Lisa J. Sotto, Chair
DHS Data Privacy and Integrity Advisory Committee

February 17, 2016

The Honorable Jeh Charles Johnson
Secretary of the U.S. Department of Homeland Security
U.S. Department of Homeland Security
Washington, DC 20528

Ms. Karen L. Neuman
Chief Privacy Officer
U.S. Department of Homeland Security
Washington, DC 20528

Re: DHS Data Privacy and Integrity Advisory Committee: Algorithmic Analytics and Privacy Protection (Report 2016-01)

Dear Secretary Johnson and Ms. Neuman:

It is my pleasure to convey to you the final report that sets forth recommendations for DHS to consider how best to address privacy protection in the conduct of “behavioral analytics” in cybersecurity programs. The Data Privacy and Integrity Advisory Committee unanimously voted to adopt the recommendations at the public meeting on February 8, 2016.

DHS’ request letters, included as Appendices A and B in the report, asked the Committee to address “how best to protect privacy while achieving the cybersecurity goals of such analysis across the various stages of the information lifecycle … and what should be included in any human review of indicators or outputs.” The report is structured into the following three sections, each of which contains advice and recommendations for DHS to consider:

- general considerations regarding the scope of the DHS inquiry;
- key considerations that impact algorithmic analytics; and
- questions to address for major categories of information handling.

This report is the result of extensive effort by Committee members working closely with the National Protection and Programs Directorate (NPPD) to research the relevant topic. We are grateful for NPPD’s cooperation in providing the necessary background and input, and for making officials with direct knowledge and expertise on the matter available to us.

We hope you will agree with these recommendations as the Department continues to refine its cybersecurity strategy.

Please do not hesitate to contact me if you have any questions regarding the report.

Sincerely,

Lisa J. Sotto

Enclosure: Report 2016-01 of the DHS Data Privacy and Integrity Advisory Committee On Algorithmic Analytics and Privacy Protection

cc: Members of the Data Privacy and Integrity Advisory Committee

Report 2016-01 of the DHS Data Privacy and Integrity Advisory Committee On Algorithmic Analytics and Privacy Protection

February 8, 2016

Report 2016-01 of the DHS Data Privacy and Integrity Advisory Committee On Algorithmic Analytics and Privacy Protection

Summary of Request and Context

The Department of Homeland Security (DHS) Data Privacy and Integrity Advisory Committee (DPIAC) established the Cybersecurity Subcommittee to explore issues where cybersecurity and privacy issues intersect in the conduct of DHS programs. Based on discussions of issues with the Subcommittee, the DPIAC advises the DHS Privacy Office about findings and recommendations on relevant topics.

On January 16**,** 2015, the DHS Chief Privacy Officer requested that the DPIAC provide guidance on how best to address privacy protection in the conduct of "behavioral analytics" in cybersecurity programs. DHS' request letters, included as Appendix A in the report, asks that DPIAC address "how best to protect privacy while achieving the cybersecurity goals of such analysis across the various stages of the information lifecycle … and what should be included in any human review of indicators or outputs".

DHS requested that the DPIAC base its guidance on fact finding conducted by the Cybersecurity Subcommittee. The Subcommittee reviewed the issues identified by DHS, and discussed conclusions with the DPIAC to inform potential findings and recommendations. This report represents the results of that review, and addresses a number of considerations related to the DHS Privacy Office's request.

In addition, the DPIAC notes that the recently enacted Cybersecurity Information Sharing Act (CISA) includes a provision that authorizes the continued assessment of innovative cybersecurity technologies, such as the programs being addressed in this report. Specifically, the statute states:

> "(c) Activities.—In carrying out subsection (b), the Secretary—
>
> …
>
> "(4) shall regularly assess through operational test and evaluation in real world or simulated environments available advanced protective technologies to improve detection and prevention capabilities, including commercial and noncommercial technologies and detection technologies beyond signature-based detection, and acquire, test, and deploy such technologies when appropriate;
>
> "(5) shall establish a pilot through which the Secretary may acquire, test, and deploy, as rapidly as possible, technologies described in paragraph (4);

Background

The DPIAC was requested to review the "behavioral analytics analysis" process and to consider how best to protect privacy while achieving the cybersecurity goals of such analysis across the various stages of the information lifecycle. DPIAC was tasked with addressing how to protect privacy while achieving the cybersecurity goals of the program, as detailed in the attached letter. This report represents the results of that review.

The report addresses a number of considerations related to the DHS Privacy Office's request. In addition to the specific findings and recommendations that follow, DPIAC has an overarching observation regarding terminology for the program. We recommend that DHS select an alternative name for the program, and suggest using either "algorithmic analytics" or "network analytics"; for the purpose of this report, we use the term "algorithmic analytics", but DHS may want to select another appropriate alternative.

We make this recommendation because DHS's term of "behavioral analytics" carries an inescapable connotation that individuals' end user or other personal actions (*i.e.*, "behavior") are to be targeted for specific review, which if accurate would understandably implicate a variety of particular privacy concerns. By contrast, we understand that the program consists mainly of automated (*i.e.*, by algorithms) analysis of government network traffic data, collected using automated means. Assuming that this understanding is correct, and end user-based targeting and surveillance is not the focus here, the BA terminology creates an unnecessary distraction and invites misinterpretation of the potential privacy impact.

In addition, the DPIAC generally has addressed in prior reports a number of the issues discussed here. While those reports are not specific to the use of algorithmic analytics, they may provide useful background and inform many of the recommendations below. Prior reports that may be of interest with respect to the topic of this paper include:

- Report 2014-02 Privacy Recommendations Regarding Auditing and Oversight of the DHS Data Framework;
- Report 2012-01 Privacy and Cybersecurity Pilots;
- Report 2011-01 Privacy Policy and Technology Recommendations for a Federated Information-Sharing System;
- Report 2010-01 The Elements of Effective Redress Programs; and
- Report 2006-01 Framework for Privacy Analysis of Programs, Technologies, and applications.[1]

[1] All DPIAC Reports are available at www.dhs.gov/privacy-advisory-committee.

This report is structured in three sections, each of which contains advice and recommendations for DHS to consider:

- General considerations regarding the scope of the DHS inquiry.
- Key considerations that impact algorithmic analytics.
- Questions to address for major categories of information handling.

1) General considerations regarding the scope of the DHS inquiry

a. Definition: What is Algorithmic Analytics (AA)?

i. AA can help identify cybersecurity threats, so as not to tip off signature-based or flow-based detective controls. Accordingly, AA involves establishing baselines for patterns of network traffic, creating and using machine algorithms to enable spotting anomalous patterns that may indicate cybersecurity threats.

1. AA is used to look for anomalous patterns in terms of activities associated with reconnaissance, compromise, or the exfiltration of data.
2. Grouping actions and associated pattern flows from normative behavior and updating these on a regular basis is critical.

ii. AA provides the ability to determine potential malicious traffic or patterns of malicious conduct without a predetermined signature.

1. Indicated by actions that align to common tactics associated with malware, hacking, or cyber kill chain types of intrusion events and occurrences.
2. Indicated by activities that do not have any apparent or reasonable alignment with organizational operations or mission or that could indicate attempts to search out or exfiltrate data.
3. Correlation-based methodology, based on “if this, then that” procedures, scoped to analyze the subsequent impact of changes made on systems and the resulting likelihood they are associated with an intrusion.
4. Analyst performs further review to determine if anomaly is associated with a potential problem event (e.g., vulnerability, threat, or incident).

5. Can include active defensive actions based on heuristic behaviors -- for example, when activity is spotted, firewalls or Intrusion Detection System/Intrusion Protection System can respond automatically with a defined action.

iii. AA is not:

1. Solely signature-based analysis, which differentiates AA from current signature-based programs such as Einstein (though AA may work best within a portfolio of cyber defenses that will include signature-based analysis (SB): a combined, or hybrid, toolkit will emerge.)
2. Assessment of individual behavior that implicates privacy concerns associated with end user identification and individual behavior analysis (such activities are not the focus of the network traffic data analysis that is the core of the work here).

b. What technologies are used to conduct these inquiries?

i. Correlation Methodology & Tools.

1. Use of Log Aggregators and other business intelligence tools to digest all logs of events so that rules can be written.
    a. Visual data analytics that alert on specific behavior.
    b. Alerts and warnings that flag for anomalous behavior.
    c. Correlation events set to trigger other devices to either passively or actively react (i.e. block) said actions.
2. Real Time Rules and queries based on current warnings elsewhere in the system to provide a real time review of system actions.
3. Thresholds for slowdowns in the system to allow more investigative time (i.e. network speed and bandwidth).

ii. Analytics and Risk Thresholds.

1. Matrices for score mapping of algorithmic traits
    a. Low, Medium, High risk scoring.
    b. Thresholds attributed to events, systems, and departments.
    c. Threshold change based on feeds from information sharing.
2. Scoring determines what level of scrutiny to pay to the activity and whether active blocking enabled (H), on watch (M), passive monitoring (L), or information for situational awareness – this allows

for standard guidelines, but gives operators and analysts flexibility to pay attention to only the most important items.

3. Preventative egress pathways outside the system based on anomalous traffic.

iii. Using Tools &Methodology to Implement AA in Security Operations[2].

1. Front End Authentication
    a. Determining end user actions are in line with their role (linked to cyber kill chain methodology).
    b. Correlating user actions across the environment to spot anomalous actions.
    c. Watching administrative users with elevated privilege or newly escalated accounts.
2. Transaction and Experience Monitoring (of actions taken on the system).
    a. Monitoring system interactions and changes to the system.
    b. Correlating items and actions occurring across the system.
    c. Correlating separate, but potentially related activities over a specified timeframe and determining whether these actions, taken together, are likely/not likely indicators of compromise.
3. Egress Methodology.
    a. Spotting exfiltration and egress packages of staged data prior to their exfiltration.
    b. Identifying network communications to external command and control or other known bad networks.
    c. Identifying the post-exfiltration activities, such as log cleansing, deletion of user admin accounts, etc. that indicate clean up efforts to defeat subsequent forensics.

c. DHS has pilots in process to assess effectiveness of this model.

i. DHS' implementation of algorithmic analytics is currently being done as part of a pilot project called "Logical Response Aperture" (LRA). DHS describes LRA as follows (see Appendix B for more detail about the LRA program):

[2] The DPIAC recognizes that current analytics programs are scoped to monitor incoming and outgoing network communications, but is providing guidance that includes internal traffic as well since these programs may evolve in scope in the future.

> “The Department of Homeland Security (DHS) Network Security Deployment (NSD) is currently engaged in an applied research task, known internally as LRA, related to automated security analytics and countermeasures. It is envisioned that this effort will substantially improve the rate and speed of both detection and response to hostile activity against the networks of US government agencies and other protected entities. In support of this effort, NSD requires collection and retention of additional network data, to drive the security analytics that will be constructed. These analytics will use computational intelligence approaches, allowing identification of attacks without signatures or indicators.”

ii. The DPIAC recognizes that the LRA pilot follows certain protocols and is limited to information that enters and exits a government system, and commends DHS for ensuring that privacy protections have been addressed throughout LRA operations. Consistent with the question posed to the DPIAC in the request letter cited above, our recommendations address how best to protect privacy in AA programs more generally, whether they emerge as future adaptations of the current LRA pilot or in other governmental settings.

iii. In expanding AA programs like LRA, DHS could map to current operations in three areas:
1. All federal systems being watched.
2. The specific baseline of an agency.
3. Flow data from major partners, such as companies participating in the DIB pilot or similar programs.

d. A number of private-sector companies are implementing similar models to combat ever-increasing and ubiquitous cyber threats. The DPIAC recommends that DHS develop benchmarks for success relative to private sector efforts.

2) Key considerations that impact algorithmic analytics.

a. Key privacy considerations -- How is PII affected in AA? Though direct linkage to PII is likely to be minimal, we recommend addressing questions that may arise about potential PII impact and protections, including:

i. How and when might PII be included?

ii. What kinds of PII are included?

iii. What other privacy issues exist in implementing algorithmic analytics? Examples may include:
    1. Mishandling AA information that can be connected to individuals.
    2. Correlation of AA data with PII.
    3. Improper alignment with FIPPs (notice, access, use, sharing disposition).

iv. Generally, AA programs may involve three categories of information, each of which calls for a different level of response with regard to protecting PII from among the recommendations below. The DPIAC finds that special or additional privacy protections should be considered for categories 2 and 3 below.
    1. All traffic entering or exiting a system – this information would generally flow unimpeded and not be retained beyond a minimum necessary time frame (see below “Retention”); basic system privacy protections would apply.
    2. Information where anomalies or related analytics demonstrate potential malware or other risks/threats – this small subset of overall traffic would be analyzed and for further investigation and potential action, necessitating special care in protecting PII.
    3. Sample data used for training purposes – these data sets should not include PII, if necessary by stripping or otherwise obfuscating privacy-sensitive material.

v. If PII questions arise:
    1. Existing protections for Federal systems and data would apply, including privacy notices and training re proper safeguards.
    2. Prior DPIAC recommendations are relevant here, specifically the good process for logging discussion in Report 2014-02 Privacy Recommendations Regarding Auditing and Oversight of the DHS Data Framework.

b. Data Quality and Integrity

i. Importance in the context of AA.
    1. The goal is for the AA initiative to grow into a predictive capability enabling a robust defense.

a. Key to this is achieving a tolerable level of both false positives and false negatives in order to minimize possible harm to innocent third parties.

b. This makes data quality and data integrity essential, to understand the possible impact of AA data on privacy in the context of the DHS AA initiative.

2. Any data analysis will only be as effective as the confidence in data quality and integrity.
3. When AA processes and software follow a rigorous logic and well-defined, executable code, any difference in a data element due to poor quality will have significant impact in a formula or computational procedure.
4. This is even more so important when dealing with the quality of data elements that carry PII.

ii. Definitions and rules.

1. As an operational issue, data quality can usually be summarized as ensuring that the data is fit for use.
2. There are several additional dimensions -- data will be fit for use if it passes a test of reflecting reality, at least with respect to accuracy, consistency, completeness, timeliness, uniqueness and validity.
3. This applies not just to the actual data but also to metadata.
   a. Most collection devices, whether sensors, the networks that transmit the data or the software that controls and manipulates it, generate a significant amount of metadata that is extremely important in providing context and key in any analytical process.
   b. Metadata can also be an important conduit for PII, either via the generation of a timestamp or geolocation in an entry.
4. The Committee on National Security Systems defines data integrity as the “Condition existing when data is unchanged from its source and has not been accidentally or maliciously modified, altered, or destroyed”. (CNSS Instruction No. 4009).

iii. Distinguishing data and metadata.

1. It seems likely that in a significant number of the instances, analysis is going to be done on metadata; that is, not the main content of a message but the accompanying data elements that provide the context.

   2. Usually, these elements cover important aspects of the message such as start and end time, originator and receiver, routing, location (where relevant and accessible such as when transmission involves mobile devices), etc. (DHS has stated that the actual content of the message transmission, what is usually considered to be the data, will be reviewed if it is executable, but is otherwise apparently not important to AA.)
   3. This should be clarified in order to avoid future confusion --thus far there have been no specific references to metadata, or distinctions made between data and metadata, despite metadata's apparent key role in AA.

iv. Protecting against threats.
   1. Threat sources may attempt to exploit certain aspects of normal IT process such as copying, versioning, backups and other handling operations.
   2. Therefore, AA should also focus on preserving the physical and logical integrity of the data.
   3. Collection, ETL and security need to be reviewed to protect integrity, as well as the backups, restores, emergency recoveries and similar operations (especially when dealing with systems of record where integrity is paramount).

v. Data governance.
   1. The AA initiative would substantially benefit from a data governance framework to address data quality and integrity issues.
   2. This would provide structure and guidance to data operations, allow confident tracking of provenance and sources, supplement security procedures and policy, and document the physical as well as logical integrity of the data.

vi. Specific issues around mobile devices.
   1. Given the possibility that some of the transmissions that will be collected and analyzed will be routed via mobile devices, the issues of location awareness and user identification, and related privacy implications could be significant.
   2. This can be both metadata as well as data, but it requires attention, process rules, and further assessment and guidance as the program expands.

vii. Systems of records -- any data quality initiatives must also take into consideration any system of records that may be involved. While this may not be the case for this AA activity it is an important consideration.

c. Accountability

i. Human oversight is vital to adding accountability for the governance of automated decision-making systems, including information-sharing systems that use advanced algorithmic analytics to identify cyber threats.

ii. Such systems have implications for individuals, sometimes directly when an individual’s PII is included in the shared information at some stage, and also indirectly because individuals are affected by the accuracy and timeliness of the transactions of the agencies whose information is subject to the analysis.

iii. Scope of human oversight in systems for sharing cyber-threat indicators.

1. Algorithms should be designed to enable review by overseers.
2. Reviews by overseers should take place on an ongoing basis.
3. Comprehensive reviews should be conducted at planned periodic intervals.
4. In non-automated “exception” cases, overseers should review decisions promptly for compliance with established guidelines and standards.
5. Where feasible, more than one human overseer should be involved in a decision based on the data and metadata reviewed.

iv. Factors overseers should address.

1. Procedural regularity in the application of the algorithms (“technological due process”).
2. Fairness in the selection of random inputs.
3. Presence and treatment of PII in information shared.
4. Avoidance of discriminatory (on protected characteristics) and other inappropriate input and processes in the algorithm’s design and implementation.

v. Who provides oversight of the overseers?

1. Those charged with providing oversight must be granted the necessary authority, which may be challenging in complex environments.

2. The recommendations of the overseers and the measures taken in response should be reviewed periodically by the DHS Privacy Office (and other appropriate DHS executives).
3. Redress must be provided for affected individuals: See DPIAC report "The Elements of Effective Redress Programs." Particularly relevant are the sections on clear ownership and accountability and integrated structure for redress.[3]

3) Questions to address for major categories of information handling:

As algorithmic analytics are applied through the lifecycle of information handling, different privacy protections may be appropriate at each stage, including collection, use, sharing, access, retention, and disposal. The DPIAC has addressed good information handling practices for protecting privacy in cybersecurity programs previously (see Report 2012-01 Privacy and Cybersecurity Pilots), and recommends that those practices be applied to the AA program as well.

a. Collection

i. "Strip and encrypt".

1. Only collect PII when necessary component of anomalous pattern;
2. Strip out (remove) PII unnecessary to performing targeted analytics, especially when traffic content is retained (unless the PII is essential to any follow-up analysis).
3. Encrypt if possible PII that is collected or retained.
4. If encryption is not possible, then protect via policy and appropriate administrative, physical, and technical safeguards.

ii. Encryption in transit and/or at rest.

1. Encrypting data may complicate the storage of the data, the speed at which the data is accessed, key rotation issues etc. If this data is going to be used in analytics engines, care should be taken to ensure additional friction is not introduced into data flows.
2. The DPIAC recognizes the issues that could be created if encryption is mandated, and instead recommends the use of

[3] Available at www.dhs.gov/sites/default/files/publications/DPIAC%20Recommendations%20Report%202010-01.pdf.

reasonable safeguards; consider a risk management-based framework, as the DPIAC recommended regarding Auditing and Oversight of the DHS Data Framework (Report 2014-02 Privacy Recommendations Regarding Auditing and Oversight of the DHS Data Framework).

iii. Since the ability to scrub the data (for example, by stripping PII) at the collection point is limited, individuals should receive notice of the program and data collection.
   1. Consider notice recommendations that could align with current, overall monitoring/end user surveillance notices and the purposes for which the data is collected and retained (*i.e.*, potential attack analysis, see "Use").
   2. Use standardized notices to explain that messages and electronic communications are subject to "inspection to review message risk and security."

iv. Criteria should be developed for selecting particular network segments for data collection, or particular risk factors.
   1. Recommend that collection site selection criteria be risk-based to avoid over-collection or low-value targeting (see also, "Retention").
   2. Consider workgroup characteristics that (a) increase Internet exposure, or (b) increase the target value of (or interest in) local assets/data for threat actors.

v. Future Considerations.
   1. DHS should consider the feasibility of identifying potential PII that is structured (and thus addressable through automated means) like SSN or account numbers–*i.e.*, items that can be automatically obfuscated by a machine.
   2. DHS should consider obfuscating the most sensitive data automatically (or at least masking the display of the data to analysts in a manner that retains speed, enhances privacy, and reduces complexity).

b. Use

i. Government uses the information collected in this system in the following ways:

1. Network traffic and analysis,
2. Detect and prevent fraud, security or other technical issues,
3. Manage government websites and other online assets,
4. Protect, enforce, or defend the legal rights, privacy, safety, operations, or property of the public or government employees.
5. Monitor interactions with government systems and networks.
6. Meet any applicable law, regulation, legal process or enforceable governmental request, and
7. Combine information about individuals about threat sources from multiple sources.

ii. Should this information be limited for use in protecting networks?

1. Should be used to protect the networks, but protect the networks also means that to the extent feasible, information is shared with others performing investigations on cyber-attacks from a criminal side as well in order to prosecute the crime, if any, constituted by the cyber-attack.

2. The DPIAC does not suggest a carte blanche approach, but recommends that DHS move from 100% defense to being in a position to prevent attacks – all under MOU and with data ownership and stewardship rights.
3. Saying that we will use the data for “law enforcement” purposes is too broad - need to narrow, only share what is relevant to the target of inquiry and not general information.

iii. DHS should develop protocols for addressing PII that may be accessible in the course of follow up analysis by following established PII handling protocols for monitoring programs (*e.g.*, Einstein).

1. Consider data set categorization for protocol development and application (*e.g.*, data collected, data identified as potentially malicious, data retained for system training purposes).

iv. Other possible uses to consider: investigations, agency internal analysis

1. Information Sharing - Use the data to learn about network attacks or threats, share this learning with other governmental agencies, ISACs, etc.
2. Network Defense - If patterns of attack are seen, blocking them has benefit, but preventing them elsewhere and shutting down the hackers is better.

c. Sharing

i. How should this system interface with CERT and other info sharing centers?
    1. Interfacing with CERT is beneficial, but with analytical attack data that shows what 5 steps to attack (i.e. the methodology of the attack); this might be better provided to ISACs, InfraGard, and ECTFs, and other specific sectors through NPPD's Sector Coordinating Councils (possible distribution through NCCIC).
    2. The DPIAC recommends reviewing how to use DHS Sector Coordinating Councils (SCC) for information sharing as well.
    3. The DPIAC supports DHS practices for sharing, including:
        a. Sharing of 90-day buffer of network traffic data
            i. This data will only be accessible by <30 analysts with controls in place to prevent abuse.
        b. Sharing of tested and verified analytics.
            i. To-date analytics have not included PII, but they might in the future.
            ii. In any case, the volume of analytics data is many orders of magnitude less than the network traffic data buffer.
            iii. Whenever possible, PII data will be removed from analytics (data minimization).
            iv. Any information sharing should be in accordance with the US-CERT Cybersecurity Information Handling Guidelines, and other US-CERT SOPs. LRA is just a new way of looking at this information, but does not change how DHS shares information and disseminates products.
            v. There is no plan for automated sharing mechanisms to be implemented within this capability.

c. Sharing of individual indicators derived from network traffic.
    i. Individual indicators generally do not include PII, but they can.
    ii. Every attempt is made to produce indicators which do not include PII, and PII is only included when essential to producing an actionable indicator.
    iii. Indicator sharing takes place only after a manual review process has occurred, wherein US-CERT should consider not only the quality of the indicator, but also any privacy or OpSec concerns.

d. Access

  i. The Committee recommends that prior DPIAC recommendations re: log/audit controls for analyst access be applied so that AA aligns with that guidance (Report 2014-02 Privacy Recommendations Regarding Auditing and Oversight of the DHS Data Framework).

  ii. Circumstances that dictate access – information from DHS.
    1. This data will only be accessible by <30 analysts with controls in place to prevent abuse.
    2. Access to buffer of network traffic data.
        a. Analyst runs query to look for malicious events.
        b. Analyst inspects related data to verify malicious event
    3. Access to tested and verified analytics.
        a. Suspected malicious traffic will be accessible by CERT as they attempt to work with the affected D/A to address/mitigate the cyber incident, in accordance with the US-CERT Cybersecurity Information Handling Guidelines, and other US-CERT SOPs.
    4. Controls.
        a. Access to buffer of network traffic data.
        b. Normal controls from NIST SP 800-53 for access control or privileges.
        c. Strong limits on volume of data download/exfiltration.
        d. Encryption of data at rest in production.
        e. Separation of duties between running queries and evaluating query results.

5. Access to tested and verified analytics, in accordance with US-CERT Cybersecurity Information Handling Guidelines, and other US-CERT SOPs.
6. Logs.
    a. Access to buffer of network traffic data.
        i. Logging of all queries by date, time, device and analyst.
        ii. Logging of all downloads by date, time, device and analyst, with limits on volume per day overall and per person.
    b. Access to tested and verified analytics, in accordance with US-CERT Cybersecurity Information Handling Guidelines, and other US-CERT SOPs.

e. Retention

i. With respect to the retention of traffic and other data obtained in connection with an AA program, the program manager should carefully consider how long data in each of the relevant categories (i.e., traffic data entering/exiting the network, data linked to malicious activity, and data obtained or maintained for the purpose of training the system) appropriately should be retained. As a rule, the data should be retained only as long as necessary to (1) serve the purposes of the program and (2) comply with relevant, existing data retention requirements.

    1. The Committee is not seeking to impose prescriptive time limits for retention but rather is recommending that the program manager thoughtfully consider, on an ongoing basis, appropriate retention periods that are the shortest necessary to serve the twin goals set forth above.

ii. Decisions regarding retention periods should be continually assessed, with an eye toward retaining the data for the shortest time period necessary to serve the purposes of the program and comply with relevant retention rules. The Committee recognizes the need for flexibility and that the "minimum necessary" period may change over time, which is the reason continuous assessments of appropriate retention periods are recommended.

iii. Based on the determinations made regarding appropriate retention limits, the program manager should memorialize in a written document, specific, concrete retention periods (e.g., 180 days for traffic data entering/exiting a network). The written document should be revised periodically as the retention periods change over time.

iv. Exceptions to the defined retention periods should be permitted only in exceptional circumstances, and then only for short, carefully considered additional periods. The exceptions process should be detailed in a written document and closely adhered to when requesting and granting extensions to the defined retention periods.

v. Adherence to the retention rules should be the subject of periodic internal audits. Such audits should be conducted at least annually and at additional intervals following any significant deviation from the rules.

f. Disposal

i. Once the defined retention period has expired, the data promptly should be securely destroyed.

ii. Protocols should be designed to ensure the security of the data upon destruction such that the data cannot be reconstructed or otherwise made legible following destruction.

iii. In securely disposing of the data, the program manager must consider all copies of the data to be destroyed. This includes original versions and all replicated copies, including backups.

iv. The methods of disposal should be revisited from time to time to ensure that the security protocols evolve as the security landscape changes.

1. The program manager should ensure that state-of-the-art destruction techniques are implemented upon disposal of the relevant data.
2. Consideration should be given to then-current government standards regarding data disposition (e.g., DOD 5220.22m or then-current NIST standards).

v. Adherence to the rules on secure destruction should be the subject of periodic internal audits. Such audits should be conducted at least annually and at additional intervals as appropriate.

Report 2016-01 of the DHS Data Privacy and Integrity Advisory Committee
On Algorithmic Analytics and Privacy Protection

Appendices

Privacy Office
U.S. Department of Homeland Security
Washington, DC 20528-0655

Homeland
Security

September 23, 2014

Ms. Lisa Sotto
Chair, DHS Data Privacy and Integrity Advisory Committee (DPIAC)
c/o Hunton & Williams LLP
200 Park Avenue
New York, NY 10166

Re: DPIAC Guidance on Behavioral Analytics in Cybersecurity Capabilities

Dear Lisa,

The Department of Homeland Security (DHS) continues to evaluate its technical and policy approach to cybersecurity initiatives. The growing sophistication of cyber threats has reduced the effectiveness of some cyber defenses and necessitated the development of new capabilities, including behavioral analysis and reputation management. Implementation of a behavioral analysis capability will require collection of certain network data elements. The Privacy Office seeks additional guidance from the DPIAC specifically on data retention and data access.

In order that the Department's cybersecurity efforts may benefit from the DPIAC's substantial experience in both technology and privacy, I request that the Committee provide written guidance on privacy best practices for DHS retention of data and access related to behavioral analysis in cybersecurity initiatives. Specifically, I ask that the Committee to review the various stages of the behavioral analysis process and consider and address ow long such data should be retained, who should access the data collected by these programs and under which circumstances these individuals are permitted access, and what should be included in any human review of indicators our outputs.

I ask that the Committee build its guidance on the fact-finding conducted by the Cybersecurity Subcommittee to produce an unclassified report and recommendations so that not only the Department but also the larger cybersecurity community can benefit from the Committee's advice. If my office can provide any assistance to you as the Committee undertakes this tasking, please do not hesitate to let Shannon Ballard or me know.

Very truly yours,

Karen L. Neuman
Chief Privacy Officer

Privacy Office
U.S. Department of Homeland Security
Washington, DC 20528-0655

Homeland Security

January 16, 2015

Ms. Lisa Sotto
Chair, DHS Data Privacy and Integrity Advisory Committee (DPIAC)
c/o Hunton & Williams LLP
200 Park Avenue
New York, NY 10166

Re: DPIAC Guidance on Behavioral Analytics in Cybersecurity Capabilities

Dear Lisa,

On October 22, 2014, the *ad hoc* DPIAC Cyber Subcommittee met via teleconference with representatives from the National Protection and Programs Directorate's Office of Cybersecurity and Communications to discuss the Behavioral Analytics in Cybersecurity Capabilities tasking issued to the Committee on September 23, 2014, seeking additional guidance from the DPIAC related to behavioral analysis in cybersecurity initiatives.

During the teleconference it was suggested that in order to provide the best possible advice, the tasking should be broadened to include review throughout the lifecycle of the cybersecurity initiatives. To that end, I am revising the tasking **(in bold)** as follows:

> Specifically, I ask the Committee to review the various stages of the behavioral analysis process and consider and address **how best to protect privacy while achieving the cybersecurity goals of such analysis across the various stages of the information lifecycle, including why data are collected, what data elements are collected, and how those data are collected; how the data are analyzed and otherwise used by the Department, including how they are shared with other Federal agencies and with outside parties;** how long such data should be retained, who should access the data collected by these programs and under what circumstances these individuals are permitted access, and what should be included in any human review of indicators or outputs.

Again, I ask that the Committee build its guidance on the fact finding conducted by the ad hoc Cybersecurity Subcommittee to produce an unclassified report and recommendations so that not only the Department but also the larger cybersecurity committee can benefit from its advice. If my office can provide additional assistance to the Committee during the undertaking of this tasking, please do not hesitate to let Sandra Taylor, Scott Mathews, or me know.

Very truly yours,

Karen L. Neuman
Chief Privacy Officer

Data Flow for the LRA Advanced Automated Analytics Prototype

26 September 2014

Overview

The Department of Homeland Security (DHS) Network Security Deployment (NSD) is currently engaged in an applied research task, known internally as LRA, related to automated security analytics and countermeasures. It is envisioned that this effort will substantially improve the rate and speed of both detection and response to hostile activity against the networks of US government agencies and other protected entities. In support of this effort, NSD requires collection and retention of additional network data, to drive the security analytics that will be constructed. These analytics will use computational intelligence approaches, allowing identification of attacks without signatures or indicators.

This document describes the flow of data through the LRA prototype. It describes the data that will be processed by the prototype, how long the data will be stored, and how it will be processed and viewed. The prototype is exploring advanced automated analytics for detecting potentially malicious behavior in computer network traffic.

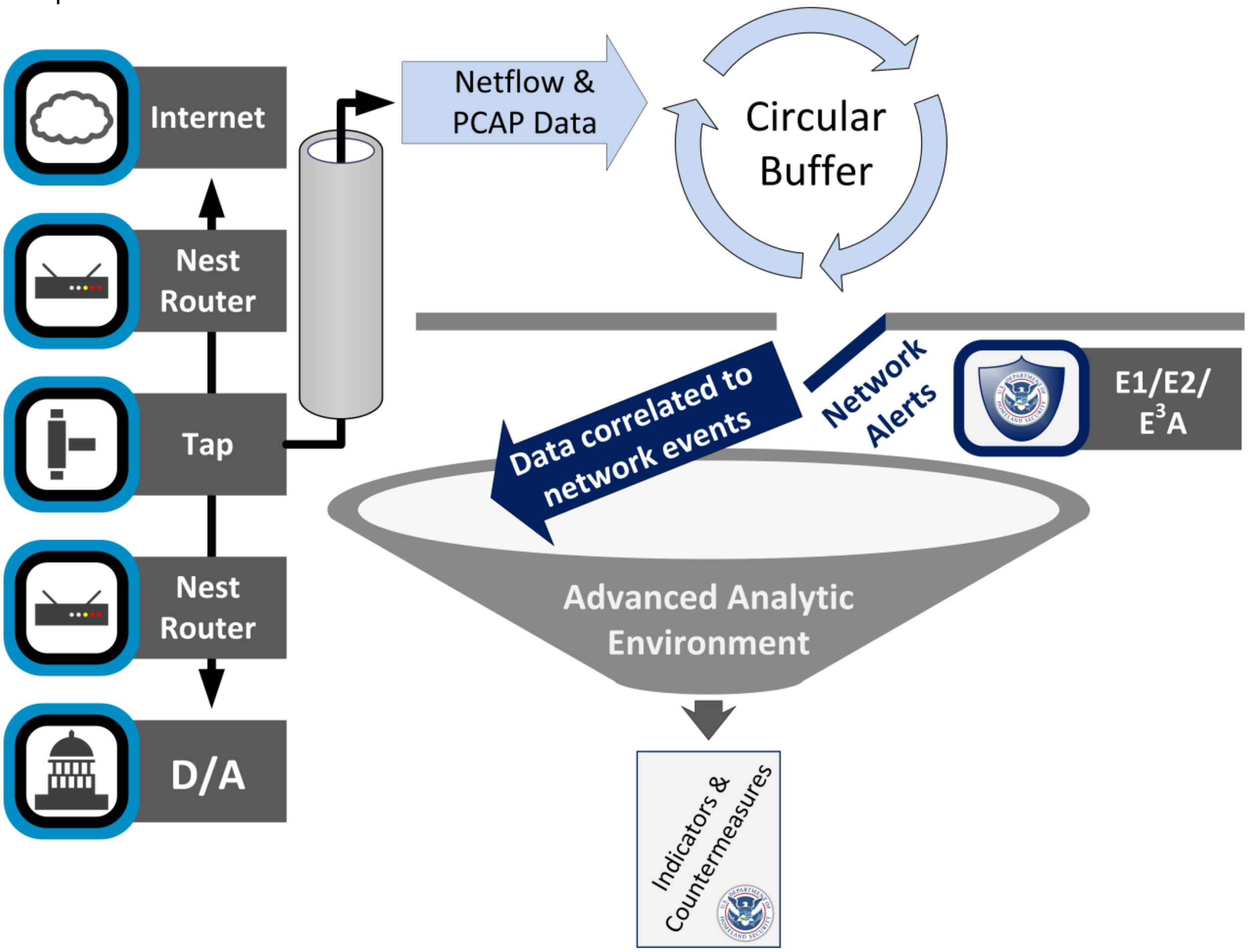

Figure 1: Overview of the LRA Prototype

Anticipated Data Flow

The prototype is intended to test the applicability of various analytic methods to the real data available at an Internet Service Provider (ISP). The technology categories listed below merely represent place holders for the technologies candidate analytics might require in a production implementation.

The network traffic from the Internet flows through the prototype as follows:

1. An externally facing router directs traffic to the router-tap-router configuration in the NEST.
2. A passive tap behind the router copies the traffic to be sent to the prototype. This may be a subset of the traffic on this network segment; if so, it will be selected by virtual local area network (VLAN) id so the prototype only has access to traffic for which it has been approved (e.g., by particular departments or agencies). Packets in the network traffic are parsed by protocol slicers (currently planned to be Bro and others). The data is stored in the Limited Duration Log Storage. Example data fields include:
    a. The domain name that is the subject of a DNS query
    b. The IP address that is the response to a DNS query
    c. The time to live (TTL) value for a name-IP address association
    d. An executable file that was attached to an email message
    e. A user agent string that was part of an http request
3. Analytics engines (based on computational intelligence and statistical algorithms), malware detectors, and other automated tools access the data from the Limited Duration Log Storage store, possibly combined with data from the External Data Feeds Storage store (containing whitelists and blacklists for domain names and executables obtained from external sources, and GeoIP data linking geographic tags with IP addresses), and produce potential indicators of malicious traffic that are stored in the Potential Indicator Storage store.
4. Humans use the Analyst User Interface (UI) to examine the potential indicators, evaluate the prototype's performance, and modify the prototype as needed to improve its performance. This is also used to produce reports of how the prototype performed and what it found.

Data Retention and Queries

The Advanced Analytics program will require access to and retention of particular data sets in order for the program to achieve mission success by executing real-time detection of and protection against network-based threats.

The data will be stored in the same format in which it was captured (PCAP) and it will be sessionized, carved, or otherwise extracted and changed in format to aid analysis. The prototype is planned to be active for a limited duration, currently estimated to be 90 days. All of this data will stored in the Limited Duration Log Storage.

The desire to keep all the data during the evaluation period stems from a belief that partway through the evaluation period we may find indicators of events that were not anticipated at the outset, and we will want to test those indicators against previously seen data. This would involve "rewinding the buffer" and possibly having to re-examine the raw data to extract additional features. Keeping past data is important for testing of new analytics because similar data is difficult to find and testing is key to reducing false positive rates, which increases the amount of automation that we can accomplish.

Staff working on the prototype will use the UI to manually query the data stores to understand what the various algorithms have found, why various indicators were proposed, and whether any malicious traffic was missed by the algorithms that could be identified by the staff (perhaps by looking at patterns known to the staff but not encoded in the algorithms, or from tips from outside sources). Examples include:

- Retrieve all DNS records for a certain duration that contain a given domain name and plot records on a timeline
- Retrieve all names of files seen as attachments for a certain duration and compare against a list of known malware filenames to find which files have been previously recognized and which have not.
- Retrieve all traffic for a certain duration and plot by time and by protocol on a scatterplot to see if there are trends that indicate potentially malicious behavior exhibited by any known protocols

Note that the above are representative examples but do not cover all possible queries. The prototype is expected to look for anomalous and malicious traffic in domain name system (DNS) data, hypertext transfer protocol (HTTP) data, electronic mail header information, and executable files (e.g., seen as attachments to electronic mail messages). The analysts will follow DHS' cybersecurity information handling guidelines. In general, analysts will retrieve results from the data stores in the following scenarios:

1. Understanding or validating the results of an analytic: Analytics often return indications of anomalous behaviors and it is up to human interpretation as to whether those anomalies are benign or malicious. Also, as the algorithms are in some cases very new, they will need to be validated. Some validation can be done automatically but some is likely to be done using manual means. The data humans will look at in those cases will be limited to that needed to validate or invalidate the specific behaviors identified by the analytic.
2. Understanding the context of malware or other malicious activity: We expect malware to come attached to an electronic mail message. To understand the intent of the malware and help identify how it works and to whom it was targeted, analysts may need to issue a string of related queries that help unravel the data stream that included the malware. In this case, the analyst will have high confidence that the electronic mail and any related data are part of malicious activity.
3. Ad hoc searches for anomalous and malicious behavior: As analyst develop new analytics, it may be natural to perform ad hoc searches for clusters of potentially malicious behavior before coding up a complete algorithm. The analyst in these cases is looking at aggregate information of sizes of clusters created based on a selection of attributes (that is, tags, not values) and only looking at specific data to understand or validate the partial algorithm represented by the ad hoc search (similar to the first case above). This will involve targeted retrieval of small subsets of the data in the prototype.

Analysts will not be reading email messages unless those messages appear to be directly related to malicious behavior, such as phishing messages. Analysts will not be issuing queries for individual records, using personally identifiable information in queries, or retrieving personally identifiable or sensitive information unless there is a priori reason to believe that the data being requested is part of a malicious behavior.