Distributed AI and Security Issues in Federated Environments

Dinesh Verma
IBM T. J Watson Research Center
Yorktown Heights, NY, USA
dverma@us.ibm.com

Seraphin Calo
IBM T. J Watson Research Center
Yorktown Heights, NY, USA
dverma@us.ibm.com

Greg Cirincione
Army Research Lab
Adelphi, MD, USA
gregory.h.cirincione.civ@mail.mil

ABSTRACT

Many real-world IoT solutions have to be implemented in a federated environment, which are environments where many different administrative organizations are involved in different parts of the solution. Smarter Cities, Federated Governance, International Trade and Military Coalition Operations are examples of federated environments. As end devices become more capable and intelligent, learning from their environment, and adapting on their own, they expose new types of security vulnerabilities and present an increased attack surface. A distributed AI approach can help mitigate many of the security problems that one may encounter in such federated environments. In this paper, we outline some of the scenarios in which we need to rethink security issues as devices become more intelligent, and discuss how distributed AI techniques can be used to reduce the security exposures in such environments.

CCS CONCEPTS

• **Security and privacy → Software and application security**; *Domain-specific security and privacy architectures*; • **Artificial Intelligence → Distributed Artificial Intelligence**; *Cooperation and coordination*

KEYWORDS

Security Policies, Federated Systems, Generative Models

ACM Reference format:

D, Verma, S. Calo, and G. Cirincione. 2018. Distributed AI and Security Issues in Federated Environments. In *Proceedings of Smart and Connected Communities Conference, Varanasi, India, Jan 2018 (SCC'18)*, 6 pages.
DOI: TBD

1 INTRODUCTION

The emergence of technologies such as the Internet of Things [1] and Artificial Intelligence [2] has enabled the creation of many sophisticated solutions for cities and communities. Using instrumented sensors, solutions that improve the safety and security of a community, manage the traffic better, reduce pollution, and improve the quality of the environment can be readily created. All of these solutions can be abstracted as a set of devices assigned to a city worker or individual which are assisting the operation of the different processes involved in a city. At the current time, we can envision a smart phone carried by an individual to be a device which helps in the creation of a smart city, along with other sensing infrastructure in the city. The current ratio of devices to humans in any city is approximately 1:1 in most environments.

As the intelligence and autonomy of the sensors and actuators increase and the cost of the devices decreases, the ratio of devices to humans increases significantly. In the not too distant future, we would envision a situation where a single human being will be assisted by dozens and in some cases hundreds of devices. Each of these devices will be like a virtual assistant for the human, capable of making a limited set of decisions in order to assist the human being in their tasks. In effect, each human is supported by a large number of virtual assistants. Unlike current devices like smartphones which are essentially information processing instruments, future devices would be more like robots and drones which can change the physical environment around us. Autonomy has to be an essential part of such systems. One way to create such systems is to follow the development roadmap described in [3], where the systems are referred to as cognitive collaborative systems.

While the creation of such systems is technically feasible, they will work in the context of a human society whose governance processes, sociological aspects and human emotions change at a significantly slower speed. Due to that, future devices will need to operate in a federated environment, where security is a key concern.

A federated environment is an environment in which devices and humans belong to more than one organization. Almost all real-life environments require us to deal with such federation. A Smarter city is a federation of different agencies, e.g., the fire department, the traffic department, the sanitation department, etc. A large enterprise is a federation of several sub-groups, such as a product unit, a service unit, a human resources unit, a strategy unit and a research unit. When countries join together for a military operation, e.g., peace-keeping in a disturbed area, the operations are conducted by a coalition that consists of soldiers drawn from different countries. Federated environments are characterized by the fact that different organizations collaborate and cooperate with each other, but the trust among them is limited even during collaboration.

In the next section of the paper, we describe how we expect the devices in this futuristic environment to be operating. In Section 3, we discuss some of the typical security and safety

issues that arise. In Section 4, we describe AI based approaches that can provide mechanisms to address those concerns, followed by a description of future work in Section 5.

2 ENVIRONMENT DESCRIPTION

In the environment that we envision, each human is assisted by several devices with computing and processing capabilities. These devices are able to communicate with each other, learn from their environment, and modify it as needed. The devices could be: wearable devices, e.g., watches and health monitors; robots performing specialized functions, e.g., intelligent cleaners and maintenance checkers; appliances like intelligent refrigerators and dish-washers; self-driving mules and drones; or, any other collection of smart devices.

From an operational perspective, the devices are controlled by a human. The human in charge of the different devices expresses a desire. Each of the devices translates that desire into an action for itself, in a manner so that the actions taken together fulfill the desire of the human. In order to do so, the devices need to coordinate among themselves, which they do following some collaboration scheme. An example is shown in Figure 1, which depicts a hypothetical future situation of a human in a house which has an intelligent refrigerator, an intelligent oven, an intelligent microwave, and an intelligent drone. When the human expresses a desire like "make dinner", the different devices work together to determine an appropriate menu, the drone fetches items from the fridge to the stove and microwave, and the appropriate menu items are cooked and prepared for the human.

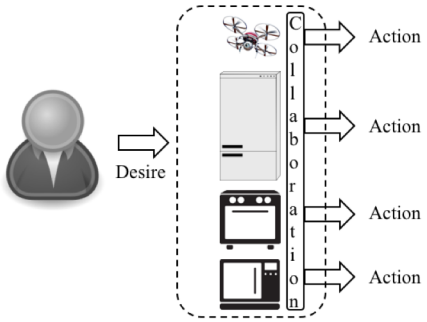

Figure 1: Mode of operation of devices in the system

The example above shows the scenario of a human in the home, but a similar scenario can be imagined for a human who is working as a city worker. A city employee charged with maintaining the roads would have different devices that monitor the state of the road. Some devices will bring in the repair material, and others will use that to repair the road. City employees charged with creating new homes would similarly have a fleet of devices at their disposal, and a policeman would have a separate set of devices for law enforcement.

Equivalently, one can imagine the military forces of tomorrow not being a set of soldiers with guns, but a set of soldiers supported by a large number of devices including drones that fly ahead to scout the surroundings, mules that carry supplies and other devices that can launch ammunition at the behest of the soldier. While the soldiers will be the ultimate decision makers, several of the supporting activities will be automated, including the driving of vehicles, deciding when to refuel, and how to schedule operations of different vehicles. The soldier may express a desire like "secure the perimeter of the rest stop" which would result in the assets for surveillance positioning themselves appropriately, and rotating their positions so that some assets may recharge themselves while others take over the surveillance. The human will be alerted when a possible intrusion of the perimeter is anticipated.

All of the scenarios will be running in a federated environment, i.e., humans belong to one or more organizations. Soldiers belong to one country, and they would collaborate with soldiers of other countries in their coalition, while they will be opposed to soldiers from yet other countries. City employees work in different agencies, and each agency has limited cooperation with other agencies. In those environments, dynamic communities of interest will continue to form, where people from different organizations cooperate together, or borrow devices from each other for some temporary period. Even in the household scenarios, humans may cooperate with their neighbors, and borrow assets from others as needed for specific tasks.

As a result, the environment looks like the abstract view shown in Figure 2. At any time, the devices that are working to

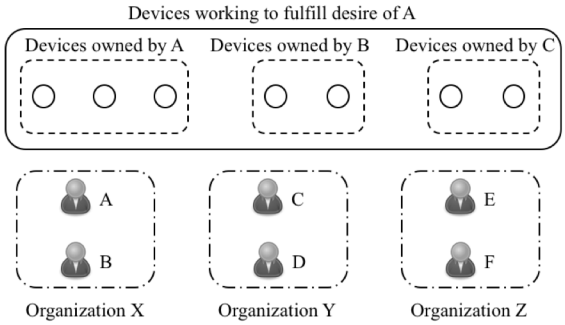

Figure 2: Devices, Humans and Organizations

fulfill the desire of a human (as shown in Figure 2) may belong to that human or they may belong to other humans. The other humans providing the devices may belong to the same organization, or they may belong to different organizations. Some of the organizations may be friendly and trusted (with varying levels of trust) while some of the organizations may be hostile, either overtly or covertly.

Each individual device can be modeled as a collection of services (or micro-services) that are running at the device. External devices can invoke those services provided they have proper authority to do so. That authority is determined by security policies which are generated automatically by each of the devices following the paradigm of cognitive generative policies [3].

The devices need to address the security concerns that arise in this environment in an autonomous manner. Note that while there are issues other than security that need to be addressed to make

the scenario envisioned above a reality, the focus of this paper is on the security issues that arise in this context.

3 SECURITY ISSUES

There are several security issues that can arise in the environment described previously. Since a group of devices involved in addressing the requirements of a specific desire may come from many different owners, either human or organizations, the set of devices need to have an efficient mechanism to decide which device is authorized to join in the group, and which is not. In many contexts, a new device may be encountered in the environment without proper authentication credentials, and the other devices need to autonomously decide whether or not to allow that unknown device in. Similarly, when an external human or device makes a request to the devices in the cohort, they need to decide whether or not to allow that request. Similarly, some of the actions that the devices may need to take may require checking with the humans, and the devices need to know when to do so and when to take an action on their own. In this section, we discuss some of these security issues.

3.1 Authentication

In order to be able to work together and collaborate, each device needs to know who are the other partners in the system that are working together with him to satisfy the desire that is expressed by the human. Let us define a cohort of devices as the set that is working together to satisfy the desire. A mechanism to decide which device is in the cohort and which one is not is needed. Sometimes devices may be borrowed from another individual or another organization. Thus, the system needs to follow a paradigm whereby a member that is entitled to be in the cohort is able to join the system easily, whereas those who should not be in the cohort are easily excluded.

One ways to approach the problem of cohort formation is to assume a central controller which is responsible for issuing credentials to each device, in a manner analogous to [4]. The controller provides a set of credentials to each member that is allowed to be part of the cohort.

Furthermore, each device in the cohort is required to run a discovery process by which it can identify and determine the presence of other systems. The discovery process could be supported by each device having a local service which would allow others to introduce their presence by invoking the service.

An integral part of the discovery service would be the need for continuous monitoring of the environment. While the devices in the cohort may follow the local standard of running a discovery service, there may be malicious devices which may be present in the environment but may not be advertising their presence. A mechanism to observe all the devices present in the environment, and observing them via visual, auditory and radio-spectrum monitoring needs to be put into place for each device (or at least a subset of devices). A device that is present in the local environment, but is not participating in the discovery process is possibly suspect, and needs to be dealt with in an appropriate manner.

Authentication among cohort devices needs to be done in a continuous manner, e.g., when the two devices come within a specific distance, or when the two devices try to communicate. Failure to authenticate is broadcast to all devices in the cohort, and devices failing multiple authentications may be quarantined.

3.2 Unknown Device

Monitoring of the physical environment may lead to detection of devices without the right credentials. Given the fact that the environment is a federated one, the mere presence of an unknown device which is not able to authenticate itself may not be dangerous in all situations. The unknown device might be a device belonging to the same human that has become damaged somehow, and is therefore unable to participate in the authentication and discovery process. It may belong to a friendly sister organization or to another human in the same organization, and just may not be able to participate in the discovery process and provide the authentication process due to some mechanical or procedural problem, e.g., its power may have run out, its communication module may have been damaged, or it may be new in the environment and not yet configured properly for participation in the cohort. It may also be a device belonging to a sister organization that is engaged in some other activity e.g. planning out how to fulfill another desire expressed by the human.

The unknown device could also be malicious, which implies that the devices in the cohort ought to be careful that their configuration is not being monitored, and be able to monitor the actions of other devices so that they can be classified as manifesting either a safe presence or a malicious presence. The physical markings on the unknown device may provide more information about its ownership, and may help in the classification. However, a better approach would be to monitor the behavior of the unknown device, and then use the behavior to determine the right security policies that should be used by the cohort. This would require extensions of approaches introduced in [5].

3.3 Information Protection

Physical devices introduce the challenge that information contained within them, if revealed to the wrong persons, can be used to create security attacks in new ways. A current risk from release of seemingly innocuous information can be seen in the information that is stored in current cars. GPS enabled car systems usually store the home information of the owner, and one of the security risks is that a criminal could steal the car, direct it to go to the home location which is likely to be unoccupied since the owner's car was at a remote location, and conduct a burglary.

If we extend the scenario where there are multiple devices working together in a cohort, the risk of revealing innocuous operation that can have a harmful impact becomes even bigger. As a result, each device needs to determine how much of the information it possesses ought to be revealed to an external entity, including both devices within the cohort, devices outside the cohort, and to humans in different organizations. We implicitly assume that the device would provide all information clearly to

the owning human, but it may need to obfuscate or modify the information when it is being provided to non-owner humans that are using it temporarily, or to an unknown human.

The decision on how information is provided depends on the context, and the device needs to determine the right information protection policy depending on the current context. If a non-owning human is trying to do something for the owning human, e.g., driving them home because the owning human is intoxicated or incapable of driving, the car should reveal the home location,; but, if the human is a suspected car thief, the home location should not be revealed.

3.4 Request Protection

Since we are modeling devices as a collection of micro-services, it is important to determine which device is allowed to invoke which micro-service. Depending on the specific desire from the human being fulfilled, and the plan of action that is created to fulfill that desire, the devices may need to determine who gets access to which micro-service running on each device. Since such access protection schemes become more complex as the number of devices increase, setting such protection rules would need to be done automatically, instead of being set manually by a user.

The approach for determining request protections automatically can be created as an augmentation to the architecture for generative policies [6] focusing on access control. While that architecture focused on access control for the network layer, analogous mechanisms can be used for protection at the services and micro-services layer. The requirement for the solution to work at a different layer would be to have the appropriate discovery mechanisms in place at the services layer which complement the network layer information. Furthermore, the action plan and the desire that is being satisfied needs to be provided so that the devices can automatically determine how to protect themselves and how to respond to access requests.

3.5 User Intervention

While the operation of the cohort of devices ought to be largely autonomous, there will be many situations where the human being is the only person who can make the determination of what is safe or not safe to undertake. An important issue in managing the security of the cohort is to decide when to ask the human or humans for input on any action to be taken.

The amount of intervention that needs to be put in place needs to be determined properly, and it would depend on the context. If too much user intervention is asked for, the system becomes burdensome. On the other hand, if too little user intervention is asked for, the system can commit mistakes which the user may not have desired.

One approach to address this would be for the devices to observe human behavior in past contexts, and then use patterns seen before to decide when to ask the user to intervene and when to have the human not intervene.

4 DISTRIBUTED AI FOR SECURITY POLICIES

The challenges described in the previous section can be addressed by using machine learning algorithms in a distributed manner. We can model the operation of each device in the system with respect to the associated security challenges in terms of policy based management, where each policy looks like an event-condition-action (ECA) rule [7]. Specifically, the different security issues can be addressed by enforcing the policies that are required in the manner described in the table below:

Issue	Event	Condition	Action
Authentication	Device Detection	Valid Credentials	Accept in Cohort
Unknown Device	Device Detection	Attribute Analysis	Accept in a category
	Device Behavior	Behavior Analysis	Adjust category
Information Protection	Access Request	Attributes of Requestor Context of Request	Obfuscation Algorithm to be used
Request Protection	Access Request	Attributes of Request	Allow/Deny
User Intervention	Pre-Action Check	Attributes of Action	Alert User/Execute Action

If we can determine the right policies for the system to support the security requirements, then the devices can enforce the right security policies themselves. The hard problem, however, is to generate the set of security policies that the devices ought to be using. In an automated environment with multiple devices, it is unreasonable to expect that a human would be able to define the right policies for each of the devices. Therefore, the devices need to be able to generate their own policies under loose high level guidance from the human.

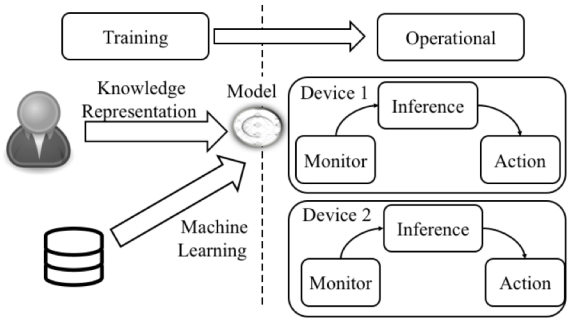

Figure 3: Typical Centralized AI Solution

In order to generate their own policies, we can adopt a variety of techniques, each of which leverages the concept that each device would monitor its environment and use that information to best estimate the policies that it should be using. The devices would be learning and sharing the policies learnt with each other.

This approach, where different devices learn from their environment independently, and share the learnt models with each other to improve them collaboratively, is an instance of distributed AI.

4.1. Distributed AI

The typical approach for building AI applications is illustrated in Figure 3. Applications consist of two different stages, a training stage, and an operational stage. During the training stage, a model is built. This model could be captured as a neural network, a decision tree, a statistical model, a set of rules, or the result of using any other appropriate modelling technique. For the purpose of security which is our focus in this paper, we can assume that the model being learnt is a set of ECA rules. During the operational stage, data is monitored from a source, e.g., to look for events or check for some conditions. The model is used to determine the action that is needed, and the appropriate recommended action is taken. During the training stage, a set of training data can be used to build the AI model automatically, which is usually referred to as machine learning. Alternatively, a human can encode and incorporate expert human knowledge into the model. Machine learning, knowledge representation, and expert systems are all subfields of AI.

In the centralized case, the models are created and used for inferencing at the same site. The data generated during monitoring can be used as training data for subsequent refinement and modification of the model. However, this approach requires a good network connection and movement of significant amounts of data to the location where training or knowledge representation can be done to create the right model.

In contrast, the distributed learning model has several devices learning the models on their own. They may be helped by an initial model that is provided by a human or machine learning process done at a central site. The initial model is modified or learnt anew by each device. Instead of exchanging data between different sites, the devices exchange models, which are iteratively learnt and refined. The setup is as shown in Figure 4, and each device has a significant degree of independence in learning and using its models. For the situation where the number of devices is large, distributed AI is the only practical approach to get autonomy for devices.

4.2 Learning from Human Behavior

Learning from human behavior is a good method for learning policies for asking for user intervention, and for reducing the user intervention amount required over time. Each device records instances where human intervention was called for, and notes the decision that was made by the human. The human input and the attributes of the action for which the intervention was requested are used to create a training set. Once a sufficient amount of training data has been accumulated, the system can build a model predicting human behavior; and, for the set of actions where the human behavior can be predicted with a high confidence level, human intervention can be replaced with the expected outcome of the learnt model.

4.3 Refinement under predefined Constraints

This approach is a good match for learning authentication, information protection, and request protection models. Each device relies on an initial set of constraints that is provided to it by a human or central environment. These constraints dictate how the device may be able to generate its own set of policies or rules, e.g., using a generative policy approach [3] [5]. The information provided as constraints might include policies about which data can be reused as training data, a set of guidelines on how to create the rules, and instructions on how the data that is seen can be converted to train a specific type of model. In some cases, even a complete initial model may be provided.

After the information is obtained, each device would create its

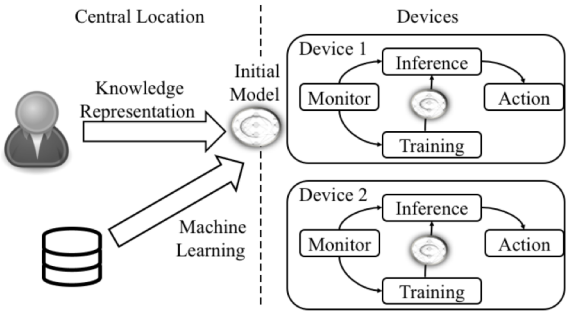

Figure 4: Typical Distributed AI Solution

own customized model for inferencing based on its characteristics. If only a set of constraints are provided, the device can generate a set of policy rules that are consistent with the constraints provided by the system. As an example, the constraints may specify that authentication must be used. On the basis of these constraints, and the capabilities available to the device, the device may create specific policies describing the type of authentication it should use with each of the other devices.

If an initial model is provided, the device can customize that model for its needs. The customization can consist of transforming the model provided to a more efficient format, e.g., converting a neural network to an equivalent set of decision rules, or replacing a neural network with a kNN clustering model. The customization can also include removing some classes or actions that may not be relevant to the needs of a device.

When the model provided is a statistical model that provides the correlation among different input parameters, the system can calculate the correlations it sees in the data that it observes. It could correlate the attributes of the devices that were previously granted access, and then use that to generate rules for devices that are not able to communicate or authenticate properly. In case the authentication fails, but the other attributes of the device match (e.g., it has physical markings that indicate it belongs to the organization), the cohort would let the device in and schedule it for maintenance and repairs.

4.4 Behavioral Analysis

Behavioral analysis creates policies dynamically based on the behavior of other devices, and is a good approach to learn policies that can handle unknown types of devices.

When an unknown device is observed, a default action would be to consider it malicious and untrusted. However, drawing a parallel from human group behavior, when a team of humans find another human, they do not automatically assume that the stranger is hostile. They observe the behavior of the newcomer, and depending on the actions being taken by the newcomer, gradually change the status and amount of trust placed in the newcomer. A similar model can be used for analyzing the behavior of unknown devices.

An unknown device in the cohort model would be a device which is not preauthorized to join the cohort but is around and is able to be discovered. Based on the amount of access that the device allows to other members in the cohort, and the behavior shown by the cohort members, the system can change the level of trust placed on the device and assign it to different roles when trying to fulfill a desire of the human.

4.5 Observation and Obfuscation

For the purpose of information protection, the devices in the cohort model cannot always be assumed to be friendly or neutral. A device may belong to a hostile organization or hostile human, and may be eavesdropping on the cohort. The system needs to guard against such devices, by including both encryption on the communications taking place, as well as controlling the information that is communicated. Since encryption takes power and cycles, the devices should be able to automatically set up policies that determine when to use encryption. When they are in the presence of an unknown device or areas where an eavesdropper can hide, devices should use encryption; while, in environments that do not have any hostiles, they may choose to communicate using more efficient clear-text protocols.

In some cases, the human desire would be for their cohorts to spy on other cohorts. In that case, the devices need to have the ability to observe other devices and infer what rules those devices are using. In the same spirit, the devices need to plan for the case where someone may be spying on them, and obfuscate their operations so that the eavesdropper cannot easily infer the policies they are using.

A similar model for obfuscation applies when information needs to be communicated to a human or another device. Each device has to decide whether the information needs to be modified before it is given out to the requesting device or human. In general, while precise information will be given out to the owning human, information may be obfuscated if it is requested by devices belonging to another organization or to an unknown human. Behavioral analysis can provide a way to generate such policies.

4.6 Learning by Sharing Models

One of the strengths of the cohort model is that devices can share policies that they have generated and learn from each other. As the devices use the different techniques above to create new policies for themselves they can also share them.

The sharing of models is most beneficial when a new device is added to the cohort. At this stage, the device does not have to learn all the policies from scratch, but can get the current policies that have been learnt by the existing devices in the cohort. This gives the new device the ability to benefit from the experiences of the previous members.

5 SUMMARY AND FUTURE WORK

In this paper, we have provided a high level overview of a future environment where devices far outnumber humans, discussed the security issues associated with such environments, and proposed a high level approach to address those security concerns. Our work is at an early stage, and we will be working to develop these ideas further to enable secure operation of devices in an environment that is going to soon become a reality.

ACKNOWLEDGMENTS

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.

REFERENCES

[1] I. Lee and K. Lee, "The Internet of Things (IoT): Applications, investments, and challenges for enterprises", *Business Horizons*, vol 58, no. 4, pp. 431-440, August 2015.

[2] H. Nakashima, H. Aghajan, and J. Augusto (Ed.). 2009. *Handbook of ambient intelligence and smart environments*. Springer Science & Business Media..

[3] E. Bertino, S. Calo, M. Toma, D. Verma, C. Williams, and B. Rivera, A cognitive policy framework for next-generation distributed federated systems: concepts and research directions. *In IEEE 37th International Conference on Distributed Computing Systems (ICDCS 2017)*, pp. 1876-1886.

[4] 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.

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

[6] D. Verma, S. Calo, E. Bertino and C. Williams, *Generative Policy approach for dynamic collaboration in coalition environments*, to appear In Proceedings of SPIE Defense + Commercial Sensing Symposium, Orlando, FL, April 2018. Jacques Cohen (Ed.). 1996. Special Issue: Digital Libraries. *Commun. ACM* 39, 11 (Nov. 1996).

[7] K. Twidle, E. Lupu, N. Dulay and M. Sloman, *Ponder-2 A Policy Environment for Autonomous Pervasive System*s, in Proc. IEEE Workshop on Policies for Distributed Systems and Networks, June 2008, pp. 245-246.