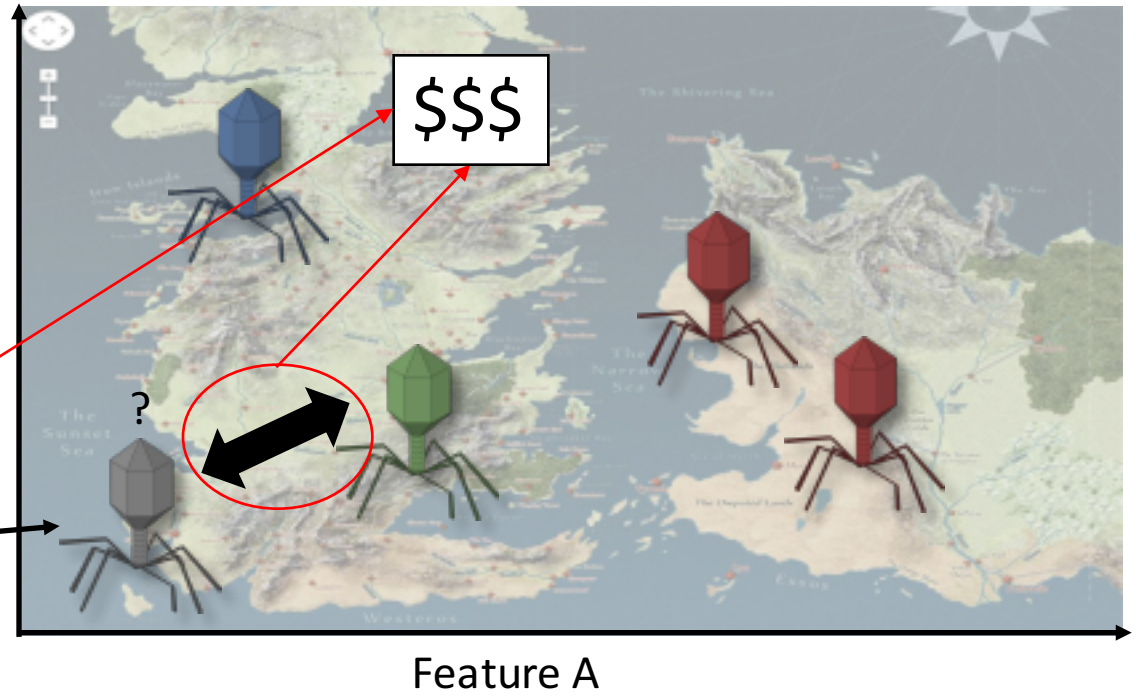
Nearest-Neighbor Classification

Attribute Extraction



Attributes
- Byte n-grams
- Opcode n-grams
- Printable strings
- System calls
- …

Attribute **Embedding**

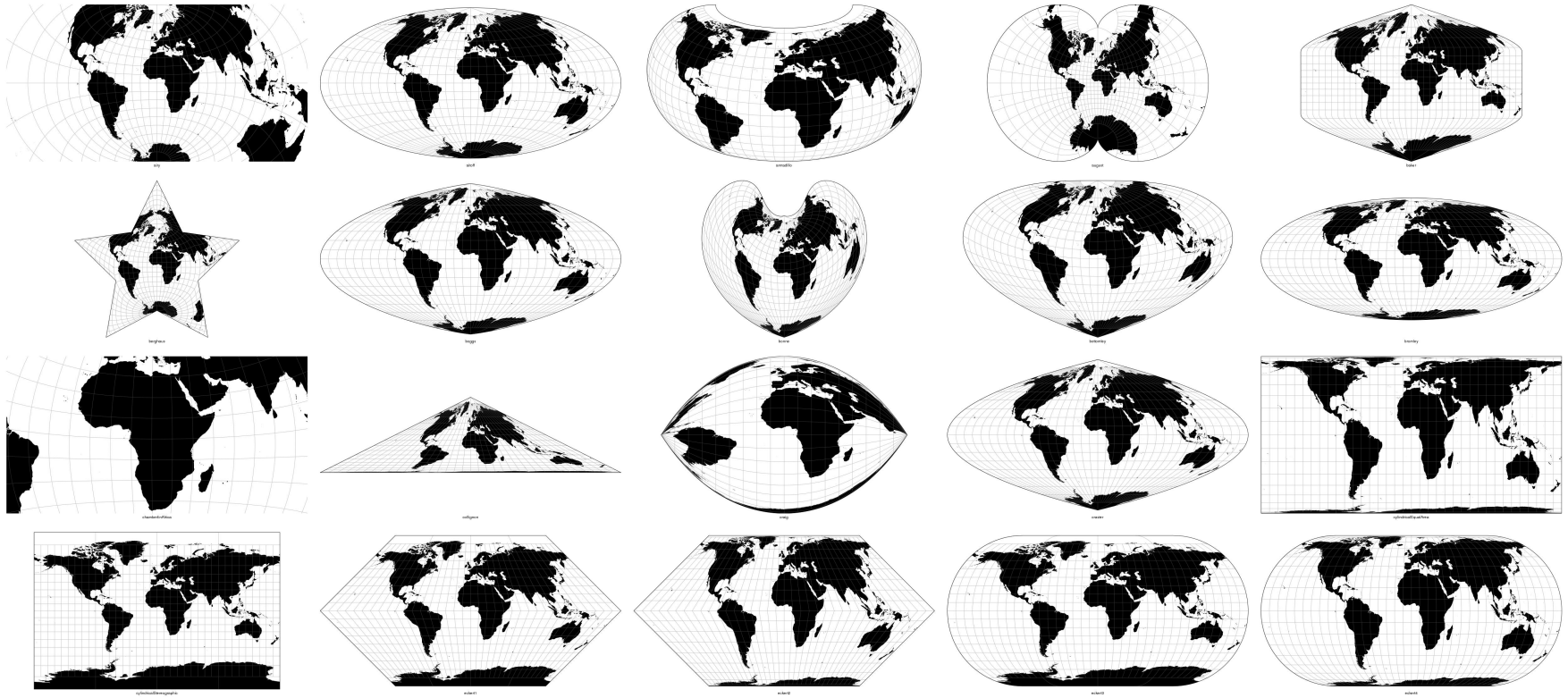


Similarity Search
- MinHash
- Feature hashing
- Other sketching
- …

Jang, Jiyong et. al. *Proceedings of the 18th ACM conference on Computer and communications security*. ACM, 2011.
Sæbjørnsen, Andreas, et al. *Proceedings of 18th international symposium on Software testing and analysis*. ACM, 2009.
Bayer, Ulrich, et al. *NDSS*. Vol. 9. 2009.
…Many more

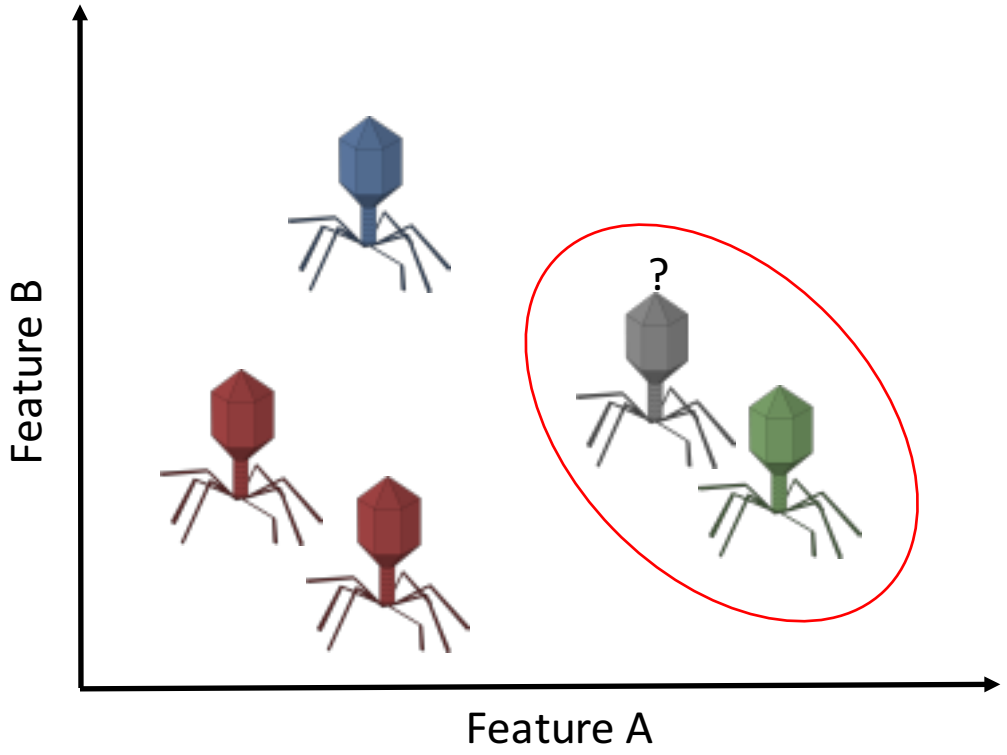
What Can Go Wrong with Embedding?

- Embeddings skew distances
  - Same embedded data can give different neighbors (ex. Alaska and Russia)

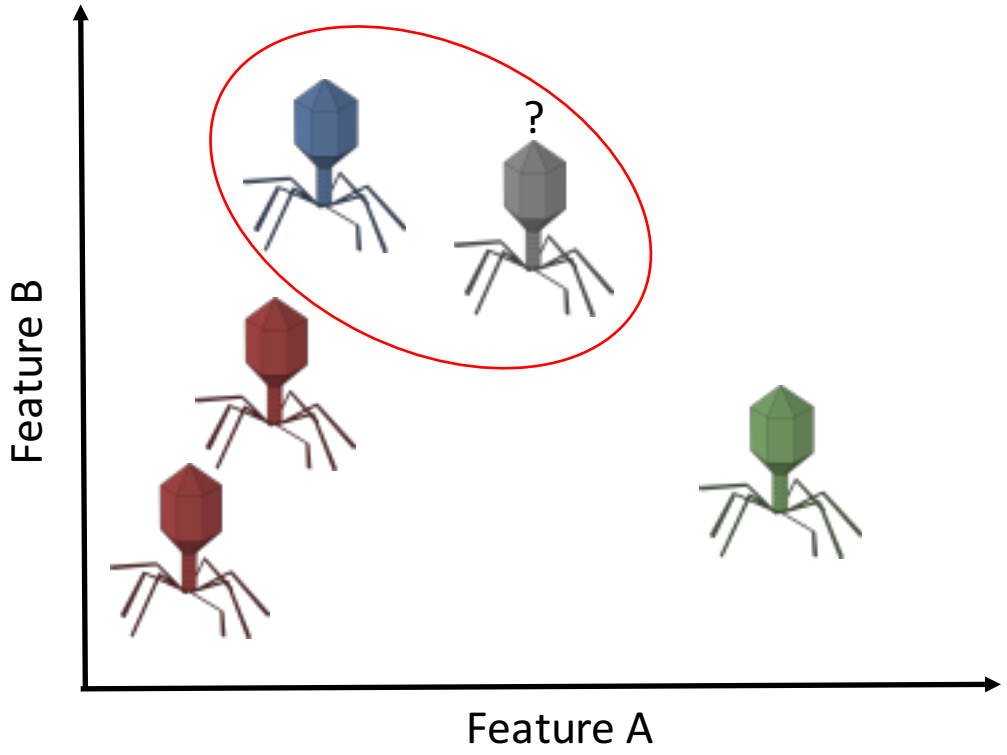


Issues with Attribute Embedding

Possible Attributes #1

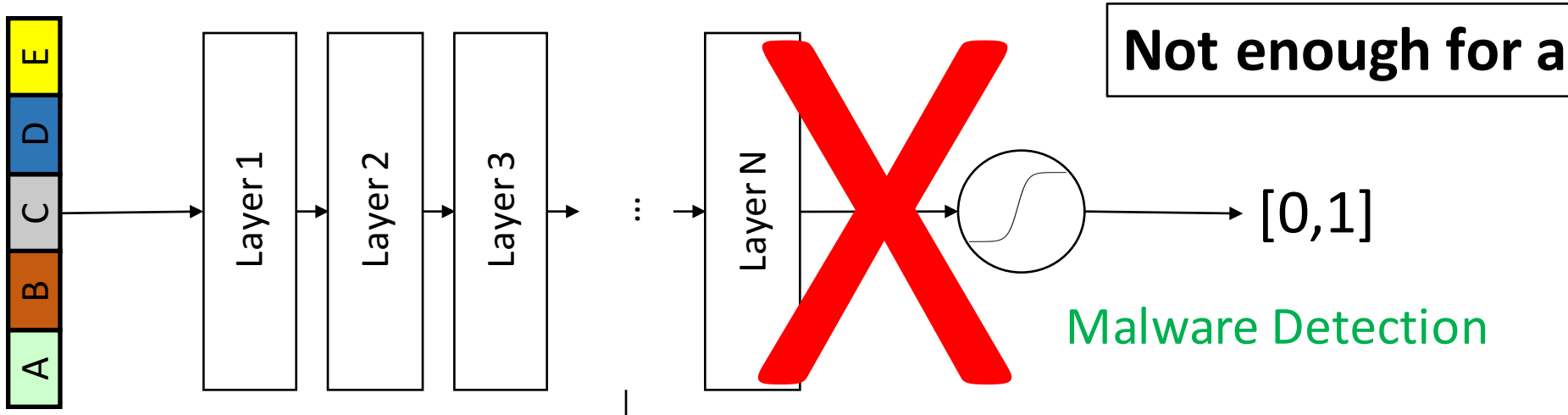


Possible Attributes #2



How to get consistent results, regardless of features?

Supervised Classification (Endpoint Solution)

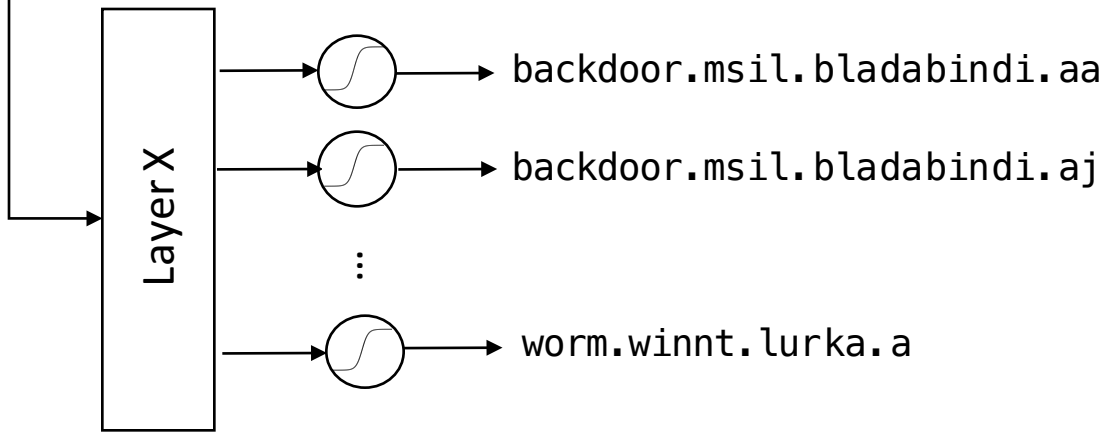


Not enough for a triage system!

Malware Detection

Joshua Saxe and Konstantin Berlin, *(MALWARE)*. IEEE, 2015.

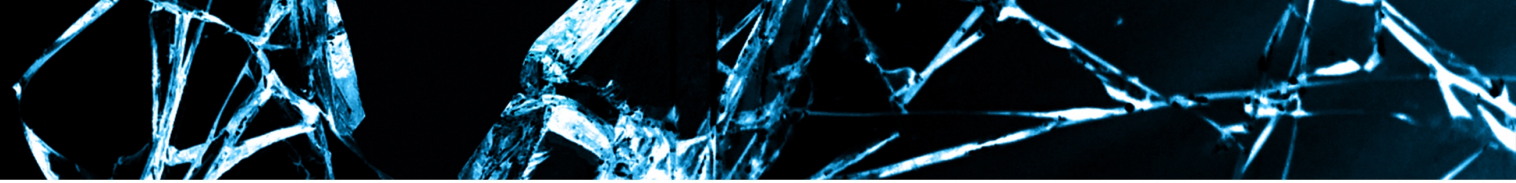
Categorical Classification



X by Invincea

0.97 F1-score (precision and recall)

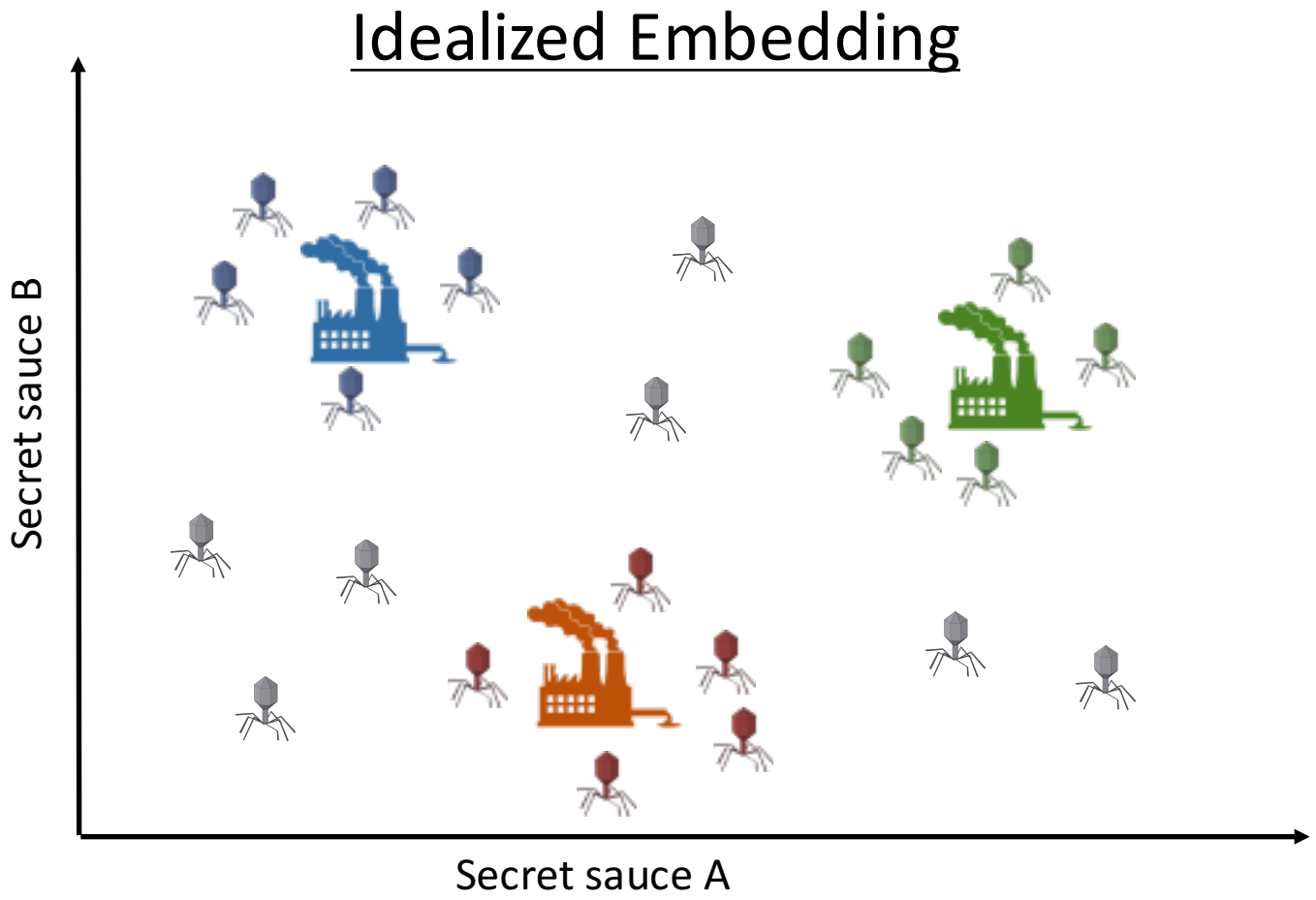
- 1500 Microsoft Families
- 2.0M Training files



How do we map malware into an **embedding** so that distances make **semantic** sense?

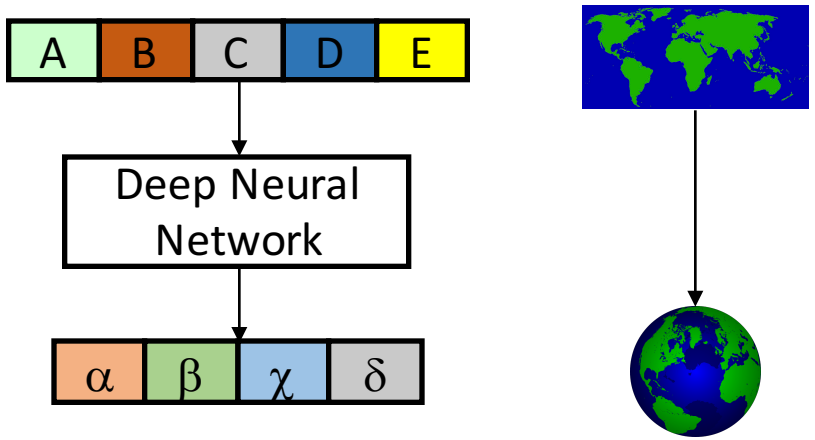
Imaginary World of Malware Factories

- Ideal World
  - Each hidden factory produces one malware family/variant
  - Factories are positioned relative to what and how they exploit vulnerabilities
- …but this not what we have!?



There is No ~~Spoon~~ Embedding…

- We created the embedding when we selected the features
- We can morph them in any way we choose
- One good way to morph the features is using a deep neural network

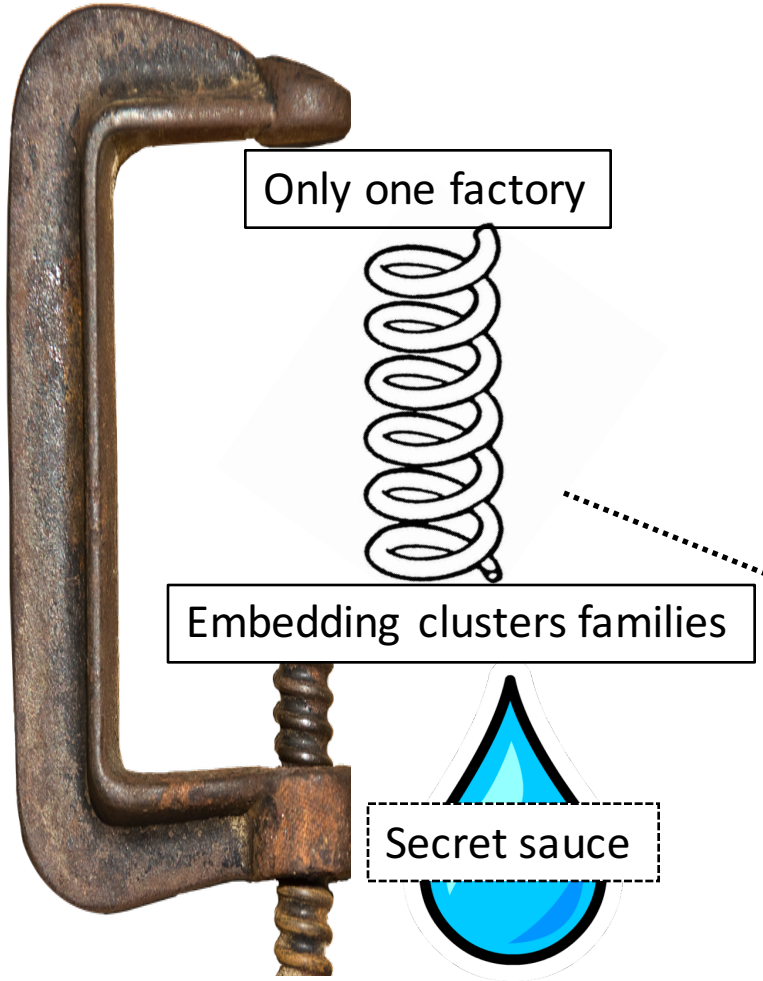


"The Matrix", 1999

Morphing the Embedding

Variational Autoencoder

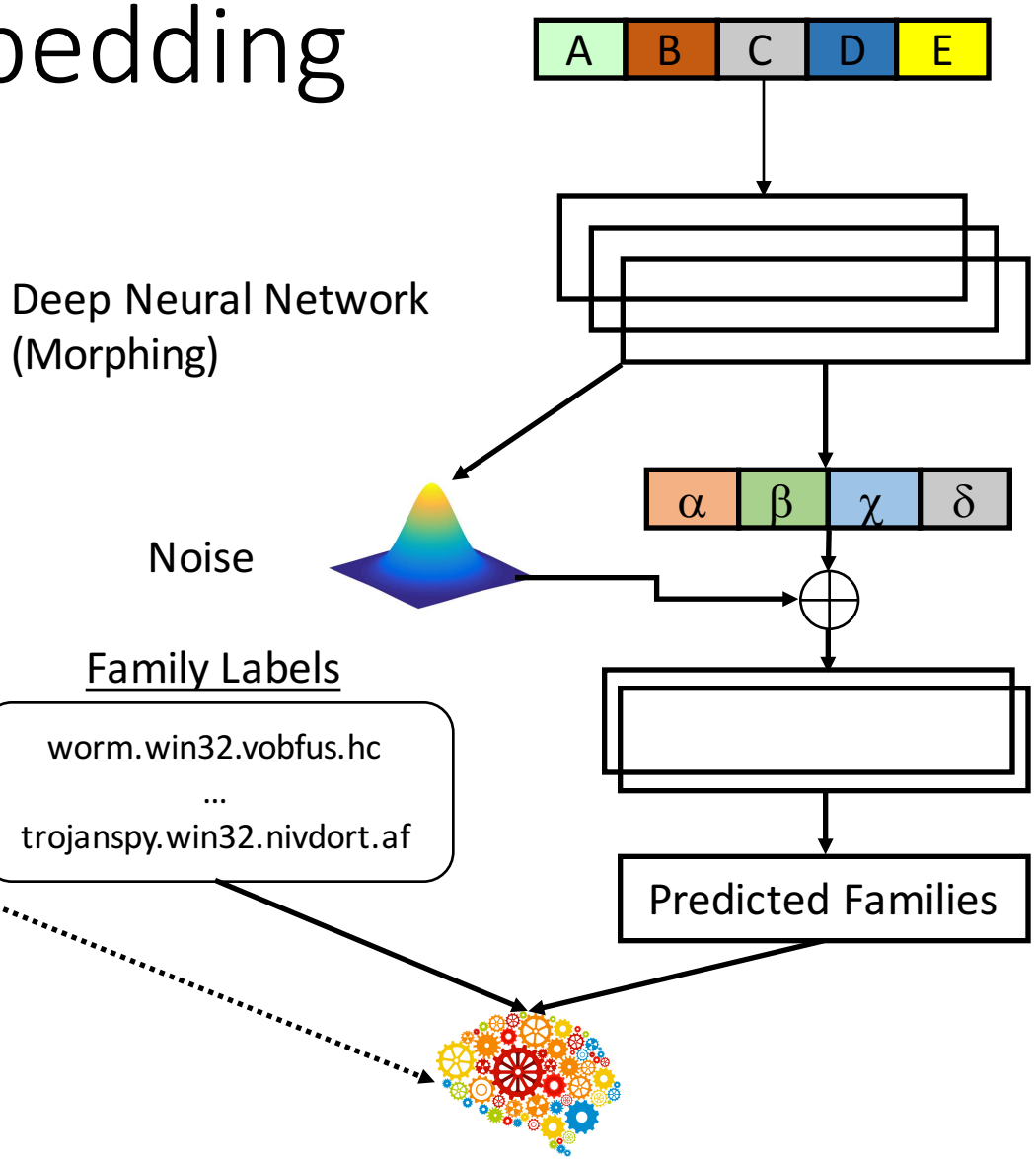
Kingma, D. P., & Welling, M. (2013). *arXiv preprint arXiv:1312.6114*.



Only one factory

Embedding clusters families

Secret sauce



Embedding Visualization

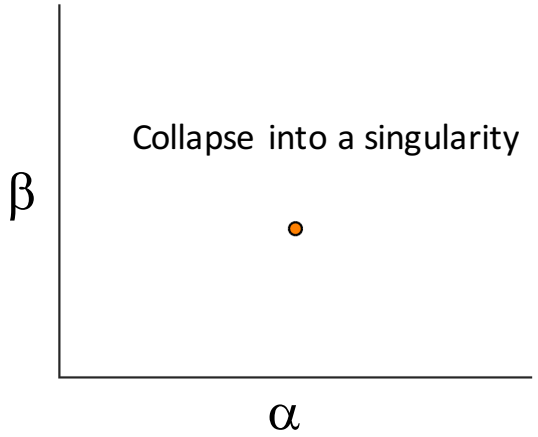
- Toy Example

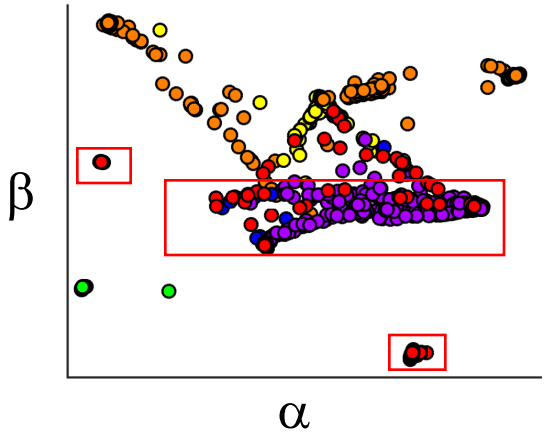
  - 8 family/variant prediction
  - 2D embedding

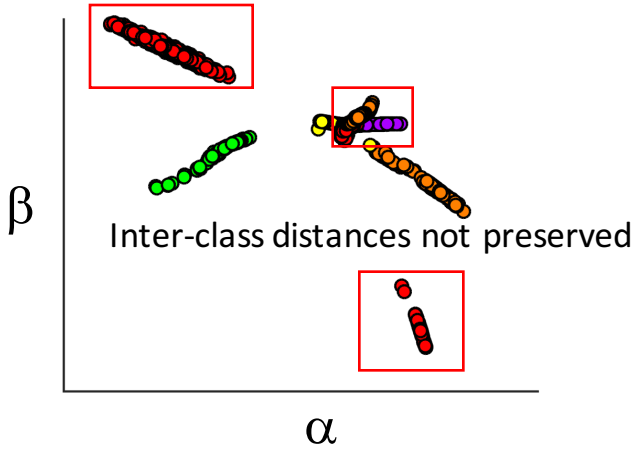
virus.win32.nabucur.d
virus.win32.ramnit.i
virus.win32.sality.at
virus.win32.shodi.i
virus.win32.virut.ae
virus.win32.virut.br
virus.win32.virut.k
worm.win32.allaple.a

Only one factory

Embedding clusters families

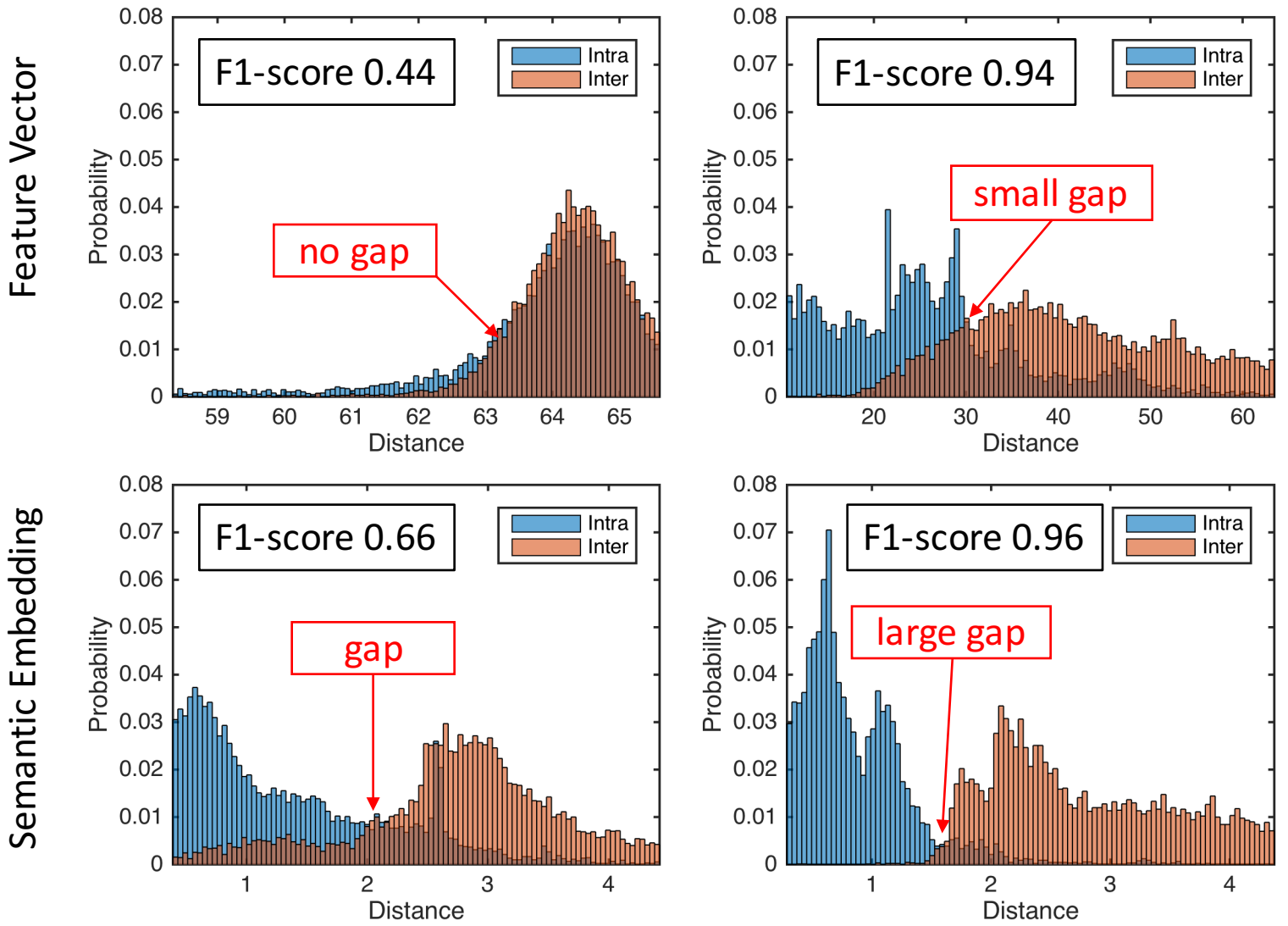






Results

- 800K samples
  - 1500 family/variants (99% coverage)
- Time-split Validation
  - Train on old data
  - Test on 30 days later
- Measure F1-score of 3-nearest neighbor classifier



Conclusion

- Developing feature extraction is expensive and requires time consuming tuning to adapt to a specific domain
- Traditional approaches to malware similarity are unsupervised and so are brittle
- Using supervised-learning approaches we can improve existing features by embedding them into a more optimized space
- **Automatic (re)tuning will improve detection rates and reduce cost**

More Information

- Email: kberlin@invincea.com
- Twitter: @kberlin