

Distributed Detection Under Information Constraints I: Computational Complexity ‡

N. S.V. Rao †

Department of Computer Science
Old Dominion University
Norfolk, VA 23529-0162
rao@cs.odu.edu

S.S. Iyengar

Department of Computer Science
Louisiana State University
Baton Rouge, LA 70803
iyengar@csc.lsu.edu

R.L. Kashyap

Department of Electrical Engineering
Purdue University
West Lafayette, IN 47907
kashyap@ce.ecn.purdue.edu

R.N. Madan

Division of Electronics
Office of Naval Research
Arlington, VA 22217-5000
madan@ocnr-hq.navy.mil

Abstract

We consider a distributed detection system consisting of a set of sensors $S = \{S_1, S_2, \dots, S_m\}$ and a set of objects $O = \{O_1, O_2, \dots, O_n\}$. There are *information constraints* on the system given by a relation $R \subseteq S \times O$ such that $(S_i, O_j) \in R$ if and only if S_i is capable of detecting O_j . Each $(S_i, O_j) \in R$ is assigned a *confidence factor* (a positive real number) which is either explicitly given or can be efficiently computed. Given that a subset of sensors have detected obstacles, the *detection problem* is to identify a subset $H \subseteq O$ that has the maximum confidence value. We consider the computational complexity of the detection problem, which depends on the nature of the confidence factor and the information constraints. Consequently, this problem exhibits a myriad of complexity levels: ranging from a worst-case exponential (in n) lower bound in a general case to an $O(m+n)$ time solvability. We show that the following simple versions of detection problem are computationally intractable: (a) deterministic formulation, where confidence factors are either 0 or 1; (b) uniform formulation where $(S_i, O_j) \in R$, for all $S_i \in S, O_j \in O$; (c) decomposable systems under multiplication operation. We then show that the following versions are solvable in polynomial (in n) time: (a) single object detection; (b) probabilistically independent detection; (c) decomposable systems under additive and non-fractional multiplicative measures; (d) matroid systems.

‡ Partially funded by an LEQSF grant from Louisiana State University Board of Regents.

† Also partially funded by Virginia's Center for Innovative Technology under grant #INF-90-015 and Old Dominion University Summer Faculty Research Fellowship for 1991.

1. Introduction

Distributed sensor systems consist of a number of sensors that cooperate, in a distributed manner, to achieve an objective based on the outputs of all sensors [21,22]. It has been realized by many researchers that there are limitations on the capabilities of a single sensor systems in a number of application-areas. Also, in several real-life systems, multiple sensors are a part of the design requirement. Often, the sensors are distributed, either geographically or functionally; thus the information obtained by the sensors has to be suitably coordinated, and a consolidated conclusion has to be distilled from the sensor data. The design, implementation and computational issues in multiple sensor systems are considerably more challenging than their counterparts in single sensor systems, for the issues due to the distributed information processing are seemingly absent in the latter. The literature on distributed sensor systems is extensive [2,9,10,20]; see [7] and references therein for some recent works. Comprehensive treatments on specialized topics such as spatial reasoning [8], sensor fusion in intelligent systems [11] also exist.

We consider a distributed detection system consisting of a set of sensors $S = \{S_1, S_2, \dots, S_m\}$ and a set of objects $O = \{O_1, O_2, \dots, O_n\}$. The *information constraints* of the system are given by a relation $R \subseteq S \times O$ such that $(S_i, O_j) \in R$ if and only if S_i is capable of detecting O_j . The sensor S_i produces an output of 1 when any object O_j , such that $(S_i, O_j) \in R$, has been sensed by S_i . Each $(S_i, O_j) \in R$ is assigned a *confidence factor* which could be a numerical value, or a value that can be computed. In general, the confidence factor is expressed as a function $f: 2^O \times 2^S \rightarrow R^+$, where R^+ is the set of non-negative real numbers. Here $f(A, B)$, for $A \subseteq O, B \subseteq S$, denotes the confidence that A is the set of objects in the work space when the sensors B produce outputs of 1. Let

$Sen(O_j) = \{S_i | (S_i, O_j) \in R\}$ and $Det(S_i) = \{O_j | (S_i, O_j) \in R\}$. Given that a subset of sensors $D \subseteq S$ have detected obstacles (i.e. produced an output of 1), the *detection problem* is to identify the subset of objects $H \subseteq O$ such that $f(H, D) = \max_{A \in 2^O} f(A, D)$. Under the assumption that $f(A, D)$ can be computed in $O(|A| + |D|)$ time, we are interested in the complexity of computing H .

The simple detection systems of Rao [14] where the detection is deterministic, are special cases of the present formulation. On the other hand, the probabilistic formulation of Demirbas [5] is a special case when $Sen(O_j) = S$, and the confidence factor is the probability measure and the events (appearances of objects in the workspace) are probabilistically independent. In terms of computational complexity, we show that this is a particularly easy case in that the detection can be carried out in $O(mn)$ time. Distributed detection problems based on probabilistic formulations have been extensively studied [3,5,17-19]. Note that in our formulation $f(\cdot)$ has lesser structure than a probability measure, and as a special case can correspond to a probability measure.

We consider the computational complexity of the detection problem; the most deciding factors are the type of information constraints and the nature of confidence factor. The detection problem exhibits a myriad of complexity levels: ranging from a worst-case exponential (in n) to an $O(mn)$ time solvability. At one extreme, this problem has an exponential (in n) lower bound for the computational complexity; it is particularly interesting to note that it seems to be computationally harder than NP-complete problems. At the other extreme, it becomes solvable in $O(mn)$ time under additive measure. Some interesting subclasses, based on Bayesian type (more specifically based on computing a posteriori probabilities) methods, fall out as part of our analysis. We are unaware of systematic complexity studies of these classes of problems; we show that these problems are polynomial time solvable under probabilistically independent events, but, become NP-hard in a general case. We also introduce the notion of *f-decomposable systems*, where $f(A, B)$ can be computed using divide-and-conquer algorithms. These systems contain some interesting subclasses; in particular, the *matroid* systems support greedy algorithms.

We show that the following simple versions of detection problem are computationally intractable (i.e. the optimization problems are NP-hard and the decision versions are NP-complete):

- (a) *deterministic formulation*, where the $f(\cdot)$ is interpreted as a probability measure on 2^O , and all detection probabilities are either 0 or 1;
- (b) *completely constrained system*, where $Sen(O_i) = S$ for all $O_i \in O$.

(c) *f-decomposable systems under a product operation* (details are given in Section 2.2).

Note that for practical purposes, the NP-Complete problems are too time-consuming to be useful for large inputs [6] (at the present time, it is not known whether these problems are solvable in polynomial time or not). The case (b) also corresponds to the case of Bayesian detection when the events are not guaranteed to be independent. We then show that the following versions are solvable in polynomial (in n) time:

- (a) *single object detection*;
- (b) *probabilistically independent detection*;
- (c) *decomposable systems under additive measures and multiplicative measures* where $f(A) \geq 1$ for all $A \in 2^O$;
- (d) systems where the family of unions of $Sen(O_i)$ forms a *matroid*.

Note that (b) corresponds to the Bayesian detection where the events are considered independent.

The organization of this paper is as follows: In Section 2.1, we consider the subclasses of the detection problem which are NP-complete. The *f-decomposable* systems that enable efficient computation of confidence factors using divide-and-conquer algorithms, are discussed in Section 2.2. In Section 2.3, we consider the cases where the detection problem is solvable in polynomial time.

2. Complexity of Detection Problem

A straight forward solution to the detection problem can be obtained by explicitly computing $\{f(A, D)\}_{A \in 2^O}$. This results in a prohibitively large complexity of $O(2^n(m+n))$. In its general formulation the detection problem is tantamount to that of picking the largest of 2^n real-numbers, and hence has a lower bound of complexity of $\Omega(2^n)$. However the properties of f and R can be utilized to expedite the computation.

Consider the set of all unions of $Sen(O_j)$ s given by

$$\xi = \{Sen(O_{i_1}) \cup Sen(O_{i_2}) \cup \dots \cup Sen(O_{i_k}) | \\ i_1, i_2, \dots, i_k \in [1, n], k \in [1, n]\}$$

and note that $D \in \xi$. For $A \in \xi$, let $\Gamma(A)$ denote the family of subsets of O such that each subset is related (under R) to the same subset A of S under the information constraints, i.e.

$$\Gamma(A) = \{B | B \subseteq O, \bigcup_{a \in B} Sen(a) = A\}.$$

Now constrain the system such that $f(B_1, A) = f(B_2, A)$ for $B_1, B_2 \in \Gamma(A)$, i.e. for each set of sensors with output 1, all the sets of potential objects will have the same confidence, and the choice of any such set of objects constitutes a solution to the detection problem. Consider the family

$\Psi = \{\Gamma(A)\}_{A \in \mathcal{E}}$. Thus in this case we have $\max_{A \in 2^O} f(A, D) = \max_{A \in \Psi} f(A, D)$. If $|\Psi| \geq c 2^n$ for some constant $c > 0$, the detection problem has a lower bound of $\Omega(2^n)$. However, if $|\Psi| \leq cp(n)$, for some constant $c \geq 0$ and some polynomial $p(n)$, the above argument cannot be applied to show an exponential lower bound. We show that this problem is NP-hard and contains several subproblems which are NP-complete. Thus, this problem is still computationally intractable.

2.1. Computationally Intractable Classes

We consider the class of NP-Complete problems that can be solved in polynomial time on a non-deterministic Turing machine [6]. Informally, in these problems it can be verified if $A=H$ for any $H \in \mathcal{O}$ and for any $A \subseteq \mathcal{O}$ in polynomial time, but there are exponential number of choices for A . We now show that the nature of f and R are orthogonal in causing NP-Completeness in certain subclasses of the detection problem; in particular we show the nature of either f or R could be damaging enough to make this problem intractable.

Multiple Object Detection:

Consider the case where detection is deterministic in that confidence function $f(A, D) = 1$ if $\bigcup_{a \in A} Sen(a) = D$ and $f(A, D) = 0$ otherwise. Thus $\max_{A \in 2^O} f(A, D)$ corresponds to finding if there exists a subset $A \subseteq \mathcal{O}$ that satisfies the information constraints imposed by R . This problem is called the *multiple object detection problem* and is shown to be NP-Complete by Rao [14] by reducing the set cover problem to this problem. Thus the nature of information constraints, R , alone is sufficient to make this problem NP-Complete.

Most Plausible Hypothesis:

At the other extreme, consider a *fully constrained system* such that $Sen(O_i) = S$ for all $O_i \in \mathcal{O}$, and let $f(\cdot)$ correspond to probability measure on 2^O . In this case computation of H corresponds to computing a most plausible explanation under the set cover model. This problem has been shown to be NP-hard by Reggia et al [15]. Thus trivializing the nature of the information constraints, R , in the detection problem does not make it any easier computationally. Notice that in this case $|\Psi| = 1$, but the detection problem is computationally intractable; thus, a polynomial bound on $|\Psi|$ does not make this problem solvable in polynomial time.

2.2. f -Decomposable Systems

We now introduce the notion of *f -decomposable systems*, where $f(\cdot)$ values can be computed using divide-and-conquer algorithms (see [1] for details on divide-and-

conquer algorithms). We say that a multiple sensor system is *f -decomposable* if, there exists an operator \square such that for every $A, B \subseteq \mathcal{O}$, $f(A \cup B, D) = \square(f(A, D), f(B, D))$, and \square is computable in $O(|A| + |B| + |D|)$ time. Now notice that for any $A \subseteq \mathcal{O}$, $f(A, D)$ can be computed in $O(|A| \log(|A|) + |D| \log(|A|))$ by using a straight-forward divide-and-conquer algorithm:

- (a) divide A into two sets A_1 and A_2 ;
- (b) recursively compute $f(A_1, D)$ and $f(A_2, D)$;
- (c) compute $(f(A_1, D), f(A_2, D))$;

The time complexity of this algorithm will be given by $T(|A|, |D|) = 2T(|A|/2, |D|) + O(|A| + |D|)$, which yields the desired complexity of $O(|A| \log(|A|) + |D| \log(|A|))$. This problem will still have $\Omega(2^n)$ lower bound if $|\Psi| \geq c 2^n$ for some $c > 0$. Now the question is the complexity of the detection problem for the case $|\Psi|$ is polynomially bounded. Note that this problem subsumes the multiple object recognition if R is unconstrained.

Consider the special case of a fully constrained system where \square corresponds to multiplication of integers and $f(\cdot)$ is integer valued. We pose a decision version of the detection problem as follows: Given R and a positive integer b , is there $H \subseteq \mathcal{O}$ such that $f(H, D) = b$? This problem can be shown to be NP-complete by establishing a polynomial time reduction from a well-known problem called the *subset product problem*, which is stated as follows [6]. Given a finite set A , size $s(a) \in \mathbb{Z}^+$ for each $a \in A$, positive integer b , is there a subset $A' \subseteq A$ such that the product of the sizes of the elements in A' is exactly b ? It is direct to see a reduction of this problem to the above version of the detection problem by identifying $f(\cdot)$ with the size function $s(\cdot)$.

2.3. Polynomial-Time Solvable Classes

Consider the single object case, where no more than one object is present in the region monitored by the sensors. In this case, we compute $f(\{O_i\}, D)$ for each $O_i \in \mathcal{O}$ and pick the object with highest $f(\cdot)$ value. This method has been used by Demirbas [4] in the special case where $f(\cdot)$ corresponds to the a posteriori probability computed from known a priori probabilities. The time complexity of this method is $O(nm)$, and the same algorithm can be employed when it is given that no more than a constant number of objects could be present in the workspace.

Now consider a completely constrained case where $f(\cdot)$ corresponds to the probability measures and the event of any obstacle being present in the workspace is independent of any other being present. Note that we allow multiple objects to be present in the workspace. Now we have $f(A \cup B, D) = f(A, D) f(B, D)$, where $A, B \in \mathcal{O}$, where $f(X, D)$ is the probability that the objects $X \subseteq \mathcal{O}$ are present in the workspace. In this case it is easy to see that $f(A \cup B, D) \leq \min(f(A, D), f(B, D))$ since each $f(\cdot)$ value is

fraction. Thus we have $\max_{A \in 2^O} f(A, D) = \max_{O_i \in O} f(\{O_i\}, D)$.

The latter can be computed in a straight-forward manner in $O(nm)$ time. In other words, this problem reduces to that of single object detection.

Now consider f -decomposable systems where the operation \square corresponds to addition on real numbers. In this case it is easy to see $\max_{A \in 2^O} f(A, D) = f(O, D)$, i.e. $H=O$. Similar result applies when \square is multiplication and $f(A) \geq 1$ for every $A \in 2^O$; but, the problem becomes computationally intractable if no constraints are placed on $f(A)$ s. Also, the above discussion holds when \square could be dynamically switched between the addition and multiplication operations.

We now state a definition of a matroid [12]. A subset system (E, Λ) is a finite set E together with a collection Λ of subsets of E closed under inclusion, i.e., if $A \in \Lambda$ and $A' \subseteq A$, then $A' \in \Lambda$. The elements of Λ are called *independent*. The *combinatorial optimization problem associated with (E, Λ)* is the following: Given $w(e) \geq 0$ for each $e \in \Lambda$, find an independent subset that has the largest possible total weight. Matroids have special structure such that a greedy algorithm will yield a solution for the combinatorial optimization problem. See [12] for a detailed treatment on matroid based algorithms. Now consider a f -decomposable system such that the subset system (S, ξ) is a matroid and the \square of the decomposable system is addition. The detection problem in this case can be solved as follows. We first compute $f(\{O_i\}, Sen(O_i))$ for each O_i . Then we use the following greedy algorithm to compute H :

```

begin
  I :=  $\phi$ ; H :=  $\phi$ ; C := O;
  while (I  $\neq$  D) do
    begin
      let  $O_k \in C$  such that
         $f(\{O_k\}, Sen(O_k)) = \max_{O_i \in C} f(\{O_i\}, Sen(O_i))$ ;
      if (I  $\cup$  Sen( $O_k$ )  $\subseteq$  D) then
        begin
          I := I  $\cup$  Sen( $O_k$ );
          H := H  $\cup$  { $O_k$ };
        end
      else
        C := C - { $O_k$ };
    end;
  output H;
end

```

The correctness of this method follows by straight-forward methods [12]. The computation of $f(\{O_i\}, Sen(O_i))$ for each O_i , has a time complexity of $O(s)$. The complexity of above procedure is $O(ns)$ since

there are n iterations and each iteration can be carried out in $O(s)$ time.

3. Conclusions

We consider a distributed detection system consisting of a set of sensors $S = \{S_1, S_2, \dots, S_m\}$ and a set of objects $O = \{O_1, O_2, \dots, O_n\}$. There are *information constraints* on the system given by a relation $R \subseteq S \times O$ such that $(S_i, O_j) \in R$ if and only if S_i is capable of detecting O_j . Each $(S_i, O_j) \in R$ is assigned a *confidence* (a positive real number) which is either explicitly given or can be computed. Given that a subset of sensors have detected obstacles, the *detection problem* is to identify a subset $\Psi \subseteq O$ that has the maximum confidence value. We consider the computational complexity of the detection problem, which depends on the nature of the confidence and the information constraints. Consequently, this problem exhibits a myriad of complexity levels - from a worst-case exponential (in n) lower bound in a general case to an $O(m+n)$ time solvability. We show that the following simple versions of detection problem are computationally intractable: (a) deterministic formulation, where confidence factors are either 0 or 1; (b) uniform formulation where $(S_i, O_j) \in R$, for all $S_i \in S, O_j \in O$; (c) decomposable systems under multiplication operation. We then show that the following versions are solvable in polynomial (in n) time: (a) single object detection; (b) probabilistically independent detection; (c) decomposable systems under additive and non-fractional multiplicative measures; (d) matroid systems.

Our study is a preliminary effort to exhibit the richness of the computational complexities of various versions of the detection problem. We feel that we barely scratched the surface of this fascinating problem which seems to embody many challenges. Some of the future investigations can be focussed in identifying practical and easily solvable versions, designing approximation algorithms for the NP-hard versions, and considering the cases where computational complexity of $f(A, D)$ is much higher than $O(|A| + |D|)$.

References

- [1] A. V. Aho, J. E. Hopcroft, J. D. Ullman, *The Design and Analysis of Computer Algorithms*, Addison-Wesley, 1974.
- [2] J. A. Barnett, DSN problems - an overview, Proc. DSN Workshop, CMU, 1978, pp. 37-40.
- [3] Z. Chair, P.K. Varshney, Optimal data fusion in multiple sensor detection systems, *IEEE Trans. Aerospace Electronic Syst.*, vol.22, no.1, 1986, pp. 98-101.
- [4] K. Demirbas, Maximum a posteriori approach to object recognition with distributed sensors, *IEEE Trans. Aerospace Electronic Syst.*, vol.24, no.3, 1988, pp. 309-313.
- [5] K. Demirbas, Distributed sensor data fusion with binary decision trees, *IEEE Trans. Aerospace Electronic Syst.*, vol.25, no.5, 1989, pp. 643-649.

- M.R. Garey, D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W.H. Freeman and Co., San Francisco, 1979.
- [7] S.S. Iyengar, R.L. Kashyap, R.B. Madan, Special Issues on Distributed Sensor Networks, *IEEE Trans. Syst. Man Cybernetics*, forth coming, 1991.
- [8] A. Kak, S. Chen (eds.), *Spatial Reasoning and multi-sensor fusion*, Morgan Kaufman Pub. Inc., 1987.
- [9] V. R. Lesser, D. D. Corkill, Functionally accurate, cooperative distributed systems. *IEEE Trans. Systems, Man and Cybernetics*, vol. SMC-11, no. 1, 1981, pp. 81-96.
- [10] V. R. Lesser, D. D. Corkill, J. Pavlis, L. Lefkowitz, E. Hudlicka, R. Brooks, A high level simulation test bed for cooperative distributed problem solving, Proc. DSN Workshop, Jan 6-7, 1982, MIT Lincoln Lab., PP. 247-270.
- [11] R.C. Luo, M.G. Kay, Multisensor integration and fusion in intelligent systems, *IEEE Trans. Syst. Man Cybernetics*, vol. SMC-11, no.1, 1981, pp. 81-96.
- [12] C.H. Papdimitriou, K. Steiglitz, *Combinatorial Optimization: Algorithm and Complexity*, Prentice-Hall, Englewood Cliffs, New Jersey, 1982.
- [13] Y. Peng, J.A. Reggia, Plausibility of diagnostic hypothesis: The nature of simplicity, *Proc. 5th Nat. Conf. on Artificial Intelligence (AAAI-86)*, 1986, pp. 140-145.
- [14] N.S.V. Rao, Synthesis of simple distributed detection networks, *Proc. IEEE Symp. on Parallel and Distributed Processing*, 1990, pp. 646-649, complete version to appear in *IEEE Trans. Systems, Man and Cybernetics*, 1991.
- [15] J.A. Reggia, D.S. Nau, P.Y. Wang, Y. Peng, A formal model of diagnosis inference. I. Problem formulation and decomposition, *Information Sciences*, vol. 37, 1985, pp. 227-256.
- [16] J.A. Reggia, D.S. Nau, P.Y. Wang, Y. Peng, A formal model of diagnosis inference. I. Algorithmic solution and application, *Information Sciences*, vol. 37, 1985, pp. 257-285.
- [17] a. R. Reibman, L.W. Nolte, Optimal detection and performance of distributed sensor systems, *IEEE Trans. Aerospace Electronic Syst.*, vol. AES-23, no.1, 1987, pp. 24-30.
- [18] F. A. Sadjadi, Hypothesis testing in a distributed environment, *IEEE Trans. Aerospace Electronic Syst.*, vol. AES-22, no.2, 1986, pp. 134-137.
- [19] R. R. Tenney, N.R. Sandell, Detection with distributed sensors, *IEEE Trans. Aerospace Electronic Syst.*, vol. AES-17, no.4, 1981, pp. 501-510.
- [20] R. Wesson, F. Hayes-Roth, J. W. Burge, C. Stasz, C. A. Sunshine, Network structures for distributed situation assessment, *IEEE Trans. Systems, Man and Cybernetics*, vol. SMC-11, no. 1, 1981, pp. 5-23.
- [21] D.P. White, DSN Communication system development, Proc. DSN Workshop, Jan 6-7, 1982, MIT Lincoln Lab., pp 151-162.
- [22] Y. Yemini, Distributed sensor network: An attempt to define the issues, Proc. DSN Workshop, CMU, 1978, pp. 53-66.