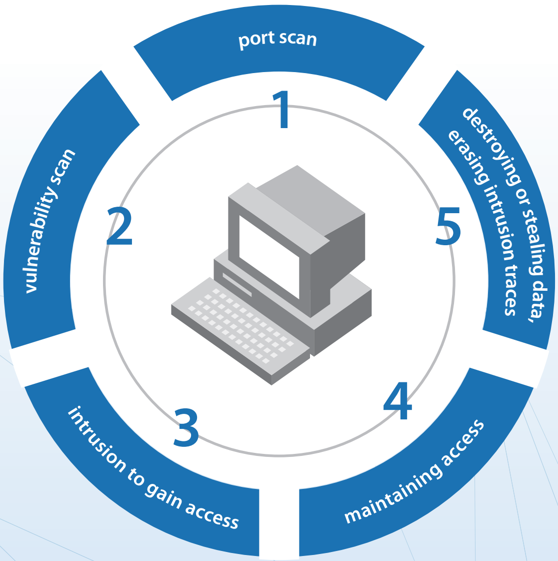
# Hillstone iNGFW White Paper: Advanced Security Correlation Analytics

Information security incidents may happen due to a lack of data availability, integrity or confidentiality. Many security vendors have conducted considerable research to determine how best to protect data. Many current security products and technologies—such as firewalls, intrusion detection/prevention systems, strong authentication access control mechanisms, virtual private networks, public key infrastructure—arose from these efforts. Businesses around the world deploy security products and technologies to identify, and defend themselves against, information security breaches.

It is imperative to gather and associate network behavioral information by using different security products and technologies. This information is required to enable analysts and administrators to accurately—and in a timely manner—identify, classify and alleviate network attacks, policy vulnerabilities and other irregularities.

The Hillstone Security Correlation Analytics Platform analyzes centrally stored threat event data acquired from the various threat detection engines of the Hillstone intelligent security architecture, enabling it to continuously accumulate disparate aspects of security incidents, and then—more importantly—to correlate and link these nuggets of information to expose conspiracies or schemes behind apparently unrelated attacks or incidents.

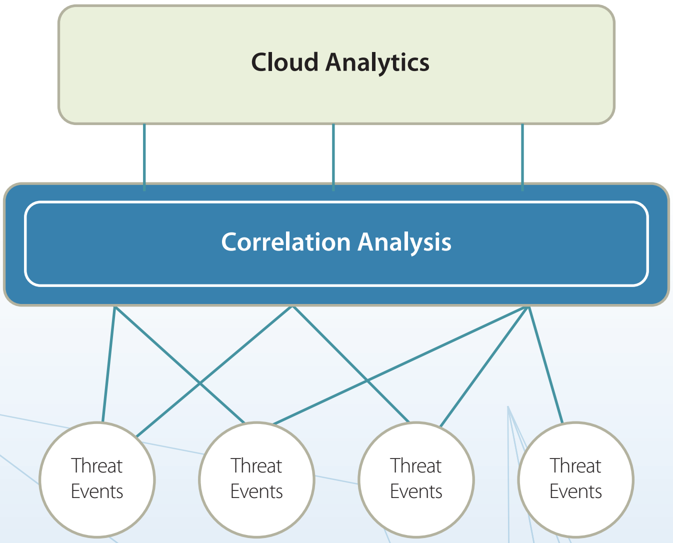
In brief, current threat scenarios and Hillstone's correlation analysis methodology can be summarized as follows:



1. Today's network attacks include multiple steps: port scan, vulnerability scan, intrusion to gain access, maintaining access, destroying or stealing data, and then erasing all traces of the intrusion. Each step is found in the evidence of attacks, necessitating continuous detection of malicious code to cover the entire threat lifecycle. Different stages and attack sources must be associated to narrow the scope of detection and expose the real attackers.
2. Evidence of network attacks are scattered across different events, logs, alarms and other data stores (such as packet captures and session statistics).
3. Single-source attack detection may lead to a higher false-positive rate.
4. It is increasingly necessary to leverage evidence from several disparate information sources to gain acceptable detection accuracy.
5. It is imperative to use efficient and intelligent correlation algorithms to detect and identify the footprints of sophisticated attacks in large and complex network environments.

## Correlation Analysis Process

The analysis process constructs a linkage between low-level information models and high-level analysis models. The low-level information models format data from diverse and independent events. The correlation engine employs a variety of consolidating and transforming algorithms to conduct state transition analysis on this seemingly disparate data. Based on the input from the low-level information models, in conjunction with analytical processing and association techniques, an artificial threat identification is made by the high-level analytical models.



Correlation engines generally use many different algorithms to detect advanced threats and anomalies. The common ingredient to these techniques is the use of multi-point, rather than single-point, detection to reduce the rate of false alarms.

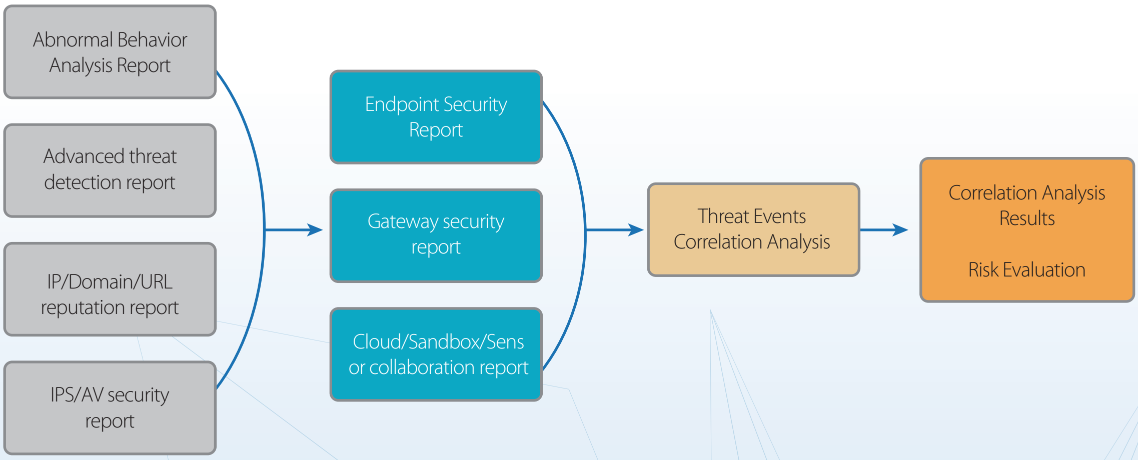
The advanced algorithms in Hillstone's intelligent security architecture Correlation Analytics Engine include deterministic rule-based, as well as probabilistic cause-based, correlations. Deterministic correlation is a traditional method relying on the creation of rules, which means that it uses pre-defined rules to analyze data identified only at a later time. Its correlation capacity therefore depends on the depth and breadth of the rule sets. Considerable expert knowledge is required to define appropriate rules. Deterministic correlation requires a large amount of conditional logic checks which does not scale well for large deployments. The innate statelessness of this method renders it unable to correlate dynamic data, therefore making it difficult to adapt to contaminated and noisy scenarios.

Probabilistic correlation, on the other hand, is an advanced method based on causal analysis and data mining, and is much better able to adapt to dynamic and noisy scenarios. Following careful study, Hillstone's intelligent security architecture adopted a Markov Process to analyze the probable and causal associations in time domains. For probabilistic correlations in spatial domains, Hillstone selected a Bayesian Network (a typical probabilistic graphical statistical model) to express random variable sets and the dependencies among them, using directed acyclic graphs to describe causal relationships.

## Converged Correlation Technology

Hillstone's intelligent security architecture introduces a systematic process to correlate context-based multi-dimensional vectors to determine a computing risk credibility for the system. The results discount simple reasoning, and instead leverages long-term considerations to establish an innovative converged correlation technology. This methodology drives decisions founded in the contexts of multiple dimensions (including time, space, and threat distribution), reduces the time and effort spent by analysts to conduct inspections or investigations, and maximizes the accuracy and reliability of identifying cyber-attacks and threats.

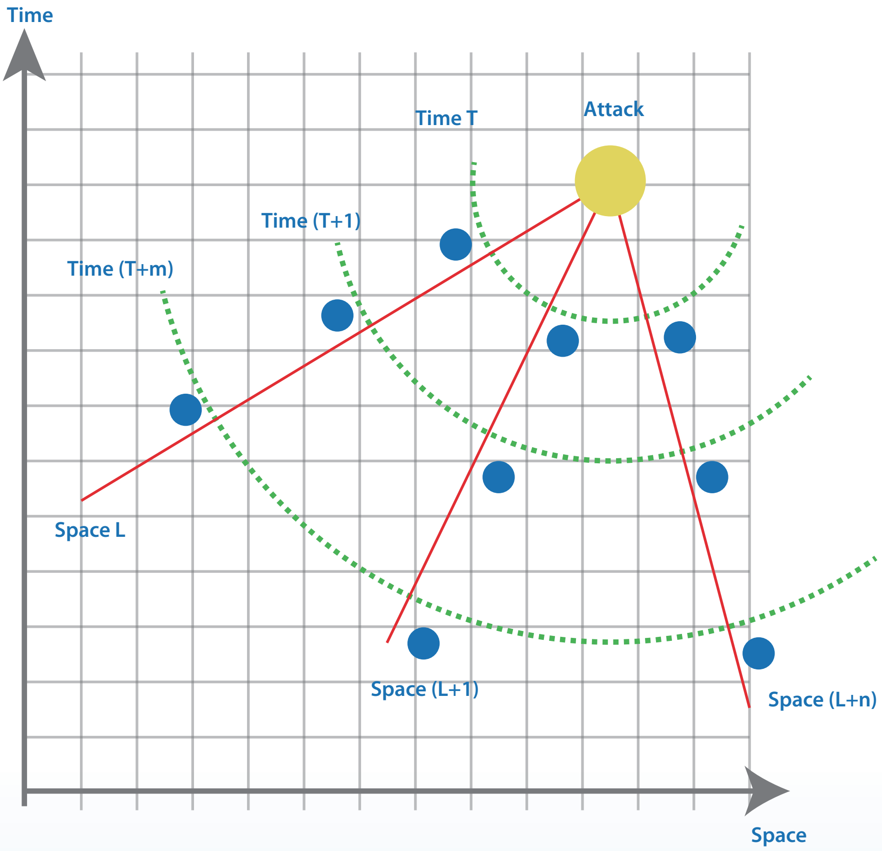
As shown below, gateway security reports, terminal behavior reports and cloud/sandbox/sensor coordination reports are correlated together to determine a final risk assessment.


Abnormal Behavior Analysis Report
Advanced threat detection report
IP/Domain/URL reputation report
IPS/AV security report
Endpoint Security Report
Gateway security report
Cloud/Sandbox/Sensor collaboration report
Threat Events Correlation Analysis
Correlation Analysis Results
Risk Evaluation

- Gateway security reports are generated by Hillstone's intelligent next-generation firewalls, and cover network behavioral levels (including abnormal network behavior reports, advanced network threat reports, domain name and URL credit reports), and intrusion prevention/anti-virus security reports.

- Endpoint behavioral reports are drawn from customer-owned servers and hosts, and used as directly associated data sources to cover terminal behavior.
- To achieve comprehensive coverage, Hillstone additionally provides cloud security services to extend correlations to cloud collaboration, cloud sandbox, enterprise security sensors and other external data sources. Cloud collaboration can provide a globally correlated high-level view, and cloud sandbox and enterprise security sensors can provide additional threat identification information to improve the overall accuracy of correlations.

The converged correlation of Hillstone's comprehensive security architecture can also be described with a time-space based network model, as shown below.



As seen in the graph above, evidence of cyber-attacks is scattered in observed events across time and space. It is more advantageous to correlate time-based reports with space-based reports. Each space-based observation can be independently correlated with all time-based reports; and similarly the time-based observations can be further correlated with space-based models.

## Conclusion

By mining comprehensive evidence from disparate reports, the grid-based converged correlation model of Hillstone's intelligent security architecture results in more accurate detections while also delivering fewer false alarms.