Research Challenges in Dynamic Policy-Based Autonomous Security

Seraphin Calo
IBM T. J. Watson Research Center
Yorktown Heights New York, USA
scalo@us.ibm.com

Elisa Bertino
Purdue University
West Lafayette, Indiana, USA
bertino@purdue.edu

Gregory Cirincione, Brian Rivera
Army Research Labs
Adelphi, Maryland, USA
{gregory.h.cirincione, brian.m.rivera}.civ@mail.mil

Emil Lupu
Imperial College London
London, UK
e.c.lupu@imperial.ac.uk

Saritha Arunkumar
IBM United Kingdom
Hursley, UK
saritha.arun@uk.ibm.com

Alan Cullen
BAE Systems
Chelmsford, UK
alan.m.cullen@baesystems.com

***Abstract*—Generative policies enable devices to generate their own policies that are validated, consistent and conflict free. This autonomy is required for security policy generation to deal with the large number of smart devices per person that will soon become reality. In this paper, we discuss the research issues that have to be addressed in order for devices involved in security enforcement to automatically generate their security policies -- enabling policy-based autonomous security management. We discuss the challenges involved in the task of automatic security policy generation, and outline some approaches based om machine learning that may potentially provide a solution to the same.**

Keywords - autonomous systems; generative policies; security

I. INTRODUCTION

Our goal is to increase autonomy of security for devices, networks of devices, information sources and analytics running on information sources. Thus autonomy is critical for effective security management in contexts that are highly dynamic and where subsystems can become disconnected from human administrators or are intended to operate autonomously An approach has been proposed for the generation of policies in the context of generic domains [1]. Our work is synergistic and consistent with that approach, but our primary focus is on addressing the unique challenges that arise in the security domain in distributed environments. It may be possible for some of our approaches to be used in other domains beyond security.

These challenges include both exploring approaches to generate security policies automatically as well as estimating the security characteristics of this approach in different coalition contexts. The work substantially extends our earlier work on security policies for security appliances [2] and in virtualized environments [1], and has two broad but inter-related thrusts: understanding vulnerabilities and developing Machine Learning (ML) analysis techniques for security, as shown in Figure 1. Vulnerabilities understanding will feed into the design of better machine learning approaches, and improved learning approaches can identify new vulnerabilities.

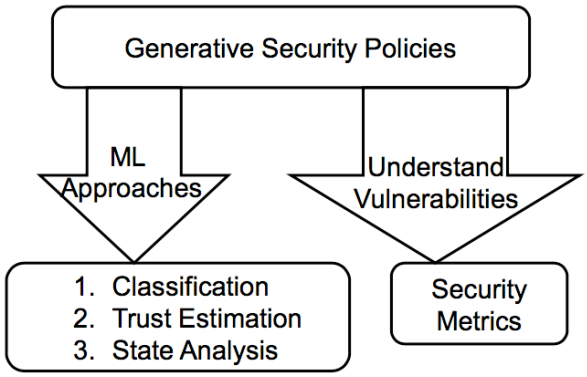

Figure 1. Major thrusts of security research

II. POLICY GENERATION APPROACHES

To explore *security-specific approaches for policy generation*, we need to investigate three different strategies to apply machine learning to generate security policies. Applying machine learning to security for policy generation requires dealing with issues such as the availability of a limited amount of data for training, and the existence of adversaries who can drive such systems astray by data poisoning measures. Other research challenges to applying machine learning in this context includes accommodating small samples with dirty data and heterogeneous data while

This research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence under Agreement Number W911NF-16-3-0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government. The U.S. and U.K. Governments are authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

enabling distributed learning with limited connectivity and under resource-constraints. Each of the suggested approaches provides different ways to address these limitations, and is applicable to a different class of security devices that are likely to be present in future coalition systems. The overall research challenge is to learn policies under various constraints and training strategies using distributed data with distributed training.

A. Classification based Security Policy Generation

Future coalition environments will consist of ubiquitous devices that incorporate significant computational capabilities, and these devices will run sophisticated machine learning algorithms. Devices tasked with security related functions in a coalition environment can use clustering and classification approaches to automatically infer the policies that ought to be used. This approach is most suitable for security environments which are partly automated and partly managed by humans. In these, a device can observe the actions of a human administrator, understand the intent of the human, and then generalize that to create equivalent policies in similar future contexts. Human input is taken as input to classify security related activities into two groups - allowed and disallowed, and techniques for classification (including case-based reasoning, KNN clustering, decision trees and neural networks) can be used to learn and automate such policies driving these operations.

The main challenge to applying machine learning in this environment is that sufficient richness to extract patterns may not be present if human actions are just studied in isolation. To extract true patterns, the context of operations and richer attributes of actors involved in any request need to be determined.

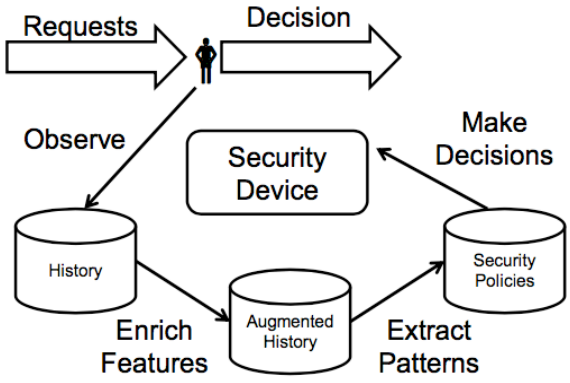

Figure 2. Learning from human decisions

If we consider requests for access to information and services being made by coalition partners, each device sees only limited aspect of each request -- such as the network addresses of requesters, and knows the decision made by the human administrator. However, to extract patterns, the device can learn much more if it also has more semantic context, such as the mission objectives, assignment of personnel to missions, current mapping of personnel to network addresses, and a way to assess the relevance of a document being requested to the mission objective. By providing this semantic context, each security device would have a history of requests that were made by coalition partners for previous missions. It can extract patterns from the accesses that were granted in the past by manual reconfiguration, and learn the features that determine whether access ought to be granted. This knowledge can then be converted into a set of access control rules that are based on the characteristics of the new incoming requests. This approach would work even when the mission objectives, personnel assignments, and relevance of documents are only partially known, since such a mapping is being built using the pattern mining/machine learning approach. Furthermore, we need to analyze and understand whether the security policies learnt in this manner are consistent and free of conflict. Humans may not always be consistent in their enforcement, and extracting a set of consistent policies from inconsistent behavior is a significant technical challenge.

The approach works well for those security domains where some information about human behavior can be learned and enriched. An example use-case will be enabling databases that allow access to coalition partners automatically by learning the right operational security policies to support dynamically formed ad-hoc communities of interest. This approach will build on the mechanisms suggested for the determination of behavioral policies [3]. A preliminary approach for generating access permissions to data has been proposed [4] and has shown that machine learning techniques can be effective in automatically administering permissions. However, this previous approach has many limitations, including the fact that it does not consider contexts and situations when generating permissions nor does it consider the contents of protected data. The ability of cognitive devices to collect context and situation information would allow one to provide the generation of fine- grained access permissions tailored to specific context, situation and data content.

The discussion has been focused mainly on access control policies, which cover a rich set of security situations. They could be for access to different resources, information and services, and access control decisions for different system elements may be linked. Other types of security policies can also be learned from observing the management decisions that are being made by humans, e.g., the requirements for authentication, security classification of information, and data protection and encryption.

B. Trust Estimation Approach for Security Policy Generation

This approach is one that is applicable for classes of security applications where a history of human operation may not be available, but an assessment of the impact of adversarial actions can be done, and damages done by an adversary can be contained. The approach consists of devices dynamically learning how much trust can be put into external

or internal entities - by observing their behaviors and putting them into different categories of trust and risk.

When human participants in a mission need to make a security decision regarding access to a resource or admission to a system, they take into account the characteristics of the person requesting access, the context, and the estimation of risk that is incurred by allowing access to that person. As an example, the doorman at a gated New York building may allow a polite and sober business-man access to a bathroom behind a locked door under some situations, but refuse the same request if it is made by an inebriated person.

Security devices can implement an analogous trust and risk estimation process to decide when to grant access. This is most suited for environments where one can define several security zones, each associated with a level of trust to access resources in the zone, the amount of risk if compromised, and a monitoring system which allows one to monitor the behavior of anyone allowed access into the system.

The trust estimation process will automatically determine, on a per request basis, whether the requester merits sufficient trust to be allowed access to a specific zone. The estimation of the trust depends on the past behavior of the requester in the current zone it has been provided access to, the closeness in properties of the requester to the new zone it wants access to, and an estimation of the risks and utility inherent in allowing the new access request. The trust estimation model would explore techniques to estimate these attributes (trust of requester, pattern based closeness to existing people with access, utility of access, and risks due to access) and study the increases in utility/vulnerability due to allowing such access. A trust approach to security management requires two types of policies: (a) trust-estimation policies providing guidelines on how to estimate trust – for example which factors to consider for estimating trust; (b) trust-based authorization and access control policies by which one can specify attribute-based policies in which some of the conditions are predicates on trust and risk. An example would be a policy by which access to a zone is allowed only if trust is greater than 0.9 (in a [0,1] range).

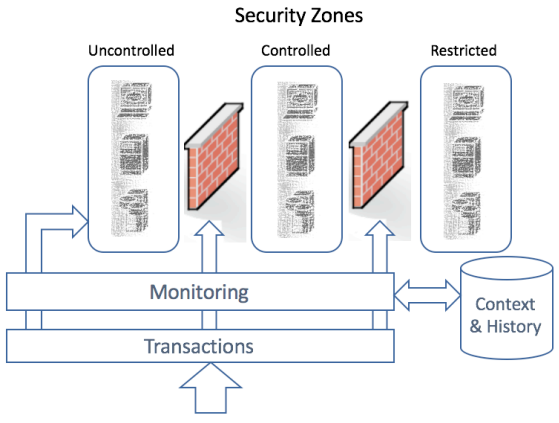

Figure 3. Trust Estimation

Security domains where this approach can be applied include future firewall systems, SDN controllers of coalition partners, and to help attain the vision of SDN based security situational awareness in coalition contexts [5].

C. *Security State Analysis*

While the previous two approaches looked at machine learning models that learnt from past behavior to determine present behavior, security state analysis is a machine learning approach that tries to forecast future behavior. It uses semantic models of system behavior along with the observed past behavior to create a forecast of the future behavior resulting from an incoming operation. The goal is to understand the impact on the state of a system being protected when an operation is requested on the system. To analyze the security state, the system continuously monitors the operations that are being invoked on it, examines how its state changes in response to those commands, and thereby learns (using machine learning techniques) the anticipated response of the system to an operation. This state learning approach can be used to reduce the manual effort involved in defining the information model or state description of a system.

Once the state transitions that happen due to the impact on the system of allowable operations are learnt to an acceptable degree of fidelity, the system can determine which of the operations are likely to put the system into a risky situation. The system can use the risk profile for an operation to determine whether the operation ought to be permitted on the system. State estimation based approaches are suitable for use in coalition cyber-physical systems.

III. NEW VULNERABILITIES

A new challenge that arises in generative security policy architectures using approaches described above is that new *vulnerabilities are introduced when devices generate their own policies*. As an example, consider a security appliance which is designed so that it can generate its own access control policies. Such an appliance allows more flexibility and requires less human touch, but also introduces new vulnerabilities. If the appliance was given a static set of access control policies, it would never violate them. However, when the appliance generates its own access control policies, there is a risk that it may provide access to an unwarranted party. An assessment of the security risk that the automatic generation of access control policies creates is important to determining whether the generative architecture provides a better system.

To address this question, we need to *define the right security metrics* to assess the effectiveness of different approaches. This will allow us to compare the generative security approach (or approaches) with the non-generative human defined approach. Such metrics can also have a significant role in improving the learning process discussed earlier, However, such security metrics need to consider not just the pure security aspect, but also the reduction in human effort for security policy definition that the generative aspect enables. Traditional security metrics need to be combined with metrics for usability, and metrics for reduction in system management effort. *Such security metrics, which*

combine potential increases in security vulnerabilities with usability, reduction in human effort, and overall improvement in system utility, have not been explored in depth before.

We propose to define and investigate the nature of security risks that are inherent in the generative architectures. We would evaluate the security risk introduced in a solution by considering *different aspects*, including *the reduction in human effort, the increase in system usability, the new security vulnerabilities, and the speed with which security mechanisms and tools can be reconfigured to deal with situation and context changes*. We will build quantitative models that evaluate the risk in specific contexts. As an example, we will consider the probability of an attacker being present in making access requests to a virtualized environment and consider the probability of such an intruder getting access with generative policy controls compared to a non-generative approach for manually defining policies. Virtualized environments pose more challenging problems in that they can be dynamic (e.g., virtual machines that can be created or destroyed depending upon traffic load). Policies must be defined or inherited for access to each new element.

In addition to the support mechanisms and knowledge required to define the policies manually, there are human factors issues in the way manual errors are made. Security vulnerabilities introduced due to manual errors will be modeled, and we will assess the impact of net utility and intrusion probability developed due to the new architectures. The human factors issues can be assessed in experiments where people exercise models of the systems to identify the types of error situations that commonly arise. These will be more associated with the manual process for defining policies. However, there may be some issues that arise in the process of providing the higher-level system constraints in the generative model as well.

This will allow us to identify the conditions under which a generative approach would be better than a manual policy definition approach. For example, we expect that a generative approach supported by machine learning techniques can be quicker in reconfiguring the security mechanisms (for example by changing firewall rules and access control rules) than a human user. In addition to the quantitative models, we will also build qualitative models for security metrics that incorporate various aspects. These qualitative models will take the form of questionnaires which provide a rating for the solutions among the different aspects, and combine the ratings in a partial order to compare the security offered by two different approaches.

IV. CONCLUSIONS

In this paper, we propose a framework by which devices involved in security enforcement can automatically generate their own security policies. We propose three different ways of applying machine learning to generate these policies: Classification based, Trust based, and State based. Each of the suggested approaches requires historical data about the operation of the system to train the models.

We note that self-generation of policies introduces new vulnerabilities and outline research directions for addressing them. We propose to define and investigate the nature of security risks that are inherent in the generative architecture. This is future work that needs to be undertaken in order to compare the effectiveness of the generative approach with respect to conventional policy specification methods, which rely on extensive human input.

We believe that the three security policy generation approaches described are applicable to many more coalition contexts beyond the ones used to motivate the approach. We will explore other coalition security scenarios, and assess the efficacy of these approaches in those scenarios.

ACKNOWLEDGMENT

Dinesh Verma and Supriyo Chakraborty helped in formulating some of the ideas and concepts contained in this paper.

REFERENCES

[1] D. Verma et Al., Dynamic and adaptive policy models for coalition operations, Proc. SPIE Defense & Security Symposium, April 2017.

[2] S. Arunkumar et Al., Next Generation Firewalls for Dynamic Coalitions, IEEE SmartWorld Congress -- International Workshop on Distributed Analytics Infrastructure and Algorithms for Multi-Organization Federations, August 2017.

[3] M. Touma, E. Bertino, B. Rivera, S. Calo and D. Verma, Framework for behavioral analytics in anomaly identification, Proceedings of SPIE Defense + Commercial Sensing Symposium, Anaheim, CA, April 2017.

[4] Q. Ni, J. Lobo, S. Calo, P. Rohatgi, E. Bertino, Automating role-based provisioning by learning from examples, Proceedings of SACMAT 2009: 75-84 .

[5] V. Mishra, D. Verma, and C. Williams, Leveraging SDN for Cyber Situational Awareness in Coalition Tactical Networks, NATO Symposium on Cyber Defence Situation Awareness, NATO-STO-IST-148, Sofia, Bulgaria, October 2016.