Integrated Reasoning
Connecting the Dots: How We Make Decisions

Reasoning is the process by which we rationalize information, reduce uncertainty, and make decisions.

A decision can be deconstructed into the influences upon it – whether those are environmental factors, bias, or an estimation on the outcome. A decision involves uncertainty, where there is more than one potential future outcome. Oftentimes, prior to making a decision we reduce uncertainty by 'collecting all the facts' or 'influences' and evaluating the evidence holistically. It is therefore the role of the decision maker to reduce the uncertainty towards an outcome and make the appropriate decision.

Humans make thousands of decisions every single day – from fast and instinctual (should I bring an umbrella today?) to complex and time consuming (should I make this investment?).

For each decision, humans are integrating many disparate influences simultaneously – in one scenario, those influences can be simple and transparent. In another, they can be opaque, and the decision may be perceived as irrational.

The influences are there – they just need to be unpacked!

Software can reason and make decisions too. First, domain experts must understand and model the decision they are looking to automate. Through integrations, all relevant evidence can be collected and simultaneously evaluated in probabilistic mathematical models -- probabilities are simply the quantification of uncertainty.

In applying this to cybersecurity and the role of the intrusion analyst, we first must identify the decision we wish to model – specifically, does the observed activity represent a malicious and actionable threat within my organization or is the activity low risk, a false positive, or authorized.

Intrusion analysts are essentially investigators connecting dots in order to prove or disprove a hypothesis and ultimately determine the most likely explanation. Decisions are made through a series of steps:

1. Establish Domain Expertise
2. Identify Relevant Evidence
3. Collect Evidence
4. Construct a Hypothesis
5. Collect Additional Evidence
6. Make Decisions
7. Build and Explain Conclusions

Each step involves careful execution to have the best chance at telling the most likely story and to take the right action. Done well, these steps significantly reduce risk.

Establish Domain Expertise

As an analyst in information security, it's difficult to make good decisions without understanding the domain. The breadth and depth required to be effective can be daunting and leave many with the thought "where do I start?"

Because information security is so complex, entire courses covering general security, such as the CISSP, as well as deeper technical courses focusing on intrusion analysis and forensics are available.

Common areas of expertise required include:

- **Network Security -** Services such as web, DNS, application and database services and many others
- **Endpoint Security -** System authentication and authorization, process and file operations and network services
- **Application Security -** Including application vulnerabilities and hardening
- **Infrastructure Management -** Authentication services, software and patch management, Identity and Access Management (IAM), vulnerability identification and remediation, Configuration Management Database (CMDB), and network and system policy enforcement
- **Incident Detection and Prevention -** Network IDS/IPS, anti-malware, web filtering, Endpoint Detection and Response (EDR), and many others

General expertise in these areas, experience with the configuration of specific technologies and an understanding of the output from these systems is fundamental to leveraging this data for decision making.

To automate the process of establishing domain expertise, Respond Software offers the Respond Analyst, an XDR Engine. The Respond Analyst is a solution that monitors and triages security data to decide if observed activity requires incident response. The Respond Analyst is seeded with domain expertise, providing an understanding of how to triage network, endpoint security, and incident prevention and detection data. The Respond Analyst is unbiased, consistent and up-to-date on the latest threats.

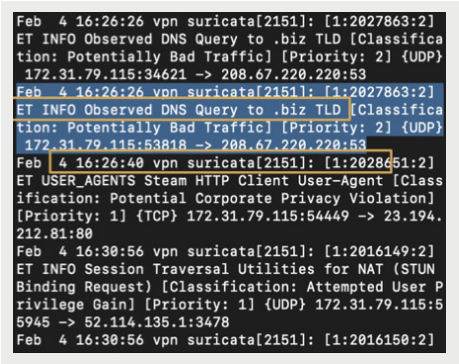
Identify Relevant Evidence

Evidence for intrusion analysis is provided in many forms. With an understanding of the domain also comes an understanding of what evidence is relevant to intrusion analysis and investigations. Unfortunately, there is a significant amount of data that is not useful mixed with important relevant evidence. It is the job of an intrusion analyst to determine what is useful, prepare that data for use and configure the appropriate means to retrieve the data when needed.

In general, data for intrusion analysis can be divided into several categories:

- **Detection telemetry -** signal through security alerts or events used to "turn our head" and may lead us to investigate an event further if we deem necessary, such as endpoint protection or network intrusion alerts.
- **Investigative context -** data used to further understand a situation, such as Dynamic Host Configuration Protocol (DHCP) for system identification, EDR to investigate detailed process activity or simply user statements provided to an analyst.
- **Risk and priority context -** data leveraged to understand system or account criticality or vulnerability, such as CMDB or vulnerability scan information.

Given the variety and magnitude of logs generated within an enterprise, not all initiate a security investigation. Some logs record authorized activity, others provide context to be used in an investigation, while a significant number of events can be discarded as low risk, informational or false positive. The Respond Analyst investigates a variety of log types including network IDS/IPS, endpoint protection, web filtering, DHCP, and endpoint detection and response. Within these events of interest, the Respond Analyst looks for specific and important characteristics to determine if additional investigation is required.



Example of a relevant Network IDS/IPS event

There are also several major challenges with evidence identification and use including:

- **Determining data usefulness -** the benefit of selective data collection based on need is important. Experience tells us exactly what types of logs are needed for intrusion analysis.
- **Data processing -** evidential data may be processed through various software to allow for automated monitoring and reporting, while data volume can present major challenges to complete processing.
- **Data storage and retrieval -** cheaper and faster data storage and retrieval options now exist and must be used as evidence is not very useful if it cannot be accessed and analyzed in some reasonable amount of time
- **Determining evidence credibility -** not all evidence has the same level of trust and intrusion analysts must consider the accuracy, completeness and risk of manipulation of data when weighing the usefulness of evidence with corroboration being key.

Compliance requirements may also mandate certain data collection and storage outside of what is typically needed for intrusion analysis, but in most cases what is needed for compliance is a subset of what is needed for intrusion analysis.

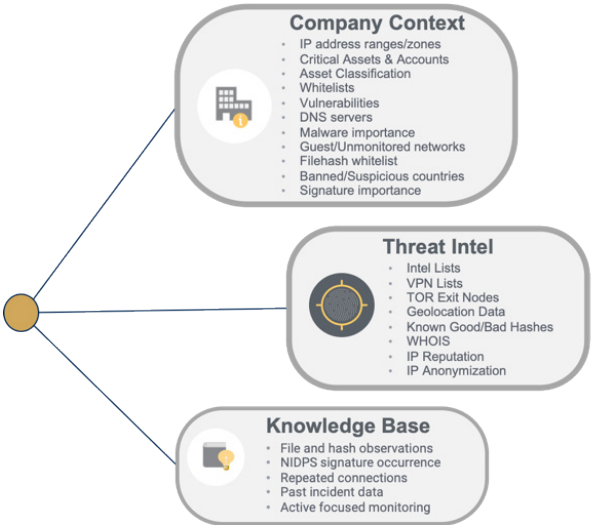
Collect Evidence

When an investigation is deemed necessary by an intrusion analyst, the analyst must collect the evidence surrounding the event. Hopefully the analyst understands what evidence will be relevant based on experience and has prepared by setting up systems in order to facilitate the collection process.

Planning is critical here as disparate systems with various ways of accessing data and with a multitude of data formats can greatly increase the time required to collect and analyze data. This has led many security operations teams to look for a “single pane of glass” solution, though in reality none exist. Even if a security operations team is successful in consolidating alerts into a single repository or case management solution, there is almost always evidence in other systems that must be accessed during an investigation. An example would include the need to access an EDR product that stores process and file operations.

Generally, initial evidence collection centers around accessing evidence for various object types, including:

- **Internal Assets -** context that describes criticality, vulnerability, and activity
- **External Systems and Domains -** context that describes the ownership, geolocation, and threat reputation
- **Accounts -** context that describes criticality and activity
- **Files -** characteristics of a file such as type, purpose and activity
- **Processes -** characteristics of a process that explain purpose and activity



The Respond Analyst collects relevant information including company context, threat intelligence and it keeps a knowledge base to assist in making escalation decisions.

This data must then be aggregated and relationships drawn between the various objects in order to construct a hypothesis to prove or disprove it.

On each event of interest, the Respond Analyst extracts the entities (accounts, IP addresses, host-names, signatures, device actions, ports, etc) and attributes the event with evidence. The Respond Analyst collects relevant information, including contextual information about the company, the criticality, vulnerability, and classification of assets, external threat intelligence, and maintains a derived knowledge base of patterns and observations. Many of the patterns are used to make decisions that go 180 days in the past, therefore accurate system identification is important to attribute the correct context and historical behaviors to the true system. The Respond Analyst uses a proprietary system identification service to determine the true systems and accounts associated with events of interest over periods of time. For example, in DHCP enabled environments, a single system can have a new IP address each time it renews a lease or rejoins a network.

Construct a Hypothesis

Once the initial evidence is collected, the analyst needs to identify and construct a hypothesis. The **MITRE ATT&CK®** framework can serve as a useful starting point for the formulation of a hypothesis. The framework maps observed attack techniques to an attack tactic, offering an explanation of what the adversary is trying to do.

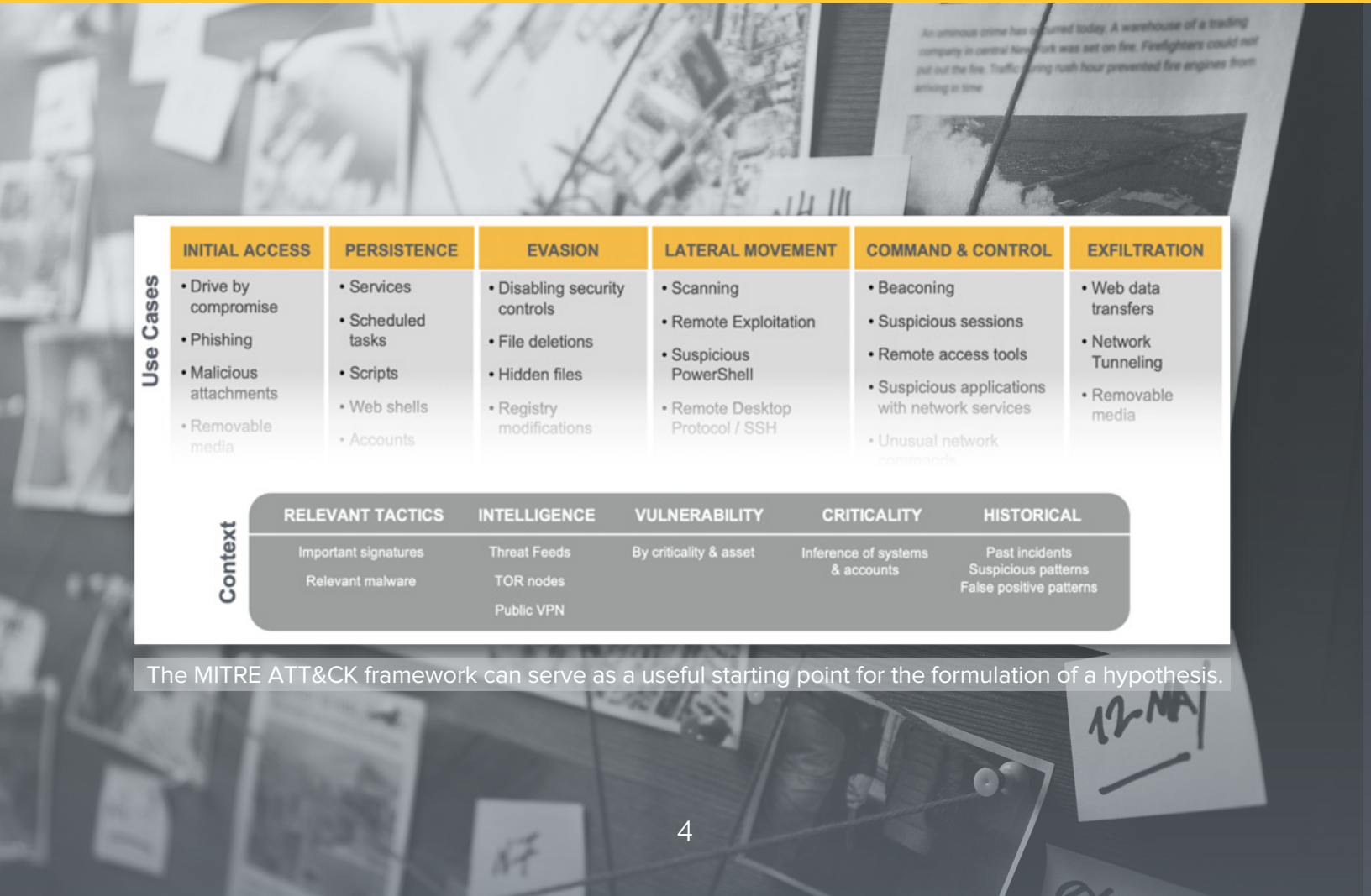
Typically, the analyst should be able to attribute an attack tactic based on the detection telemetry and supporting investigative context. For example, if a network intrusion alert is in-bound from an external system, the tactic could be initial access, if the alert is between two internal systems, the tactic could be lateral movement or discovery, or if the alert is outbound the tactic could be command and control or exfiltration.

For alerts generated by endpoint protection platforms and endpoint detection and response, the action taken by the endpoint agent to prevent a malicious file from executing or blocking a suspicious process can help differentiate between the early stage tactic of Initial Access and later stage tactics of Execution, Persistence, or Collection.

Applying relevant investigative context reduces uncertainty to prove or disprove the hypothesis. For example, vulnerability data can reduce uncertainty if the attack was both relevant and successful. Oftentimes, detection telemetry alerts on normal administrative activity investigative context of the type and function of the accounts and systems helps differentiate between normal and malicious behavior.

Depending both on the event type and context gathered in the previous steps, the Respond Analyst decides which use case to investigate, while attempting to prove or disprove the hypothesis that the activity is malicious and actionable. Example use cases could be an Initial access and execution on a system, lateral movement, or command and control.

[The MITRE ATT&CK® Framework]



The MITRE ATT&CK framework can serve as a useful starting point for the formulation of a hypothesis.

Collect Additional Evidence

Gathering more evidence about a given attack may be necessary, but will require more time and has a cost for both humans and systems.

For example, a suspicious outbound network connection, alerted through either the network IPS or a web filtering system, may cause an analyst to determine, based on the context of the evidence, that the attack is outbound command and control malware. But the analyst still has to prove or disprove this hypothesis. The analyst may then determine that more questions need to be asked.

Does the domain look suspicious?

- Registration information gleaned from WHOIS lookups will return domain ownership and registration dates. If the owner is an authorized entity whom you do business with, perhaps this alert is a false positive or they themselves have been compromised. Oftentimes, adversaries will anonymize their ownership information, however a recent registration date increases the likelihood the domain is malicious.
- In addition, threat intelligence solutions like VirusTotal provide insight if the larger security community believes the domain is malicious. However, external sources of intelligence need to be taken with caution and validated - which takes time to do properly.

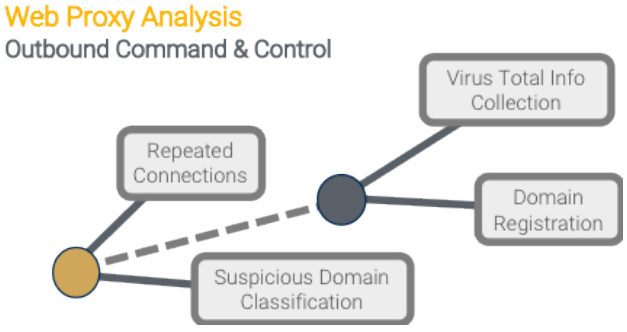
Is this a repeated connection? Is there a suspicious pattern?

- Command and Control traffic often repeats at robotic intervals as the malware checks-in to the adversary-controlled server. The analyst will have to pivot and query a data repository to understand the nature of the communications between these two systems.
- In addition, the analyst must understand the scope of the incident. If many other internal systems are communicating to this domain, perhaps the activity is authorized and the alert is a false positive, or the analyst may have found a more pervasive intrusion.

Does the internal asset show signs of compromise?

- For the system to be beaconing to a malicious domain, there is likely to be malware on the system - typically in the form of a malicious file or process. Pivoting to the endpoint, the analyst should first check for known malware found on the system. Subsequently, using the endpoint detection and response solution to evaluate the process data to understand if any anomalous behavior has occurred.

The Respond Analyst has the ability to ask additional questions based on the evidence that it has previously collected similar to how a human analyst does, but in an automated, scalable and much faster fashion. In the example below, the Respond Analyst is mining high volume web filtering data (which can reach upwards of 50 million events per day) to identify suspicious patterns of repeated connections resembling command and control. If the criteria is met, the Respond Analyst executes additional queries to determine if the domain has suspicious or recent registration or a malicious threat reputation.



In this outbound command & control example, the Respond Analyst mines high volumes of web filtering data to identify suspicious patterns such as repeated connections.

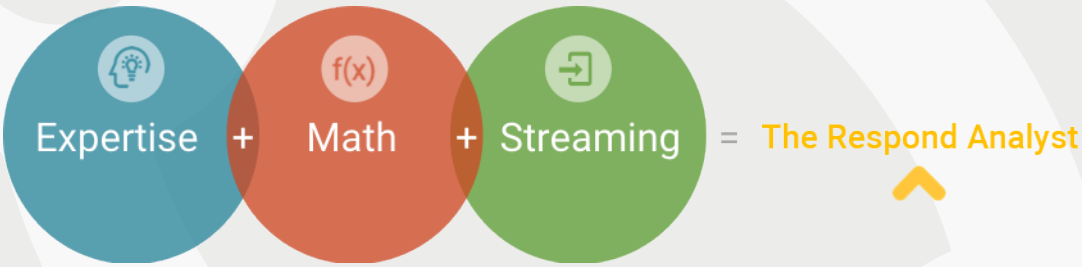
The above questions are examples and only represent part of the investigation an analyst must perform. As mentioned earlier, performing a deeper analysis is not required for each investigation. Analysts should be able to discard a large number of false positives with preliminary investigative context -- freeing up much needed time to collect and evaluate additional evidence.

Make Decisions

After all of the evidence is gathered and the right questions are asked, a decision needs to be made. If the analyst determines the event is normal activity, it can safely be ignored and discarded. However, if the determination is that the event or series of events, are malicious and actionable, the incident will scoped and escalated to build a case which will be discussed further in the next section, **Build and Explain Conclusions**.

The Respond Analyst uses decision automation derived from built-in security expertise and probabilistic mathematics to determine the liklihood of streamed security events being mailicious and actionable.

The Respond Analyst decides if the activity requires incident response or if it can be safely ignored – saving security teams time by not chasing false positives. The Respond Analyst employs a variety of modeling approaches given the evidence attributed to the case (in the previous steps) to prove or disprove the hypothesis, including probabilistic graphical models, logistic regression, and decision trees. The Respond Analyst picks the appropriate modeling approach given the upstream use cases and event types.



The Respond Analyst builds-in years of SOC experience and expertise, leverages math and probability to make decisions about the requirement to escalate or suppress millions of streaming events.

Build and Explain Conclusions

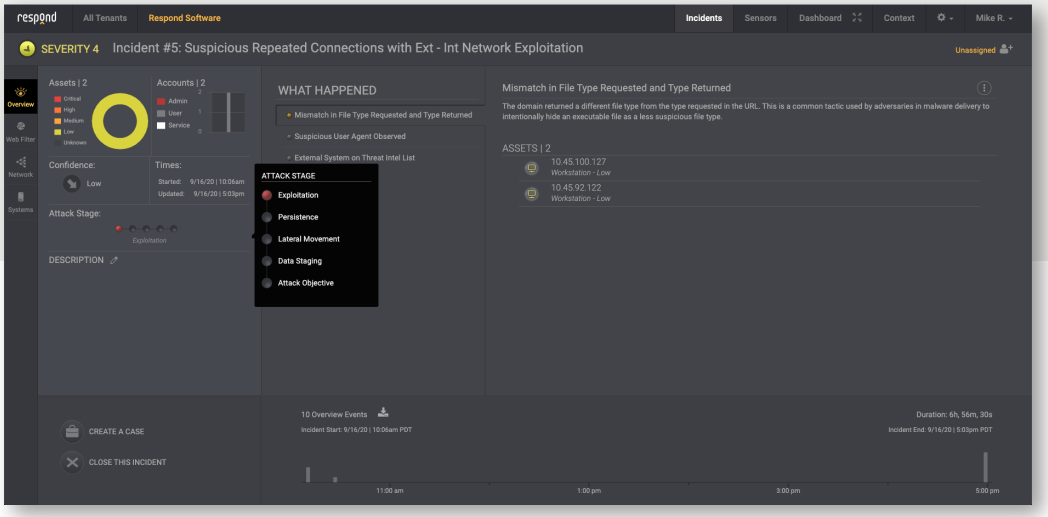
After the analyst determines there is enough information to prove or disprove the hypothesis, a case can then be built. Each piece of evidence will be evaluated simultaneously and carry a unique influence on the analyst's decision.

In many organizations, the person who performs the investigation (the analyst) is not the same as the person who performs the remediation (the incident responder). Therefore, the analyst needs to document their findings in a concise, logical, and easy to understand briefing -- the case.

Using the outbound malware example - if the source of the malware has a vulnerability, that might increase the chance that there is something malicious and actionable to remedy. However, the incident might still be escalated even if the source does not have a vulnerability because it is outbound malware.

There are several reasons for escalating this including the criticality of the asset involved or it might be a zero-day attack.

The Respond Analyst scopes incidents to include all related systems and activity for the duration of the attack. The incident may span a few seconds or many days. Next, the Respond Analyst prioritizes the incident, factoring in the scope, asset criticality, attack stage, and confidence in the escalation. All of the supporting evidence and context is succinctly summarized and explained in the user interface. Next, the Respond Analyst notifies the user about the new incident – via email, text, or phone call. Subsequently, users can push incidents into a case management or SOAR platform to track the remediation of the incident.



The Respond Analyst exposes relevant information about an incident that requires remediation.

The expectation that human analysts have the ability and capacity to monitor, triage and potentially escalate the multitude of events that are generated in today's SOC is not reasonable or sustainable. SOC teams need an automated solution that removes them from the tedious task of weeding through endless false positives, and instead enables them to investigate real incidents that require remediation to keep their organization safe and secure. The Respond Analyst's integrated reasoning capability enables it to consider all of the sensors, company context, threat intelligence and vulnerability information required to build an incident, scope and escalate it for remediation. The Respond Analyst is the answer to removing human analysts from ineffectually and endlessly staring at a console with mediocre results and instead empowers them to become threat hunters. The Respond Analyst is designed to automate the decision-making steps necessary to protect today's digitally driven business.