

NRL Report 7162

Probabilities of Successful Data Association for Tracking in the Plane when Using Maximum Likelihood Association Procedures

M. E. B. OWENS

*Information Systems Branch
Mathematics and Information Sciences Division*

February 2, 1971



PLEASE RETURN THIS COPY TO:

NAVAL RESEARCH LABORATORY
WASHINGTON, D.C. 20375
ATTN: CODE 2628

Because of our limited supply you are requested to return this copy as soon as it has served your purposes so that it may be made available to others for reference use. Your cooperation will be appreciated.

NDW-NRL-5070/2616 (1-84)

NAVAL RESEARCH LABORATORY
Washington, D.C.

CONTENTS

Abstract	ii
Problem Status	ii
Authorization	ii
INTRODUCTION	1
THE PROBLEMS IN GENERAL	1
PROBLEM I: ASSOCIATION OF INACCURATE POSITION DATA WITH ACCURATE POSITION DATA	2
The Maximum Likelihood Estimate of the True Position of \mathcal{U}	3
The Probability the Maximum Likelihood Procedure Will Yield the True Position of \mathcal{U}	3
PROBLEM II: ASSOCIATION OF POSITION DATA WITH SPEED INFORMATION	5
The Maximum Likelihood Estimate of the Position of \mathcal{U} at Time t_2	6
The Probability the Maximum Likelihood Procedure Will Yield the True Position of \mathcal{U}	6
PROBLEM III: ASSOCIATION OF POSITION DATA WITH COURSE AND SPEED INFORMATION	8
The Maximum Likelihood Estimate of the True Position of \mathcal{U} at Time t_2	8
The Probability the Maximum Likelihood Will Yield the True Position of \mathcal{U}	8
EXAMPLES	10

ABSTRACT

Data association is the determination of which subset of data contained in a larger set is related to a specific entity. The problem of data association is examined in an ocean surveillance environment in which the entities are vessels and the available data referring to the state of these vessels is either position, position and speed, or position, course, and speed. It is assumed that the vessels are uniformly and independently distributed over the ocean region of interest and that the total number of vessels in the region is known. For each data mix the maximum likelihood procedure for associating data referring to a particular vessel is given. Then a derivation is given for the probability that the data associated using this procedure actually refers to the state of the particular vessel. These probabilities are not only a function of the data mix available but also a function of the accuracies of the data and the number of vessels in the area of interest.

PROBLEM STATUS

This is an interim report; work on the problem is continuing.

AUTHORIZATION

NRL Problem B01-06
Project XR008-05-01, Task No. 50501

Manuscript submitted July 15, 1970.

PROBABILITIES OF SUCCESSFUL DATA ASSOCIATION FOR
TRACKING IN THE PLANE WHEN USING MAXIMUM
LIKELIHOOD ASSOCIATION PROCEDURES

INTRODUCTION

During the analysis and evaluation of proposed ocean surveillance systems, the analyst is often asked to study the feasibility of associating data which is related to vessels at sea and which is to be generated by one or more sensors at the same or different times. By data association we mean the determination of which subset of data contained in a larger body is related to a specific vessel. The problem of data association is most difficult when the sensors under consideration produce only geometric data (e.g., position, course, and speed) that is related to unidentified vessels, and in this situation data association is usually based on a maximum likelihood procedure*. To determine how well data can be associated in a given environment, the analyst usually resorts to oversimplified analytic models or to large-scale computer simulation. In most cases the oversimplified analytic models yield incomplete results. The implementation of computer simulations is time consuming, and simulation models are sometimes employed to answer questions which are out of the context of the model's original purpose.

We will formulate and solve analytically three problems which correspond to three different data mixes. In each case the solution is the probability that the data related to any one given vessel will be associated correctly when the maximum likelihood procedure is used. In those situations where the problems are applicable, the analyst can use these probabilities to examine the effects of the availability and accuracy of different types of information on the association process. These probabilities also should have application in the design of data association models in which they could be employed as prior probabilities.

THE PROBLEMS IN GENERAL

In each problem we approximate the region of the ocean by a planar region, and we assume that the vessels in the region are uniformly and independently distributed with a known density (number of vessels per unit area). To be specific, if there are n vessels in the region and the region has area a' , we define the vessel density as

$$\rho = n/a'.$$

We assume that a' and n are large and that ρ is small relative to unity.

Among the vessels in the region, there is one of particular interest; hereafter, this vessel will be called U . Suppose at an initial time t_1 we are given information concerning vessel U , the information depending on the problem according to Table 1. Suppose we

*For example, see R. W. Sittler, "An Optimal Data Association Problem in Surveillance Theory," I. E. E. E. Trans. MIL-8 (No. 2) April 1964

Table 1.
Available Information Concerning Vessel \cup at Time t_1
as a Function of Problem Number

Problem Number	Position	Accuracy of Position	Speed Information	Course Information
I	Yes	Poor	No	No
II	Yes	Good	Yes	No
III	Yes	Good	Yes	Yes

are then given the set of positions at time t_2 , where $t_2 = t_1 + \tau$, of all the vessels in the region, but we are not told which of these positions, is the position of vessel \cup . Let this set be denoted by

$$\{(x_i, y_i): i = 1, 2, \dots, n\},$$

and assume that this set is a sample from the set of random vectors

$$\{(X_i, Y_i): i = 1, 2, \dots, n\}$$

which are uniformly and independently distributed over the region.

Using the information concerning the state of \cup at time t_1 , we can predict the position of \cup at time t_2 , and with this predicted position we can select the maximum likelihood position of \cup at time t_2 from the set

$$\{(x_i, y_i)\}.$$

In each of the three problems, we will outline this maximum likelihood selection procedure. After this trivial task is completed, we will then calculate the probability that the position to be selected using this procedure is the true position of \cup at time t_2 or, equivalently, that the data referring to the state of \cup will be associated successfully.

PROBLEM I: ASSOCIATION OF INACCURATE POSITION DATA WITH ACCURATE POSITION DATA

Here

$$t_1 = t_2.$$

Along with the set $\{(x_i, y_i)\}$ we are given a vector (x_0, y_0) together with the fact that (x_0, y_0) is a sample from the random vector (X_0, Y_0) which has the distribution

$$N \left\{ \begin{pmatrix} a \\ b \end{pmatrix}, \Sigma \right\}.$$

We assume that (a, b) , the true position of vessel \cup , is unknown and that

$$\Sigma = \begin{pmatrix} \sigma_X^2 & 0 \\ 0 & \sigma_Y^2 \end{pmatrix},$$

where σ_X and σ_Y are known.

At this point we should examine the physical interpretation of this problem. At a fixed moment in time we are given the positions of all the vessels in the region of interest, but we do not know which vessel is at a particular position. We are also given the

"approximate" position of vessel U . On the basis of this "approximate" position, we will attempt to select the position of U from the set of exact positions.

The Maximum Likelihood Estimate of the True Position of U

First we apply the transformation

$$T = \begin{pmatrix} \sigma_X^{-1} & 0 \\ 0 & \sigma_Y^{-1} \end{pmatrix}$$

to the plane, and we adopt the notation

$$\begin{pmatrix} x'_i \\ y'_i \end{pmatrix} = T \begin{pmatrix} x_i \\ y_i \end{pmatrix}$$

for $i = 0, 1, 2, \dots, n$ and

$$\begin{pmatrix} a' \\ b' \end{pmatrix} = T \begin{pmatrix} a \\ b \end{pmatrix}.$$

Under this transformation, we note that (x'_0, y'_0) is a sample from the random vector (X'_0, Y'_0) , which is distributed according to

$$N \left\{ \begin{pmatrix} a' \\ b' \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

and that the density is now

$$\rho' = \sigma_X \sigma_Y \rho.$$

Observe that the random vectors (X'_i, Y'_i) , for $i = 1, 2, \dots, n$, are uniformly distributed over the transformed region.

It follows that the maximum likelihood estimate of (a', b') , or equivalently (a, b) , is that point (x'_m, y'_m) selected from the set $\{(x'_i, y'_i)\}$ which is nearest to (x'_0, y'_0) .

The Probability the Maximum Likelihood Procedure Will Yield the True Position of U

Given that the true position of U is (x_m, y_m) , define the random variables A and R by

$$A = \min_i \left| (X'_i, Y'_i) - (X'_0, Y'_0) \right|,$$

where i ranges over the set $\{1, 2, \dots, m-1, m+1, \dots, n\}$, and

$$R = \left| (X'_0, Y'_0) - (X'_m, Y'_m) \right|.$$

We see that A is the random variable denoting the distance from the "approximate" position of U to the nearest position which is not the true position of U . Also we see that

R is the random variable denoting the distance from the true position of \mathcal{U} to the "approximate" position of \mathcal{U} . If P_I denotes the probability that the maximum likelihood procedure will yield the true position of \mathcal{U} , it follows that

$$P_I = P(R < A) .$$

Before we can calculate P_I , we must know the distributions of A and R . It is well known that the probability density function of R is

$$f_R(r) = re^{-r^2/2}$$

for r positive. As for the random variable A ,

$$F_A(a) = 1 - P(A \geq a) ,$$

and

$$P(A \geq a) = P \left[\begin{array}{l} \text{none of the random vectors } (X'_i, Y'_i) \\ (i = 1, 2, \dots, n; i \neq m) \text{ is within a} \\ \text{circle of radius } a \text{ centered at} \\ (X'_m, Y'_m) \end{array} \right]$$

$$\approx \left(1 - \frac{\pi a^2 \rho'}{n} \right)^n$$

$$\approx e^{-\pi a^2 \rho'}$$

for a positive. Under the assumptions in the previous section, these approximations introduce a negligible error; hence, we will assume

$$P(A \geq a) = e^{-\pi a^2 \rho'}$$

Differentiating F_A with respect to a , we see that

$$f_A(a) = 2\pi a \rho' e^{-\pi a^2 \rho'}$$

for a positive.

Now

$$P_I = P(R < A)$$

$$= \int_0^\infty \int_0^a f_A(a) f_R(r) dr da$$

$$= \int_0^\infty 2\pi a \rho' e^{-\pi a^2 \rho'} (1 - e^{-a^2/2}) da$$

$$\begin{aligned}
&= 1 - \int_0^{\infty} 2\pi a \rho' \exp[-a^2 (\frac{1}{2} + \pi \rho')] da \\
&= 1 - \frac{\pi \rho'}{\frac{1}{2} + \pi \rho'} \\
&= \frac{1}{1 + 2\pi \rho'}.
\end{aligned}$$

Upon examination of

$$P_I = \frac{1}{1 + 2\pi \rho \sigma_X \sigma_Y}, \quad (1)$$

we see that P_I is bounded by zero and one and that P_I is a decreasing function of ρ , σ_X , and σ_Y . Therefore, as the density and/or the errors in the "approximate" position of \cup increase, P_I decreases.

PROBLEM II: ASSOCIATION OF POSITION DATA WITH SPEED INFORMATION

Suppose we are told that at time t_1 the position of vessel \cup is (a_1, b_1) and that (a_1, b_1) is a sample of the random vector (X_*, Y_*) which has the uniform distribution over the region. Further, suppose we are told that during the time interval (t_1, t_2) \cup travels with constant velocity and that the speed of \cup is s , where s is a sample of the random variable S which has the normal distribution with mean μ_S and variance σ_S^2 . We assume that μ_S and σ_S^2 are known, s is unknown, and μ_S and σ_S^2 are such that the probability that s is negative is negligible. Since no information about the course of \cup is known to us, we implicitly assume that the true course of \cup is a sample from the uniform distribution over the interval $(0, 2\pi)$.

Next we are given the set $\{(x_i, y_i)\}$ of the unidentified positions of all the vessels in the region at time t_2 . Let (a_2, b_2) denote the position of \cup at time t_2 . We know that

$$(a_2, b_2) = (x_m, y_m) \in \{(x_i, y_i)\}$$

for some m such that $1 \leq m \leq n$. It was previously remarked that the set $\{(x_i, y_i)\}$ is a random sample of the set of random vectors $\{(X_i, Y_i)\}$, which are uniformly and independently distributed over the region. From the discussion in this section we know that

$$P \left[\left| (X_*, Y_*) - (X_m, Y_m) \right| = \tau S \right] = 1, \quad (2)$$

where the random vector (X_*, Y_*) is uniformly distributed over the region. Can (X_*, Y_*) and (X_m, Y_m) both be uniformly distributed over the region and Eq. (2) be true? Strictly speaking, the answer is no. But if the region is large relative to the expected distance traveled and we disregard the problems encountered at the boundary of the region, the answer to this question is yes for all practical purposes; hence, we will suppose the answer is yes.

The Maximum Likelihood Estimate of the True Position of \cup at Time t_2

Since we are assuming implicitly that the course of \cup is a sample from the uniform distribution over the interval $(0, 2\pi)$, no consideration needs to be given to the course in estimating (a_2, b_2) . Hence, the maximum likelihood estimate of (a_2, b_2) , or (x_m, y_m) , is $(x_{\hat{m}}, y_{\hat{m}})$, where \hat{m} is such that

$$\left| \left| (x_{\hat{m}}, y_{\hat{m}}) - (a_1, b_1) \right| - \tau\mu_S \right| \leq \left| \left| (x_i, y_i) - (a_1, b_1) \right| - \tau\mu_S \right|,$$

for $i = 1, 2, \dots, n$. In other words the estimate is that point which belongs to the set $\{(x_i, y_i)\}$ such that the distance from the point to the circle of radius $\tau\mu_S$ centered at (a_1, b_1) is minimized.

The Probability the Maximum Likelihood Procedure Will Yield the True Position of \cup

Define the random variable R by

$$R = |\tau S - \tau\mu_S|$$

and note that R has the probability density function

$$f_R(r) = \frac{\sqrt{2}}{\sigma_S \tau \sqrt{\pi}} \exp \left[-r^2 / (2\tau^2 \sigma_S^2) \right]$$

for r positive. Next we define the random variable A by

$$A = \min_i \left| \left| (X_i, Y_i) - (a_1, b_1) \right| - \tau\mu_S \right|,$$

where i ranges over the set $\{1, 2, \dots, m-1, m+1, \dots, n\}$. Now

$$P(A \leq a) = 1 - P(A > a)$$

$$= P \left[\begin{array}{l} \text{none of the random vectors } (X_i, Y_i), \\ i \neq m \text{ is within distance } a \text{ from a} \\ \text{circle of radius } \tau\mu_S \text{ centered at} \\ (a_1, b_1) \end{array} \right]$$

$$\approx 1 - \exp \left\{ -\pi \rho \left[(a + \tau\mu_S)^2 - (a - \tau\mu_S)^2 \right] \right\}$$

$$= 1 - \exp(-4\pi\rho\tau\mu_S a)$$

for a positive. Hereafter, the above approximations will be taken as equality. Differentiating with respect to a , we find that

$$f_A(a) = 4\pi\rho\tau\mu_S \exp[-4\pi\rho\tau\mu_S a]$$

for a positive.

Let P_{II} denote the probability we seek, and note

$$\begin{aligned}
 P_{II} &= P(R < A) \\
 &= \int_0^\infty \int_r^\infty f_A(a) f_R(r) \, da \, dr \\
 &= \int_0^\infty \frac{\sqrt{2}}{\sigma_S \tau \sqrt{\pi}} \exp \left[-r^2 / (2\tau^2 \sigma_S^2) - 4\pi\rho\tau\mu_S r \right] \, dr \\
 &= 2 \exp \left[8 (\pi\rho\tau^2 \sigma_S \mu_S)^2 \right] \int_{4\pi\rho\tau^2 \sigma_S \mu_S}^\infty \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \, dz .
 \end{aligned} \tag{3}$$

From Eq. (3) we see that

$$P_{II} = 2e^{c^2/2} \int_c^\infty \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \, dz , \tag{4}$$

where

$$c = 4\pi\rho\tau^2 \sigma_S \mu_S .$$

It follows that $P_{II} \rightarrow 1$ as $c \rightarrow 0$, and by an application of L'Hospital rule and the fundamental theorem of calculus, we have $P_{II} \rightarrow 0$ as $c \rightarrow \infty$. Figure 1 is a graph of P_{II} as a function of c .

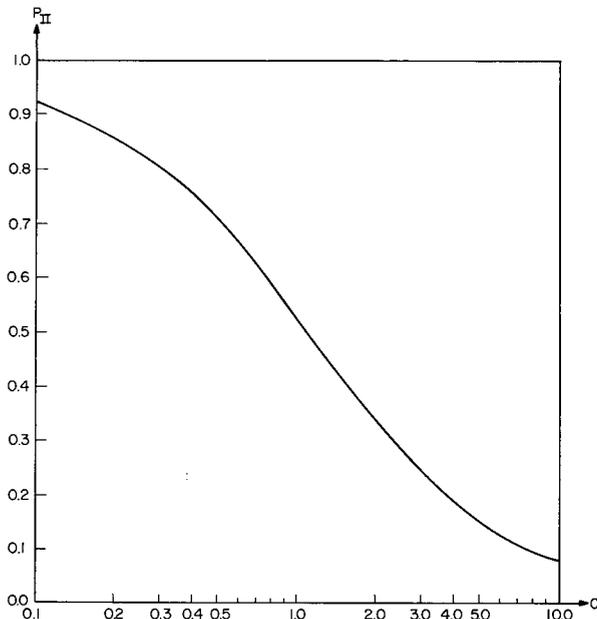


Fig. 1 — P_{II} as a function of the product $c = 4\pi\rho\tau^2 \sigma_S \mu_S$

**PROBLEM III: ASSOCIATION OF POSITION DATA WITH COURSE
AND SPEED INFORMATION**

Problem III is exactly the same as Problem II except we are told that the unknown course of vessel \mathcal{V} is a sample θ of the random variable Θ , which is normally distributed with known mean μ_{Θ} and known variance σ_{Θ}^2 . We assume that Θ and S are independent and that

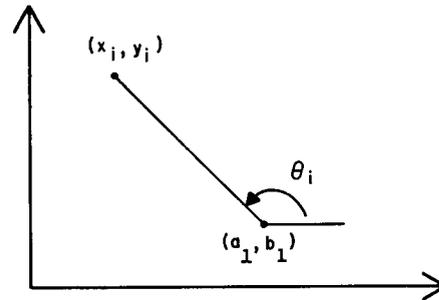
$$P(|\Theta - \mu_{\Theta}| > \pi)$$

is negligible.

The Maximum Likelihood Estimate of the True Position of \mathcal{V} at Time t_2

Let θ_i , for $i = 1, 2, \dots, n$, be as in Fig 2.

Fig. 2 - The definition of θ_i



and let each θ_i be adjusted so that

$$\mu_{\Theta} - \pi \leq \theta_i < \mu_{\Theta} + \pi .$$

The maximum likelihood estimate of (a_2, b_2) , or equivalently (x_m, y_m) , is $x_{\hat{m}}, y_{\hat{m}}$, where $\hat{m} (1 \leq \hat{m} \leq n)$ is such that

$$\left[\frac{[|(x_i, y_i) - (a_1, b_1)| - \tau\mu_S]^2}{\tau^2\sigma_S^2} + \frac{(\theta_i - \mu_{\Theta})^2}{\sigma_{\Theta}^2} \right]$$

is minimized when $i = \hat{m}$.

The Probability the Maximum Likelihood Will Yield the True Position of \mathcal{V}

Define the random variable R by

$$R = \left[\frac{(\tau S - \tau\mu_S)^2}{\tau^2\sigma_S^2} + \frac{(\Theta - \mu_{\Theta})^2}{\sigma_{\Theta}^2} \right]^{1/2}$$

and note R has the probability density function

$$f_R(r) = r e^{-r^2/2}$$

for r positive. Next define the random variable A by

$$A = \min_i \left[\frac{[|(x_i, y_i) - (a_1, b_1)| - \tau\mu_S]^2}{\tau^2\sigma_S^2} + \frac{(\theta_i - \mu_\theta)^2}{\sigma_\theta^2} \right]^{1/2}$$

where i ranges over the set $\{1, 2, \dots, m-1, m+1, \dots, n\}$ and θ_i is the random variable which corresponds to (X_i, Y_i) as the number θ_i corresponds to (x_i, y_i) . It follows that

$$P(A \leq a) = 1 - P \left[\begin{array}{l} \text{none of the random vectors} \\ (X_i, Y_i), i \neq m, \text{ fall within} \\ \text{the region } Q \end{array} \right];$$

where

$$Q = \left\{ (x, y): \frac{[|(x, y) - (a_1, b_1)| - \tau\mu_S]^2}{\tau^2\sigma_S^2} + \frac{(a - \mu_\theta)^2}{\sigma_\theta^2} \leq a \right\}$$

where a corresponds to (x, y) as θ_i corresponds to (x_i, y_i) . To calculate this probability, we must first calculate the area of Q , and

$$\begin{aligned} \iint_Q dx dy &= \iint_Q r dr da \\ &= 2 \int_0^{a\sigma_\theta} \int_{\tau\mu_S - g(a)}^{\tau\mu_S + g(a)} r dr da, \end{aligned}$$

where

$$g(a) = \tau\mu_S (a^2 - \theta^2/\sigma_\theta^2)^{1/2}.$$

Calculation of the above integral yields

$$\iint_Q dx dy = \pi\mu_S\sigma_\theta\tau^2 a^2.$$

Hence,

$$P(A \leq a) \approx 1 - \exp(-\pi\rho\mu_S\sigma_\theta\tau^2 a^2)$$

for a positive, and we will accept the approximation as an equality. The probability density function of A is

$$f_A(a) = 2\pi\rho\mu_S\sigma_\theta\tau^2 a \exp(-\pi\rho\mu_S\sigma_\theta\tau^2 a^2).$$

for a positive.

As before, let P_{III} be the probability that the maximum likelihood estimate of (a_2, b_2) will be (a_2, b_2) . It follows that

$$\begin{aligned}
 P_{III} &= P(R < A) \\
 &= \int_0^\infty \int_r^\infty f_A(a) f_R(r) \, da \, dr \\
 &= \int_0^\infty r \exp \left[-r^2 \left(\frac{1}{2} + \pi \rho \mu_S \sigma_S \sigma_\Theta \tau^2 \right) \right] dr \\
 &= \frac{1}{1 + 2\pi \rho \mu_S \sigma_S \sigma_\Theta \tau^2}.
 \end{aligned} \tag{5}$$

We see that P_{III} is a decreasing function of the various parameters and it is bounded by zero and one.

EXAMPLES

Consider a region of the ocean in which the shipping density is 40 vessels per 100,000 sq. naut mi, and suppose we will be given the unidentified positions at time t_2 of all the vessels in the region.

To illustrate Problem I, suppose we will be told that the position of vessel \mathcal{V} (which is of interest to us) at time t_2 is (x_0, y_0) and that the true position of \mathcal{V} is within 30 naut mi of (x_0, y_0) with 90% confidence. Given that this confidence region was derived from a circular normal distribution, it follows that

$$\chi_{(2, 0.9)}^2 \pi \sigma_X \sigma_Y = \pi (30)^2$$

and

$$\sigma_X = \sigma_Y = 13.98 \text{ naut mi.}$$

From Eq. (1) the probability that the maximum likelihood position of \mathcal{V} , which is selected from the set of unidentified positions, will be the true position of \mathcal{V} is

$$P_I = \frac{1}{1 + 2\pi \rho \sigma_X \sigma_Y},$$

where ρ is the density per unit area. For this example a simple computation yields

$$P_I = 0.67.$$

To examine another situation, suppose we will be given the position at time t_1 of vessel \mathcal{V} , where

$$t_2 = t_1 + 3,$$

and that the speed of \mathcal{V} from time t_1 to t_2 will be a sample from a normal distribution with mean 15 knots and standard deviation 3 knots. Given that \mathcal{V} will maintain a constant velocity during the time interval of interest, we will be able to select the maximum likelihood position at time t_2 from the set of unidentified positions. By Eq. (4) the probability that the selected position will be the true position of \mathcal{V} is

$$P_{II} = 2e^{c^2/2} \int_c^\infty \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz,$$

where

$$c = 4\pi\rho\tau^2\sigma_S\mu_S,$$

and in this example

$$P_{II} = 0.319.$$

The final example is exactly the same as that of the preceding paragraph except we will be given that the course of \mathcal{U} is a sample from a normal distribution with mean $\pi/2$ radians and standard deviation $\pi/6$ radians. We recall from Eq. (5) that the probability the maximum likelihood position of \mathcal{U} will be the true position of \mathcal{U} at time t_2 is

$$P_{III} = \frac{1}{1 + 2\pi\rho\mu_S\sigma_S\sigma_\Theta\tau^2}.$$

It follows that

$$P_{III} = 0.652$$

for this case.

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY <i>(Corporate author)</i> Naval Research Laboratory Washington, D. C. 20390		2a. REPORT SECURITY CLASSIFICATION Unclassified	
		2b. GROUP	
3. REPORT TITLE PROBABILITIES OF SUCCESSFUL DATA ASSOCIATION FOR TRACKING IN THE PLANE WHEN USING MAXIMUM LIKELIHOOD ASSOCIATION PROCEDURES			
4. DESCRIPTIVE NOTES <i>(Type of report and inclusive dates)</i> An interim report on a continuing NRL problem.			
5. AUTHOR(S) <i>(First name, middle initial, last name)</i> M. E. B. Owens			
6. REPORT DATE February 2, 1971	7a. TOTAL NO. OF PAGES 15	7b. NO. OF REFS 1	
8a. CONTRACT OR GRANT NO. NRL Problem B01-06	9a. ORIGINATOR'S REPORT NUMBER(S) NRL Report 7162		
b. PROJECT NO. XR008-05-01 Task No. 50501			
c.	9b. OTHER REPORT NO(S) <i>(Any other numbers that may be assigned this report)</i>		
d.			
10. DISTRIBUTION STATEMENT Approved for public release; distribution unlimited.			
11. SUPPLEMENTARY NOTES		12. SPONSORING MILITARY ACTIVITY Department of the Navy Naval Electronic Systems Command Washington, D. C. 20390	
13. ABSTRACT Data association is the determination of which subset of data contained in a larger set is related to a specific entity. The problem of data association is examined in an ocean surveillance environment in which the entities are vessels and the available data referring to the state of these vessels is either position, position and speed, or position, course, and speed. It is assumed that the vessels are uniformly and independently distributed over the ocean region of interest and that the total number of vessels in the region is known. For each data mix the maximum likelihood procedure for associating data referring to a particular vessel is given. Then a derivation is given for the probability that the data associated using this procedure actually refers to the state of the particular vessel. These probabilities are not only a function of the data mix available but also a function of the accuracies of the data and the number of vessels in the area of interest.			

DD FORM 1 NOV 65 1473 (PAGE 1)

13

S/N 0101-807-6801

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Correlation Data association Tracking Maximum likelihood						