Summarizing my DHS Internship Experience

D. L. Roberts

September 7, 2006

Summarizing my DHS Internship Experience

David L. Roberts
College of Computing
Georgia Institute of Technology
Atlanta, GA
robertsd@cc.gatech.edu

1 Introduction

In this paper, I have been asked to address four main topics: 1) A description of my internship topic; 2) A description of my contributions to the project; 3) A discussion of research directions beneficial to the Department of Homeland Security (DHS); and 4) A discussion of the impact the internship experience had on my career aspirations. I feel the first three points can best be addressed using the contents of a paper my mentor, Dr. Tina Eliassi-Rad, and I have published based on our work this summer [Roberts and Eliassi-Rad, 2006]. Sections 2 - 5 are intended for this purpose and have been excerpted from that paper. I will conclude this paper in Section 6 with a discussion of the fourth point.

2 A Position on Value of Information for Evidence Detection

The sheer size and complexity of data sets in real-world applications have prompted many efforts on management and analysis of large complex networks. These networks store multi-source data with multi-modal and multi-relational properties in a human-understandable way. One of the associated challenges is to find evidence in support of or against a particular hypothesis. It is common to try to detect patterns in the data and use these patterns to perform inference or update beliefs in a set of hypotheses.

The published literature is full of algorithms for pattern matching, pattern mining, link analysis, and many others on such networks [Gallagher, 2006; Getoor and Diehl, 2005; Washio and Motoda, 2003]. A logical next step is to take these algorithms further and use their patterns/results as evidence. In this paper, we argue that any approach to evidence detection should consider the *value of information* (VOI) for the output of the data sources. Without such measurement, all sources and their data are considered "equal." This assumption is clearly false. For instance, using patterns that were generated from an unreliable data source is clearly not desirable. Our proposed VOI framework enables evaluation of detected patterns based on their qualities.

Throughout the remainder of this paper, we highlight related work from a number of fields. We go on to provide anecdotal evidence illustrating how VOI can be used to identify poor quality or incorrect information and therefore

avoid the inevitable reduction in accuracy of inference or beliefs. We provide an idealized definition of VOI as well as a technique for estimating it.

The problem of determining VOI is well-studied in various fields dating back to the 1940s. All existing approaches solve a variant of the following problem. Given a set of sources, which is the best (or best set) to obtain an observation (or a set of observations) from? In other words, an agent must determine the optimal "activation schedule" for the sources of information to maximize (or minimize) some objective. There are a number of approaches to solving this problem, based mostly on decision theory and/or information theory. However, measuring VOI for evidence detection differs from previous work in several ways.

First, prior work typically makes inherent assumptions about reliability of information [Horvitz and Rutledge, 1991; McCarthy, 1956; Shannon, 1948]. In particular, the traditional approaches of information and decision theory assume sources to be fully reliable. Data for evidence detection is often not fully reliable or even relevant. It originates from multi-modal sources – each with varying characteristics that can change as the world evolves.

Second, previous work typically characterizes the value of querying an information source which is a *pull* problem (as data is pulled from sources) [Heckerman et al., 1993; Zhang et al., 2002]. For the purpose of evidence detection, we are interested in understanding how to interpret data that we have already obtained or has been *pushed* to us. To understand this distinction, consider the following situation. You are buying a new car from a company that is known to produce a high quality product. Unfortunately, the company has redesigned the car for this year and you do not know if it is up to the usual standards. There are two ways you can proceed. One is to proceed under the assumption that the company's reputation is sufficient and they will likely not produce a bad car. Alternatively, you can test drive the car and determine if the product is consistent with what you know about the company. The first case, when you make the decision solely based on reputation, is an example of a pull problem. The second case, when you decide after test driving the vehicle, is an example of a push problem. The key distinction is whether the value of what the source (or company) produces is determined prior to (pull) or after (push) inspection of the product (or information). Our position is that in detecting evidence within a body of existing information, push methods are what need to be utilized. To use the methods developed for the pull problem, where inspection of the current information does not occur when evaluating quality or value, imposes an artificial handicap on the detection of evidence. The pull approach is warranted when there is a cost for obtaining information as is typically assumed in decision theory. Since evidence is available without cost, we can exploit the opportunity to inspect the current information and develop a measure for VOI more attune to shifts in information quality.

Third, we found little work that attempted to learn VOI and/or its components. Generally speaking, the majority of work in this area comes from the Information Fusion community [Rogova and Nimier, 2004]. The goal in information fusion is to combine multiple sources of information into one coherent representation. Often, the pre-fusion information is missing values, pertains to disjoint concepts, or may be unreliable. All of these as well as other properties of
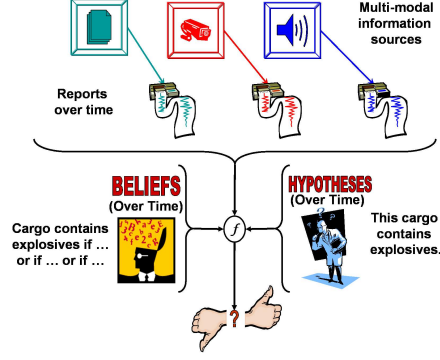
Figure 1: VOI Framework for Evidence Detection

the pre-fusion information must be taken into account when designing a fusion operator.

Moreover existing approaches in the fusion community, such as Delmotte et al. [1996], generally do not involve learning to characterize the quality of data. In this work, it is noted that the quality of knowledge produced by fusion is influenced by adequacy of the data, quality of the uncertainty model, and quality of the prior knowledge. Much of this work has focused on the improvement of an uncertainty model and has completely ignored the reliability of information. When it is considered, two measures of reliability are discussed: 1) The relative stability of the first order uncertainty; 2) The accuracy of the beliefs. It is assumed that the fusion operator will not introduce any residual uncertainty that is not due to the data itself. A fair amount of research has been devoted to the incorporation of reliability into fusion rules. In this research, the reliability measure comes in one of three forms: 1) It is encoded by external sources (*e.g.* context or an expert); 2) It is learned using training data; 3) It is constructed based on agreement of sources or consensus [Delmotte et al., 1996; Parra-Loera et al., 1991]. It is not, however, estimated prior to fusion. Historically, consensus models of reliability have taken one of the following two forms: 1) A degree of deviation between measurements of each source and the fusion result (*e.g.* posterior belief); 2) A measure of "inner trust" based on a pairwise degree of "likeliness" of agreement (or consensus) between sources. While this work is interesting and in some cases can improve inference performance, there is a notable problem with consensus measures of reliability. Specifically, lack of consensus is sufficient for low reliability but not necessary. Further, for VOI to be useful for evidence detection, it must be computed prior to fusion–and not as a function of its output.

Measuring VOI for evidence enables a solution to the problem of detecting information (*i.e.* evidence) with the highest *qualitative* value to either confirm a true hypothesis or disconfirm a false hypothesis.[1] VOI for evidence detection should: 1) capture information from possibly unreliable sources; 2) characterize the value of a body of existing information (a push problem); and 3) be tailored to improve inference and allow data triage.[2]

Figure 1 depicts our framework for VOI in evidence detection. At the top of the Figure there are a set of sources

---

[1]For a discussion on the degree to which a hypothesis is confirmed or disconfirmed, see Fitelson [2001].

[2]Note that even with anonymized data, one needs to evaluate the source.

3

Media	news-paper *1*	web paper *2*	news-paper *3*	web paper *3*	news-paper *5*	news-paper *6*	blog *7*
Reput.	very reliable	very reliable	somewhat reliable	unknown	somewhat reliable	very reliable	unknown

Figure 2: A Set of News Sources and Their Reputation

that are pushing reports (or information). The different icons in the sources are intended to represent the multi-modal data. The reports they generate evolve over time as the environment evolves. In the center of the Figure, there is an *f* inside a circle. To the left is a representation of a set of *beliefs* (*i.e.* prior knowledge or knowledge obtained earlier) that are evolving over time and to the right are the set of *hypotheses* that are (possibly) evolving over time. The *f* represents the combination of information, beliefs, and hypothesis into a quality metric. Note that this is not the same as fusion. In the fusion framework, the information from the sources would be combined to produce "manipulated" data. In our VOI framework, the information from the sources is combined to produce a measure of quality. This measure of quality can be used to detect evidence, inform inference, or even inform fusion. This quality (or VOI) measure is represented by the thumbs-up-thumbs-down icon at the bottom of the Figure and generally lies somewhere in between the extremes of useless and useful (*i.e.* it is not a binary measure of goodness).

To recap, by removing the data-equality assumption, information can be judged based on its quality. This should improve our ability to detect evidence, test hypotheses, and perform data triage [Kent, 1994; Schrage, 2005]. In the next section, we address the following questions: What should value of information mean for evidence detection? What are the basic parameters needed to characterize value of information? How should these parameters be measured and combined to compute a value for information? Finally, how should value of information be used in evidence detection?

3   An Anecdote in Favor of VOI

On Wednesday July 5th, 2006, the day after North Korea (DPRK) test fired missiles, countless news sources were publishing numerous articles with a great deal of information of varying quality. The following are two examples, loosely based on facts obtained from seven news sources in the early morning hours after the tests, of how VOI can help in the detection of quality evidence. The type and reputation of these sources are presented in Figure 2.

3.1   Missiles Crashed?

In this example we illustrate that multi-modal information sources can often seem to be reporting unrelated information but still support the same hypothesis and that the reputation of the sources has an influence on the conclusions that should be drawn. Of the seven sources consulted, sources 1, 3, and 5 stated that missiles landed in the Sea of Japan, sources 4 and 7, which are of unknown reliability, stated that missiles crashed after 62 seconds, and the remaining two

sources, 2 and 6, stated that missiles had flown 150 miles.

Despite the fact that the sources were saying three slightly different things and some of them were sources of unknown quality, we can easily infer (correctly) that missiles were fired and they did crash. Perhaps more interestingly, we can infer the speed of the missiles at approximately 8,700 MPH. However, with prior beliefs about the speed of intercontinental ballistic missiles (which travel approximately 15,650 MPH for the first few minutes of flight[3]) or prior beliefs about the location of the Sea of Japan relative to the DPRK, we can determine that some of the information is incorrect. Clearly, the most suspect information is the 62 second flight time reported by sources 4 and 7—both with unknown reputations. Since the other information from the sources with known reputations is corroborated by our beliefs, we can confidently discount the information from the suspect sources. Using this, we can attribute a low reliability to these sources and count their information with less weight in the future.

3.2   Japan and the DPRK at War?

The seven sources shown in Figure 2 were consulted over a period of a few hours in the order they appear. Over that period, the information given changed. Specifically, the first three sources each stated that no sanctions had been brought against the DPRK. The fourth source quoted a DPRK spokesperson as saying that "the imposition of sanctions would equate to a declaration of war." Finally, the last three sources reported that Japan had canceled its weekly ferry route to the DPRK which was responsible for a majority of the trade between the countries. However, the last source reported that this action by Japan was a form of sanctions.

This example illustrates two important issues that arise when dealing with evidence detection: timeliness and inaccuracy. First, strictly speaking, no conclusion can be drawn about whether or not sanctions have been brought against the DPRK since the first three sources and the last three sources conflict. However, as the information provided by the first three sources is older than that of the last three, it can be considered less accurate. This is a common assumption made when dealing with belief systems [Benferhat et al., 2002]. Second, ignoring the reputation of the sources, it would be logically sound to draw the conclusion that Japan and the DPRK are at war. However, given that two key pieces of information that enable this conclusion come from sources whose reputation is unknown, we can use our better judgment (or existing beliefs that war is not declared on such short notice and without explicit news coverage) to infer that they are actually not at war and the reliability of at least one of the unknown sources is low.

4   VOI for Evidence Detection

To measure value of information, we first need to define what constitutes information. A piece of information is the smallest amount of data needed to update the probability distribution of a hypothesis within a database of hypotheses.

---

[3]http://en.wikipedia.org/wiki/ICBM

Naturally, information is conditional on a chosen domain topic (e.g. Avian influenza) or a selected taxonomy (e.g. viral infections). This conditionality manifests itself in hypotheses and beliefs.

In information theory, VOI is a measure of uncertainty reduction (entropy). In decision theory, VOI is a measure of influence over a decision or choice of actions (value for taking an action). In either case, however, the measures are traditionally defined to characterize the expected value of *querying* an information source and not to provide an interpretation of the information the source has *already produced*. In contrast, for evidence detection, we define VOI to be a measure of the potential increase/decrease in believing a true/false hypothesis. Specifically, VOI for evidence detection is a measure of the ability for a body of information to *increase belief in a true hypothesis* (or set of hypotheses) or, conversely, *decrease belief in a false hypothesis* (or set of hypotheses). In the absence of an oracle that can inform us of the truth of hypotheses, we need algorithms to approximate VOI over time.

4.1 Temporal Tracking

The way in which time is considered when it comes to VOI for evidence detection is one way in which our approach is set apart from existing work in VOI. Specifically, we use time to distinguish between a *process* across a range of multiple time steps and an *event* at a single point in time. When dealing with evidence, we have the ability to both consider the total (or partial) history of information produced by a source as well as the information it has produced most recently. This allows us to characterize the expected value of a source of information over time in a way similar to the information and decision theoretic measures where the value is based on the distribution of information to be produced by the source. Additionally, we can also characterize the quality of the information that has been produced by the source at the current time step without the need for an expert to encode contextual knowledge about the reliability and/or validity of the source. For example, even completely reliable news sources have to print a retraction on rare occasions. If we look beyond simply characterizing the reliability of the source and make use of our existing beliefs, then we may be able to tell ahead of time that information from a reliable source is potentially faulty and may later be retracted. Using our VOI for an event, the information that will eventually be retracted can hopefully be identified early, before it may be a detriment to inference.

4.2 Components for Estimating VOI

We have identified three separate influences that we feel are responsible for how information is interpreted and that we will use as component measures to estimate VOI. They are *reliability R*, *independence I*, and *coherence C*. Each of these can be measured in a variety of ways.

We first define some notation: $S = \{S_1, \ldots, S_n\}$ is a set of information sources. $T = \{t_x, \ldots, t_y\}$ is a time interval and t_i represents a particular point in time. $H = \{H_1, \ldots, H_n\}$ represents a set of hypotheses. $B =$

Measurements on Sources, S, over Time Interval, T				
Context →	Objective		Subjective	
Memory →	Process	Event	Process	Event
Reliability, R	$R(S_i \mid H, T)$	$R(S_i \mid H, t_k)$	$R(S_i \mid H, B, T)$	$R(S_i \mid H, B, t_k)$
Coherence, C	$C(S_i, S_j \mid H, T)$	$C(S_i, S_j \mid H, t_k)$	$C(S_i, S_j \mid H, B, T)$	$C(S_i, S_j \mid H, B, t_k)$
Independence, I	$I(S_i, S_j \mid H, T)$	$I(S_i, S_j \mid H, t_k)$	$I(S_i, S_j \mid H, B, T)$	$I(S_i, S_j \mid H, B, t_k)$

Figure 3: A Taxonomy of Basic Components for VOI

$\{B_1, \ldots, B_n\}$ is a set of beliefs. Lastly, E represents evidence or the information produced by a source or set of sources at the given time step or time interval. This can be a single value (in the case of one source at one point in time) or as complex as a set of sets (in the case of a set of sources over a time interval) or anywhere in between. Figure 3 depicts the taxonomy to which our basic components for estimating VOI belong. As mentioned above, each of these components can be measured in a variety of ways. The categories are based on how memory/time (process vs. event) and beliefs/context (objective vs. subjective) are considered during the computation of the component measures. These component measures are described in slightly more detail below. In some cases, we provide an illustrative example of how the measure may be computed.

We add a few comments on objective and subjective measures. The process of calculating a component measure of VOI (or VOI itself) conditioned on a belief set is known as a subjective measure. If the computation is performed without a belief set, it is known as an objective measure. This is reflected in Figure 3 where B appears in the subjective measures but not in the objective measures of the taxonomy. In either case, all computations are made with respect to a set of hypotheses (see Figure 3). This conditioning on hypotheses is to provide an appropriate frame of reference (*e.g.* the value of knowing that a car is blue is zero with respect to the hypothesis that the car is of high quality).

4.2.1 Reliability

Reliability is intended to capture how frequently a source of information agrees with ground truth. In the context of news sources, reliability is interpreted as a measure of how frequently a particular news source reports information that turns out to be true. To assess the reliability of a source, we can consider statistical models that use the rate at which the source produces true positive or false positive information.

4.2.2 Coherence

Coherence captures agreement among sources. The interpretation of coherence among news sources is a measure of how frequently they agree. For example, take two conservative talk show hosts: Pat Robertson and Rush Limbaugh. As they are motivated by current events and politics, they frequently discuss many of the same topics on their shows. They are clearly independent as they are both free thinkers. However, they agree on many things, which makes them coherent sources of information.

4.2.3 Independence

Independence captures the causal dependence between sources (rather than simple agreement). It has a very intuitive interpretation. For example, articles from the New York Times and Al Jazeera that both report the same thing are likely to be independent but two articles from the Associated Press on the same topic are likely not independent. In the context of the probability model of uncertainty, given a belief set, independence is a conditional measure that captures the degree of autonomy between sources.

5 Concluding Thoughts

We have argued for the importance of considering a qualitative value of information content for evidence detection. We outlined how to value information based on three time-dependent components: reliability of a data source, independence between data sources, and coherence among data sources. We also raised several points with respect to temporal tracking. The bottom line is that evaluating information from real-world, dynamic data sources will provide us with a much-needed qualitative metric for both evidence detection and hypothesis testing.

We have also outlined a framework that allows the computation of VOI for both processes and events. This distinction between processes and events appears to be both novel and very powerful. The resulting potential increase in ability to inform inference is a fertile area for future research. Other work along these lines includes investigating appropriate learning methods for estimating each component function individually as well as learning how to best combine them. We also plan to conduct an extensive experimental study using additional data sets and begin looking at some of the alternate models of uncertainty (such as possibility theory or Dempster-Shafer theory).

Lastly, to fully enable triage, we have to look at how, specifically, VOI is used to select among the many possible combinations of sources. In our preliminary work, we assumed a fixed number of sources to be selected for inference. Ultimately, rather than selecting the strongest evidence produced by n sources, we would like this to be adapted at each time step to select the strongest evidence possible. Even with the limited experimentation described in this position paper, we see very encouraging results. We are excited to experiment with the full potential for this approach and pursue the many avenues of future research.

6 Career Aspirations

One of the most exciting things that arose as result of the internship environment provided by LLNL was the opportunity to get to know other DHS fellows and scholars from other fields and backgrounds. At times stretched well beyond my scientific comfort level, we debated what I felt were profound questions of scientific and philosophical importance to the homeland security community. Specifically, we sought to answer such questions as: Why do we

need security? What is the right policy for scientific funding for securing the homeland? How do we know when we have accomplished our goal (or can it be known)? What is our ethical responsibility as scientists in securing the homeland? As many of these questions have no clear cut answers, the thought experiments that are required simply to debate them intelligently are demanding.

At this point you may find yourself wondering what this has to do with my career aspirations. To this, I respond with the following: I feel that my internship this summer and my relationship with the DHS that has formed over the past year is the beginning of a lasting partnership. This partnership is one that I envision enabling me to continue to focus on answering these tough questions (as well as others) in both a philosophical manner as well as a scientific manner. In a philosophical capacity, I see myself continuing to debate and discuss both the questions mentioned above as well as many others. In the continued debate I see myself maintaining and nurturing a scientific interest in some of the many problems of importance to the DHS community.

It has long been my goal to be a research faculty member at the university level and my experience this summer has served to open my eyes to the opportunity to contribute to the DHS mission as an academic. Through learning about DHS centers of excellence and the new Institute for Discrete Sciences at LLNL, I have begun to think about how homeland security can continue to play a significant role in my ongoing research agenda. Most notably, my exposure to the interests of the DHS community at LLNL this summer has lead me to identify a potential dissertation topic. Further, due to my summer mentor's interest in the topic, I have taken the appropriate steps to keep the option of returning to LLNL next summer open. Looking a bit further into the future, while I still maintain my goal of working in academics, I have begun to think about the possibility to spending some time as a postdoc at a national lab or applying for a AAAS fellowship where I can spend more time thinking about both the philosophical and scientific goals of the DHS community as well as how my scientific pursuits can help to further enhance our nation's ability to achieve those goals.

7 Acknowledgments / Auspices

References

Salem Benferhat, Didier Dubois, Henri Prade, and Mary-Anne Williams. A practical approach to revising prioritized knowledge bases. *Studia Logica*, 70:105–130, 2002.

Francois Delmotte, Laurent Dubois, and Pierre Borne. Context-dependent trust in data fusion within the possibility theory. In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, volume 1, pages 538–543, October 11-14 1996.

Branden Fitelson. *Studies in Bayesian Confirmation Theory*. PhD thesis, Univ. of Wisconsin – Madison, 2001.

Brian Gallagher. The state of the art in graph-based pattern matching. In *Papers from the AAAI Fall Symposium - Capturing and Using Patterns for Evidence Detection*, Oct 2006.

L. Getoor and C.P. Diehl. Link mining: A survey. *SIGKDD Explorations - Special Issue on Link Mining*, 7(2):3–12, 2005.

David Heckerman, Eric Horvitz, and Blackford Middleton. An approximate nonmyopic computation for value of information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(3):292–298, 1993.

Eric Horvitz and Geoffrey Rutledge. Time-dependent utility and action under uncertainty. In *Proceedings of the Seventh Conference on Uncertainty in Artificial Intelligence*, pages 151–158, San Mateo, CA, 1991. Morgan Kaufman.

Sherman Kent. *Sherman Kent and the Board of National Estimates: Collected Essays*, chapter Words of Estimative Probability. Center for Intelligence Studies, 1994. URL `http://www.odci.gov/csi/books/shermankent/6words.html`.

John McCarthy. Measures of the value of information. *Proceedings of the National Academy of Sciences*, 42(8): 654–655, September, 15th 1956. URL `http://www.pnas.org/cgi/reprint/42/9/654.pdf`.

Ramon Parra-Loera, Wiley Thompson, and Ajit Salvi. Adaptive selection of sensors based on individual performances in a multisensor environment. In Vibeke Libby, editor, *Proceedings of SPIE – Volume 1470: Data Structures and Target Classification*, pages 30–36, August 1991.

David L. Roberts and Tina Eliassi-Rad. A position paper: Value of information for evidence detection. In *Papers from the AAAI Fall Symposium - Capturing and Using Patterns for Evidence Detection*, Oct 2006. URL `http://www.cc.gatech.edu/~robertsd/papers/roberts06-evidence.pdf`.

Galina L. Rogova and Vincent Nimier. Reliability in information fusion: Literature survey. In Per Svensson and Johan Schubert, editors, *Proc. of the 7th International Conference on Information Fusion*, volume II, pages 1158–1165, Mountain View, CA, Jun 2004. International Society of Information Fusion. ISBN 91-7056-116-8. URL `http://www.fusion2004.foi.se/papers/IF04-1158.pdf`.

Michael Schrage. What percent is 'slam dunk'? The Washington Post, February, 20 2005.

Claude E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 1948. URL `http://cm.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf`.

T. Washio and H. Motoda. State of the art of graph-based data mining. *SIGKDD Explorations - Special Issue on Multi-Relational Data Mining*, 5(1):59–68, 2003.

Yongmian Zhang, Qiang Ji, and Carl G. Looney. Active information fusion for decision making under uncertainty. In *Proceedings of the Fifth International Conference on Information Fusion (Fusion 2002)*, July 2002. URL `http://www.ecse.rpi.edu/homepages/qji/Fusion/fusion02_zhang.pdf`.