BAYESIAN NETWORK

Ransomware Infection in a Financial Institution

Casaleggi Elena - 3319326
Fatigati Veronica - 3151595
Lo Giudice Matteo - 3303092
Salutari Tommaso - 3173886

RoadMap

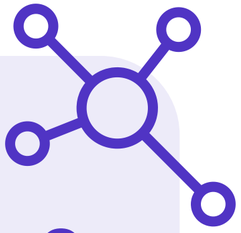
- What, Why & Strategy
- Model
- Weaknesses
- Scenarios

What are we analysing?

*"How do **internal security practices** and **employee behavior** influence the probability and consequences of a **successful ransomware attack** on a financial institution."*

Why is it relevant?

- **Financial institutions** are prime ransomware **targets**.
- Attacks cause major **operational**, **financial**, and **reputational harm**.
- **Human error** and **defenses** strongly affect breach risk.
- Understanding these links supports better **risk management**.



Why use Bayesian Networks?

- Handles **complex interdependencies.**
- Supports **probabilistic reasoning** under uncertainty.
- Enables **scenario simulation** for proactive **decision-making**.
- **Visual models** make it easier to see how threats and defenses are connected.

Data Inputs

Probability-Based Variables: all nodes defined by conditional probability tables (CPTs)

Hybrid Parameterization

- Expert Elicitation: used AgenaRisk's NPT editor to encode our **team's judgments** where data are scarce.
- Public Benchmarks: populated CPT entries with **industry statistics** and **regulator reports**.

Real-world Constraints

- Confidentiality: True **incident data** are **proprietary**; we use proxy probabilities instead.
- Model Complexity: Hundreds of interdependent variables mean CPTs represent **best-available estimates**, not absolute truths.

Modelling Strategy

1. Defined the Risk Event

Defined a full ransomware path from phishing email to business impacts.

2. Built the Bayesian Network

Converted that path into a causal network linking triggers, controls, consequences and mitigations.

3. Populated Probabilities

Blended expert judgment with published stats, noting key uncertainties.

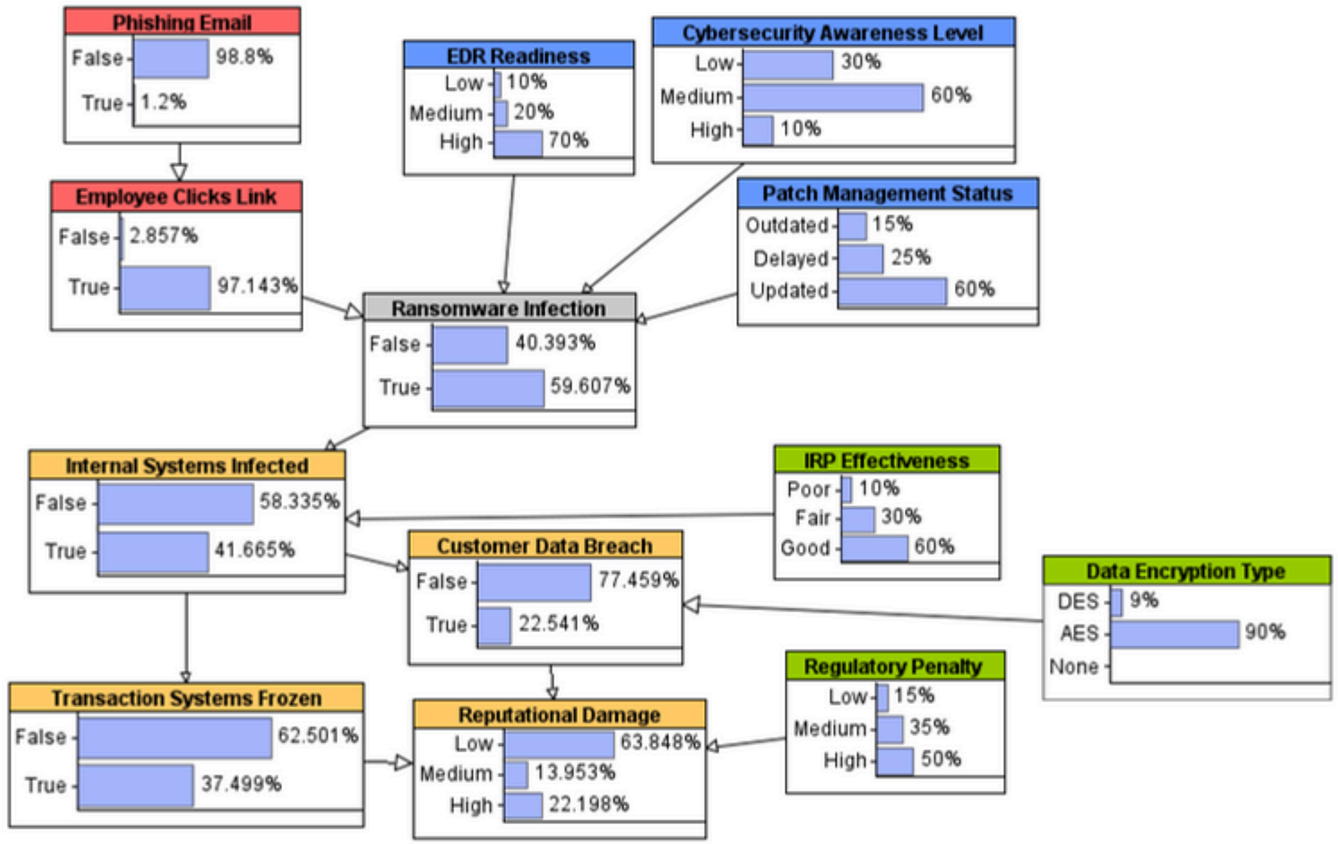
4. Ran “what-if” Simulations

Tested different scenarios to update risk odds.

5. Identified top Fixes

Used sensitivity results to single out the controls and mitigations that delivered the greatest risk reduction.

The Model



STRUCTURE:

- **Two triggers** – Phishing Email and Employee Clicks Link
- **One Risk Event** – Ransomware Infection
- **Three Controls** – EDR Readiness, Cybersecurity Awareness Level, and Patch Management Status
- **Four Consequences** – Internal Systems Infected, Transaction Systems Frozen, Customer Data Breach, and Reputational Damage
- **Three Mitigations** – IRP Effectivenessl, Data Encryption Type, and Regulatory Penalty

Triggers

PHISHING EMAIL

Boolean: True/False.

This node represents whether an email is a phishing attempt or not, effectively capturing the binary nature of the event. We opted for a Boolean type because it allows us to model the ratio of phishing emails among all received messages as a probability of the “True” state. This approach aligns well with the underlying concept and simplifies integration into the broader risk model.

EMPLOYEE CLICKS LINK

Boolean: True/False.

This node represents whether an employee clicks on a link in an email, capturing a clear binary outcome. A Boolean type is appropriate because the event can only occur in one of two states — the link is either clicked or not. This simplification effectively supports modeling human behavior in the phishing attack pathway.

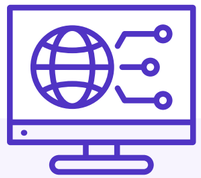
Risk Event

RANSOMWARE INFECTION

Boolean (True/False)

This node captures the occurrence of a ransomware infection, a binary outcome that depends on the presence of certain triggers (e.g., phishing and user behavior) and the effectiveness of relevant controls. A Boolean type is appropriate here, as the event either takes place or it doesn't. This simplifies the modeling of the risk event while still allowing for nuanced influence from upstream variables.

Controls



EDR READINESS

Ranked (Low/Medium/High)

This node represents the organization's level of preparedness in deploying and managing EDR tools. We chose a ranked type because it reflects a clear hierarchy of effectiveness, something a Boolean or labelled type would fail to capture. In cybersecurity, EDR Readiness is a key control factor in preventing and containing endpoint infections, and ranked variables are ideal for modeling such progressive capabilities.

CYBERSECURITY AWARENESS

Ranked (Low/Medium/High)

This node reflects how well employees understand and respond to cyber threats. A ranked type captures the progressive nature of awareness and its impact on reducing risk.
As with EDR Readiness, the hierarchical structure allows us to model increasing levels of protection and control effectiveness in a realistic way.

PATCH MANAGEMENT STATUS

Ranked (Outdated/Delayed/Updated)

This node represents how effectively the organization applies software updates to fix vulnerabilities.
A ranked type is appropriate, as improvements in patch management correspond to progressively stronger protection against infections, aligning well with its role as a control variable.

Consequences

INTERNAL SYSTEMS INFECTED

Boolean (True/False)

This node represents whether internal systems have been compromised by malware or unauthorized software. A Boolean type is appropriate to keep the model straightforward, focusing only on whether infection occurs, without adding complexity related to the extent or severity of the compromise.

CUSTOMER DATA BREACH

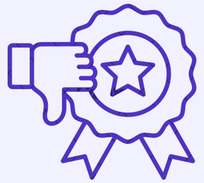
Boolean (True/False)

This node represents whether unauthorized access to sensitive customer data has occurred. We chose a Boolean type to keep the model simple and focused—either a breach happens or it doesn't. While we considered ranking the node to reflect breach severity, the lack of sufficient data made a binary approach more practical and consistent with the model's level of abstraction.

TRANSACTIONS FROZEN

Boolean (True/False)

This node represents whether critical business systems, such as payment or processing platforms, become inoperable due to a cyberattack. We selected a Boolean type to maintain model simplicity—either the systems are frozen or they are not. While more granular modeling was possible, a binary approach aligns with the overall abstraction level and available data.

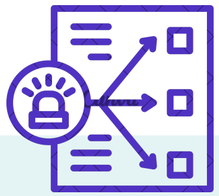


REPUTATIONAL DAMAGE

Ranked (Low/Medium/High)

This node captures the degree of trust and credibility loss an organization may suffer after a cyber incident, potentially leading to customer loss, negative publicity, and lasting brand impact. We chose a ranked type to reflect varying levels of severity, which are meaningful to estimate and interpret. Compared to other variables, assigning probability distributions across these levels was more feasible.

Mitigations



IRP EFFECTIVENESS

Ranked (Poor/Fair/Good)

This node represents how effectively the organization can detect, respond to, and recover from cyber incidents through its Incident Response Plan. A ranked type is appropriate because IRP effectiveness progresses in clearly ordered levels, reflecting real-world improvements in preparedness and coordination. This structure captures expert judgment and uncertainty without requiring precise numerical inputs, making it ideal in this case.

DATA ENCRYPTION TYPE

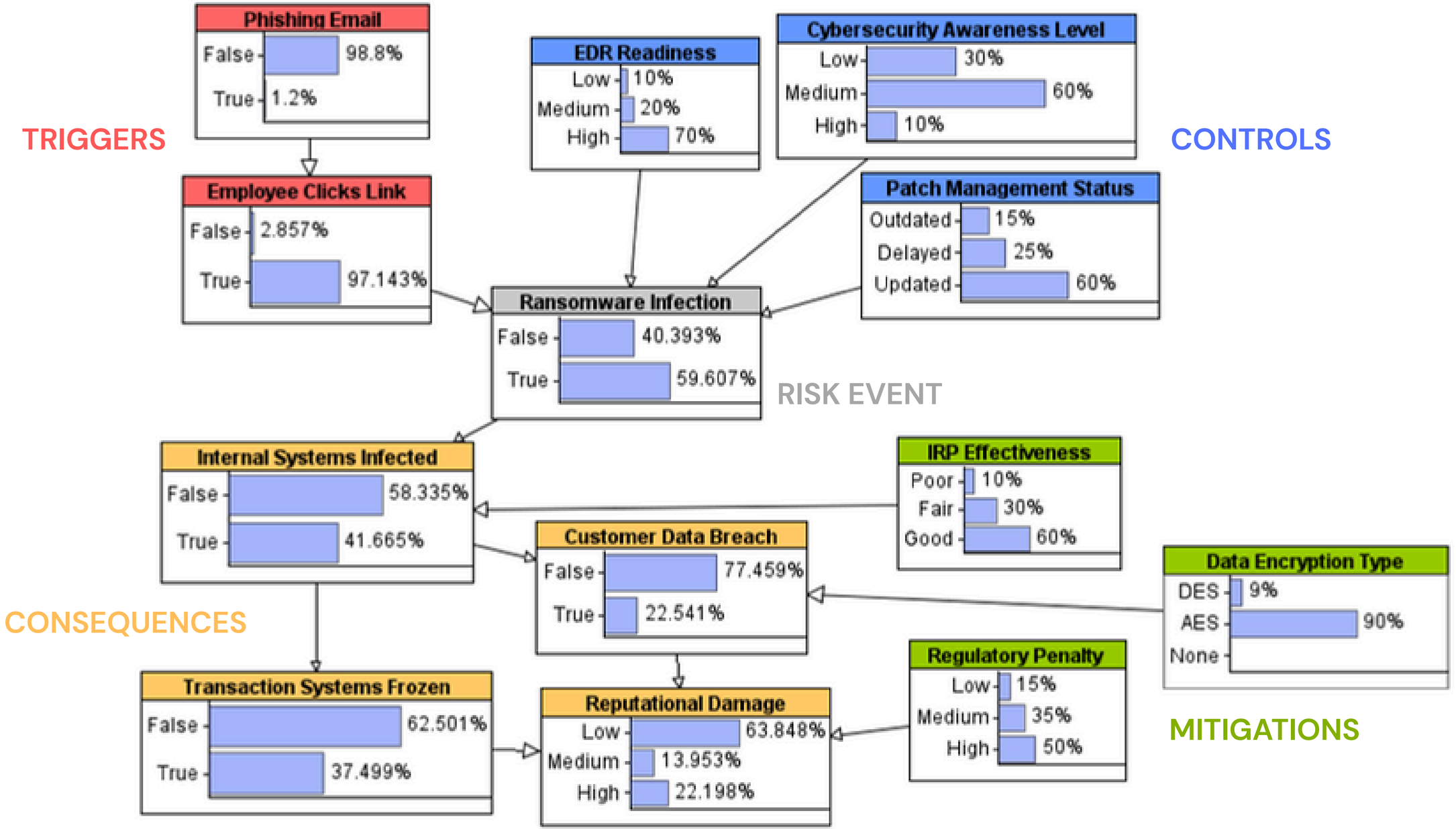
Labelled (None/DES/AES)

This node represents the cryptographic standard used to protect data. DES is an outdated symmetric algorithm with weak security, while AES is the current standard, offering much stronger protection. We chose a labelled variable because the values are distinct and non-ordered, there's no numerical or ordinal relationship between them, allowing to model the impact of each method on risk without implying hierarchies.

REGULATORY PENALTY

Ranked (Low/Medium/High)

This node captures the potential financial or legal consequences an organization may face for non-compliance. A ranked type is appropriate, as penalties generally fall into clearly ordered categories reflecting increasing severity. While exact numerical values may vary, this structure enables a realistic and interpretable representation of risk, especially when relying on qualitative assessments.


TRIGGERS
CONTROLS
Phishing Email
False 98.8%
True 1.2%
EDR Readiness
Low 10%
Medium 20%
High 70%
Cybersecurity Awareness Level
Low 30%
Medium 60%
High 10%
Employee Clicks Link
False 2.857%
True 97.143%
Patch Management Status
Outdated 15%
Delayed 25%
Updated 60%
Ransomware Infection
False 40.393%
True 59.607%
RISK EVENT
Internal Systems Infected
False 58.335%
True 41.665%
IRP Effectiveness
Poor 10%
Fair 30%
Good 60%
Customer Data Breach
False 77.459%
True 22.541%
Data Encryption Type
DES 9%
AES 90%
None
CONSEQUENCES
Transaction Systems Frozen
False 62.501%
True 37.499%
Reputational Damage
Low 63.848%
Medium 13.953%
High 22.198%
Regulatory Penalty
Low 15%
Medium 35%
High 50%
MITIGATIONS

Our Assumptions

To estimate the probabilities of the two triggers, we used data from StationX e JumpCloud.

- **Phishing Email** node, we noted that approximately 1.2% of global email traffic is classified as malicious, which amounts to around 3.4 billion phishing emails sent every day
- **Employee Clicks Link** we found that in finance and insurance sectors employees show a phishing email click rate of 26.6%, indicating a heightened exposure to these types of cyber threats.

Cybersecurity Awareness

- **Low (30%)**: Often found in traditional financial institutions with older staff or passive/non-participants in training programs.
- **Medium (60%)**: A baseline level is expected due to regulatory pressures (e.g., DORA compliance).
- **High (10%)**: Typically seen in IT departments, where specialized knowledge and awareness are consistently strong.

EDR Readiness

- **Low (10%):** Rare in the financial sector; indicates either absence or minimal deployment—generally unacceptable.
- **Medium (20%)**: Seen in institutions currently upgrading systems or in early digitalization phases.
- **High (70%)**: Common in well-structured, digital-native, or large financial institutions with mature cybersecurity postures.

Patch Management Status

- **Outdated (15%)**: Rare but possible in legacy systems or institutions with limited IT resources; indicates irregular or manual patching processes.
- **Delayed (25%)**: Common in institutions undergoing digital transformation, where patching is scheduled but not fully automated or prioritized.
- **Updated (60%)**: Found in mature or digital-first financial institutions, with automated, centralized patch management ensuring timely updates and reduced exposure.

Sources:
- StationX
- JumpCloud

Our Assumptions

RANSOMWARE INFECTION

In estimating the overall risk event, we applied a weighted approach based on the type of control. Technical controls, such as patch management status and EDR readiness, were given greater weight due to their direct impact on an organization's ability to prevent or contain cyber threats. In contrast, cybersecurity awareness was considered less influential, as its effects are more indirect and dependent on individual behavior.

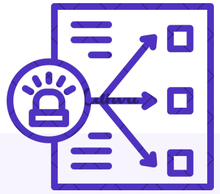
Our Assumptions

On the Consequences

For the **consequence values**, we relied on manual estimations. This approach allowed us to incorporate **contextual knowledge** and **sector-specific judgment**, especially where quantitative data was limited or where professional evaluation provided more accurate insight into potential impacts.

In particular, we considered that both transaction frozen and customer data breach directly depend on the status of internal systems, further justifying a tailored, qualitative assessment.

Our Assumptions



IRP EFFECTIVENESS

- **Poor (10%):** No formal plan or untested response. Slow and uncoordinated response, unacceptable in most financial institutions.
- **Fair (30%):** Plan exists and is partially tested, but with execution gaps. Typical of institutions still evolving.
- **Good (60%):** Fully developed and regularly tested IRP with clear responsibilities. Typical of mature or large institutions.

DATA ENCRYPTION TYPE

- **None (1%)**: No encryption in place. Considered a worst-case scenario and highly improbable in financial institutions.
- **DES (9%)**: Use of DES encryption. Still found in some legacy systems but considered outdated and less secure.
- **AES (90%)**: AES encryption implemented. A strong, modern standard widely used to protect sensitive financial data.

REGULATORY PENALTY

- **Low (15%):** Minor breaches with quick response and full cooperation. Rare in finance due to strict compliance.
- **Medium (35%):** Issues like mishandling customer data or late reporting. Moderately likely in a regulated sector like finance.
- **High (50%):** Breaches of sensitive data or repeated failures. High risk in finance given GDPR's strict fines.

Information we don’t have

LIMITED INSIGHT INTO TARGETING CRITERIA

Our model does not account for **how** or **why** attackers select **specific targets**, nor does it include details about the **attacker’s identity** or **profile**. While this abstraction helps us focus on the **general impact** of a ransomware attack, it delimitates our ability to assess how factors like attacker strategy or target attractiveness influence the likelihood and nature of the incident.

EMPLOYEE’S PSYCHOPHYSICAL STATE

Stress, **fatigue**, or **distraction** at the time of email arrival can override even strong training and awareness. These **human conditions** are **unpredictable** and **unmeasurable** in real time, making them a persistent uncertainty and a key limitation in modeling cyber risk.

THIRD-PARTIES

Our model does not capture **third-party vendor dependencies**, which are common in financial institutions and can serve as alternative **ransomware entry points**. Without detailed data on vendor access and controls, this critical infection pathway remains unmodeled, adding uncertainty that stems from the complex and often opaque nature of supply chain security.

Omitted variables

Email filtering was not included in the model, despite being a critical first layer of defense that can block phishing attempts before they reach users, reducing the likelihood of ransomware initiation.

Faster detection and response times reduce ransomware spread, directly enhancing incident containment and influencing the effectiveness of the overall response strategy.

Clear and timely communication, both internal and external, helps preserve public trust, contain reputational damage, and reduce potential regulatory penalties.

Frequent and well-structured training serves as a key preventive factor, improving users' ability to detect phishing and lowering the likelihood of ransomware activation through user interaction.

Reliable and up-to-date backups support rapid system recovery, limiting operational disruption and reputational impact, while reducing pressure to consider ransom payment.

Mismeasurement Issues

Generalized Model Design for Broad Applicability

Cybersecurity Awareness Level Reduced to Fixed Categories

Data Breach Modeled Without Granularity

Mismeasurement Issues

Cybersecurity Awareness Level Reduced to Fixed Categories

Data Breach Modeled Without Granularity

Generalized Model Design for Broad Applicability

To enhance **flexibility** and ensure **broad applicability**, we adopted a generalized model structure not tied to any specific financial institution.

This design choice allows the network to accommodate a variety of organizational contexts, even though it naturally abstracts from certain institution-specific details and configurations.

Generalized Model Design for Broad Applicability

Data Breach Modeled Without Granularity

Mismeasurement Issues

Cybersecurity Awareness Level Reduced to Fixed Categories

While the model simplifies this node into **fixed categories**, this simplification abstracts from real-world complexity, it effectively captures the key behavioral patterns relevant to phishing risk without overcomplicating the model.

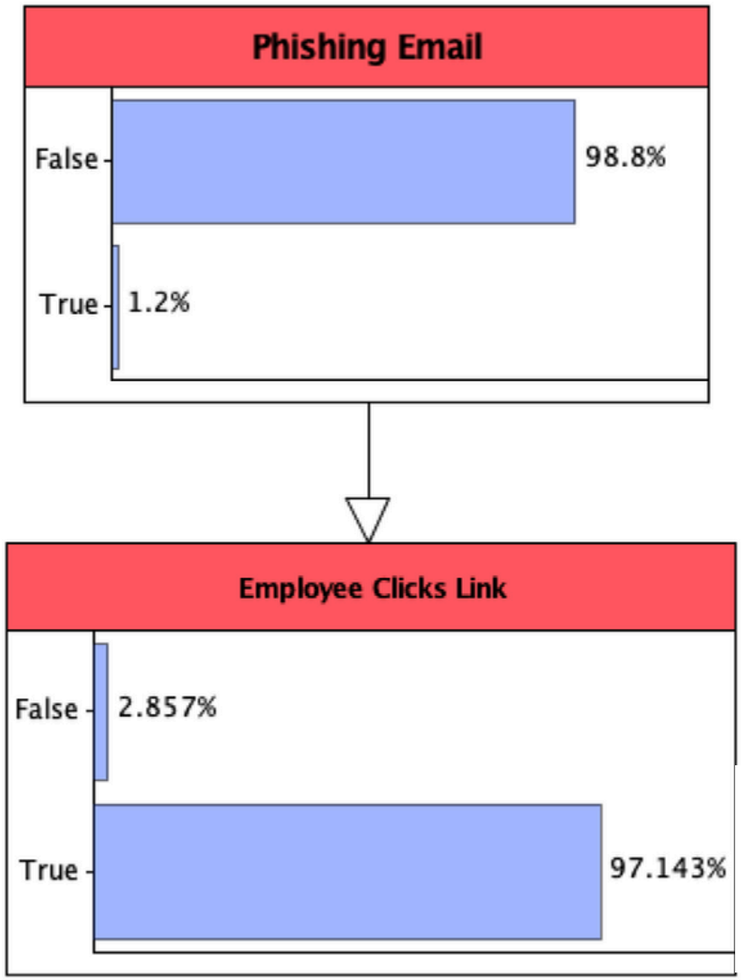
Mismeasurement Issues

Generalized Model Design for Broad Applicability

Cybersecurity Awareness Level Reduced to Fixed Categories

Data Breach Modeled Without Granularity

While this abstraction does not differentiate between types of compromised data, it effectively supports scenario-level reasoning by focusing on the **occurrence of a breach** and its potential regulatory and reputational consequences.



Impact of our ASSUMPTIONs

To reflect a **realistic attack** flow, the network positions ***"Phishing Email"*** before ***"Employee Clicks Link"***, following a logical and causal sequence.

This structure distinguishes between email delivery and user interaction, allowing us to model technical controls and human behavior separately.

In *agenaRisk*, this separation enables **clearer conditional probability** definitions and improves the **accuracy of scenario propagation**.

It mirrors the actual timeline of a phishing attempt, enhancing both model **interpretability** and **outcome** estimation.

Impact of our ASSUMPTIONs

To keep the model focused and interpretable, we deliberately **excluded ransom payment dynamics**, such as whether the organization decides to pay or receives decryption keys from the attacker.

This choice reflects the fact that such **decisions are highly variable**, depending on legal, ethical, and operational factors that differ across institutions.

In *agenaRisk*, this assumption allows for a c**leaner network structure** and **more reliable propagation**, although it may slightly overestimate the duration or severity of certain outcomes where ransom payment might have led to faster restoration.

Scenarios Overview

1. Worst-Case Scenario

- **100% phishing success**, low controls
- Full **ransomware infection**
- Major system outages, **data breach**
- **High reputational damage** and **high regulatory severity**

2. Best-Case Scenario

- **Strong controls** and **high awareness**
- **Low click rate** and rare infection
- Minimal system and data impact
- Very **low reputational damage**

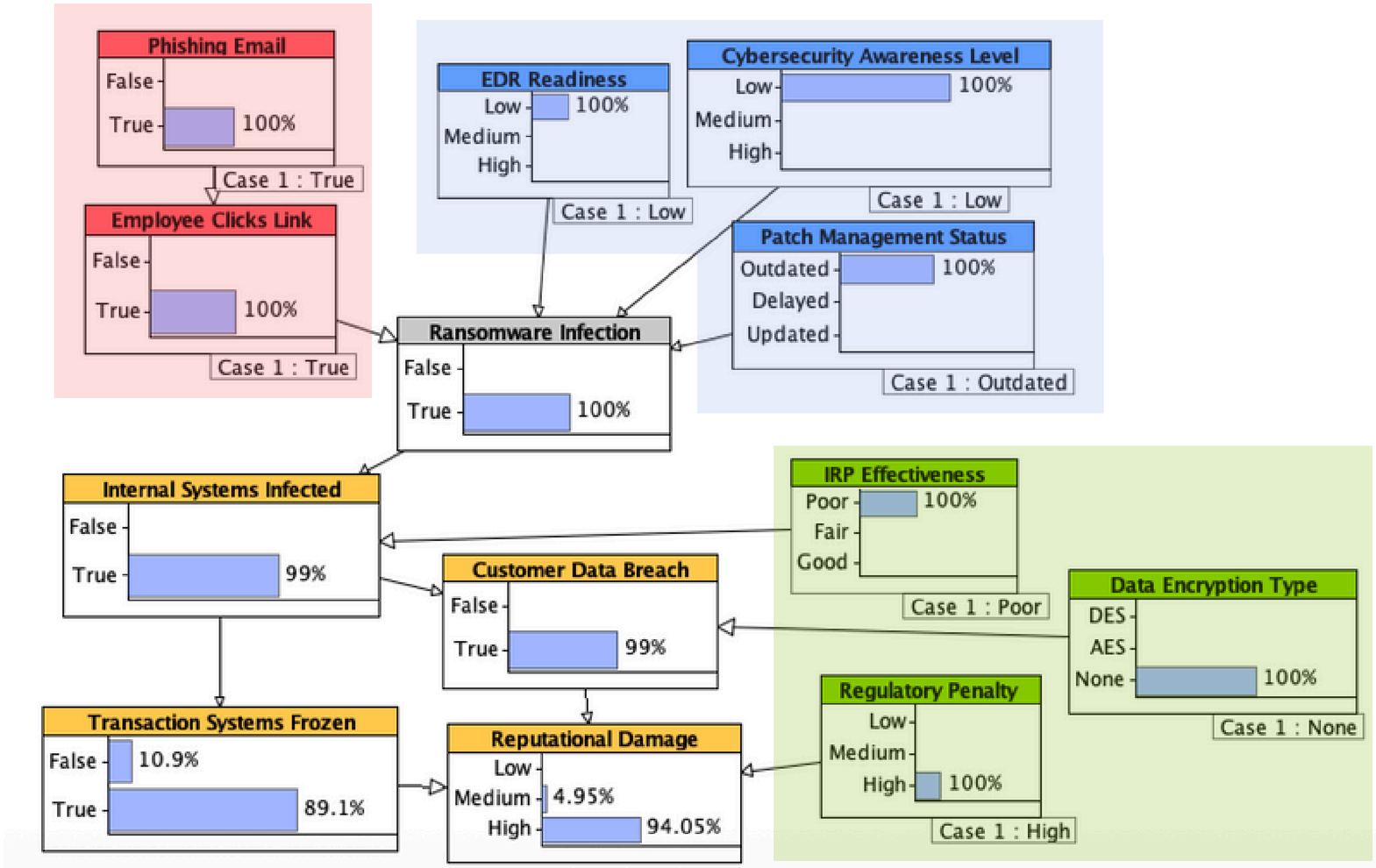
3. Controls' Impact

- Tested **low EDR**, **awareness**, and **patching** individually
- Each **weak control raised infection** and spread
- Showed where **strengthening controls** delivers the most **benefit**

4. Mitigations' Effectiveness

- Compared poor, fair, and good mitigations
- Better **IRP** and **encryption** sharply **reduced spread**
- **Lowered** data breach, downtime, and penalties
- Strong mitigations as critical **risk reducers**

1. Worst-Case Scenario



Motivations

Reveal the ceiling of loss:

- shows the institution's maximum financial, operational, and reputational exposure.

Surface hidden dependencies:

- illustrates how one human error can domino through every weak control.

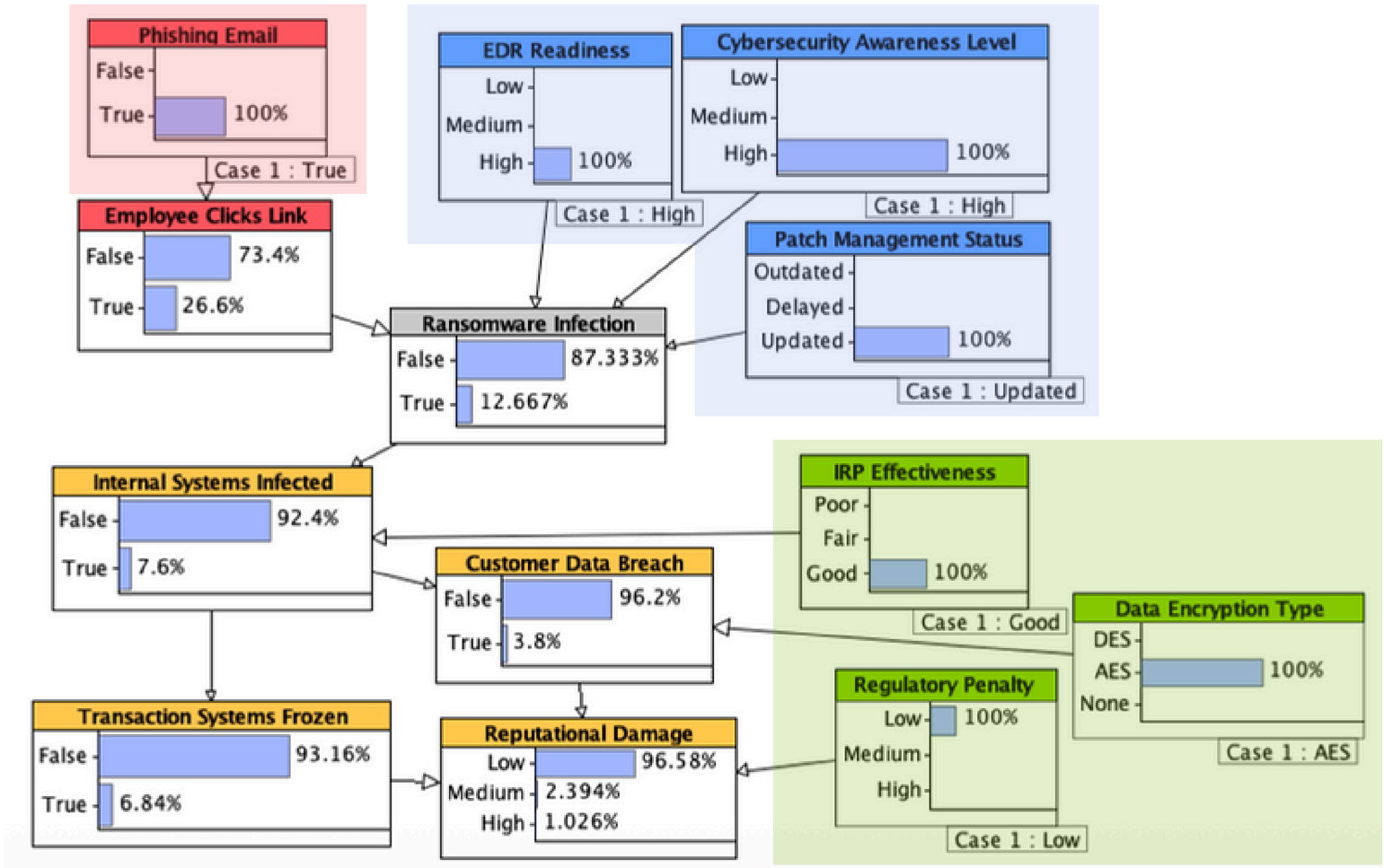
Stress-test risk appetite:

- helps leadership judge whether current reserves, insurance, and controls are sufficient for a black-swan event.

Create urgency:

- hard numbers on catastrophic impact make the case for minimum baseline controls.

2. Best-Case Scenario



Motivations

Proves the payoff of strong security:

- quantifies how mature controls drive orders-of-magnitude risk reduction.

Defines an aspirational target state:

- offers concrete metrics the organisation can measure progress against.

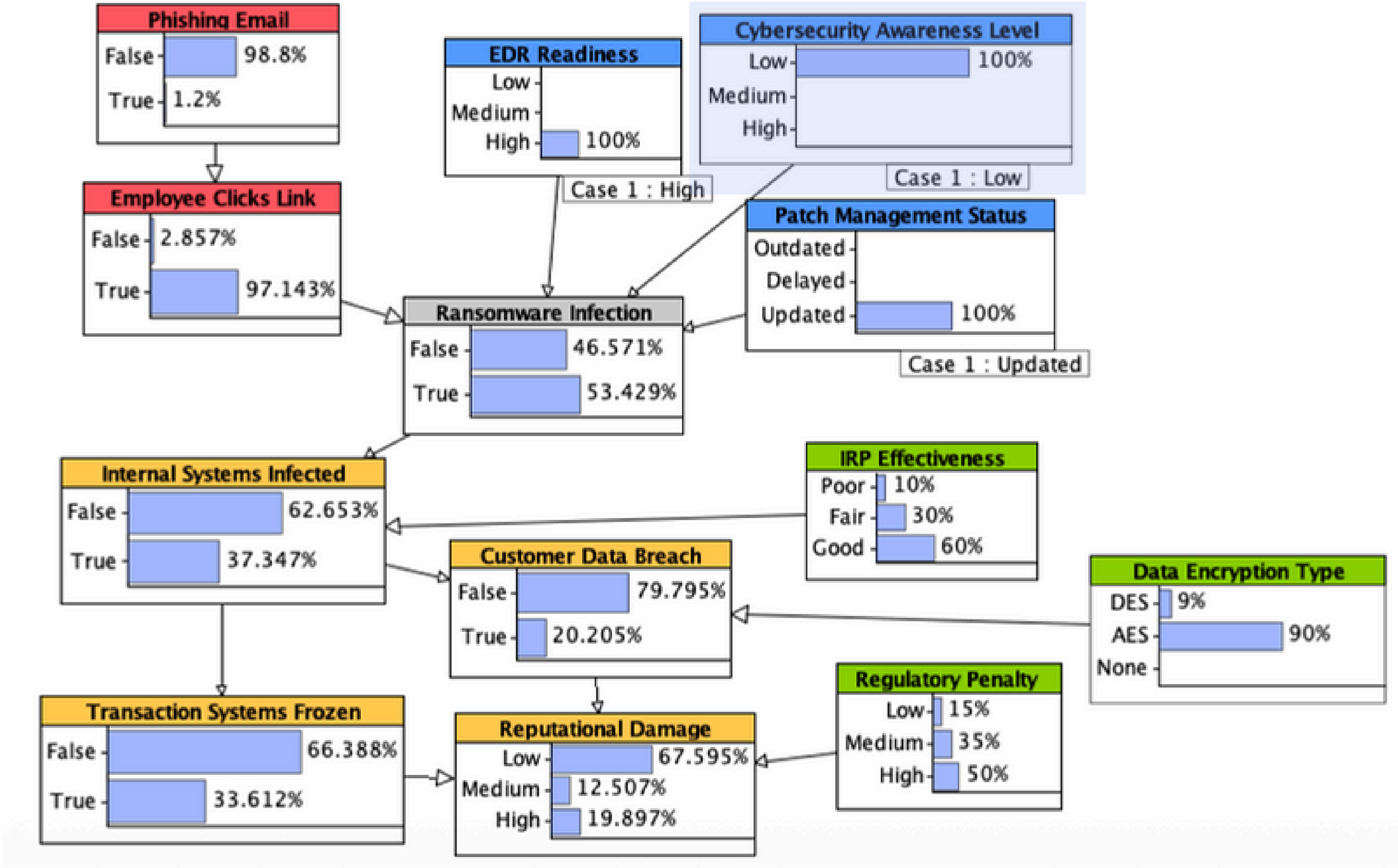
Frames ROI conversations:

- contrasting with worst-case provides a clear cost-versus-benefit narrative for security budgets.

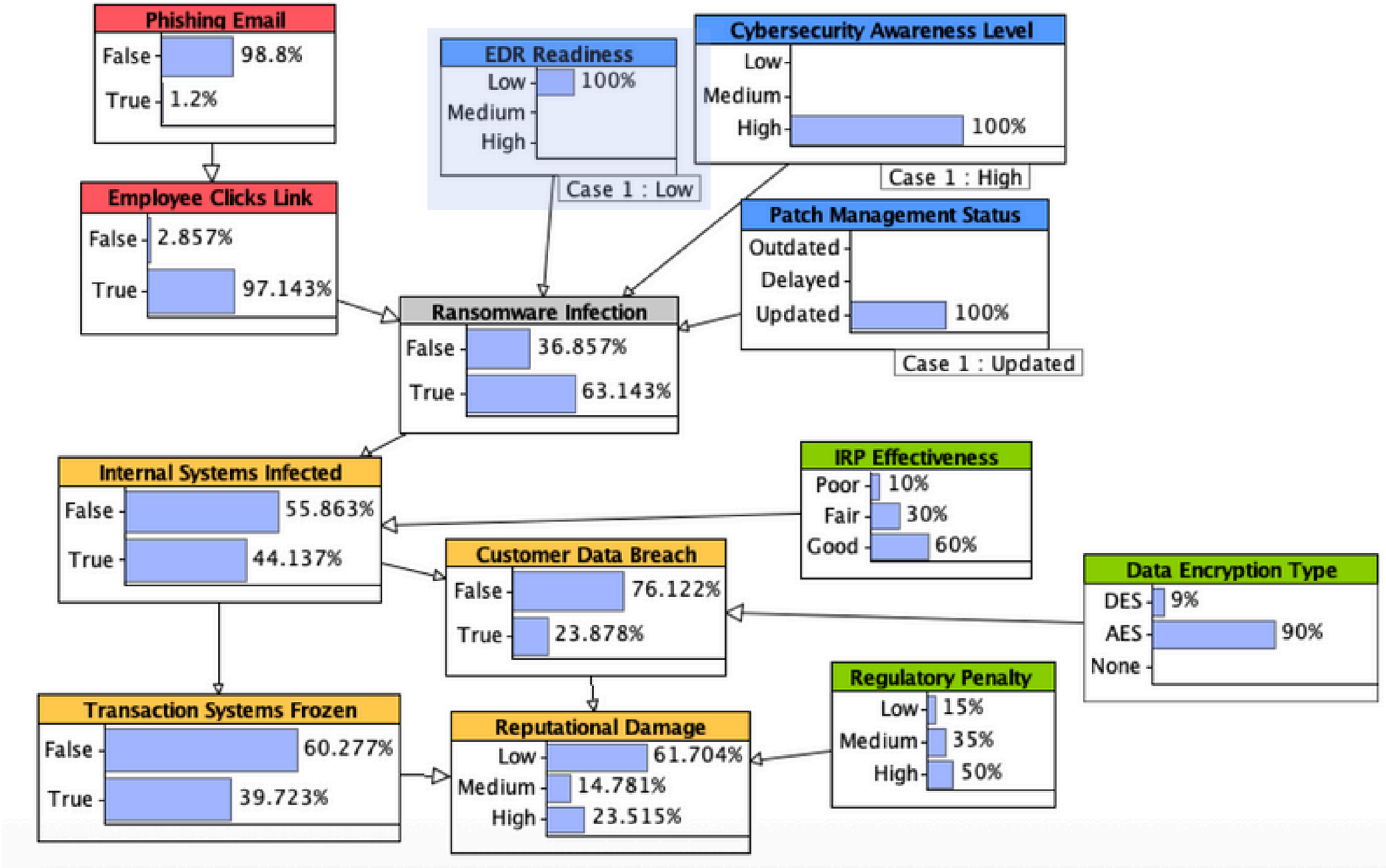
Builds confidence:

- shows stakeholders that ransomware risk can be contained, not just endured.

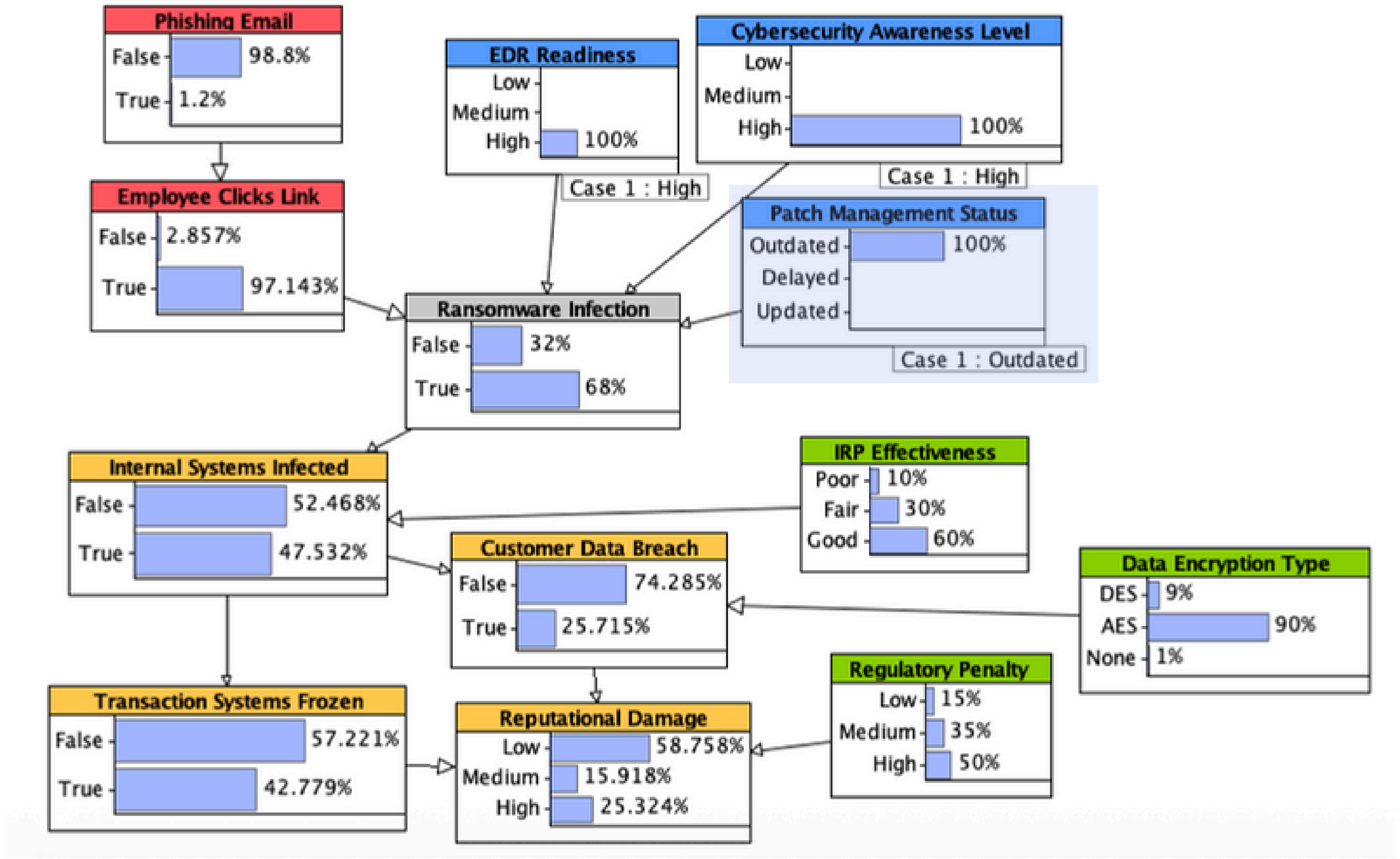
3. Controls' Impact – Low Cybersecurity Awareness



3. Controls' Impact – Low EDR Readiness



3. Controls' Impact – Outdated Patch Management



3. Controls' Impact: Motivations

Pinpoint high-leverage investments:

- identifies which single control changes (EDR, patching, awareness) cut risk the most.

Supports risk-based prioritisation:

- guides where limited resources deliver the greatest marginal benefit.

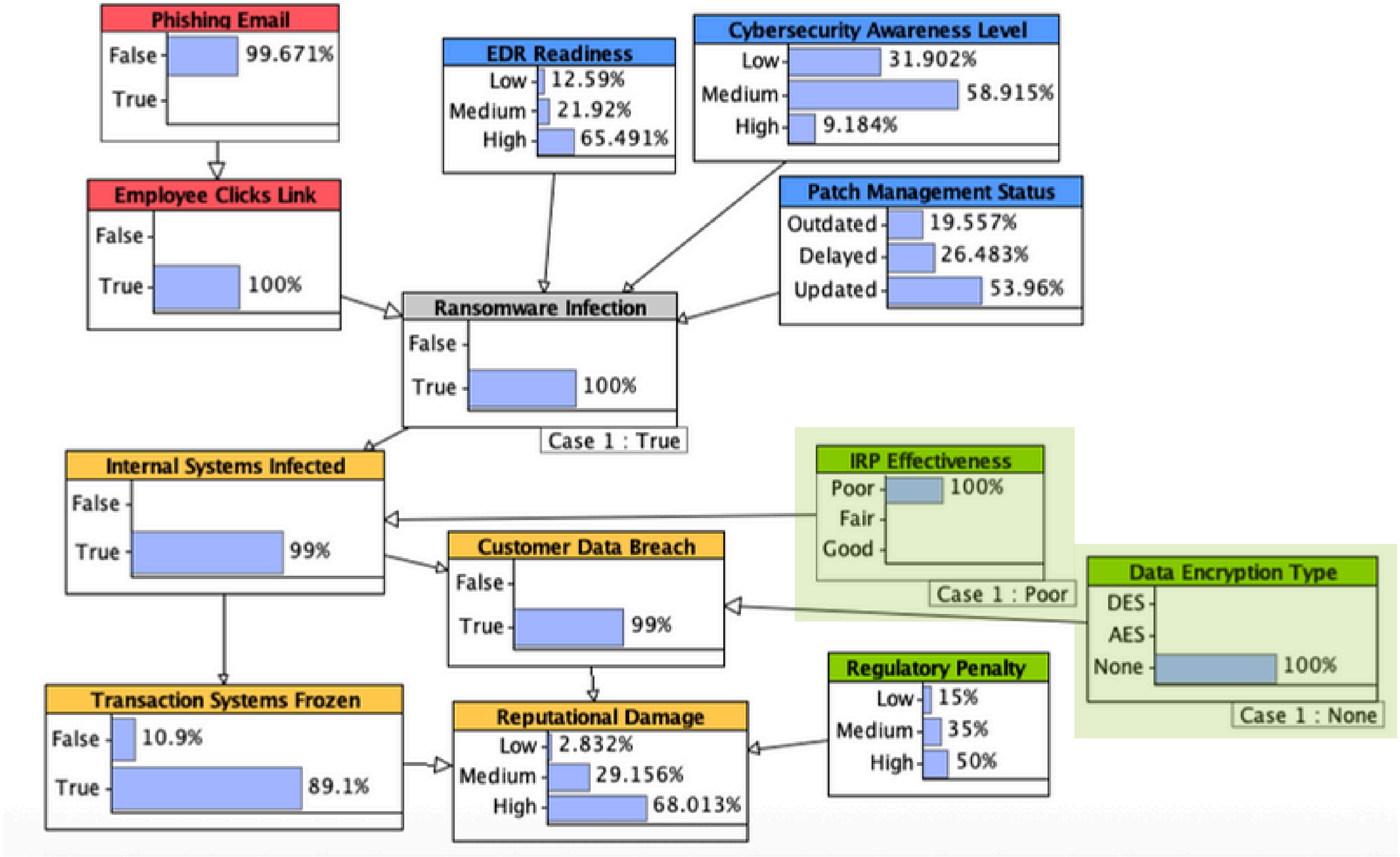
Highlights "weakest-link" dynamics:

- demonstrates that one neglected control can undermine an otherwise strong posture.

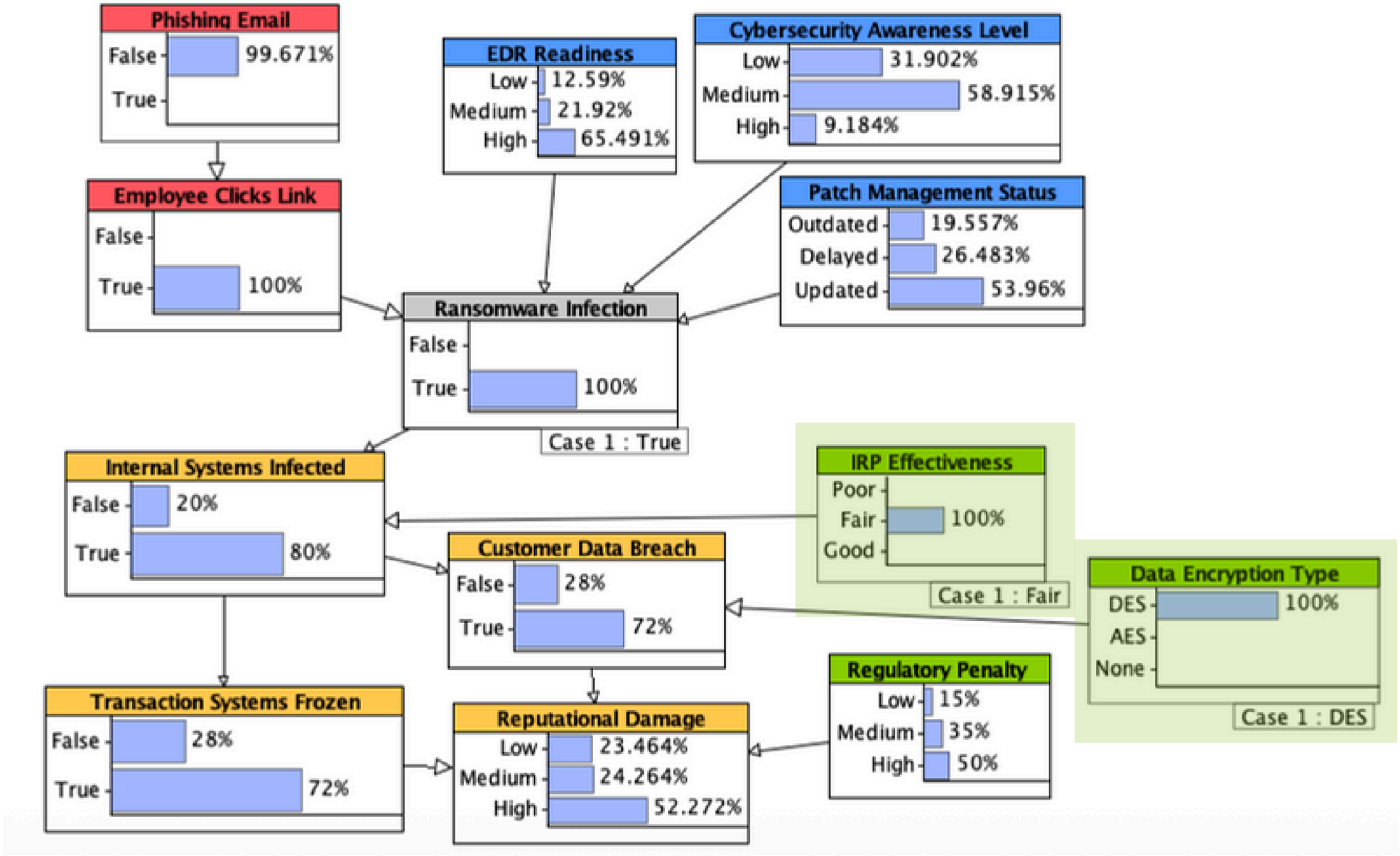
Informs a phased roadmap:

- lets us target quick wins first and plan longer-term improvements.

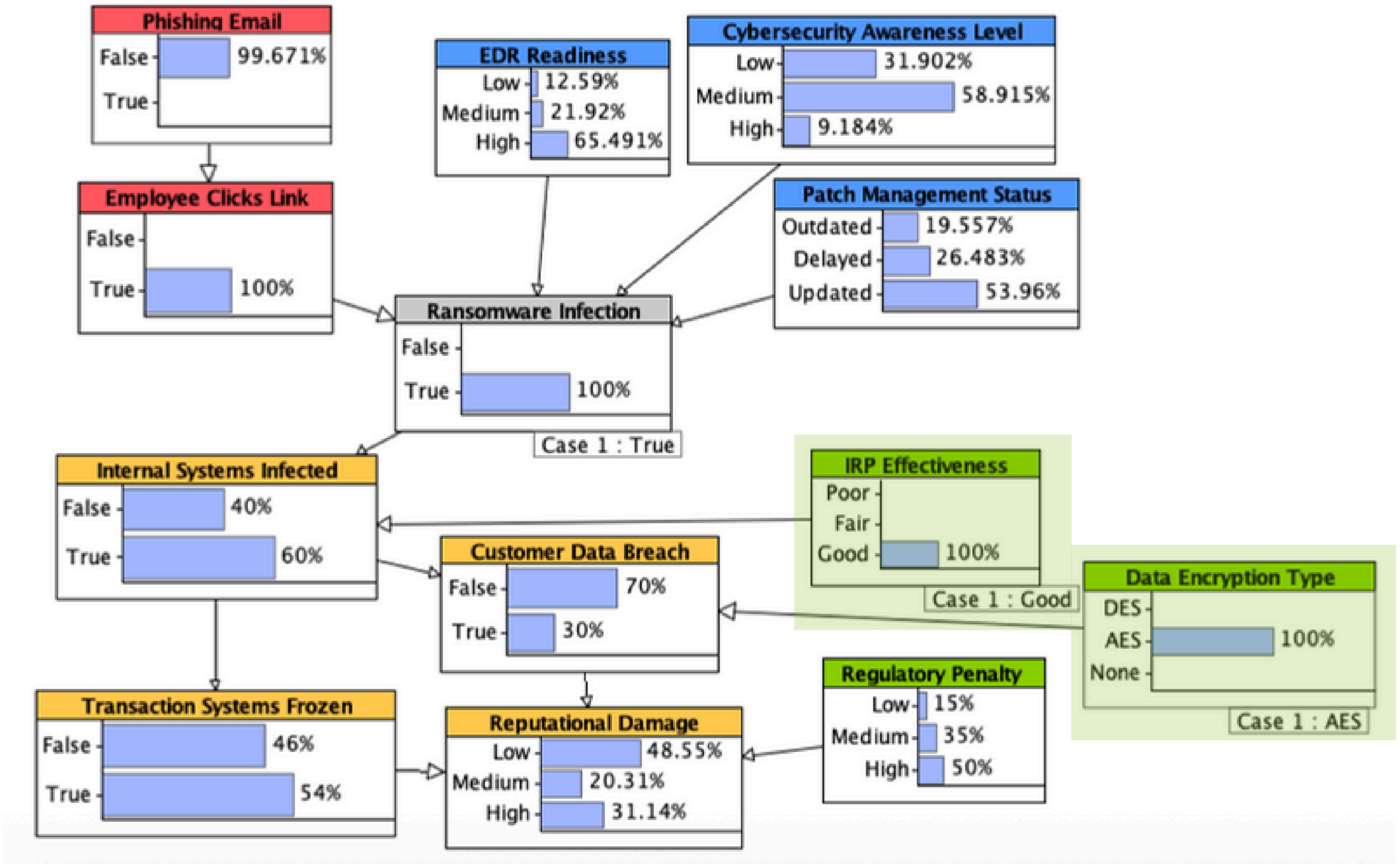
4. Mitigations' Effectiveness – Poor/No Mitigations



4. Mitigations' Effectiveness – Fair/DES Mitigations



4. Mitigations' Effectiveness – Good/AES Mitigations



4. Mitigations' Effectiveness: Motivations

Validate defence-in-depth:

- shows how incident response and encryption cap damage after prevention fails.

Quantify residual-risk reduction:

- proves good IRP slashes downtime and AES cuts penalties to near zero.

Justify reactive-capability spend:

- balances the security budget between proactive controls and recovery readiness.

Meet stakeholder expectations:

- aligns with regulator and cyber-insurance demands for demonstrable response strength.

THANK YOU.