Inference Management: Trust and Obfuscation in Coalition Settings

Chatschik Bisdikian[†], Murat Sensoy[*], Nir Oren[*], Christopher Burnett[*]
Timothy J. Norman[*], Mani B. Srivastava[§] and Lance Kaplan[¶]
[†]T. J. Watson IBM Research Center, New York, USA
[*]University of Aberdeen, Aberdeen, UK
[§]University of California, Los Angeles, USA
[¶]US Army Research Laboratory, Adelphi, MD, USA

***Abstract*—In modern coalition operations, decision makers must be capable of obtaining and fusing information from diverse sources. The reliability of these sources can vary, and, in order to protect their interests, the information they provide could be altered, e.g., obfuscated, to limit the inferences that can be made with it. The trustworthiness of fused information depends on both the reliability of these sources and their inference management policies. Information consumers must determine how to evaluate trust in the presence of inference management techniques, such as obfuscation, while information providers must determine the appropriate level of obfuscation in order to ensure both that they remain trusted, and do not reveal any private information. In this paper, we present and formalise trust in the context of inference management and discuss the relationships between the two. We illustration the pertinent concepts via a multi-party coalition scenario, and present numerical examples, using subjective logic computational techniques, of how trust and obfuscation can influence belief levels in information gathered.**

I. Introduction

In "traditional" sensing applications, such as localization, monitoring, and identification, sensing systems are deployed to support and enhance one's capabilities to observe the world of interest and estimate its state. Typically, these sensor-enabled applications (which act as *information consumers*, or, simply *consumers*) are deployed along with their own dedicated sensing resources for making the observations that they consume (i.e., process). In these tightly-coupled (closed) deployments, the consumer has reasonable knowledge of (or access to) the capabilities and deficiencies of the deployed sensing resources. Figure 1 shows at a high-level the structure of a traditional sensor-driven "inference" system that makes inferences regarding the world under observation.

In the context of modern coalition operations, new sensing opportunities and information collection paradigms have proliferated. Tasks like providing humanitarian relief to the injured after a disaster or defeating insurgents in a war-torn country require a collaborative effort of different organizations (e.g., non-governmental organizations(NGOs) and the military). To achieve these tasks, critical information–such as the location of injured people–has to be sensed and exchanged between information providers and consumers under different

This work was presented at the *SPIE Defense, Security, and Sensing Symposium*, Baltimore, MD, USA, April 23-27, 2012.

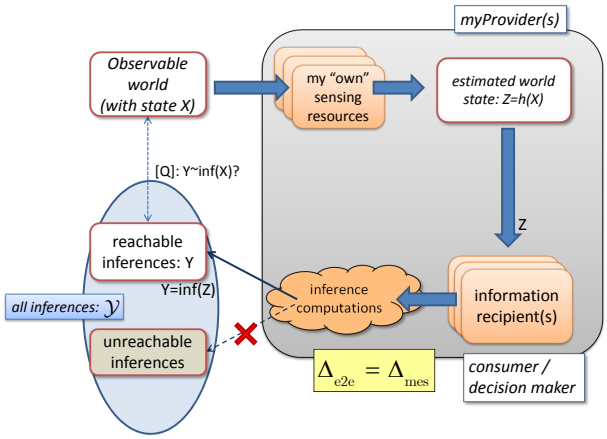


Fig. 1: A traditional (tightly-coupled) sensor-enabled inference system.

administrative domains. In these cases, the closed, single-administrative-domain association between the information producers and consumers is challenged by more open and hence more complex and unpredictable collaborative multi-administrative (and even no-administrative) domain associations. Knowledge of the capabilities of sensing entities can be unavailable and unknown, or policy-constrained, and certainly of questionable reliance. In addition, shared information can be deliberately manipulated for various reasons. As a result, it becomes harder to quantify the value of the fused information and the risks associated with acting on subsequent inferences.

Figure 2 shows an example of a collaborative, end-to-end information processing system where the information providers and consumers interact only as necessary. In such a setting, policies that affect the sharing of information, including any deliberate alteration of it through techniques such as obfuscation [1], and the underlying trust between the parties involved, play a key role in mediating effective collaboration.

Naturally, transacting parties in these loosely-coupled collaborative settings are *primarily* concerned with whether the information exchanged between them, and its quality (QoI) [2], are sufficient to satisfy their needs. However, there is a *secondary* concern as to whether the exchanged information,

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
20 SEP 2012		**00-00-2012 to 00-00-2012**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Inference Management: Trust and Obfuscation in Coalition Settings	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **IBM T.J. Watson Research Center,19 Skyline Drive,Hawthorne,NY,10532**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Presented at the Annual Conference of International Technology Alliance (ACITA) 2012 held 17-21 Sep in Southampton, UK. U.S. Government or Federal Rights License

14. ABSTRACT

In modern coalition operations, decision makers must be capable of obtaining and fusing information from diverse sources. The reliability of these sources can vary, and, in order to protect their interests, the information they provide could be altered, e.g., obfuscated, to limit the inferences that can be made with it. The trustworthiness of fused information depends on both the reliability of these sources and their inference management policies. Information consumers must determine how to evaluate trust in the presence of inference management techniques, such as obfuscation, while information providers must determine the appropriate level of obfuscation in order to ensure both that they remain trusted, and do not reveal any private information. In this paper, we present and formalise trust in the context of inference management and discuss the relationships between the two. We illustration the pertinent concepts via a multi-party coalition scenario, and present numerical examples, using subjective logic computational techniques, of how trust and obfuscation can influence belief levels in information gathered.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**8**	

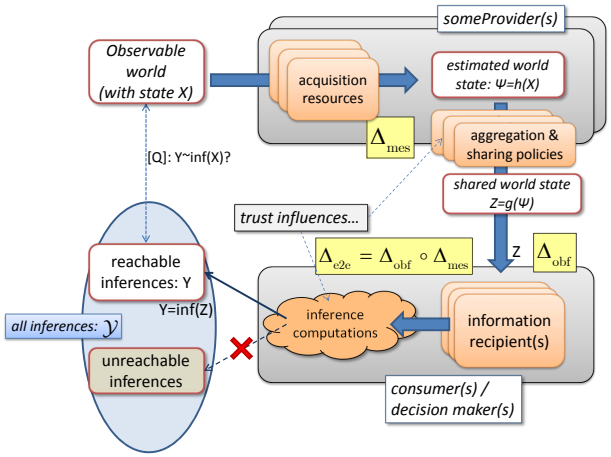


Fig. 2: A loosely-coupled, collaborative end-to-end inference system.

as a whole or in part, will be utilized only for some stated or implied purpose, and not for some other, possibly illicit, ends. In this context, obfuscation serves as a mechanism for protecting against the inappropriate use of shared information. It is one of the process by which information providers deliberately conceal and in general alter aspects of the information they provide to protect sensitive information by affecting the ability of making certain inferences while allowing information consumers to still hopefully derive value from the information they receive [1]. These techniques can range from manipulating the information content directly, e.g., removing a name entry from an information record; translating, generalizing, or adding noise to a location entry; altering features or dimensionality of the information, e.g., truncating Fourier coefficients, concealing variance information, etc. They aim to influence, i.e., *manage*, the set of inferences that can be made with [the parts of] the information shared. In this paper, we will summarily refer to such *inference managing* techniques as *obfuscation*, even though, strictly speaking, obfuscation (e.g., as in [1]) typically refers to a specific subset of these techniques.

This paper builds upon our early work on trust and obfuscation [3] where we introduced and defined the pertinent concepts and terminology. In this paper, we take these early concepts and instantiate them through a concrete usage scenario drawn from the area of combined military coalition and humanitarian aid operations. More specifically, while we borrow from our previous work [3] as necessary to provide context, the main contribution in this paper is enriching our previous work through the development and presentation of an exemplar case where a number of actors in the context of the aforementioned scenario share information for various purposes. We show how evidential analysis techniques can be employed to ascribe levels of belief to outgoing (derivative) pieces of information when fusing incoming information that was shared by actors of varying trust levels. We also discuss how the quality of the derivative information can negatively effect the trust and propose methods to avoid this.

The organization of the paper is as follows. Section II introduces our scenario. Section III defines and discusses *trust* and *obfuscation* from different perspectives through examples based on the introduced scenario and Subjective Logic [4]. Section V discusses the interplay between trust and obfuscation. Lastly, Section VI concludes the paper.

II. Example Scenario

In this section, we describe a simple scenario that will be used to describe the relationship between obfuscation and trust in the remainder of this paper. Consider a situation in which a military coalition from two nations, $\mathcal{A}$ and $\mathcal{B}$ are attempting to stabilise a war-torn country, coming up against insurgent forces $\mathcal{F}$. We assume that the coalition has operated in the country for some time, and has begun training up a local army, which we denote by $\mathcal{C}$. At the same time, humanitarian organisations (denoted by $\mathcal{R}$) are attempting to provide aid to civilians within the country, while a news agency $\mathcal{N}$ broadcasts information from anonymous sources in the area.

Now in order to act as efficiently as possible, the humanitarian organisation must interact with all other actors in the environment. On the other hand, $\mathcal{A}$ and $\mathcal{B}$ have no formal contact with $\mathcal{F}$, while they suspect that $\mathcal{C}$ may have been infiltrated by some elements of $\mathcal{F}$. It should be clear that $\mathcal{A}$ and $\mathcal{B}$ have a high level of trust between them, and that while $\mathcal{C}$ has a high level of trust in both members of the coalition, $\mathcal{A}$ and $\mathcal{B}$ do not trust $\mathcal{C}$ as much (the concept of trust is discussed more formally in Section III). The trust relationships between the parties are summarized in Figure 3, where the thickness of the directed edges represents the degree of trust the edge source assigns the edge target.

The type of trust we have discussed so far is to some extent universal. For example, members of the military coalition ($\mathcal{A}$ and $\mathcal{B}$) trust each other in all contexts. The military coalition has little trust in $\mathcal{C}$ and $\mathcal{R}$ as information providers, but places a high degree of trust in the information provided by $\mathcal{N}$. However, $\mathcal{C}$, $\mathcal{N}$, and $\mathcal{R}$ trust the coalition as an information provider. Although insurgents ($\mathcal{F}$) are enemies of both the military coalition ($\mathcal{A}$ and $\mathcal{B}$) and the local army $\mathcal{C}$, they may leak some useful information to humanitarian organisations ($\mathcal{R}$), the local army ($\mathcal{C}$), and the news agency ($\mathcal{N}$). For instance, to avoid extreme reactions due to civilian casualties, $\mathcal{F}$ may warn $\mathcal{R}$ about the regions with high IED concentrations. In these situations, $\mathcal{R}$ may trust $\mathcal{F}$ with regard to information about potential IEDs, but little else.

III. Trust and Obfuscation Definitions

Sensor-enabled applications collect sensory information about their "world" (i.e., surroundings) to support reasoning and inferences about the world's state and the evolution of alternatives. The significance and effectiveness of the inferences made and of the ensuing actions taken depend on the *quality of information* received and the value it brings to the tasks at hand [3]. These in turn are influenced by the relationships that are developed between information providers

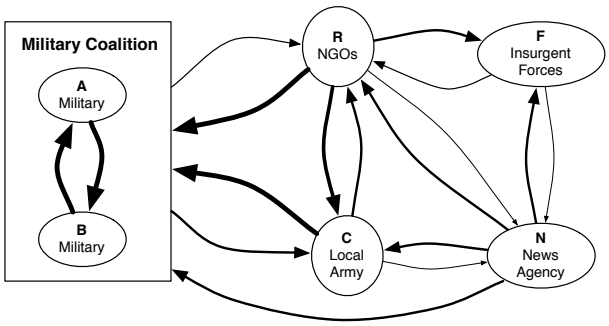


Fig. 3: Trust relationships between parties.

and consumers. Now these relationships described by levels of *trust* and instantiated (at least, in part) by *obfuscation*. In this section we examine these concepts more closely and define them.

A. The consumer's view of trust

Trust can be broadly defined as the willingness of one party (the *trustor*) to rely on the actions of another party (the *trustee*) [5]. Trust is critical in enabling interactions between parties in uncertain and constantly changing environments. In our case, these parties are the information *consumers* and *providers*. For a consumer, we define trust as follows:

Definition 1: Trust (from the consumer's viewpoint) represents the consumer's degree of belief that she can rely on the information that a provider has provided her with. ▮

Although not necessarily tied to, the reference to "degrees of belief" in the above definition allows us to naturally leverage the rich modeling and computational toolsets developed for belief-based analysis. We would shortly show some of these based on subjective logic techniques.

If a consumer has a low level of trust in the information it receives, ensuing inferences may be considered unreliable. Trust in a body of available information is influenced by many factors. These include the means by which information has been collected, i.e., its provenance, or how different pieces of information relate to, and corroborate each other, i.e., consistency in the body of information. For example, if information derived from a third-party provider conflicts with established facts and knowledge, then the level of trust in, and the value subsequently ascribed to such information would be low.

1) Trust modeling and assessment with subjective logic: In the literature, trust is most often described from the consumers' point of view [6]. Most trust and reputation systems model the consumer's trust in a provider via logic-based or probabilistic approaches [6]. A well-established framework for trust in this context is Subjective Logic (SL) [4]. In subjective logic, the opinion of an information consumer c about the validity of a proposition α is represented as $w^c_\alpha = \langle b^c_\alpha, d^c_\alpha, u^c_\alpha \rangle$, where b^c_α represents c's belief in α, d^c_α represents c's disbelief in α, u^c_α represents c's uncertainty about α, and $b^c_\alpha + d^c_\alpha + u^c_\alpha = 1$. Applying this to our definitions, let p be an information provider and α be the proposition "p is trustworthy." Given this, w^c_α represents the opinion of c about the trustworthiness of p. Opinions about the trustworthiness of information providers could be hard-coded or dynamically computed over time. For instance, w^c_α can be computed based on Equation 1, where r is proportional to the amount of positive evidence for α (i.e., number of times p has provided satisfactory information), and s is proportional to the amount of negative evidence (e.g., number of times p failed to provide satisfactory information). $W \geq 2$ is then a constant representing the non-informative prior weight.

$$b^c_\alpha = \frac{r}{r+s+W}, d^c_\alpha = \frac{s}{r+s+W}, u^c_\alpha = \frac{W}{r+s+W} \quad (1)$$

Naturally, trust in information providers affects the trust in the information they produce. If a provider has historically behaved in an untrustworthy manner, for example due to the consistent use of inappropriate or faulty sensors to gather information, any subsequent information provided may also be considered untrustworthy. Hence, the trust in providers and the information they produce are both highly correlated and build upon each other. SL exploits this correlation through the *discounting operator* $\otimes$. Let w^p_β be the opinion of p about proposition β. If p shares its opinion with c, the provided opinion is converted to a normalised opinion (i.e., w^{cp}_β) based on c's opinion about the trustworthiness of p (for convenience represented as w^c_p instead of w^c_α) as in Equation 2. This equation emphasizes that distrust with regards to p and uncertainty about the trustworthiness of p both contribute to the uncertainty about u^{cp}_β.

$$\begin{aligned} w^{cp}_\beta &= w^c_p \otimes w^p_\beta = \langle b^{cp}_\beta, d^{cp}_\beta, u^{cp}_\beta \rangle \\ b^{cp}_\beta &= b^c_p \times b^p_\beta \\ d^{cp}_\beta &= b^c_p \times d^p_\beta \\ u^{cp}_\beta &= d^c_p + u^c_p + b^c_p \times u^p_\beta \end{aligned} \quad (2)$$

Trust with respect to information cannot always be derived from trust regarding the information source. For example, when information is provided by a third-party information aggregator, provenance information is required in order to make the derivation, and often, such information may be incomplete or absent altogether.

As described above, frameworks such as SL allow for the expression of uncertainty about propositions. Now when various opinions about a phenomenon or object exist, such uncertainties are critical in deriving a *fused* opinion from the individual opinions. SL provides several operators for such fusion, including the *consensus operator* $\oplus$ which is used to fuse opinions about β from information sources s_i and s_j as defined by Equations 3 and 4. These equations imply that the more uncertainty one has about some individual piece of information, the less it affects the overall fused opinion. A provider could therefore mitigate the distrust placed in it due to poor information reporting by associating a high degree of uncertainty with this information. Such highly uncertain

information will limit the weight the consumer will assign to it during fusion, preventing them from being misled when making decisions based on this information.

$$w_\beta^{c(s_i,s_j)} = w_\beta^{cs_i} \oplus w_\beta^{cs_j} = (w_{s_i}^c \otimes w_\beta^{s_i}) \oplus (w_{s_j}^c \otimes w_\beta^{s_j}) \quad (3)$$

$$\begin{aligned} w_\beta^{(a,b)} &= w_\beta^a \oplus w_\beta^b = \langle b_\beta^{a,b}, d_\beta^{a,b}, u_\beta^{a,b} \rangle \\ b_\beta^{(a,b)} &= (b_\beta^a \times u_\beta^b + b_\beta^b \times u_\beta^a)/\kappa \\ d_\beta^{(a,b)} &= (d_\beta^a \times u_\beta^b + d_\beta^b \times u_\beta^a)/\kappa \\ u_\beta^{(a,b)} &= (u_\beta^a \times u_\beta^b)/\kappa \\ \kappa &= u_\beta^a + u_\beta^b - u_\beta^a \times u_\beta^b \end{aligned} \quad (4)$$

B. Obfuscation and the provider's view of trust

The dictionary definition of *obfuscation* is "to make so confused or opaque as to be difficult to perceive or understand"[1]. Common examples of obfuscation include GPS's selective availability, where adding noise and witholding certain data results in only low-accuracy location and time information being made available to civilian users. Similarly, civilian users can often access only low resolution satellite imagery, whereas others could be provided with more fine grained images.

In computing, obfuscation refers to the process of hiding certain data, while maintaining the usefulness of the data for an intended purpose, i.e., allowing authorized inferences [1]. It has typically been used to prevent leakage of personal information that would, for example, enable the identification of a patient in medical records, or the identification of a person and/or creating permanent records of a person's exact location when location-based information is made available to third parties for marketing purposes. Techniques such as anonymization and location abstraction are commonly used for this purpose [1], [7].

However, when opportunities abound for collecting and fusing information derived from multiple providers (including physical sensors, knowledge bases, human observers, experts, etc.), sufficient knowledge may be gained and inferences drawn that could go beyond the intentions that caused the gathering of information in the first place. For instance, by providing NGOs with dates of danger for civilians in an area, the military coalition may implicitly enable NGOs to reason about the dates of military operations in the area. This leads us directly to the provider's view of trust, which we define as follows.

Definition 2: Trust (from a provider's view) represents the provider's degree of belief that a consumer will use her information only for expressed purposes. ▮

This definition suggests that obfuscation can be used not only to protect a specific piece of information such as the position of an object of interest, but also meta-information regarding the domain, such as the capabilities of some sensors. In the context of the coalition operations wherein coalition members develop ad hoc and transient alliance relationships, partners may desire to share localisation information. Now assume that one partner has high quality sensor resources and would like to provide another partner with the information gathered from these sensors, *without revealing the location and nature of these sensors*. Obfuscation could be used to hide this latter meta-information while no obfuscation would take place Therefore, we can define obfuscation more broadly as follows:

Definition 3: Obfuscation is the process that an information provider uses to *consciously* influence the set of inferences that could be made involving the information she provides. ▮

The above definition underscores our broader view about obfuscation mentioned in the Introduction. It focuses on the ultimate intention of a provider to manage the inferences that can be made with the information it shares with an aim to deliberately control the range of uses that the shared information may have. Figure 4 comprises a *conceptual* representation for this situation. It depicts an area $\mathcal{I}$, called *inference space*, representing the set of all possible inferences (or, hypotheses) that can be asserted. The latter should be interpreted within the general context of the interactions between the information suppliers and providers. Certain subsets of these inferences may be more critical to an information provider than others possibly entailing an uncomfortably high level of risk for the provider. These critical inference subsets are represented by the red region on the right, while the green region on the left represents risk-free inferences. As an example, the red region could contain information revealing the whereabouts and/or assets of a specific very important person. The figure also shows an inference region R_d representing the region of inferences a consumer has expressed a desire for and was the reason for a particular act of information sharing. For a piece of information X that is shared, the figure shows the subset of inferences R_X that can be deduced from X. Clearly, X will be of value to the consumer if R_d and R_X intersect; the figure shows a case where $R_d \subset R_X$. Should some of these inferences in R_X fall with the critical reddish area, the provider may instead decide to share an obfuscated version $Z = h(X)$ (e.g., Z could be a noisier, translated or dimensionally-diminished variant of X) which leads to a corresponding inference region R_Z located away from the red area. Likewise, R_Z would be of value to the consumer if R_d intersects with R_Z.

The transformation of inference regions due to the obfuscation may imply strict disallowance of certain inferences. It may also imply the mere modification of the likelihood that, using the shared information, a particular inference can be made with an acceptable level of confidence. As such, the level of confidence in the inferences made using Z may not be the same as those made using X even for inferences that fall within region R_d. Therefore, the value of information that the consumer obtains when using the obfuscated information Z could be lower than when using X.

[1] *http://www.thefreedictionary.com*

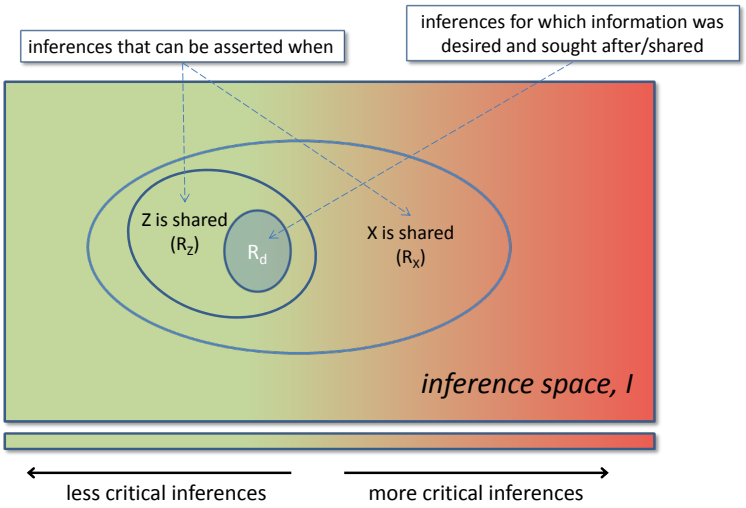


Fig. 4: An illustrative example of influencing inferences that can be made.

IV. Numerical examples

In this section, we give examples from our scenario to demonstrate the concepts introduced in the previous sections.

A. Trustworthiness of information providers

In our scenario, the local army $\mathcal{C}$ places a high level of trust in the information that it receives from the coalition members $\mathcal{A}$ and $\mathcal{B}$, however, the coalition members do not place as much trust in information they receive from $\mathcal{C}$. The evidence against $\mathcal{C}$'s trustworthiness as an information provider could be based on invalid/inconsistent/misleading information it provided in the past.

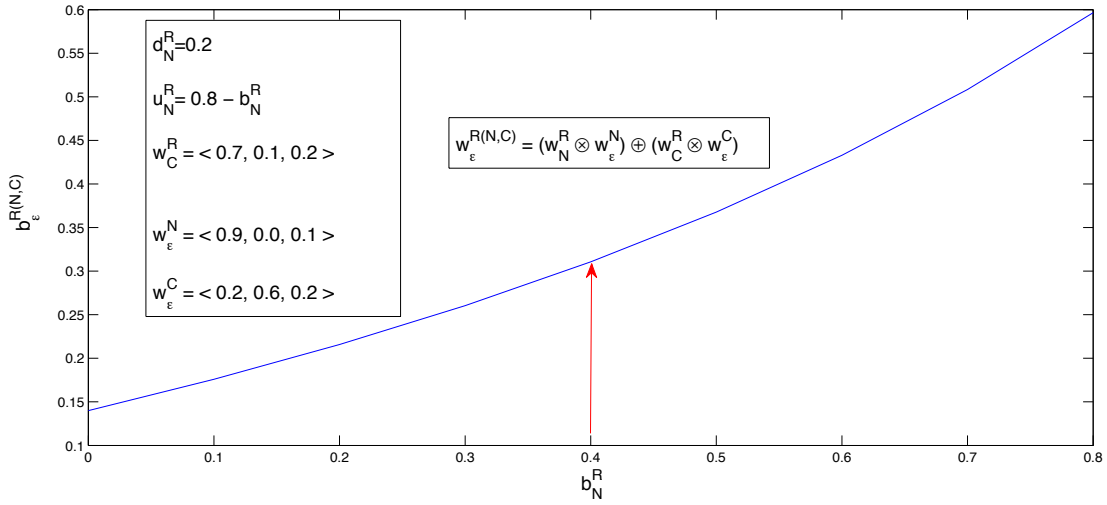
For example, suppose $\mathcal{C}$ (the local army) provides to the coalition members the intelligence report i "insurgents will attack city X on day D" with an expressed opinion $w_i^{\mathcal{C}} = \langle 0.9, 0.0, 0.1\rangle$, signifying a high level of confidence in the validity of the report. If, at some point, the coalition finds that this intelligence report was false, the provided information becomes significant evidence against $\mathcal{C}$'s trustworthiness. On the other hand, if the opinion shared by $\mathcal{C}$ had much more uncertainty (e.g., $\langle 0.5, 0.0, 0.5\rangle$), the false intelligence would not act as significant evidence against $\mathcal{C}$'s trustworthiness as an information provider, and would not impact on $\mathcal{A}$ and $\mathcal{B}$'s trust in $\mathcal{C}$ as much.

The military coalition in our scenario may provide information with high variance or uncertainty, which may still perceived as trustworthy by the NGOs and the local army, provided that the high variance or uncertainty in the information was clearly stated by the provider and understood by the consumer. In other words, even if the information provided may not have been of high quality, the provider was trusted in that it did not conceal its low quality.

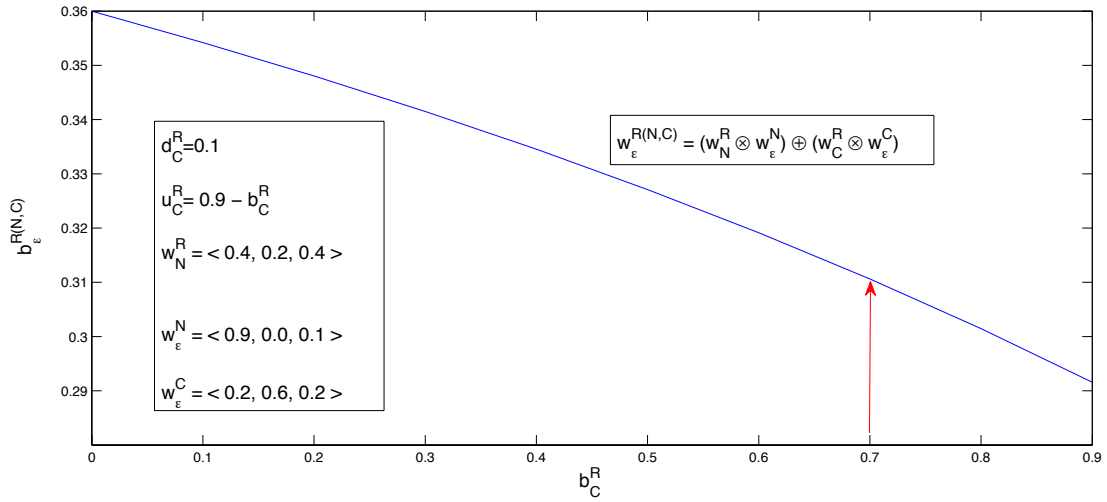
Assume that the military coalition has provided its opinion to NGOs about the proposition $p_{\{a_0:d_0\}}$ "there is a danger in area a_0 on date d_0" as $w^{\mathcal{A},\mathcal{B}}_{p_{\{a_0:d_0\}}} = \langle 0.29, 0.01, 0.7\rangle$. If it turns out that there was no danger on date d_0 in the area a_0, the trustworthiness of the military by the NGO would not be greatly reduced due to the high degree of uncertainty of this proposition.

Now let us assume that the news agency $\mathcal{N}$ disseminates information from anonymous sources. Reliability of the information cannot be measured easily. Even though the actual information source may be reliable, since the source is kept anonymous by $\mathcal{N}$, the information from $\mathcal{N}$ is regarded as untrustworthy by the military coalition in our scenario. This could be due to a policy adopted by the coalition stating that any anonymous information is untrustworthy. However, $\mathcal{C}$, $\mathcal{R}$, and $\mathcal{F}$ may have a much more relaxed trust policies about anonymous information sources. Hence, their trust in $\mathcal{N}$ may vary based on their own assessment of the information and their own trust criteria (e.g., evidence for/against).

The computed trust plays a significant role during information fusion. If $\mathcal{R}$ receives conflicting opinions from $\mathcal{C}$ and the military coalition, it firstly normalises these opinions using the discounting operator in Equation 2 and fuses them using the consensus operator in Equation 4. For instance, let $w_\varepsilon^{\mathcal{N}} = \langle 0.9, 0.0, 0.1\rangle$ and $w_\varepsilon^{\mathcal{C}} = \langle 0.2, 0.6, 0.2\rangle$ be opinions of the news agency and the local army about some event ε. Assume these very different options are received by $\mathcal{R}$ whose degree of trust in $\mathcal{N}$ and $\mathcal{C}$ are represented as opinions $w_{\mathcal{N}}^{\mathcal{R}} = \langle 0.4, 0.2, 0.4\rangle$ and $w_{\mathcal{C}}^{\mathcal{R}} = \langle 0.7, 0.1, 0.2\rangle$, respectively. After normalisation based on trust, the opinions become $w_\varepsilon^{\mathcal{RN}} = \langle 0.36, 0.0, 0.64\rangle$ and $w_\varepsilon^{\mathcal{RC}} = \langle 0.14, 0.42, 0.44\rangle$. These two normalised opinions are fused as $w_\varepsilon^{\mathcal{R}(\mathcal{N},\mathcal{C})} = \langle 0.31, 0.34, 0.35\rangle$. The overall opinion about ε is significantly more influenced by the opinion of $\mathcal{C}$ since thanks to its higher trustworthiness with regards to $\mathcal{R}$. To clearly demonstrate this influence, Figure 5, shows $w_\varepsilon^{\mathcal{R}(\mathcal{N},\mathcal{C})}$ as a functions of the opinion of $\mathcal{R}$ about $\mathcal{N}$ and $\mathcal{C}$. Specifically,



(a) Belief of $\mathcal{R}$ in $\mathcal{N}$ vs. $b_\varepsilon^{R(N,C)}$.



(b) Belief of $\mathcal{R}$ in $\mathcal{C}$ vs. $b_\varepsilon^{R(N,C)}$.

Fig. 5: How belief of $\mathcal{R}$ in ε changes as its belief in $\mathcal{N}$ and $\mathcal{C}$ varies.

Figure 5a shows $w_\varepsilon^{\mathcal{R}(\mathcal{N},\mathcal{C})}$ against the belief of $\mathcal{R}$ in $\mathcal{N}$ and Figure 5b shows $w_\varepsilon^{\mathcal{R}(\mathcal{N},\mathcal{C})}$ against the belief of $\mathcal{R}$ in $\mathcal{C}$. In the figures, we highlight the specific values from the example above using arrows.

B. Trustworthiness of consumers and obfuscation

In our example scenario, military coalition members $\mathcal{A}$ and $\mathcal{B}$ would (typically) have no problem sharing operational information with each other, but would prefer not to give $\mathcal{C}$ and $\mathcal{R}$ all this information, as this could allow $\mathcal{F}$ to obtain these details. However, some information *must* occasionally be shared (for example to protect civilians from being caught up in operations). Therefore, the coalition might inform $\mathcal{R}$ that they should not enter some city between certain dates while having full knowledge of when they would conduct some operation in that city, and precisely which area of the city this operation would be conducted in. The coalition might obfuscate details of the attack further by giving a selection of cities rather than a single one, together with a range of dates.

This represents an example of *obfuscation by abstraction* or *generalization*, where the pieces of information that are shared are supersets of what is actually known, thus, rendering the shared information less specific. Thus, with regard with the spatiotemporal context of a shared piece of information (describing the *when* and *where* contained in the information) the specific time of an event may be abstracted, for example, to the day of the event or a range of days. Similarly, the specific location of the event–for example a specific street corner–may be abstracted to a city block, or the entire city, or county.

Let $p_{\{a_{0:i},d_{0:j}\}}$ represent a proposition "there is danger in an area $a \in \{a_0 \dots a_i\}$ on a day $d \in \{d_0 \dots d_j\}$". For instance, $p_{\{a_{0:3},d_{0:5}\}}$ represents the proposition "there is danger in one of the areas $\{a_0 \dots a_3\}$ on one of the days $\{d_0 \dots d_5\}$". If the military coalition conducts an operation on day d_0 in the area a_0, it may want to alert $\mathcal{R}$ about the danger on the area in the specific day. However, sharing a strong opinion with $\mathcal{R}$, such

as $w_{p_{\{a_0,d_0\}}} = \langle 0.9, 0.0, 0.1\rangle$ may lead $\mathcal{R}$ to reveal that there will be a military operation in a_0 on d_0 to $\mathcal{F}$. In such situations, an obfuscation strategy could involve sharing a more general opinion such as $w_{p_{\{a_{0:3},d_{0:5}\}}} = \langle 0.9, 0.0, 0.1\rangle$ together with providing highly uncertain opinions for the specific areas and days, e.g., $w_{p_{\{a_0,d_0\}}} = \ldots = w_{p_{\{a_3,d_5\}}} = \langle 0.05, 0.05, 0.9\rangle$. The provided information would make $\mathcal{R}$ keep away from the dangerous areas without revealing sensitive information such as the locations and dates of the military operations to $\mathcal{F}$. This type of obfuscation is *successful* in the sense that the coalition's sensitive information is kept secret while allowing the information consumer ($\mathcal{R}$) to pursue its plans. Furthermore, the applied obfuscation does not lead to any decrease in the trustworthiness of the military coalition.

Shared sensory information can also be used to infer the type and location of sensors. Therefore, the military coalition can obfuscate the information before sharing it with $\mathcal{C}$ to hide the positions of its sensors and dates/locations of its classified activities. Similarly, insurgents ($\mathcal{F}$) obfuscate the information before sharing it with $\mathcal{R}$ to hide information about their activities and locations; the news agency ($\mathcal{N}$) obfuscates the news before broadcasting to hide the identity of its information sources.

Having provided some examples of the uses of obfuscation, we turn our attention to the interplay between trust and obfuscation and discuss reasoning under obfuscation.

V. Interplay between Trust and Obfuscation

A. Relationship between obfuscation and trust

There is a clear relationship between the level of obfuscation and trust between a provider and a consumer. That is, the provider's level of trust towards the consumer determines the level of obfuscation the provider can apply. In our scenario, the high level of trust between them means that military coalition members need to obfuscate information that they share with each other only slightly, while they will obfuscate more drastically information shared with the local army ($\mathcal{C}$) and NGOs ($\mathcal{R}$) to avoid revealing sensitive information. Lower levels of obfuscation (i.e., permitting more inferences) entail more *risk* for the producer than higher ones, as additional inferences could allow the consumer the knowledge necessary to take undesirable actions. That is, lower level of obfuscation may allow a consumer a spectrum of inferences that include sensitive ones (such as the dates or locations of classified activities), i.e., inferences that fall in the critical (red) area on the right-side of Figure 4. However, the producer might *trust* that the consumer will not exploit these opportunities, and when such a trust relationship exists, the level of obfuscation could be lowered.

Now, excessive obfuscation may make information received less useful for the consumer, which would then result in a decrease in the perceived trustworthiness of the information provider. Furthermore, certain types of obfuscation, such as the injection of errors in the information shared to hinder certain inferences, can blur the border between the obfuscation and *deception*. Deception is defined as the act of misleading another through intentionally false statements or fraudulent actions[2]. When obfuscation crosses the line into deception, the trust of the party being deceived towards the party deceiving will be significantly reduced, leading to a vicious cycle of decreasing trust, increasing obfuscation, and ultimately, decreased information sharing. Avoiding this cycle is therefore an important consideration for information providers and consumers who depend on long-term interactions with others. An interesting use study about personal deceptive practices, i.e., lying, driven by a Bayesian-network model can be found in the literature [8].

B. Reasoning under obfuscation

Obfuscation allows two parties to communicate and cooperate through sharing information while still protecting sensitive information. Without obfuscation, the risk of revealing sensitive information would prevent such communication and cooperation from occurring, leaving both parties worse off. Obfuscation is therefore a valuable tool in maintaining cooperation between parties in highly heterogeneous coalition settings. Given the utility of obfuscation, coalitions should be capable of reasoning in the presence of obfuscation. Obfuscation-aware reasoning may involve various steps where each aspect of information about the obfuscation must be considered. Consumers can have varying degrees of knowledge about the obfuscation of a particular piece of information. We can roughly categorize this knowledge as follows:

- complete knowledge of the obfuscation type and its extent;
- knowledge of obfuscation type but not its extent;
- knowledge that obfuscation is taking place (nothing else); and
- no obfuscation knowledge is available.

Awareness of the level at which obfuscation is taking place could suggest the remedial actions necessary when utilising obfuscated information. Contracts between parties explicitly specifying information sharing policies can provide a possible source of such knowledge.

Information providers and consumers often agree upon sharing specific pieces of information. Such agreements could be formalized as information sharing *contracts*, i.e., *QoI-level agreements* (QLAs), which express QoI expectations for received information. These QLAs may be further enriched by expressing the intended uses of information, i.e., the inferences to be made with the information provided. This, in turn, allows the provider to determine appropriate levels of obfuscation *a priori* based on the stated inferences and the trust the provider has toward the consumer.

In this way, such contracts determine the context in which the consumer should evaluate trustworthiness of the information provider. For example, if the consumer (e.g., the local army $\mathcal{C}$) states in the QLA that the information provided shall be sufficient to to estimate the position of a sniper within 20 meters of its true position with 10 seconds of the event of a

[2] *http://dictionary.law.com*

sniper shooting (these are QoI expectations), the provider (e.g., the military coalition) will be contractually responsible for providing information to support this localization task at the stated quality level. He will not be obliged to divulge anything more. In this case the consumer may not have grounds to rate the provider as untrustworthy if he does not deliver information with stricter QoI levels. This allows the provider to obfuscate the information to hinder other inferences, such as inferences that allow the positions of its sensors to be determined, without risking its trustworthiness as long as this obfuscation does not violate the agreement. Note that a consumer may always rate the provider along other dimension as well such as cost-effectiveness, capability, etc. He may then rank them higher or lower to other similarly trusted providers that may provide lower or higher QoI; further discussion on this is beyond the scope of this paper.

VI. Conclusions

Inference management techniques, which we summarily refer to somehow imprecisely as obfuscation, are necessary, and often desirable, serving to encourage information sharing in diverse coalitions by reducing the perceived risk for information providers while delivering value to its consumers. However, such techniques may also become problematic when assessing one's trust in the information received and its providers. There exists then a need for techniques which facilitate the identification of appropriate obfuscation methods, as well as techniques for reasoning with obfuscated information. Addressing this need requires an understanding of the complex relationship between obfuscation and trust. In this paper, we described the concepts of trust and obfuscation, and examined their relationship through a coalition scenario. We demonstrated how information consumers evaluate their trust in information providers using the subjective logic operator toolset. Then, we discussed, through our scenario, how obfuscation can be utilised to prevent sensitive inferences without decreasing information consumers' trust in the providers. Lastly, we introduced the concept of *QoI-level agreements* to serve as contracts between information consumers and providers to determine the context of information sharing and the level of obfuscation.

As a future work, we would like to leverage existing knowledge representation approaches to represent and reason about QoI-level agreements. We also want to examine how these agreements can be utilised within normative distributed systems, i.e., multi-agent systems, to promote information sharing without compromising sensitive information.

Acknowledgements

The authors would like to thank Supriyo Chakraborty for enlightening discussions during the course of this project.

This research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence and was accomplished under Agreement Number W911NF-06-3-0001. The views and conclusions contained in this document are those of the author(s) 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] D. E. Bakken, R. Parameswaran, D. M. Blough, A. A. Franz, and T. J. Palmer, "Data obfuscation: Anonymity and desensitization of usable data sets," *IEEE Security and Privacy*, vol. 2, no. 6, pp. 34–41, 2004.

[2] C. Bisdikian, L. M. Kaplan, M. B. Srivastava, D. J. Thornley, D. Verma, and R. I. Young, "Quality of sensor-originated information in coalition information networks," in *Network Science for Military Coalition Operations: Information Exchange and Interaction*, D. Verma, Ed. IGI Global, 2010, pp. 15–41.

[3] C. Bisdikian, M. Sensoy, T. J. Norman, and M. B. Srivastava, "Trust and obfuscation principles for quality of information in emerging pervasive environments," in *Workshop on Information Quality and Quality of Service for Pervasive Computing (IQ2S12)*, 2012.

[4] A. Jøsang and T. Bhuiyan, "Optimal trust network analysis with subjective logic," in *SECURWARE'08*, Cap Esterel, France, Aug. 25–31 2008.

[5] D. Gambetta, *Trust: Making and Breaking Cooperative Relations*. Blackwell, 1990.

[6] A. Jøsang, R. Ismail, and C. Boyd, "A survey of trust and reputation systems for online service provision," *Decis. Support Syst.*, vol. 43, pp. 618–644, 2007.

[7] M. Duckham and L. Kulik, "A formal model of obfuscation and negotiation for location privacy," in *3rd Int'l Conf. on Pervasive Computing, PERVASIVE 2005*, Munich, Germany, May 8–13, 2005.

[8] X. An, D. Jutla, and N. Cercone, "Reasoning about obfuscated private information: who have lied and how to lie," in *5th ACM Wkshp on Privacy in Electronic Society (WPES'06)*, Alexandria, Virginia, USA, Oct. 30–Nov. 3, 2006.