Automated Risk Assessment of Sensor Information Disclosure in Coalition Operations

Marco Carvalho, Carlos Perez and Jeff Bradshaw
Institute for Human and Machine Cognition
15 SE Osceola Ave., Ocala FL 34471, USA
{mcarvalho, cperez, jbradshaw}@ihmc.us

***Abstract*—In this work, we introduce a new approach for automated risk estimation of source disclosure in tactical coalition operation environments. Our approach seeks to mitigate the unintended disclosure of data sources (i.e. sensors) and sensor capabilities through the selective release of information to other parties in a tactical network. While the proposed approach can be used for the automatic filtering of information based on pre-defined levels of acceptable risk, we introduce an application where the risk estimation is defined in terms of likelihood of unintended source disclosure and presented to a human for decision support.**

We experimentally demonstrate our approach on simulated network environments based on military exercises conducted at the U.S. Army National Training Center. Our current results were based on a subset of a broader exercise (161 nodes) with a ratio of approximately 30% of the nodes representing the sensor field. The results shown in this paper are focused on risk reduction as a function of the level of abstraction of the information disclosed, and assume a fully informed and rational adversary.

I. Introduction

In mission-critical and tactical operational environments, the timely collection, processing and dissemination of information is a critical capability and a differentiating factor for mission success. Generally, in such environments, information is gathered from multiple sensor networks highly diverse capabilities and levels of security classification. Once fused and classified, the information is then disseminated to the appropriate interested parties, including coalition partners and collaborators.

The dissemination of different types of information must be regulated and properly constrained, which is normally done through filters and access control policies that define which partners or collaborators have access to what kind of information. This is to protect not only access to potentially classified or restricted information, but also to protect the presence of classified sensors or capabilities.

While previous research efforts on automated (or semi-automated) algorithms for information release have primarily focused on the filtering of information provided by classified sources (i.e. the classification of information based on its source), it is possible that even under the assumption that such constraints are not violated, the release of any kind of information may lead to the unintended disclosure of potentially compromising details about restricted sensors.

In this work we introduce a new approach designed to estimate the risk of unintended sensor capability disclosure associated with information release. The goal is to provide an risk estimate that can be used by a information manager (human or software) to quickly determine the help mitigate the unintended disclosure of data sources (i.e. sensors) and sensor capabilities through the selective release of information.

Our focus scenario is a tactical sensor network deployed in a military operational environment where all challenges associated with node mobility and dynamic topologies play an important role. We consider a decision maker in the loop, who can make determinations about the appropriate release of information for a given estimated risk. For a human-in-loop, the estimate risk matrices will be provided in the form of a table, and for automated decision processes, the risk can be used as thresholds for different information release rules.

Our review of the related literature focused on the general methods for privacy preservation and unintended information, although we will also identify

some of the non-filter based strategies for selective information release. After introducing and describing our proposed approach, we will demonstrate its application on synthetically generated datasets for different network configuration and distribution of sensors. We also present preliminary results of our approach on networks created from military exercise datasets collected from the Army National Training Center.

Underlying our approach, is a probabilistic estimation of risk of information disclosure through association, assuming a fully informed 'adversary'. This is the contribution of our work, and while our motivating scenario and focus of discussion will be based on a military operational scenario, the proposed methodology can be easily extended to other domains and applications.

II. Problem Description and Motivation

Let us consider a scenario where some sensors are deployed in an geographical area. Under a common administrative domain, the sensor field will collect, process and distributed tactical information to a coalition of partners. The sensors may have different resources, capabilities or levels of classification. Sensor information is aggregated and used to detect and identify the classes of targets or entities in the area.

Let us also, consider that some of these sensors or their capabilities are classified, and restricted to a sub-set of the adversaries, or coalition members. Their presence and/or location cannot be disclosed to the agents to whom you are providing the information to (i.e. soldiers or allies).

The release of information collected by the sensors may unintentionally compromise policies by unnecessarily disclosing enough information that would allow an adversary or unauthorized partner to correctly infer the existence, capabilities and/or location of the classified sensors. Conversely, the blocking of critical information to prevent unauthorized correlations or inferences could greatly compromise the mission.

Finding a balance between information release and capability disclosure is not a simple task, and often relies in subjective human judgement and experience. The challenge, however, is further aggravated by the complexity of overlapping and heterogenous sensors networks in coalition operations or adversarial environments. The scale, dynamism and complexity of such environments often makes impossible for decision makers to properly estimate the higher order disclosure effects and risks associated with the release of sensor information, even when properly authorized for the consumer.

In this work, we propose an automated decision support system for the risk boundary estimation of unintended sensor capability disclosure. The goal is to provide human decision makers with worst case estimates of trade-offs between the level of specificity in information released to coalition partners, and the likelihood of sensor capability disclosure.

Our approach leverages our previous work on distributed resource discovery and data dissemination infrastructures [1] [2], to automatically build an maintain a map of the capabilities, resources and locations of sensors in a complex sensor network. That information is combined with an ontology of the military field equipment used to provide a hierarchical organization of objects in the field and to offer generalized descriptions of specific entities.

In this paper, we introduce an automatic approach for sensor disclosure risk estimation. We propose the automatic construction of a Bayesian Network (BN) from an equipment ontology and sensor resource capabilities. The Bayesian Network proposed in this work will be dynamically updated with changes or resources or topology in the field, and will be used both to estimate the disclosure risks associated with information release, and to recommend generalized descriptions of objects detected in the field to satisfy given risk requirements.

III. Related Work

Preserving privacy while revealing information has been a concern for decades in statistics [3], [4] and more recently in data mining [5]. In statistics, this problem has received several names: Statistical Disclosure Control (SDC) [6], Statistical Disclosure Limitation (SDL) [7] and inference control [8]. In data mining, this problem is called Privacy Preserving Data Mining (PPDM) [5]. SDC, SDL and inference control attempt to control the risk of disclosure of information about specific individuals from statistical summary results or aggregates. SDC has developed theories that attempt to characterize

the trade off between privacy and usability of statistical databases, and it has found boundaries for the minimum level of noise required to ensure a minimum level of privacy [9], [10].

PPDM attempts to develop algorithms and techniques for extracting knowledge from large amounts of data while protecting sensitive information. PPDM has developed some techniques and definitions of what composes a privacy preserving disclosure of information. PPDM mainly focuses in two aspects of information disclosure: disclosing data to be used by data mining algorithms (also known as microdata), and disclosing results of data mining algorithms, also known as aggregates. PPDM has develop definitions for privacy requirements when publishing microdata (k-anonymity [11], ℓ-diversity [12], t-closeness [13] and m-invariance [14]). And it has also created metrics for quantifying the protection provided, the failure to hide and the data quality of different privacy preserving data mining algorithms [15], [16].

In the context of the present paper, we are interested on privacy protection when disclosing aggregates. An attacker using information from aggregates might be able to create an inference channel that could allow him/her to determine the identity of individuals [17]. We propose using a BN for aggregating the measurements from different sensors to infer the presence of an entity, and also to assess the risk of disclosing the presence of those sensors by disclosing the presence of the inferred entity. In [18], the authors also propose using a BN for disclosure control, in the context of query restriction through auditing. The BN is used for representing the user's knowledge about private associations in the data after a sequence of min and max queries, this is, queries about the minimum and maximum elements of a subset of the database. Then, after making probabilistic inferences on the BN, the auditor of the database decides whether or not to answer the query, if it is determined that responding the query creates a privacy breach.

In [9] and [10], the authors found theoretical boundaries for the level of noise that needs to be added to the results when the results are represented as sums of the entries in the database. They define a statistical database as a query-response mechanism that allows users to access its content via statistical queries. They focus on binary databases, where the content consists of n binary entries. A statistical query on the database specifies a subset of those entries, and the answer is the number of entries having value 1 among the entries specified in the query. Before returning the answer to the user, the result is perturbed by a database curator that will add some noise to the result. In this scenario, the main question is how much noise needs to be added to the results from the queries so an attacker cannot reconstruct most of the database.

In [9], the authors showed that a polynomial adversary, this is, an adversary that can make a polynomial number of queries to the database (polynomial in relation to the size of the database), can potentially reconstruct most of the database, if the perturbation of the results is lower than $\Omega(\sqrt{n})$. The strategy consists on finding the candidate bit string that minimizes the error for all the results already returned by the database. In [10], the authors go even further, and propose an algorithm that requires only a fixed number of queries for each bit revealed.

The problem of disclosure control proposed in the present paper can be partially modeled by the representation proposed in [9], [10]. The bits in the database could represent the presence or absence of a sensor in the terrain. The attacker should try then to determine this sequence of bits. The main difference is that queries performed in the database cannot be modeled as sums of bits, but instead as an inference performed with those sensors. It is still an open question to us, if the polynomial adversary proposed in [9] can also be used in this scenario. It is still unclear to us if the assumptions needed for the proof of the algorithm, also apply in this new scenario.

The notion of disturbance in the results proposed here is different from the one proposed in [9], [10]. While in [9], [10] models the disturbances as noise added to the response, here we propose modeling the disturbance as providing a more general answer to the user, which even being a more general answer is still a correct answer.

IV. Proposed Approach

Our approach will be based the automatic construction of a Bayesian Network (BN) from sensor

information collected form the field, and target descriptions that can be used to abstract and generalize target detections. The BN, in our approach, is used for both to classify a target given signal provided by multiple sensors, and also to assess the risk of sensor exposure by disclosing the target information. Furthermore, the same model can be used to propose alternative descriptions (generalizations) of the target to reduce the risks associated with sensor exposure.

A Bayesian Network is a directed acyclic graph where nodes represent variables and links represent dependency relationships among the variables. Each node is annotated with a conditional probability table or distribution for the values of the node given the values of its parent nodes.

The relationships in a BN follow a special property known as the Markov condition, which states that a node's probability distribution is independent of its non-descendant nodes given its parents. The Markov condition is essential for performing efficient inferences in a BN, because it allows to exploit the conditional independencies between the nodes to avoid calculating unnecessary relations.

The BN that will be used has 3 types of nodes: **entities**, **features** and **sensors**. All nodes are binary, meaning that they can only take on two values: *true* or *false*. The entities are nodes representing the objects being discovered by the sensors. One or multiple ontologies will be used for the representing the entities and the relationships between them (Figure 1). There will be a link from any sub class entity to its super class entity, as expressed by the ontology.

Entities will have descriptive features that may or may not be unique to that entity. To represent this relationship, there will be a link from entities to the features that they posses. Features can also be arranged in ontologies, where there are links from a more specific feature to a more general feature that includes it. For example the feature known as *heavy* will have a link to the feature known as *weight*.

A feature can be measured by one or more sensors. To represent this relationship, there will be a link from features to the sensors that measure the feature. Nothing precludes a sensors from sensing multiple features. Sensors do not form an ontology among themselves. Figure 2 shows an example of a BN with entities, features and sensors.

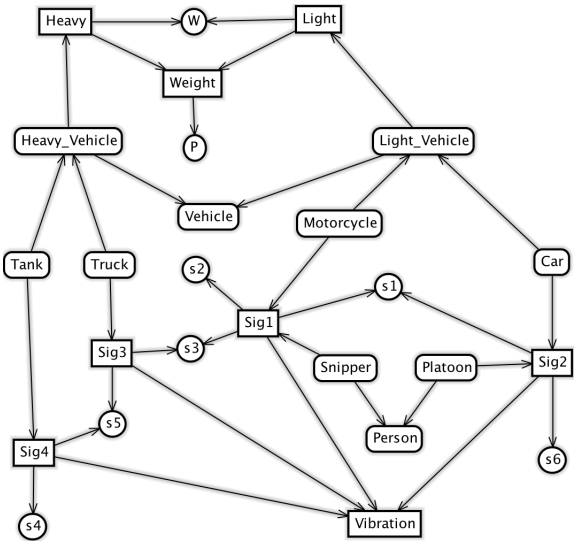

Fig. 2. Example of BN with entities, features, and sensors.

For computing the probability tables in the BN, a model known as the Noisy OR-Gate Model [19, p. 158] will be used. In the Noisy OR-Gate Model, the relationships in the BN are considered to represent causal mechanisms. So, if all of the parents of a node have the value false, then the node will also get the value false. But it at least one of the parents has the value true, then the node might get activated depending on the strength of the relationship with the active parent. That is why it is called a Noisy-OR Gate model, because it resembles the behavior of an OR Gate.

By default, all relationships among the nodes from the ontologies are considered deterministic, this is, they have a probability of 1, meaning that if the entity is of certain type, then with probability 1, the node is also of the type of the super class node, in the ontology. But nothing precludes the use of different probability values.

Using the Noisy-OR gate model, the probability of a node being true given its active parents is determined by the following formula:

$$P(n = true|p_1, \ldots, p_m) = 1 - \prod_{p_j \in A} \Big[1 - P(n = true|p_j = true)\Big], \quad (1)$$

where $P(n = true|p_j = true)$ is the strength of the link going from p_j to n, and A is the set of

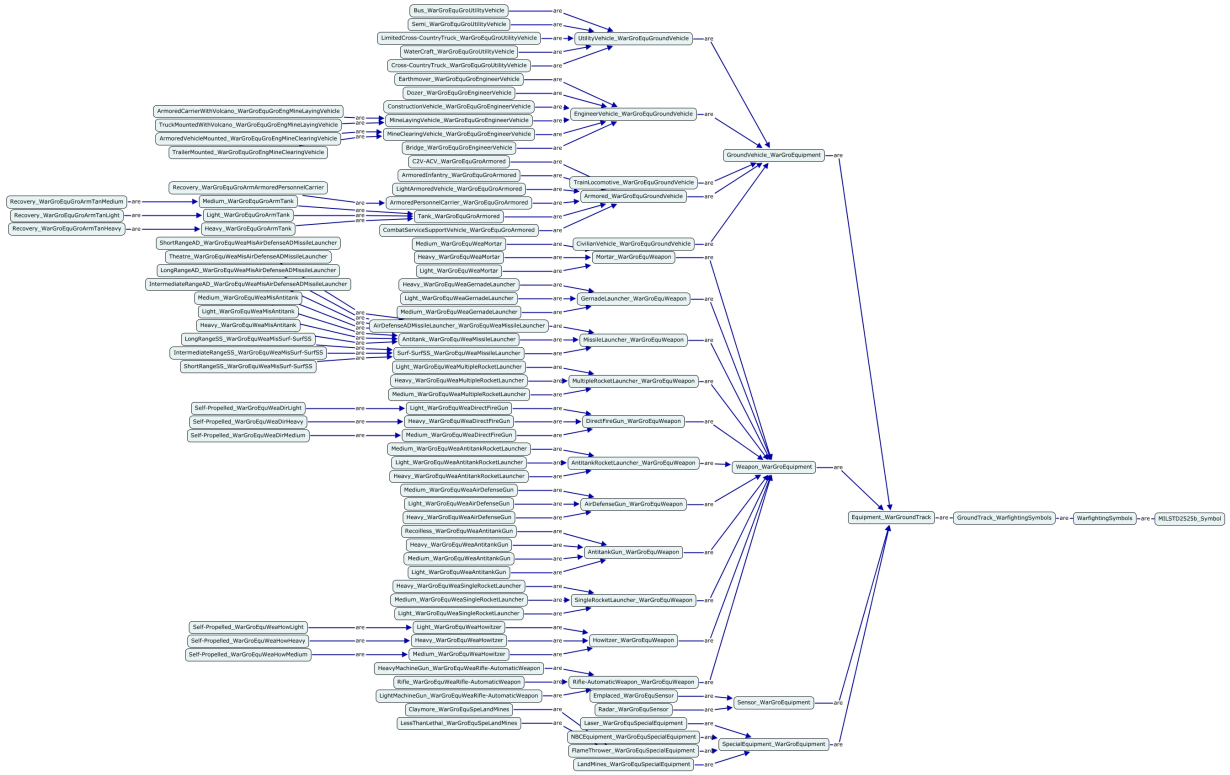

Fig. 1. Example of Military Equipment Ontology (ground vehicles only).

active parents, this is, the set of parent nodes that have a value equal to true.

For every node in the BN, that it is a root node (node that does not have any parents) the probability of having the value true is by default 0.5. But it can be easily allowed to modify these prior probabilities to impose certain beliefs or prior knowledge.

Finally, entity nodes have an inhibitory mechanism in their probability tables. If there is already another entity with a value true that is not a descendant or an ancestor of the entity, then the value for the entity must be false, because, there could only be a single object being observed or inferred at the same time.

V. Inference

The algorithm that will be use for inference in the resulting Bayesian Network is called **Logic Sampling** [19, p. 210]. This algorithm is an approximate inference algorithm for Bayesian Networks, that uses simulation for calculating an estimate of the probability of the variables given some evidence. Algorithm 1 shows the pseudo-code of the logic sampling algorithm.

The general idea behind logic sampling consists of setting the values of the observed variables according to the evidence, and then generating several times the values for the remaining variables (the unobserved variables) with a pseudo random number generator using the conditional probability table for each variable. Then, for each unobserved variable, a ratio is calculated for each of its values, by dividing the number samples for which the variable took each value by the total number of samples.

The main type of inferencing that will be done on the BN will consist of determining what entity or entities can be inferred given an evidence provided by some sensors. For example, consider the evidence shown in figure 3. This evidence tells us that two sensors detected features. Sensor P detected the feature $Weight$ and sensor $s1$ detected the feature $Sig2$. After running the inference algorithm we can say that the object that activated the sensors is a Car. We could also say that the object is a $Light\_Vehicle$ or a $Vehicle$ (Figure 4).

It is important to consider here that we might

Input: N (Nodes in the Bayesian Network in ancestral ordering)
Input: n (Number of nodes in N)
Input: A (set of nodes $A \subseteq N$ with the evidence)
Input: m (number of samples for the simulation)
$c[1 \ldots n] \leftarrow 0$;
for $r \leftarrow 1$; $r \leq m$; $r \leftarrow r + 1$ **do**
  $s[1 \ldots n] \leftarrow false$;
  $j \leftarrow 1$;
  while $j < n$ **do**
    $pa(j) \leftarrow \{(N[i], s[i]) :$ $N[i]$ is parent of $N[j]\}$;
    if random $([0, 1]) \leq P(N[j]|pa(j))$ **then**
      $s[j] \leftarrow true$;
    end
    if $N[j] \in A \wedge s[j] = false$ **then**
      $j = 1$;
    else
      $j + +$;
    end
  end
  for $i \leftarrow 1$; $i \leq n$; $i \leftarrow i + 1$ **do**
    if $s[i] = true$ **then**
      $c[i] \leftarrow c[i] + 1$;
    end
  end
end
$P[1 \ldots n] \leftarrow 0$;
for $i \leftarrow 1$; $i \leq n$; $i \leftarrow i + 1$ **do**
  $P[i] \leftarrow \frac{c[i]}{m}$;
end
return P

Algorithm 1: Logic Sampling

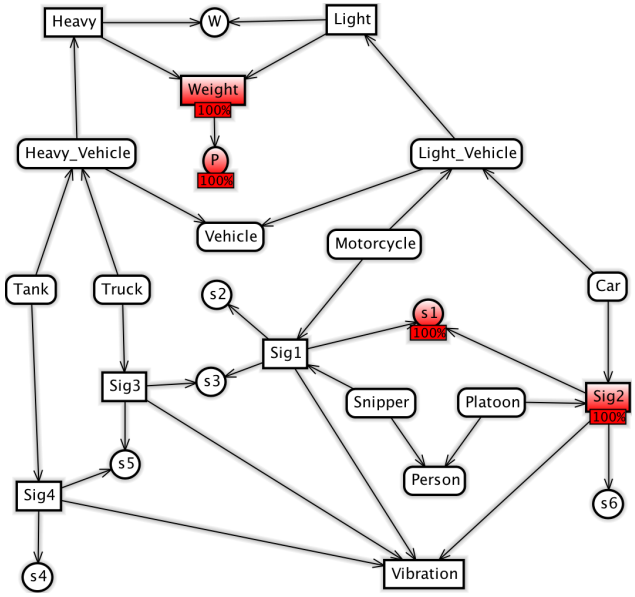

Fig. 3. Example of BN with evidence entered.

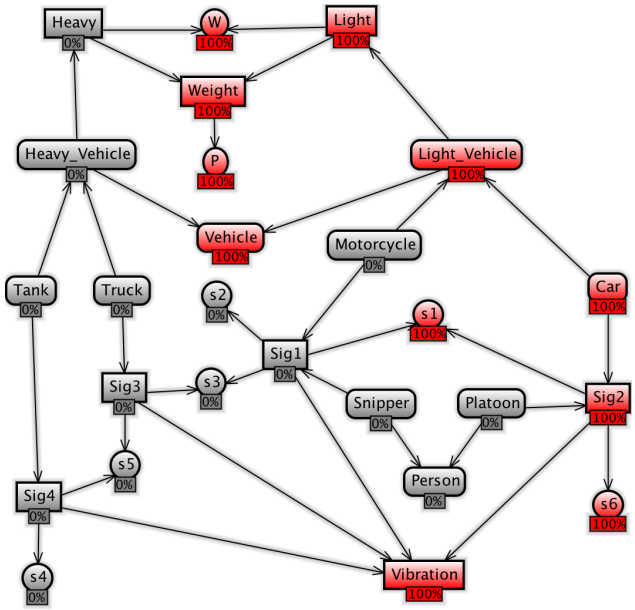

Fig. 4. Example of BN with inference results.

also feed the inference algorithm with negative information, this is, information about the sensors that did not activate. Using this information could give us more accurate inferences, but as we really do not know if the sensor did not activate because the object did not have a feature that can be detected by the sensor, or simply because the object is out of range from the sensor, then this type of information is not reliable. So from know on, we will consider that all evidence will only come from sensors that activated. But the algorithms could easily be extended to use negative information.

VI. Risk Assessment

After determining the type of object being observed using the evidence provided by the sensors, we need to asses the risk of disclosing this information, in case that one or more of the sensors in use is a classified sensor. Assessing the risk,

here means assessing the probability of disclosing the existence and/or location of a classified sensors that are on the field after disclosing the existence of an object detected by the sensors. The proposed approach for assessing this risk consists of finding all the combinations of sensors that would allow to detect the entity being disclosed, and then computing the conditional probability of the classified sensors across all the combinations.

We will call a Entity Detection Sensor (EDS) set, a set of sensors that are able to detect together a given entity. For example, in the BN from figure 2, sensors P and $s1$ form an EDS set for entity Car. But for assessing the risk we are not interested on all EDS sets, but only on the minimum Entity Detection Sensor (mEDS) sets, which are EDS sets that do not have a proper subset that is also an EDS set for the same entity. The reason for this is that once we find an mEDS set, then any other set that includes all the sensors in the mEDS set will also be able to detect the entity. This definition will not only provide more accurate measurements of risk, but will also reduce the search space.

Algorithm 2 finds the mEDS sets for a given entity. This algorithm first runs an inference using as evidence the entity e being disclosed, and creates a set F with all the features that get activated for this inference. The activation is determined by a threshold t that is passed to the algorithm. Then using this set, we construct another set S with all the sensors that can detect this features.

Then iterating over the power set of S, and inference is done using each combination of sensors for each of the respective features that they measure. If for any of the combinations, the probability of the entity e is higher than the threshold t, then this combination of sensors is added to the list of mEDS sets for the entity. A very important feature of the algorithm is that it explores the power set space increasingly, first trying subsets with 1 element, then subsets with 2 elements, etc. This approach helps to reduce the search space when the mEDS sets are small. The worst case occurs when all sensors are needed, in which case the number of combinations explored is $O(2^n)$, being n the number of sensors. Applying the algorithm over the BN of figure 2 with entity Car produces the following list of mEDS sets: $\{P, s1\}$, $\{P, s6\}$, $\{W, s1\}$ and $\{W, s6\}$. In this case the risk of each of this sensors is easy to compute, since each of them occur in two of the mEDS sets, so the risk is 50%.

Algorithm 2 does not consider the mEDS sets for sub classes of the entity in the ontology. This is not important when searching for the mEDS sets for an entity with no sub classes, like for example, for entity Car in figure 2, but it is important for entities that have sub classes, because when assessing the risk, we need to consider that the recipient of the information might suspect that the reason why we are disclosing a more general piece of information is for hiding the existence of a classified sensor, so, he/she will try to explore the mEDS sets for all sub classes of the entity that we provided. A simple modification of the algorithm, that recursively iterates over the sub classes will provide the complete set of mEDS sets needed for assessing the risk. Applying this modified version of the algorithm over the BN of figure 2 with entity $Light\_Vehicle$ produces the following list of mEDS sets: $\{P, s1\}$, $\{P, s2\}$, $\{P, s3\}$, $\{P, s6\}$, $\{W\}$, $\{W, s1\}$, $\{W, s2\}$, $\{W, s3\}$ and $\{W, s6\}$. In this case, the risk for some sensors $s1$ and $s6$ decreases to 22%, and for sensor P decreases to 44%, while the risk for sensor W increases to 56%.

VII. Evaluation

To validate the hypothesis that a more generalized description of entities (based on the ontology hierarchy) will help reduce the risk of unintended sensor information, we randomly generated some networks of different sizes, and for each of those networks we assessed the risk of disclosing each of the entities that have no sub classes. Then we assessed the risk of each of the immediate super classes. Finally, we subtracted the risk for each sensor when providing more general information from the risk for the same sensor when providing more specific information. Averaging these differences across all cases, we checked if our proposed heuristic is statistically valid.

Under the null hypothesis of no risk reduction occur, the average of the computed differences (i.e. entity disclosure risks, and immediate parent description disclosure risk) should be close to zero. Rejecting the null hypothesis would tell us that the differences are not zero, which could be good or bad

Input: N (ASP Network)
Input: e (Entity to be disclosed)
Input: t (Activation threshold)
$P \leftarrow$ **DoInference**$(N, \{(e, true)\})$;
$F \leftarrow \{f : f$ is feature $\wedge\, P[f] \geq t\}$;
$S \leftarrow \{s : s$ is sensor $\wedge \exists f \in F(s$ detects $f)\}$;
$F[s \in S] \leftarrow \{f : f \in F \wedge s \in S \wedge s$ detects $f\}$;
$mEDS \leftarrow \{\}$;
/* 2^S is the power set of S */
for $C \in 2^S$ **do**
  indexes$[1 \ldots |C|] \leftarrow 0$;
  $i \leftarrow 1$;
  while $C \notin mEDS \wedge i \geq 1$ **do**
    indexes$[i] \leftarrow$ **indexes**$[i] + 1$;
    if indexes$[i] \geq$ **max**$(1, |F[C[i]]|)$ **then**
      indexes$[i] \leftarrow 0$;
      $i \leftarrow i - 1$;
    else
      if $i = |C|$ **then**
        $E \leftarrow \{\}$;
        for $j \in \{1 \ldots |C|\}$ **do**
          $s \leftarrow C[j]$;
          $E \leftarrow E \bigcup \{(s, true)\}$;
          $f \leftarrow F[s]$;
          if indexes$[j] \leq |f|$ **then**
            $E \leftarrow E \bigcup \{(f[$**indexes**$[j]], true)\}$;
          end
        end
        $P \leftarrow$ **DoInference**(N, E);
        if $P(e) \geq t$ **then**
          $mEDS \leftarrow mEDS \bigcup \{C\}$;
        end
      else
        $i \leftarrow i + 1$;
      end
    end
  end
end
return $mEDS$

Algorithm 2: Minimum Entity Detection Sensor (mEDS) sets search

depending on which direction the risk moves. What we are really looking for is a positive difference, meaning that the risk of providing more general information is lower than the risk of providing more specific information.

Computing the differences for each sensor on each risk assessment result provided us 574 different data points. The average value is 0.3092819, and the variance is 0.06492386. Performing a t test over the data gives a p-value which is smaller than $2.2e-16$.

Based on the calculated p-value we can reject the null hypothesis and claim that, *on average*, the results support the hypothesis that providing more general information about a target will help protect the disclosure of its information sources. It is important to note that these results are dependent on the structure of the ontologies or equipment and features used for our analysis. Our decision support system will only proceed with recommendation to a user if there is statistical support for the proposed underlying hypothesis described here.

The algorithms were also tested on networks created using a subset from military exercise datasets collected from the Army National Training Center. The subset consists of 161 nodes, where 51 nodes were selected as sensors and the remaining 110 nodes were selected as entities. The 51 sensor nodes were grouped into 18 different sensor types, and the 110 entity nodes were also grouped in an ontology containing 63 nodes.

Table I shows the results for the risk assessment algorithm over 3 different entities in the scenario. Each row in the table shows the risk assessment for each of the 18 sensor types in the scenario. For each entity the risk assessment was performed over the entity type itself and over the parent entity and the parent's parent entity. Their respective results are arranged in this same order.

The parent of entity type *P07* is *P0* and the parent of entity type *P0* is entity *P*. A first observation to be made on these results is that moving up in the ontology only helps to reduce the risk of the sensors involved in discovering the lower level entity type. For example, sensor *SNH* has zero risk when discovering an entity of type *P07*, but it has a risk of 36% when discovering entities of the higher level type *P*. This behavior is just a direct consequence of the approach, which is, that the

Sensor	Entity								
	P07	P0	P	P82	P8	P	4UZ	4U	4
SEH	14.3%	12.5%	8.0%	33.3%	33.3%	8.0%	18.2%	18.2%	17.0%
SQ3IR	14.3%	12.5%	24.0%	33.3%	33.3%	24.0%	36.4%	36.4%	22.6%
SEC	14.3%	12.5%	4.0%	0.0%	0.0%	4.0%	18.2%	18.2%	11.3%
SNH	0.0%	0.0%	36.0%	0.0%	0.0%	36.0%	100.0%	100.0%	52.8%
SW	0.0%	0.0%	12.0%	0.0%	0.0%	12.0%	0.0%	0.0%	7.6%
SQ5	14.3%	12.5%	4.0%	0.0%	0.0%	4.0%	9.1%	9.1%	22.6%
SNU	100.0%	87.5%	48.0%	0.0%	0.0%	48.0%	0.0%	0.0%	3.8%
SN8	0.0%	12.5%	4.0%	0.0%	0.0%	4.0%	0.0%	0.0%	0.0%
SW0	0.0%	0.0%	24.0%	100.0%	100.0%	24.0%	0.0%	0.0%	7.6%
SEY2	14.3%	12.5%	4.0%	0.0%	0.0%	4.0%	27.3%	27.3%	11.3%
SN9	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2.0%
SN90	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	2.0%
SW3	0.0%	12.5%	12.0%	0.0%	0.0%	12.0%	0.0%	0.0%	15.1%
SQ3I6	57.1%	50.0%	56.0%	33.3%	33.3%	56.0%	36.4%	36.4%	22.6%
SWP	0.0%	0.0%	12.0%	0.0%	0.0%	12.0%	0.0%	0.0%	17.0%
SEYO	57.1%	50.0%	16.0%	0.0%	0.0%	16.0%	36.4%	36.4%	37.7%
SWN	0.0%	0.0%	12.0%	0.0%	0.0%	12.0%	0.0%	0.0%	9.4%
SQ3Z	14.3%	12.5%	4.0%	0.0%	0.0%	4.0%	18.2%	18.2%	20.8%

TABLE I
RISK ASSESSMENT FOR SUBSET OF FOR THE ARMY NATIONAL TRAINING CENTER MILITARY EXERCISE SCENARIO.

reason why other sensors get their disclosure risk reduced is because other sensors are added into the mix.

Let us consider, for each entity, the sensors that have a disclosure probability risk greater than zero for the lower level entity, the cells highlighted in black show reductions in the risk, while cells highlighted in gray show increased risk of disclosure.

It is important to notice that under some conditions (gray cells) there is an increase in sensor disclosure risk for less specific entity descriptions (see, for example sensor *SQ5*, for entity *4* in the last column). It is also possible that an increase in risk may occur after a decrease in risk has been observed for the same entity hierarchy (see, or example, sensor *SQ3IR*, for entities *P07, P0, P*).

This behavior is consistent with the fact that some sensors detect more general features, that is, features that are present in more general entities. Specialized sensors allow to effectively detect very specific features. As we generalize the description of some entities the value of specialized features diminish in comparison with more general features, which could lead to the increase of probability disclosure of more general sensors. In general, however, more abstract features are detected by a greater number of sensors which should compensate for the effects in most (but not all) cases. In the counter examples shown in table I, the sensors that measure more general features start gaining relevance in comparison with more specific sensors and their likelihood of disclosure increases.

VIII. CONCLUSIONS AND FUTURE WORK

Intuitively, providing a more general and vague information about a given target should help to protect the sources of the information that made the target detection (and identification inference) possible. The approach proposed in this work builds on such intuition to quantitatively estimate the risk of source disclosure as a function of information release.

Our experimental results showed that most often, a small generalization of the target description may yield a significant risk reduction for unintended sensor information disclosure. Furthermore, we have illustrated that such gains can be quantified as probability of unintended disclosure per sensor, under the worst-case assumption of a fully informed and rational adversary. This information could provide valuable run-time insight to a human operation for information release decision support.

However, we have also noted in our results that while the hypothesis is often true, that is not always the case. There are conditions in which a more general description of a target may in fact, increase

the probability of disclosure of unintended sensors. As discussed in our analysis, such cases are due to the uneven distribution of feature detection capabilities between different sensors and different levels of target abstraction. However, even in such cases, the proposed risk assessment framework properly indicated the risks associated with each level of abstraction, recommending the release of the appropriate description.

The quantitative estimate of disclosure risk can provide human operators with a reference for decision making. That information could be taken into account with other mission-level requirement and operational contexts to help support a decision for which level of detail to be released for different conditions. The experimental results shown in this paper were obtained by running the algorithms several times over small sections of a simulated network environment based on actual troop movements for exercises conducted at the US Army National Training Center.

Our approach relies on there critical pieces for information for analysis. a) the network topology including sensor location, descriptions and capabilities, b) an ontology describing the hierarchy of classes for the entities being tracked or monitored by the sensor field (the military vehicle ontology, in our example), and c) the correlations among the ontologies for entities and ontologies for entity features. Network topology and sensor details can be efficiently discovered using cross-layer monitoring infrastructures, as we have shown in our previous work. The second and third requirements are also very realistic and should not represent a limitation to our approach.

As part of our continued research in this domain, we will investigate new methods to improving the efficiency of the proposed algorithms for large scale problems. We are currently exploring several heuristics that can be applied to simplify the problem including temporal and spatial associations between events, and the introduction of higher level policies for information release, which could help prevent the exploration of solutions in prohibited spaces.

REFERENCES

[1] M. Carvalho, N. Suri, V. Shurbanov, and E. Lloyd, "A cross-layer network substrate for the battlefield," in *Proceedings of the 25th Army Science Conference*, November 2006.

[2] M. Carvalho, A. Granados, W. Naqvi, A. Brothers, J. P. Hanna, and K. Turck, "A cross-layer communications substrate for tactical information management systems," in *Military Communications Conference (MILCOM)*. IEEE, Nov. 2008.

[3] J. Achugbue and F. Chin, "The effectiveness of output modification by rounding for protection of statistical databases," *INFOR*, vol. 17, no. 3, pp. 209–218, 1979.

[4] N. R. Adam and J. C. Worthmann, "Security-control methods for statistical databases: a comparative study," *ACM Comput. Surv.*, vol. 21, no. 4, pp. 515–556, 1989.

[5] R. Agrawal and R. Srikant, "Privacy-preserving data mining," *ACM Sigmod Record*, vol. 29, no. 2, pp. 439–450, 2000.

[6] T. Dalenius, "Towards a methodology for statistical disclosure control," *Statistik Tidskrift*, vol. 15, pp. 429–444, 1977.

[7] D. Rubin, "Statistical disclosure limitation," *Journal of Official Statistics*, vol. 9, no. 2, pp. 461–468, 1993.

[8] F. Chin and G. Ozsoyoglu, "Auditing and inference control in statistical databases," *IEEE Transactions on Software Engineering*, pp. 574–582, 1982.

[9] I. Dinur and K. Nissim, "Revealing information while preserving privacy," in *22nd ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. ACM New York, NY, USA, 2003, pp. 202–210.

[10] C. Dwork and S. Yekhanin, "New efficient attacks on statistical disclosure control mechanisms," *Lecture Notes in Computer Science*, vol. 5157, pp. 469–480, 2008.

[11] L. Sweeney, "*k*-anonymity: A model for protecting privacy," *International Journal of Uncertainty Fuzziness and Knowledge Based Systems*, vol. 10, no. 5, pp. 557–570, 2002.

[12] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, "l-diversity: Privacy beyond k-anonymity," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 1, no. 1, p. 3, 2007.

[13] N. Li, T. Li, and S. Venkatasubramanian, "t-closeness: Privacy beyond k-anonymity and l-diversity," in *Proc. of ICDE*. Citeseer, 2007, pp. 106–115.

[14] X. Xiao and Y. Tao, "M-invariance: towards privacy preserving re-publication of dynamic datasets," in *Proceedings of the 2007 ACM SIGMOD international conference on Management of data*. ACM, 2007, p. 700.

[15] D. Agrawal and C. Aggarwal, "On the design and quantification of privacy preserving data mining algorithms," in *Proceedings of the twentieth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. ACM New York, NY, USA, 2001, pp. 247–255.

[16] E. Bertino, D. Lin, and W. Jiang, "A Survey of Quantification of Privacy Preserving Data Mining Algorithms," *Privacy-Preserving Data Mining. Springer, US*, 2008.

[17] M. Kantarcioglu, J. Jin, and C. Clifton, "When do data mining results violate privacy?" in *Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining*. ACM, 2004, p. 604.

[18] G. Canfora and B. Cavallo, "A bayesian model for disclosure control in statistical databases," *Data & Knowledge Engineering*, vol. 68, no. 11, pp. 1187 – 1205, 2009, including Special Section: Conference on Privacy in Statistical Databases (PSD 2008) - Six selected and extended papers on Database Privacy.

[19] R. E. Neapolitan, *Learning Bayesian Networks*, ser. Series in Artificial Intelligence. Upper Saddle River, NJ: Pearson Prentice Hall, 2004.