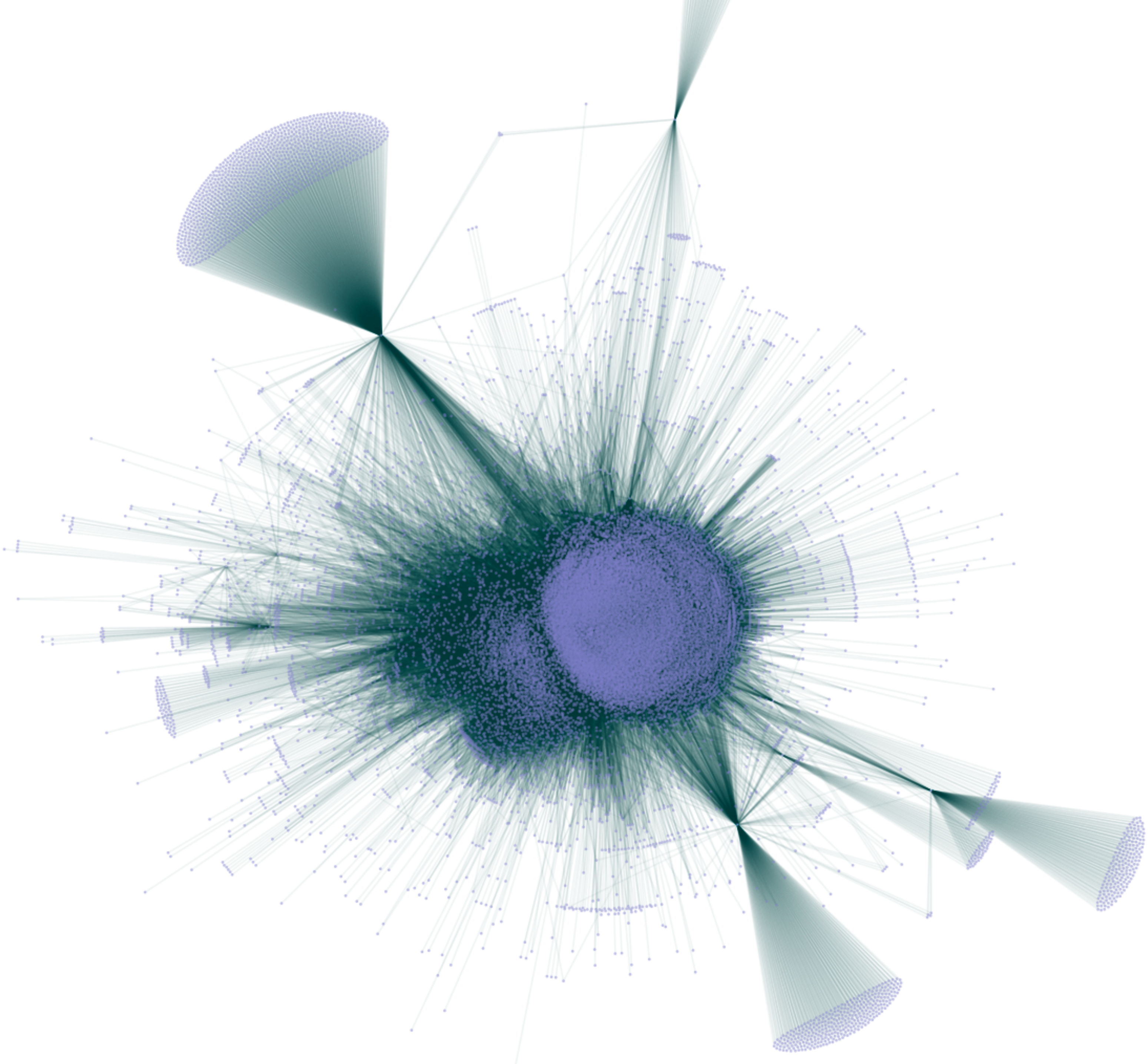


De-mystifying Data Science for Cyber Security

Joshua Neil

Principal Data Scientist Lead

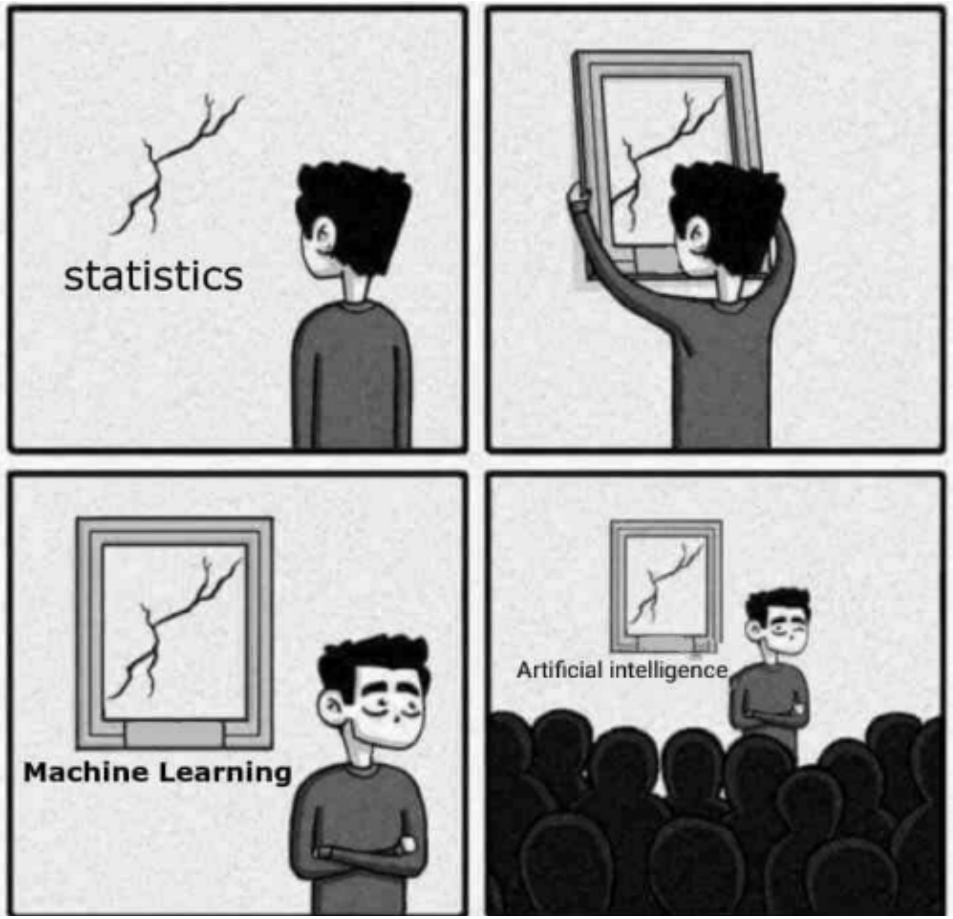
Microsoft Defender
Advanced Threat Protection

Outline

We do AI on graphs	AI is really just code and data
The Kill Chain is a Graph	I’ll explain graphs, and post breach attacks from the data science viewpoint
Learning dynamics of graphs	Modeling the nodes and edges
Finding the attack in the graph	Greedy approach (there are others, but I won’t kill you with statistics)

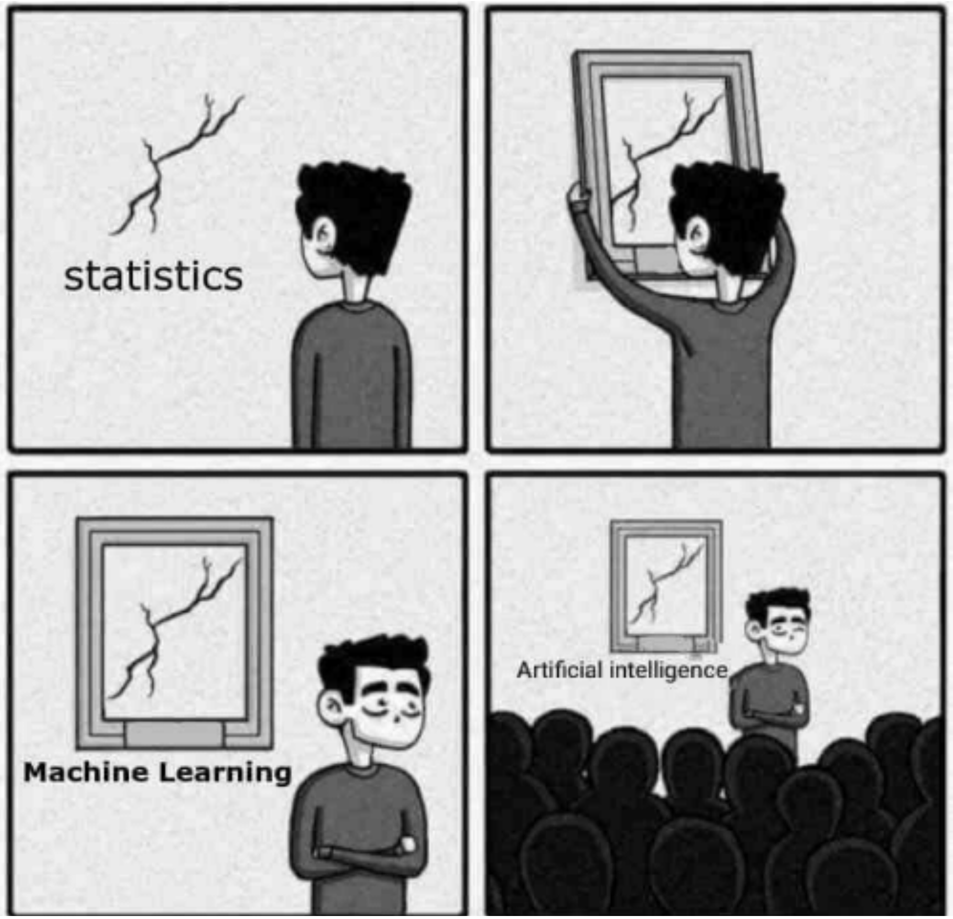
We do AI*!

*statistics



We do AI*! On Graphs**!!

*statistics **nodes and edges



+

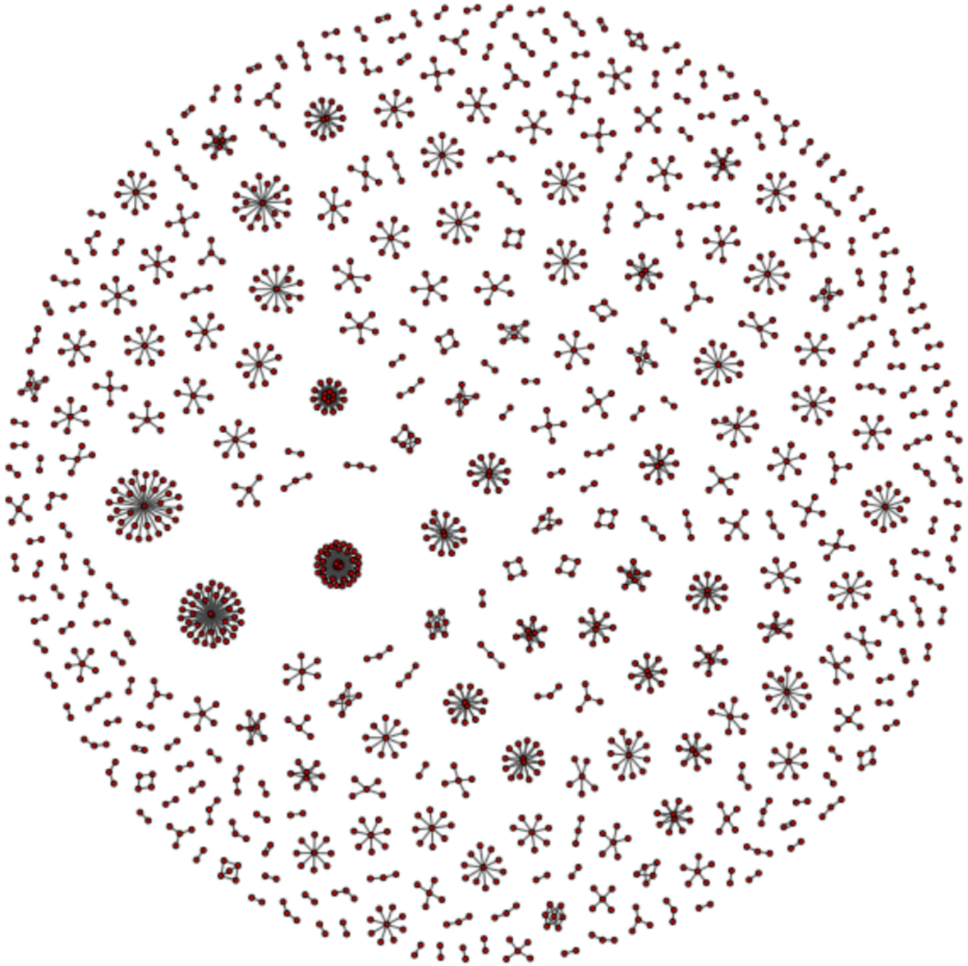
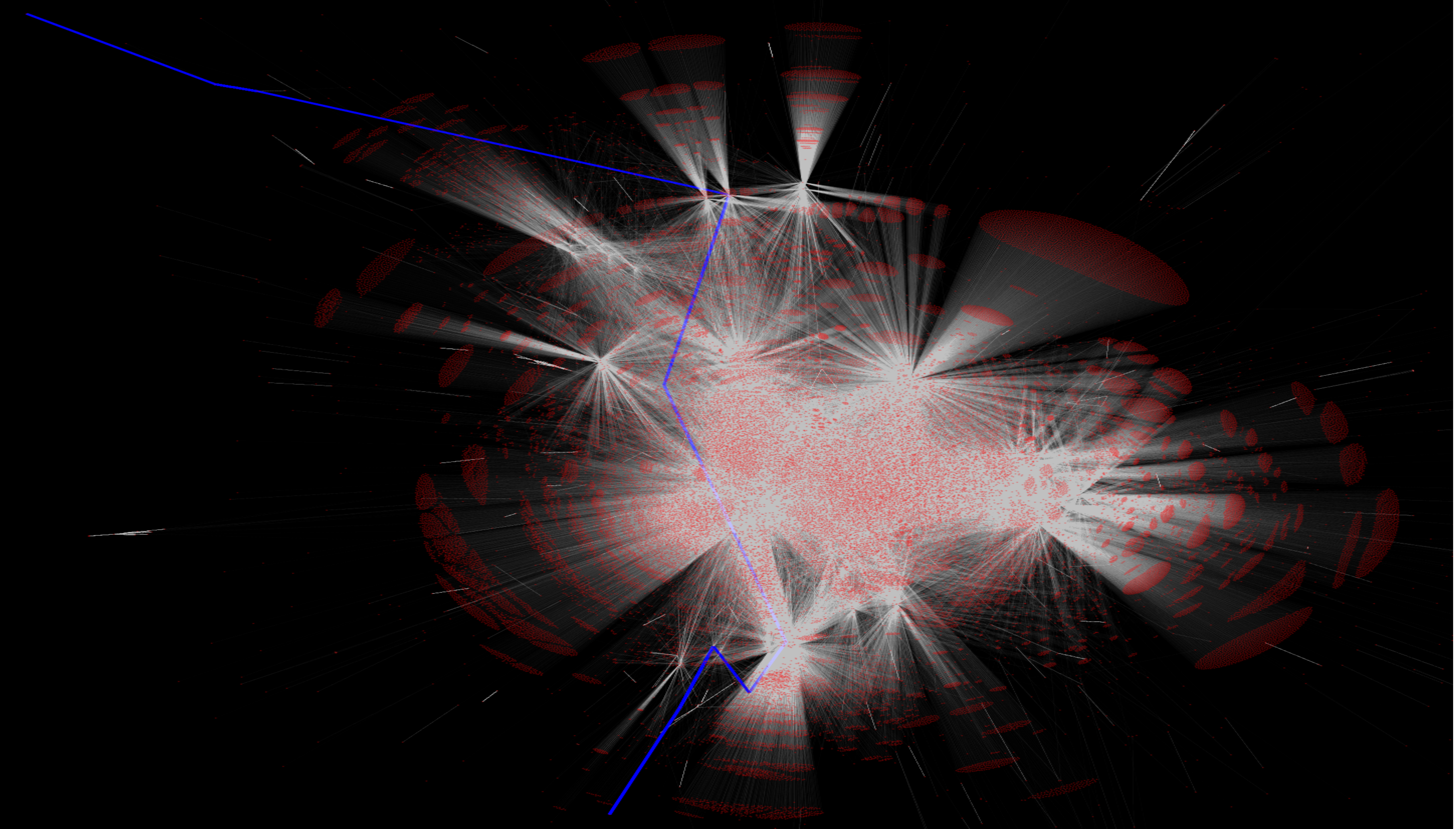
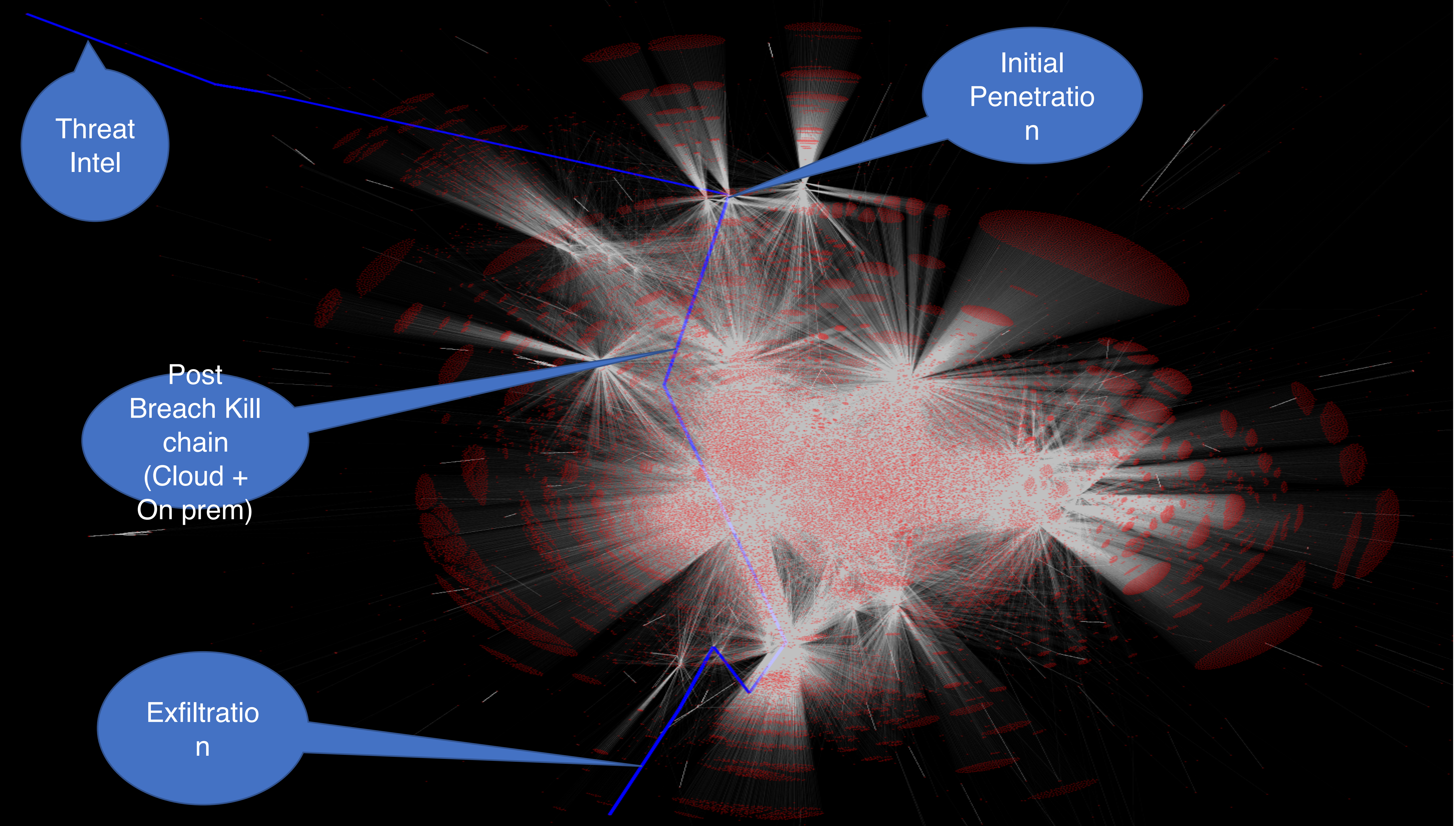


Image Credit: Evan Argyle, MSET





Threat
Intel

Initial
Penetratio
n

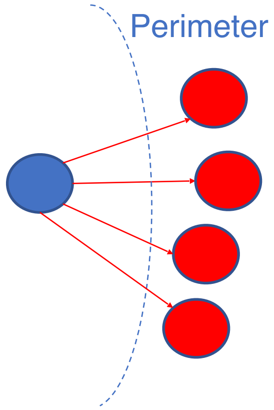
Post
Breach Kill
chain
(Cloud +
On prem)

Exfiltratio
n

Attack behaviors and Enterprise Graphs

Initial penetration

- Deviations in Email behavior due to phishing barrage



* Red indicates deviations the attacker has introduced in the normal behavior of the endpoints and communications

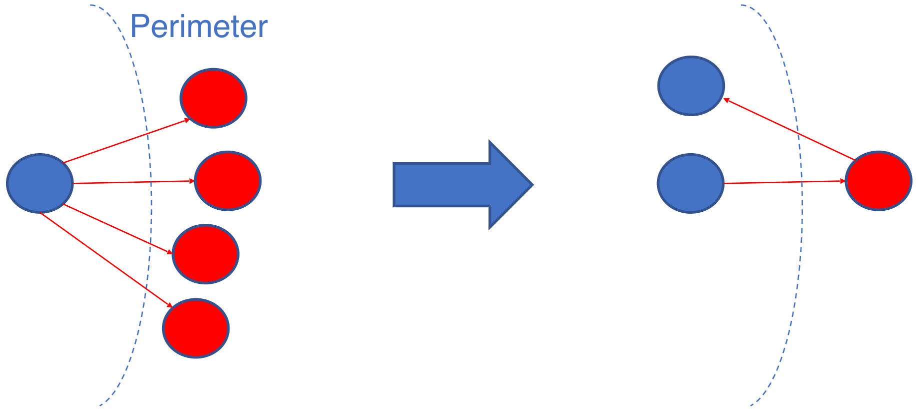
Attack behaviors and Enterprise Graphs

Initial penetration

- Deviations in Email behavior due to phishing barrage

Persistence and callback

- processes, command lines, registry, scheduled task, etc
- Deviations on network, low reputation, beaconing, etc
- Credential deviations



* Red indicates deviations the attacker has introduced in the normal behavior of the endpoints and communications

Attack behaviors and Enterprise Graphs

Initial penetration

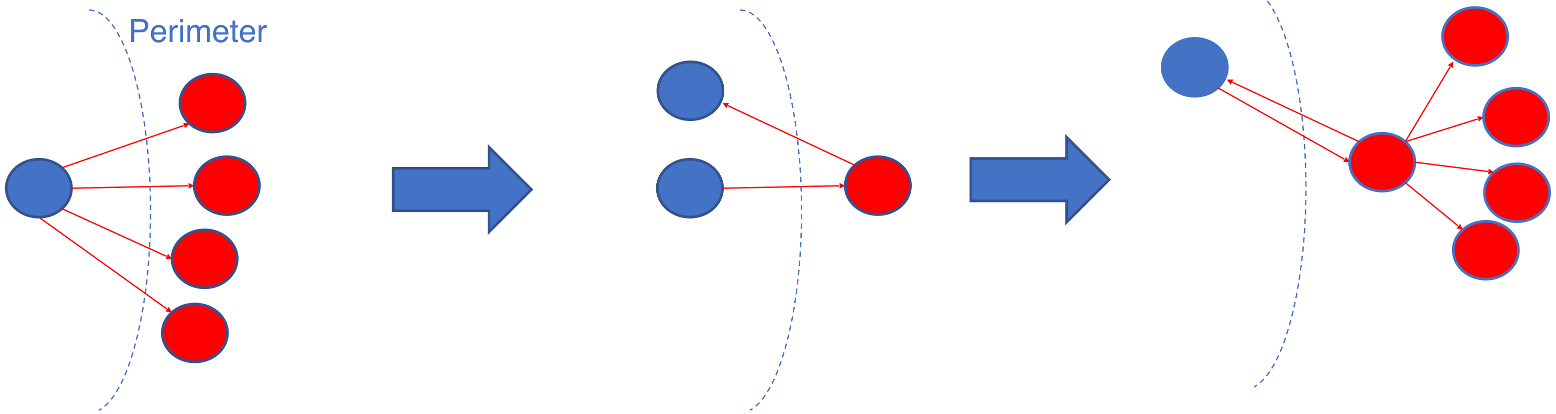
- Deviations in Email behavior due to phishing barrage

Persistence and callback

- processes, command lines, registry, scheduled task, etc
- Deviations on network, low reputation, beaconing, etc
- Credential deviations

C2/Recon

- Deviations in perimeter network comms and internal comms graph
- Internal Port deviations for horizontal and vertical port scanning between machines
- Deviations in HTML/DNS requests for covert channel C2



* Red indicates deviations the attacker has introduced in the normal behavior of the endpoints and communications

Attack behaviors and Enterprise Graphs

Initial penetration

- Deviations in Email behavior due to phishing barrage

Persistence and callback

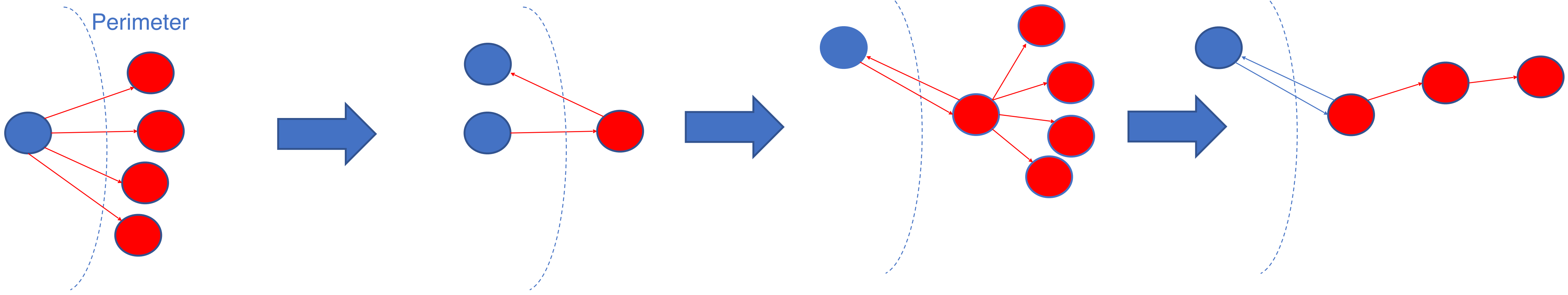
- processes, command lines, registry, scheduled task, etc
- Deviations on network, low reputation, beaconing, etc
- Credential deviations

C2/Recon

- Deviations in perimeter network comms and internal comms graph
- Internal Port deviations for horizontal and vertical port scanning between machines
- Deviations in HTML/DNS requests for covert channel C2

Lateral Movement

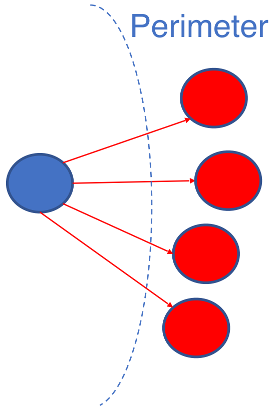
- Network and OS deviations
- Credential anomalies
- Insider/pattern of life anomalies



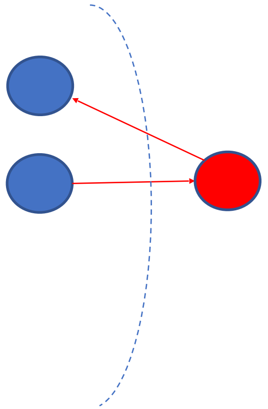
* Red indicates deviations the attacker has introduced in the normal behavior of the endpoints and communications

Attack behaviors and Enterprise Graphs

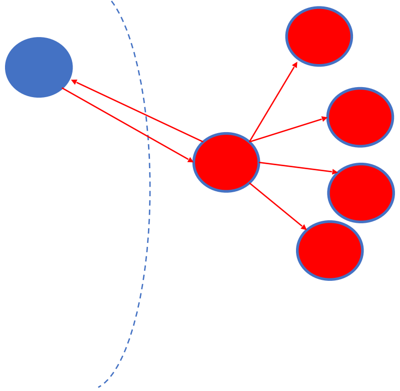
Initial penetration
- Deviations in Email behavior due to phishing barrage



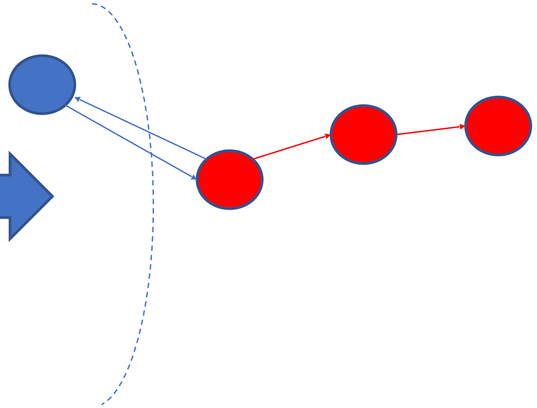
Persistence and callback
- processes, command lines, registry, scheduled task, etc
- Deviations on network, low reputation, beaconing, etc
- Credential deviations



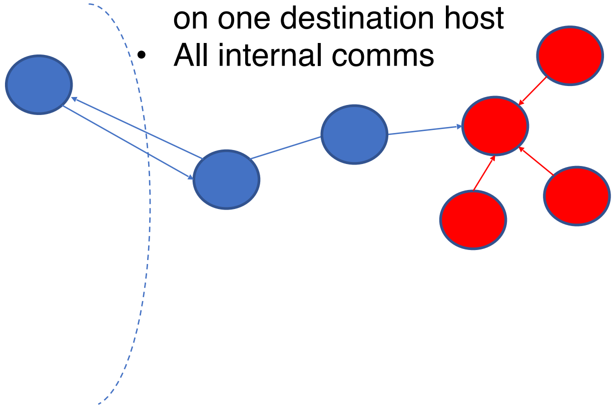
C2/Recon
- Deviations in perimeter network comms and internal comms graph
- Internal Port deviations for horizontal and vertical port scanning between machines
- Deviations in HTML/DNS requests for covert channel C2



Lateral Movement
- Network and OS deviations
- Credential anomalies
- Insider/pattern of life anomalies

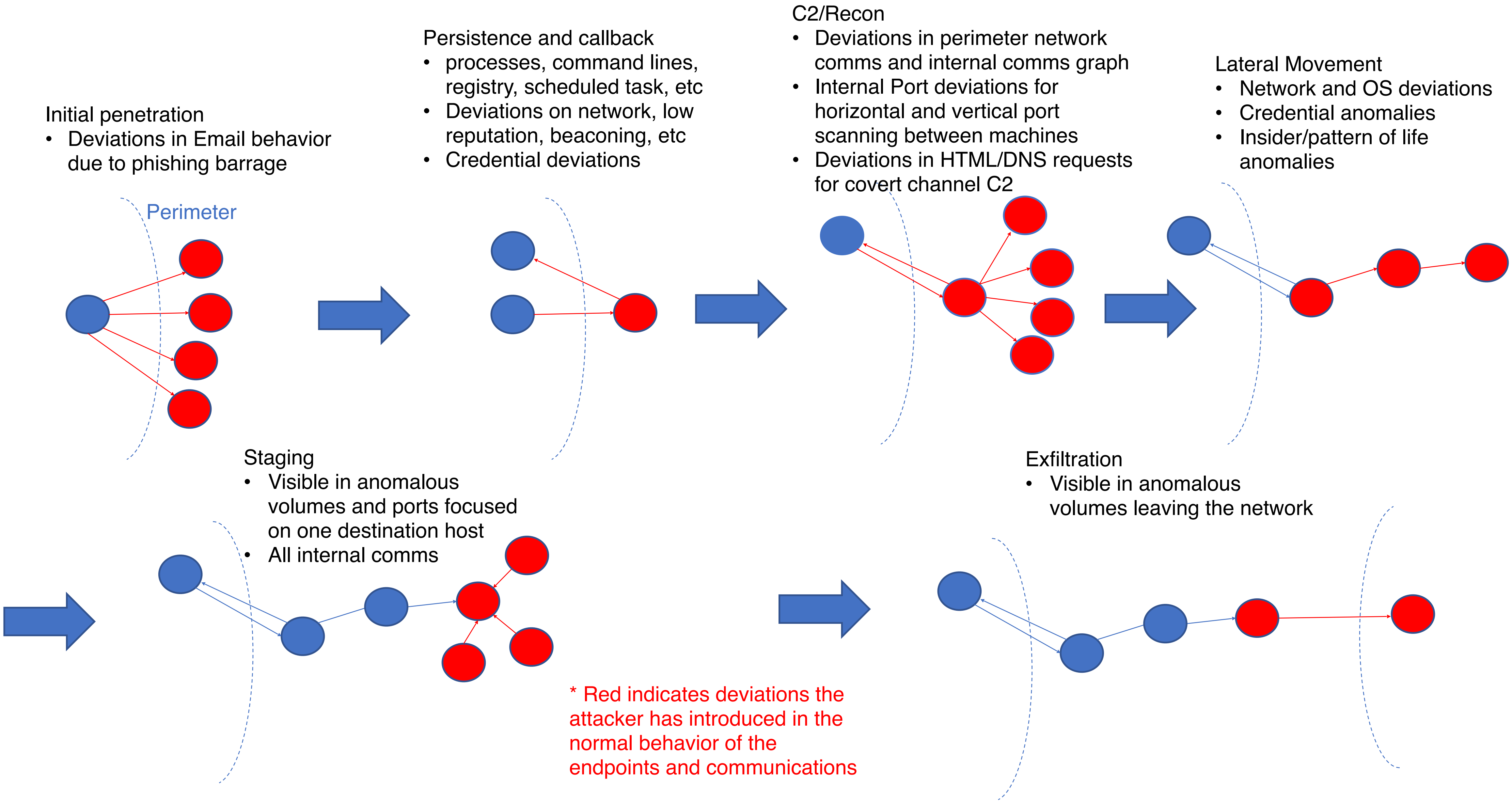


Staging
- Visible in anomalous volumes and ports focused on one destination host
- All internal comms



* Red indicates deviations the attacker has introduced in the normal behavior of the endpoints and communications

Attack behaviors and Enterprise Graphs



Initial penetration

- Deviations in Email behavior due to phishing barrage

Persistence and callback

- processes, command lines, registry, scheduled task, etc
- Deviations on network, low reputation, beaconing, etc
- Credential deviations

C2/Recon

- Deviations in perimeter network comms and internal comms graph
- Internal Port deviations for horizontal and vertical port scanning between machines
- Deviations in HTML/DNS requests for covert channel C2

Lateral Movement

- Network and OS deviations
- Credential anomalies
- Insider/pattern of life anomalies

Staging

- Visible in anomalous volumes and ports focused on one destination host
- All internal comms

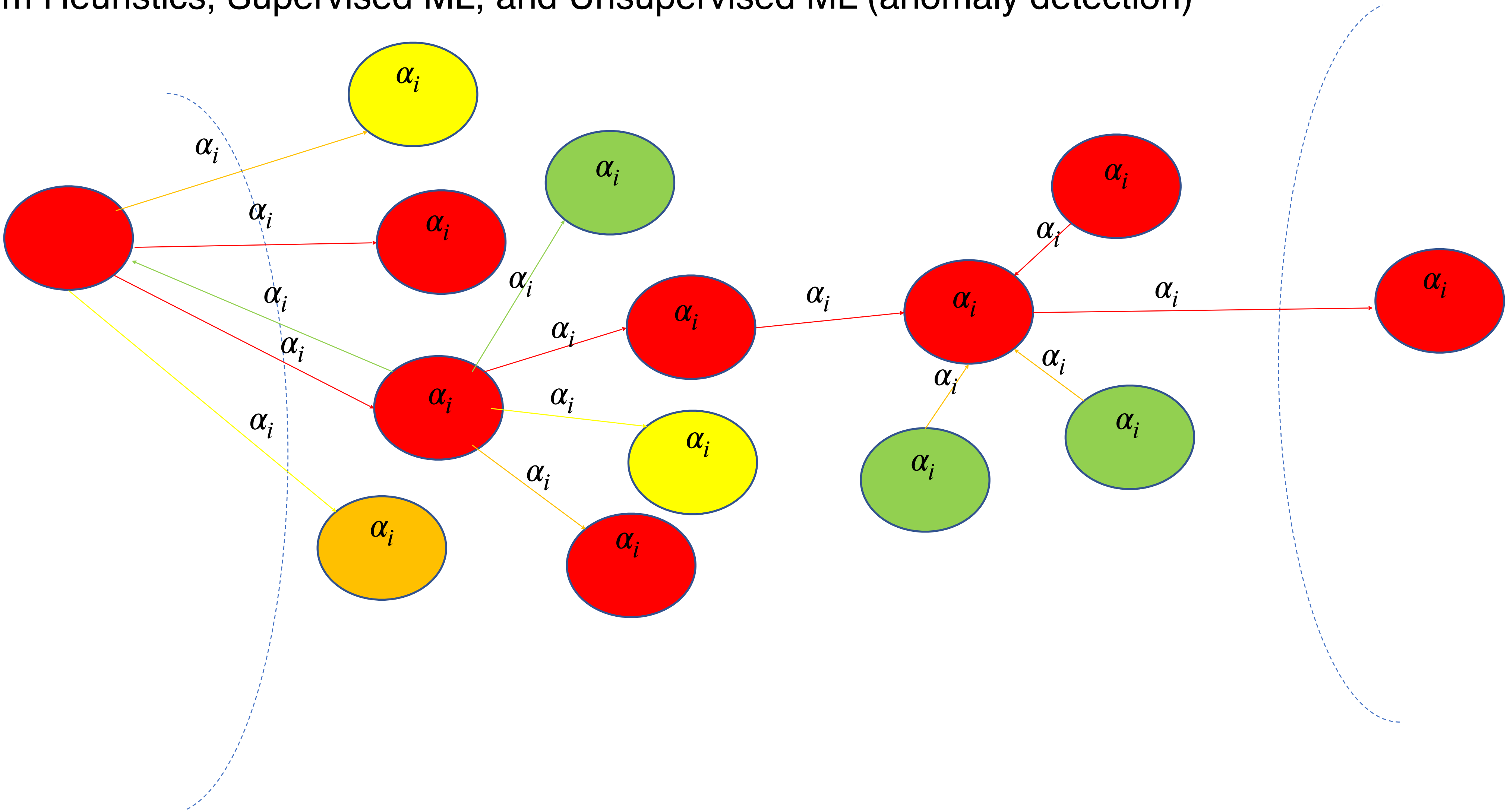
Exfiltration

- Visible in anomalous volumes leaving the network

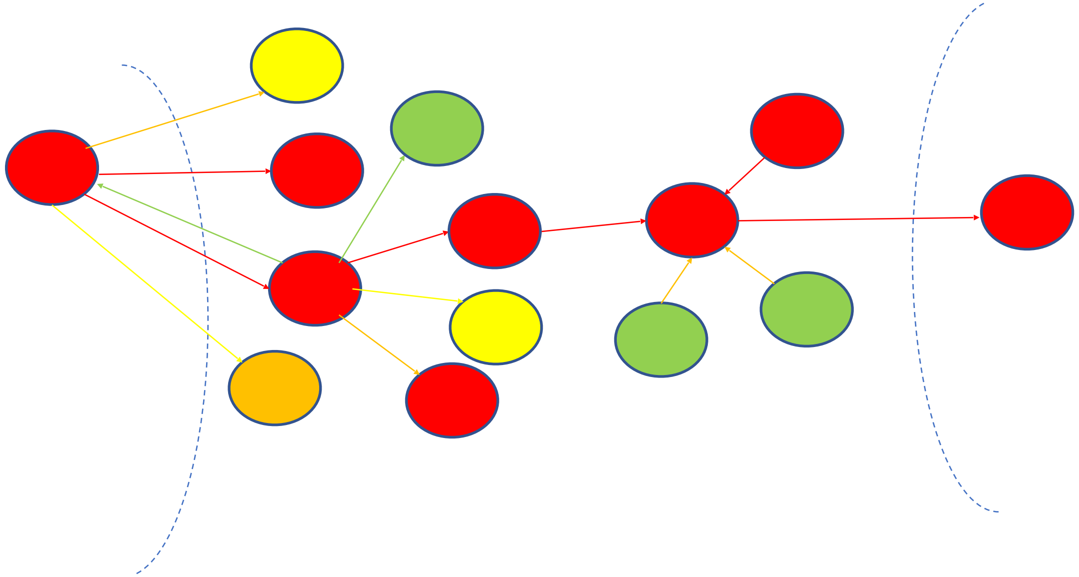
* Red indicates deviations the attacker has introduced in the normal behavior of the endpoints and communications

Score nodes and edges

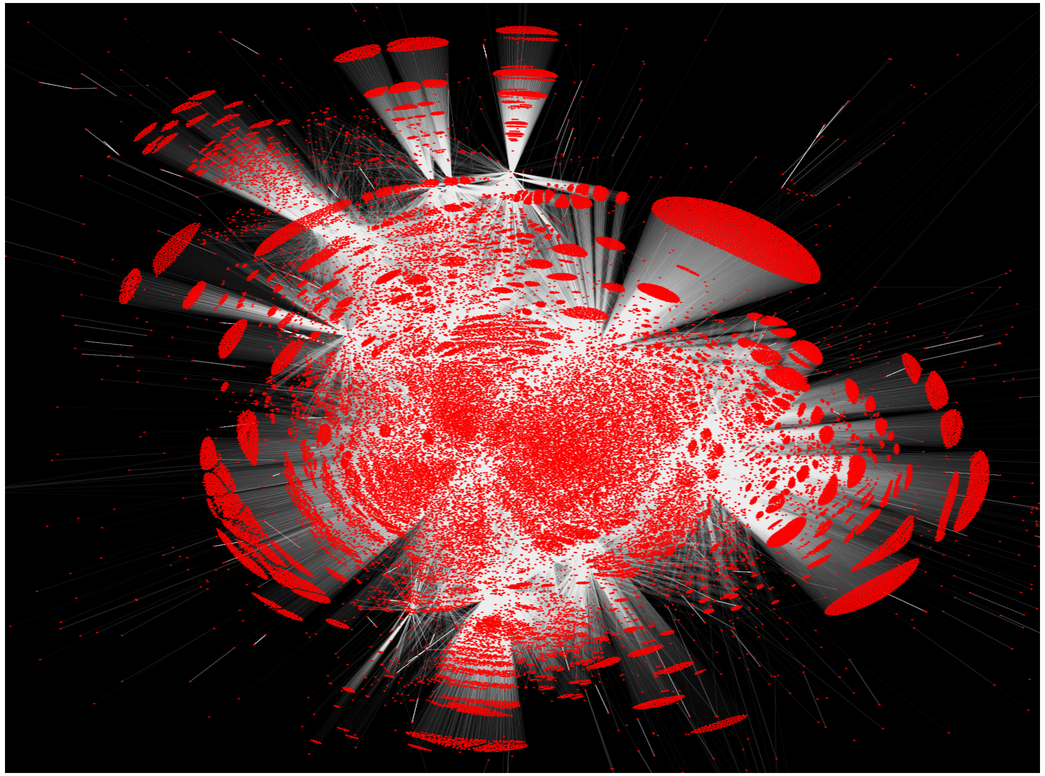
From Heuristics, Supervised ML, and Unsupervised ML (anomaly detection)



How do we find this:



In here:

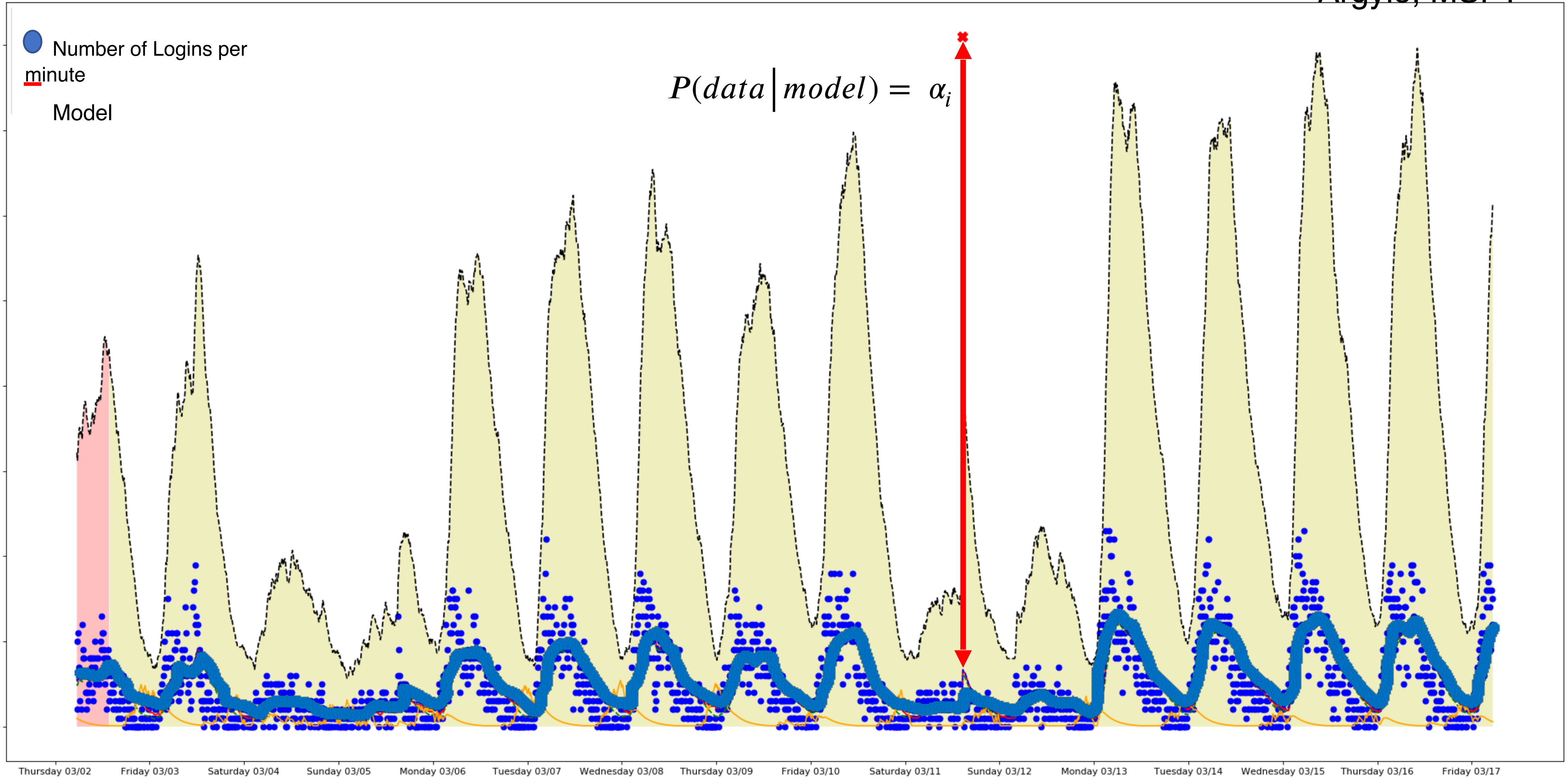


If we have scores (α_i) on each node and edge?

Interlude: Scoring the Edge

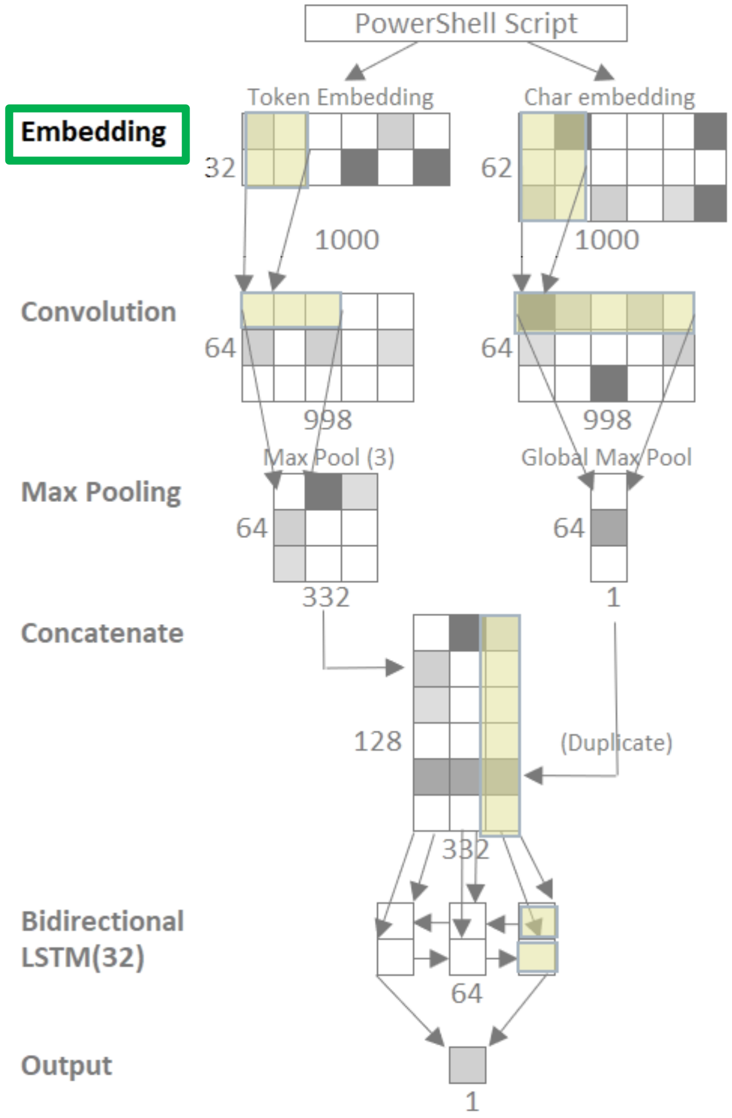
Number of incoming login attempts

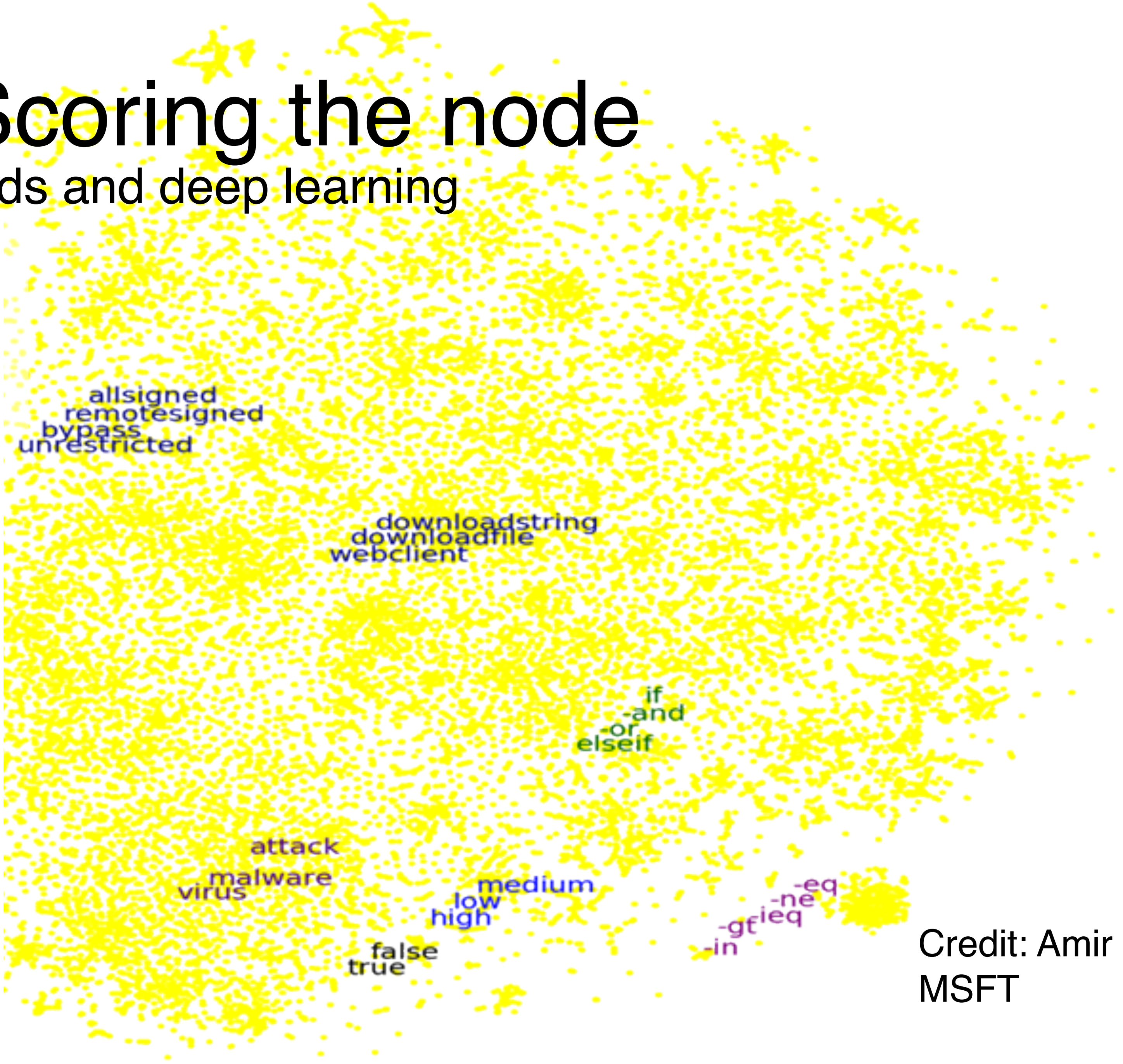
Credit: Evan Argyle, MSFT



Interlude: Scoring the node

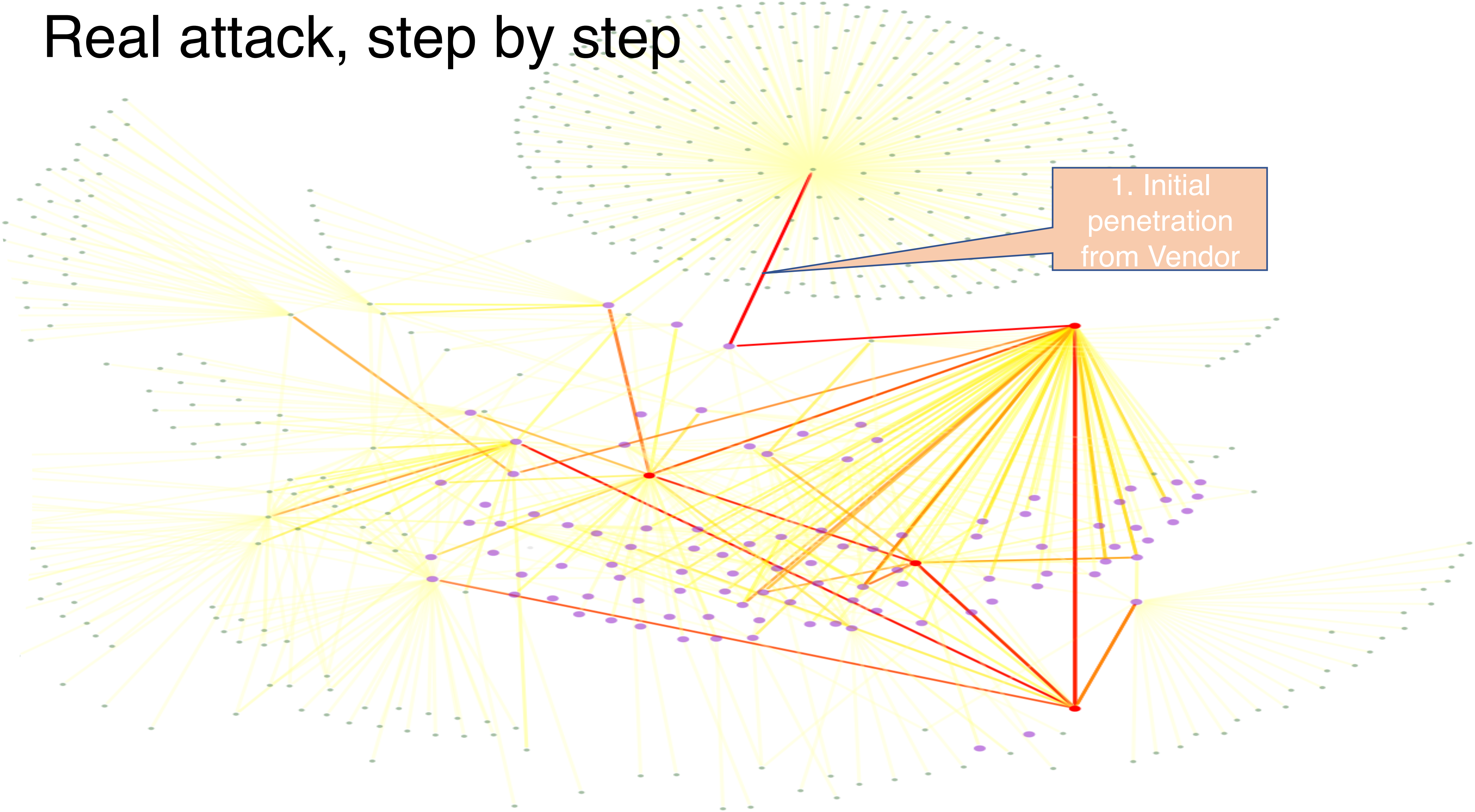
Powershell commands and deep learning





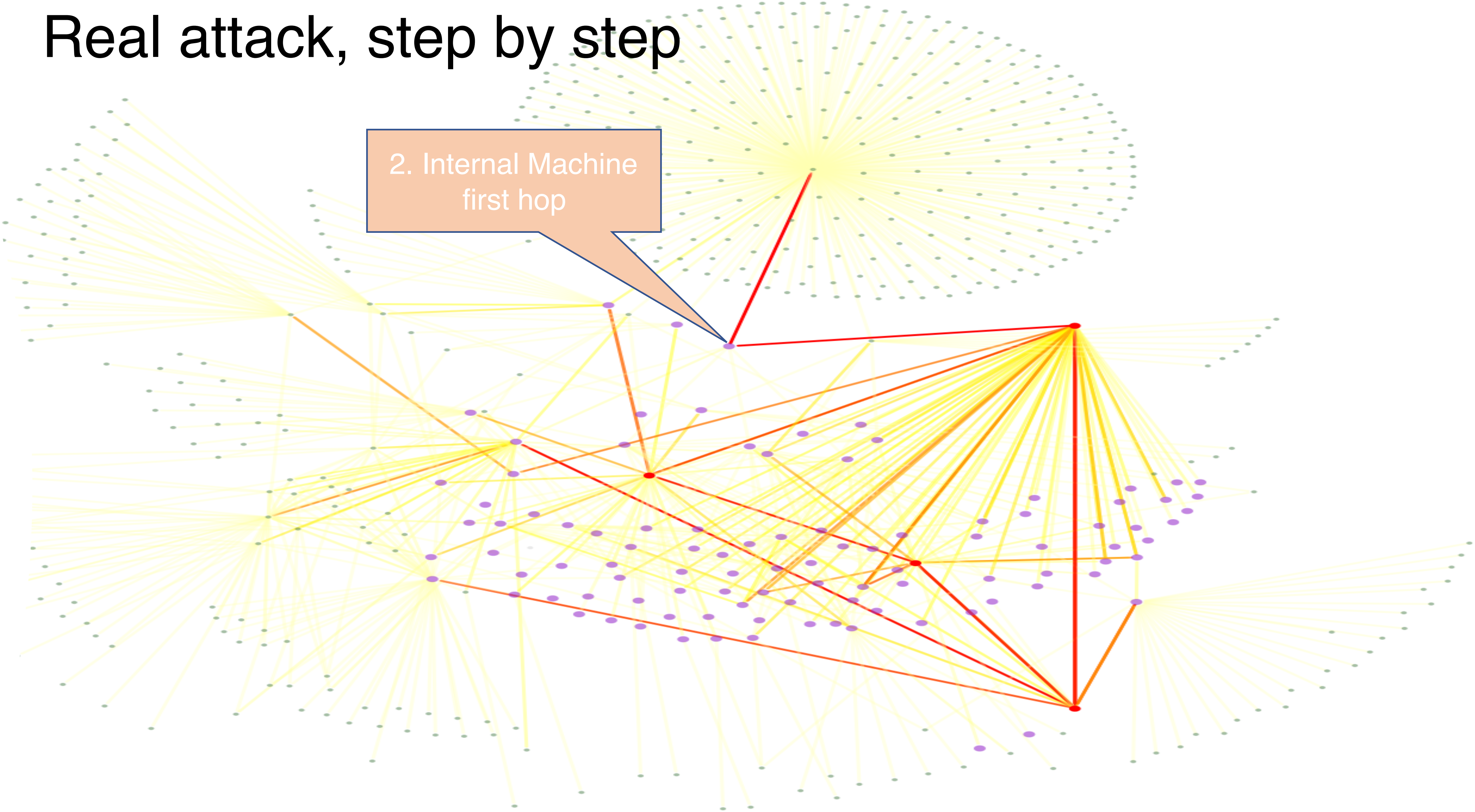
Credit: Amir Rubin, MSFT

Real attack, step by step



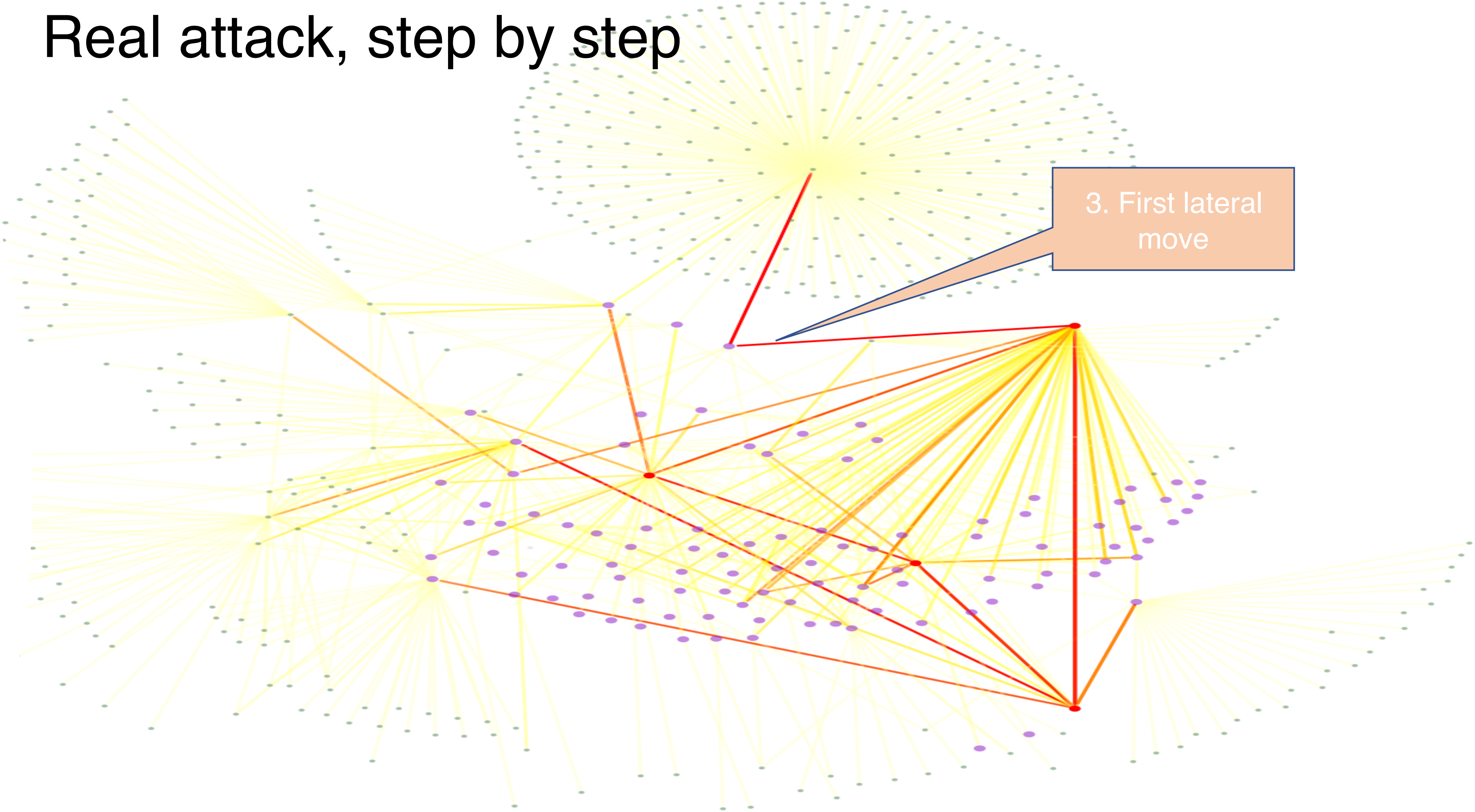
1. Initial penetration from Vendor

Real attack, step by step



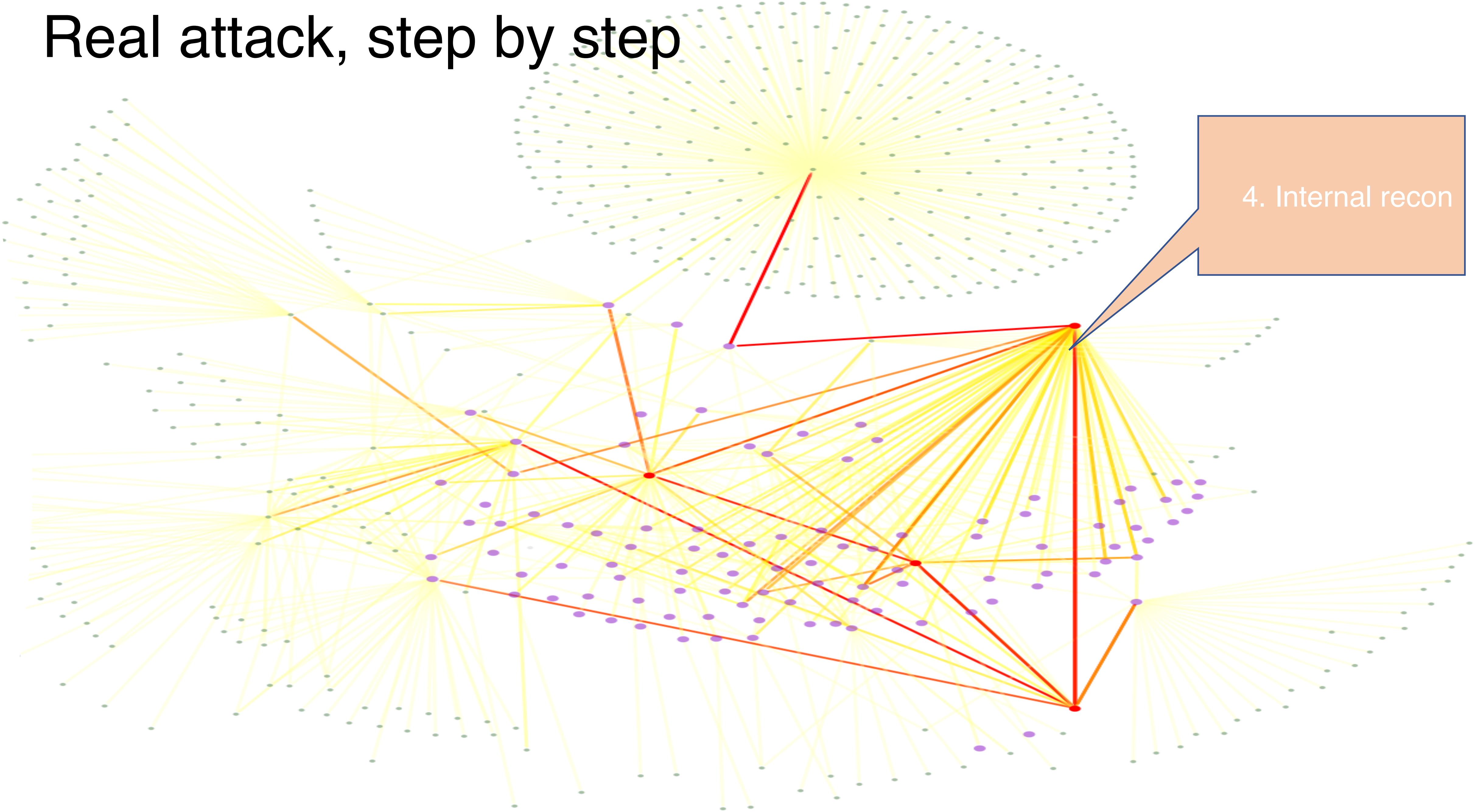
2. Internal Machine first hop

Real attack, step by step



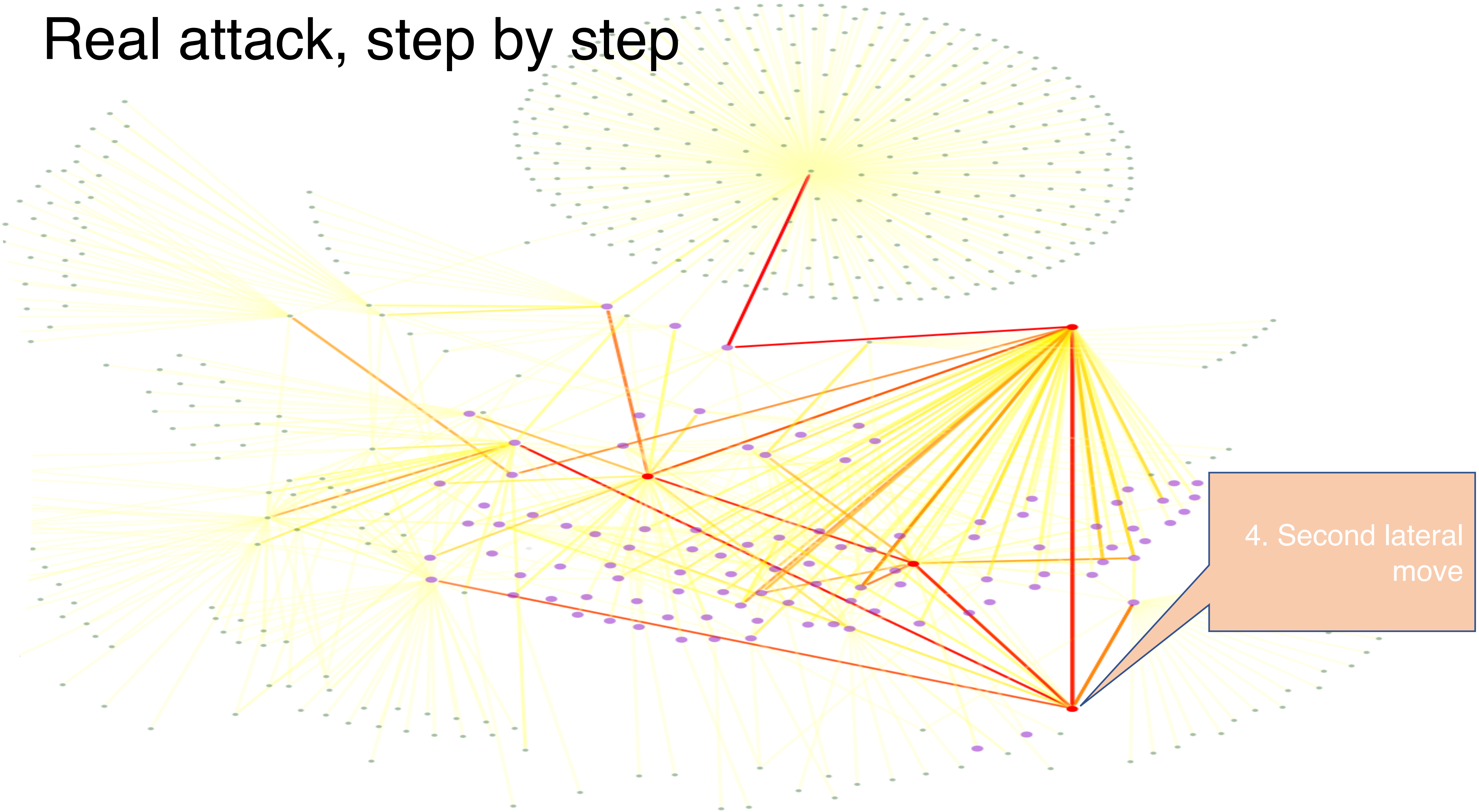
3. First lateral move

Real attack, step by step



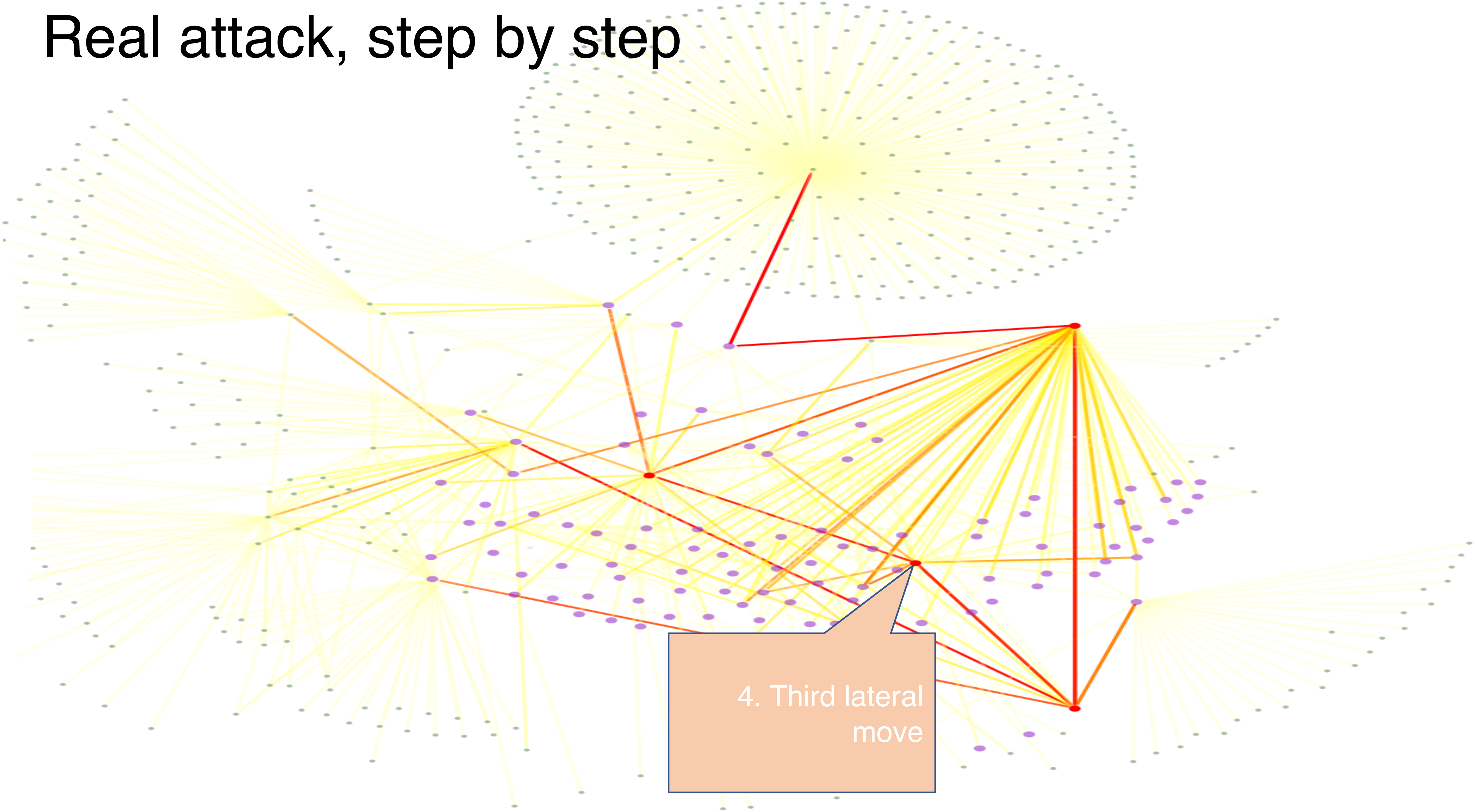
4. Internal recon

Real attack, step by step



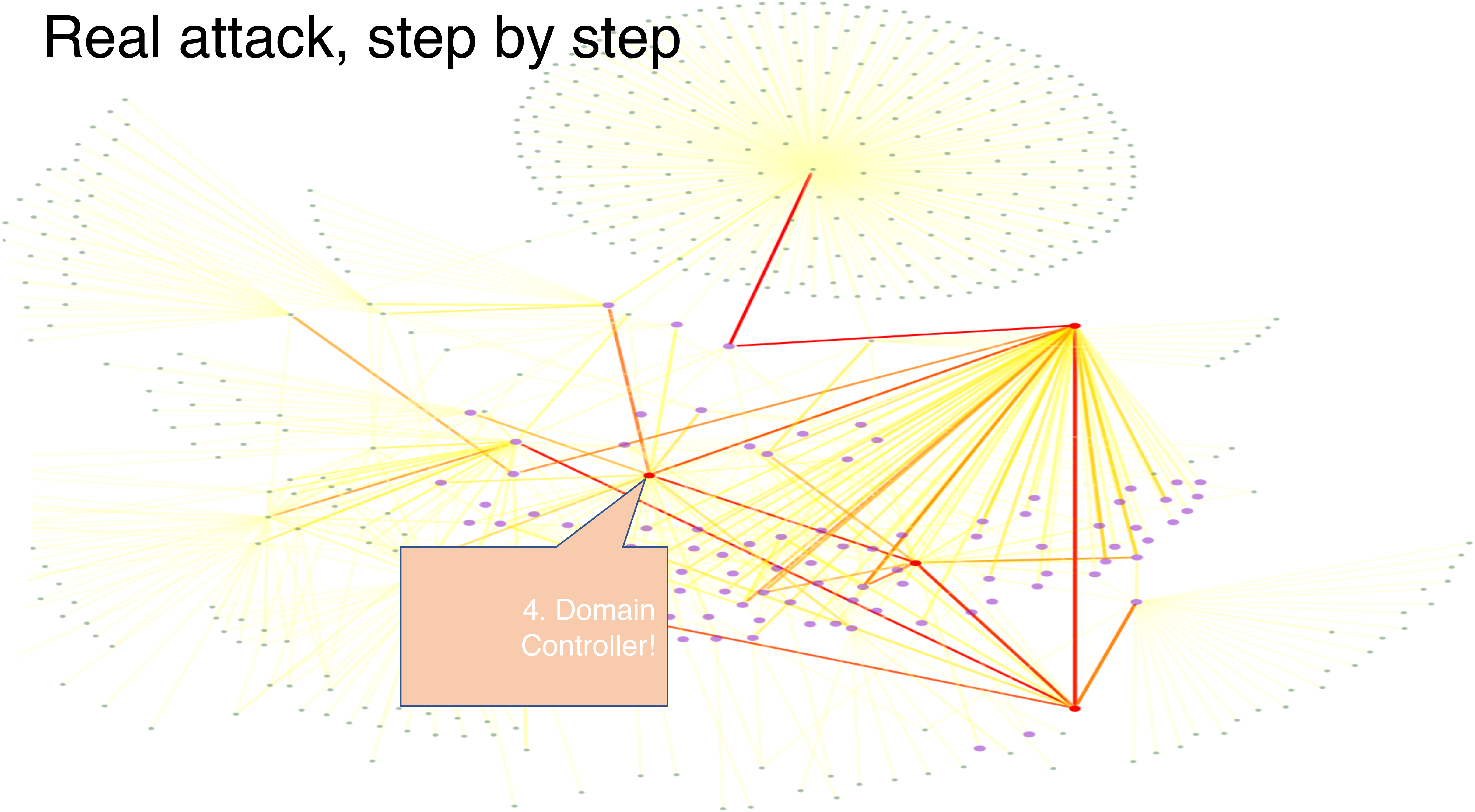
4. Second lateral move

Real attack, step by step

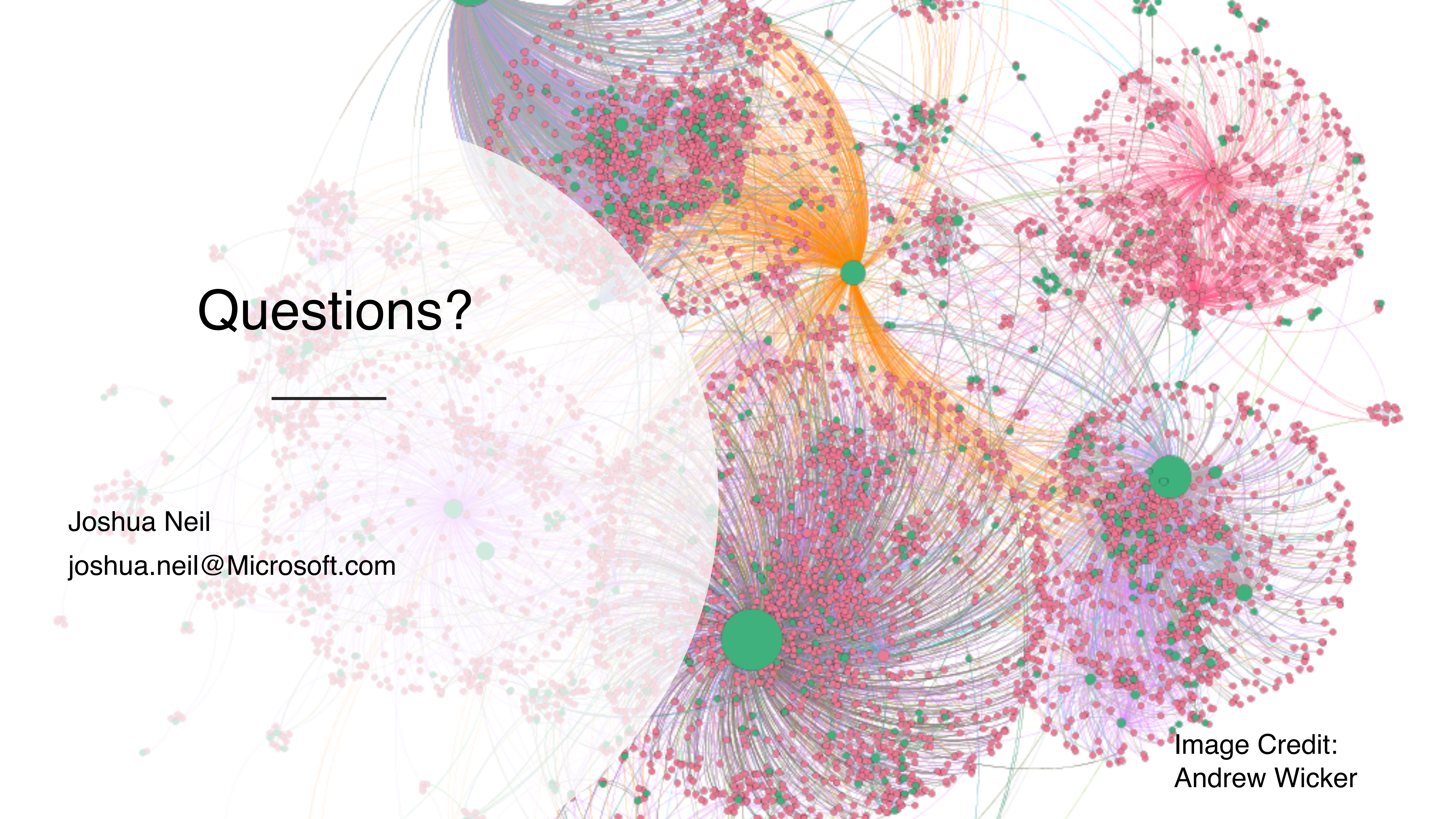


4. Third lateral move

Real attack, step by step



4. Domain Controller!



Questions?

Joshua Neil

joshua.neil@Microsoft.com

Image Credit:
Andrew Wicker