Probabilistic Graphical Model for Cyber Defensive Policy Assessment and Facilitation of Intuitive Optimal Policy Selection

Pranab Banerjee[*] Thomas Allen[†]

Abstract

The age of cyber warfare necessitates effective defensive plans for operational integrity of networked security assets. Under a cyber attack, a decision maker needs to select the most effective defensive action (policy) from a set of feasible policies brought forth by domain experts and/or automated policy generators. However, selecting an optimal policy is non-trivial in practice because of complex dependencies among constituent components of a critical operational system; temporally dynamic mission goals; and uncertain knowledge about the states of some components. To address these issues, a Bayesian network based probabilistic framework was developed to assess the impact of a policy on mission success. At the core is a probabilistic graphical mission model built on top of the assets terrain using domain knowledge. The framework quantifies the probability of mission success under a policy as a score, and intuitively explains the propagation of policy effects leading to the mission outcome, thus facilitating optimal policy selection. For a mission composed of temporally ordered sub-tasks, the Bayesian network is dynamically pruned based on currently completed steps.

Key Words: Cyber defense policy, Cyber mission model, Optimal policy selection, Probabilistic graphical model, Bayesian network

1. Introduction

Sophisticated cyber defense tools, capable of reacting in real time to cyber attacks, are in their infancy. One of the key challenges facing a cyber infrastructure defense team is determining how to dynamically deploy and control emerging cyber defense tools so as to ensure resiliency of a mission and avoid cyber fratricide. While many, if not most, cyber defense tools are designed to operate according to a set of (usually) pre-defined policies, these tools and policies are typically developed in isolation from one another and with the primary goal of addressing a single type of attack, or preserving a single type of process. Given the interconnectedness of our networks and the large number of dynamically changing missions, cyber commanders need the ability to manage multiple instances of multiple cyber defense tools to maintain mission resilience. Further, they need the ability to define and dynamically adjust the policies that control cyber defense tools. Currently, there is limited understanding of both (1) the impact of our cyber defense tools and resources on mission success, and (2) the impact of mission priorities and tasking on the behavior and use of cyber defense tools.

This paper describes a Bayesian network (Darwiche 2010) based probabilistic graphical model (Koller 2009) to capture the causal relationships among cyber assets to facilitate probabilistic inferencing in order to assess the impact of a cyber defensive policy on carrying out a desired mission. Decisions about how, when, and where to deploy cyber defense tools should be made in the context of the missions that the overall cyber system is supporting, taking into account mission needs (information, bandwidth, access to specific applications, etc.), priorities, and the cyber threat environment. Defensive actions may be proactive (e.g., active shaping of the cyber terrain to protect specific cyber resources in order to meet key mission needs) or reactive (e.g., responding to an attack in real time), with a goal of ensuring robust or resilient mission execution. The probabilistic graphical model developed here captures the priorities, goals, and cyber

[*] Boston Fusion Corp., 70 Westview Street, Suite 100, Lexington, MA 02421

[†] Boston Fusion Corp., 70 Westview Street, Suite 100, Lexington, MA 02421

system needs of on-going missions, and use that information to quantify the feasibility of a cyber defense policy (generated by a domain expert or an external policy generation engine) as the probability of mission success under the policy. This framework allows a cyber commander to control cyber tools in a way that is appropriate for the overall goals of a mission, and not merely in a way that is appropriate for the local goals of computer resource (platform, network, application, data) protection.

2. Probabilistic Graphical Model

Probabilistic Graphical Models provide an efficient way to make inferences under uncertainty through a combined application of probability theory and graph theory (Cowell 1999). These models have a rigorous theoretical foundation, and provide a formal mechanism for exploiting conditional independence among cyber terrain components (the nodes in the graphical model) for efficient computation of marginals of the joint distribution over all the components, the states of which are treated as random variables. This formalism is directly aligned with the domain of cyber defensive policy assessment where we are primarily interested in inferring the probability of the mission being successful under a policy, which essentially sets a specified set of cyber resources to desired states (ON or OFF). This inference task essentially entails marginalization of the joint distribution over the states of the mission and the relevant cyber terrain components that the mission depends on. A Bayesian Network is a class of probabilistic graphical model represented as a directed graph with explicit parent-child dependencies among the nodes. Such a parent-child dependency allows the network to model causal relationships which is important for the application domain of interest in this paper, since the goal here is to assess the impact of a set of cyber resources on a mission on the whole.

2.1 Bayesian Network for Policy Assessment

Availability of relevant resources in a timely manner is key to carrying out a cyber mission. All resources may not be equally critical, however, for completing a mission. For example, some of the resources may have redundancies which would make the mission more robust against the failure of such an entity. In some cases, the unavailability of a set of resources may not completely jeopardize a mission, but may result in partial failure or performance degradation. An effective way to quantify such a spectrum of outcomes is in a probabilistic setting where the probability of success of a mission is computed as an outcome based on the state of the underlying cyber terrain. Hence, a probabilistic graphical model encapsulating the causal inter-dependencies among the cyber resources and the overall mission, with the ability to assess the impact of the loss of a set of resources on achieving the goals of the mission is highly beneficial in exploring various resource allocation and management policies. In such a graphical representation, a directed edge is drawn from a vertex A to another vertex B if the state of B is influenced by the state of A. No edges are drawn between two nodes if there is no influence of either one on the other. The majority of such resource dependency graphs for a cyber mission are acyclic in nature. This motivated the use of a Bayesian network, based on a directed acyclic graph, for performing *what-if* analyses on a cyber terrain model should a set of resources become unavailable. This framework allows cyber defensive policies to be evaluated in terms of their consequences in the context of a mission.

The estimation of the effect of availability of resources on the completion of a mission can be cast as a top-down or causal reasoning problem where the states of the resources (the nodes that cause an outcome - the mission status) are specified and the effect (the probability of success of the mission) is to be inferred. A natural way to formulate this task is to treat each of the resources and the goals of the mission as random variables, and define causal relationships among them. If $\mathcal{R} = \{r_1, r_2, \ldots, r_{N_R}\}$ denotes the set of N_R resources, and M denotes the mission, then we are interested in estimating the probability $P(M|\mathcal{R}')$ for some $\mathcal{R}' \subseteq \mathcal{R}$. The obvious way to compute $P(M|\mathcal{R}')$ is to compute the joint distribution $P(M, \mathcal{R})$

and then marginalize over $\mathcal{R} \setminus \mathcal{R}'$. In practice, N_R can be quite large, making the direct marginalization over the joint distribution computationally impractical. However, this computation can be simplified by taking advantage of any conditional independencies among the resources. An efficient way to compute the probability of a mission conditional on a set of resources is through belief propagation in a Bayesian network model exploiting all such conditional independencies.

In our approach, a Bayesian network, which is a directed acyclic graph is used to model the relationships among the resources and the mission tasks in an intuitive way, where a directed edge $A \rightarrow B$ represents the relationship that the cyber resource B depends on the resource A. Nodes A and B are conditionally independent given a node C if $P(A|B,C) = P(A|C)$. This is denoted by the notation $A \perp\!\!\!\perp B|C$. In this case, no edges are drawn between A and B. This graph structure is embedded in a probability space to quantify the effect of a parent node on a child by associating a conditional probability distribution with each child node in the graph. The simpler the structure of such a graphical model, the simpler will be the inferencing process. The current research focuses on resource dependencies without any cycles, which cover the overwhelming majority of cyber missions. Hence, acyclic constraint of a Bayesian network does not pose any limitation. In our application domain, the number of vertices in the graph is dictated by the number of resources and subtasks, and hence fixed for a given cyber terrain. Hence, the only way to simplify the graph structure is by reducing the number of edges when appropriate, and the key is to exploit all conditional independence properties. To this end, it is assumed that the graphical model satisfies the Markov condition

$$v_i \perp\!\!\!\perp nd(v_i)|pa(v_i) \forall v_i \in \{M, \mathcal{R}\} \tag{1}$$

where $nd(v_i)$ denotes the non-descendant nodes of the vertex v_i in the Bayesian network, and $pa(v_i)$ denotes the parents of v_i. Equation (1) states that any node in the network is independent of its non-descendants given its immediate parents. This is a reasonable assumption to make for the resource models in the current application domain. These independence properties influence the structure of the Bayesian network. The dependency structure of the Bayesian network allows us to factorize the multivariate joint distribution $P(M, \mathcal{R})$ into univariate statistics (the conditional probabilities of the nodes), and the causal Markov condition (Equation (1)) significantly simplifies the specifications of these node level conditional probabilities. In essence, we can now compute the joint distribution as

$$P(M, \mathcal{R}) = \prod_{i=1}^{N} P(v_i|pa(v_i)) \tag{2}$$

where $\{v_i\}_{i=1}^{N}$ is the set of all the nodes in the Bayesian network, and $pa(v_i)$ refers to the immediate parents of the node v_i.

2.1.1 An Example Scenario

Figure 1 illustrates a Bayesian network for a simplistic and fictitious cyber mission. In this scenario, a company has offices in San Francisco and Boston, with their main file server located in Boston. They also maintain a backup file server on the cloud. The mission is to access a file using a computer at the San Francisco location. The Bayesian network in Figure 1 models this mission. This is a directed graph with arrows pointing from parents to children where a child node is functionally dependent on its parents. Thus, the directed arrow from the node “Internet” to “BOS Router” implies that the router at the Boston office depends on its access to the Internet in order for it to be functional. The cyber assets are shown as ovals, and logical operations are depicted as gray double circles. The node corresponding to the mission is a rectangle. By default, all the assets are shown in green, indicating that they are all fully functional. If any of the nodes become inoperational, they are colored red.

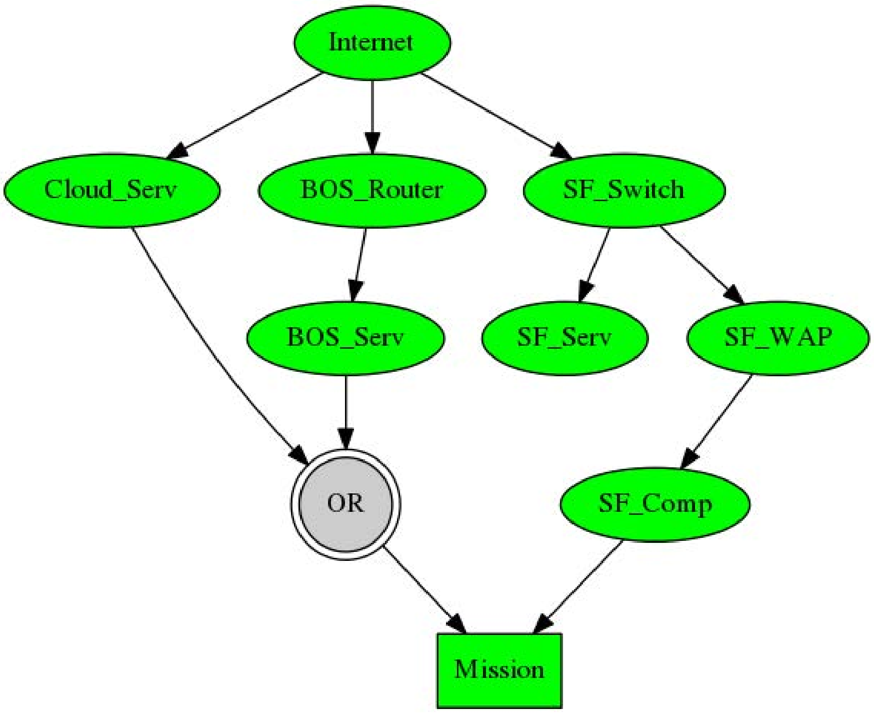


Figure 1: Example of a simplistic mission

According to this network, in order for the above described mission to succeed, the computer being used to support the mission (node "SF_Comp") has to be functional, as well as either the cloud based backup file server (node "Cloud_Serv") or the file server at the Boston office (node "BOS_Serv") has to be operational. These parent nodes of the mission in turn have their own dependencies. For example, the functionality of "SF_Comp" depends on the state of the wireless access point in the San Francisco office, represented by the node "SF_WAP". Using the Markov property of Equation (1), the joint probability distribution

$$P\Big(Internet,\ Cloud\_Serv,\ BOS\_Router,\ SF\_Switch,\ BOS\_Serv,\ SF\_Serv,\ SF\_WAP,\ SF\_Comp,\ Mission\Big) =$$

$$\Big(P(Internet).P(Cloud\_Serv \mid Internet).P(BOS\_Router \mid Internet).P(SF\_Switch \mid Internet).$$

$$P(BOS\_Serv \mid BOS\_Router).P(SF\_Serv \mid SF\_Switch).P(SF\_WAP \mid SF\_Switch).(SF\_Comp \mid SF\_WAP).$$

$$P(Mission \mid Cloud\_Serv, BOS\_Serv, SF\_Comp)\Big)$$

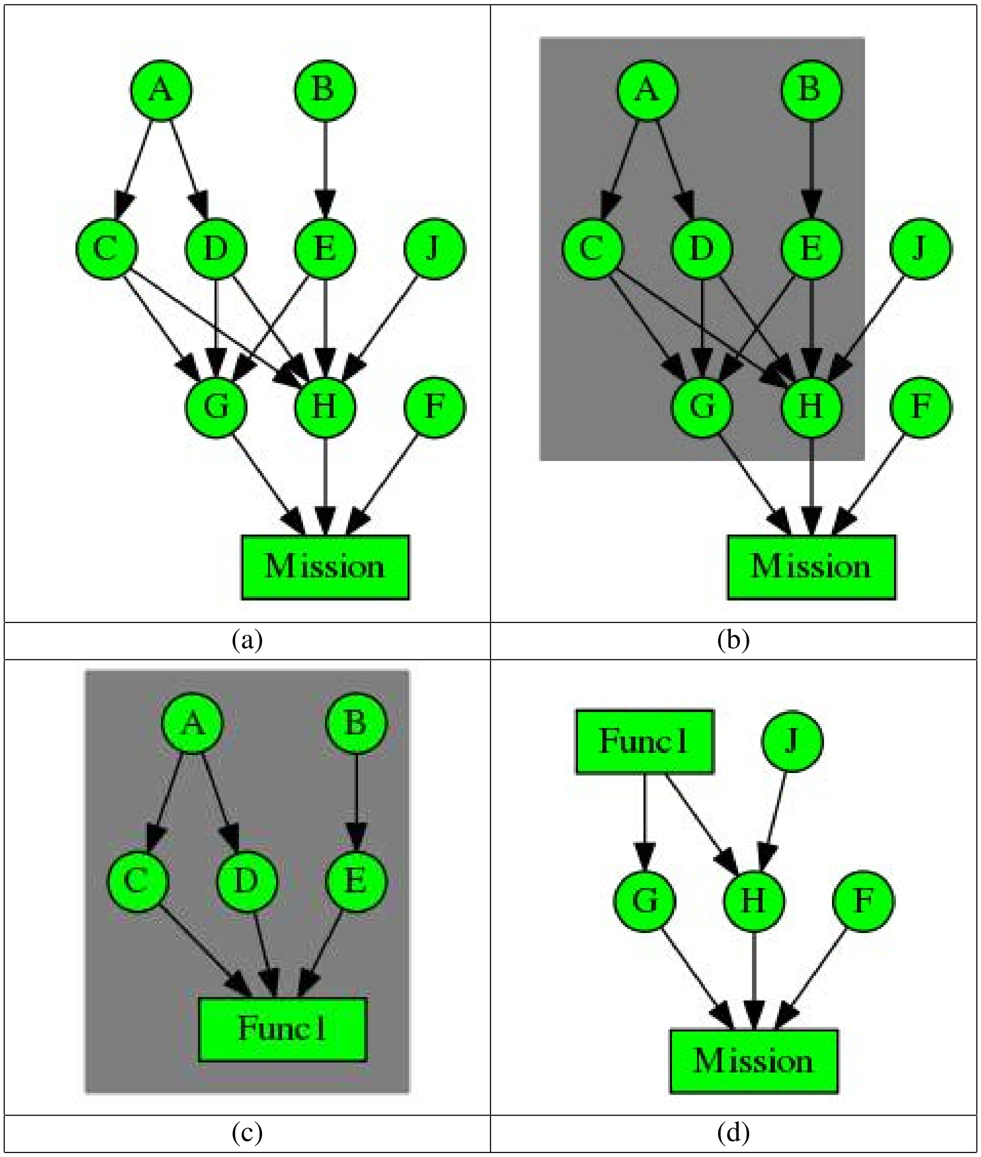
2.2 Bayesian Network Construction

The Bayesian network describing a cyber mission model over a specified cyber terrain is constructed based on knowledge of domain experts. The dependencies among the raw cyber resources relevant to a mission are obtained from domain specifications. The state space for each resource is assumed to be discrete and finite. A majority of the resources have a binary state space, and they can either be **ON** or **OFF** corresponding enumerated values of 1 and 0 respectively. However, arbitrary number of states are allowed. This is useful for modeling system degradation, where a resource may not be fully operational (**ON**) or completely inoperational **OFF** but may function at a reduced level of efficiency.

There are two main steps to defining a Bayesian network to model a cyber mission. The first step is to define the causal dependencies among the resources. This determines the topology of the network by constructing directed edges between (*parent*, *child*) tuples where the parent has a causal effect on the child. The second step is to define the conditional probability table (Darwiche 2010) for each node, which encodes the strengths of the causal influences of all the parents on the node. In our case, both of these steps are carried out using the experience and expertise of a domain expert, since usually there is not enough relevant data available for automated structure learning for the Bayesian network.

2.2.1 Modular Bayesian Network

In a complex Bayesian network, it is not uncommon to find multiple nodes dependent on a set of common ancestral nodes in an identical way. Figure 2 illustrates a case.



(a)

(b)

(c)

(d)

Figure 2: Example of object oriented Bayesian network

Fig 2(a) shows a directed graph representing a mission model consisting of resources shown in green circles. Here, the nodes **G** and **H** have common dependencies as shown in the grey box in Figure 2(b).

They both depend on the resources **A**, **B**, **C**, **D** and **E** in the same way. When such a common dependenciy is frequent enough in a Bayesian network, it is encapsulated as a higher level function module, and the Bayesian network is subsequently described in terms of these high level modules. Figure 2(c) shows a function module called **Func1** that encapsulates the common dedpency depicted in Figure 2(b), and Figure 2(c) shows the original Bayesian network in Figure 2(a) redefined in terms of the function module **Func1**.

Thus, the original causal network is decomposed into a set of logical functional units, and a set of raw resources that do not fit into any functional unit. The functional units constitute high level building blocks for the model. Such a decomposition allows higher level description of the mission in terms of these functional units, resulting in significantly more compact representation of the model compared to one in terms of raw resources only. Such a high level representation is also easier to comprehend and debug since the higher level functions resulting from the aggregation of a set of raw resources abstract away low-level non-intuitive inter-dependencies. In a complex dependency graph, it is also likely that the graph obtained by the first level of modularization manifests potential for defining even higher level functions composed of the first tier function modules and/or raw resources. The Bayesian network framework in our implementation allows arbirary levels of nested function hierarchies for a fully generalized network structure specification.

In addition to the above mechanism, higher level function blocks are also defined to encapsulate functions that can be used to abstract away many behind the scene dependencies that may not be directly relevant in defining a specific mission model.

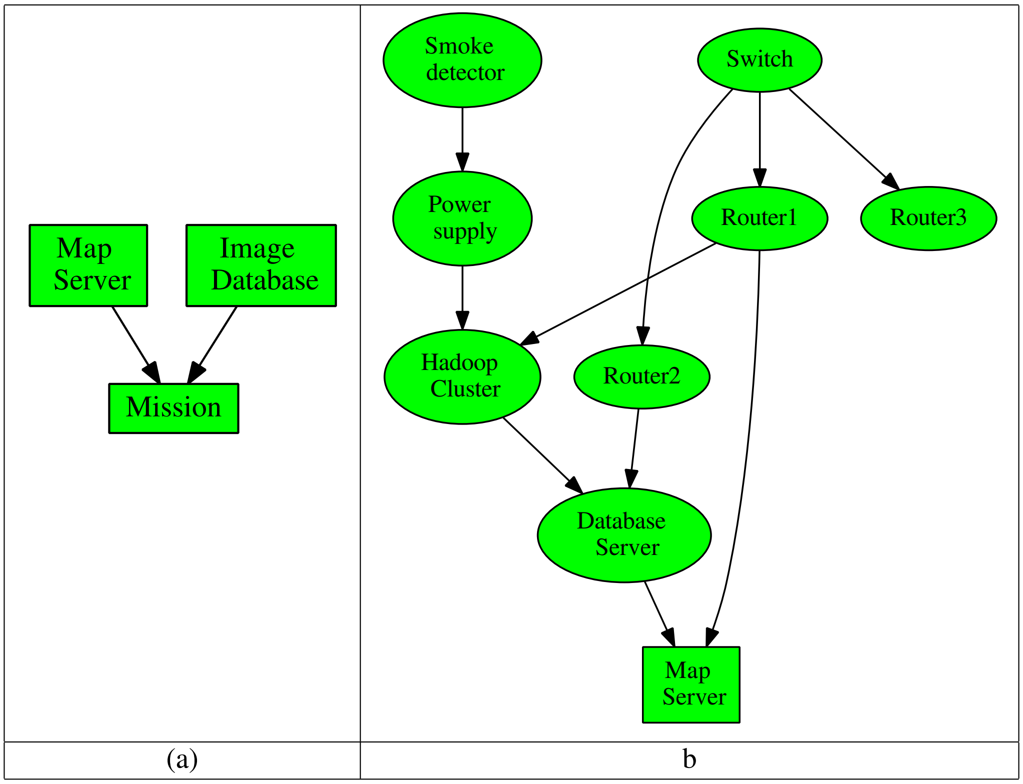


Figure 3: Example of abstraction of low level dependencies via a high level function. (a) a high-level Bayesian network describing a mission model, (b) expanded view of the high level function *Map Server* used in (a).

For example, consider a mission where the goal is to perform geo-spatial analysis of an image that requires the availability of a map server and an image database server. The Bayesian network in Figure

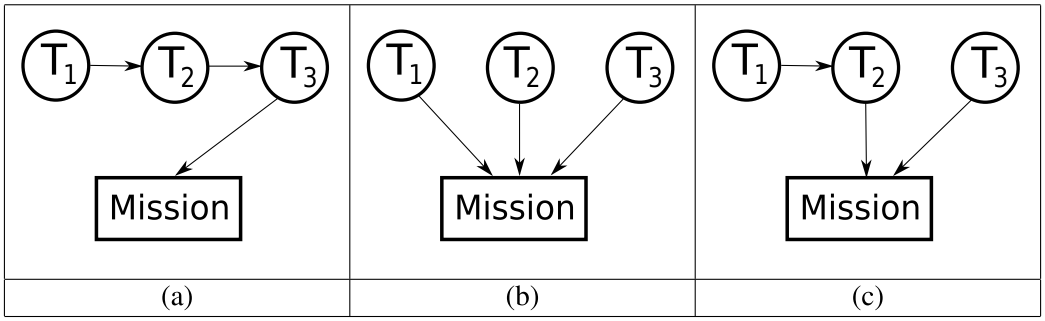
3(a) represents this mission model. Figure 3(b) shows the Bayesian network behind the high level function *Map Server* used in the mission model. Since the *Map Server* is likely to be useful in other missions as well, defining this high level function allows a compact representation of any mission that depends on the *Map Server*. It is envisioned that a library of such re-usable high level functions would be developed for an application domain based on the expertise of domain experts. The construction of the high level functional modules is currently more of an art than science. The determination of the functional granularity is subjective and is determined manually.

2.3 Inferring The Impact Of A Policy

The goal of the Bayesian network presented here is to quantify the impact of a cyber defense policy on a mission, and compare a set of policies in order to select the one that is most suitable to achieve the goals of the specific mission at hand. A policy is defined as state specifications for a set of cyber resources that the mission depends on. For example, if an attack on a web server is detected, a policy may be to turn off web services on the host machine. In the current context, the state of a resource is binary. It is either On or Off. The Bayesian network allows incorporation of contextual knowledge about any resource, such as its reliability, via the conditionanl probability tables or the probability of a root node to be in an ON state. Also, the framework is perfectly general to allow incorporation of resources with an arbitrary number of discrete states. Once the Bayesian network is constructed for a given mission M that depends on a set of resources $\mathcal{R} = \{r_1, r_2, \ldots, r_{N_R}\}$, the impact of a policy that requires setting the state of a subset of resources $\mathcal{R}' \subseteq \mathcal{R}$ is the probability of mission success $P(M|\mathcal{R}')$. This is computed using the belief propagation algorithm proposed in the seminal work by Pearl (Pearl 1982) and extended by Kim and Pearl (Kim 1983). This probability is thresholded by a domain dependent threshold parameter to flag the mission as failed or successful. The probability of mission success is the metric used to compare different policies and select an optimal one.

2.4 Incorporating Current State of The Mission

In many cases, a mission can be decomposed into a set of functional sub-tasks, where a mission is considered to be complete once all the sub-tasks are completed. These sub-tasks may have sequential temporal dependence; may be independent of each other; or the mission may depend on a mixture of both types of subtasks.



(a) (b) (c)

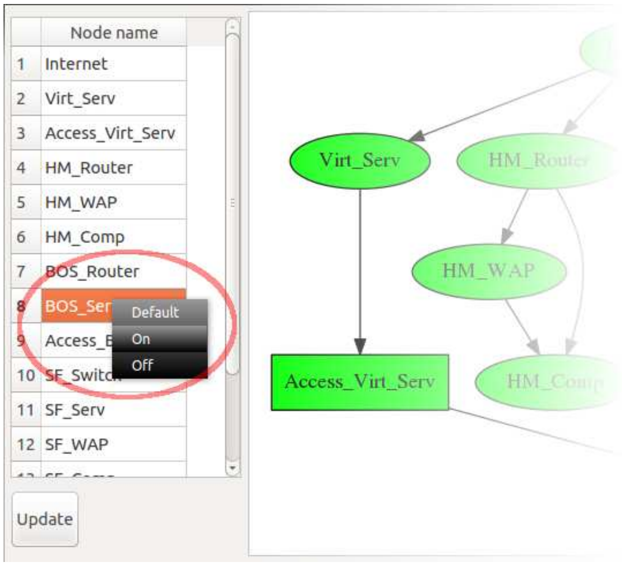
Figure 4: Dependencies of a mission on parent sub-tasks

Figure 4 illustrates these different scenarios. In Figure 4(a), the mission is decomposed into a set of sub-tasks with sequential dependence. The sub-task T_3 depends on the sub-task T_2, which in turn depends on T_1. The overall mission depends on T_3. Figure 4(b) shows a case where the mission depends on a set of sub-tasks T_1, T_2, and T_3 that are independent of each other. Figure 4(c) is an example of a mixed case, where the subtask T_2 is dependent on the sub-task T_1 but T_3 is independent of T_1 or T_2.

In our framework, a dynamic status is maintained of all the sub-tasks, and the Bayesian network for the entire mission is pruned by collapsing the sub-network corresponding to a completed sub-task to a single node having the state of that sub-task (success or failure). For example, if the task T_2 in Figure 4(b) is already completed and it failed, then the subgraph $G(T_2)$ corresponding to the dependencies of T_2 on cyber resources is collapsed to the single node T_2 with its status. This reduces the complexity of computing the impact of a set of resources on the overall mission, because, from now on, any resource r_k such that $r_k \in G(T_2) \setminus \Big(G(T_2) \cap \big(G(T_1) \cup G(T_3)\big)\Big)$ can be ignored in terms of its influence on the future of the mission.

3. Results

A graphical user interface (GUI) was developed to allow a user to interactively explore the impact of any policy from a feasible collection of policies generated by an external policy generation engine or by a domain expert. A policy essentially corresponds to setting the states of a set of cyber resources to desired values. The GUI allows one to set the states of any set of cyber resources and analyze the effect on the overall mission.



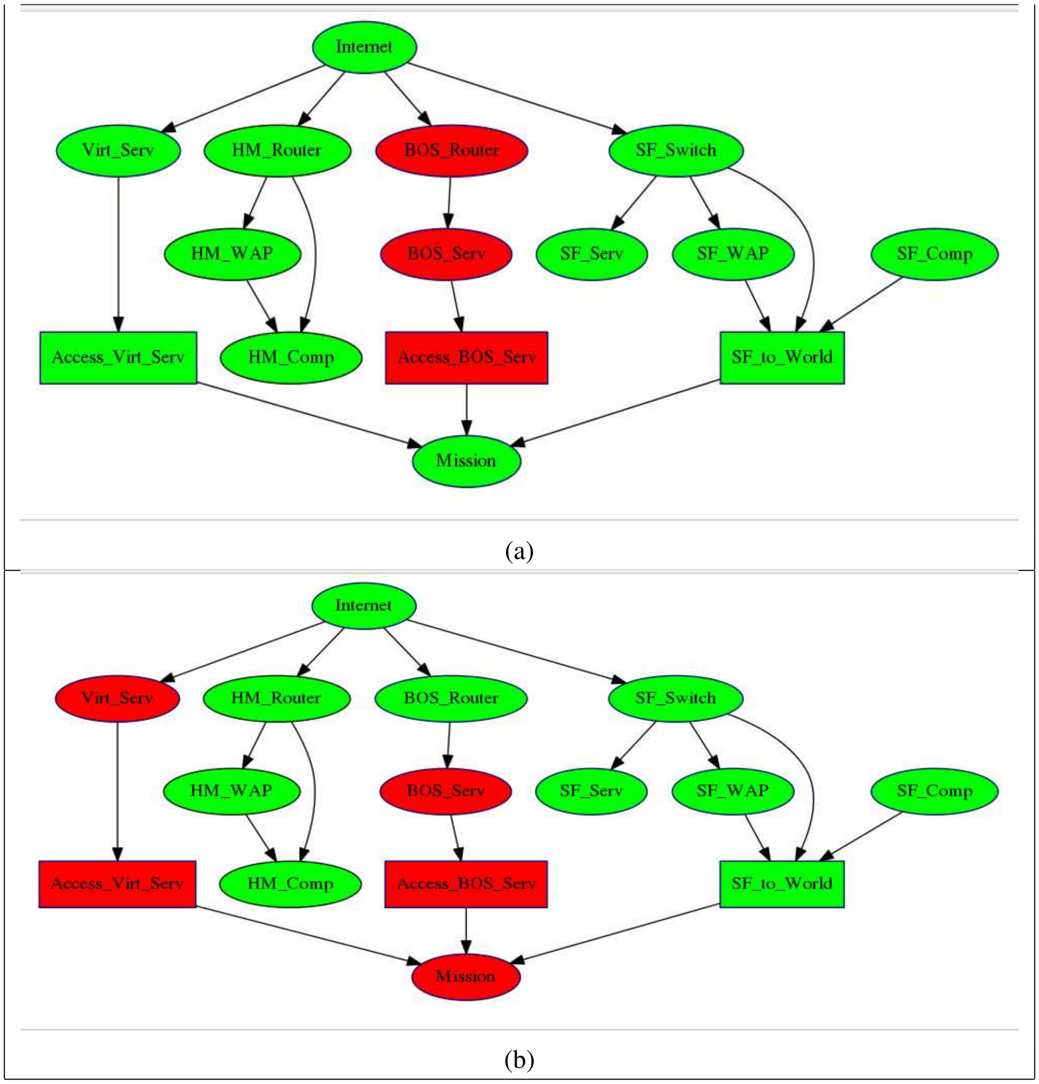
	Node name
1	Internet
2	Virt_Serv
3	Access_Virt_Serv
4	HM_Router
5	HM_WAP
6	HM_Comp
7	BOS_Router
8	BOS_Ser
9	Access_E
10	SF_Switc
11	SF_Serv
12	SF_WAP

Figure 5: GUI to set the states of resources

Figure 5 shows this user interface. The set of available cyber resources relevant to the Bayesian network is shown on the left. Right clicking on a resource displays the available states in a pop−up window (as shown in the region marked with a red ellipse in the figure), from which a desired state can be selected by clicking on it. Once the states of all the desired resources are set to relevant values corresponding to a policy, the effect of this policy can be visualized by clicking the "Update" button below the list of resources.

Figure 6(a) and 6(b) show examples of such exploratory visualization for a sample mission that succeeds if the function module *SF_to_World* succeeds, and at least one of the function modules *Access_BOS Serv* and *Access_Virt_Serv* succeeds.

Figure 6(a) shows the effect of a policy that requires turning off the node named *BOS_Router*, which causes the node *BOS_Serv* and the function module *Access_BOS_Serv* to fail, as indicated by the red color for these nodes. However, the mission itself succeeds because the function modules *SF_to_World* and *Access_Virt_Serv* reach successful states. Figure 6(b) shows the effect of turning off the nodes *BOS_Serv* and *Virt_Serv*. In this case, the overall mission fails and the reason is clearly visualized as the propagation of failures along red colored paths leading to the *Mission* node. The real value of this exploratory tool is appreciated for more complex networks where it is practically impossible to manually ascertain the impact of setting the states of a set of resources to specific values.



(a)

(b)

Figure 6: Visualization for *if-then* exploratory analysis

4. Conclusion

This paper has presented a probabilistic graphical model based on a Bayesian network to capture the interdependencies among the cyber resources that a given cyber mission depends on. The directed acyclic graph based representation makes it intuitive for human cognition to comprehend the inter-dependencies, and the probability space imposed on the graph facilitates a disciplined way to quantify the impact of a cyber defense policy on a mission. Under adversarial cyber attack, usually there is more than one defensive policy that can be adopted to mitigate the immediate risk, but not all of them have the same impact on the mission as a whole. Some policies may cause unnecessary cyber fratricide, hence jeopardizing the mission, while others may be overly conservative and cause avoidable delays, thus making the mission fail. The probabilistic framework presented here allows an analyst to quantify the impact of a policy on mission success and thus prioritize the different feasible policies in the context of the mission at hand. The user interface allows an analyst to not only quantify the impact as the probability of mission success, but it also allows one to visualize the chain of events that affect the mission in an adverse manner.

The directed acyclic graph based model used here addresses the vast majority of cyber missions in practice. However, there are rare situations where cyclic dependency relations need to be taken into account. This is the focus of our future research. The goal is to expand the framework to relax the constraint of acyclic dependency without sacrificing the advantages of a probabilistic graphical model.

5. Acknowledgment

The authors gratefully acknowledge the support of the Air Force Research Laboratory (AFRL) & Defense Advanced Research Projects Agency (DARPA) under Contract No FA8750-14-C-0281 for this research. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Air Force Research Laboratory (AFRL) & Defense Advanced Research Projects Agency (DARPA).

REFERENCES

Cowell, R. G., Dawid, P., Lauritzen, S. L., and Spiegelhalter, D. J. (1999), *Probabilistic Networks and Expert Systems*, Springer-Verlag, New York.

Darwiche, A. (2010), *Bayesian networks*, Communications of the ACM, Volume 53 Issue 12, December 2010, pp 80-90.

Kim, J. H. and Pearl J. (1983), *A computational model for combined causal and diagnostic reasoning in inference systems*, Proceedings of the Eighth International Joint Conference on Artificial Intelligence. IJCAI-83: Karlsruhe, Germany. pp. 190193.

Koller, D. and Friedman, N. (2009), *Probabilistic Graphical Models: Principles and Techniques*, The MIT Press.

Pearl, J. (1982), *Reverend Bayes on inference engines: A distributed hierarchical approach'*, Proceedings of the Second National Conference on Artificial Intelligence. AAAI-82: Pittsburgh, PA. Menlo Park, California: AAAI Press. pp. 133136.