Agent HairDO

Jeremy Kackley
Noetic Strategies, Inc
P.O. Box 22225
Huntsville, AL 35814
jeremy.kackley@gmail.com

James Jacobs
Jackson State University/CDID
1230 Raymond Rd.
Jackson, MS 39204
jjacobs@c-did.com

Paulus Wahjudi
Weisberg Division of
Engineering & Computer
Science
Marshall University
Huntington, WV 25755
wahjudi@marshall.edu

Jean Gourd
Louisiana Tech University
Center for Secure Cyberspace
P.O. Box 10348
Ruston, LA 71272
jgourd@latech.edu

ABSTRACT

Blah.

1. INTRODUCTION

One of the weakness in the mobile agent infrastructure for detecting and combating compromised platforms (ref to original paper) is the inability to detect compromised nodes that are passively intercepting information.

The ability to identify passive attackers, compromised agency nodes, compromised agents and /or systems in the DCCP framework requires the ability detect anomalies and /or changes in the mobile agent network and associate these changes to a specific threat. Several techniques exist in the mobile agent research community to eliminate the ability of the network to change by hardening the network. Other techniques rely on the ability to detect, trace and eliminate the anomalies in either the network or the agents.

Most of the work for Mobile agent security is focused on hardening the agent or the agencies against attack. The focus is on encryption or encapsulation techniques to harden the agent or agencies from the ability. These techniques are demonstrated in [1,4]. The problem with encryption or hardening techniques is that they only buy time before the information is decoded or cracks develop in the hardened shell compromising the mobile agent network. The general technique to by extra time is to change the defense faster than the time it takes to crack the system. The problem with this approach is the in ability to actually detect a failure in the defense in a timely manner. For example, a passive node could simply collect data for a period of time long enough to break the defense and after which have little chance of being detected.

Other techniques for Mobile Agent protection rely on identifying and eliminating the threat. The identification mechanism will mark the agent or the transmission of the agent with watermarks [2] or packet tags[]. The watermarks or tags can be used to identify the movements of agents by leaving traces to follow back to the source of the anomaly with passive tracking.

The problems with these techniques are the limited focus does not look at patterns for the entire network. They only compare changes for individual instances. In order to address passive nodes the ability to determine changes in agents and the agents movement, the agencies and the agencies intent Other tracking requires active involvement and transmission of activity to determine the mobile agent's status [3]. Although the active tracking does give a better picture of the mobile agent network status the active scheme mandates a lot of communication overhead.

The proposed pollination scheme is a passive system to allow minimum overhead with active monitoring to provide near real-time discovery of the mobile agent networks status. Pollination involves the exchange of pollen between the mobile agent and the agency to provide a tracking and a pattern mechanism for use with inference modeling. The pollen allows form tracking an individual agent's movements and intentions and the pollination patterns in both the agent and the agencies all for network and agency status to be inferred. The inference model will then classify the intent and the security protocols in DCCP will be in acted based on the perceived intent. Scaling of the pollination model allows for the overhead to be minimized to the level of the threat.

The concept of pollination is designed to create a series of trail markers on both the nodes visited by a mobile agent and mobile agent itself. The trail markers allow immediate identification of what node the agent has visited by simple inspection of the pollen the agent is carrying. The location of the node inspection is trivially the destination (last location) and by traversing the trail of pollen back to the source Node one can trace the historical record of where the Agent has been.

The information provided by pollination is meant for both historical and active. Historical information can be used to determine the sequence of event after and event has occurred. Active information is used form real time inspection to determine if and event has occured.

For example, in a lot of cases the data and the code that processes the data by themselves are not sensitive. However, the ability to get both the data and the code has the potential to cause harm the company. If I mark both of these nodes and have each node in the network sensitive to this situation a mobile agent contain pollen for both locations can apprehended.

Security in information system is an important aspect that cover for all applications that covers three main components: data security, machine security and network security. One of the key problem is authentication of an entity in relation to their access to various resources. A trusted entity might become compromised, and thus untrustworthy, despite being positively identified. Determining if an entity has been compromised is an important but complicated process. This paper introduces the pollination concept that extends the Detect and Combat Compromised Platforms (DCCP) framework [ ] capability to detect compromised platforms that passively intercept information.

2. FRAMEWORK OVERVIEW

The framework to Detect and Combat Compromised Platforms in a Mobile Agent Infrastructure [ ] laid out an agent based framework for detecting compromised platforms. The key aspect of this framework was the concept of threat levels. Threat levels correspond to a global view of how dangerous the situation is, as well as a controlling factor for the operation of the framework. These levels range from One, which can be considered "situation normal" where strictly passive observation occurs, to Four where action is taken against suspected nodes.

2.1 Threat Level One: Network Observation

This corresponds to situation normal and is the default threat level. The key action that takes place at this threat level is establishing and maintaining a network of "probe" agents. These probes can be thought of as a distributed set of eyes and ears. This threat level also sees the establishment of a Central Authority Node (CAN). This node can be thought of as the nerve center of the framework; and as agents percolate through the network they carry reports generated by the probes. These reports ultimately are carried to the CAN which makes judgements based on them. In this way, the CAN can monitor and maintain a somewhat out-of-date view of the entire network at a relatively low cost to performance. This view can be used to search for anomalies, such as a disproportionate number of agents arriving to those leaving a given node. Anomalies are domain-dependent. A certain level of anomalies is expected as a by product of network behavior, thus a threshold value T_1 is defined which indicates the maximum amount of anomalies to be expected in an non-compromised network. Rates above this value would constitute an elevation of the threat level.

2.2 Threat Level Two: Network Suspected Compromise Investigation

At this threat level the network is suspected to be compromised. This leads the CAN to generate new agent types: Commander and Detective Agents. Commander agents can be considered a localized CAN; the purpose of which is to reduce report latency. Detective agents are proactive versions of the probes that communicate observations directly to their commander agent. The concept here is to blockade the suspected node or nodes and investigate incoming and outgoing traffic to see if the anomalies are still occurring greater than the threshold value. There is a network effect whereby any nodes that can only communicate through a suspected node are of course also suspect and cannot be trusted; thus the virtual blockade could comprise a major section of the network. The CAN takes into account this network effect when placing Detective agents so as not to compromise the aggregate data. Another important point is that this is merely an investigative roadblock; communication is investigated and monitored, but is not stopped. Again, anomalies are domain dependent, but it makes logical sense that there would be more types of anomalies defined at this level. Additionally, T_2 is the second threshold value of anomaly detection prior to elevation to the third threat level.

2.3 Threat Level Three: Network Compromise Confirmation

This threat level sees the creation of an additional type of agent, the Secret Agent. The Secret agent is something of a sacrificial agent. Its actions (and the expected results thereof) are predefined however; ergo it can be send to the compromised node, and if the results are not observed exterior to the node, or its communication of its observed effects do not match the observations of detective agents, then an inference can be made that compromise has occurred. It is possible that the Secret agent would never be heard from again, in which case this process must be repeated until either the agent survives, or a set number of agents have been sacrificed. It is possible at this layer either to elevate to level four, or to deescalate if the Secret agent is not interfered with.

2.4 Threat Level Four: Network Compromise Resolution

For the final threat level, the assumption is that compromise has occurred. There are a variety of actions that can be taken at this point. The appropriate action is very domain dependent; if resource availability is more important than information security, then simply alerting a human while continuing to gather information is the appropriate response. Alternatively, if information security is more important than availability, or redundant resources exist, then automated responses are possible. The least severe response would be rerouting requests from the compromised node to a sandbox for future analysis and to prevent the compromise of the rest of the network; presumably without making it obvious to the attacker that he has been detected. A more severe action would be to blockade that node from the network, preventing any requests from leaving or going to that node. The most severe would be attempting to remove that node from the network and/or crash it, such as via a distributed denial of service attack or some out of band signal.

DCCP infrastructure is an effective method to detect and combat compromised platforms. The progressive network threat level allows for a dynamic and adaptive detection with varying degree of response. Once a platform is suspected of being compromised, the node and any nodes that comes in contact will be thoroughly investigated before a verdict is given. Depending on the algorithm used to observe the network traffic, DCCP can be adapted to observe the smallest piece of data or it can focus the network as whole. However, DCCP is dependent on one key element which is the active data packets send by the compromised platform that the framework can then intercept and detect for irregularities. In the event that a compromised platform is passive and focus only in intercepting information routed through it without actively sending packets to other nodes as an attempt to obtain information and/or to infect other nodes, DCCP will never suspect that node as being compromised. This weakness will be addressed in this paper through the utilization of a Mobile Agents Pollination (MAP) technique.

3. MOBILE AGENTS POLLINATION (MAP)

Mobile Agents Pollination (MAP) is a concept to address some of the issues with combating compromised platforms previously discussed [] by identifying a mobile agentâĂŹs movement and actions within a network. MAP uses pollen to uniquely identify the agencies or groups of agencies within a network and uses pollination to form a trail map defining the path an agent utilized when visiting the agency nodes. In addition to the trail map, the pollen and map properties can also be utilized as action indicators to infer the meaning of the agent's visitations and the agencies intent within the network.

MAP is similar to the natural process of flower pollination by bees. Bees traverse a field of flowers to acquire nectar. The bee stops at each flower and inadvertently collects pollen and distributes pollen to the flower. A pollen grain represents a flower and a collection of pollen on the bee represents all the flowers the bee has visited. The pollen collection left by the bee at each flower represents the sequence of flowers visitations before acquiring nectar at the current flower. The analogy follows for MAP with the mobile agent as the bee, the nectar as the information and the pollen as an identification marker of a particular agency node. The mobile agent traversing a network of connected nodes to acquire information. MAP defines the process were the mobile agent inadvertently collects pollen from the current node and distribute pollen from the previous nodes visited. The pollen can then be used to quickly map where the mobile agent has traveled and the sequence of travel. Other traits related to the map, the pollen, and the pollen grains relationship to the agency node can be utilized to expand the perceptibility of the mobile agent's activity in the network.

The main purpose of MAP is to utilize pollination for tracking of mobile agent's activity in the network and use the tracking information to infer the intent of the mobile agent and the agencies involved in the network. Pollen is a marker used to identify a specific node a mobile agent may visit. Pollination is the process of exchanging pollen to provide a mechanism to reconstruct where the agent has been and infer the actions the agent performed.

The pollen can be unique to each node or the pollen can be unique for a section depending on requirements. The pollen variation throughout the network is similar to the DNA of a flower. The DNA for a specific node type is mostly the same between entities with minor variations making the sequence unique. The pollen from two different node types will have a greater variation in the predominating factors. The pollen is an identification of a node and should be a dependent on the agency configuration to make the process of spoofing an agency difficult. The information provided by pollination is meant for both historical and active. Historical information can be used to determine the sequence of event after an event has occurred. Active information is used form real time inspection to determine if an event has occurred.

The two key observable parameter categories involved in the process of a mobile agent collecting information from the agencies are spatial and temporal. The space for MAP refers to the network space in direction used by the agency node, the connection, the mobile agent, and the pollen. The spatial category could be expanded to include distance for the connection and the agents travel, however at this time it is unclear on the usage and necessity of the metric.

The network spatial reference is depicted with standard network nodal reference with connections showing the relation between the nodes. The pollen associated with the nodes and the agents will be represented by patterns. The pollen grains cared by the mobile agent and distributed to the agencies maintain the pattern of the origin node. The sequence of pollen grains creates a series of trail markers which define the spatial movements of the agent through the network. A direction depicted by the arrow is inherent in the sequence as the agent moved from the preceding nodes to reach the current node location. The pattern allows immediate identification of what node(s) the agent has visited by simple inspection of the pollen the agent is carrying. Traversing the trail of pollen defined by the sequence will lead back to the source node.

Time for MAP is used to refer to the amount of processing the mobile agent expended at the agency node in terms of number of pollen grains. The number of grains refers to the action time of the agent, such as time spent, data examined, or a ratio of time spent and data examined. The number is associated with a grain of a specific node and the combination of all the grain counts will indicate the time taken for the sequence to complete. A common time reference will maintain a standard gauge for use with the analytics. The actual time interval is up to the implementation, however the granularity of the interval should represent the difference between operation actions of the mobile agents for the operations we want to identify. The temporal reference is depicted by a count below the grains.

We acquire a time-sequence pattern by applying both the spatial and temporal components. The time-sequence pattern maps the networks meaning to the goal of the mobile agent's movement. The spatial, temporal and spatial-temporal observations are used to infer the meaning of the

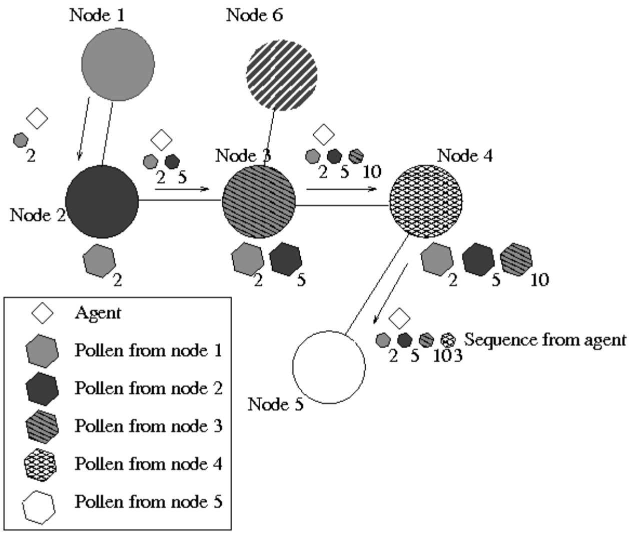

Figure 1: Pollination and Mobile Agents

action with regard to the agent and nodes intent. Associating the nodes content description with the pattern may allow for further identification of the mobile agents actions.

The process of pollination leaves two distinct time-sequence trails as the mobile agent moves from node to node. The first trail is the set of pollen attached the agent from the nodes visited. The second set is distributed along the path of nodes traveled were each node has a snapshot of the previous places the agent has been. The sets have a number of key attributes including: node references, number of pollen spores, sequence of spores and the order (pattern) of nodes visited, amount of pollen attached. The pollination concept is depicted in Fig. 1.

Example: Figure 1 has an example of an agent moving from Node 1 to 5. The agent at each node exchanges pollen with the agency in the process of doing work. The amount of work is quantified with the temporal reference. The agent reaches Node 5 and the resulting pattern is depicted with the temporal references (2, 5, 10, and 3). Note that Node 6 is not path and no pollen is exchange with either the node or the agent. The example could be used as base pattern or a composite of patterns can be used to determine anomalies that occur with either the agent's or the node's standard operation.

3.1 Usage for DCCP

In order to address the passive attack we simply have analytics to infer the meaning of the change in pattern. For example, we have a passive node that inserts itself into the network. The node will have to understand that pollination occurs otherwise the agency node would not be accepted. We will assume they are smart enough to spoof a node to insert it to continue with more advanced security concerns.

LetâĂŹs say we mark the path agents would take from Node 1 to Node 4 in Fig. 2 by marking each of the intermediate nodes with different pollen. Every-time and agent reaches Node 4, I have a sequence of Pollen that corresponds to the path taken. I verify this with the sequence that the agent was suppose to follow to determine if the agent has been compromised by either a additional passive node or a violation of the agent code. The passive node is identified by added pollen, incorrect pollen, or the lack there of in the sequence inspected at the end point. A Commander / Detective team can be used determine which node is the passive node and eliminate it. For compromised agents I can simply eliminate them at Node A or Node B on arrival and trigger the inspection. For added security I can make every node along the path check the pollen sequence and in essence have a passive defense mechanism against corruption.

3.2 Implementation

The pollination process works at the application level where the agencies represent the nodes and the mobile agents move throughout the network hoping from agency to agency. The agencies each have their own pollen definition. Implementing the pollen requires the ability to attach pollen and transport pollen with the mobile agent and read and written by the agencies. It is proposed that process of attaching the pollen to the mobile agent is outside of the agent itself to stick with the inadvertent nature of pollination. The agent himself should not know or care about the pollination. We envision the use of the manipulation of the Open System Interconnection (OSI) model's transport layer for both attaching the pollen to the agent and transporting the pollen and agent to the destination agency. In the OSI model the application data to be transported is broken into packets and transmitted from the source to the destination. We can add additional packets by appending the pollen to the data stream or we can manipulate the packets using packet tagging. Adding additional packets can be accomplished at the application by simply appending to the end of the mobile agent in that data stream. The addendum is removed at the agency and the pollen is recovered.

Packet tagging is accomplished Packet tagging marks packets with identifiers for local purpose [http://www.openbsd.org/faq/pf/ The actual mark is part of the packet and cannot be removed. The mark can be modified or replaced with another mark which maybe an issue if the agent is transmitted external to the system. However, for our purposes the mobile agent is assumed to remain internal.

In addition to the pollen tag itself the count representing action at each node is used to acquire further insight into the meaning. It is expected that a node with little information to share we require less time for the mobile agent to visit the node (less activity). Using the activity gauge we can infer some of the intent of both the agent and the agency. In the example of a passive node scenario we can expect the sharing to be minimal with the agent as described by the 0 for the agent with regard to the passive nodes pattern.

The reverse is also possible with the relationship between the node pattern and the data stored at the node. Using this relationship between the information of the node with the highest activity we can infer the type of information the agent is interested in gathering and the information with little interest. The entire concept is depicted in Figure 2.

The purpose of pollination is to allow and easy identification

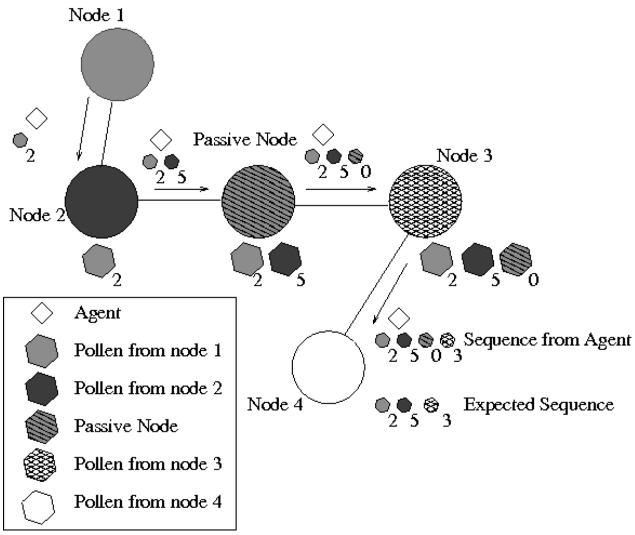

Figure 2: Network with Passive Node

method for activity within a mobile agent environment using pollination patterns. Any standard inference model, Fuzzy Logic, Neural Network, Bayesian can be utilized to trigger DCCP security events from the pollination patterns.

Variation to the scheme can be accomplished to acquire different levels of security through the network. At a low level we are only general concerned with sensitive data or application. For this level we only need to pollinate those locations and track the movement of agents caring that pollen. Agencies are always active to interpret the meaning of the Mobile agents and the surrounding agencies. The limitation of the pollination to a subset of the network has the effect of greater focus ability on only the things that matter. It also reduces the overhead associated with pollination in both time and space.

We can change the pollination patterns associated with the network on a periodic basis to ensure security. This change can either be notified to the DCCP security team in advance or the change could simply trigger and event and allows the team to determine the appropriate action. The later is preferable as the addition of a mechanism to disable security for changing patterns could be an exploit. The event triggered by the pattern change can be used as a test of system integrity as the process goes through the security levels and back to level 1 afterward.

The standard state of the system is in DCCP Level 1 most of the time, however as security concerns increase the number of pollinated nodes increase to match the threat. The increase can either be focus tactically on the relative sensitivity of the data or the increase can be distributed throughout the network to get a big picture look at the secured environment. At the highest level all nodes will be pollinated and nodes without pollen or agent's not containing pollen will be apprehended. The pollination paths not adhering to the required patterns will be examined to determine what event took place by the DCCP security agent team.

3.3 Detecting passive compromised nodes with pollination

Once pollination is in the framework, we can then proceed to setup a set of traps that will assist in detecting compromised passive nodes. The idea of the trap is to lure passive nodes to actively search for a prized data that in turn will expose their cover. The Central Authority Node (CAN) will randomly select a set of strategic node of interest (SNI) throughout the network as a the host of the prized data. Each trap will have a designated area of effect that determines the number of nodes that are affected. CAN will then send agents to each nodes with the objective of broadcasting the existence of a crucial data in the SNI. The CAN will then observe network activities and validate the pollens of every agents that visited the SNI, when an invalid pollen pattern is found, CAN will then raise the network situational awareness into threat level two and then proceed to investigate the path taken by the agent to reach the SNI. Every nodes that the suspected agent have come in contact with will then be considered as a suspected compromised node. The Commander Agent and Detective Agents will then proceed the investigation to confirm whether a node is compromised through level three.

4. MOLE EXAMPLE

5. CONCLUSION AND FUTURE WORK

The DCCP framework provides the ability to detect , investigate and deal with compromised hosts utilizing a multi agent scheme in an effective and efficient network intrusion mechanism. However, DCCP relies heavily on nodes that actively tries to obtain data that they are not entitled to and tipping their hand in the process. In the event that an intruder remains passive and only collects data that it came across, DCCP will be unable to detect it. To cover such deficiency, we incorporate Pollination into DCCP which provides, the ability to detect passive attackers. As an added bonus, Pollination also provides additional forensic capabilities to DCCP..

The role of the pollens can be expanded to include various additional functionality. A heavy burst of pollens in short intervals can affect a compromised node, making it ineffective and possibly shutting it down or forcing the node to restart which could purge the compromising elements. Another plausible feature is the pollen's ability to affect agent's functionality or to modify the data and rendering it useless. A predetermined set of pollen colors could have different hidden emergency message that will be relayed by each host to the CAN. The simple coloring scheme can act as a silent alarm that notifies the CAN for a possible breach in one or more nodes. [8]

6. REFERENCES

[1] B. Al-Duwairi and G. Manimaran. A novel packet marking scheme for IP traceback. In *Parallel and Distributed Systems, 2004. ICPADS 2004. Proceedings. Tenth International Conference on*, pages 195–202. IEEE, 2004.

[2] A. Belenky and N. Ansari. Tracing multiple attackers with deterministic packet marking (DPM). In *Communications, Computers and signal Processing,*

2003. PACRIM. 2003 IEEE Pacific Rim Conference on, volume 1, pages 49–52. IEEE, 2003.

[3] Y. Bhavani and P. Reddy. AN EFFICIENT IP TRACEBACK THROUGH PACKET MARKING ALGORITHM. *International Journal*, 2.

[4] Y. Djemaiel and N. Boudriga. A global marking scheme for tracing cyber attacks. In *Proceedings of the 2007 ACM symposium on Applied computing*, SAC '07, pages 170–174, New York, NY, USA, 2007. ACM.

[5] Z. Gao and N. Ansari. Tracing cyber attacks from the practical perspective. *Communications Magazine, IEEE*, 43(5):123–131, 2005.

[6] M. Goodrich. Efficient packet marking for large-scale IP traceback. In *Proceedings of the 9th ACM Conference on Computer and Communications Security*, pages 117–126. ACM, 2002.

[7] D. Lin and T. Huang. A mobile-agent security architecture. In *e-Business and Information System Security (EBISS), 2010 2nd International Conference on*, pages 1 –4, May 2010.

[8] D. Peng, Z. Shi, L. Tao, and W. Ma. Enhanced and authenticated deterministic packet marking for ip traceback. In M. Xu, Y. Zhan, J. Cao, and Y. Liu, editors, *Advanced Parallel Processing Technologies*, volume 4847 of *Lecture Notes in Computer Science*, pages 508–517. Springer Berlin / Heidelberg, 2007. 10.1007/978-3-540-76837-1_55.

[9] D. Peng, Z. Shi, L. Tao, and W. Ma. Enhanced and authenticated deterministic packet marking for ip traceback. In *Proceedings of the 7th international conference on Advanced parallel processing technologies*, APPT'07, pages 508–517, Berlin, Heidelberg, 2007. Springer-Verlag.