Integrating Defenders & Attackers into Cyber Security Risk Models

Varun Agarwal
Diane Henshel, Alexander Alexeev, Mariana Cains

SRA Annual Meeting, Arlington VA, 10-14 December 2017

INDIANA UNIVERSITY
SCHOOL OF INFORMATICS, COMPUTING, AND ENGINEERING

SCHOOL OF PUBLIC AND ENVIRONMENTAL AFFAIRS
INDIANA UNIVERSITY

Outline

- Motivation & Research Goals
- Technical Approach & Conceptual Framework
- Statistical Application
- Preliminary Results
- Conclusions & Future Work

Motivation & Research Goals

- Current cyber security risk modeling frameworks include only hardware and software
- Importance of human factors is under-represented in major risk assessment frameworks

So, we propose a model

The goal is to:

- Incorporate human factors (attackers & defenders) in cyber risk models
- Model risk ***dynamically***
- Identify minimum number of ***necessary and sufficient variables*** that capture the dynamic system risk
- Finally, evaluate cause of high-risk situations

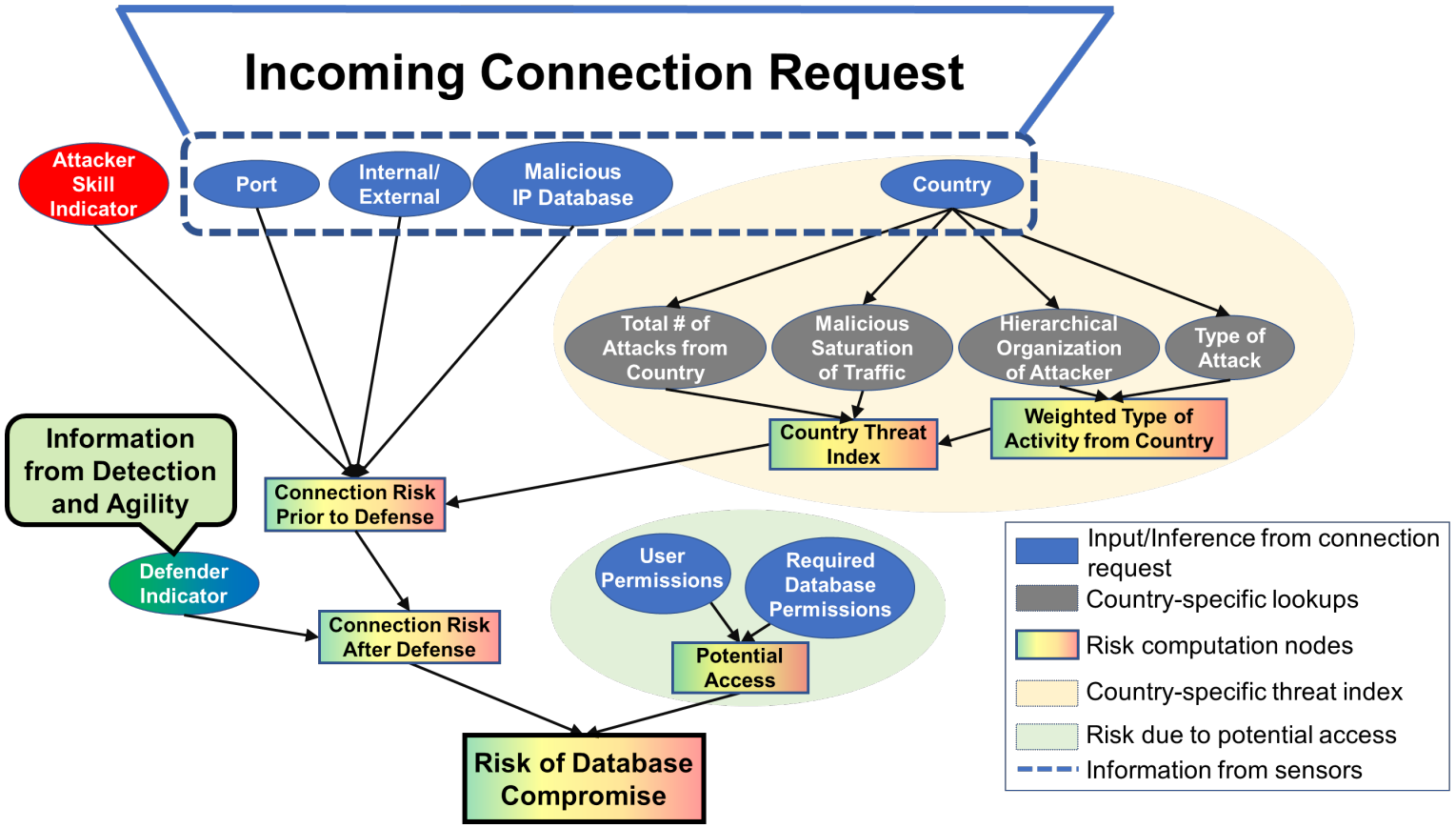
Technical Approach

To achieve our goal

- We use hybrid Bayesian network to build our risk model
  - Reason - Bayesian networks allow for causal inference
  - Graphical models are more suitable for assessing risk in complex systems
- Presented model is built around modeling risk to a database server

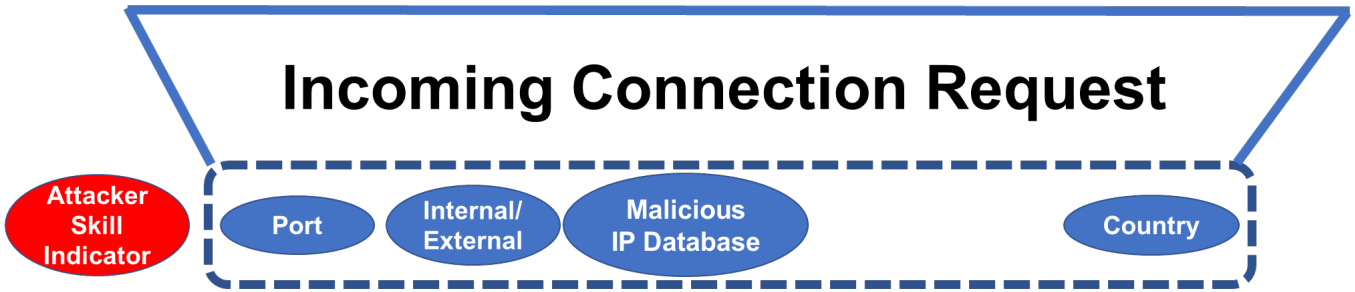
Conceptual Risk Framework

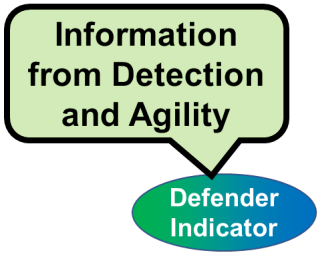
Framework outputs risk associated with an ***Incoming Connection Request***

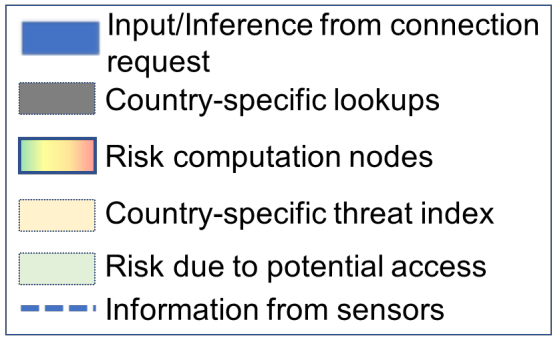


Step 1 - Incoming connection request detected

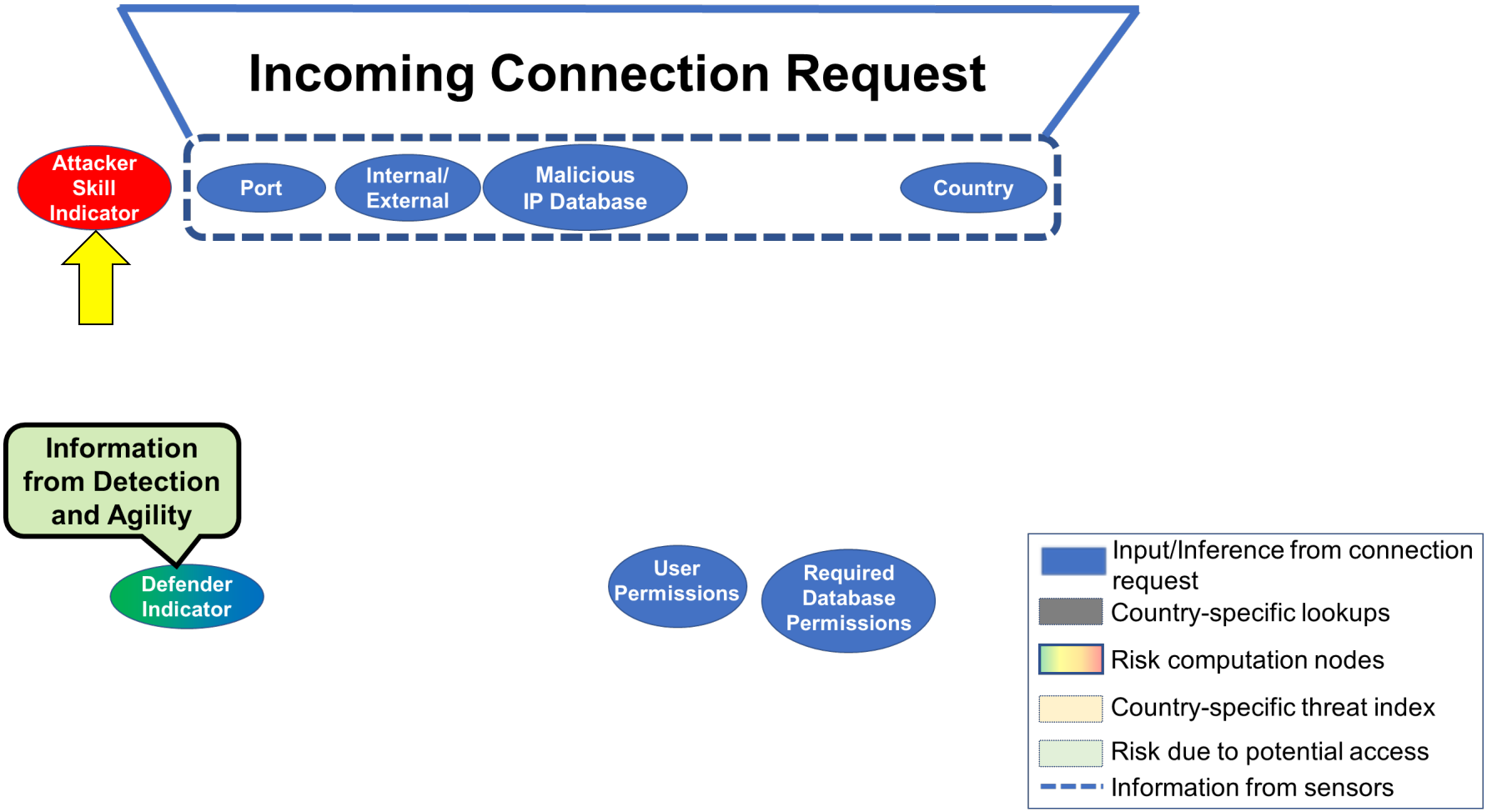
Step 2 - Set evidence for inferences from connection request



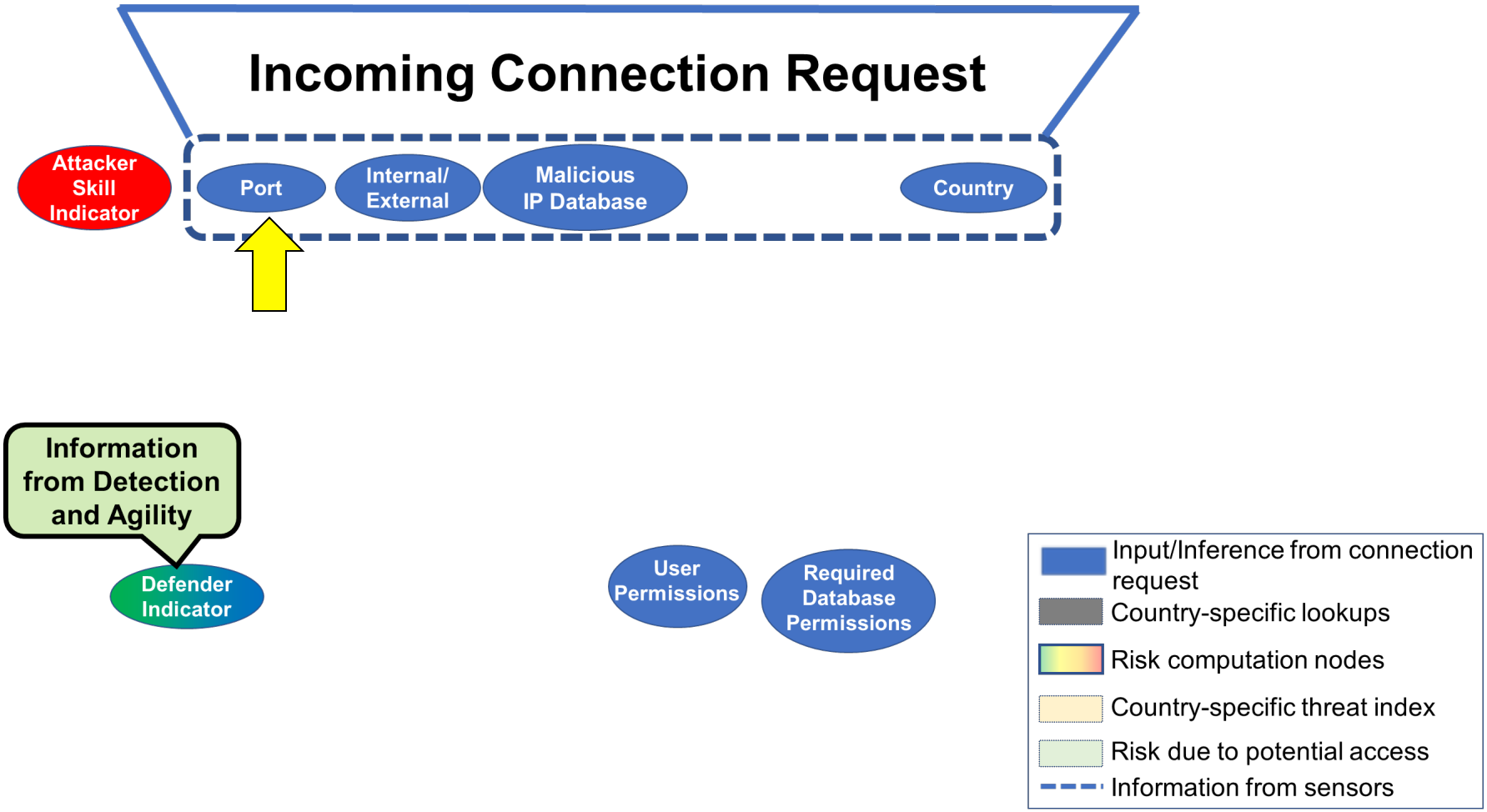




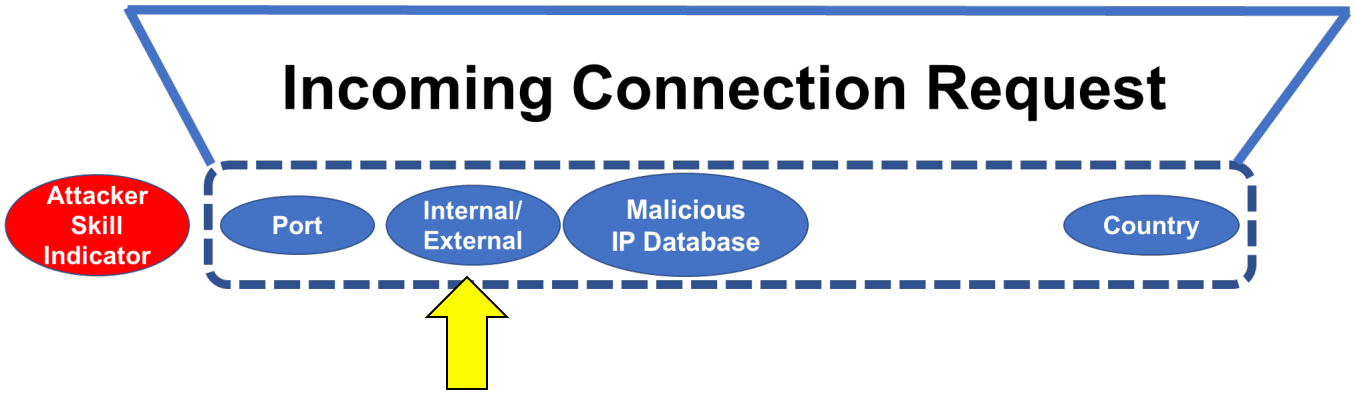
Attacker Skill – Distribution informed by prior experience

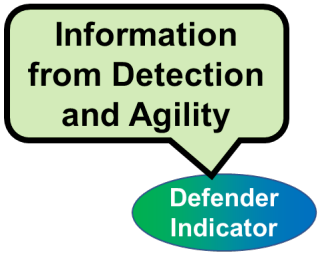


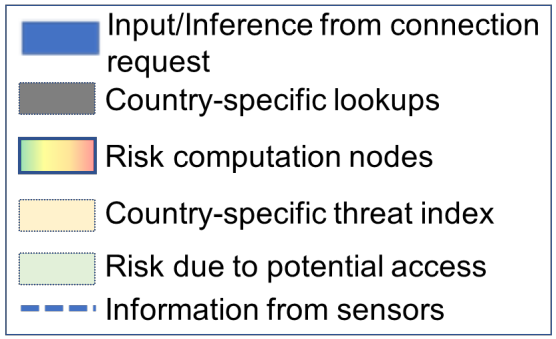
Port – port through which connection request comes in (e.g. Port 80 for HTTP, Port 22 for SSH)



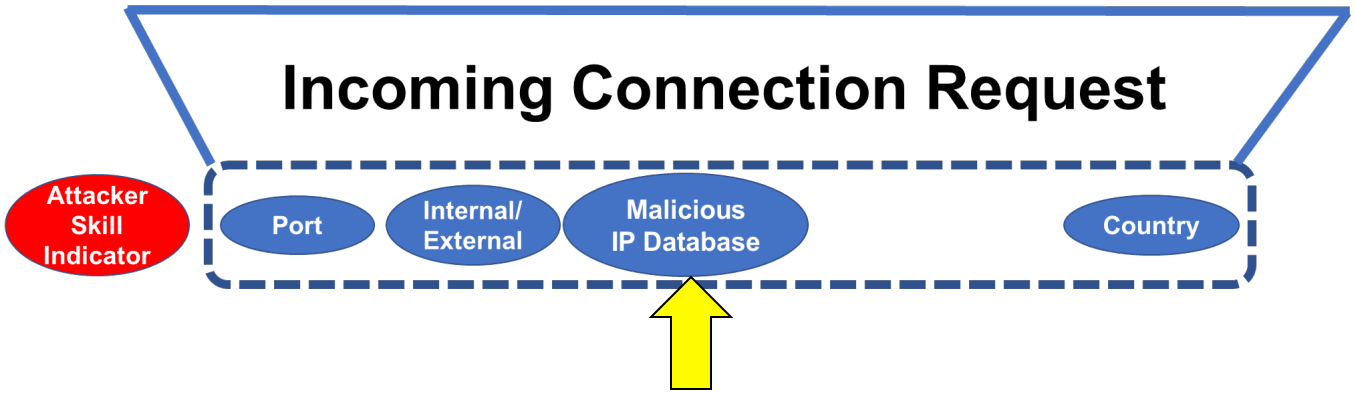
Internal/External – Is origin of connection Request internal to server's network?

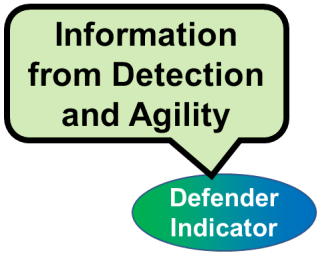


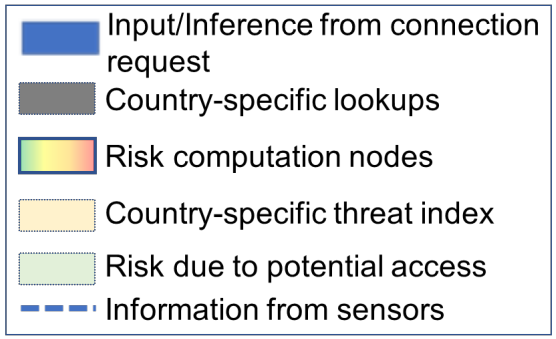




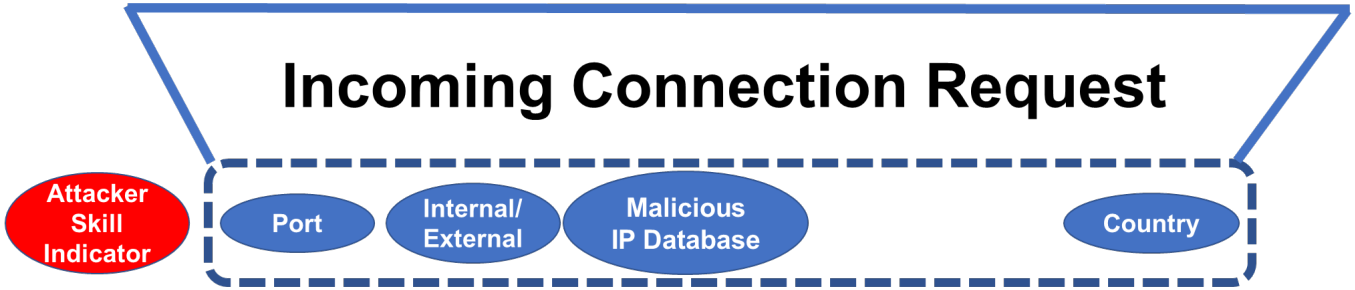
Malicious IP Database – Is IP listed in online malicious IP databases?

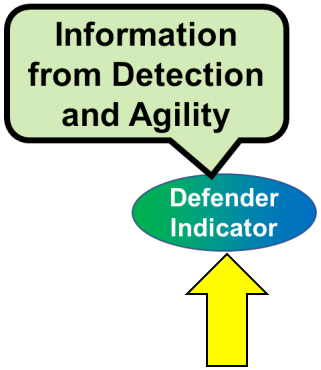


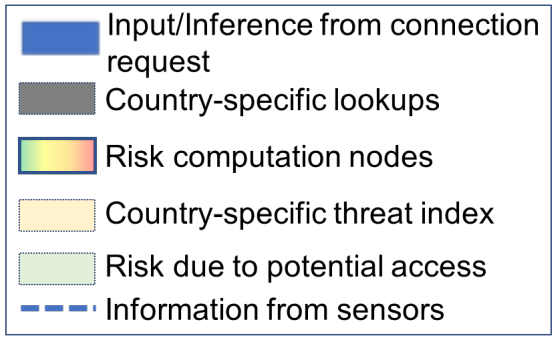




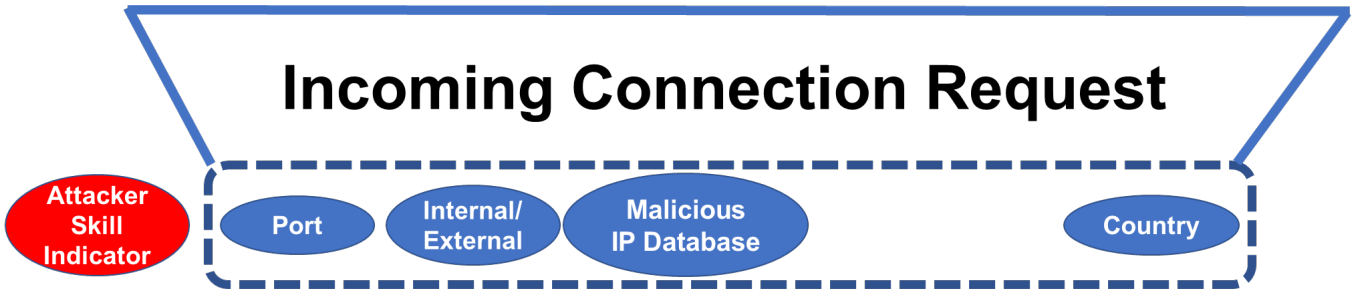
Defender Skill – Can be measured through internal assessments of cyber security experts/defenders (Low skill, Medium skill, High skill)

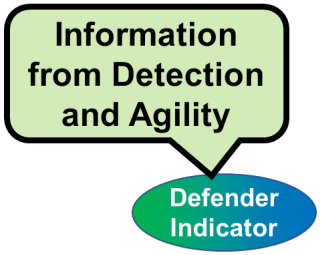


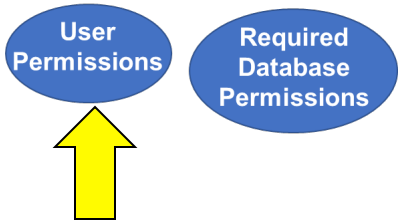


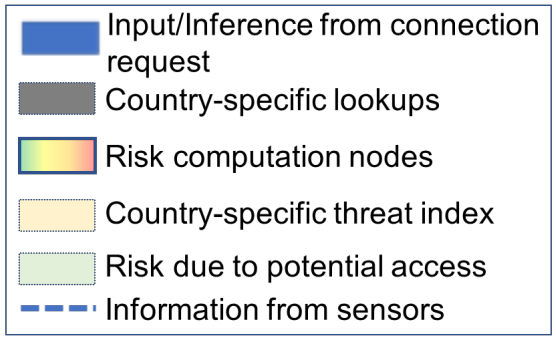


User Permission – Access level that the user possesses (Low, medium, high)

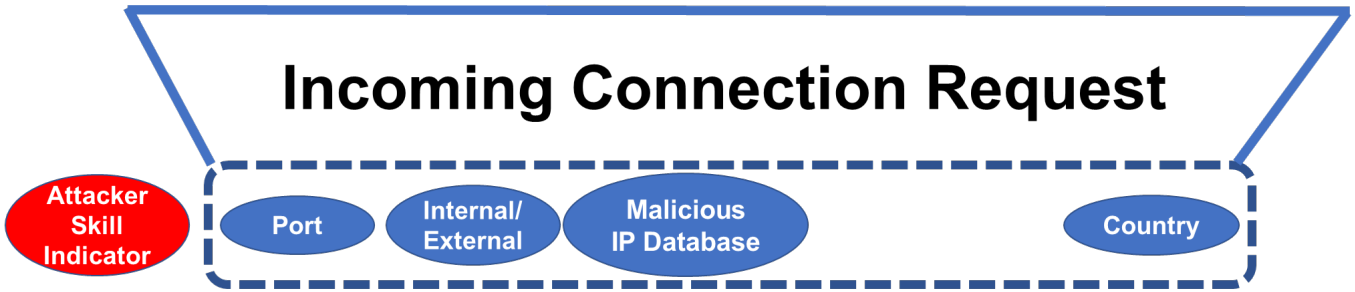


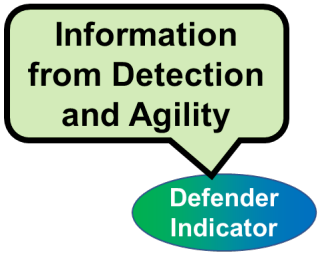


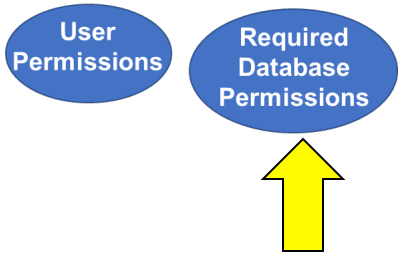


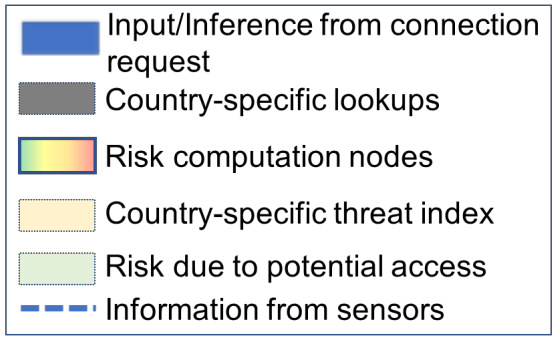


Required Permission – Access Level required to communicate with server

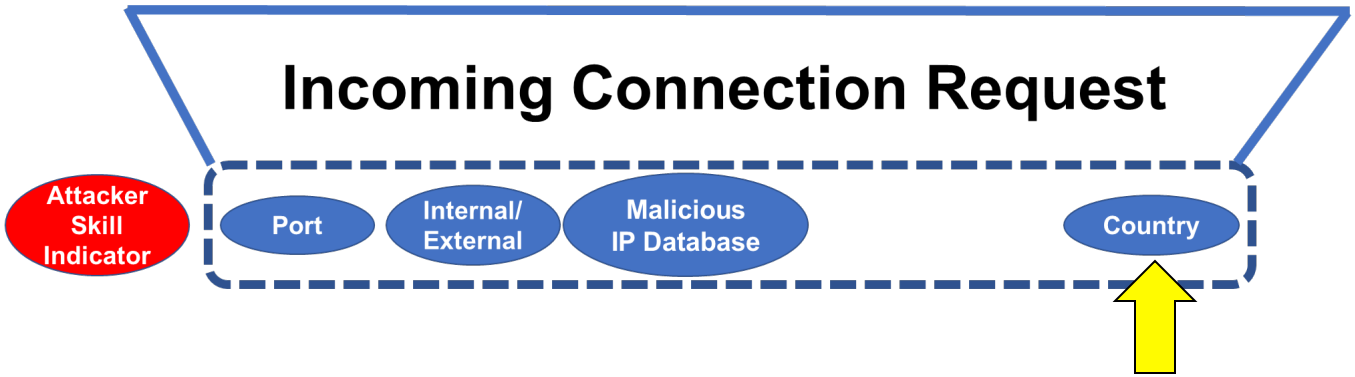


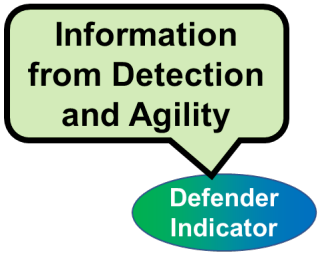


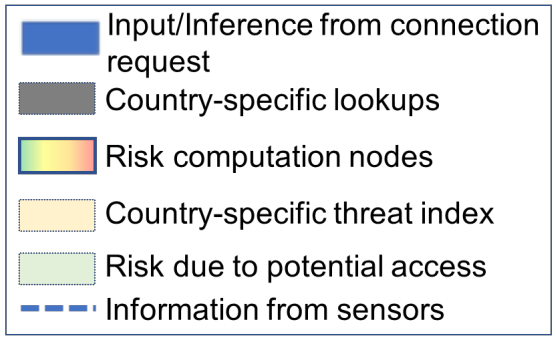




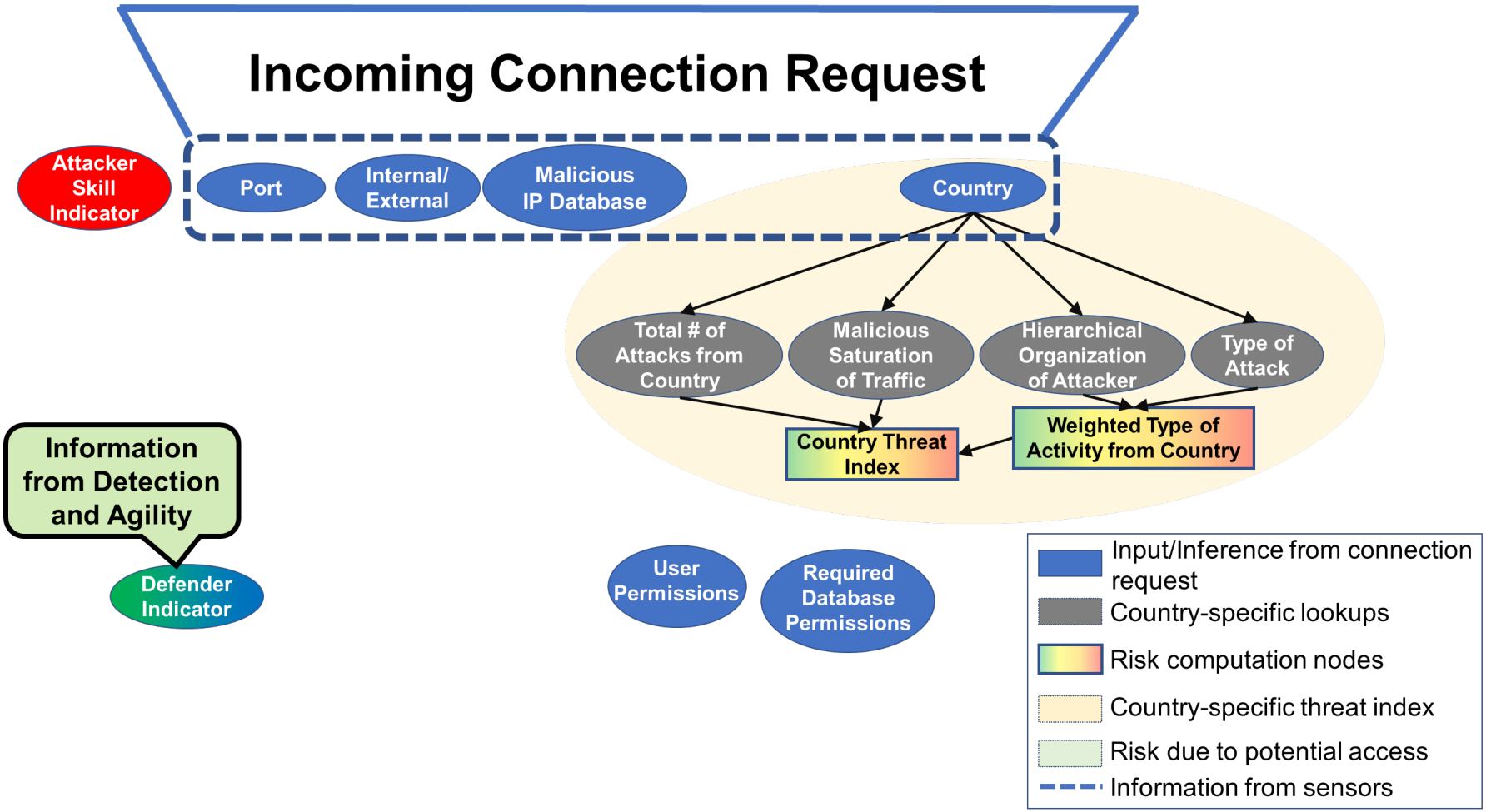
Country – Geographical origin of the connection request, as identified by IP



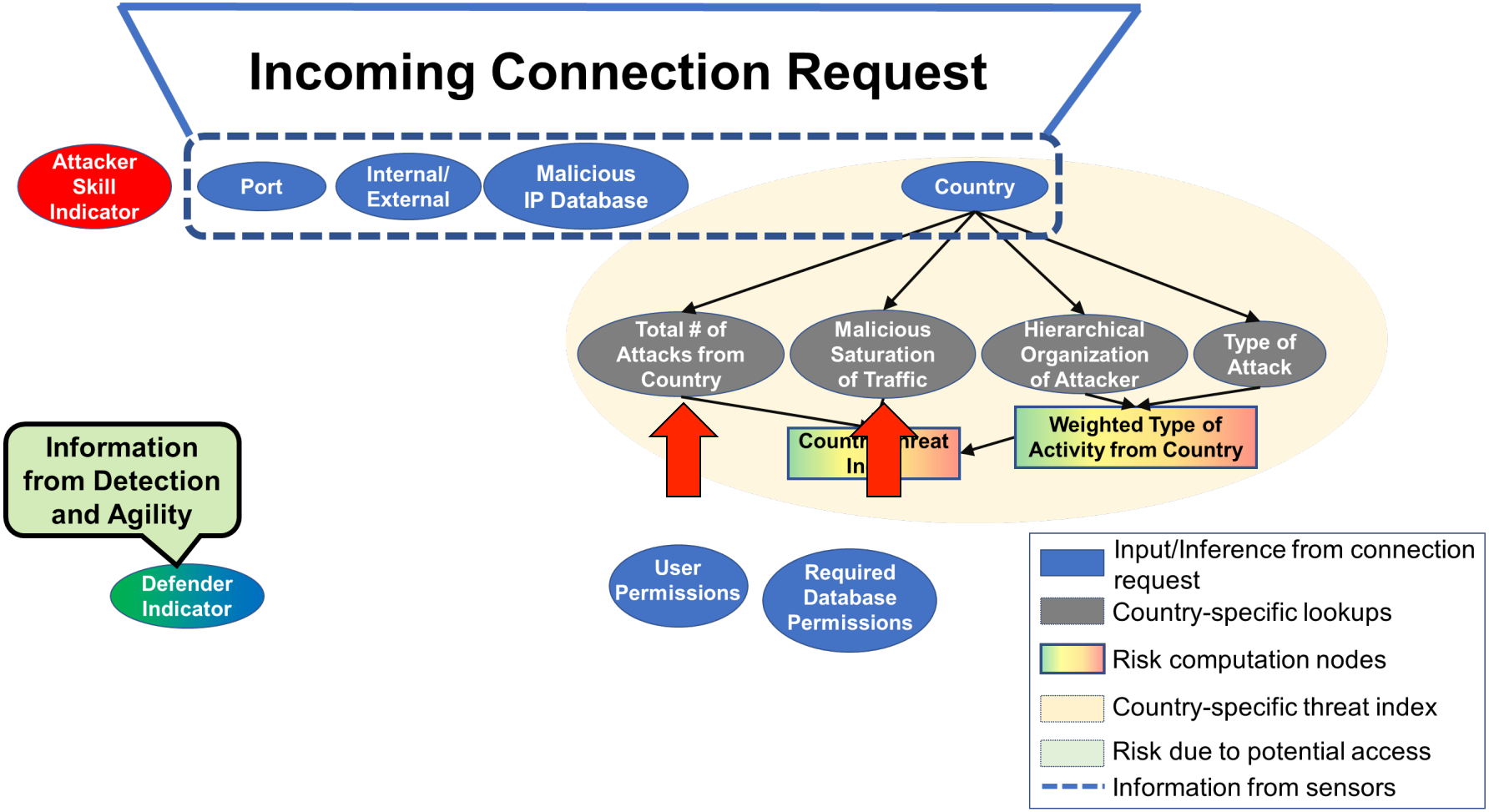




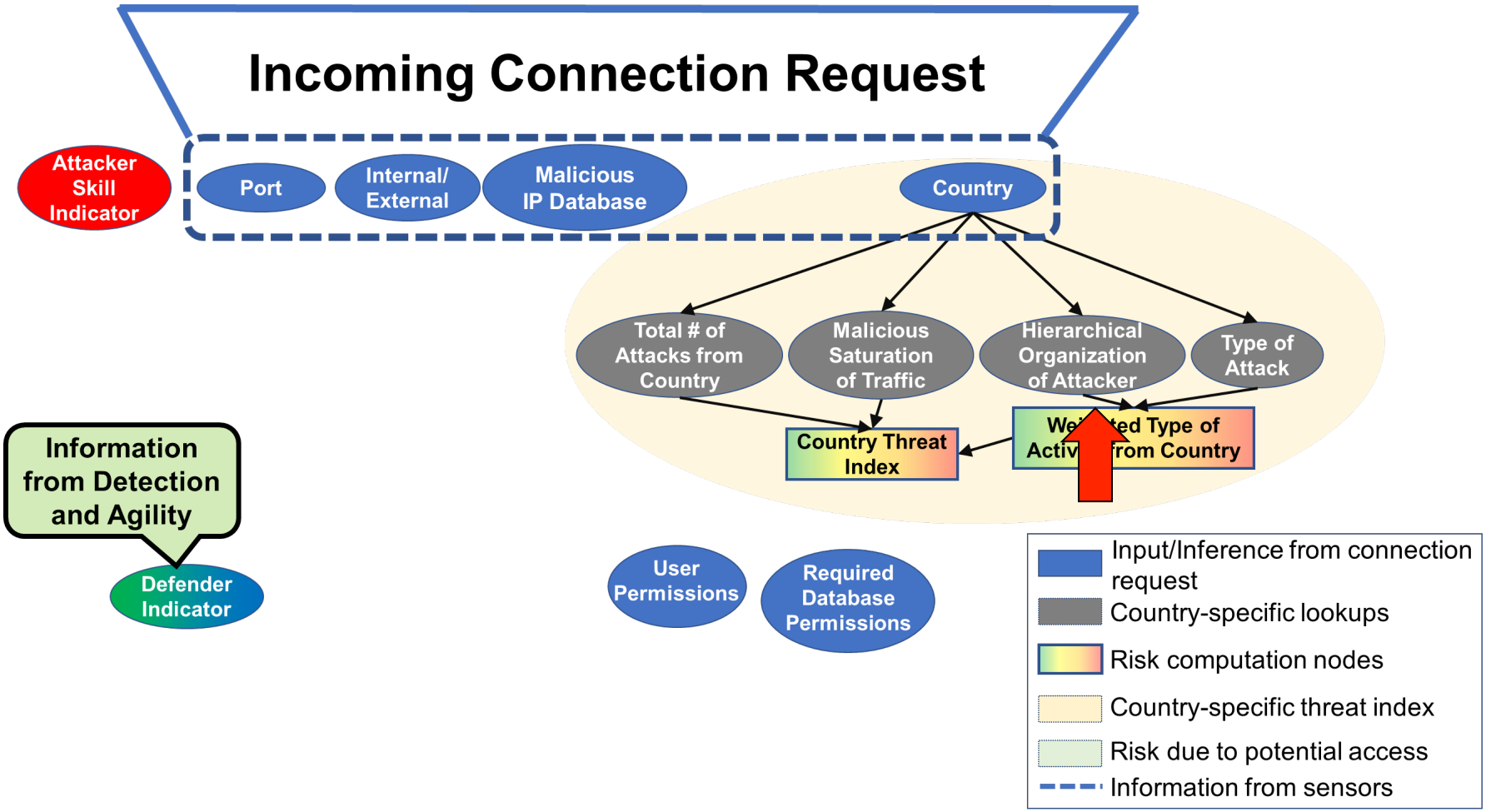
Step 3 – Include country-specific lookups



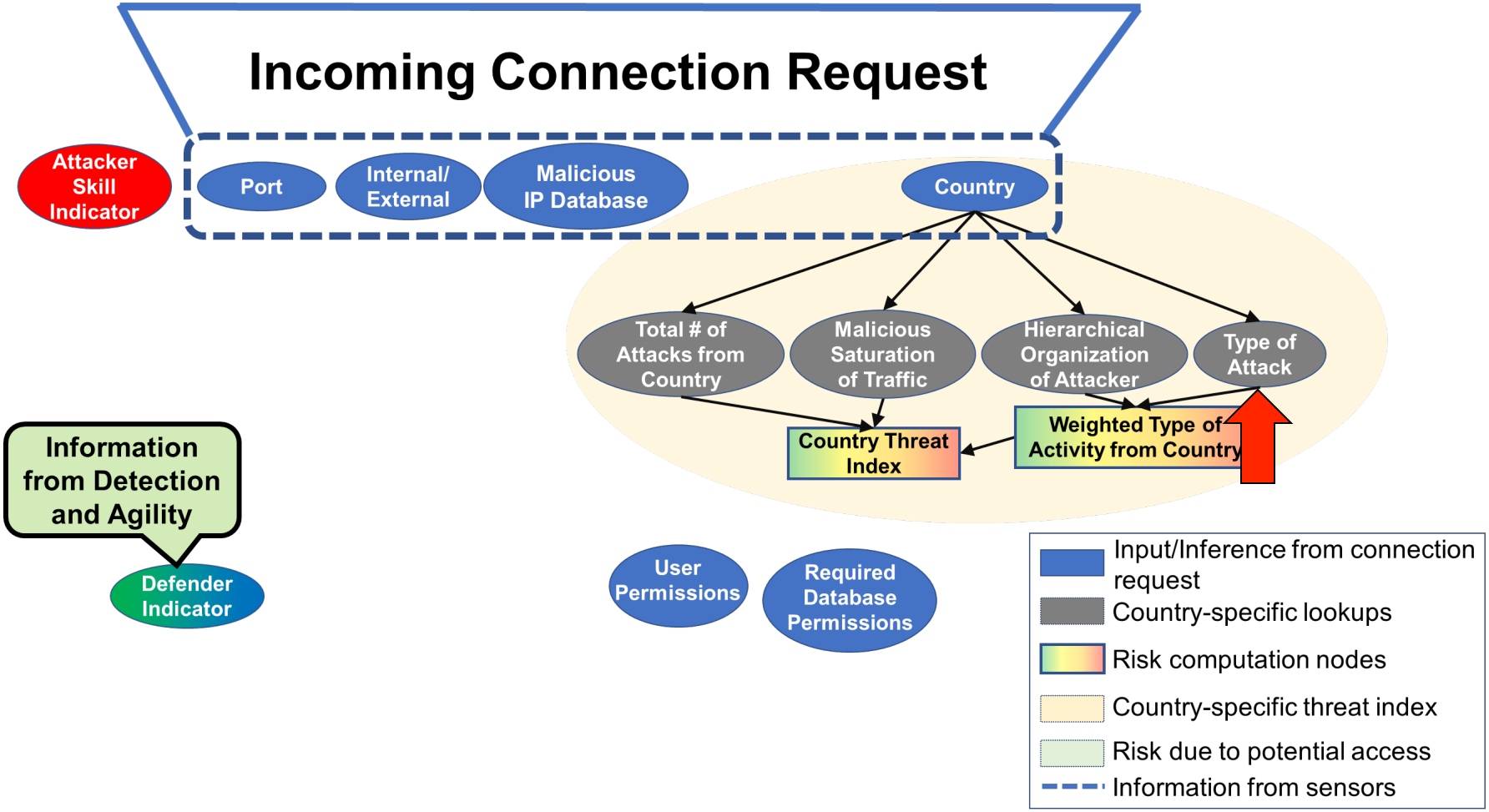
No. of attacks from country – Total logged attacks from a country in a year
Malicious Saturation of Traffic - % of traffic which is malicious



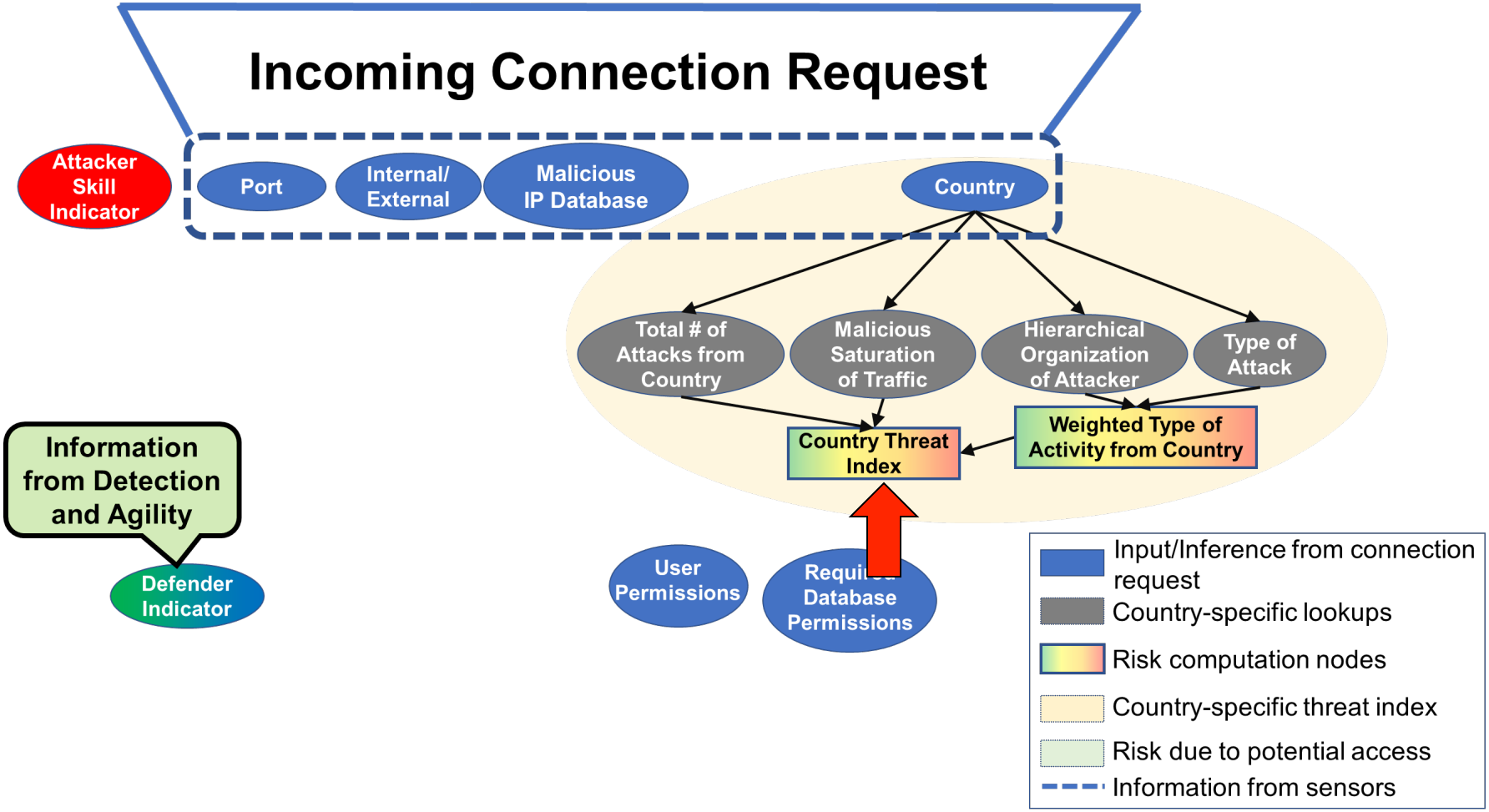
Hierarchical Organization of Attacker – Individual, Independent group, State Tolerated, State Funded attackers



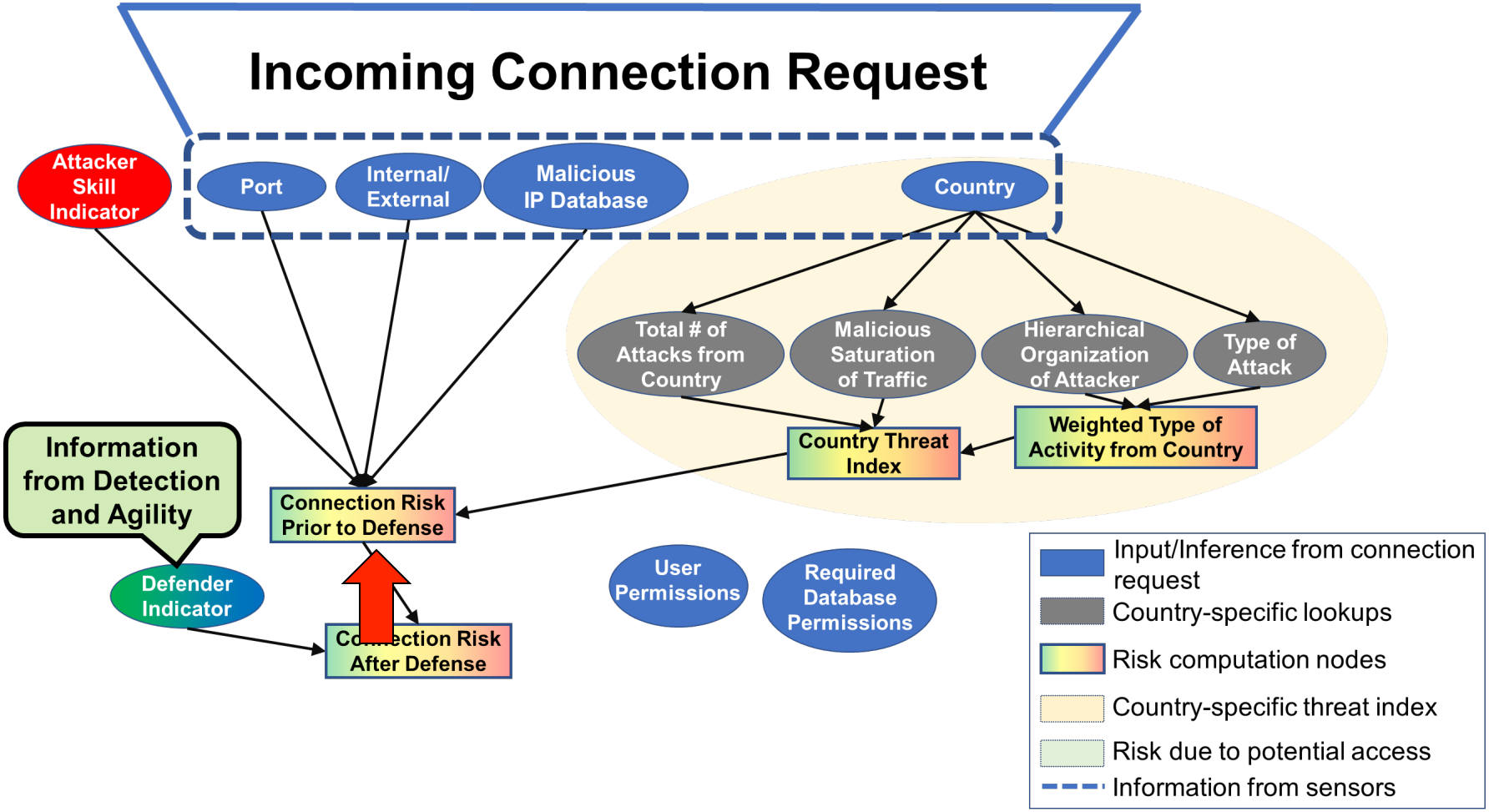
Type of Attack – Captures risk associated with type of attack (Botnets – low risk, Phishing – Medium Risk, APT – High Risk)



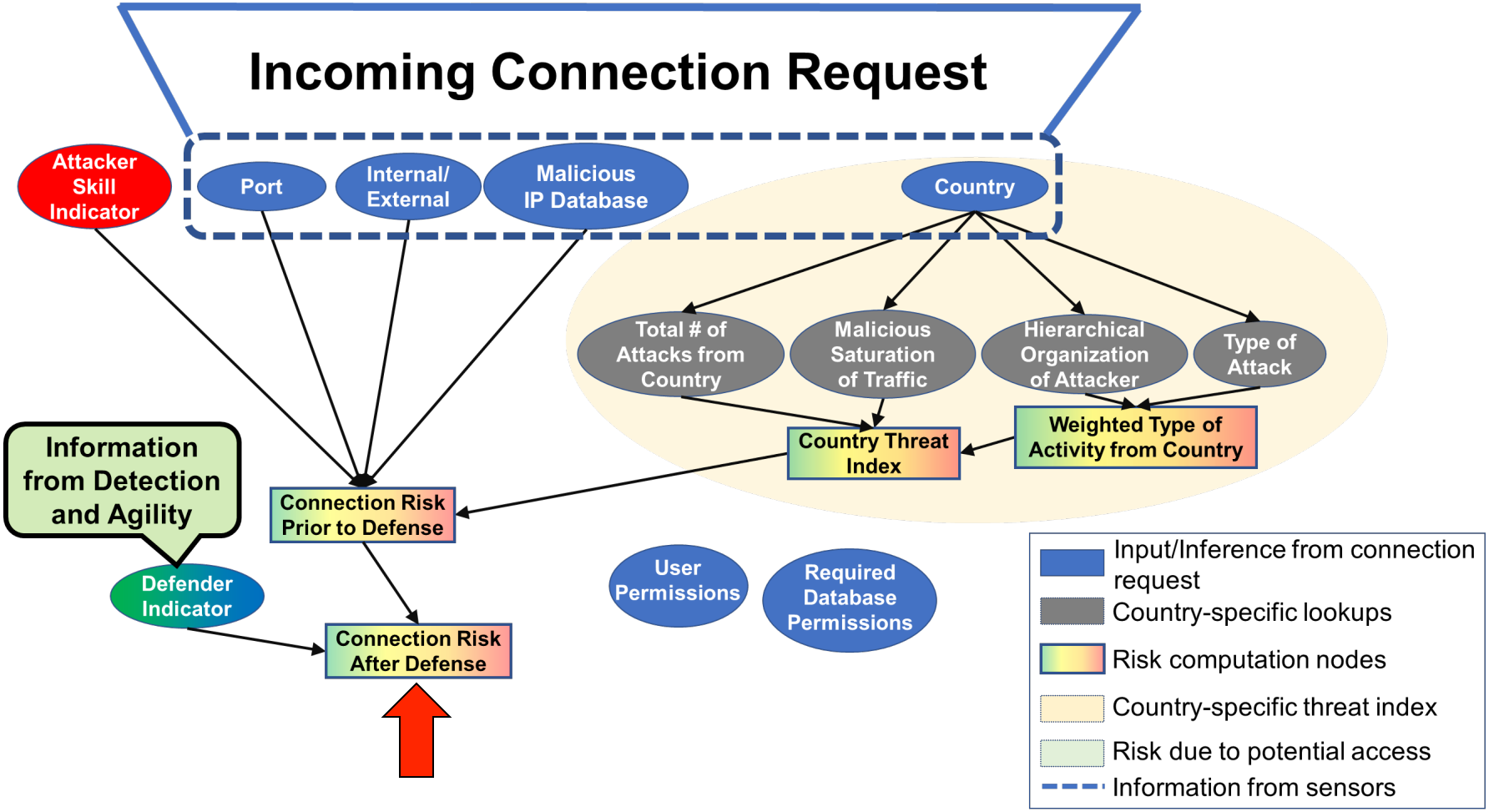
Country Threat Index – Aggregates and measures risk due to country-specific metrics



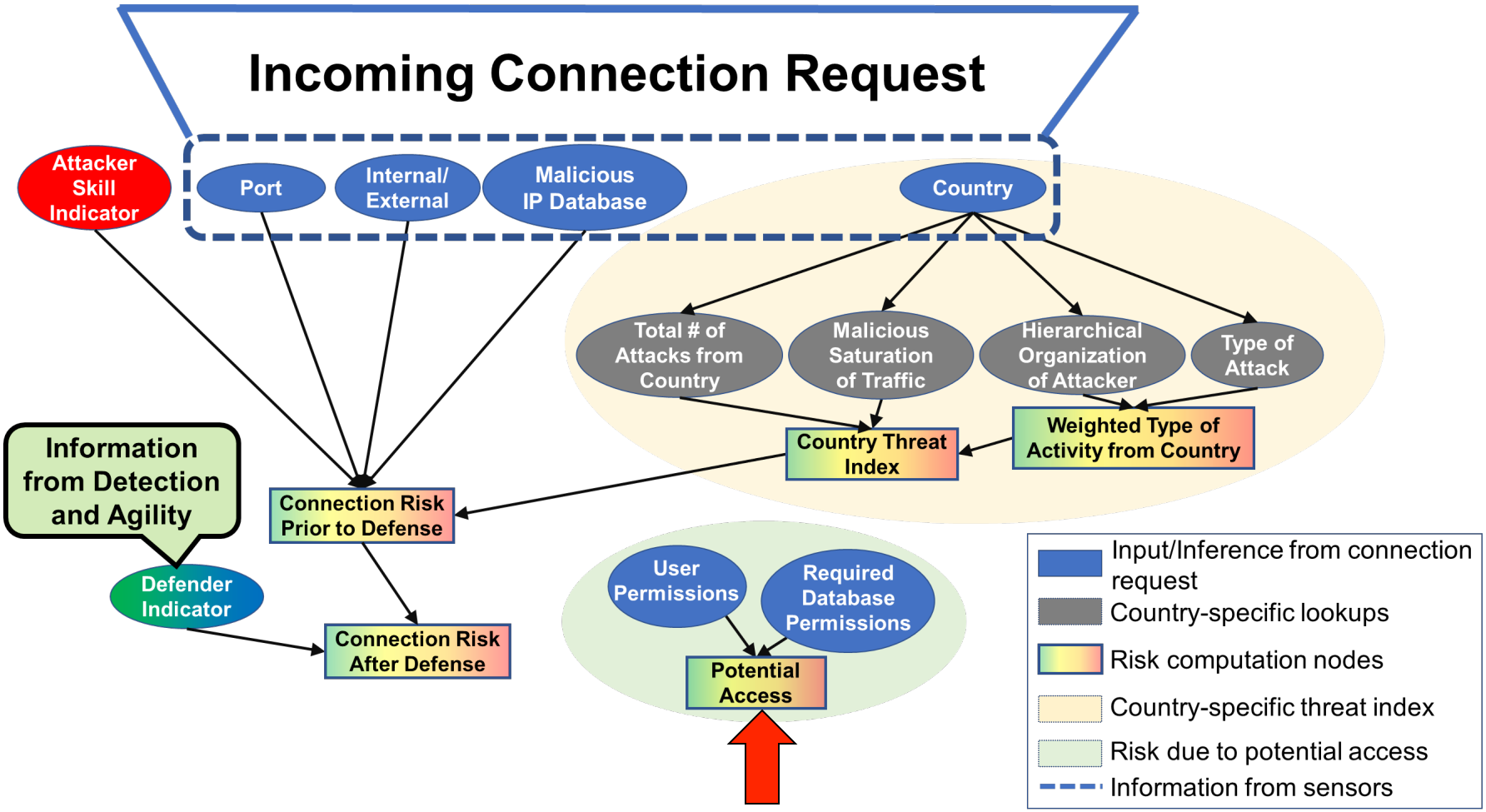
Connection Risk Prior to Defense – Aggregates risk from the connection metrics, before defender skill metric is accounted for



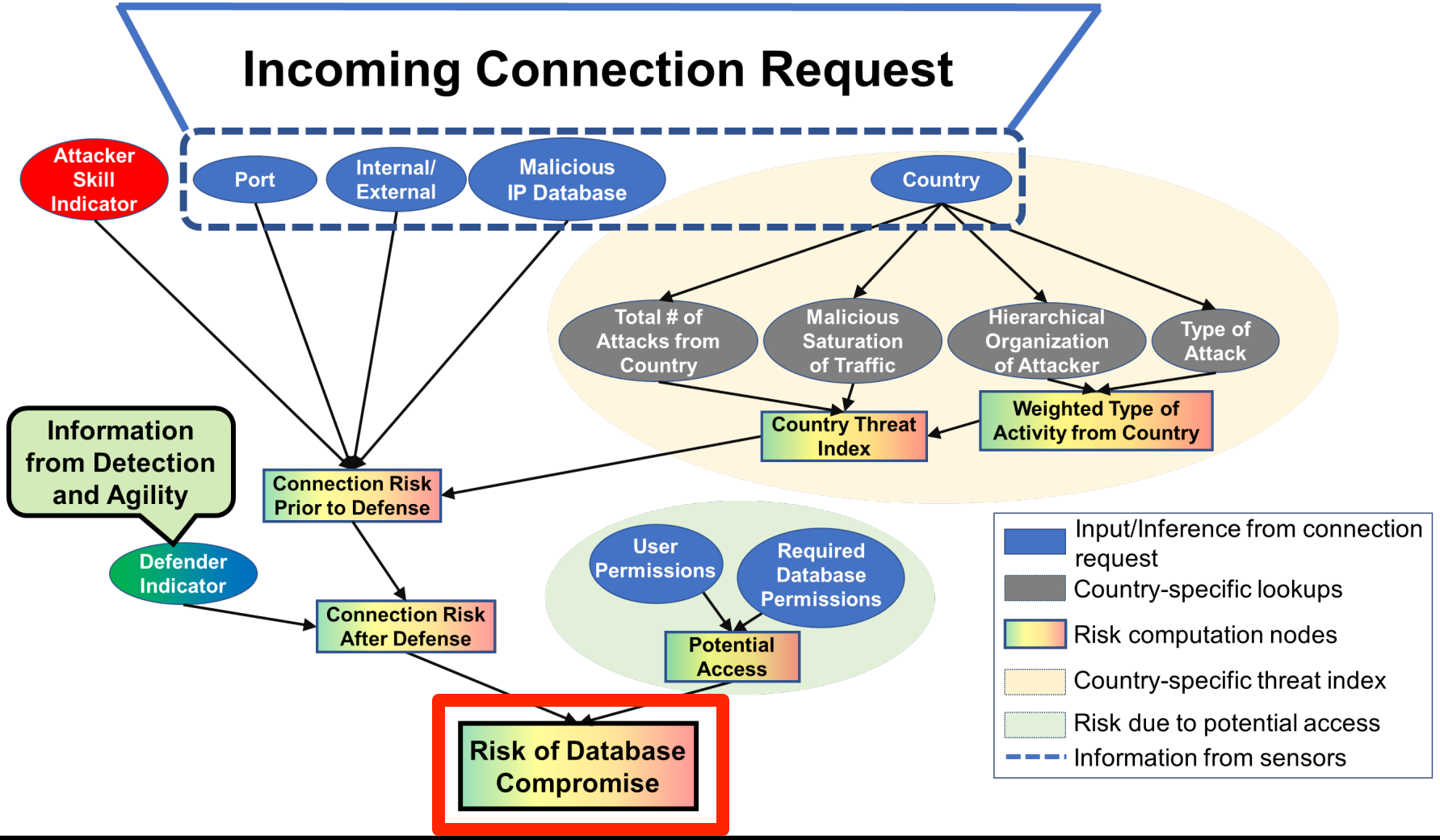
Connection Risk After Defense – Aggregates risk after accounting for defender skill metric



Potential Access – What is the potential that the query is successful?



Final Step – Aggregate risk due to all the accounted metrics in final risk node

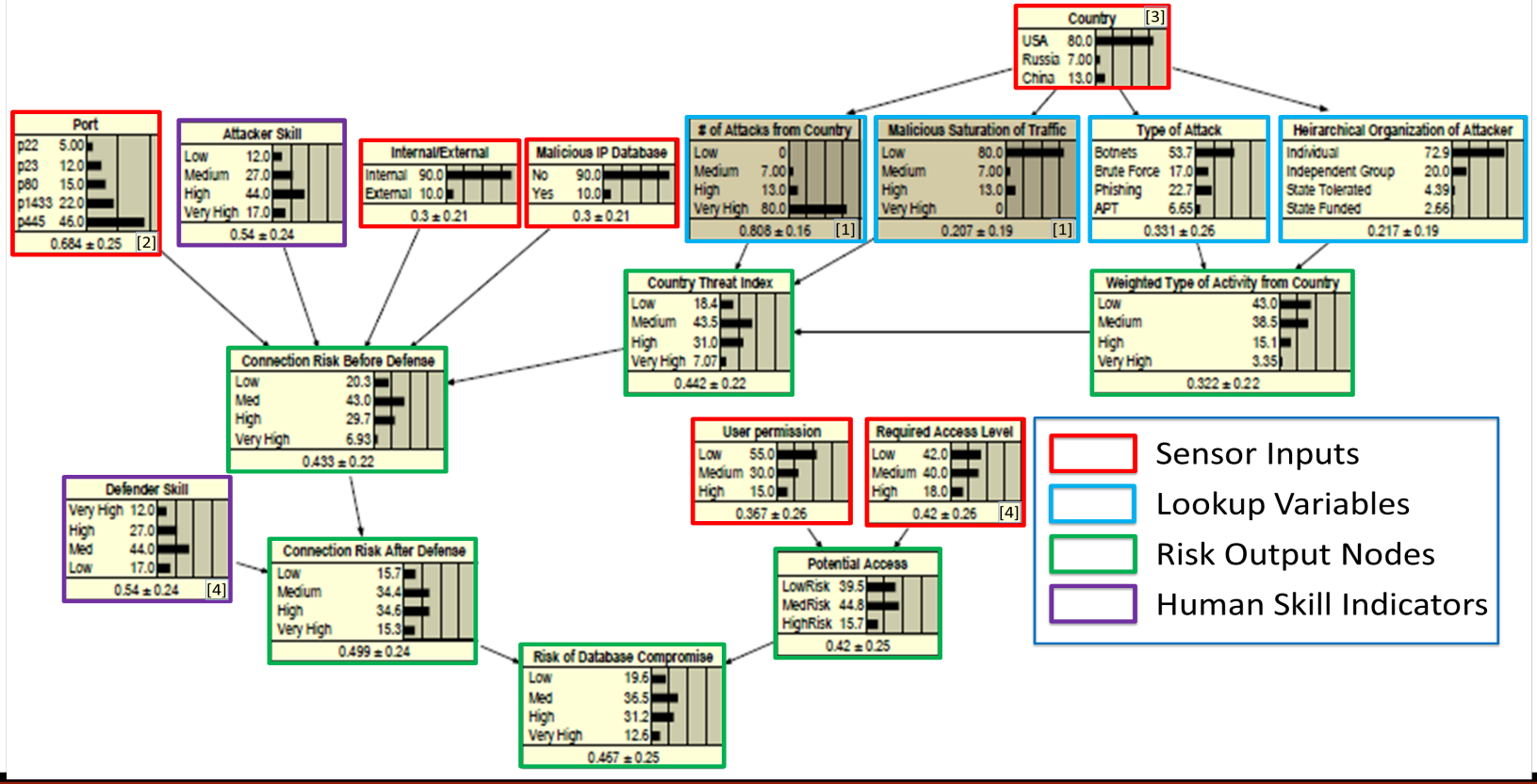


Sources of Uncertainty

- Skilled attackers can spoof IP address and appear to be on the internal network
- First true origin of the connection request might be untraceable
- Spoofing user permissions presents risk to the database
- Specification bias in the model

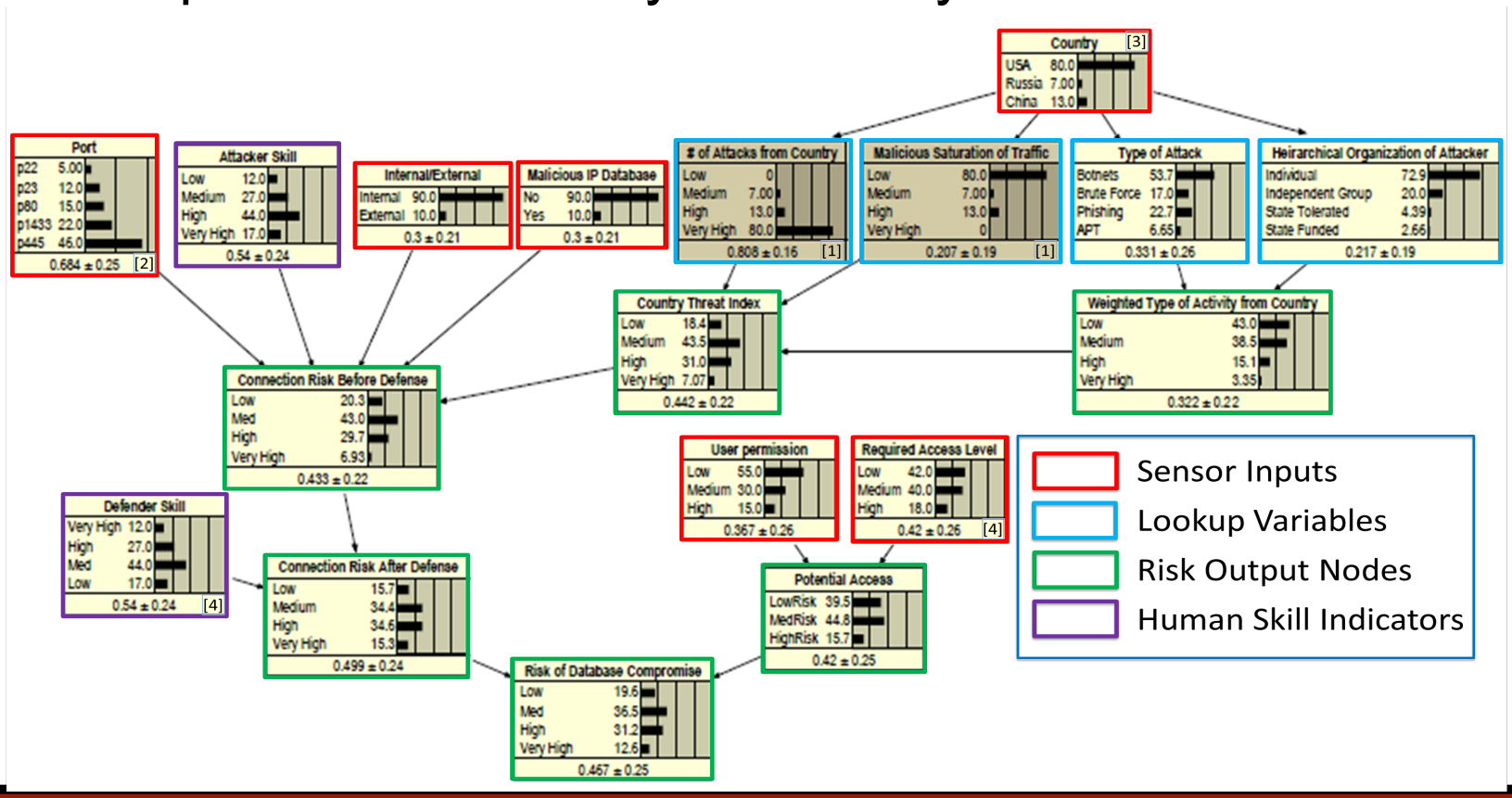
Statistical Application

- Implemented conceptual framework as Bayesian Network
- Directed edges represent dependencies
- Figure shows marginal distribution for each node



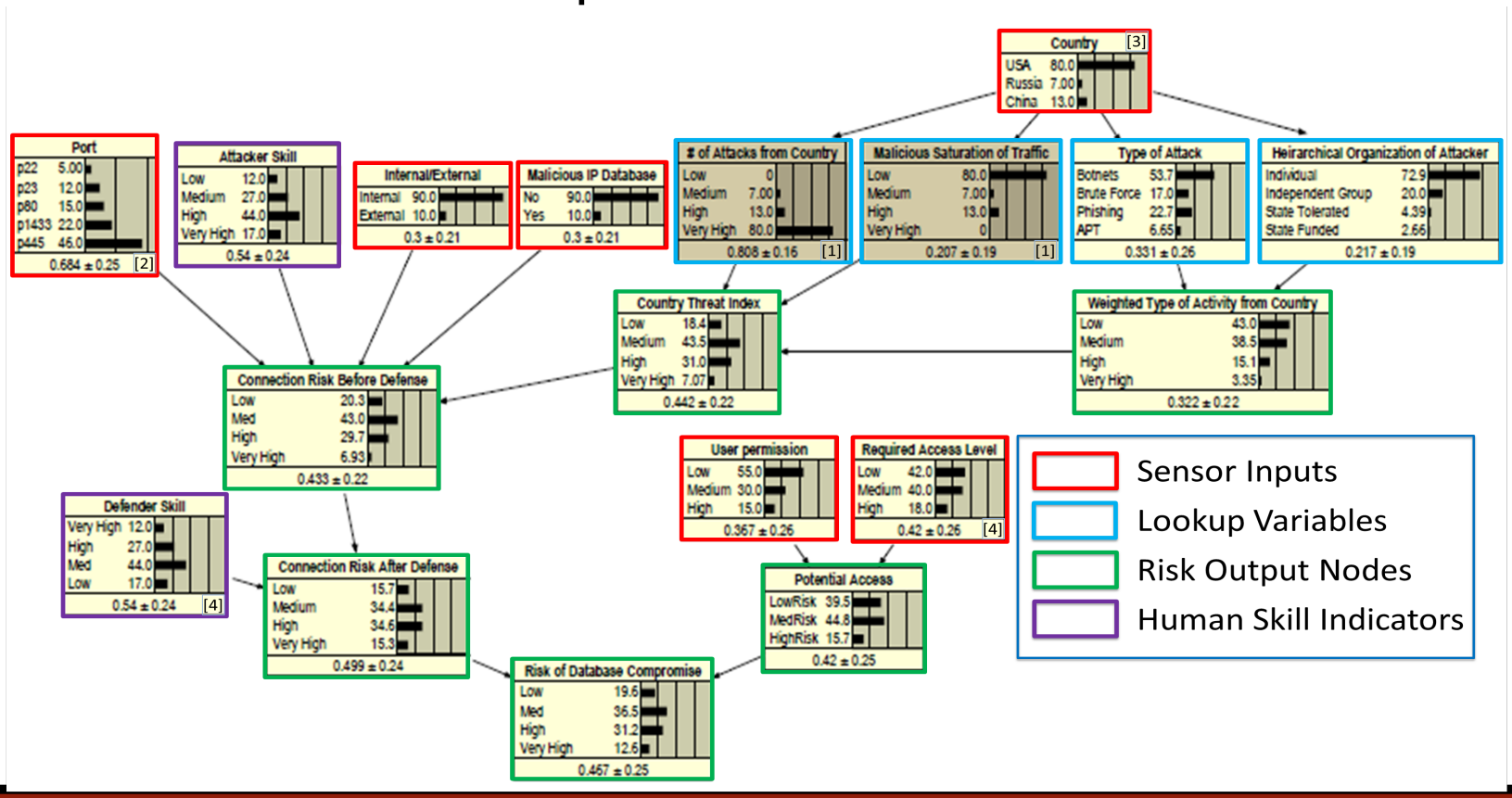
Statistical Application

- Priors for Sensor inputs inducted from cyber reports
- Conditional probability tables hypothesized by collaborating with experts in risk and cybersecurity



Statistical Application

- Risk to database calculated by conditional probability P(R|S)
- S is the input state of the model – observed by setting evidence for sensor inputs and human skill indicators



Results

- Evidence set for ***hypothetical scenarios***
- S_1 (Low – Medium Risk)
- S_2 (High Risk)

Variable	State S_1	State S_2
Port (P)	p80 (Medium risk)	p22 (Very high risk)
Attacker Skill (AS)	Medium Skill (Medium to high risk)	Medium Skill (Medium to high risk)
Connection (C0)	Internal, (Low risk)	External, (High risk)
Malicious IP Database (IP)	Not listed, (Low risk)	Malicious Listed IP, (High risk)
Country Threat Index (CTI)	P(L| Country = USA) = 0.203 P(M| Country = USA) = 0.457 P(H| Country = USA) = 0.289 P(VH| Country = USA) = 0.051	P(L| Country = China) = 0.061 P(M| Country = China) = 0.308 P(H| Country = China) = 0.445 P(VH| Country = China) = 0.185
Defense (D)	High Skill (Medium to low risk)	High Skill (Medium to low risk)
User Permission (UP)	Low, (Low risk)	High, (High risk)
Required Access Level (RAL)	Low, (Low risk)	High, (High risk)
Risk of Database Compromise (R)	**$P(L\|S_1) = 0.383$** **$P(M\|S_1) = 0.376$** **$P(H\|S_1) = 0.161$** **$P(VH\|S_1) = 0.08$**	**$P(L\|S_2) = 0.098$** **$P(M\|S_2) = 0.215$** **$P(H\|S_2) = 0.508$** **$P(VH\|S_2) = 0.179$**

Variable	State S_1	State S_2
Port (P)	p80 (Medium risk)	p22 (Very high risk)
Attacker Skill (AS)	Medium Skill (Medium to high risk)	Medium Skill (Medium to high risk)
Connection (C0)	Internal, (Low risk)	External, (High risk)
Malicious IP Database (IP)	Not listed, (Low risk)	Malicious Listed IP, (High risk)
Country Threat Index (CTI)	P(L| Country = USA) = 0.203 P(M| Country = USA) = 0.457 P(H| Country = USA) = 0.289 P(VH| Country = USA) = 0.051	P(L| Country = China) = 0.061 P(M| Country = China) = 0.308 P(H| Country = China) = 0.445 P(VH| Country = China) = 0.185
Defense (D)	High Skill (Medium to low risk)	High Skill (Medium to low risk)
User Permission (UP)	Low, (Low risk)	High, (High risk)
Required Access Level (RAL)	Low, (Low risk)	High, (High risk)
Risk of Database Compromise (R)	**$P(\text{Low Risk}\|S_1) = 0.383$** **$P(\text{Medium Risk}\|S_1) = 0.376$** **$P(\text{High Risk}\|S_1) = 0.161$** **$P(\text{Very High}\|S_1) = 0.08$**	**$P(\text{Low Risk}\|S_2) = 0.098$** **$P(\text{Medium Risk}\|S_2) = 0.215$** **$P(\text{High Risk}\|S_2) = 0.508$** **$P(\text{Very High Risk}\|S_2) = 0.179$**

Conclusions and Future Tasks

- Quantitatively integrated humans as risk factors in network risk calculations
- Developed a metric to indicate relative risk by a country
- Model provides a reasonable estimation of risk for different conditions of the network

Future Tasks

- Validation of analysis
  - Validate against DETER testbed with modelled attackers and defenders
  - Assess model performance dynamically

Thank you!

varagarw@indiana.edu