



governmentattic.org

"Rummaging in the government's attic"

Description of document: United States Cryptologic History, Sources in Cryptologic History, Volume 4, A Collection of Writings on Traffic Analysis, Vera R. Filby, Center For Cryptologic History, National Security Agency, 1993

Requested date: 02-August-2012

Released date: 11-June-2013

Posted date: 24-June-2013

Source of document: National Security Agency
Attn: FOIA/PA Office (DJ4)
9800 Savage Road, Suite 6248
Ft. George G. Meade, MD 20755-6248
Fax: 443-479-3612 (ATTN: FOIA/PA Office)
[On-Line Request Form](#)

The governmentattic.org web site ("the site") is noncommercial and free to the public. The site and materials made available on the site, such as this file, are for reference only. The governmentattic.org web site and its principals have made every effort to make this information as complete and as accurate as possible, however, there may be mistakes and omissions, both typographical and in content. The governmentattic.org web site and its principals shall have neither liability nor responsibility to any person or entity with respect to any loss or damage caused, or alleged to have been caused, directly or indirectly, by the information provided on the governmentattic.org web site or in this file. The public records published on the site were obtained from government agencies using proper legal channels. Each document is identified as to the source. Any concerns about the contents of the site should be directed to the agency originating the document in question. GovernmentAttic.org is not responsible for the contents of documents published on the website.



NATIONAL SECURITY AGENCY
CENTRAL SECURITY SERVICE
FORT GEORGE G. MEADE, MARYLAND 20755-6000

Serial: MDR-68578
11 June 2013

This responds to your 2 August 2012 request for the declassification review of the following NSA document: *A Collection of Writings on Traffic Analysis*. The material has been reviewed under the Mandatory Declassification Review (MDR) requirements of Executive Order (E.O.) 13526 and is enclosed. We have determined that some of the information in the material requires continued protection from public release.

Some portions deleted from the document are currently and properly classified in accordance with E.O. 13526. The information denied meets the criteria for classification as set forth in Section 1.4 subparagraphs (c) and (d) and remains classified TOP SECRET, SECRET and CONFIDENTIAL as provided in Section 1.2 of E.O. 13526.

Additionally, section 3.5 (c) of E.O. 13526 allows for the protection afforded to information under the provisions of law. Therefore, the names of NSA/CSS employees and information that would reveal NSA/CSS functions and activities have been protected in accordance with Section 6, Public Law 86-36 (50 U.S. Code 402 note).

Since your request for declassification has been partially denied you are hereby advised of this Agency's appeal procedures. Any person denied access to information may file an appeal to the NSA/CSS MDR Appeal Authority. The appeal must be postmarked no later than 60 calendar days after the date of the denial letter. The appeal shall be in writing addressed to the NSA/CSS MDR Appeal Authority (DJ5), National Security Agency, 9800 Savage Road, STE 6881, Fort George G. Meade, MD 20755-6881. The appeal shall reference the initial denial of access and shall contain, in sufficient detail and particularity, the grounds upon which the requester believes the release of information is required. The NSA/CSS MDR Appeal Authority will endeavor to respond to the appeal within 60 working days after receipt of the appeal.

Sincerely,

A handwritten signature in black ink, appearing to read "Blake C. Barnes", followed by a horizontal line.

BLAKE C. BARNES
Chief
Declassification Services

Encl:
a/s

sources in
cryptologic
history
number 4

~~TOP SECRET~~

NO. 114

united states cryptologic history

A Collection of Writings on Traffic Analysis



~~THIS DOCUMENT CONTAINS CODEWORD MATERIAL~~
~~NOT RELEASABLE TO CONTRACTORS~~

Declassified and approved for release by NSA on 06-11-2012 pursuant to
E.O. 13526 - MDR 68578

CH-E32-93-05

national security agency
central security service

~~TOP SECRET~~

This monograph is a product of the National Security Agency history program. Its contents and conclusions are those of the authors, based on original research, and do not necessarily represent the official views of the National Security Agency. Please address divergent opinion or additional detail to the Center for Cryptologic History (E324).

Contents of this publication should not be reproduced or further disseminated outside the U.S. Intelligence Community without the permission of the Director, NSA. Inquiries about reproduction and dissemination should be directed to the Center for Cryptologic History, National Security Agency, Fort George G. Meade, MD 20755-6000, ATTN: E324.

UNITED STATES CRYPTOLOGIC HISTORY

Sources in Cryptologic History

Volume 4

A Collection of Writings on Traffic Analysis

Vera R. Filby



CENTER FOR CRYPTOLOGIC HISTORY

NATIONAL SECURITY AGENCY

1993

Table of Contents

	Page
Foreword	vii
Preface	ix
A Capsule History of Traffic Analysis	1
The First Quarter Century	
Fundamentals	
What Good is Traffic Analysis?	5
Establishing Communications Norms	11
Chatter Patterns: A Last Resort	15
A Note about NRs	25
Views of the Automated Future	
Some Thoughts concerning Traffic Analysis	
Mechanization	27
Cleaning the Augean Stables or How Much	
TA Can a Computer Do?	33
Traffic Analysis or [] Data Transmission	
Systems	41
Traffic Analysts of the Future	45
Callsign Stories	
[]	49
Recovery of a Vietnamese Communist Callsign	
System	71
True Base: Two Tales	73
Maybe It's Related to the Phases of the Moon	81
[]	87
Miscellany	
Area Studies and Their Place in Traffic	
Analysis	101

The 1970s

Fundamentals

P.L. 86-36
EO 1.4.(c)

The Reality of Communications Changes		105
A Note about Organizing TA Problems		109
		111

Mechanization

Address to Traffic Analysis Mechanization		
Forum	MAJ. GEN. JOHN E. MORRISON	113
Introduction to Traffic Analysis Mechanization ..	ROBERT S. BENJAMIN	115
Automation of a TA Process	TIM MURPHY	129
How Clean Does a Database Need to Be?		139
The Hand Is Not Quicker than the Eye		141

Traffic Analysis and Mathematics

The Lost Indicator	DR. RALPH W. JOLLENSTEN	143
Applications of Set Theory to Traffic Analysis		151
TA, CA, Logic, Math - Where Do They Intersect (TACALOMA)		155

Anxieties

Let's Not Lose Our TA Skills		171
Letters		175

Miscellany

Simplicity in Color	C. GAROFALO	179
The Impact of ARDF on Traffic Analysis	ALLEN L. GILBERT	181
Barometer - Readers' Comments		183
How Many Angels Can Stand on the Head of a Case Notation?		185

The 1980s and Beyond

Fundamentals

Traffic Analysis: A Current Perspective	JAMES HOPPER	187
---	--------------	-----

TA and Computers	
There's a New World Coming--Are You Ready?	195
In Pursuit of: Faster Horses, Younger Women, Older Whiskey, and More Money	199
Computerizing Traffic Analysis	203
Professional Concerns	
Traffic Analysis: Specialty without a Portfolio	215
New Areas of Interest	
[Redacted Box]	221
Hail and Farewell	
The Future of Traffic Analysis	241
Valedictory of a Traffic Analyst	245
JOSEPH STARR	
Author's Biography	249
Sources and Additional Readings	251

Foreword

If the professional field of traffic analysis has had its ups and downs, the need for its skill and art has never really diminished. Like Mark Twain, its early demise has been reported at regular intervals, but, happily, has always proved premature. Even though the discipline may change as a career field, the value to the cryptologic mission of contributions made by traffic analysis - and traffic analysts - continues.

Captain Duane Whitlock, USN, a veteran of cryptologic activity on Corregidor before World War II, speaking at the 1992 Symposium on Cryptologic History at NSA, emphasized and reemphasized the indispensability of traffic analysis. In that period of crisis before Pearl Harbor, traffic analysis worked closely to support cryptanalysis, and during periods when systems were unreadable, TA constituted the only means of following developments in the armed forces of a potential enemy.

Half a century after the attack on Pearl Harbor, we need to understand that despite more sophisticated collection and processing, traffic analysis still makes basic and vital contributions to the national mission.

Vera Filby has collected a number of thought-provoking articles on traffic analysis into one handy volume. This reader serves many purposes, not the least of which is to stimulate us into considering how the field has changed, what roles traffic analysis has played in the past, and what it can do for the cryptologic mission in the future.

Reduced to slogans, this is a good book for times like these, and these are good times for a book like this.

DAVID A. HATCH
Director,
Center for Cryptologic History

In 1989 NSA Senior Executive [] proposed the creation of a series of readers in cryptology to make more readily available some of the literature in the several cryptologic disciplines that has accumulated over the years in the National Security Agency. The present collection of occasional articles on traffic analysis (TA) is offered to the cryptologic community as a first contribution to what it is hoped will be a continuing program. The articles reflect the times and circumstances of their creation and thus present a picture of traffic analysis across a span of nearly half a century.

Sources were mainly the *NSA Technical Journal* (published 1956-1981), *COMMAND* (1968-73), *Cryptolog* (1974-present), and *Cryptologic Quarterly* (1981-present). Articles on various aspects of TA, along with letters to the editor and TA problems, published in these and other journals, amounted to some 300-350 pages, making selection imperative. Several of the articles describe TA techniques. Many concern the problems, both technological and professional, attendant on the automation of TA. Others recount TA stories and successes. Some originated as briefings or addresses.

Choices for this collection were hard since almost all the candidates were worthy of inclusion. The selections are arranged by subject within three chronological periods: the first twenty to twenty-five years, the 1970s, and the 1980s. The TA problems were omitted for lack of space. It is hoped they may eventually be published separately.

The items included may seem a scant production for forty-odd years. But they are personal writings, voluntary offerings to communicate ideas, to teach, to present an argument, to tell a story. They are on a quite different plane from the technical and related production of traffic analysis: SIGINT technical reports; astronomical numbers of technical records, supplements, files, messages, working aids, and weekly TA notes; professional studies, reviews, and development plans; and training aids and manuals. A training manual, *Radio Traffic Analysis*, by Robert S. Benjamin, published in 1955, was a major literary production. This exhaustive treatise was the training standard and bible for traffic analysts for many years.

VERA R. FILBY

A Capsule History of Traffic Analysis

THE FIRST QUARTER CENTURY

During the early years of the Armed Forces Security Agency and NSA, little was published to record the development of TA. But in the mid-1960s, evidence of an attention to TA as a discipline and career and concern for its status and future began to appear in written articles. This was in tune with the Agency's growing concern for the state of cryptologic disciplines, which resulted in the establishment of the professionalization program. In 1967 the beginnings of a TA library were established, and in 1968 the Techniques and Standards group, P1, and the TA Career Panel sponsored a TA workshop. Papers presented at its several sessions were later collected in response to a call for TA documentation. In October 1968 the first issue of *Command* was published, and the following year the Communications Analysis Association was formed.

Computers have been used in TA applications since the 1940s, but it was in the mid-1960s that the first wave of the computer revolution began to cause tremors in the work force. Many analysts were distrustful and apprehensive, both for their discipline and for their careers; but others were ready and eager to participate in the development of automation. Communication between analyst and computer programmer was a serious problem. Part of the eventual solution was in training traffic analysts to become computer programmers.

THE 1970s

The momentum begun in the late 1960s led to the Traffic Analysis Mechanization Forum in February 1970, where twenty-four briefings were presented. This was followed by the Traffic Analysis-Mathematics Symposium in May 1971 with twenty briefings. Compilations of these briefings were published in P1.

The level of energy, thought, and activity in the study of traffic analysis declined in the mid-1970s. Automation, decreased need for TA in targets where the communications structure had been largely recovered or where communications had become exploitable, and limitations in Agency fiscal resources reduced the size and depressed the morale of the TA work force. Early retirements, resignations, and moves into data systems, management, and other areas thinned out the ranks. Training requirements fell so low that the sole TA instructor remaining in the National Cryptologic School (NCS) had time for other duties.

~~CONFIDENTIAL~~

THE 1980s AND BEYOND

Although the decline in Agency strength had reversed by the beginning of the 1980s, TA as a career remained at the same level, despite the fact that the volumes of traffic pouring into the data banks continued so vast that much of it could not even be looked at, while at the same time new communications and advanced systems appeared and had to be dealt with. Many analysts were still trying to get used to – or resigned to – computers, and they suffered prolonged frustration and dissatisfaction with systems they found unhelpful. Nevertheless, by the mid-1980s new waves of computerization had swept through the Agency, and a computer environment eventually prevailed.

A growing awareness of a need for cryptologic renewal characterized the early 1980s. The prospect of a future of struggle against new and increasingly complex signals underscored the need to enhance the vigor, skills, and creativity of the work force. It became clear also that future problems would call for multidisciplined approaches.

In the 1980s studies were undertaken to assess the situation and plans formed to deal with it. Programs and projects, both major and minor, were set in motion. One of these was RELOAD. Project RELOAD's stated purpose was to restore analytic and reporting responsibilities to the field, where during many years of centralization analytic skills had weakened. RELOAD became the driving force for rejuvenation and change both at headquarters and overseas, and its effects continued to proliferate throughout the decade.

Under RELOAD, a Traffic Analysis Working Group (TAWG) began in 1986 to examine the traffic analysis career field. The group's final report, presented as the Traffic Analysis Focus Plan in 1989, recorded its finding that major systemic reforms were needed and provided a list of proposals and actions taken. One of the proposals was the creation of an Intelligence Analysis career field combining TA and Intelligence Research -- an idea that had been put forth in earlier years but discarded. This time it was accepted, and beginning in 1987 working groups developed career structures, established professional standards, and identified training needs. Their prolonged efforts culminated in the inauguration of the Intelligence Analysis career field on 1 January 1991.

RELOAD and technical change resulted in a greatly increased need for training. To meet it, the NCS recruited more instructors and greatly expanded the use of adjunct faculty. By 1989 the TA faculty were teaching or developing a total of nineteen courses, and more than forty adjunct instructors were certified to teach one or more courses. To recognize outstanding talent, the TAWG established the Gold Nugget Award; the first was presented in 1989.

The NCS held a TA Curriculum Review in mid-December 1990. During the course of the presentations, it appeared that the prospect for TA in the 1990s ranges from the conventional to the experimental, from traditional net reconstruction and callsign recovery of low-level military targets to invention of strategies for solving the most sophisticated digital signals structures. The need for TA is undiminished, and the challenge for the profession is greater than ever.

~~CONFIDENTIAL~~

~~HANDLE VIA COMINT CHANNELS ONLY~~

~~CONFIDENTIAL~~

P.L. 86-36

Acknowledgment

Thanks are due to for her encouragement and support in the making of this collection. Without the access to her files she so generously offered, it could not have been produced.

VERA R. FILBY

~~CONFIDENTIAL~~

~~HANDLE VIA COMINT CHANNELS ONLY~~

The First Quarter Century

What Is Good Traffic Analysis?¹

ROBERT H. SHAW²

What is good traffic analysis? The posing of this question in these words by Mr. G. H. Vergine has raised in my mind a large number of points that while by no means original, seldom seem to be available in one place at one time. This paper is likely, therefore, to take on something of the appearance of a list. I do not apologize; I think a list is needed, and I hope that some of the topics enumerated here will be expanded in workshops to come.

1. To develop the concept of good traffic analysis, it is necessary first to be clear on the prior question of what traffic analysis is, and what its relationship may be to other SIGINT disciplines.

1.1. Historically, SIGINT (or COMINT, as it was almost exclusively at first) was initially conceived as a unified whole, the response of vaguely named communications intelligence entities to the presence of exploitable transmissions. The great growth of communications during and after the Second World War and the great diversification of techniques of transmission, encryption, keying, and read-out, led to an inevitable specialization of the analytic people, and in due course cryptanalysis, traffic analysis, signal analysis, substantive analysis, intercept, collection, computery and all the rest became, first, fields of primary concern to different units, and, later, disciplines of their own. Ineluctably, SIGINT came to be looked at as the sum of autonomous parts rather than as a conveniently divided whole; and now interdisciplinary working is something to be striven for, rather than the obvious circumstance of professional life. All these disciplines are interwoven. A traffic analyst has no vested interest in staking out a preserve of his own called traffic analysis; he has a vested interest in contributing by his own techniques to what is known through all techniques about a target.

1.2. The classic definitions of traffic analysis date from the Second World War. Typical might be that of the U.S. Navy (DNSC5, *Communications Instructions*, 1944, Chapter 4): "Traffic analysis is the obtaining of intelligence from communications by means other than cryptanalysis." This sort of early definition brigades many newer disciplines in as constituent parts of the discipline of traffic analysis and thus creates much jurisdictional controversy today. What about SIT (Special Identification Techniques) substantive analysis, and so on? What about collateral? The confusion is obvious in today's world, where efforts to restrict traffic analysis are resisted by the traffic analyst, and where the main object of SIGINT gets lost in a slather of hurt feelings and administrative slack. It is worth considering that the naming of what one does is of much less consequence than the doing.

1. This article was originally published as *P1 Informal No. 10 (T/A #2)*, dated March 1967.

2. Dr. Shaw, one of the "brightest and best" at NSA and its predecessor agencies, died in early March 1971. See the necrology in the *NSA Technical Journal*, Vol. XVI, No. 2, Spring 1971.

1.3. Technically speaking, the one set of techniques always excluded from traffic analysis is cryptanalysis. At the outset, no one was in any doubt about what that meant: it meant that traffic analysts did not try to decrypt texts of messages. Often they gave information to the people who did; often they gave information to the people who tried to solve the callsign systems; but traffic analysts were concerned with what were called message externals – everything that could be learned about a message and its many kinds of meaning by studying that message in the form in which it was uttered by the originating command or bureau. The raw material of traffic analysis is traffic: messages, chatter, callsigns, times, frequencies, signal strengths, addresses, authenticators, net structures, message lengths, precedences, plaintext transmissions, etc.

1.4. Certain of these raw materials of traffic analysis turned out in the long run to be of sufficient complexity to require the creation of subdisciplines to deal with them. Perhaps the first to come into prominence was the matter of callsigns. Endless reams have been written arguing whether cryptanalysis may sometimes be involved, but that such cryptanalysis is heavily dependent on traffic analytic identification of commands through uses of various callsigns (so-called "continuities"), and that the results of such cryptanalysis are of interest to the SIGINT producing agencies only. The mere labeling of traffic is in itself of no more lasting interest than the discovery that a dash followed by two dots represents the letter *D*. Good traffic analysis identifies continuities, and perhaps even equates them with commands; good cryptanalysis recovers callsign systems from those continuities; the resulting identifications are useful in many ways to text cryptanalysts, to substantive analysts, and to traffic analysts. The nonproducing world is unruffled by even the most brilliant results in this area.

1.5. Traffic analysis has been nibbled away at in another way: a highly automated discipline of signal analysis and SIT has grown up. Spreading from advanced collection procedures on the one hand and from advanced identification and location techniques on the other, this discipline both answers questions and asks them. To the traffic analyst it is a source of very refined raw material. He pays for it by having to pass his results back so that the procedure may be improved. Dealings among traffic analysts, signals analysts, and SIT people are also of local concern only and are valuable only insofar as they enhance the value of the information about locations, movements, actions and intentions, which is the real product of SIGINT and the ultimate justification for a SIGINT effort.

1.6. A principal function of traffic analysis is the determination of the activity of a country, or of a function of that country, by deduction from the characteristics of signals. Plaintext signals can be of considerable help in what they say, in the relationships they imply, and in the essentialia that are not said. It follows that traffic analysts should give grave thought to all the plain text they see, in headings, texts, or chatter. On some circuits there is so much plain text that it can be used as a source of intelligence in its own right, neglecting its origin. It was to deal with this situation that the discipline of substantive analysis was created and set apart from traffic analysis. The relationship of a traffic analyst to substantive analysts should therefore be that of information exchange: the traffic analyst refers a wordy circuit to the substantive analyst, who in return tells the

traffic analyst what commands, or what kinds of commands, are talking, and what those commands are up to. Once again each helps the other, and their mutual conversation in its raw form is not of interest outside the producing community.

1.6.1. On circuits only lightly involved in plain text, the traffic analyst must expect to do both jobs.

1.6.2. Particular note should be made of the great desirability of [REDACTED]

[REDACTED] Cryptanalysts, traffic analysts, and substantive analysts have a common interest in this project and should have equal opportunity to contribute to it. The results may show up anywhere, so that the files should be generally accessible within the producing community.

1.7. And that reminds us that people are forever telling traffic analysts something relating to their problem. This so-called "collateral" is intended to assist the traffic analyst in the interpretation of what he sees. There is a tendency on the part of many traffic analysts to busy themselves at once to confirm or deny the collateral. This rather misses the point, although occasionally it is possible and valuable to do so. Collateral is *already* known; one must not expect to make the welkin ring by repeating it aloud, even if it is not contradicted by indications in the traffic. The collateral is made available to the analyst so that something new may emerge. *Verbum sap*

2. What, then, should good traffic analysts be doing? When we know that, we can evaluate their work meaningfully.

2.1. They work toward the identification of stations and, through them, of commands. In the course of this, they may be lucky enough to uncover previously unknown subordinations (so-called "order of battle"), although there is a continuous and active danger of confusing communication with command. The real purpose of these identifications, apart from an occasional new and unexpected recovery, is.

2.2. The establishment of continuity. The purpose of continuity is to permit the following from day to day of the traffic of a known command, in order to find out what that command is up to. The study of the activities of known commands is a principal justification for traffic analysis.

2.3. There are by-products at this point, prominent by-products that have a tendency to become primary objects in the minds of analysts. One of these by-products is the material necessary to solve callsign systems. No one outside the producing community cares about this, in spite of the number of reports one sees. (Have you ever considered the probable reaction of the secretary of defense on learning that today the Second Field Army is known as 7X03? He is unlikely to declare a national holiday.) Another is the material necessary to analyze the activity of the commands. This also is of internal interest only. (Go ahead; *tell* the secretary of defense that 7X03 had 114 messages today but only 111 yesterday, and see what he says.) If consumers demand reports of such material, they are

very specifically criticizing the performance of the traffic analysts; they are, in fact, telling the producing agency that it has done an incomplete or incompetent job.

2.4. Traffic analysts should make every effort to be current. Working with the superficialities of traffic, they are generally dealing in transient operation phenomena, and nothing is so old as yesterday's newspaper.

2.5. They support other SIGINT activities. They provide cryptanalysts with link data and with busts; they provide signal analysts with confirmation; they provide substantive analysts with material. They must not stop reading when they come to a BT; to do so is to create a jurisdictional stupidity which, if persisted in, can undo any meaningful understanding of the traffic.

2.6. They produce intelligence (or, in the common barbaric jargon, "intelligence information"). That means something beyond listing the raw material from which intelligence may be generated, something beyond wild guessing about the meaning of observed phenomena. (Navy readers will recall with glee the WWII TA report that the U. S. Navy had got a new, secret ship capable of incredible speed, "Heard off Cuba at 0913A, it was heard again at 1753Z at Port Lyautey; its callsign today is NERK.") It is not even the writing up of an intellectual marking signal: "We are still here, but we can't think of anything to say." Silence is golden; it takes more than gold, therefore, as openers for the silent mouth.

3. It might be said, therefore, that good traffic analysis is

3.1. *Traffic analysis*. It is not reporting, remembering, reviewing, or regurgitating. It is not a listing of phenomena. It is not the publication of intelligence of interest only to producers. It is information about the doings of organizations, information pertinent to the decision-making function of executive departments of our government.

3.2. *Defensible*. Any piece of TA output should have an *apparatus criticus* that will be fully explanatory and that will stand painstaking examination. Without it, a TA report is an unsubstantiated guess that will simply promote wild speculation and bring its originator into disrepute.

3.3. *Timely. Verbum sap.*

3.4. *Informative*. It may even contain predictions; this is greatly to be desired, if the predictions are carefully thought out and carefully justified. Predictions of routine communications/changes are, of course, beneath contempt except as working aids; predictions of fluctuation of target activity can be of the greatest value. Defensive traffic analysis predicted the Okinawa invasion and permitted the JCS to reschedule what might otherwise have been a disaster.

3.5. *Comprehensive*. The traffic analyst has the broadest scope of any SIGINT professional. His raw material is anything transmitted (except the plain texts of encrypted messages; if he can get them, he can use even those). It should never be said of a TA result that it failed to consider some known fact. A good traffic analyst, therefore, is a

person who knows how to balance comprehensiveness and timeliness: he is not panicked into publishing raw material, nor is he coerced into over-researching an item whose value lies in its immediate delivery

3.6. *Concise.* There is little to say that cannot be said in ten sentences. The footnoting, of course, can (and probably often should) take up a dozen pages, but the point of the item should be seen at once.

3.7. *Accurate. Verbum sap.*

4. What I am really suggesting is that the traffic analyst should begin each day with sixty seconds of meditation on what he is going to do, and end it with sixty seconds of meditation on what he has done. And honesty is a great virtue.

Establishing Communications Norms

P.L. 86-36



Communications norms can be described as what traffic can be expected, when it can be expected, and how it can be expected. The establishing of communications norms (referred to for the remainder of this paper as norms) is one of the most underrated procedures in COMINT. It is probably because the techniques in establishing norms appear so self-evident that analysts seldom "have time" to really know their target.

One might ask, why establish norms? In a period of high intercept volume and low analytic resources, it is often said, "I don't have the time - other things are more important." Well, let's take the time here and see if other things really are more important. To do this, you would just have to pin the analyst down as to what are the other things that are more important. After much hesitation and more vague and nebulous statements, you will probably find that the analyst is referring to what I'll call the production of "obvious intelligence." Obvious intelligence is what can be derived from a message after it has been converted to plain language. This is not intended to imply that obvious intelligence is not important, but frequently (too frequently) our translations do little more than provide a consumer with the same information he has either already received or will soon receive from a non-COMINT source. Somehow, competition with the press doesn't seem to be one of the Agency's objectives.

Obvious intelligence as defined above is available to anyone who can read. Once received or reduced to plain language, there is not special talent in deriving intelligence. On the other hand, there is very valuable intelligence to be derived by what is not said or in what is said in a different manner than usual. This type of information I'll call "subtle intelligence." As communications stand-downs are often the first indications of an imminent crisis, abnormal communications also have great potentials as sources of valuable intelligence. The unique and most important aspect of abnormal communications is that usually they provide intelligence not otherwise available - even by reading the messages.

Communications norms can only be established over a period of time as opposed to instant intelligence, which can be derived from reading a message. This is another reason why most analysts prefer and practice the "get rich quick" technique of rapidly searching through haystacks to find the one needle (which may or may not even be there) that will spell out "WAR WILL BE DECLARED TOMORROW." Should such a message ever be found, it would probably still be necessary to turn to a traffic analyst to find out who sent it.

While all abnormal communications may contain intelligence, it may not always be valuable intelligence. Each abnormality must be subjected to an evaluation of why it deviated from the norm. The traffic analyst would be well advised to consult other

professionals prior to reaching any conclusions as to why. At the same time, however, he should be very conscious of the fact that in the vast majority of cases he is better qualified to interpret COMINT phenomena than the consumer. He should take care to insure that his interpretation is restricted to COMINT facts.

Having discussed what norms are and why it is desirable to establish them, let's turn to how we establish norms. In the words of one of our former directors, "First there must be a signal." However, because of the restricted subject of this paper, all future discussion will be made under the assumption that the analyst is working with individual groups of homogeneous traffic.

Some care must be taken to distinguish between establishing norms and recovering technical data such as those used for TEXTA.¹ For example, a group composed of control and six outstations may appear as control and four outstations when examined from a "normal working" point of view. Likewise, certain frequencies may appear on TEXTA that would not appear on a normal frequency study. A large volume of traffic is not necessary to recover TEXTA data, but norms are established more easily, and they are more dependable with large volumes of traffic.

Since it was assumed that the analyst is working with traffic identified to case notation, recovery of calls, frequencies, etc., will already have been made and recorded on TEXTA. The analyst's job is now to determine the normal working. The different types of sorts in which the traffic can be put are numerous, and it would be beneficial to investigate various machine programs to do the bookkeeping.² Short of using a machine program, however, a good basic approach is to put each case in date order. A further breakdown of time within date is helpful and often recommended in training manuals, but past experience has revealed that the benefits gained are usually not as great as the cost in the additional time required. Specific norms are not established with the same degree of ease, and the types of sorts and logs used are dependent upon what norm(s) the analyst is attempting to determine.

Initially, traffic from each case notation should be examined separately. A log listing date/time of intercept, frequency, and actual contact is a good start. Having logged the above information, the analyst should begin analysis of the logs in an attempt to answer questions that would enable him to determine the norm. The specific questions to be answered are limited only by the imagination of the analyst, but discretion should be used to insure that the answers to each question will be worth the effort required to find out. The above-mentioned discretion is a very important ability of an analyst. Intelligence of some sort can be obtained from every bit of information, but its value is a factor of the probability of the occurrence of the conditions under which it will be useful. In other words, don't analyze solely for the pleasure derived in performing analysis.

P.L. 86-36
EO 1.4.(c)
EO 1.4.(d)

1. TEXTA means Technical Extracts of Traffic Analysis. A system of recording and maintaining a database of technical information on communications targets.

2. Traffic Analysis Processing System (TAPS) and TANGERINE.

First, let's look at some of the items that can be determined by looking down the date of intercept column. A simple visual examination will reveal any "missing" dates during the period being examined. The missing dates will often show a pattern. A common pattern is no traffic every seventh day, and a comparison of the pattern and a calendar will probably reveal that the missing days are Sundays (or whatever day of the week the people of the target country consider the Sabbath). Don't be concerned if there are no missing days of intercept; communications of more advanced targets are active seven days a week. Either way you have just established your first norm - the group normally operates seven days a week or the group normally does not work on Sunday. At this point the analyst is cautioned to make sure he understands the working of his intercept source. This common pitfall can best be illustrated by the analyst who had established the norm that the group he studies operated only five and a half days a week closing down at noon on Saturday until Monday.

REF ID: A66386
EEO 1144 (C)

Examination of the time of intercept column of the log will show the normal hours of operation and schedules. Once again as mentioned above, care should be taken not to recover the intercept source's hours of operation. Often the difference between the time the group begins operation and the time it stops is the same as the hours of a regular day's work or at most maybe sixteen hours a day (two eight-hour shifts).

Schedules are not as easy to establish as hours of operation. The main reason for this is that rarely does the analyst have the benefit of cast-iron coverage. A good intercept operator often makes up for sparse cover. He notes when he hears his target, and the next day he will usually be looking for that case at the same time. This results in the same approximate time being recorded for several days in a row, and it also establishes another norm. The other side of the coin is also true, however. Certain elements in collection management in our agency have caused a great deal of emphasis to be placed on the "productivity" of our intercept positions. Often this results in undesirable pressure being brought to bear at the source of intercept. In their efforts to make their positions productive, analysts spend less time searching for new schedules and placing more and more reliance on the few already recovered.

Another hindrance in establishing schedule norms is that operators at [] sites will often copy only "messages." Several receivers are used simultaneously and known schedules are checked out, but unless a message is being passed there is usually nothing to indicate to the analyst that the group was active at such and such a time. The disadvantage here is twofold. First, while the intercept operator may know when to look for a certain group, the odds are high that the analyst is not aware of the schedule. Second, after the norm has been established, a missed schedule by the target (which should have high intelligence value) would appear to an analyst exactly the same as the group coming

up on schedule but not passing a message (probably of no great intelligence value) and therefore heard but not copied by the intercept operator.

Schedules are much more difficult to an analyst than they are to the intercept operator. Often a good tech exchange letter to the source expressing the analyst's views and asking specific questions will help establish schedule norms very quickly.

Analysis of the frequency column of the log will increase the list of norms, which should be getting fairly large by now. The frequency normally employed can be easily determined by visual observation, but to establish norms beyond this, it is usually necessary to analyze the frequency column in conjunction with the date, time of intercept, and/or call-up columns. The same is true of the call-up column. In isolation, analysis will show only those stations that are normally contacted, but these norms can be refined greatly by further analysis between two or more columns of the log.

With relatively little effort so far, the analyst has managed to build up quite a set of norms. The log can now be expanded to include message headings and a column for type of traffic. The columns to be used for the message headings will vary, but almost all will include message number, precedence, date/time of origin, addresses, originator, and group count.

Without going into the detail given in the first part of the log, it should be noted that certain norms can be established by analysis of each column in isolation, but the possibilities are almost unlimited when analyzing each column with one or more other columns.

As was stated earlier in this paper, different norms are established with varying degrees of difficulty. There will be absolutely no doubt about the validity of certain norms. Others may be questionable, but the exercise of trying to establish norms will allow the analyst to identify those areas that need more intercept and analysis. This should result in better tech exchange and support and thereby make better use of collection resources.

Once the norms have been established, the analyst can quickly scan the daily intercept (without logging) and with a high degree of confidence separate the small percentage of abnormal traffic. The remainder of his time can be sent interpreting the abnormal traffic and thereby coming up with valuable unique intelligence.

P.L. 86-36

[] in a lecture to our senior intelligence class, once said if he had three wishes they would be (1) to know everything in the world, (2) to be able to report it to parties capable of doing something about it, and (3) to have them believe it. Establishing good communications norms will enable a traffic analyst to continue to provide the intelligence community with unique intelligence when the linguist and cryptanalyst are truly stumped. At his point, the problems are just beginning as we are faced with the problem of "having them believe it," but, oh, what a wonderful problem with which to be faced.

Chatter Patterns: A Last Resort

WAYNE E. STOFFEL

A possible method of identifying radio operators by their reaction to standard situations occurring in chatter, for use when conventional techniques fail.

BACKGROUND

The success or failure of most traffic analysis problems depends primarily upon the analyst's ability to achieve continuity.¹ Simply defined, continuity involves bridging a communications change by equating a given element appearing before the change with a different element appearing after it. The term continuity refers to the discovered relationship between the given element and its replacement, *without reference to the underlying meaning*. For example, we may by various methods discover that callsign ABC during November was replaced by DEF during December, and thus achieve continuity from ABC (November) to DEF (December). Note that the time factor is intimately involved in the relationship between ABC and GHI is more accurately termed an equation or co-location and is not a continuity in its pure sense. Continuity can exist between ABC and DEF without any knowledge of the location, identity or function of ABC or DEF. The importance of the distinction between continuity and other forms of equation lies in the fact that once any knowledge is gained about ABC, it automatically applied to DEF (and vice versa). If we discover that ABC served the Chief of Staff, 12th Division, Greenville, for his contacts with subordinate regiments on the Division administrative/logistic net, this information applied to DEF, in toto. On the other hand, about GHI we can only say (with any certainty) that it is located at Greenville. (Depending upon the type of equation made between ABC and GHI, we may further be able to say that GHI also serves 12th Division or that it also serves an administrative/logistic function).

A direct cryptanalytic analogy to continuity can be recognized by considering a simple substitution system involving a matrix with changing coordinates. For example, the following matrix has been recovered for 1 April:

1. A number of countries today go to surprising lengths to suppress in their communications system distinctive characteristics that might serve to disclose their identity. Among the more common methods of suppressing characteristics is that of frequently changing certain communication elements, such as callsigns, frequencies, schedules, procedure, routing, and address symbols. Since it is often necessary for the traffic analyst to study several months of material on a given net before concrete intelligence results can be developed, and since communications elements may change as often as twice each day, he must, somehow, find a way to nullify the effect of these frequent changes in order to pull homogeneous material together for study. He may note certain characteristics that do *not* change frequently (for instance, that a given station sends a distinctive service message each day at 1100), which can serve as identifying features. When he is successful in nullifying a communications change, the traffic analyst refers to the result as *continuity*.

UNCLASSIFIED

	4	2	1	8	•	•	•
2	A	D	R	L	•	•	•
9	P	B	O	C			
3	M	-	E	T			
1	-	K	N	-			
•	•						
•	•						

On 2 April, assumption of the probable word "ATTACK" yields.

47	23	23	47	63	55
A	T	T	A	C	K

	7	5	-	3	-	-	-
4	A	D	R	L	•	•	•
6	P	B	O	C			
2	M	-	E	T			
5	-	K	N	-			
•	•						
•	•						

It can then be stated, if the assumed word "ATTACK" proves correct, that row coordinate 4 on 2 April is *continuity of* row coordinate 2 on 1 April. It can also be shown that cipher value 57 on 2 April is *continuity of* cipher value 14 on 1 April. In this second case, we have achieved continuity without knowing what the actual plain value is. Finally, we can say that cipher value 43 on 2 April is continuity of cipher value 28 on 1 April. In this instance, when $43c(2\text{April}) = 28c(1\text{April})$ is proved, and $28c(1\text{April}) = Lp$, then $43c(2\text{April}) = Lp$.

The more frequently an element changes, the more important continuity becomes (since it is virtually the only consistent method for achieving enough depth on a given element so that a study of its underlying nature and purpose can be undertaken), and the harder it is to get. Most of us can sympathize with the unfortunate analyst whose formerly stable problem suddenly adopts twice-daily changing callsigns, frequencies, addresses and discriminants.²

On problems involving fast-changing elements, continuity is usually achieved by means of whatever characteristics are available that can be trusted to be unique. If many are available, the easiest, fastest or most economical methods are, of course, tried first,

2. Traffic analysts will recognize that, for the sake of simplicity, the complexities of the various classes of equations and their accompanying validations have been avoided in this presentation. Other readers are warned that many validities have been avoided in this presentation. Other readers are warned that many a "Donnybrook" can and does develop between traffic analysts on these very factors.

UNCLASSIFIED

while the more intricate and time-consuming methods are held in reserve for tough cases. It often happens that certain nets develop a stubborn streak that defies description (in mixed company) and, despite application of the most time-consuming routines, manage to remain intact and featureless.³ Where all else has failed, the analyst may well find the following proposed routine useful.

INTRODUCTION

Most people are creatures of habit, particularly when performing a routine task, and radio operators are no exception. There have been considerable experimentation with and study of the variable characteristics of a Morse operator's transmitting habits or "fist" in an effort to develop a systematic process of recording and analysis which would permit ready recognition of the individual at the key. There is, however, a large area of variable operator habit which has remained virtually unexplored during recent years: habitual operator characteristics as displayed in routine chatter exchanges.

A good many traffic analysts can recall a specific instance where a unique or rare procedure signal was consistently used by a certain net or station and, in the last resort, could thus be relied upon to identify its user. There may be few, however, who can recall conducting a comprehensive and systematic search for such characteristics in order to achieve continuity and identification.

What follows is an outline proposal for a routine of systematic search for unique chatter or conversation characteristics that can be used for continuity or co-location purposes. For the most part, specific details are avoided except for examples, since they will vary from problem to problem. It will be seen that the routine is not readily usable on large problems, and may, in fact, be suitable only on limited problems where the area of inquiry is relatively small and all standard methods of achieving continuity have failed. An obvious prerequisite would be a significant volume of activity transmitted by the stations under study, with some assurance that a fairly complete (preferably verbatim) copy of chatter has been recorded by the intercept operator.

BASIC ASSUMPTIONS

It can be empirically demonstrated that regardless of the degree of conformity enforced by the target's COMSEC service, different operators use different operators use different combinations of procedure signals to express the same ideas, but that each operator tends to be consistent with himself.

3. This situation tends to exist to a greater or lesser degree on most problems, although it can be appreciated that the point is ordinarily glossed over in discussion unless the words "additional personnel" are injected into the conversation at a suitable point.

The writer's contention is that these habits are more widespread than is generally supposed and that, under admittedly special circumstances, a systematic routine will disclose a sufficient number of them to permit continuity to be developed.

Expert chatter readers will recognize that operator chatter must be treated as a distinct, albeit peculiar, language.⁴ Despite the best intentions of the signal officer who compiles an extensive set of procedure signals for radio operations, the "plain" side of his "code" is generally restrictive. In actual operational use, a given procedure signal (prosign) tends to lose its rigidity and takes on a more general *concept* or *idea* form (particularly where it is used so often as to be easily recognized without "looking it up"). Thus the prosign QTR can be shown to have the fixed meaning ("The correct time is - hours"), whereas in actual usage among experienced operators, it embodies the general concept of *time* and is so used in a wide variety of contexts.⁵ Complementing the tendency of experienced operators to generalize prosign meanings is the equally strong tendency to minimize and abbreviate words and prosigns in order to conserve both time and energy. "Ham" chatter displays this quite clearly.⁶ It is not difficult to visualize how a relatively isolated segment of a radio network could gradually evolve a "local dialect" distinct from that of the rest as a result of improvisation under these pressures. Certainly a regimented COMSEC system with a firm domination over the radio schools could suppress some of this variation, but if we confine ourselves to studying experienced operators, it is likely that some recognizable variance and individuality will occur.

A SAMPLE PROBLEM

If distinctive operator habits do, in fact, exist, how do we go about finding and recording them? Evidently, if a way can be found to catalog the *situations* that confront a radio operator most frequently, we can collect his *responses* to any given recurring situation and by observation determine whether his reaction is fixed by habit or is variable. For example, we might select as a favorable starting point several hours of intercept between station A and station B during which a number of messages were sent by each station. As a recurring situation, we might select *message transmission* and further restrict our examination to the station responses during the period immediately before starting each message. We might find

4. A more precise analogy has been suggested that compares chatter to a codebook usage wherein (a) the vocabulary is not precisely suited to the material being encoded and (b) the code is large enough so that code clerks tend to use combinations of common, memorized groups in preference to rarer but more precise and economical groups that must be looked up each time they are needed.

5. For example, the interrogative for "QTR?" is listed as "What is the correct time?" The prosign QSY means "I shall send on ____ kilo[hertz] and its interrogative form (QSY?) is interpreted as "On what frequency shall I send?" or "Should I change frequency?" The compound "QTR QSY?" may well be used to mean "When should I change frequency?"

6. For example, the prosign "CUL" is a "Ham" contraction of "See or contact you later."

Example 1

A: QTC (I have traffic for you.)
 B: GA (Go ahead.)
 A: C AS (Yes, stand by.)
 B: C (Yes.)
 A: BT (Break Sign – attention, etc.)
 A: NR (Goes into preamble.)

Examination of the same basic situation a short time later when station A was again about to transmit a message showed that after receiving "GA," station A again said "C AS" (Yes, stand by) and after receiving the affirmative from the other end began his transmission with a break sign. A third message still later in the same schedule begins with the same exchange, and it now begins to look as if we have found a starting point.

A quick look at the activity of station B shows that the two messages it sent were also preceded by identical chatter exchanges.

Example 2

B: QTC AAA (I have an "AAA"⁷ message for you)
 A: AS (Stand by)
 B: C (YES)
 A: GA (Go ahead)
 B: C AS (Yes. Stand by)
 A: C AS (Yes. I'll stand by)
 B: C (pause) BT (Yes. Break sign – attention)
 (Then into preamble)

Let us now examine what we have so far in the way of possible habits:

- (a) When offering a "QTC," both station A (Example 1) and station B (Example 2) sent "C AS" after receiving "GA" from the other end. Each then preceded the preamble with "BT," but station B (Example 2) used the compound "C (pause) BT."
- (b) When receiving a "QTC," station A (Example 2) responded with "AS" before giving the "GA," while station B (Example 1) gave "GA" immediately. When responding to "C AS," station B (example 1) gave the brief answer "C," while station A (Example 2) used what may be a variant form – "C AS."

Later the same day, another exchange of messages is found between stations A and B. During this later schedule, two messages from station A are preceded by

7. "AAA" in this instance refers to type or priority of message (e.g., "2nd priority" or "service").

UNCLASSIFIED

Example 3

A: QTC
B: GA
A: C (pause) VVV QTC
(goes into preamble).

and one message from station B is preceded by

Example 4

B: QTC
A: GA
B: C AS
A: C
B: C (pause) BT
(goes into preamble)

It is quickly seen that the behavior of station B is essentially unchanged but that of station A shows no parallel with what went before. Our choice at this point is quite simple – either station A has changed operators or the “habit” is not sufficiently strong. The resourceful analyst would study carefully the chatter exchange during the opening of this second schedule for any evidence of a new operator at station A (extensive tuning, authentication, etc). If the “new operator” hypothesis does not appear sound, other types of habit must be sought. On the other hand, if it *does* appear sound, examination of suspected continuities from previous or successive dates should show whether the time of change is fixed (i.e., the end of one duty tour and the beginning of another). It would appear that once the duration and change times of operator shifts can be established, analysis can proceed at a much faster rate, since the change times will allow the analyst to sort activity for any given date into tentatively homogeneous groups.⁸

Thus far, our accumulated results are far from impressive. Where can we look for other habits? Two situations obviously related to the one examined above would be the area immediately following the message (message closure and receipting exchange) and any “in-text” servicing (receiving station interrupting to ask for repeats while the message is still being transmitted) or “post-text” servicing (after the message is finished but before receipt is acknowledged), but there must surely be other areas that could be equally profitable.

TYPICAL SITUATIONS

We may find it useful to consider a typical schedule between two stations and examine the successive situations that confront the radio operator. Since certain of these will tend to recur within the same schedule (e.g., opening traffic, as in the example above), while

8. Some care must still be exercised in watching for cases where extra operators are put to cope with heavy traffic volumes, or for any other situation having the same effect.

others by their very nature will tend to occur only once in any given schedule, it is convenient to distinguish between the two *types*, since the former is much more useful as a starting point (one is bothered less by possible operator changes, and only one schedule is generally needed for initial isolation of a tentative habit) while the latter comes into use, for the most part after some initial foothold has been achieved. For purposes of convenience, we shall call the former *primary* and the latter *secondary habits*.

2. *Tuning*

Immediately after initial contact, various adjustments of tone, power, and frequency must usually be made before reception is considered good enough for the transaction of business. The exchanges may range from a short, terse and businesslike operation to a long, temperamental and often humorous argument. Unless they occur frequently, these longer-winded battles are of little use to the type of study being described,⁹ and attention should be concentrated upon the shorter and more lucrative exchanges.¹⁰ The first schedule after a frequency change usually contains much more tuning chatter than do subsequent schedules on that same frequency.

3. *Recognition*

Recognition exchanges may occur with or without a specific system such as an authentication chart or table of challenges and responses. They are most often seen on the first schedule after a new operator comes on duty, although some signal plans seem not to require their use unless messages are to be exchanged, while others obviously specify such use on every schedule. Many experienced operators prefer to rely on aural recognition of "fist" characteristics and frequently ask the other end to "send V's" (QSV) or adopt some other device toward the same end.¹¹

4. *Opening Traffic*

The exchanges treated in some detail (see examples 1-4) may be preceded by statements from *both* operators that they have traffic to be transmitted. In this situation, agreement must be reached on an order of transmission, and such an exchange may be a good source of secondary habits.

5. *Preamble and Text Handling*

This category embraces a wide variety of characteristics, some of which are generally recognized as useful. In order to find *operator* habits, one must recognize that the operator

9. Except, of course, for the laudable purpose of recreation.

10. In analyzing these exchanges, it is useful to remember that frequently the operator does not have direct access to the transmitter itself and must relay adjusting instructions to a remote transmitter site by telephone.

11. This use of QSV should not be confused with the more extensive use during tuning or equipment adjustments. When the sending occurs early in the schedule, it is not always easy to distinguish between the two, but its use in the recognition sense is usually unmistakable when, during later operations, consistent mis-encipherment of procedure, etc., arouses clearly recognizable operator suspicions about an operator's identity.

is here working from a printed or written record, so that the order of preamble elements, for example, is controlled (in most problems) by their arrangement on the message form while breaks and separators may generally be attributed to the operator himself. Here also belongs the situation where the operator realizes he has mis-sent a portion of the text, sends an error sign, and corrects the mistake. In this category, one is most definitely at the mercy of the intercept operator, and one is likely to find him completely absorbed in copying the text (to the exclusion of nontextual transmissions).

6. *"Break-in" Servicing*

The receiving operator, under signal instructions, is allowed to "break-in" during text transmission to ask for verification or repeats of certain passages which he has missed or which seem doubtful. Where this happens (and where the intercept operator provides a verbatim record of the exchanges), primary characteristics may be found, since a number of prosigns are usually available for use in this situation, and requests for repeats can and do take several forms.

7. *Closing Traffic*

Most signal instructions will provide for some prosign such as **BT**, **BK** or **K** to mark the end of text, but some operators use additional compounds for emphasis, or to remind the other end that there are still more messages to be transmitted. As special case, traffic sent by broadcast methods is usually sent twice, and the procedure used to separate the two consecutive transmission frequently shows strong habit patterns.

8. *Post-Message Servicing and Receipting*

From the transmitting operator's point of view, a given message has not been "cleared" until the other end officially receipts for it. If the other end is not satisfied that his "copy" is correct, he will not give a receipt (**QSL**) until he has verified the questionable passages. Although the situation is slightly different from that described in "Break-in Servicing," habits found in one situation would be likely to show up in the other. As a special case, servicing may be asked for during a later schedule and, if it can be shown that the message has already been "cleared" (i.e., that a **QSL** was given), this "late" servicing may well result from an inability to decrypt the message.¹² The servicing request in this instance may differ from "break-in" or "post-message" servicing only to the extent that the involved message must be clearly identified (i.e., by serial number or other unique indicator).

9. *Breaks, Waits, and Interruptions*

We are here concerned, not with pauses that appear to be a fixed part of habits rising out of other situations (i.e., the pause before message transmission as shown in the first examples above), but rather with the nonroutine or unexpected interruptions that cause temporary or permanent breaks in a given schedule. Among the situations that can be

¹². Such information might be particularly useful to the cryptanalyst.

expected to produce habitual responses¹³ are intervention of other schedules, equipment failures, interference, operator changes, shortage of transmitters and interruptions by other stations.

On especially long waits, the transmitting operator may key certain characters or compounds to "hold" the other end, in the general sense of "Hang on, I'm still here" or "Keep listening, I'll only be another minute or so." The actual signals sent during this "hold keying" may well be unique to each operator, but again we are dependent upon verbatim intercept copy if this characteristic is to be used.

10. *Next-Appearance Discussions*

Once the business of a given schedule has been transacted and the schedule is about to be terminated, some mention is usually made of the next appearance. Where contact times and frequencies are predetermined by the signal instructions, this mention is not likely to exceed a very perfunctory "Watch for me; I'll watch for you." On the other hand, the discussions may well involve times and frequencies. Either situation will yield useful secondary habits. As a special case, satisfactory contact may not have been achieved, and ensuing discussions about another time and frequency may yield significant habits if the situation recurs.

11. *Sign-Off*

The actual termination of a schedule frequently involves a little ritual that is difficult to describe to one who has never heard it. Between operators who are used to working with each other, it is usually fairly rapid and highly stylized.¹⁴ While this area should not be ignored as a source of habits, a departure from the routine specified in the signal instructions is frequently the result of tacit agreement between *both* operators and must be treated accordingly.

12. *Special Circumstances*

The above categories obviously do not complete the list of situations which may be useful on any given problem. If the net under examination regularly changes frequencies in midschedule, the chatter exchanges before and after each change merit some observation. Another special case involves the use of a matrix or table for prosign encipherment. Aside from the obvious benefits such a system can provide where local usage makes it effective for net or complex identification, the *use* of each *cell* in the matrix can be likened to the use of a comparable *prosign*. Thus, habitual use of certain cells or the formation of various compounds is just as useful as the prosigns themselves. This principle

13. Obviously, interruptions caused by flood, fire and other emergencies cannot be expected to appear often enough to be a fruitful source of habitual responses.

14. A typical exchange sometimes used by U.S. personnel, where conformity to COMSEC regulations is not rigidly enforced, involves the transmission EF (dit, di-di-dah-dit) and the answer EE (dit, dit), which approximates the rhythm of the familiar "Shave and a haircut..."

UNCLASSIFIED

also applies to related systems, such as authentication, wherever habits can form as a result of allowing the operator a free choice in selection among a number of variables.

CONCLUSIONS

It will be evident that the proposed approach to maintaining continuity through chatter analysis has application only in limited cases. Because of its complexity, it may well be attempted only as a last resort and would undoubtedly require the services of a skilled chatter reader.

P.L. 86-36

On some problems, one or two distinctive habits may be sufficient, while on others a wide variety of situations may need to be examined before individual operators can be distinguished. It may be found useful, when looking for habits, to keep a similar running record of those responses *which are the same for all operators*, on the theory that such responses have been specified by the signal instructions or form a "local dialect." Such a list would be helpful in later examinations of a related net or complex, since it would define situations where habits are *not* likely to be found. (It might also become a useful *net identification* tool.)

It should be emphasized that the "habits" we seek in this approach are not tendencies to act in a given manner, but are more nearly instinctive reactions or reflexes to recurrent stimuli. Where these reactions are

found to be quite variable, it may be assumed that the operator concerned lacks sufficient experience to have developed such habits or that the situation is rare enough so that he has not developed a reflexive response.

The approach may be useful, not only for continuity development in selected areas, but for inter-net equations after other evidence has narrowed the area of search to reasonable proportions, and to bridge communications changes where continuity is available both before and after, but not across, the change.

UNCLASSIFIED

A Note about NRs

P.L. 86-36



A number range is an NR range, or so I have always thought. It is, of course, useful in recovering continuities and colocations, and the rate at which its member numbers are used up gives a relative measure of the tempo of activity of the user of the NR range. Recently, however, several interesting anomalies in NR usage on the Vietnamese Communist (VC) problem have triggered some thoughts about the behavior of NR ranges.

I recall sermonizing about my conviction that fewer errors occur in the NR (you may read "Radio Station Serial Number" for NR throughout this article) than anywhere else in the message. Errors, in this context, are those committed by the *target* and not by our ability to perceive what was sent (i.e., not intercept, or logging, or forwarding, or machine errors on *our* part). The argument underlying this belief centers on the use of the NR as a control or reference number: one refers to a message by means of that NR. Any ambiguity or conflict would cause difficulties in even the most primitive filing system. For that reason, one can imagine the keeping of a log book; each sheet of the log could have, down the left side of the page, the string of consecutive numbers that will become NRs as each message is logged onto the sheet. A neat, pretty picture – but not always correct.

In the course of looking at one particular VC NR range, it became apparent that, from time to time, *duplicate* NRs on consecutive messages did occur. At first, each instance was attributed to an intercept or recording error on our part, but in some instances the messages were re-sent on later dates with the same offending duplicate NRs. There is at least one explanation for such duplicates. The VC evidently do not like to send long messages and usually break up long texts into short messages of roughly 75–100 groups. Each part then gets its own NR and group count, and one of the parts will carry the file date and time for the whole batch. We might see five consecutive messages, only one of which carries a file date and time; it is as if all five messages were block-filed at one time. It seemed significant that none of the duplicates were *within* one of these batches of messages. Where duplicates occurred, they were invariably between the first message of a batch and the message preceding it. Often the earlier member of the duplicate pair was a single message filed late in the evening (later than the "typical" file times for that station) and carrying a high precedence. From time to time, we have captured radio station documents, and one frequently encountered is a small log book that is a rough diary of schedules met, significant (to the VC operator) chatter, and message preambles sent or received. Where duplicate NRs have occurred, it seems possible that the first message, carrying high precedence and needing to be sent out "after normal hours," simply was not logged in properly (such things have been known to happen around here, too) or that no log, as such, was kept at all by the station – the activity log book contained all preambles and could easily be checked to see what the next unused number should be. The pages of

these books are quite small, and the last previous message would often be logged on some earlier page. Instead of logging in, say, NR 76 on the line immediately below the entry for NR 75, in a formal NR log, one might simply leaf back through the activity log book for the last outgoing message. It would be easy to miss the entry of a single message, if most message entries contained strings of from five to fifteen messages, or if the message was not properly logged in. Thus, in the example, the leafing back for the last-used number, NR 75 might be missed, either because it was only a single entry and easily overlooked or because it was mishandled after hours and perhaps was not even logged. In any event, it seems clear that a *separate* NR log was not being maintained at that station.

Another aspect related to how the NR log was kept was evident in the circular NRs of one station: he used the block 100-399 to one set of outstations and 400-699 to another set. Each time one or the other ranges reverted, it was evident that the station was undecided about whether the ranges began at 100 and 400 or at 101 and 401. Over a long period, both starting points were used about equally often. Clearly, no preprinted logs could be the basis for such indecision. It could be accounted for either by the small operator log mentioned above or by a *blank* NR log (i.e., not preprinted with numbers).

A third "happening" seems, at first glance, to clash with the behavior described above. On 1 January 1969, the station that had been using the 100-399 and 400-699 circular series changed a number of aspects of its operation, including its NR patterns. In particular, both of these circular ranges were replaced by 1-100 ranges. In one circular series (but not the other), the number 44 was omitted. The omission is easily proved, since other serials occur in crypt indicators that do not omit 44. On the face of it, this could be explained by assuming that the station used a preprinted log containing an error in it. And that may indeed be the real explanation. But there is another possibility. Many of the VC documents captured and turned in to us for study bear the unmistakable signs of repeated usage. VC units in the field are short of paper and have learned, by writing lightly with pencil and later erasing, to stretch the life of each available scrap of paper. (Whether this fundamental difference between VC and U.S. methods has had any impact upon the course of the war will be left to historians and philosophers to debate.) It seems at least possible that some sort of log was kept during the *first* "pass" from 1 to 100, during which 44 was omitted in error. When the series reverted to 1, out came the trusty eraser and all but the NR was erased on each successive line as needed.

An NR range is an NR range, and its principal uses are still related to continuity, colocation, and tempo of activity. But subtle variations in *how* the range is used may, now and then, give us a little insight into the activities of the humans who use the NR range. All this presumes, of course, that the target country has not yet extended its COMSEC activities into the area of NRs.

Some Thoughts Concerning Traffic Analysis Mechanization

ROBERT S. BENJAMIN

Paradoxically, traffic analysis mechanization *will save manpower* but should not decrease and may even increase the amount of "real" traffic analysis work that remains to be done. NSA and the cryptologic community have always had far more things to do on most analytic problems that could be done with available human resources. Many of these things that have to be done prior to important analysis have, in the past, involved hours of logging, summarizing, counting, verification, some degarbling, and other similar operations. These are the things that machines can do best – once material is in a *machinable medium* that permits its being fed to a computer for automatic data handling. On the other hand, only people are capable of assessing the *meaning* of the various phenomena discovered in the process of organizing and reducing the data.

The scope of traffic analysis ranges between the first recorded results at an intercept station and the issuance of SIGINT end-product reports that detail traffic analytic findings in a form suited for intelligence consumers. Practically anything that goes on in the long series of complex processes between antenna and product, wherein "data" (signals) are gradually being transformed into increasingly meaningful information, is of concern to the traffic analyst. The traffic analyst works with signals collection people in directing intercept cover; he stands ready to assist the cryptanalyst in the latter's efforts to cryptanalyze message texts; he writes end-product reports for intelligence consumers or contributes basic data for integration into such reports. He works similarly with a multitude of other experts in various other SIGINT and COMSEC specialties.

The nature of traffic analysis is going to change; the change will come slowly at first, but we will be shifting gears in the next few years, and we must be ready for it. Material will be coming to us in predigested form – the more clerical parts of our job will in the near future be done to an increasing extent by machines. And we must develop and exchange ideas concerning the new working methods now appearing. A deeper order of creative thinking will be called for. Once mechanization systems are begun and databases are established, the use to which the data are put will be limited only by the imagination of those using the particular system. The *manual* part of analysis will be done more and more by machine; the *mental* part must be done by people.

~~CONFIDENTIAL~~

SHOULD ANALYSTS LOOK AT HARD COPY?

People faced with machine output in one of its many forms often wonder whether they should continue to examine hard copy intercept. The answer to this question depends on a number of considerations – two, principally:

- a. Hard copy should be examined until the analyst convinces himself of the accuracy and utility of the machine product.
- b. Hard copy should be examined *if and only if* there is a real advantage to be gained.

A complete version of the intercept will probably have to be available for "back-up" use in many instances – just as the New York telephone directory must be available even though one doesn't read it page-by-page. This "back-up" traffic may be in hard copy, microfilm, or in a machine recorded form, and should be retained for a reasonable period. It should be available for emergency use to permit analytical reentry into traffic external systems (callsigns, etc.) when the machine "digests" are insufficient and to permit occasional spot checks of the efficiency and reliability of the system.

One thing is sure—extensive recopying and counting from hard copy should *not* be done. Summaries will be prepared by machine, as will detailed files concerning data elements and data items under consideration.

The traffic analyst is no longer limited to the traditional "sort and list" output of traffic analytic data. Machine output now may be in many other forms. To cite a few examples: summary listings (prints), printed matrices (coordinate tables of various kinds), printed "net reconstructions," decks of punched cards with "interpreted" (printed) information on them, answers received at a remote inquiry station (electric typewriter) communicating with a data bank in a remote computer, and "soft copy" displayed on a scope resembling a television screen (which may be photographed, if need be). Every few months, new forms of useful output seem to be appearing, resulting from computer industry developments, and from cooperation within NSA between so-called "machine people," R/D people, and analysts.

LARGE PROBLEMS VS. SMALL PROBLEMS

All those techniques of traffic analysis that depend upon collecting, summarizing and analyzing masses of data for achieving answers will lend themselves to mechanized treatment. But creativeness and judgment will be necessary

- a. to design the logic for the mechanization in the first place,
- b. to interpret the results obtained from machine output,
- c. to improve and extend the capabilities of the particular system by discovering new ways to use it and new ways of producing even better information, and

~~CONFIDENTIAL~~

~~HANDLE VIA COMINT CHANNELS ONLY~~

- c. to know when to go to hard copy intercepted traffic for answers that the machine system cannot give. Hopefully, proper interpretation of machine output will give a good part of these answers.

The one-of-a-kind type of occurrence is trickier and does not lend itself to machine analysis so readily. Such an occurrence may be recognized by a human being at an earlier point (in the field, for instance) and appropriately indicated, tagged, or reported in some suitable reporting medium. Or a scan of the hard copy material by an analyst *knowing what to look for* and taking fullest advantages of what has already been done by machine, may show up one-of-a-kind occurrences that may be significant. Whether a "hand scan" of hard copy will be profitable will depend upon the target, the requirements, and the circumstances.

The very, very small analytic problems are often cited as examples of things that need not be mechanized. This remains to be seen. Even in these small problems, however, certain data are necessarily a part of overall SIGINT requirements for coverage accounting reasons and must be mechanized. To the extent that an analyst can profit from what is already being mechanized, so much the better.

PENDING DEVELOPMENTS

There is a need to apply to traffic analysis methods like those called "management by exception" techniques now used in mechanized data handling in the business world. However, before such "analysis by exception" techniques can be used, norms must be established so that the computer can recognize occurrences or phenomena outside these norms. Admittedly, the establishment of norms is not an easy task. Machine scanning and selection methods involving the search for material containing certain desired "words" or the elimination of material containing unwanted "words" are a class of "exception" techniques.

Many people are working on developing many projects that relate to traffic analysis mechanization. Some projects that come to mind are

STRUM (Standard Technical Report Using Modules)

TAPS (Traffic Analysis Processing System - a flexible family of programs)

TIPS (Technical Information Processing System)

SPECOL ("Special Computer Oriented Language" - an experiment in data correlation and retrieval)

TEXTA Mechanization

ZITO (produces soft copy of a scope)

AG-22 (Paper-tape producing typewriter).

TRAPEZE (System to produce formatted reports from AG-22 paper-tape output).

~~CONFIDENTIAL~~

Each of the above projects is significant enough to deserve being treated at great length, but that is beyond the scope of this article. And there are many other important projects that could be mentioned.

We urge individual analysts to learn more about machine processing. The National Cryptologic School offers several excellent courses to introduce the "layman" to the general idea of using machines, and other such courses will be devised in the months and years ahead.

Learning a programming "language" is another way to gain familiarity with machines. These so-called "higher programming languages" are machine-independent in that *compilers* are available that enable programs that have been written using these languages to be run on different computers, even computers made by different manufacturers. FORTRAN is one such language, but FORTRAN lends itself more readily to mathematical computation than to the handling of traffic analytic data. Showing more promise for data handling is COBOL, a language designed for business applications, which can handle alphabetical as well as numerical information. Courses in either of these languages can provide the traffic analyst with a good conceptual base. There are other "programming languages" in existence, or being developed, that are suitable for information handling.

As TIPS Pilot is developed, simple formatted interrogative statements through a RYE remote station will enable any analyst or manager to ask a wide range of questions and get rapid answers from the RYE computer concerning information in the TIPS files. Subsequent versions of TIPS - TIPS I and TIPS II (TIPTOES) - will add many refinements in future years.

However, analysts need not learn computer programming to use mechanization, but they should become familiar enough with machines to feel comfortable when confronted with the idea of mechanized data. Machine prints will not look like conventional "hard copy," but they will spare the analyst many, many hours of drudgery and permit him to spend his time really analyzing and synthesizing. Whole new areas of techniques will open up; for example, machine prints of TAPS output have shown phenomena not at all readily apparent in the hundreds of input entries that were combined, reduced, and presented to analysts in summary form. Interpreting these phenomena calls for a high order of thinking and judgment.

In summary, analysts on analytic problems that aren't already using machines should be exploring ways to use them, particularly since STRUM, AG-22 and the like will soon present a machinable input for practically all problems. And, of course, machine people need to know more about traffic analytic problems; at best, it isn't easy to dredge meaning out of the mass of confusing data facing the traffic analyst, and the more that machine people understand this, the more help they will be to this important aspect of SIGINT production.

~~CONFIDENTIAL~~

~~HANDLE VIA COMINT CHANNELS ONLY~~

~~CONFIDENTIAL~~

P.L. 86-36

Acknowledgments

I should like to express my appreciation for the review and helpful comments by Mr. James Taylor, C4103; Mr. Robert Grove, C416; Mr. George Hicken, C03; [REDACTED] [REDACTED] A03; and Dr. Walter W. Jacobs, E.

~~CONFIDENTIAL~~

~~HANDLE VIA COMINT CHANNELS ONLY~~

Cleaning the Augean Stables or How Much T/A Can a Computer Do?

FRED MASON

One of the labors of Hercules was to clean the stables of King Augeas. These housed three thousand very healthy oxen and hadn't been cleaned in twenty years. Rather than trying a shovel, Hercules was smart enough to turn two rivers through the stables.

T/A has a comparable task. Manually, T/A is a desperate attempt to clean up the constantly augmented flood of intercept, in the belief that there may be one or more valuable needles in the original haystack. And the collectors are so damned healthy!

The computer seems to offer a way of doing it by dropping the input into electronic rivers. These, if properly directed, can multiply the human effort and wash, massage, organize, compare with the past, and separate the input—putting the debris here and a small pile of shining needles there.

How much of this can the computer, in fact, do? Part I attempts to define T/A, and Part II attempts to answer the computer question.

PART I. TRAFFIC ANALYSIS

1. Glittering Generalities

The purpose of SIGINT is to make available to supported commanders a reconstruction of the target Order of Battle (O/B) and changes thereto as derived from target emissions — and in time to permit educated command decisions to be made.

Order of Battle answers the command question about the opponent, "What forces oppose me? What are their titles, types, echelons, command relationships, equipment, morale, strength, locations, deployments?" It is a static description of the target at a moment in time.

Changes in O/B update the answer by establishing a vector; he was here then, now he's there; his O/B change is describable and sometimes measurable in amount and direction. Such intelligence vectors assist the commander to understand target changes in capabilities and to estimate hostile intentions.

2. *Serendipity*

SIGINT attempts to accomplish its purpose in several ways. One school of SIGINT thought has it that in combat all is flux; that little can be equated from day to day; and that the valuable SIGINT answers will be derived from exploitation of readable traffic (cryptanalysis) and the unusual, the compromise, the momentarily understandable in an ever-changing environment (traffic analysis).

This approach is Serendipity – the happy faculty for finding valuable things unexpectedly. Cryptanalysis has put a massive effort into reading as much encrypted text as possible, using substantial computer support to break into and exploit traffic. (The best way to find the most four-leaf clovers is to examine the entire lawn.)

Traffic analysis is governed by fewer rules and varies more in its application. Serendipity is also systematized by using a small number of full-time omnivorous readers who see all intercepted traffic, all reports on it, and all friendly correspondence relating to the situation. Ideally, their memories are photographic and their knowledge of O/B encyclopedic. They ignore all that does not excite their curiosity but seize anything with known indicators or with interesting unknown features. They then assemble all other traffic from that net and from any other net with the same SOI or displaying the same indicators or features. From this section of input is deduced the activity, often only after extensive study. But the reading of new material must continue, so that higher priority compromises, or additional data on the one being exploited, won't be missed. And collection must be redirected in order to pick up more of the activity.

SIGINT tries to take "unexpectedly" out of the Serendipity definition by collecting as much as possible and by examining it all for exploitability. From the serendipitous approach comes most of the SIGINT headlines – the front page stories.

And there is no question of its value. But it is unpredictable. It produces immense detail on a minor unit one day and on the entire O/B the next, but by its nature it cannot sustain its effort on any one subject. There is always a new fire to put out. Often, by itself it cannot answer the commander who asks, "Do all of the forces that opposed me yesterday still oppose me? How have they changed? Are there new ones?"

3. *Continuity*

As opposed to Serendipity (Crypt or T/A), which monitors the entire stream of intercept but selects only a small portion for exploitation, Continuity, i. e., traffic analysis, must process it all. It must sustain the effort, carry knowledge gained from compromise forward into overtly different communications structures, keep continuity, keep track of all of the units opposing us yesterday, and identify new ones. This is cleaning the stables.

Intercepted communications come in "sets," messages and chatter exchanged between two or more call signs representing command or staff functions of one or more military units. The most repetitive terms in each set are the calls used, and the most useful term (because it is not arbitrary) is the frequency applied by the intercept operator. These are

the built-in terms labeling a set. They often are enough to identify the set, but the terms occurring in messages and chatter must usually be used as well.

Since calls and frequencies (and other terms) change frequently, the same *net* may be presented by many *sets* over a period of time. And each major unit controls many nets, an army group or military district often having as many as one hundred nets potentially active at the same time (plus the nets of its subordinates). These represent the different command and staff functions of the unit, working with different combinations of organic and subordinate command and staff functions, and with many alternate nets (morse, voice, and printer parallels).

Fortunately, the collection of terms and the typical formats into which they fit (like sentence structure) are usually sufficiently revealing to permit the intercept operator and traffic identification analyst to label them with some accuracy—at least as to country and service.

So traffic analysis – the proper labeling of each set of intercept, of each radio station, and of each activity – begins at the intercept set. Based on the initial set label (case notation), the intercept is sorted, and more or less homogeneous piles are given to each analyst for further processing.

Continuity is the "Alias" problem; set 542 is the same as set 821, which was the same as set 127, etc. Done exhaustively, it reveals only the number of nets active. It does not reveal who the stations are, where they are, what functions they represent, or where the net fits into the network. It does permit the application of the answers in one set to all of the other sets of which the set is a continuity.

Serendipity, either as a separate function or that of the continuity analyst in the course of his processing, provides the who, where, what, and structural location of the net. Continuity carries it forward.

4. *Battlefield Vocabularies*

Most of the categories of military communications terms are encoded for electrical transmission. A list of arbitrary terms is *generated* (calls, addresses, codewords, etc.) according to some sort of system. From the semifixed list are *selected* subsets of terms for each period according to other, local systems. And the selected subsets are *allocated* to various users according to still other systems. Frequencies follow the same scheme, except that they are not arbitrarily generated; they are selected and allocated according to locally devised systems.

For each period, then, the selected subset is matched against a list of plaintext meanings. The plaintext list (units in O/B order, command and reporting phases, etc.) does not change very much over a long period of time. The encoded terms for transmission change frequently and usually provide for a number of alternate terms each period. Change in the plaintext list of units usually indicates change in O/B.

There are two stages in the traffic analytic attack on the problem of battlefield vocabularies. The first is the isolation of terms in a category which presumably come from one subset: calls used by one or more related nets that possibly come from one call sign book; addresses used in an area presumably selected according to a common set of rules from a master list; frequencies used by related nets and presumably allocated from the same subset; etc. The collections of terms so isolated are input in the second stage, that of solving the systems of *generation* of the master list, of *selection* from the master list for the period being examined, and the *allocation* of the selected subset to the corresponding plaintext list of terms.

Generally, there is little point in trying the second stage (solution) until the first stage is accomplished with some accuracy. And the first stage depends upon the success of the continuity effort, since a single set of data will not, normally, provide enough terms to permit successful attack.

Success in any part of the attack makes the whole much easier, (e.g., solution of the call sign book generation system permits set and nets to be associated rapidly, which then provides collections of data for attack on the frequency problem).

5. Activity

Continuity and the solution of battlefield vocabularies, as discussed above, provide the skeleton of Order of Battle. Changes in O/B from one time to the next provide inferential information about target intentions: e.g., this division has been resubordinated from Army A to Army B; the capability of Army B is thereby measured and inferences may be drawn about its intentions, etc.

Better information is often available. The cryptanalytic solution of enciphered text and the traffic analytic solution of plain text and brevity codes often give more direct and valid information about capabilities and intentions. Grids give relative unit positions and artillery (tube, missile, and bomber) targets. Codewords give implementation times for activities and may, if solved, define the activity. ELINT cuts may reveal concentrations of radars in defense of troop concentrations or locate beacons for airborne drops.

The Serendipity approach acts, almost entirely, on the isolation of such revealing texts and then finds all related nets to build that part of the network for the period. And it may well be that this is all that is possible in a fluid battle area of some size.

Continuity attempts to account for all units, both active and inactive. If it can be accomplished, this complements serendipity by providing information on the existence of reserves and uncommitted and inactive units, and by providing the skeletal O/B to which interesting activity can be related.

PART II. THE COMPUTER

1. *The Computer is an Idiot*

So, too, is a shovel, a lever, and a quill pen. The difference lies in the time lag between human control and mechanical or electronic action. A shovel, lever, or quill pen is used directly and immediately by a human being; a computer is programmed for use in the future.

Directing action at second hand and in the future is very difficult, particularly if it is impossible to predict what may be sensed at the time the action is required.

You drive your car at some speed towards an intersection. Without conscious thought you identify all fixed scenery and all moving objects. You compute their probable positions at the point of time at which you will be in the intersection and, semiautomatically, you hit the brake, accelerate, or continue at your previous speed. But just try to get the car safely through the intersection by briefing a non-driver (the programmer), who will in turn instruct (program) a rote passenger on what to sense and what, in each possible circumstance, to instruct the blind driver to do.

The difficulty lies in trying to explain what you have done, step by little step, when you have written a sentence with your quill pen or driven your car.

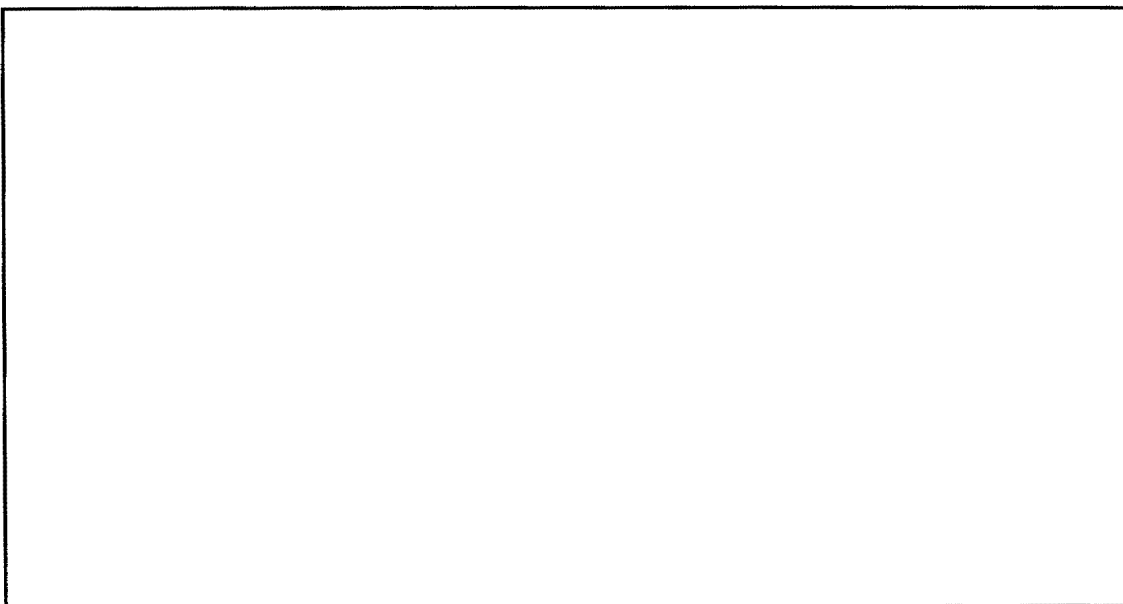
If it can be done, it multiplies your power very considerably. The computer, properly instructed, can drive a hundred cars as well as one. For this reason, it is a challenge and well worth trying.

2. *What is the Computer Now Doing For T/A?*

Programs now in being, or shortly to be implemented, accomplish a great deal of the clerical work previously done by traffic analytic personnel. In most cases the computer attempts no identification of call, or set, or net with specific military units or nets, but groups like things together and applies the best label added to the intercept by humans. The following process illustrates the type of computer-analyst interplay now in being.

P.L. 86-36
EO 1.4.(c)

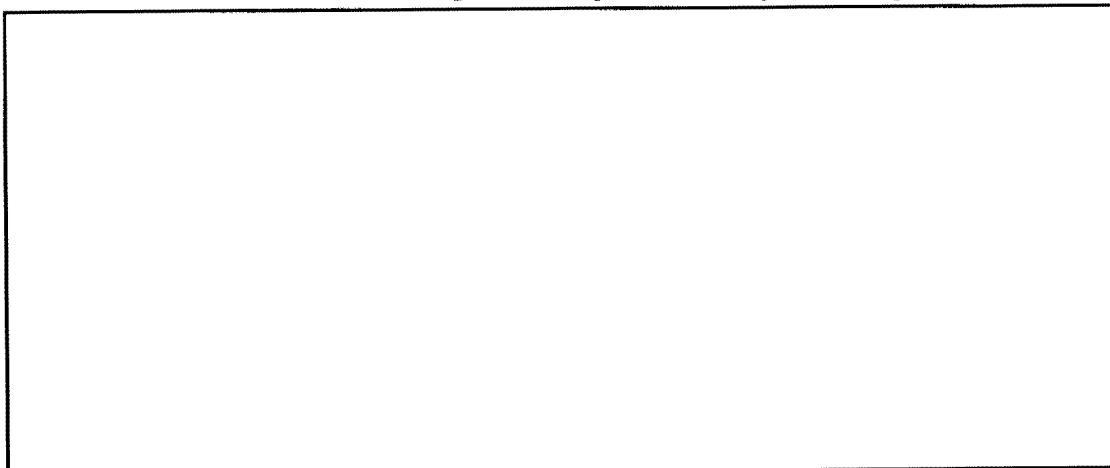
~~SECRET SPOKE~~



The characteristic of the above type of process is the control of the system by the analyst; if it is used as one of his tools and all of the tables and files in the system are maintained currently, it will work well for him; if the tables are not maintained, it will cause the input to degenerate. The analyst is necessary to the system.

3. What Could the Computer Do?

The manual T/A payoff is in the identification of O/B and of activity. Much of this is concealed by the periodically changing battlefield vocabulary, which humans deal with as "The Game": "sounds like"; "in this context may be"; "the activity around it suggests it may be a synonym for"; etc. It is not believed that this is within the computer capability at this time, only because human beings cannot explain what they do at this point.



It is conceivable that the computer could build its own vocabulary of terms and formats from its input, applying to each the best identification of the set from which it came. A

~~SECRET SPOKE~~

new set could then have its terms and formats compared to the computer vocabulary and all identifications of previous sets retrieved. The retrieved identifications might then be examined statistically. The computer output might be

This set is (65%) division X

It is (20%) artillery, (5%) SSm

Callsign ABCD is (25%) the umpteenth motorized rifle regiment

Best guess as to case is.....(42%), or.....(29%)

This set is a continuity of sets.....(82%),.....(74%),.....(68%)

4. The Computer in Combat

Those functions now accomplished or suggested for batch processing could be done in much smaller batches, approaching on-line processing. The computer could materially assist in sorting, selecting, and identifying intercept sets and activity and could, as well, guide intercept, if the callsign history could be directly queried. (The previous identification and analytic comment – "want" or "don't want" – could be made immediately available).

Portability is feasible with solid-state computer components, and size is approaching the manageable. The problem lies in analytic use. Unless analysts are trained in and appreciate the multiplication of their power possible through intelligent use of the computer, it should not be used. The machine will only hamper the successful manual operation.

5. Present Problems

The feeling still persists that the computer is an expensive, mystic black box and that it must operate efficiently at all costs. Its priesthood still carries this belief to the point where its efficiency is emphasized at the expense of the T/A problems it serves. It is not comparable to a quill pen, kept sharpened and ready for instant use when needed, as well as being adaptable to the needs of the moment. It is more like a railroad, with trains leaving on schedule (more or less). If you happen to want to go to that point at that time, climb aboard (after having applied for a ticket well in advance). But to change the composition of the train or its destination or its scheduled times is almost beyond the passenger's capability.

Not because of characteristics inherent in the computer, but rather because of the management of this tool, the computer is not adaptable, not flexible, not responsive to T/A needs – except in the case of those needs that are continuing and that can be conceived whole and perfect when the process was first requested.

6. Conclusion

The computer can perform many of the functions that traffic analysts do manually. It can group data into homogeneous piles. It can apply previous annotations to sets of data and to terms. It can sort and list. It can select data or discard them, based on previously determined criteria. It can provide a sizable memory with total recall. It can continue a process whose rules have been defined

PRL.866386
EEO1144 (4b)

The computer cannot, at the moment, solve the unusual. A new system or a nonstereotyped sentence cannot be given meaning by the computer until an analyst instructs the computer. (It can, however, collect the unusual for the analyst.)

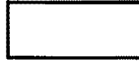
The computer cannot, at the moment, be used to solve the little T/A problems, such as solving a frequency rota or running out a callsign pattern or solving a key sequence. Most of these problems are hand generated by target signal personnel and change frequently. By the time the analyst explains to the computer what to do, he can solve the problem himself.

The computer, then, is a tool, a big tool. It can materially assist the traffic analyst in his work and do most of the clerical labor associated with T/A. The system does not now permit it to be used as a little tool or a fine tool. And it is no more than a tool.

Build a road with a stream shovel; don't use the shovel to open a door! And yet I think that maybe, if I could tinker with it a bit and try a few things, I could open that door with the computer. Certainly, a better road could be built.

Traffic Analysis on [] Data Transmission Systems

P.L. 86-36



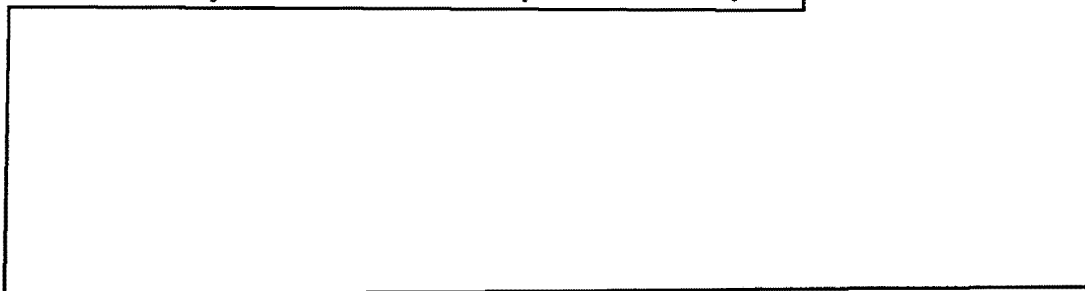
P.L. 86-36
EO 1.4.(c)

Conventional traffic analysis has never been a glamorous job. It conjures up a picture of untidy desks piled high with traffic of various hues and drawers bursting with much-thumbed tools of the trade - callsign tables, network diagrams and, of course, that inevitable odd file of mysterious unidentified bits and pieces that is never thrown away because it is such a convenient place to put more unidentified bits and pieces. Amidst this scene we find that most maligned representative of the COMINT community, the traffic analyst, continuing his never-ending battle against the nonstop flow of paper and wiping his carbon-stained fingers all over his clean white shirt in a gesture of despair. His plight is known to all of us, and many have done sterling work to try to emancipate him. But despite some machine aid, the lot of the traffic analyst working on conventional communications, morse, radioprinter, and voice has not changed very much in twenty years.

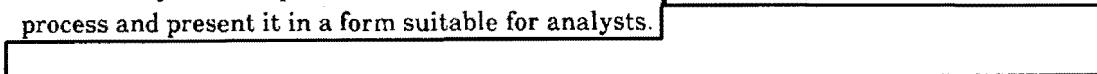
The advent of [] data transmissions threatened an even greater amount of paper unless new techniques were developed. As a result of these new techniques, there appeared on the scene a new breed of clean-shirted traffic analyst whose previously cluttered desk is now rarely used except perhaps for lunch.

What are these data transmission systems that so successfully changed the face of traffic analysis? They are very much in vogue these days, and the air and air defense forces of many nations, including the United States and USSR have turned to such systems to augment and sometimes replace their existing facilities for air surveillance reporting and weapons control. Basically, a data transmission system permits the automatic transmission of formatted information by means of digital coded impulses and gives a greatly increased traffic capacity coupled with speed and accuracy.

Two such systems are in current operational use by the []



To analyze and exploit the information transmitted in a data system, it is necessary to process and present it in a form suitable for analysts. []



[REDACTED] And so the idea of traffic analysis from a display was born in 1961.

It is probably necessary at this juncture to placate the skeptics amongst the old hands who will have none of this nonsense and allay the fears of the new entrant who feels he was not born with enough hands. What is this traffic analysis that can be done at a display and how is it carried out?

[REDACTED]

It may seem that the new concept of traffic analysis demands a super breed of analyst – a combination of radar operator, machine specialist and data traffic analyst. It is true that more is demanded of the analyst, but experience has shown that the novelty of the new approach and the results obtained from it have whetted the appetite of those morse traffic analysts who have been transferred to data analysis, with the result that they have adapted themselves to the new techniques with a minimum of training. Air and air defense analysts are, of course, weaned on the interpretation of tracking, and machine handling has been made very simple as almost every requirement of the analyst has been covered in programming while operating consists of little more than pressing the required pushbuttons on the control panel.

Although the technique of traffic analysis from displays is still in its infancy, its

[REDACTED]

Although it was an air/air defense data system and the threat of unmanageable amounts of paper that resulted in this new method of traffic analysis using an immediate combination of man and machine and replacing the paper by displaying the data, it should not be construed that analysis by display will be limited to the air/air defense world or even to data systems. Traffic analysis involves looking at paper, a lot of paper. It involves sorting this paper and inevitably re-sorting. Obviously such procedures call for machine aid, and in many areas this help is available or planned on a larger scale. But in almost every case this help will be in the form of machine listings, perhaps alleviating the sorting problem, but replacing one lot of paper with another. Perhaps the reader can see a use for displays instead of listings for his own problem. Perhaps a "quick look" would permit enough discrimination to reduce the volume of listings to a minimum. It is hoped that this introduction to traffic analysis or [REDACTED] data transmission systems, traffic analysis from displays, will open up a new world to the many who did not know of its existence and that it will stimulate further ideas on how to handle traffic.

Traffic Analysts of the Future

P.L. 86-36



Let us take a glimpse into the future – a future much different from today. I want to convey a man-machine concept that visualizes a fairly complete computer complex from intercept collection to the storage and retrieval of intelligence by the consumers from their desks. This sort of data processing assembly line would contain numerous computers with process computers controlling other computers but with humans at certain vital points along the way. Traffic analysis is one of the subsystems in this assembly line. I want to fire the imagination of the traffic analysts, managers, data systems analysts, R & D types, and others so that some day we can put the pieces of the assembly line together. Certain operations in the assembly line can best be done by humans until we advance the state of the art in artificial intelligence. In the meantime, we want to make it as easy as possible for the humans.

In this assembly line concept, the output of one stage must be compatible with the input to the next stage or process without extensive conversion, reformatting, or other changes.

The concept visualizes that as the data flow through the system on a real-time basis, a human can call data from mass storage, state the processes (programs) to which they should be subjected, and display the results at any stage desired. Based on the results, the analyst would then specify additional processes until he is satisfied with the results.

Let us visualize how a traffic analyst of the future may work. First, he will work from a console instead of a desk. The console would probably contain at least two CRTs or TV-type displays plus an electric typewriter and light pen. All of these would be connected to a remote computer for the proper man-machine relationship.

Now suppose our traffic analyst wants to get continuity on a very important net for day X. All data are in a mass storage device to which the computer has access. He states his problem on the keyboard in stereotyped English, e.g., "Get all unidentified data from file X for day Y." (This can be further defined as data from certain intercept stations.) "Build and compare net structure with XXMT50101 using Program _____." "Display net structure of XXMT50101 on scope 1 and best match other nets on scope 2." After visual observation of all possibilities, the analyst wants more information on new net number 2. He asks for a display of the crypt systems used by both nets if messages were passed.

Then he asks for a comparison of traffic volume, address groups, etc. Having satisfied himself that the net is correctly identified, he types on the keyboard, "Identify new net structure number 2 as XXMT50101." This activates a program that updates all records in the master file with the new information.

After going through all his assigned nets and maintaining continuity, he calls for the display of all remaining unidentified nets. One group of control and three outstations has now appeared for several days. The analyst calls for a display of the group on a geographic mask with all stations geographically located by the previous day's direction finding. Then he reviews callsign, frequency, cryptosystems, usage and other elements to determine the echelon of the group. With two or three possibilities, he requests an order-of-battle (OB) display showing units located on a geographic overlay. In this way he can compare net structure with order-of-battle of units known to be in the same geographic area. After several tries, he finds a unit which fits so far as echelon structure, but one of the outstation location does not fit. Considering the possibility that the OB may not be up to date, the analyst requests a display of the OB overlay for the last two months on the second display scope showing day-by-day overlays at a two-second interval. With this technique he notes that one unit has been moving. By projecting the movement up to the latest information, he concludes that the unidentified outstation and the one unit are the same. In addition, he now has later information as to the location of the unit. He adds this information to the OB database so that tomorrow when the intelligence consumers interrogate the OB files for new information, this new location of a military unit will be produced.

Identifications are then made on the keyboard, and all records in the file are brought up to date.

Barometer-Readers' Comments

P.L. 86-36

The future envisaged by [] bothers me, in that it presumes a traffic analyst (and a consumer) who controls the situation, who asks the right questions of the computer. He queries the computer for new data on his assigned nets (or units) and then asks for all other unidentified nets, from which he selects and develops a new net. This suggests that, if he is preoccupied with other tasks, he needn't ask and *nothing* will arrive on his desk (or CRT). TA need not be done?

If this is not true--and, in fact, he must ask--then why the query at all? A set of mandatory queries ("Give me all, massaged") need be programmed only once and the computer will provide the data routinely. This is, I believe, the TA function. The target, not the analyst, controls the problem. The analyst (collectively) is *required to examine all of the intercept* and do something with it. If a mountain of material comes in, he must cope with a mountain, not just pick out bits he predetermines as being his.

It's like a police department. Does it sit in its police station asking questions about the status of Hollywood and Vine? Or does it patrol the whole city? Who can judge in advance where a crime will occur?

You cannot ask *all* of the right questions unless (1) there are a very limited number of possible questions and (2) you know all the possible questions.

Sincerely, Fred Mason

Sir:

In reference to Mr. Mason's letter in *COMMAND*, dtd March 1969, I don't believe Mr. Mason has heard about management by exception. The advantage of computers is that they can be programmed to look at the unusual. We do have a mountain of intercept, but it is not necessary for a human to look at it all. The computer looks at it all and finds the unusual or unidentified to which the human can devote all his intellect. Let's use machine for the routine and free humans to be analysts. Most analysts like to piddle around with the intercept because they can see tangible progress - that is, process so many papers. Most junior traffic analysts would not know what to do if they were faced with real analysis.



FRL 886386

Mr. Mason's rebuttal:

Management by exception? Maybe. Analysis by exception? I don't believe in it. Along with our primary mission of reporting all changes in O/B is our continuing obligation to reaffirm the status of all known target forces - to always expand, amplify, flesh out the static O/B skeleton. As a premise, we have never fully described O/B; there is always more we could say about a target. If this is true, neither we, nor our consumers, nor the computers, know enough about our target to say "All is normal, don't report." The fact of "normal" activity is reportable.

It has to do with indications of imminence. No one asks that we report everything, but no one permits us to suppress anything either. It is a very thin line we walk, this Pearl Harbor complex. For instance, it is "normal" for any one of fifty people to be in my little cubicle. At what point is it abnormal? When ten are here at once? Twenty? Fifty? It is now "normal" for Soviet crews to man missile sites in Egypt. If there are now X such crews, 2X is a 100 percent increase, surely reportable. But 3X is a 50 percent increase over 2X probably reportable. 4X is only a 33 1/3 percent increase over 3X; 5X a 25 percent increase; 6X = 20 percent; 7X = 16 2/3 percent; 8X = 14.3 percent. Is there a threshold below which we stop reporting? Or keeping a count? Or let the computer withhold intercept from us?

I'll go a step or two down the road. The computer is fully qualified (with a lot of analytic guidance) to summarize, to reduce the daily mountain to a comprehensible molehill. But the molehill must be regularly and exhaustively examined, refined, compared with previous molehills - and by humans. The validity of this map must be confirmed and reconfirmed by comparison with the mountain it represents (and the mountain itself is only a small sampling of the voluntary overt utterings of secretive "bad guys.")

Fred Mason

A Brief Quarter Century of Soviet Crypto/Traffic Analysis



INTRODUCTION

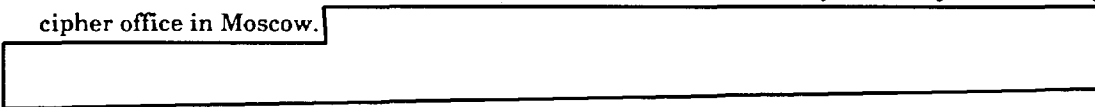
The earliest piece of Soviet communications intercepted by the U.S. was dated early 1944,¹ according to my crypto/traffic analysis (C/TA) records. Since that time we have experienced many varied and complex problems in the TA field. We all know that TA deals with the items of communication that make it possible for the transmission of a message—the analysis of the text of which becomes cryptanalysis. In turn, C/TA is that special area where a cryptanalytic attack is made on the elements of TA in the attempt to assemble order, primarily military order of battle.


I have chosen as the base for C/TA the twenty-five years NSA has devoted to Soviet communications. The progressive improvements of these systems from an elementary state to certainly the most sophisticated yet indicates the degree and effort the Soviets have taken to deny the SIGINT community C/TA success. Also, we can be reasonably sure of the existence of strong Soviet influence in the communications of other Communist countries.

This paper is confined to callsign systems analysis.

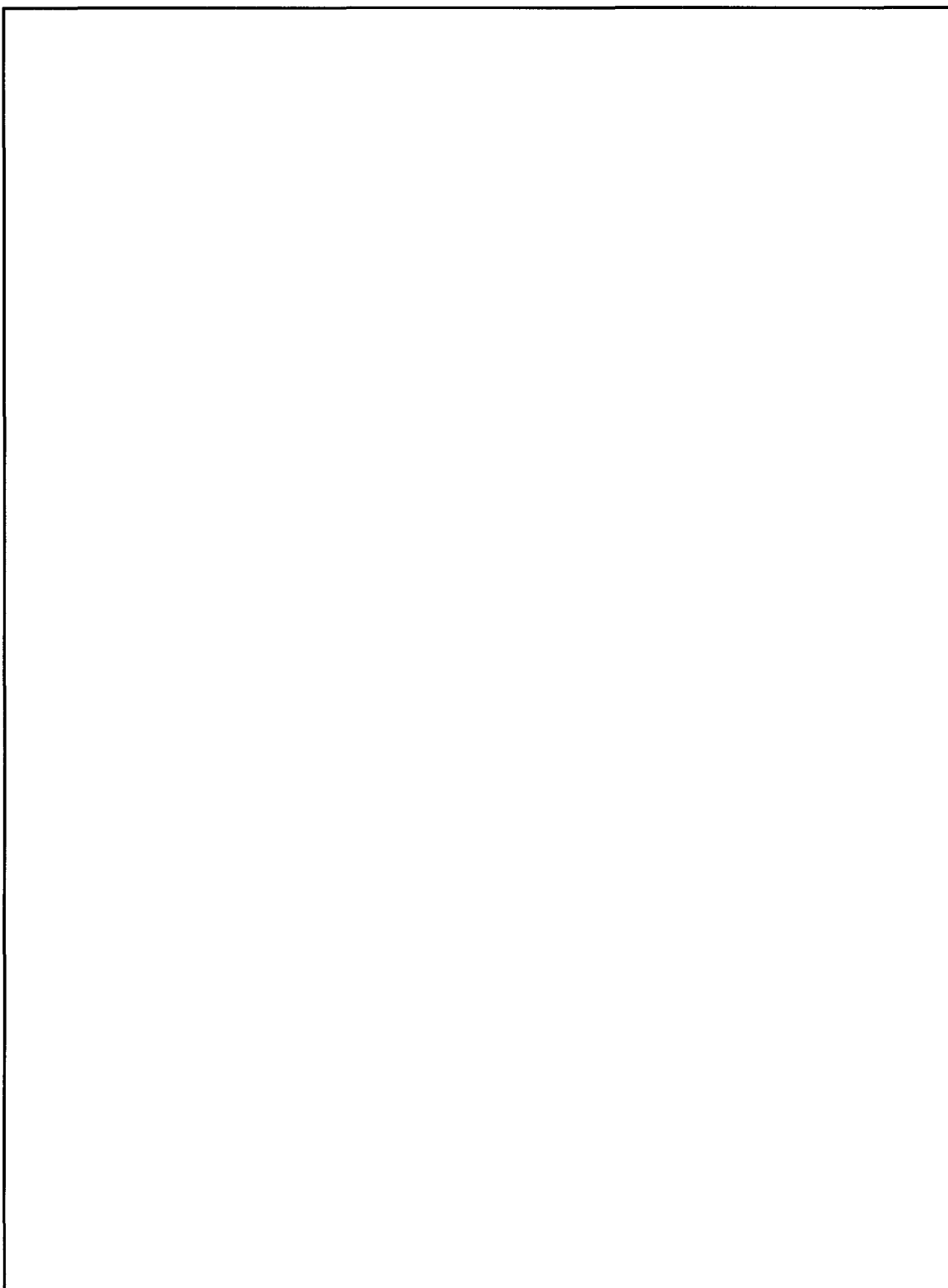
THE PRIMITIVE YEARS

While the United States C/TA effort did not actually begin until about 14 February 1944, we glean an insight into the earlier Soviet posture from captured German documents. We may summarize the period prior to 1 November 1948 as one in which very limited control was exercised over communications of the military forces by the military cipher office in Moscow.



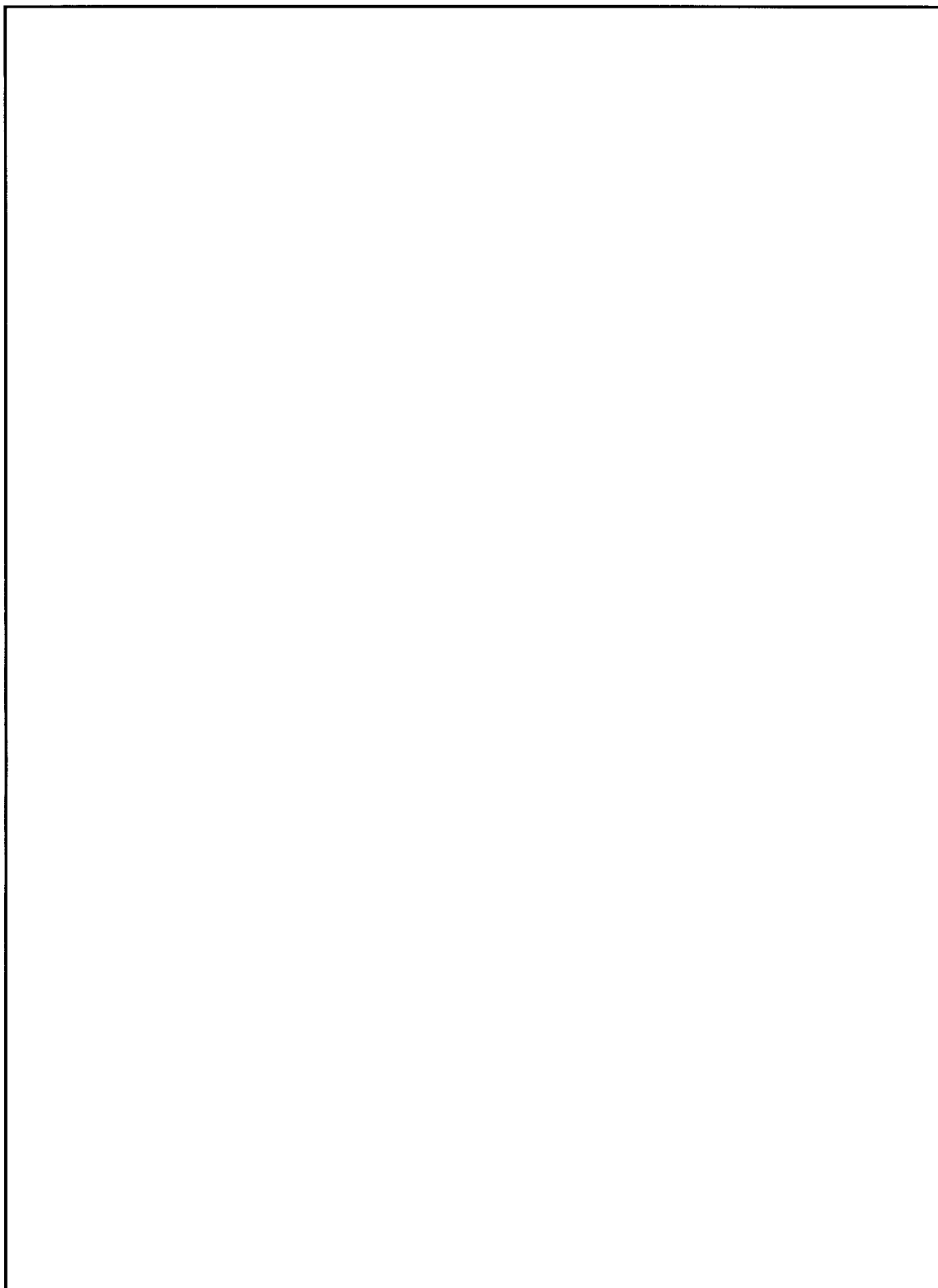
1. The U.S. Navy was already working on Russian traffic in Alaska in 1943. Mr. Frank Raven  and, according to him, the British began passing Russian traffic to the U.S. in 1940 or 1941. The author cannot further document this earlier period and, therefore, would be grateful for assistance from anyone possessing a documentation or recalling the contents thereof to set the record straight.

~~TOP SECRET UMBRA~~



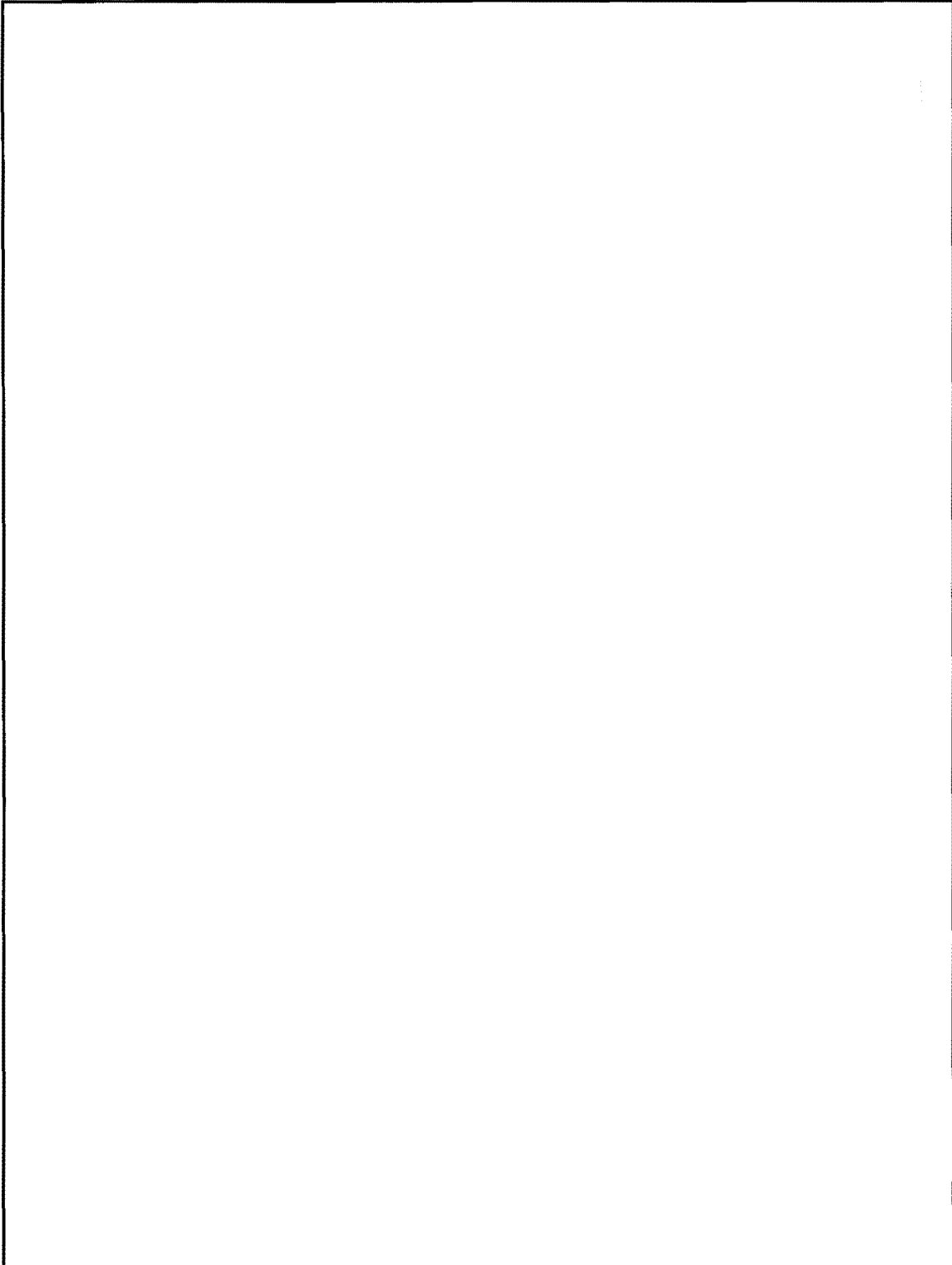
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

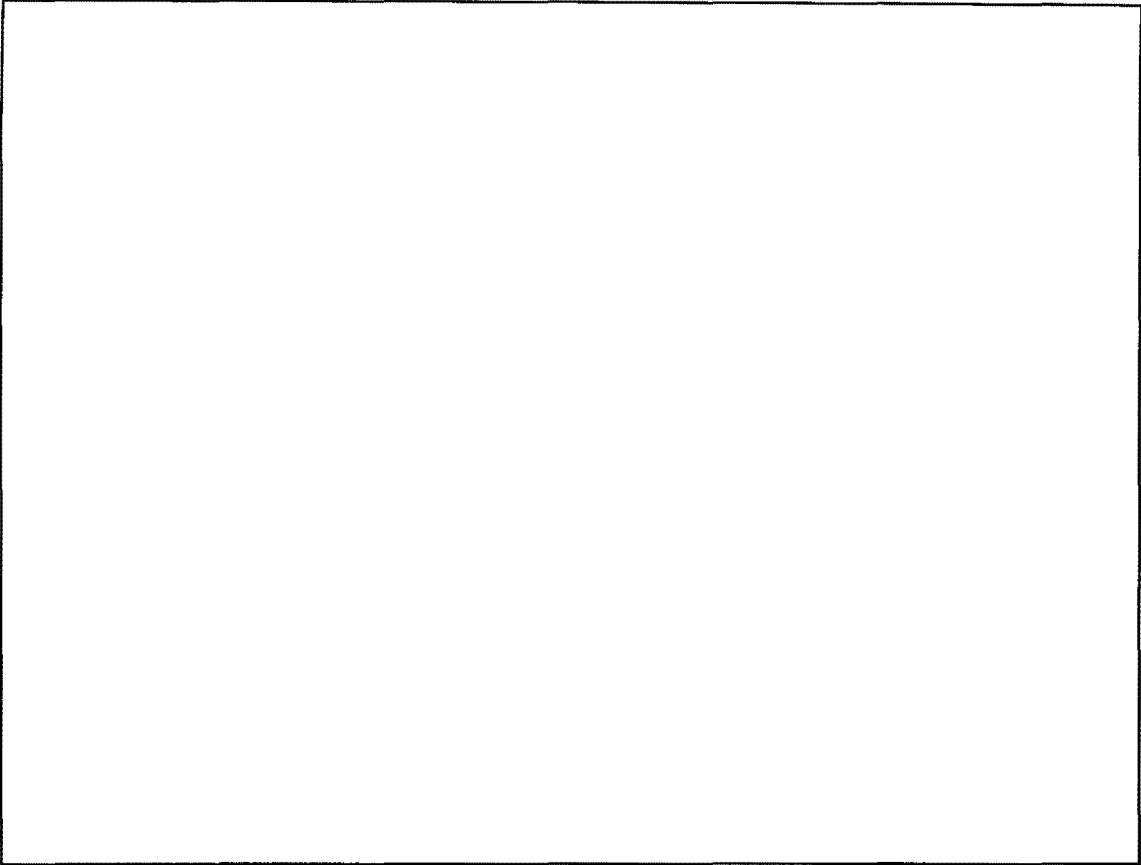
~~TOP SECRET UMBRA~~



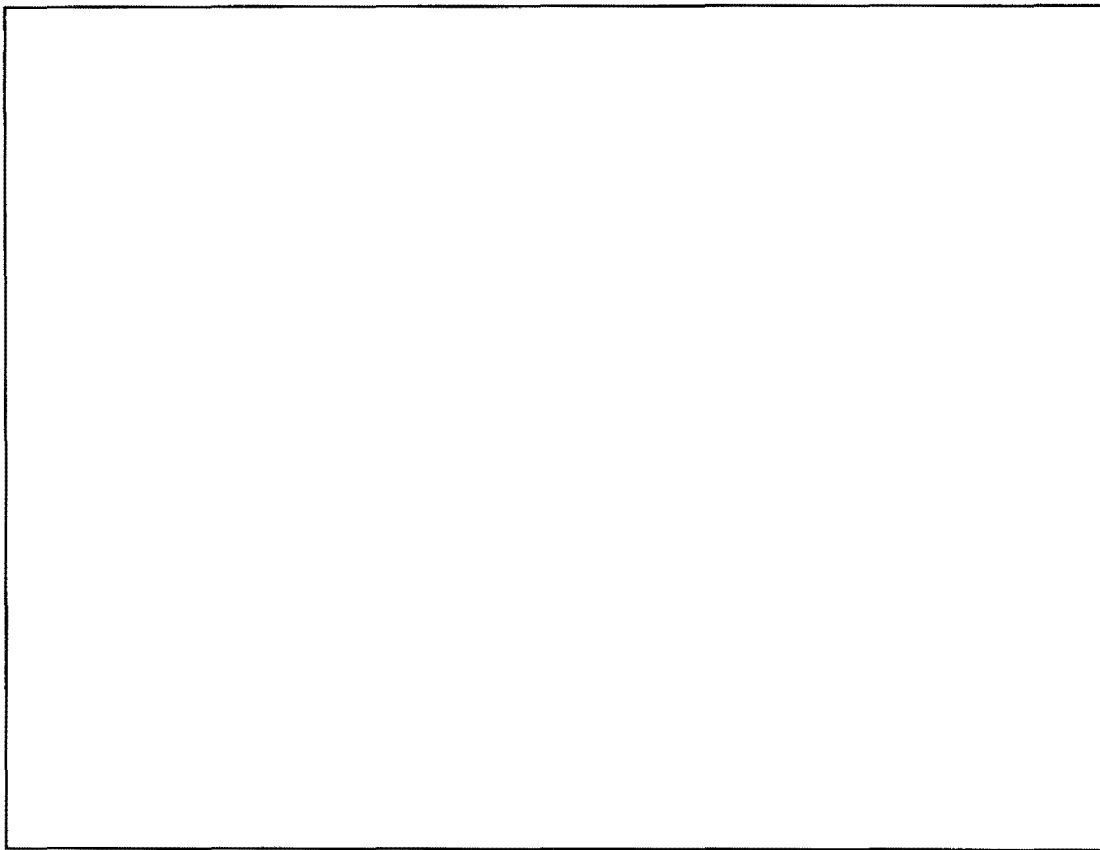
These have been the highlights, briefly told, yet symbolizing the culmination of many sympathetic efforts on these varied systems--each of which represented a greater challenge than its predecessor. I am proud to have served the C/TA effort during the past quarter century, and I envy the role of C/TA of the future because of the inevitable technical challenges.

~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



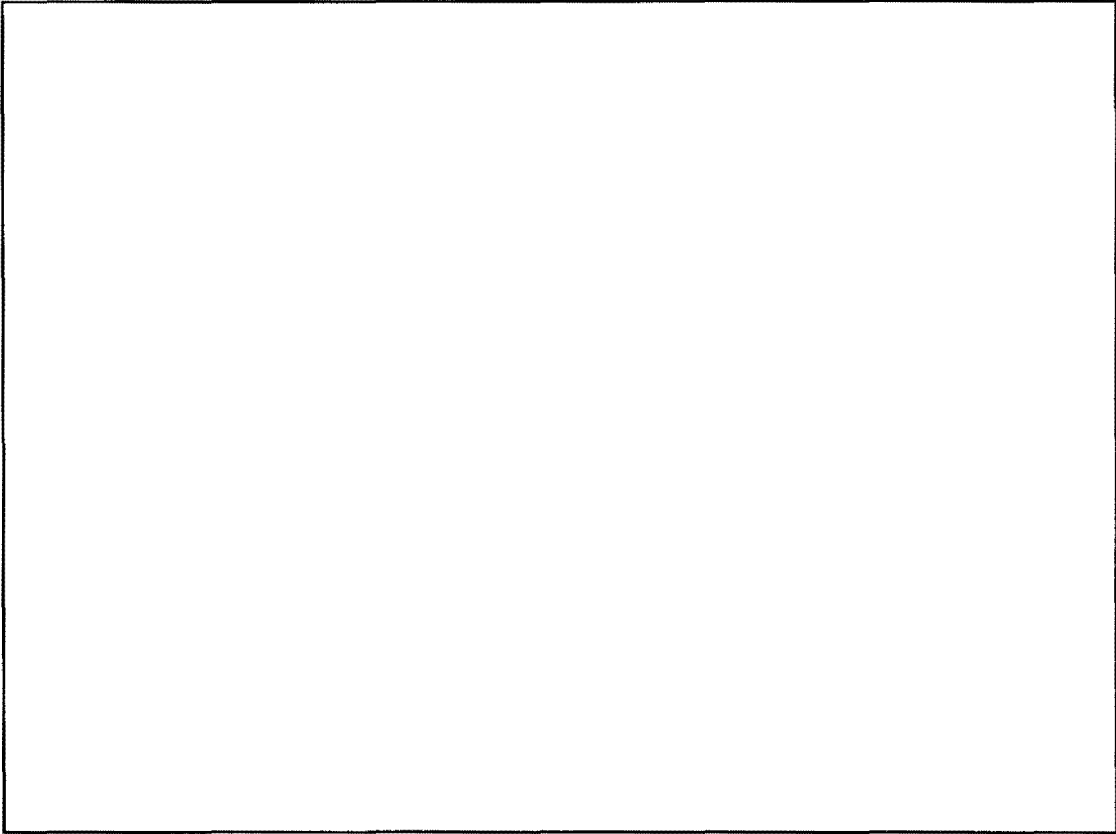
~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

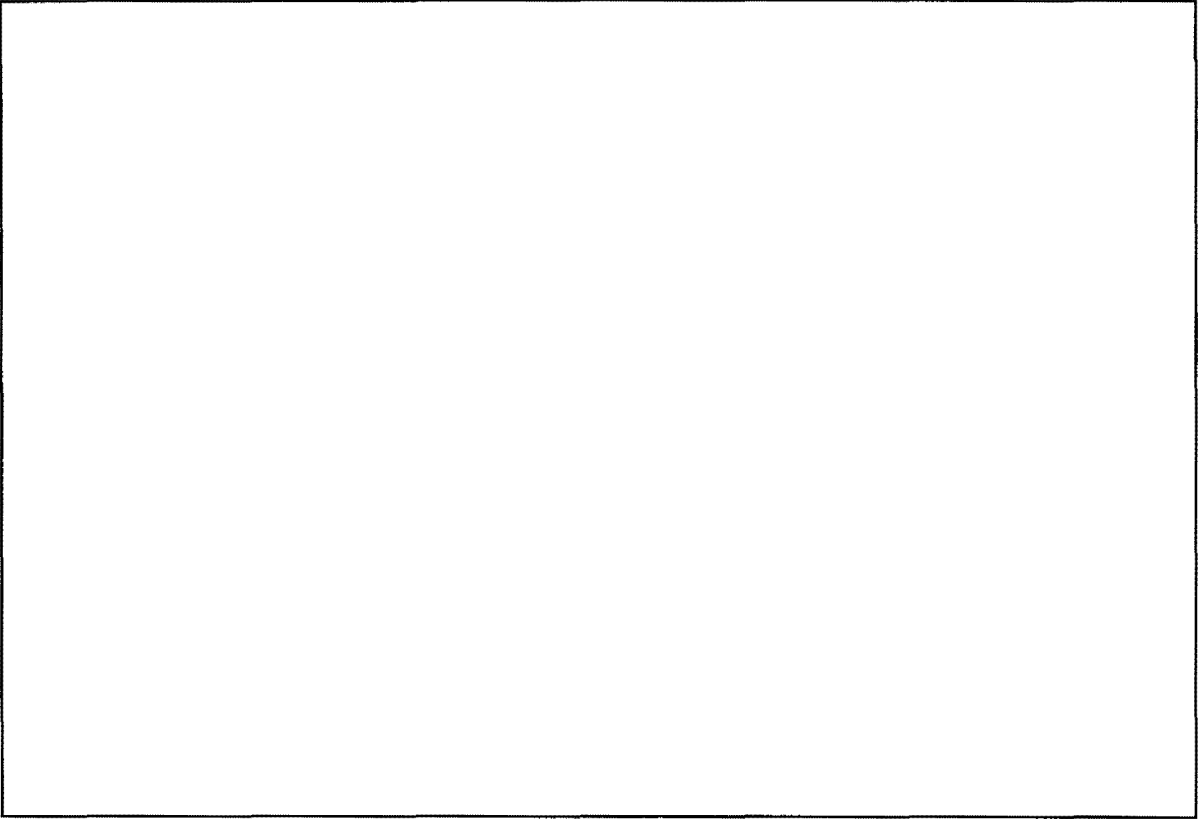
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



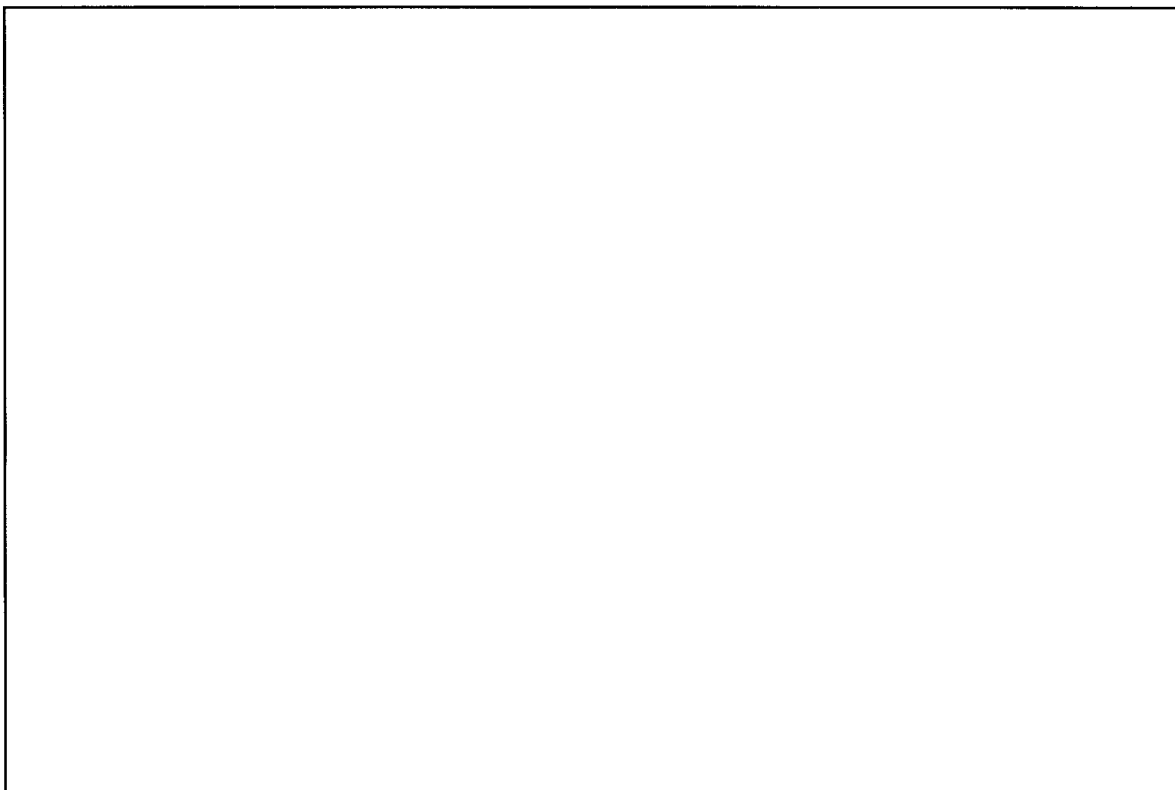
~~TOP SECRET UMBRA~~

~~TOP SECRET UMARA~~



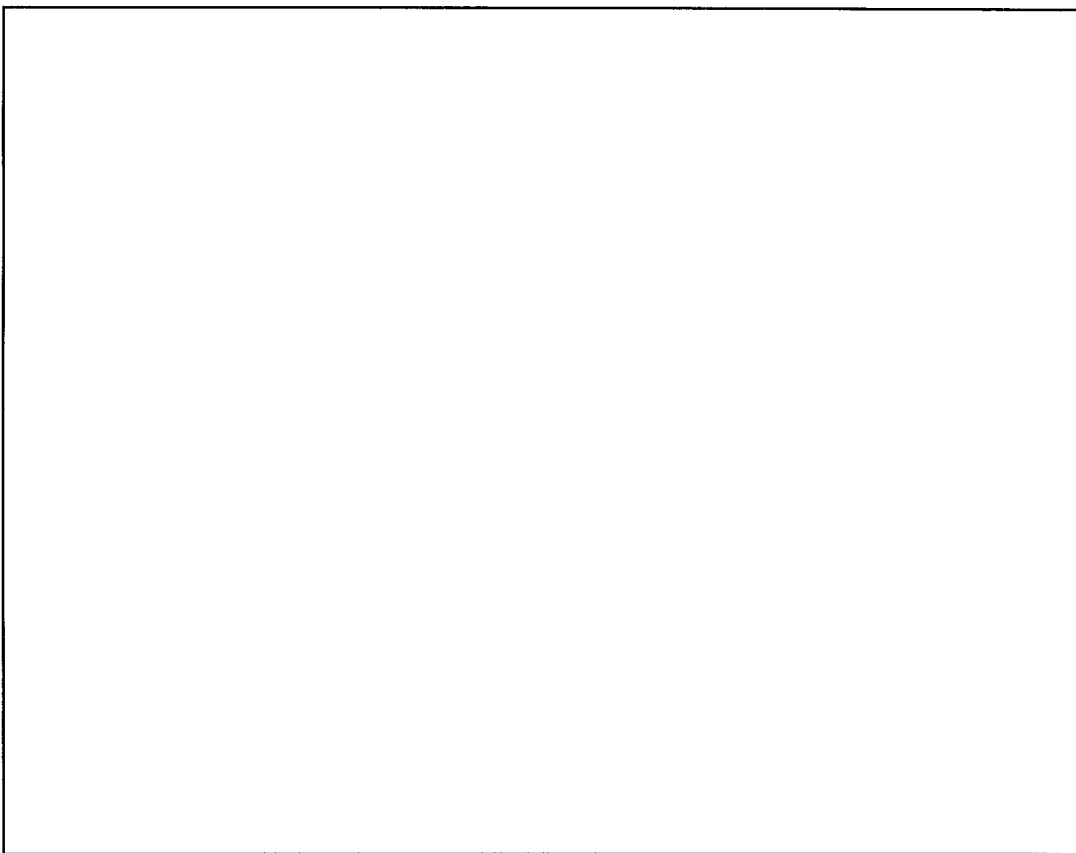
~~TOP SECRET UMARA~~

~~TOP SECRET UMBRA~~



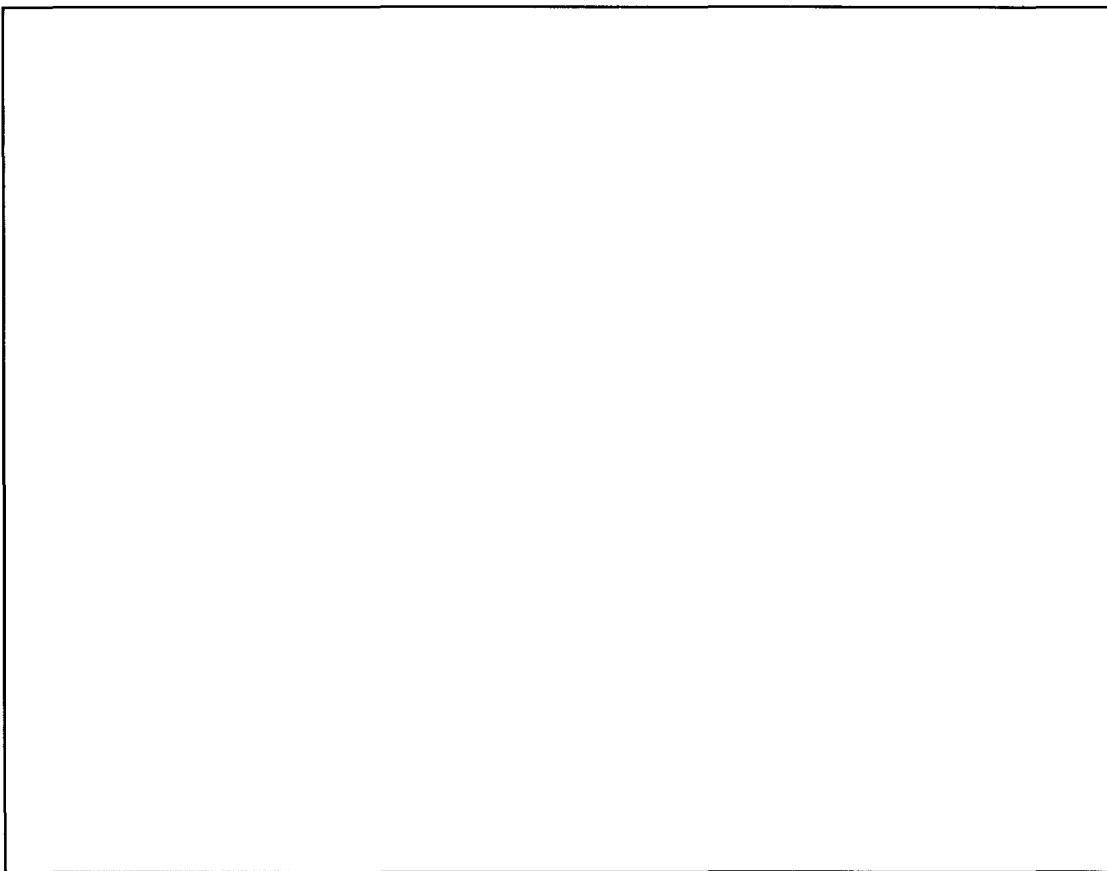
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~

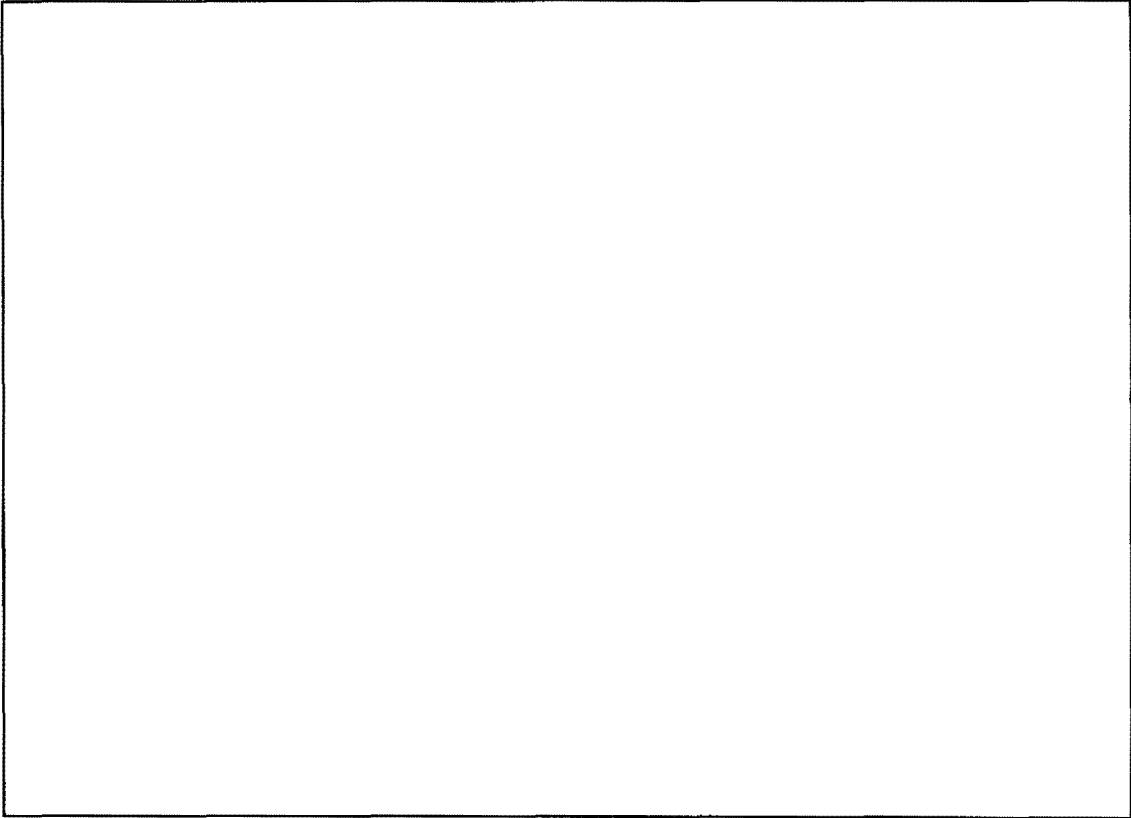


~~TOP SECRET UMBRA~~

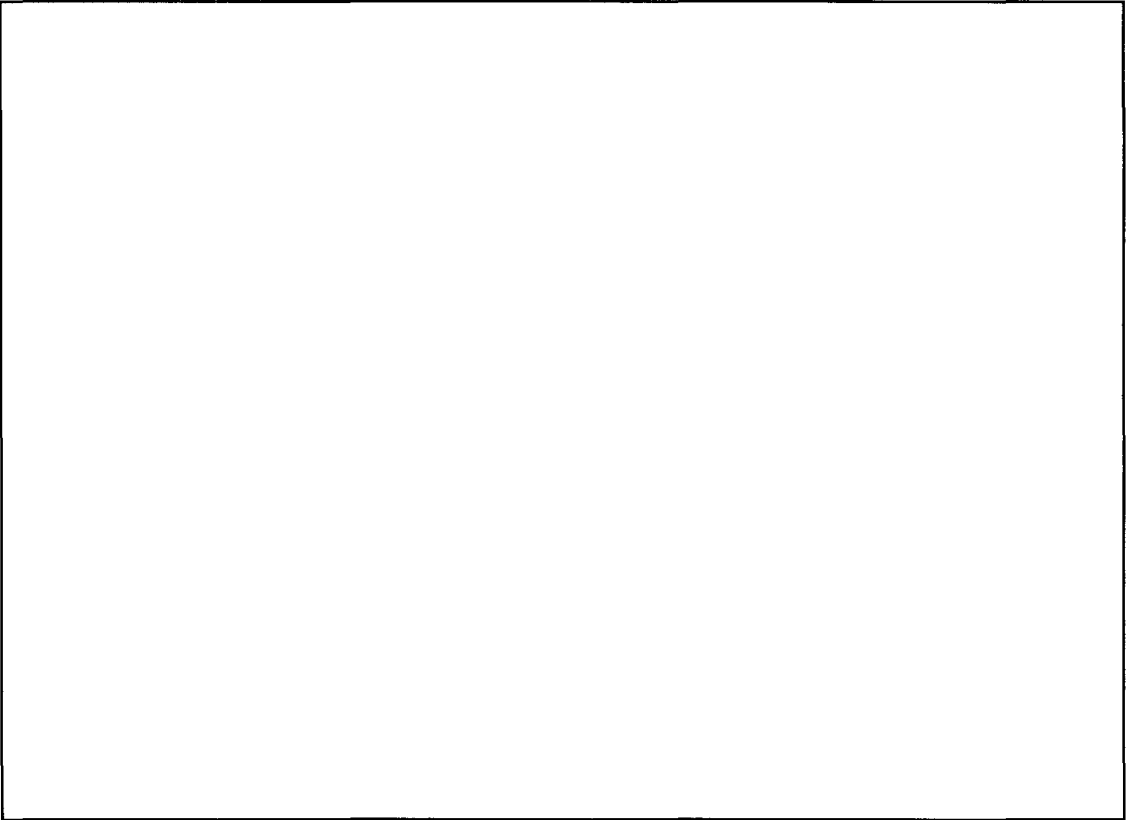
~~TOP SECRET UMBRA~~



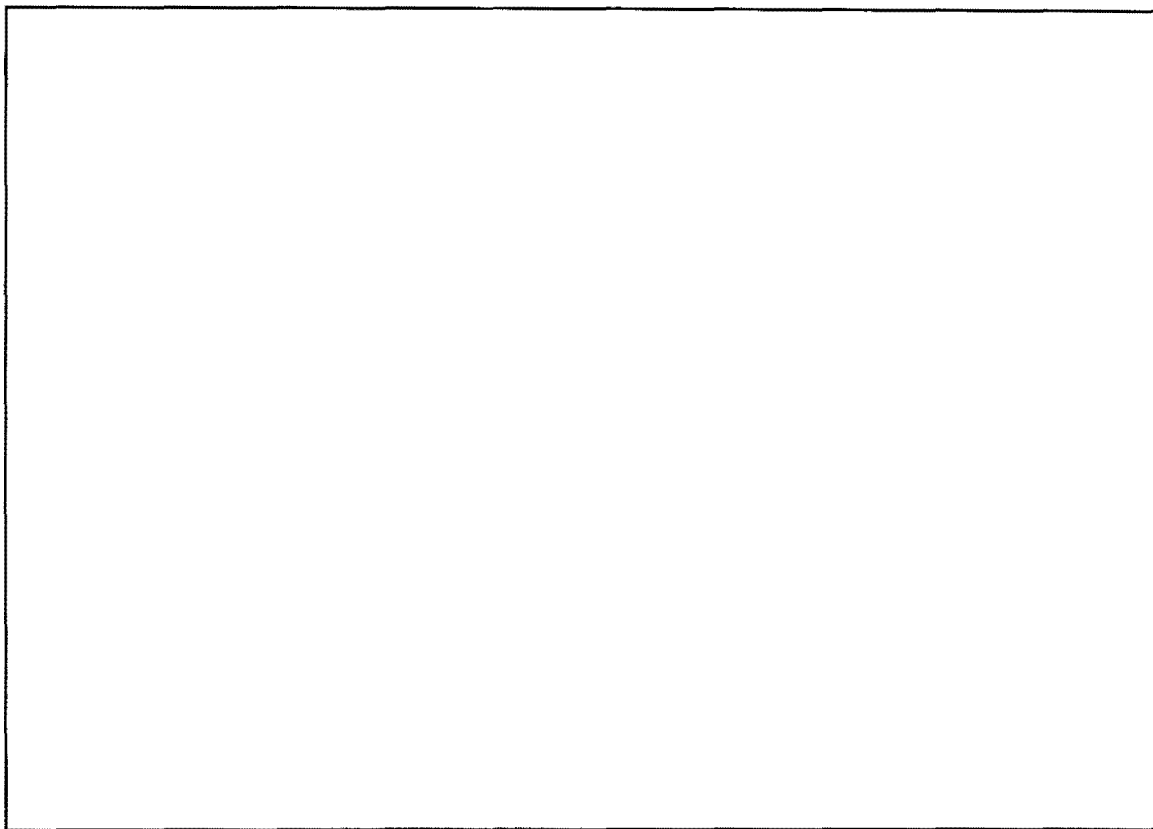
~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~



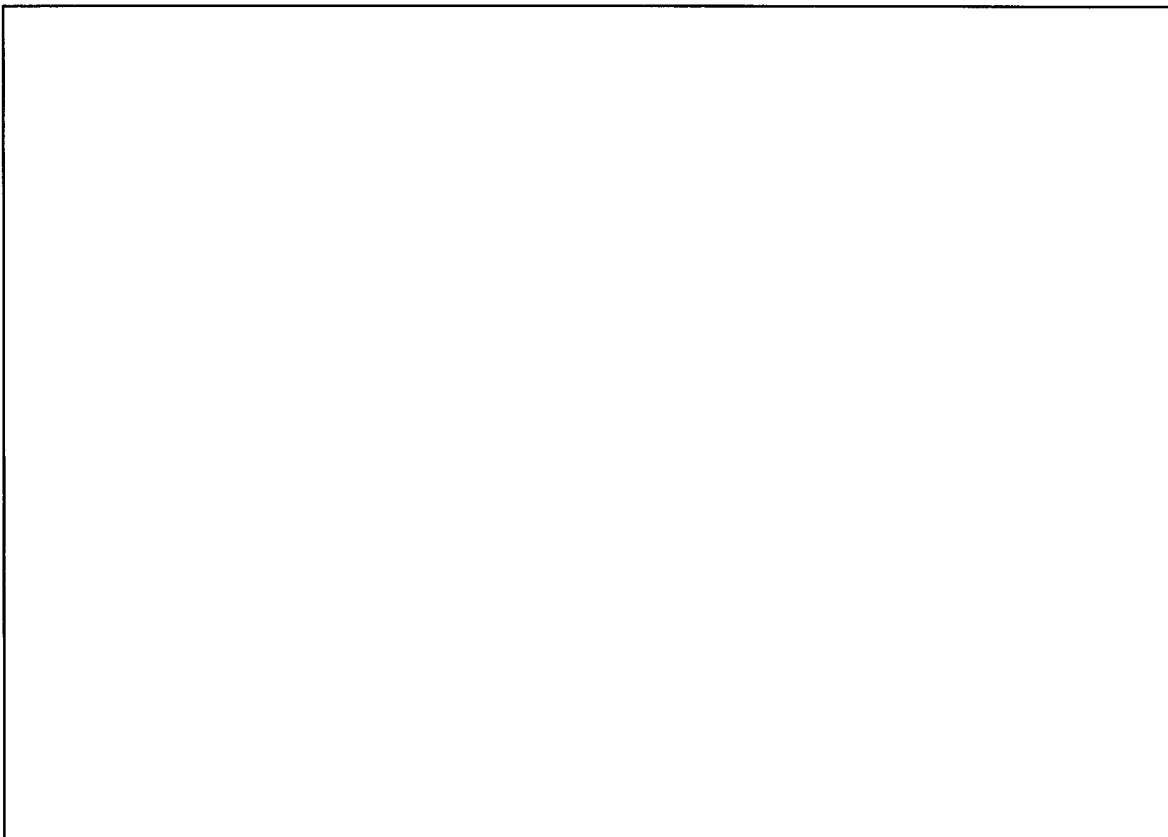
~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

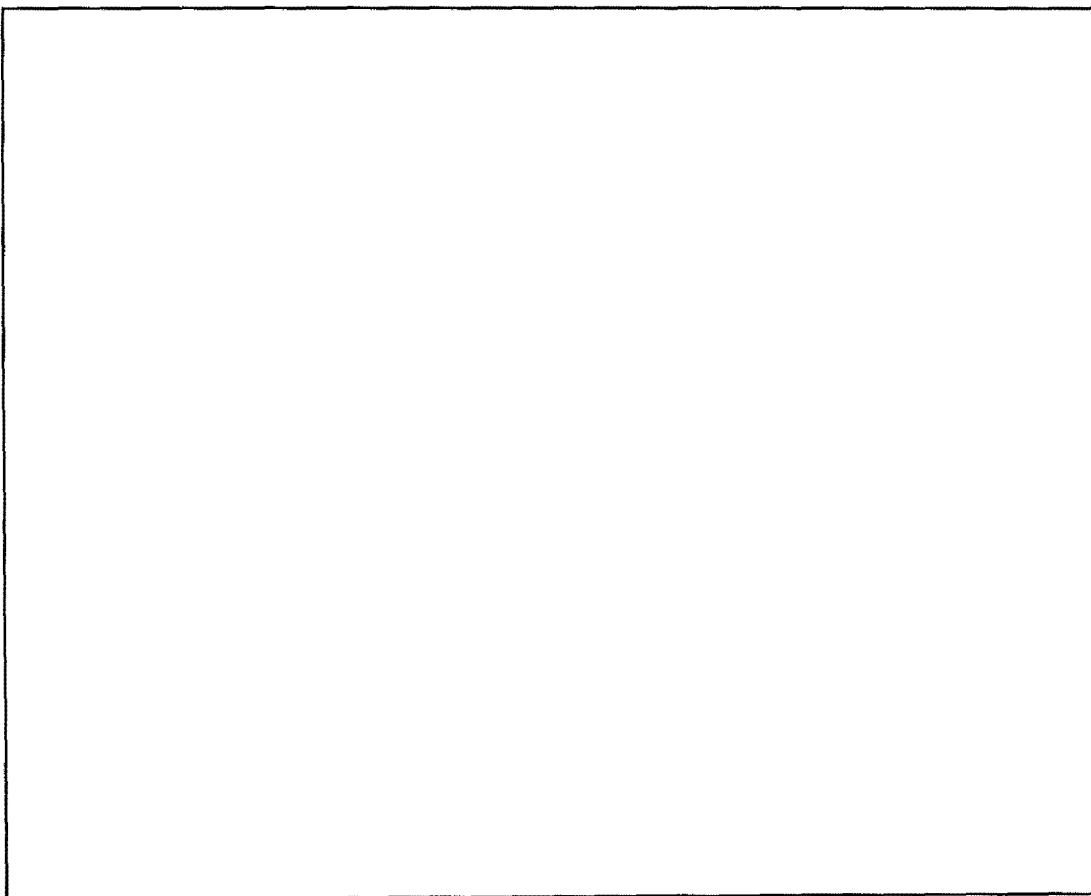
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



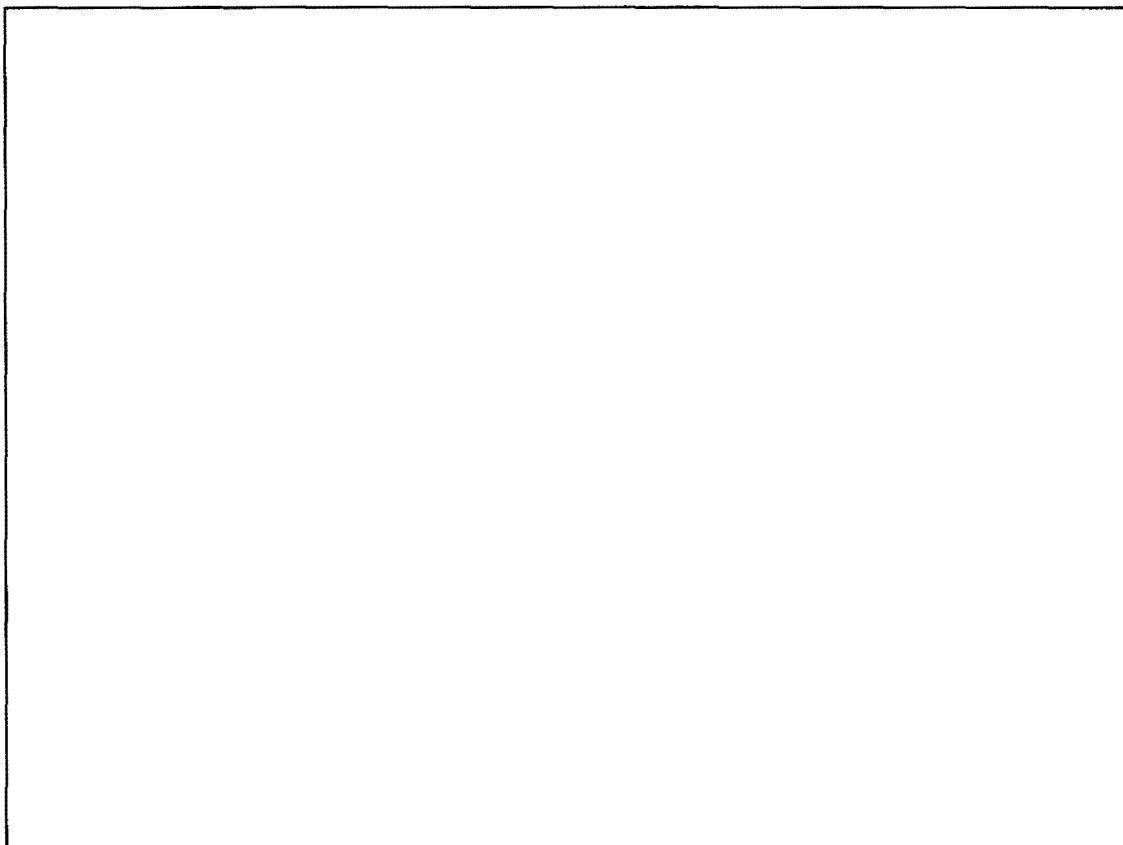
~~TOP SECRET UMBRA~~

~~TOP SECRET UMARA~~



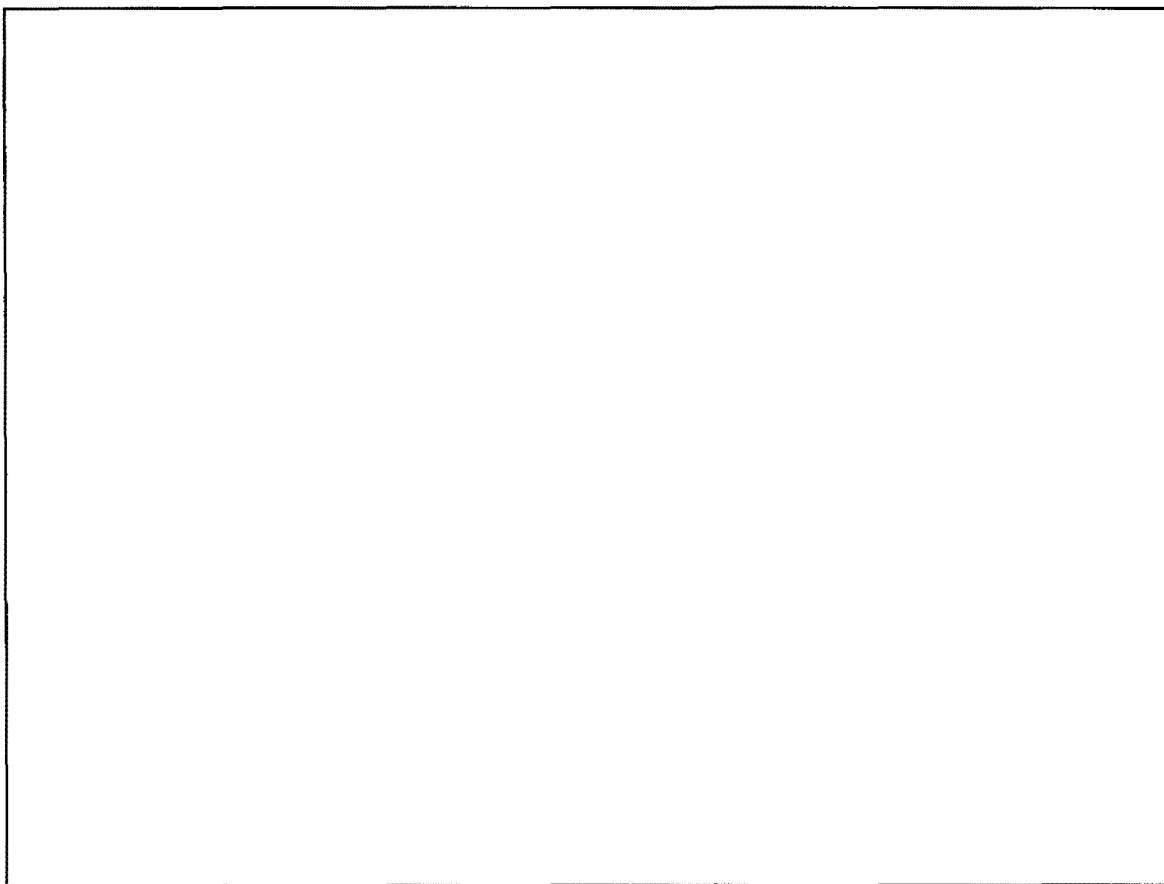
~~TOP SECRET UMARA~~

~~TOP SECRET UMBRA~~



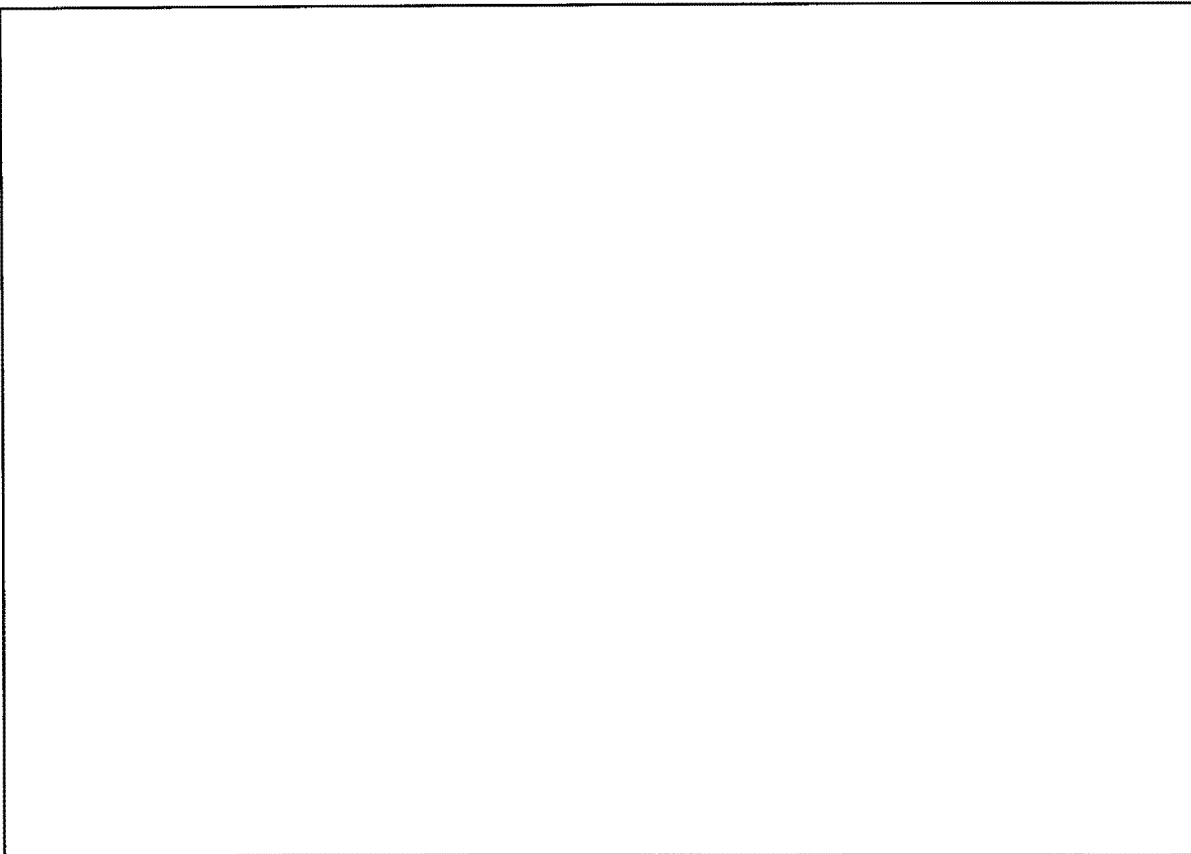
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~

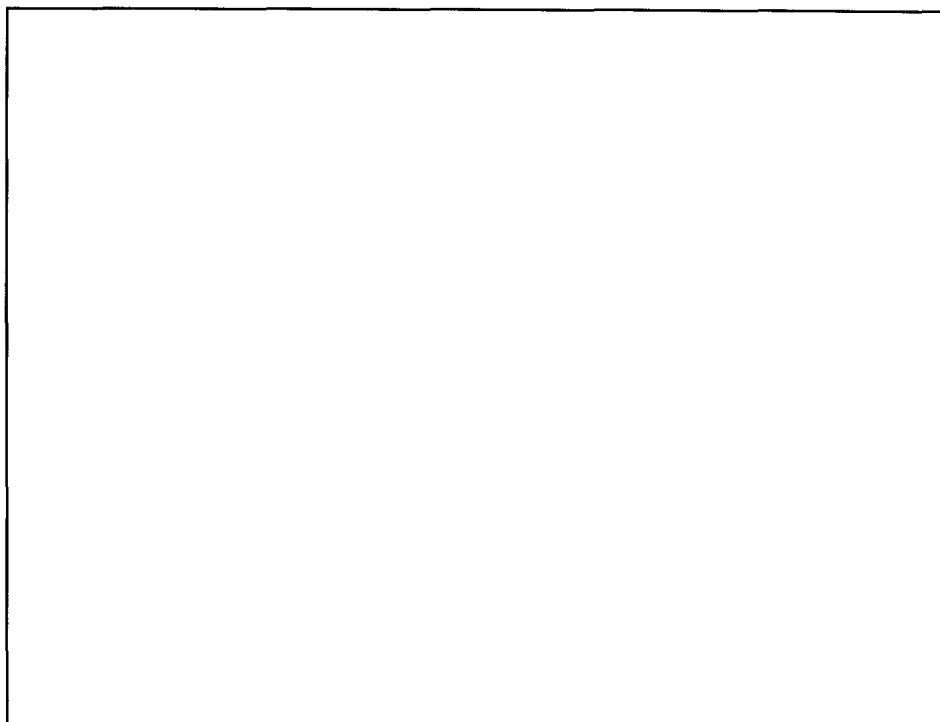


~~TOP SECRET UMBRA~~

~~TOP SECRET UMARA~~



~~TOP SECRET UMARA~~



~~TOP SECRET UMARA~~

~~TOP SECRET UMARA~~

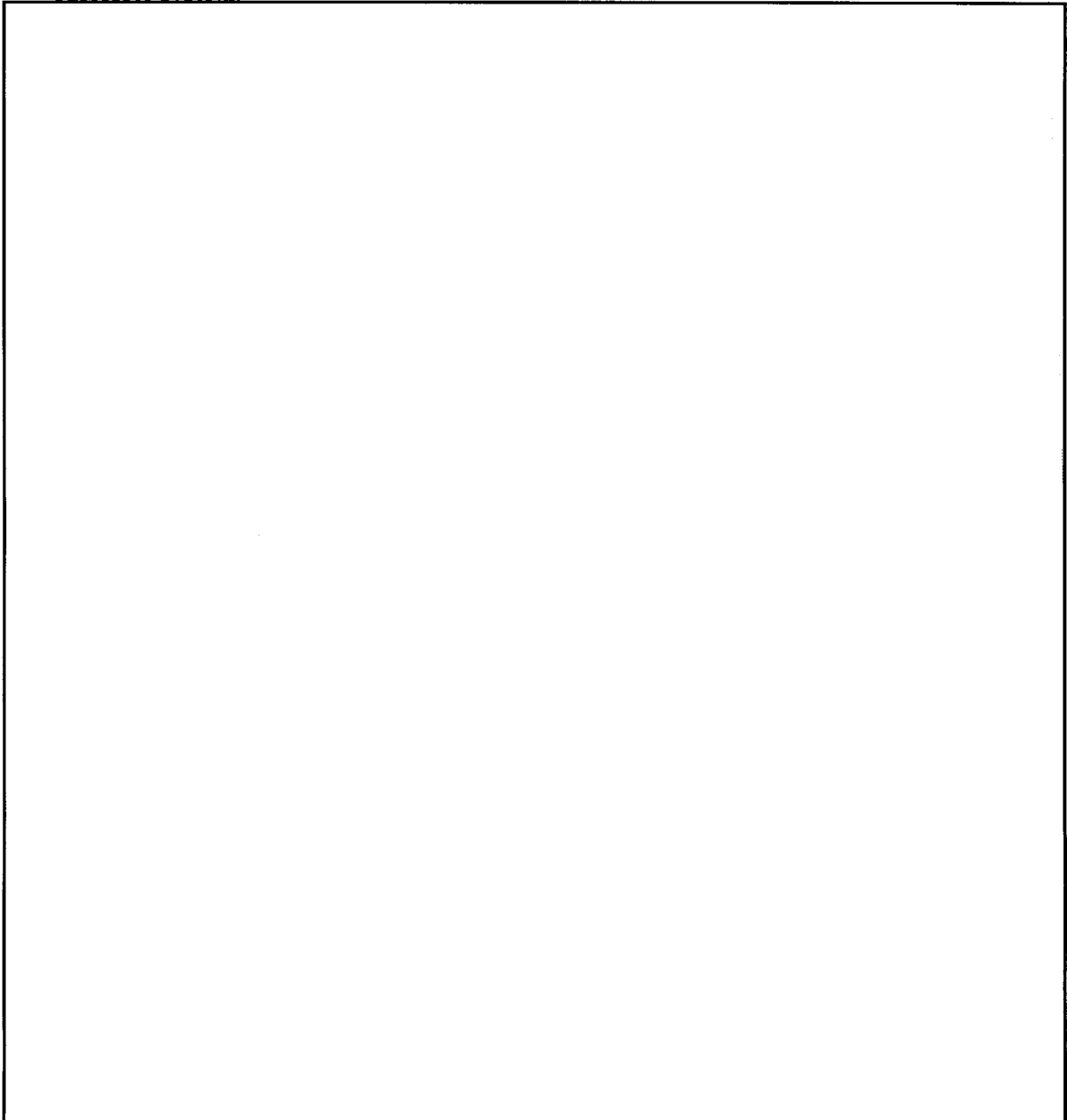
Recovery of a Vietnamese Communist Callsign System



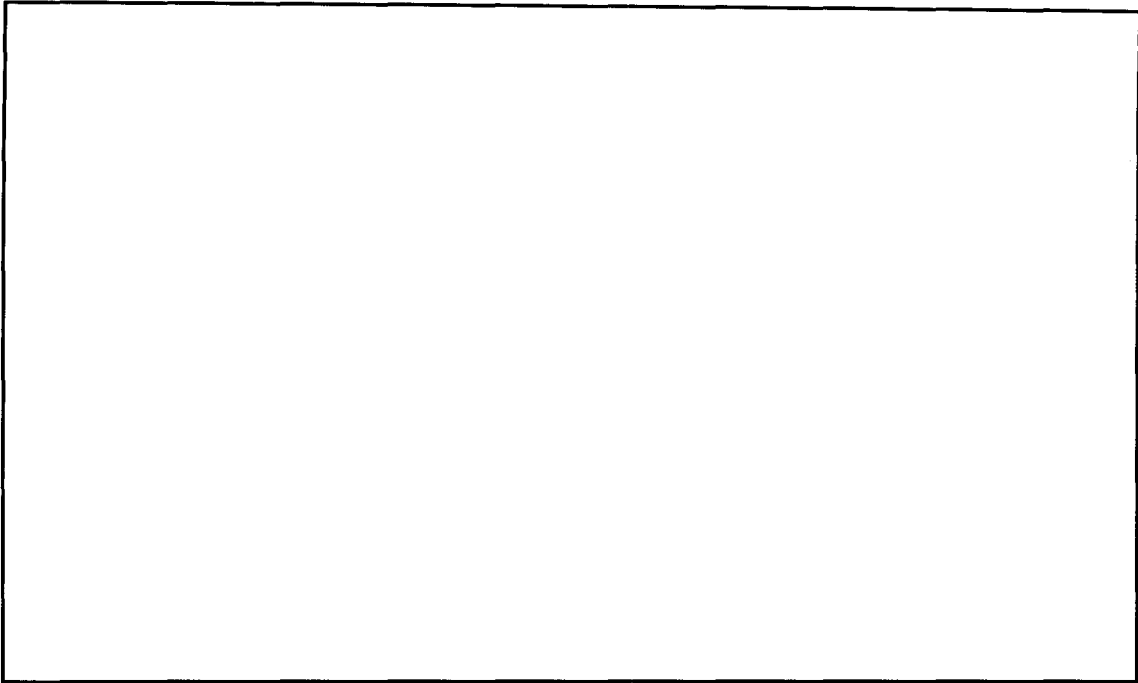
P.L. 86-36

The recovery of the Vietnamese Communist callsign system now titled AO1 (V61A, Abigail-1) clearly demonstrates the value of historical research toward the full understanding of a presumably obsolete signal plan, in that it facilitated recovery of its successor system.

P.L. 86-36
EO 1.4.(c)



~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

True Base: Two Tales (U)

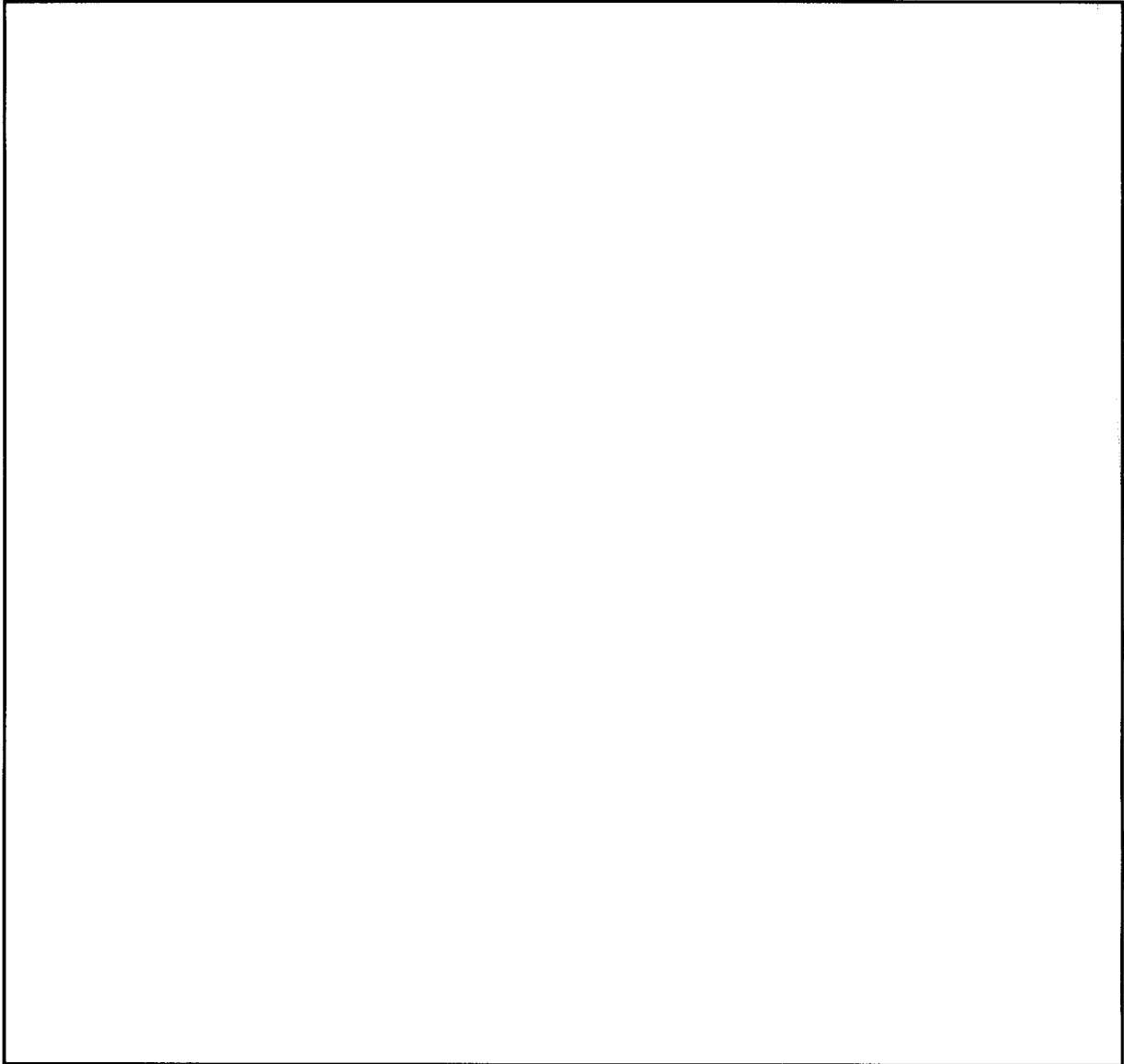


P.L. 86-36

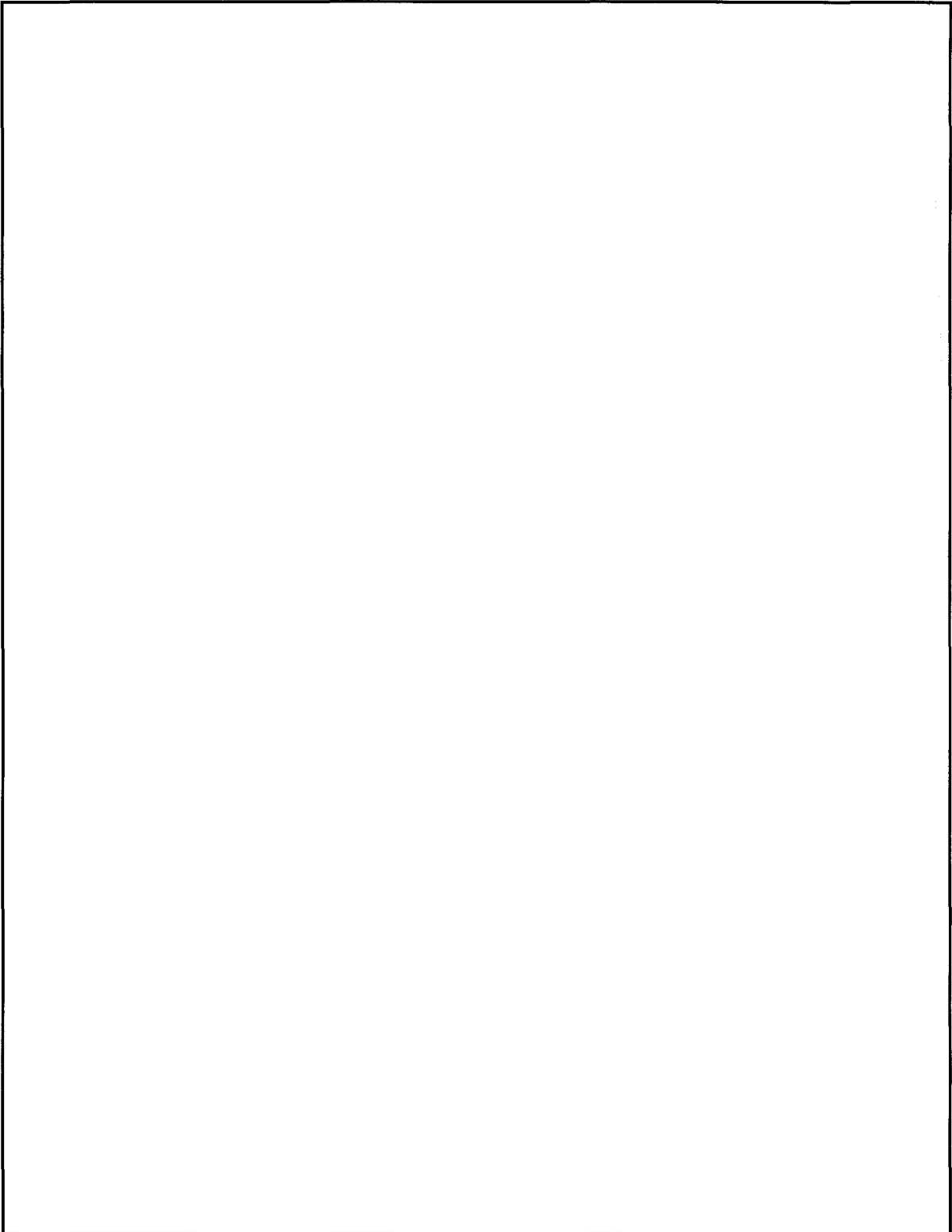
Both of these tales are true. Some of the details have been changed, but not necessarily to protect the innocent. They both have to do with the same thing: what happens to our knowledge about a target when we put one of its systems on true base.

P.L. 86-36
EO 1.4.(c)

FIRST TALE



~~TOP SECRET UMBRA~~



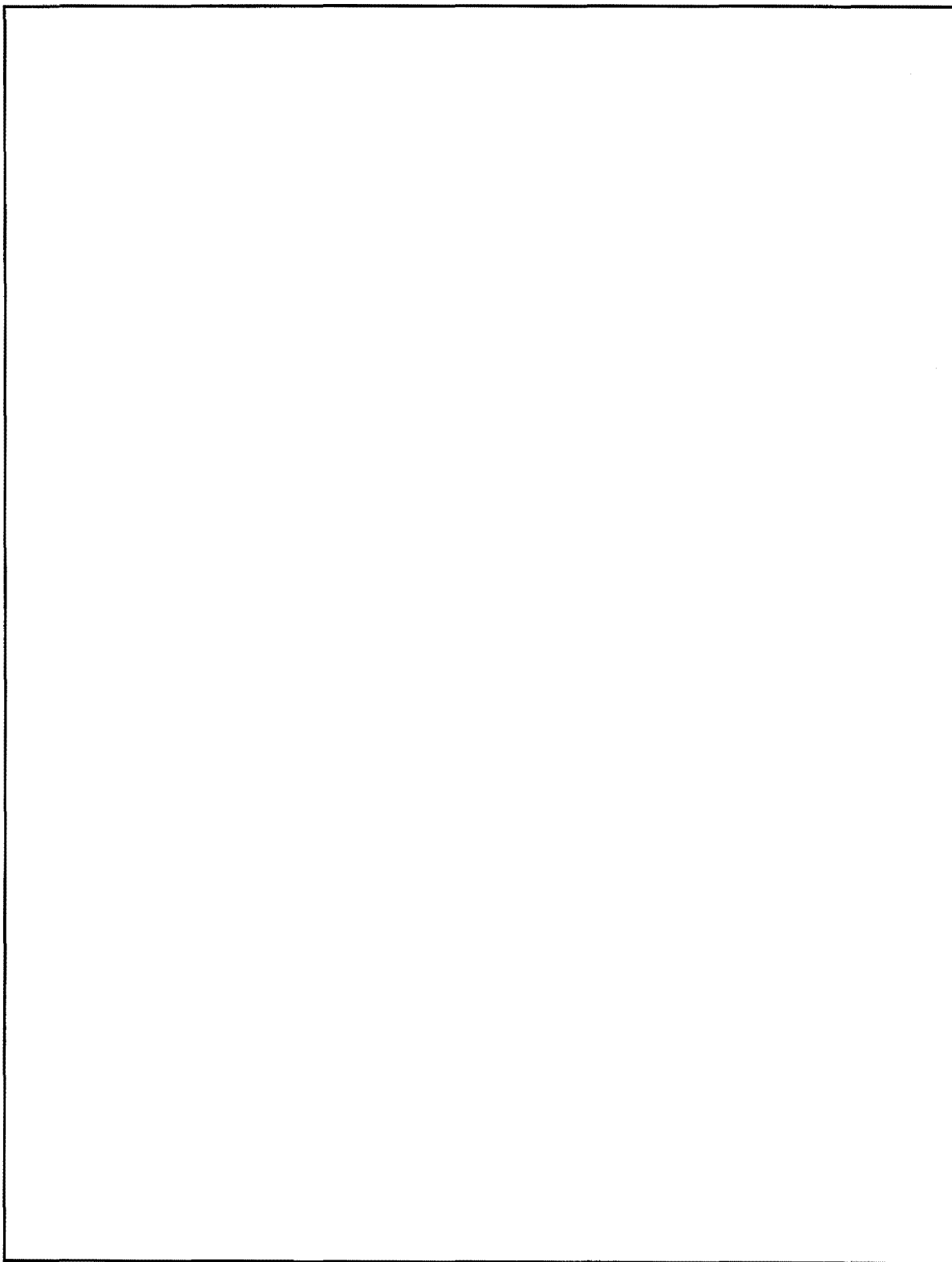
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



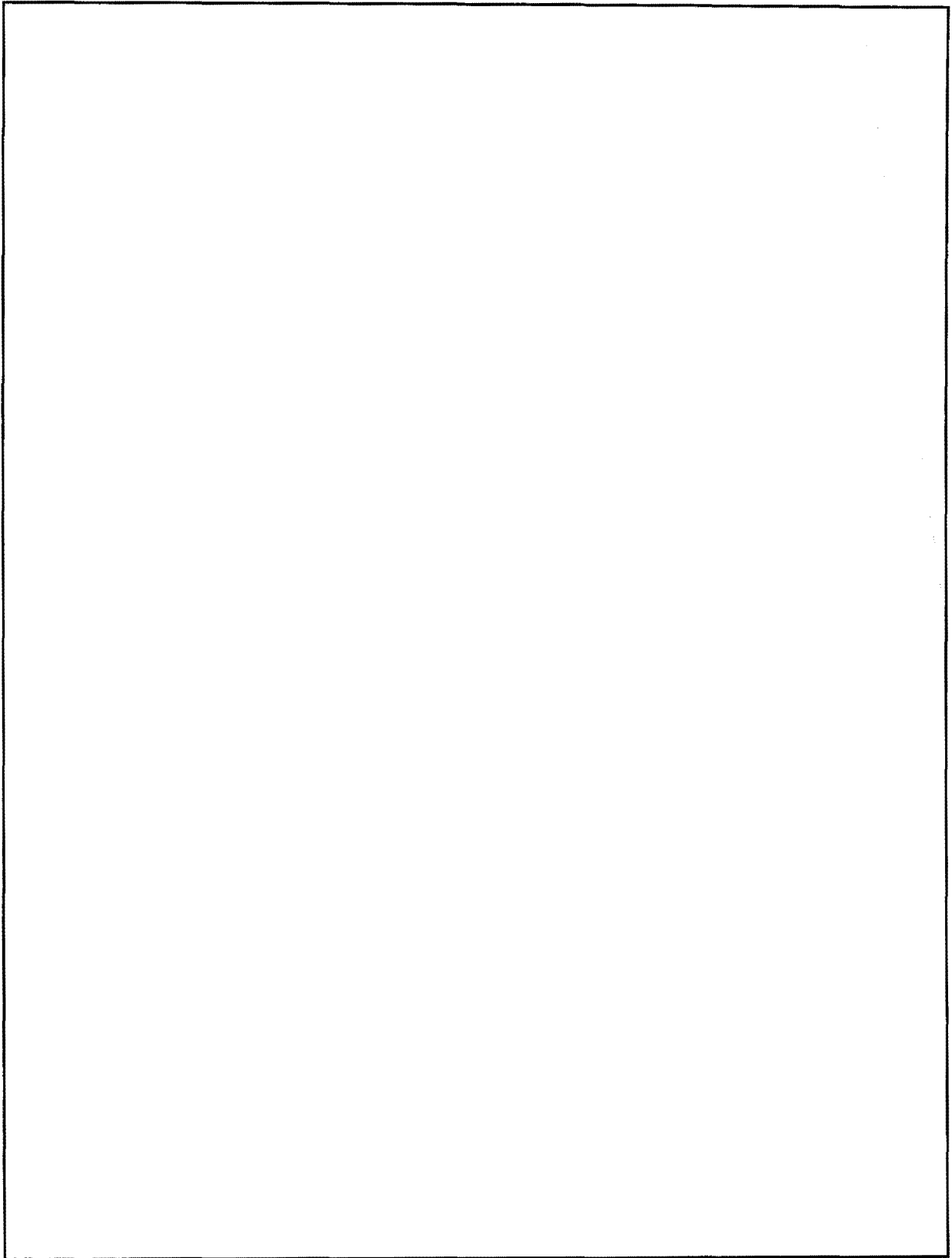
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



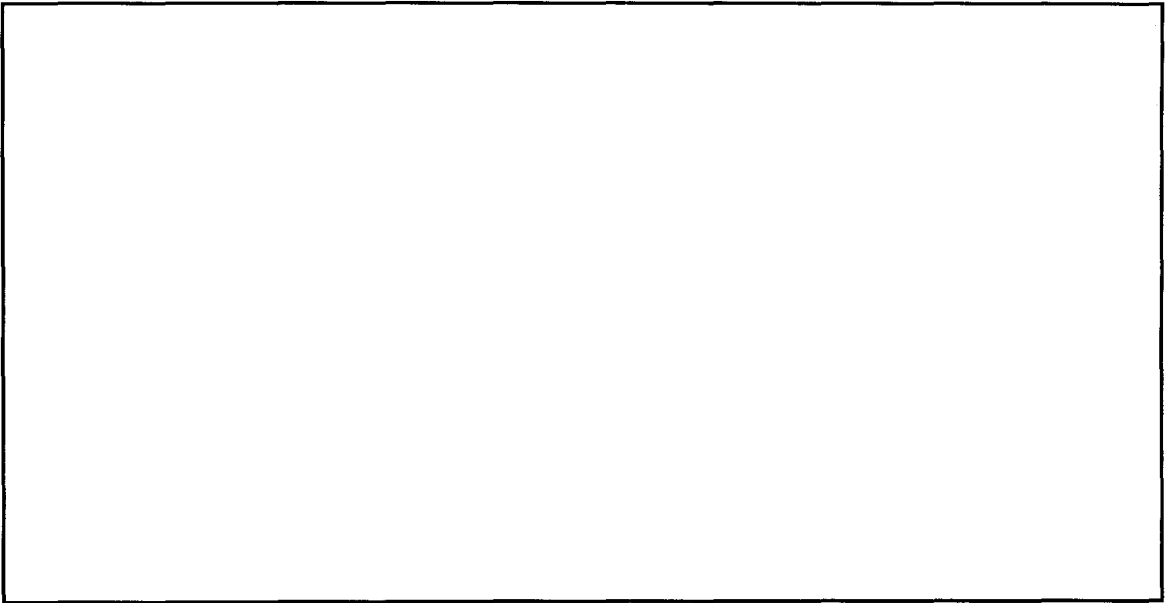
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

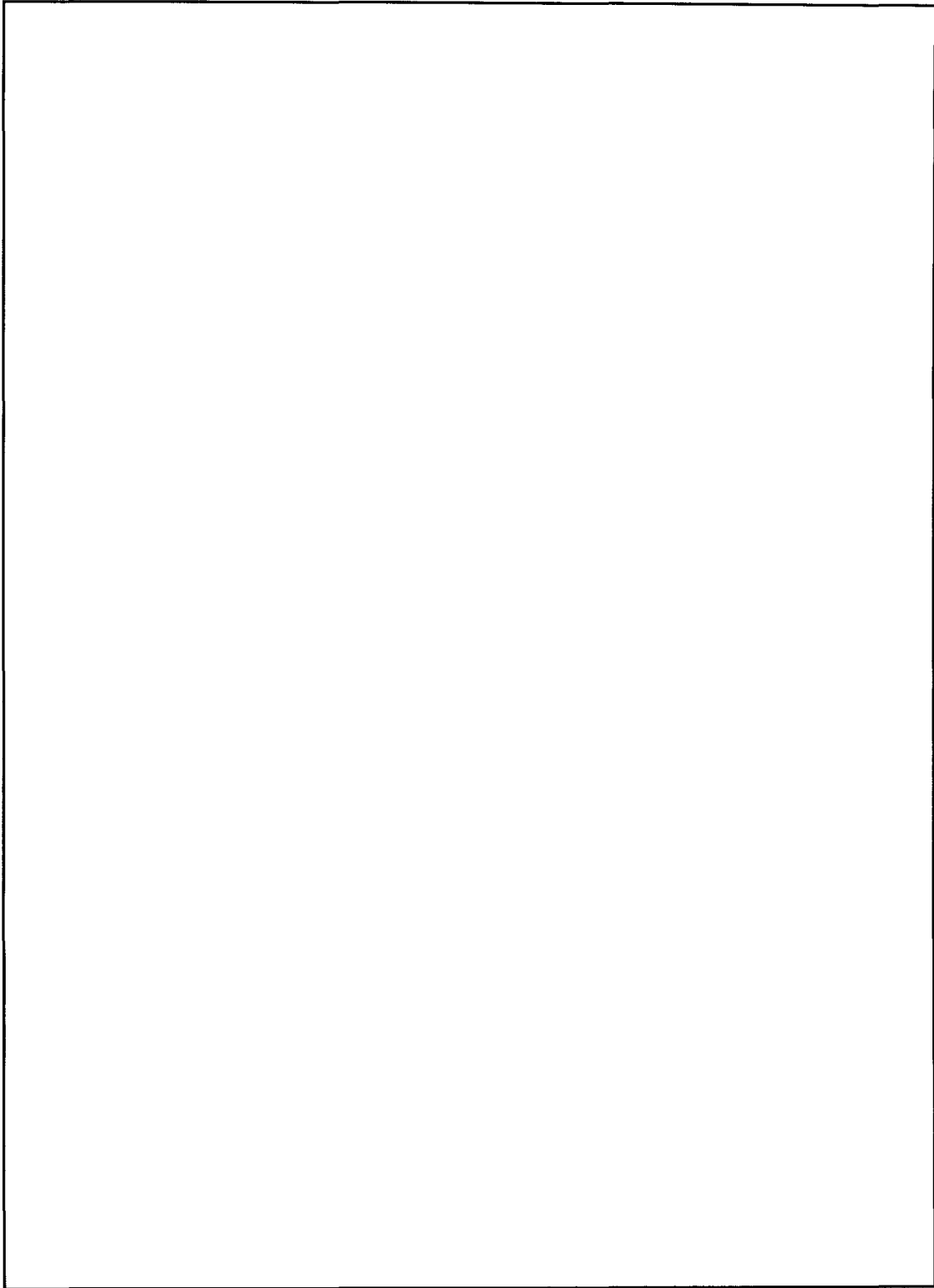
Maybe It's Related to the Phase of the Moon

P.L. 86-36
EO 1.4.(c)

This story of the dissection of a callsign system proves the validity of that old saw "Many a true word is spoken in jest." It proves a lot of other things, too - among them that it ill behooves the cryptanalyst to dismiss the word spoken in jest too quickly. But you may ask what a cryptanalyst is doing "dissecting" a callsign system in the first place - isn't that a job for a traffic analyst? Well, in case some of us haven't yet learned the lesson that you can't really draw a line between the work of the cryptanalyst, the traffic analyst, and the linguist, this story provides a bit more proof of that, too.

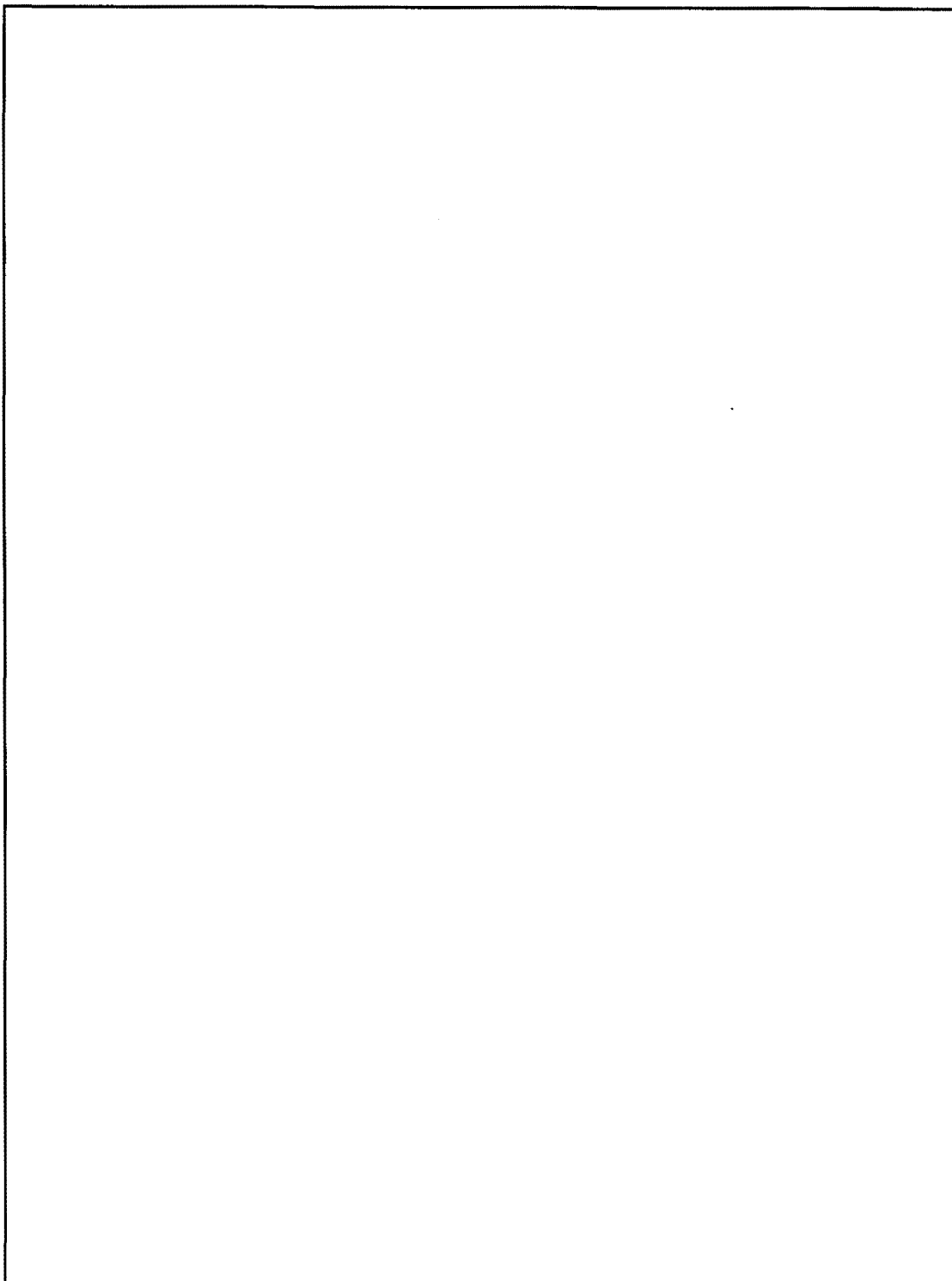
The reader has probably guessed by now that the title of this piece was the "true word spoken in jest." But it wasn't really spoken entirely in jest, because we knew that many of the

~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



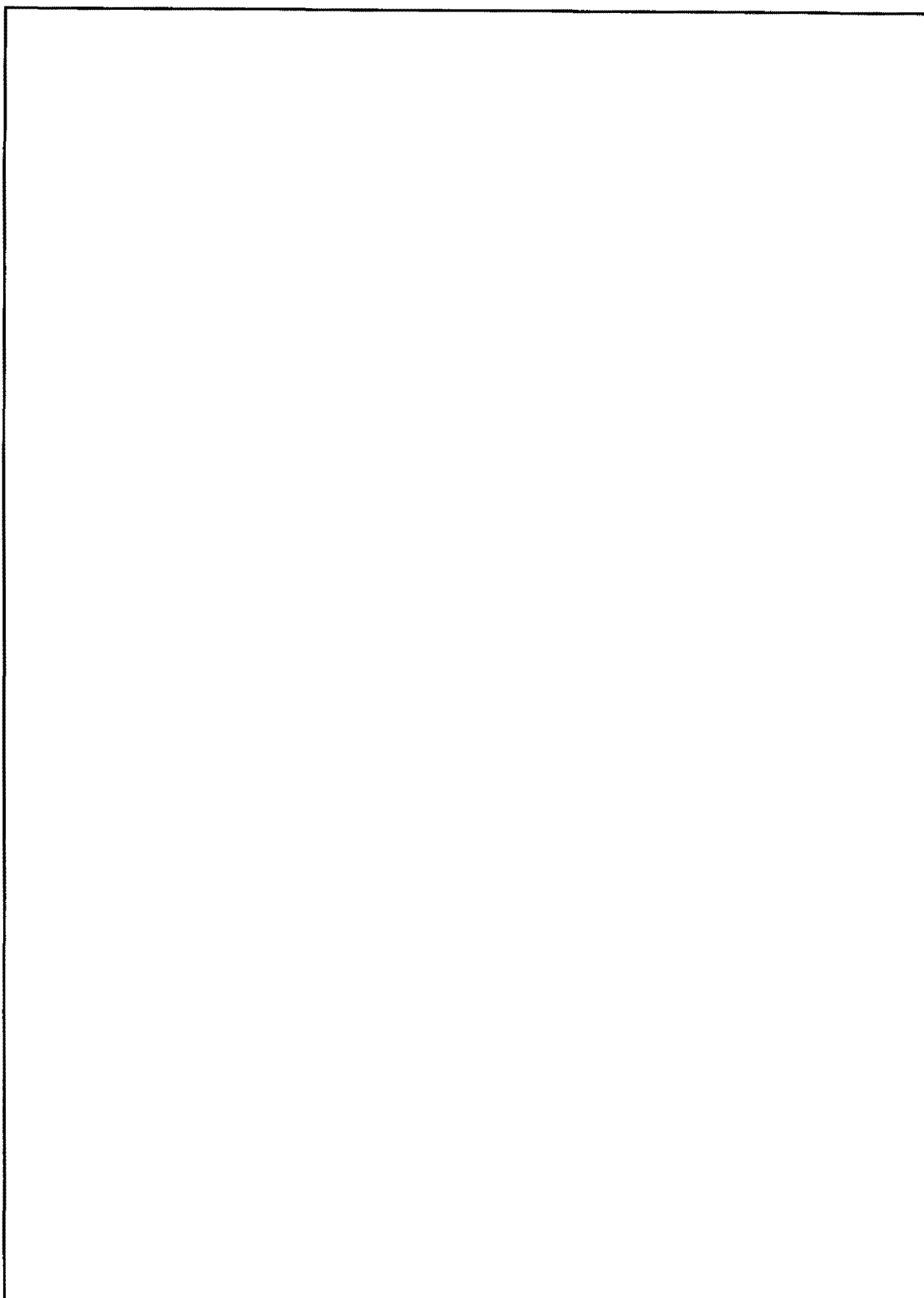
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



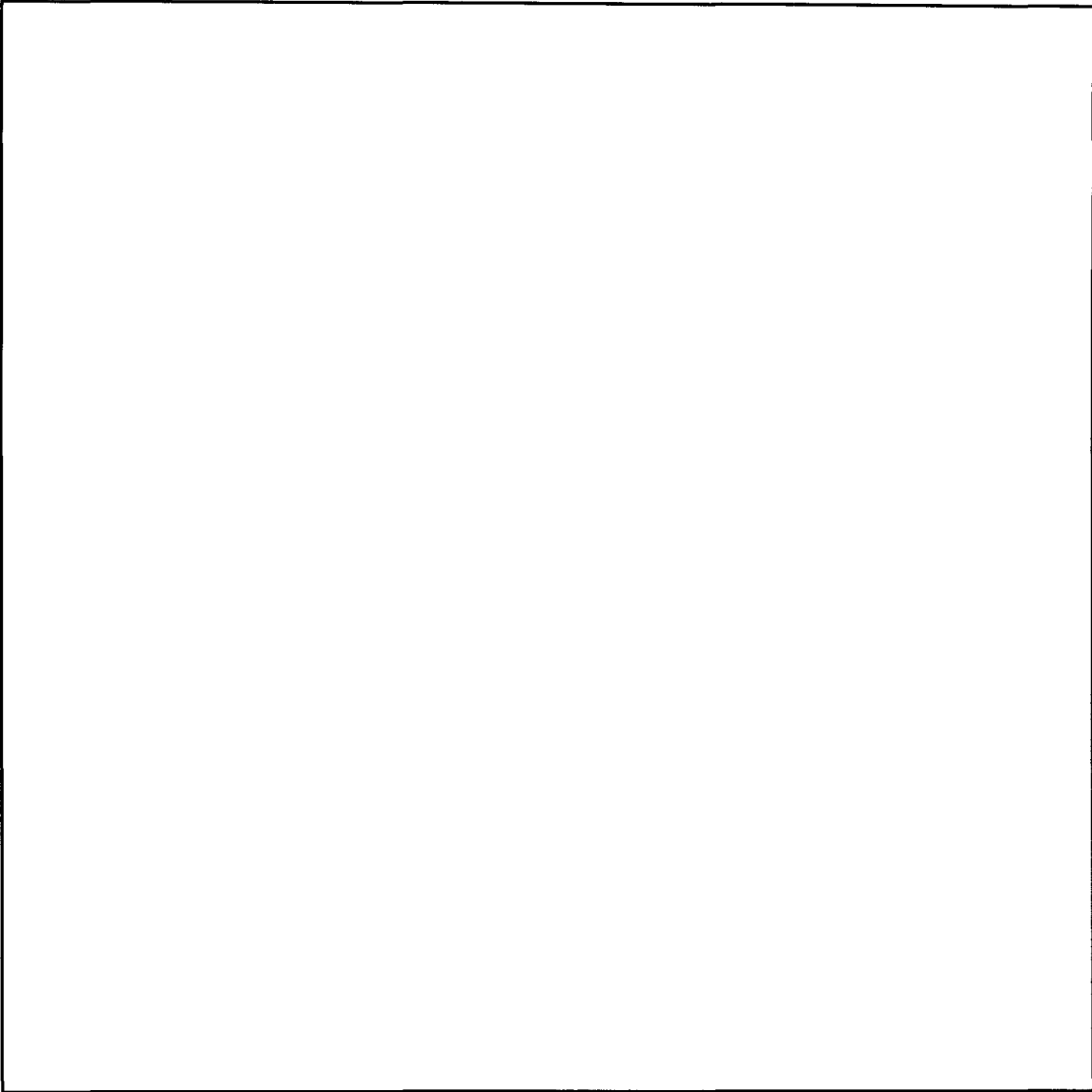
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~

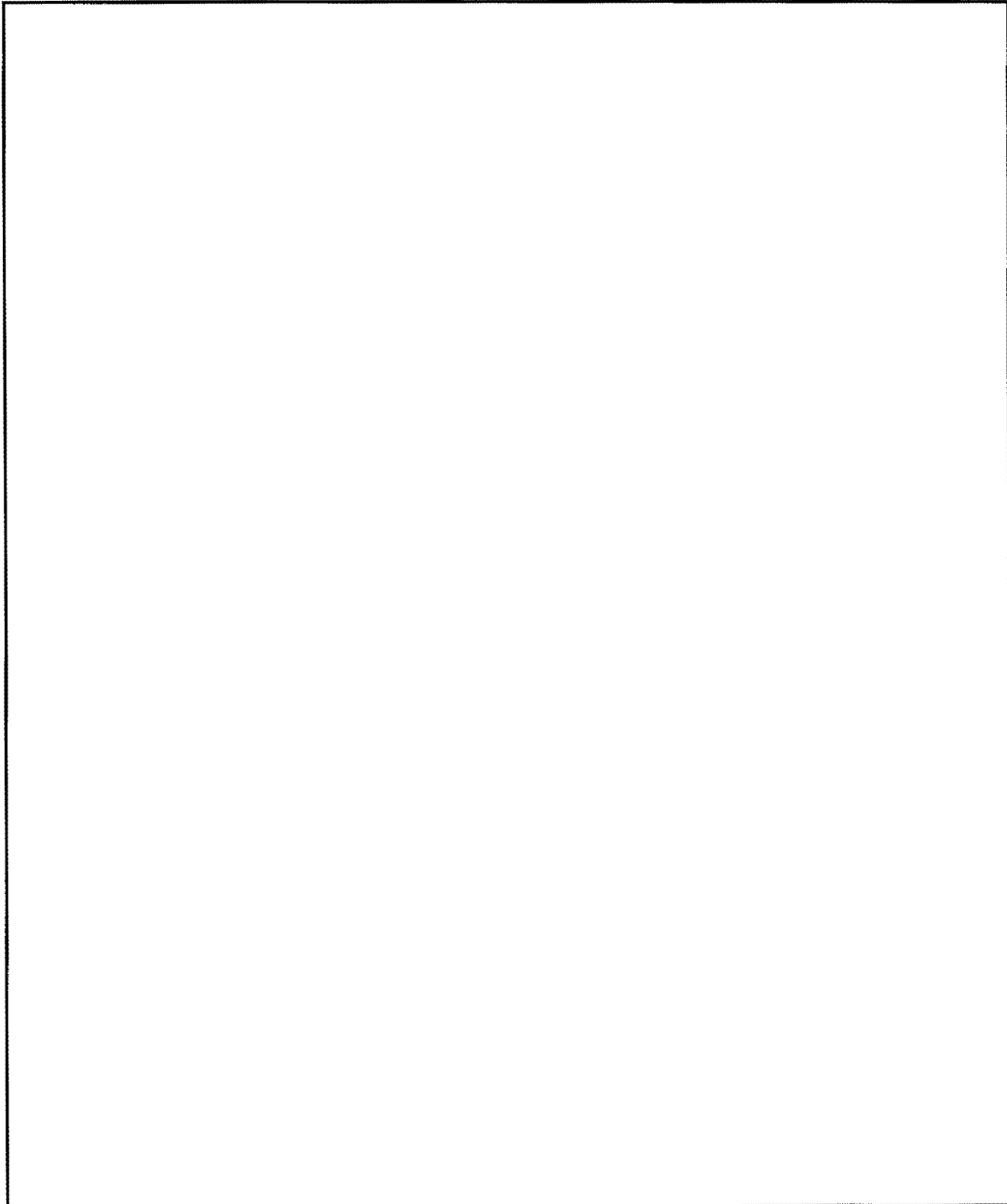


~~TOP SECRET UMBRA~~

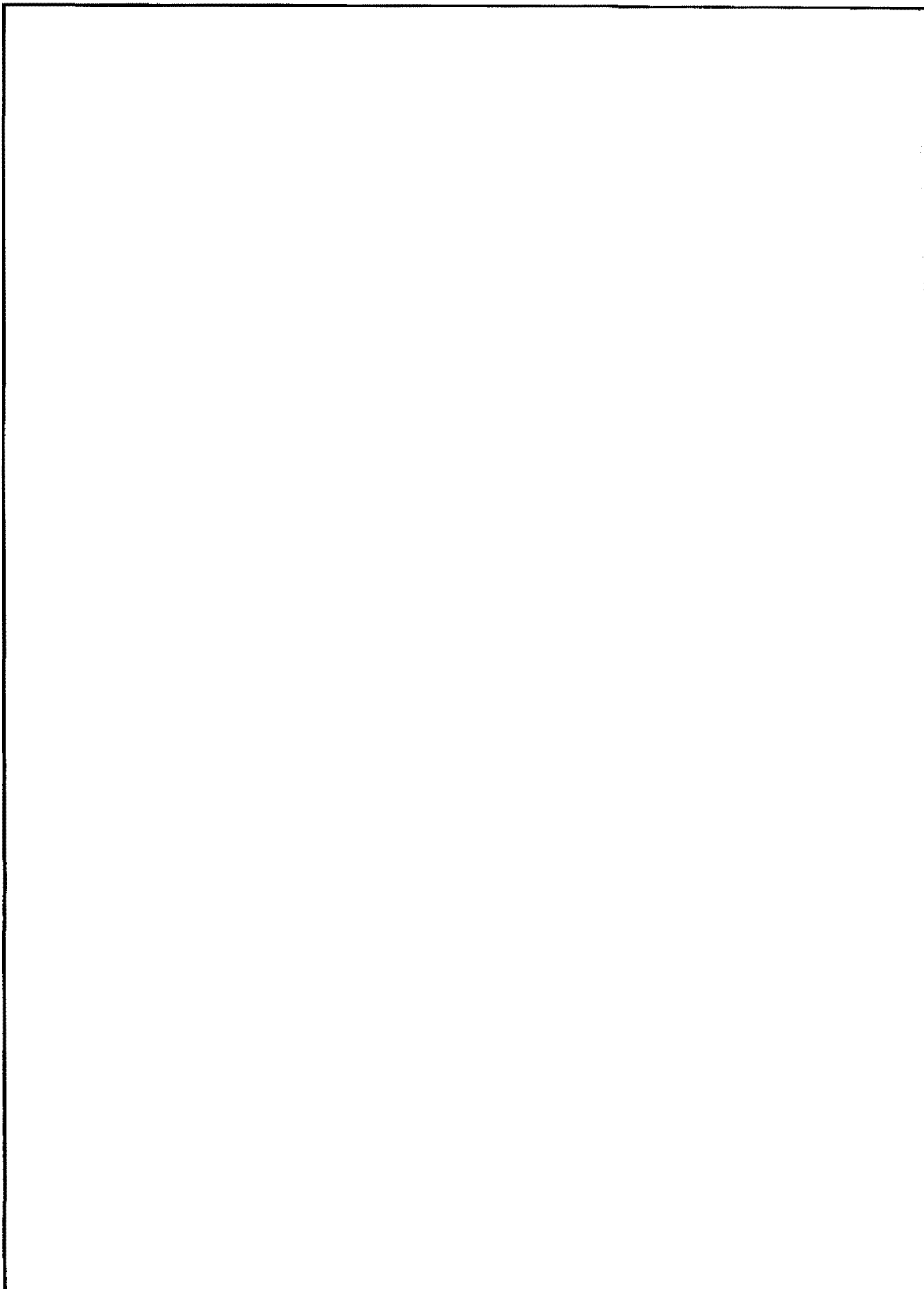


P. R. 86-336
EO 1.1.42 (c)

JAMES D. DELANEY

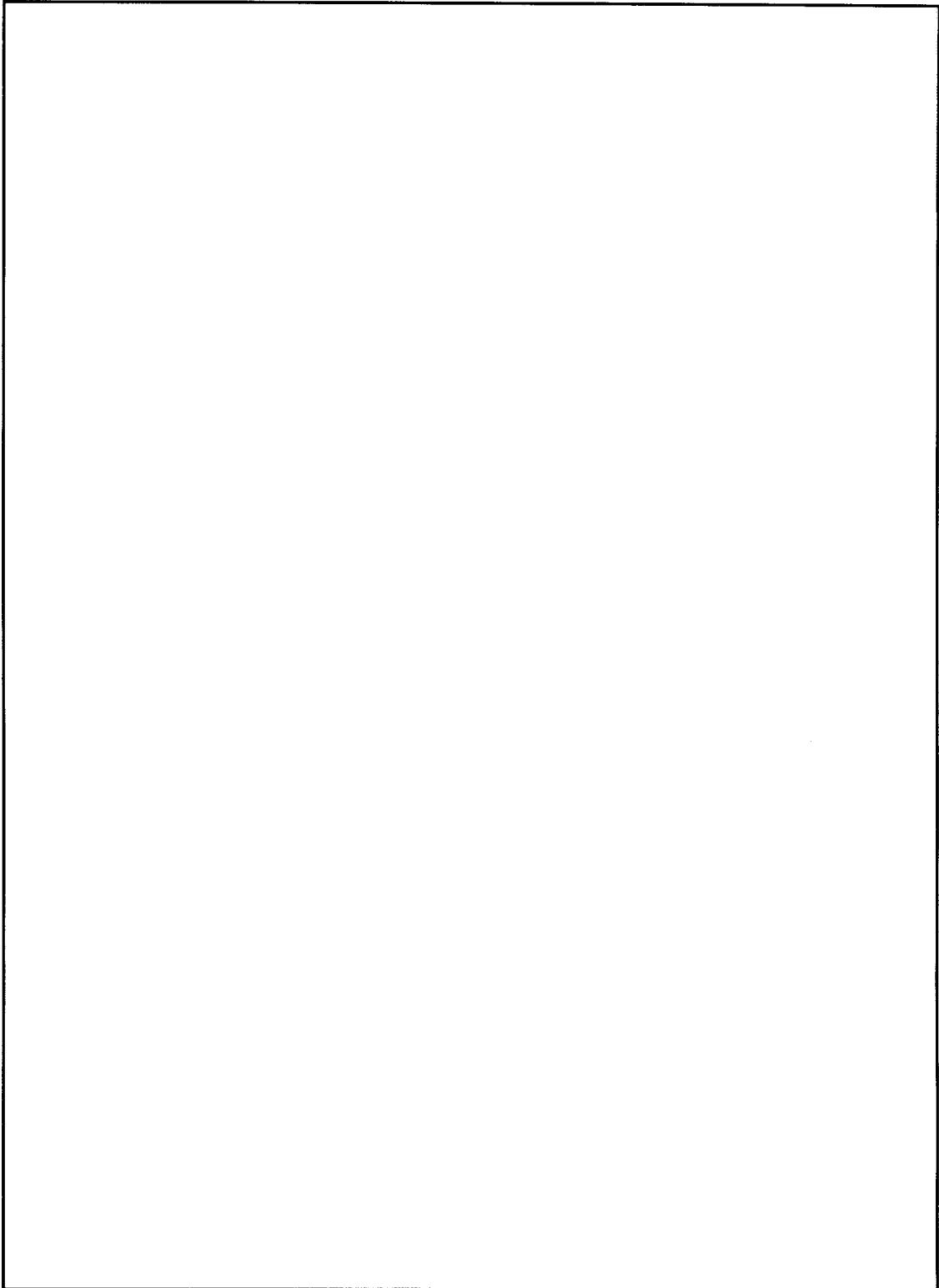


~~TOP SECRET UMBRA~~



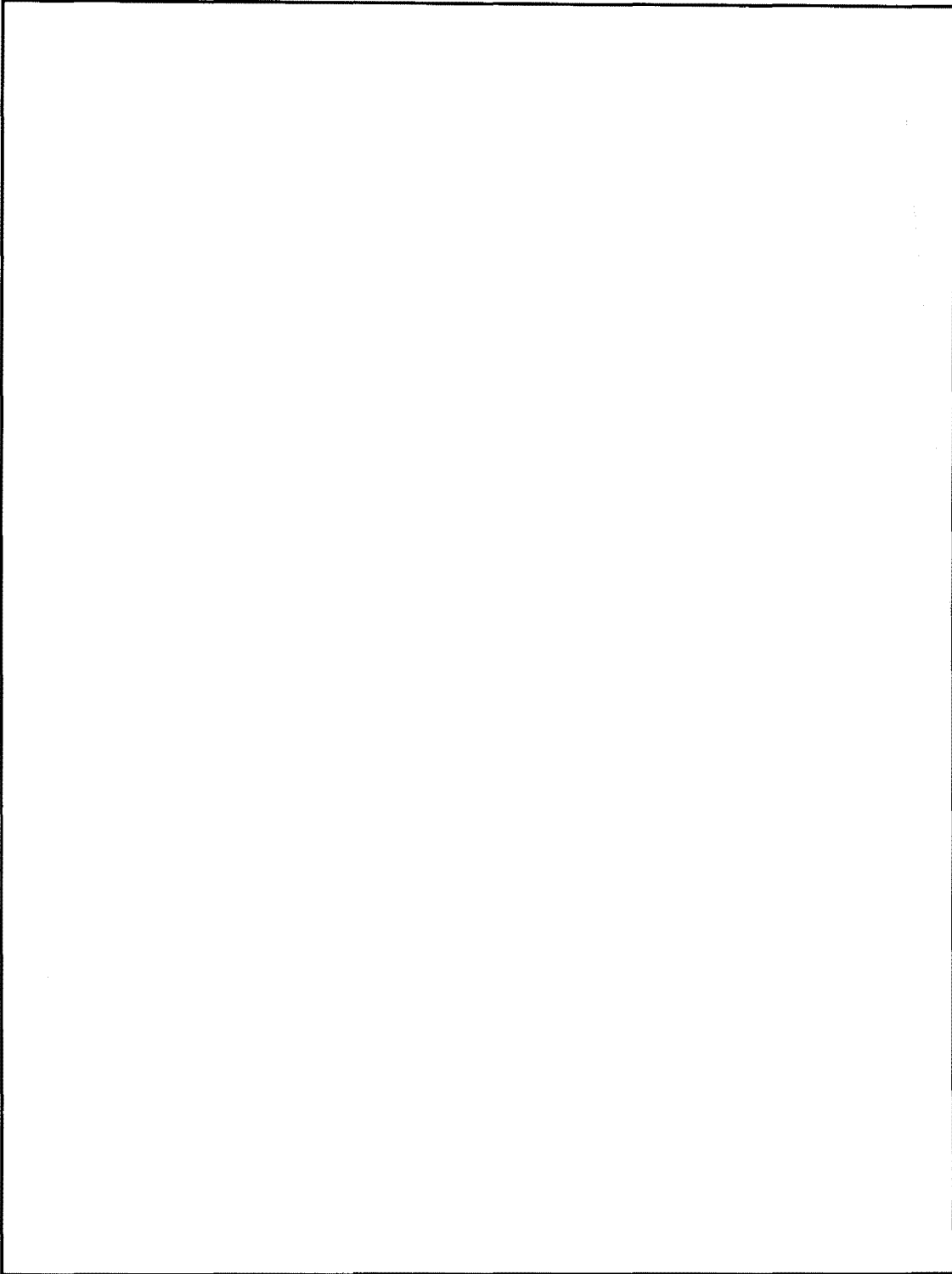
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



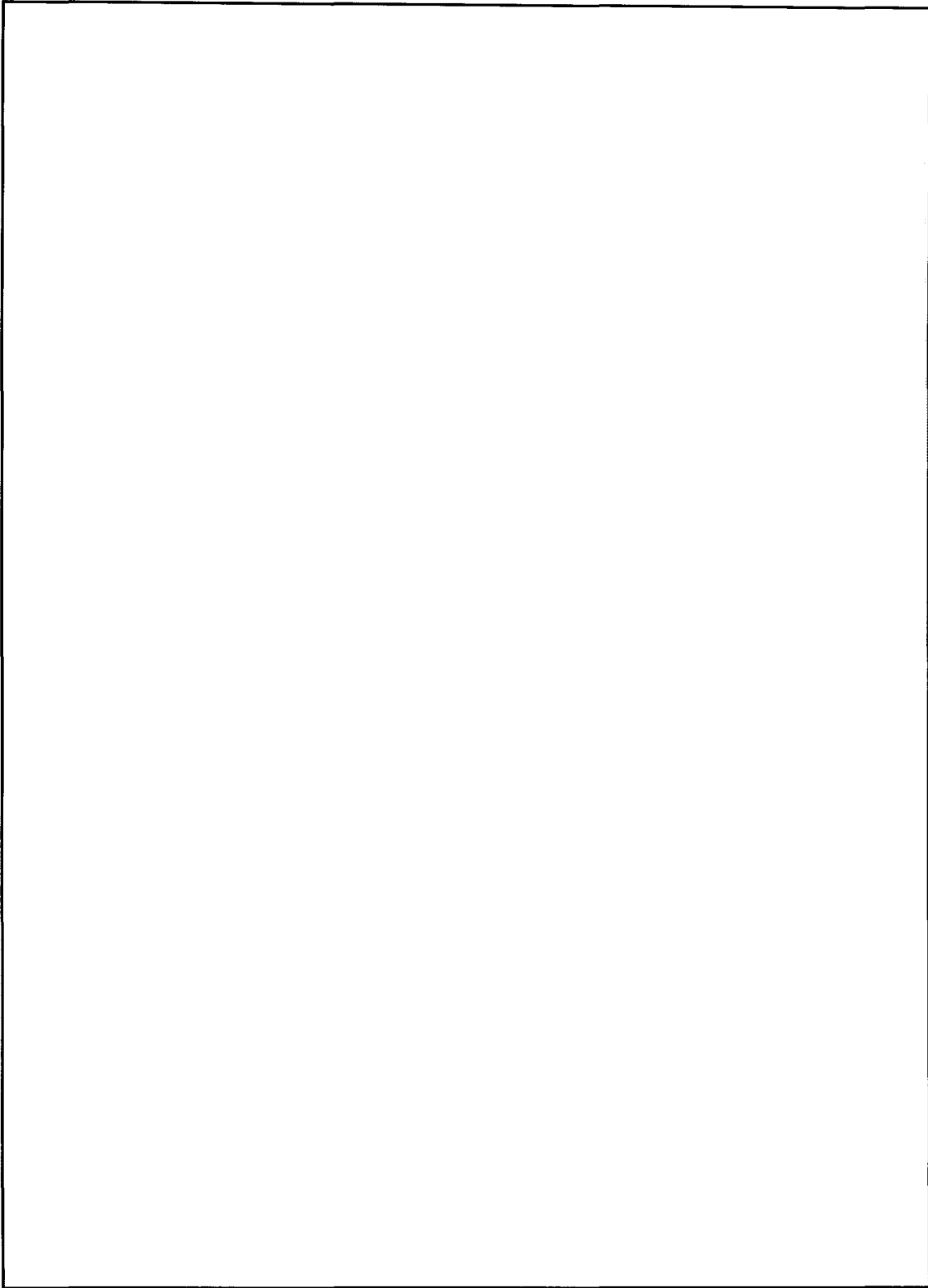
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



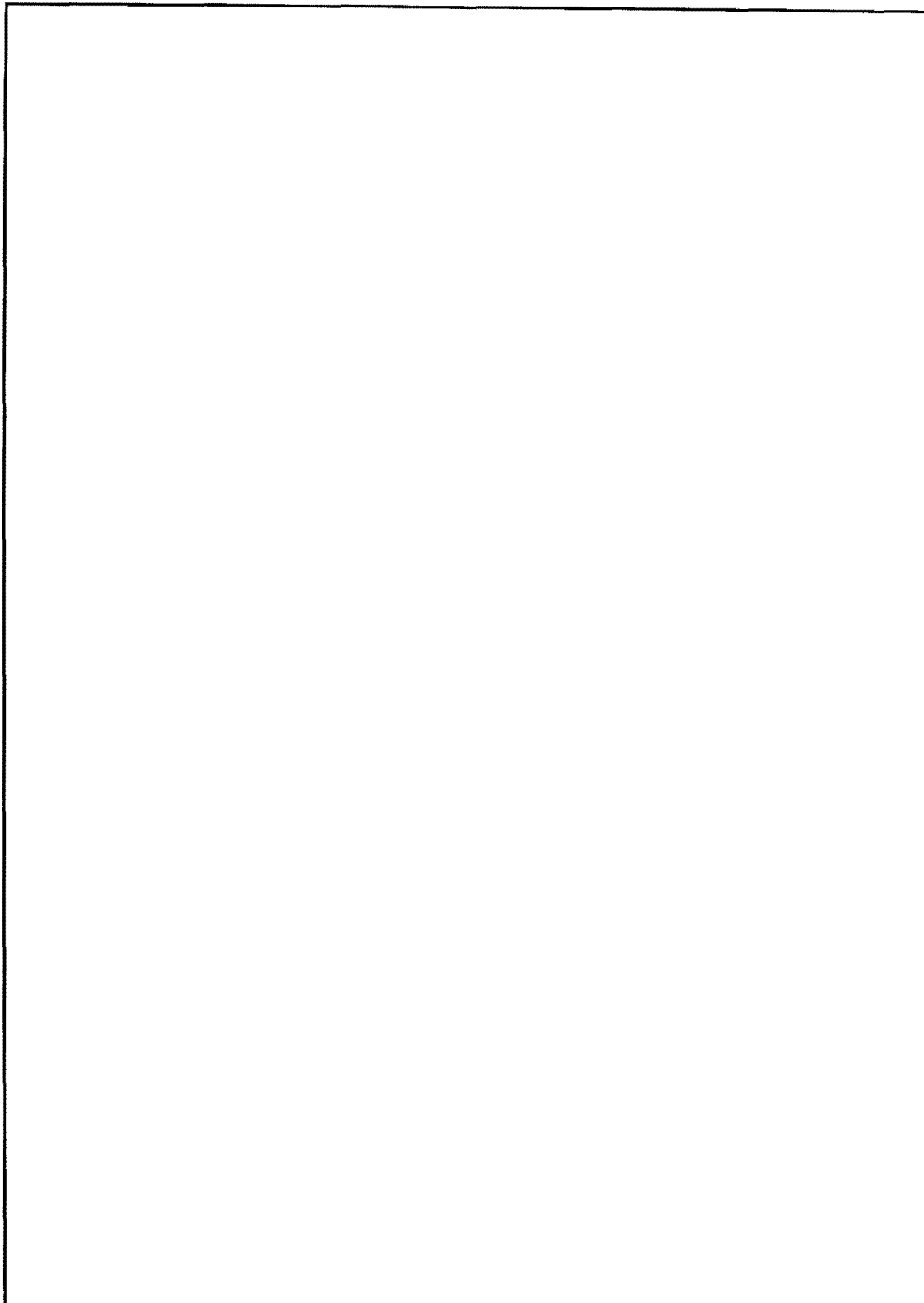
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



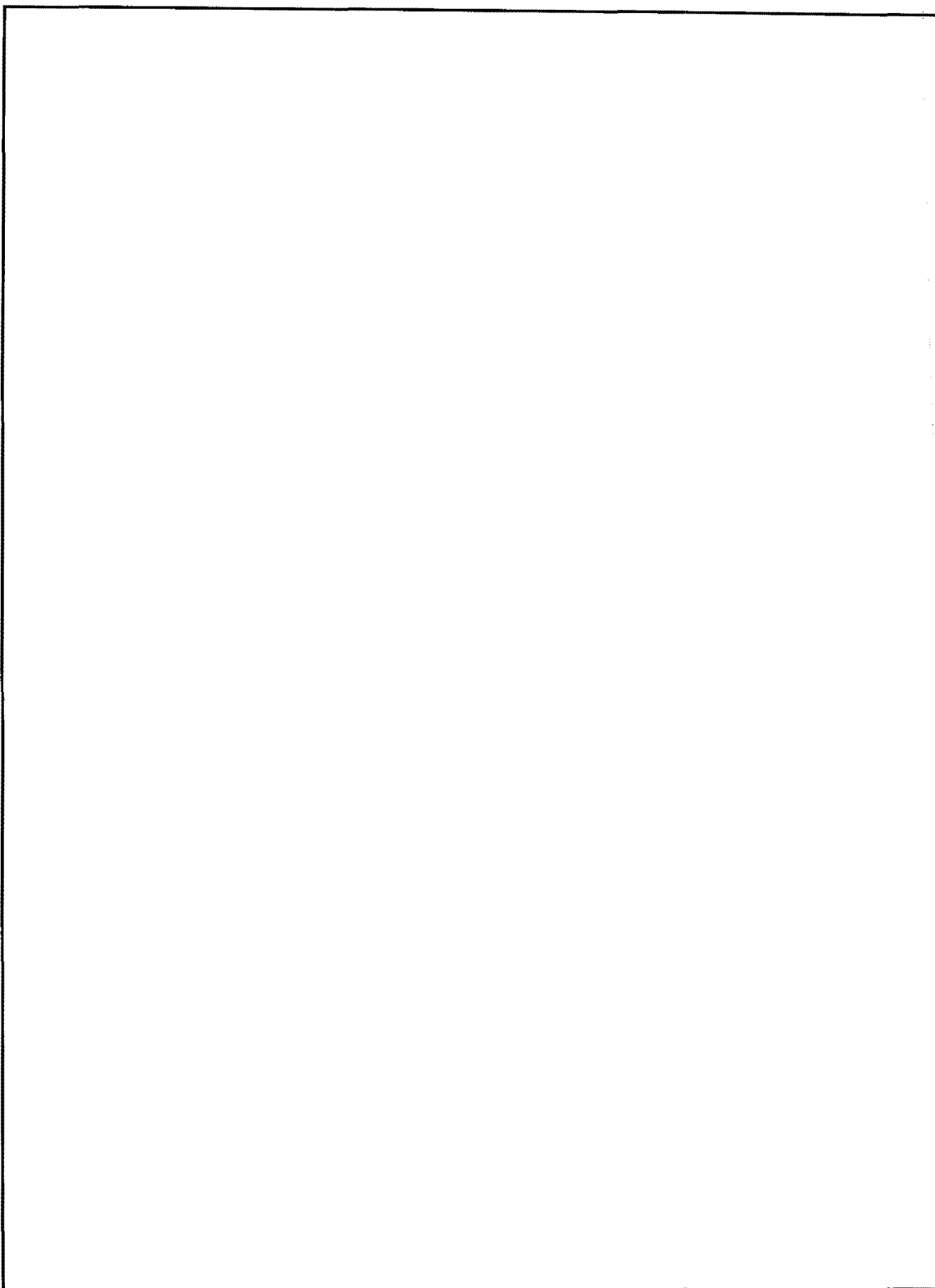
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



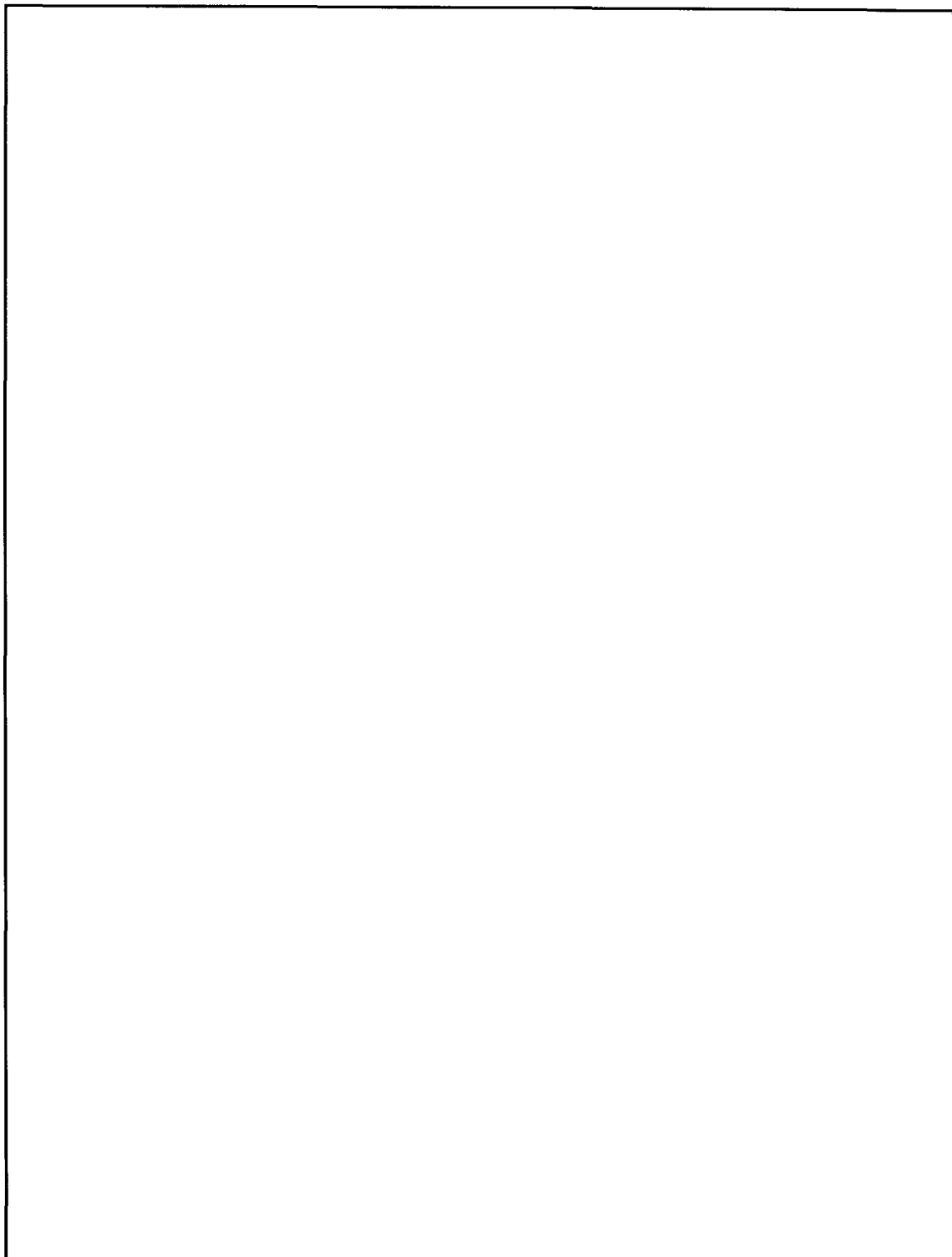
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

Acknowledgments

In recording the story of type [] foremost recognition belongs to Mr. C. Garofalo (P1), who so capably directed the exploitation effort (as on every other [] system) during the first year. Recognition also belongs to the many other dedicated members of the [] Task Group, including analysts from P1, A61, and the former A2. I took over the task group at the end of the first year, after the initial trauma was past and the system was well on its way to recovery.

~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

Area Studies and Their Place in Traffic Analysis

P.L. 86-36

The area study of a target country and its foreign influences provides essential support to the traffic analyst in developing a knowledge of existing communications and in selecting methods for the exploitation of these communications. Area studies include a broad and diversified range of factors influential in communications development.

Initially one should become familiar with the physical terrain of a country. If there an outlet to the sea? If so, what communications serve the area? If not, is the country dependent upon another area to import and export material? Is transportation limited to natural means such as rivers and narrow roads, or are there adequate transportation facilities by which movement within the country can be accomplished? Is the area mountainous, thereby interrupting long-range communication, or is the land basically flat, aiding in long-range contact? What is the climate like? Is it extremely hot or cold? Is the weather a factor in the method used for communications, or does it have little effect?

In Africa an important developmental element is the European colonial power that ruled the area prior to independence. Each power imported its own communications equipment and trained its subjects in its own methods. Whether the colonial power was Britain, France, Belgium, Portugal, or Spain made a considerable difference in the procedures introduced and the level of sophistication achieved. The influence of the colonial power can be seen not only in the communications but in many other aspects of the developing countries. Of significance also is the present-day rapprochement between the independent country and its colonizer, in the type of aid agreements and the dependence of the new country.

The level of the communications is also dependent in part upon the education level of the technicians. Areas of low educational opportunities cannot be expected to employ sophisticated operating procedures. However, it has been noted that each year they become more advanced.

The administrative structure of a country is an especially valuable factor in the understanding and exploitation of the communications structure. If facilities are limited, it is to be expected that they will be placed where they are most needed. Frequently different organizations employ the same set of facilities for communication purposes. A normal pattern in Africa is to have the control station located in the capital city of the country with the outstations in the provincial capitals and any subordinate stations located within the area encompassed by the province. A knowledge of the administrative structure greatly facilitates the reconstruction of these networks. Along the same lines, a basic understanding of the organization of the police and military in each country is particularly useful.

Having acquired a basic understanding of the rudimentary background of a country, it is most helpful to have a knowledge of the attitudes of the people. This enables a traffic analyst to understand why various acts are performed and to anticipate the possible outcomes of any given events. To understand the people of a region, one must closely examine their historical background. What are their social patterns? How have they evolved and what part do these play in the daily life of the people? What are the religious beliefs and how strongly do they influence the people? Is the population, on the whole, stable, or are the people inclined to rise up easily against control? While these factors may not be readily discernible in countries of a high level of sophistication, they play an increasingly important role in the developing countries.

The political make-up of a country also is a gauge in the development of communications. Is the government popularly elected? Do tribal or party affiliations play important roles? Is the government a military dictatorship? How susceptible is the civilian government to a coup d'état by the military or some other factor? What is the party structure? Is the country ruled by a one-party government? Is the government corrupt, and what are the chances for subversive activity? If subversive elements exist in the country, what is their background? Are they communist-inspired, or is their *raison d'être* one that can be more easily remedied?

Bordering countries should be studied on the basis of not only their colonizer but also their relationship with the country of primary interest. If dissident elements exist, are they receiving support from a neighboring country? Is the country interested in the downfall of the government of its neighbor? Are there mutual defense pacts or customs unions between the areas? Often the relationship of a country with its neighbors is a determining factor for the structure of its communications. Is it necessary for the country to use transportation facilities of its neighbors? This aspect may involve the development of a communications link between the two countries.

The level of the country's economy and stability indicators should be studied. Are the large concerns operated by the nationals of the country or by foreign technicians? Do the companies have their own communications? Are they within the country only, or do they have a link with the colonizer? If foreign nationals operate the industry, is this a cause for unrest? What is the commercial importance of the country's products on the world market? Do fluctuating prices play an important role in the economic stability of the country? Does the country have important mineral resources so that it might be considered a prize in the East-West struggle? Is there no industry and is the economy dependent upon primary produce for its revenues?

Another area of importance in considering methods of exploitation is the language employed. In the African countries, one of the most secure mechanisms is the use of indigenous languages and dialects for sending messages. This is of importance to the traffic analyst in the area of station locations and military unit locations. An unknown language, like an unsolved cipher system, can place a handicap on the traffic analyst.

~~CONFIDENTIAL~~

Only a few of the many aspects of area studies have been discussed in the foregoing pages. However, it should be readily apparent that area studies play a vital role in traffic analysis.

~~CONFIDENTIAL~~

~~HANDLE VIA COMINT CHANNELS ONLY~~

The 1970s

The Reality of Communications Changes

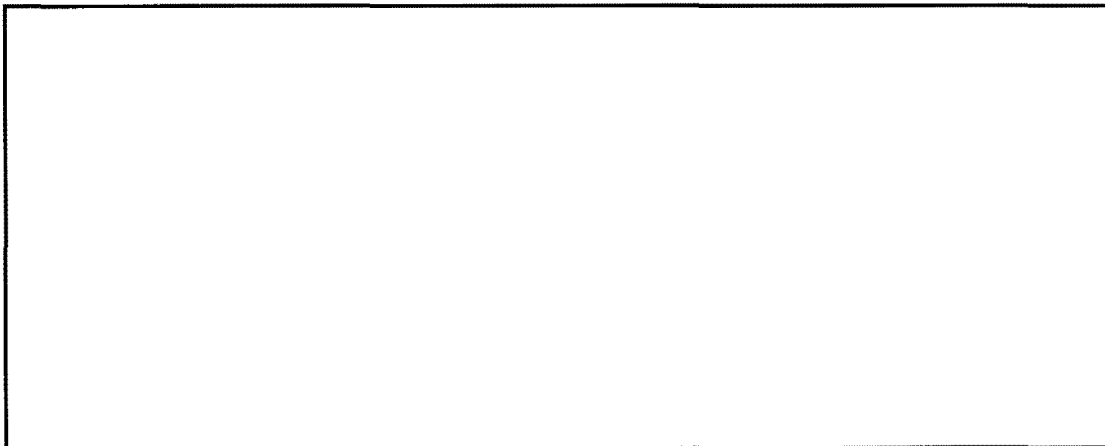


P.L. 86-36

All analysts and managers of analytic efforts must constantly face both the possibility of a communications change on their targets and the consequences of such a change. The term "communications change" frequently causes unnecessary apprehension – the change does not inevitably signal adverse consequences on target identification, maintenance of continuity, and production of SIGINT. Many changes (introduction of new callsigns, frequencies, etc.) on most targets are routine; they occur regularly and are only slight hindrances to the proficient analyst. On the other hand, some communications changes are *not* routine and do have an adverse effect on SIGINT production. They can result in reduction, or even total loss, of capability to identify and maintain continuity on target communications nets and the specific associated terminals. The latter type of communications change is the subject of this article.

* * * * *

Changes that might affect exploitation capability will vary greatly for different targets, depending on the extent of current exploitation and on the complexity of the newly introduced operational procedures. However, knowledge of the relationship between various communications features can greatly assist in prediction of future operational usage. Some features that should be considered follow:



P. 868636
(E.O. 4.4)(c)

3. *Sudden versus gradual change:* Many changes (e.g., newly allocated frequencies) can be implemented immediately upon receipt. Other changes require "live" testing and extensive operator training and orientation. The following changes, for example, would probably require an extended period for implementation:

a. *Introduction of a more sophisticated mode of communications:* Equipment procurement is usually limited, and testing and training are required before the new system becomes operational.

b. *Use of a new Morse cut number system:* Operator training is obviously required prior to full implementation.

c. *Introduction of abbreviated language for chatter:* The nationwide replacement of English by abbreviated [] as the vehicle for operator chatter is an example of a change requiring extensive operator training.

P. E. 868636
E.O. 4.4(c)

Some indicators of an impending communications change are

1. Temporary extension of the normal period of use of existing SOI materials;
2. Limited testing of new procedures on existing links/nets or on supplementary communications;
3. Direct references in chatter to new procedures. Such references could consist of anything from a casual implication to a statement of the effective date and type of new SOI materials;
4. Trends toward standardization or diversification, whichever is applicable;
5. Use of or references to more sophisticated modes of communication.

Although the ability to predict impending communications changes is a distinct advantage, recovery of continuity on target communications is greatly expedited by contingency planning that defines actions to be taken following introduction of new SOI materials. Contingency planning in preparation for subsequent analytic recovery must be realistic and flexible. Consideration should be given to the following factors:

1. *Timely field station reporting of deviations from the norm:* As the mission of most collection sites is limited in scope, this reporting permits higher echelon to make an early assessment of the overall extent of the communications change, to advise all elements concerned, and to issue necessary instructions.
2. *Target recognition/identification:* Even though such things as callsign and frequency usage have changed, the best source of target recognition/identification is the operator who has copied the target in the past and who will probably recognize it in the future. Operator identifications should be considered valid unless disproved. These identifications should be provided, in a format usable for traffic identification, to other field sites that are tasked with similar targets and that are likewise encountering difficulty in isolation and identification of mission targets. Thus, time will not be wasted in copying communications that are another site's mission.
3. *Establishing procedures for early continuous follow-up collection on potentially mission-associated communications:* Although these communications may not be

identified beyond nationality, establishing procedures for early collection will prove most advantageous.

4. *Determining possible methods of attack as a means of associating homogeneous intercept and performing follow-on analysis:* In making this determination, we must ask, "What would we do if the old tried and proven analytic techniques and aids were no longer available?" A definitive answer to this question will probably not be found, but alternate approaches can be devised. For example, if callsigns cannot be exploited, related intercept can often be associated on the basis of [] features. These features may, therefore, need examination very early after a communications change.

P.E. 868636
E.O. 4.4(c)

5. *Once the possible methods of attack have been determined, developing detailed procedures for quick implementation:* These procedures include issuing instructions to be followed in the event of a communications change, outlining processing (preferably in conjunction with a flow chart), and devising the machine software that would be needed for machine processing. Processing of data after an extensive communications change does not require completely new procedures, although some alternation or expansion of existing standard procedures will probably be necessary. Maximum retention of established procedures, which are already well known to all operating elements, will cause minimum confusion following a communications change and will aid in early recovery.

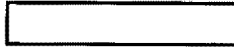
6. *Maintaining continuous documentation on all special processing or analytic actions taken and the type, extent, and data of actual changes in target SOI:* This documentation will aid in keeping all elements currently informed and in preparing for later SOI changes.

* * * * *

If this article succeeds in stimulating more realistic planning for future communications changes, deterioration of SIGINT production after such changes will be minimal.

A Note about Organizing T/A Problems

P.L. 86-36



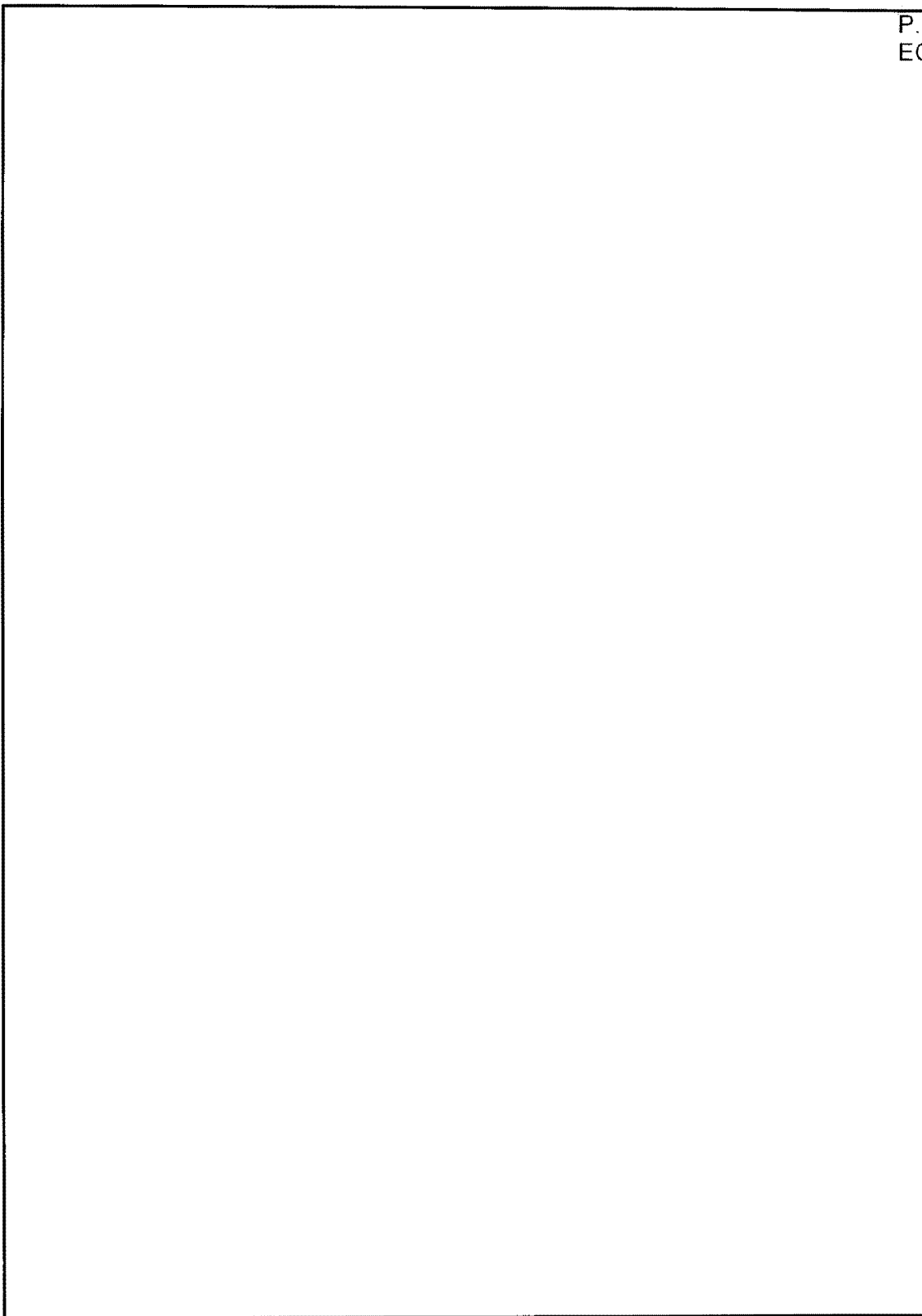
The ancients (Greeks, I think) believed that there was a central and fundamental Truth that underlay all worldly things and that man's goal should be the discovery of that Truth. In my salad days, as I wrestled with questions about how to organize a T/A effort, it seemed to me that there ought to be some basic fundamental T/A System; given that System, it ought to be possible to organize any T/A problem correctly.

As my arteries have begun to harden, my beliefs about these questions have shifted. It now seems to me that the most effective organizations (T/A and others) have been those that concentrated less upon adherence to some discovered Truth and more on defense against catastrophic failures. Consider all of the possibilities (in order of likelihood, if you wish) that could wipe out the average T/A shop: many-fold increase in traffic, major comm change, loss of key personnel, comms outage, destruction of database, etc. Organizations that can survive are those designed to adapt and adjust – to roll with the punch. Those that can't adjust go quickly into imbalance and overload, and major surgery (reorganization and restaffing) is necessary.

The principal difference between these two viewpoints is in their levels of complexity. The central Truth idea is really a two-body problem: if you can measure where you think you are and where you think the center of Truth is, then you can form some notion about whether you are heading toward your goal or away from it. However, the catastrophe defense idea is more of a multibody problem – strengthening your defense against one possibility often weakens your possible response to others, and some balance is clearly necessary.

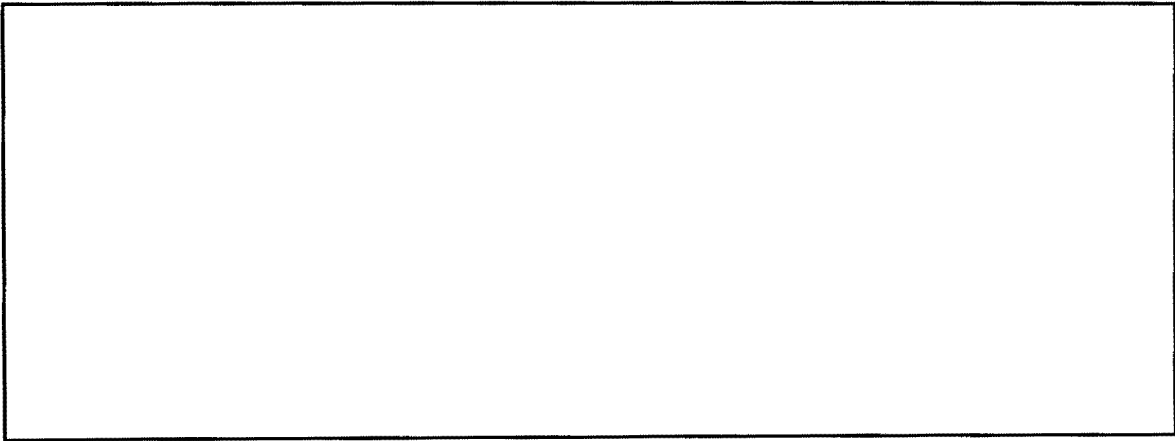
Unlimited defense against all possible contingencies is, of course, impossible without unlimited resources. The scariest part of this notion, however, is not that defenses must be limited; it is rather (as any computer programmer will tell you) that none of us is smart enough to think of all the bad things that can possibly happen. Surprises do come along, and it is, I suppose, the continuing possibility of the unexpected that keeps this business interesting.

P.L. 86-36
EO 1.4.(c)



~~TOP SECRET UMBRA~~

CRYPTOLOGIC QUARTERLY



~~TOP SECRET UMBRA~~

Traffic Analysis Mechanization Forum

OPENING ADDRESS BY MAJOR GENERAL JOHN E. MORRISON, ADP

Good Morning, Ladies and Gentlemen, and representatives of the Service Cryptologic Agencies.

I see a man in the back there, [REDACTED], who reminds me of a story I tell frequently – but Charlie is rarely around to hear it. I joined the cryptology business in February 1942, when the Signals Intelligence Service was holding up in the Munition Building on Constitution Avenue – Charlie was my crypt instructor, using the orange handbook written by Billy Friedman. Charlie and I are still around, hustling at the old stand. Charlie has stayed in the business consistently through the years. And, lest I travel under false colors, let me say that half of my career has been in communications and half in the signal intelligence business, starting as far back as 1942 in the SIS, which later became the Army Security Agency. P.L. 86-36

I'm glad to welcome the SCA [Service Cryptologic Agency] people here – I see some old friends from the field, old friends from PROD, and I see number of people I do not recognize, who, I assume, are from the PROD organization.

I've looked at your three-day schedule; it's an exciting schedule, but I don't know how you'll ever get through it. But from a topical viewpoint there is a great deal of interest in this particular schedule. It will provide us all a common insight, I think, into the kinds of problems we are confronted with today, and most particularly how do we get a bigger bang out of the traffic analysis discipline?

Now we all know that CA is, I guess, the queen of the disciplines and has been for some time. It carries with it a mystique that permits its acceptability, its respectability, in almost any circle – even in uninformed circles. Also when you get into telemetry analysis, people understand that; and signals analysis, people understand that; but this TA business – what can you get out of TA?

What do you get out of TA? We've made a living out of TA!

We've done an awful lot of magnificent things in an ongoing war – a real war in Southeast Asia! That's what we get out of TA! We've made a living out of it.

We handle a tremendous volume of information in the computer center, and I guess because of the diagnostic problems, and the probability problems, and all those things that go on in connection with cryptanalysis, one might say that the lion's share of the electronic data processing capabilities in the basement of this building is devoted to cryptanalysis. When I'm making a "beg bug" lecture on the subject of what we do in PROD – I usually "whomp up" this comparison – we have general purpose computers and we have special purpose computers, and we use our special-purpose computers 90 percent of the time in

~~TOP SECRET~~

cryptanalysis, and we use our general-purpose computers 60 percent of the time in cryptanalysis.

The point is that we spend a lot of time on CA, and the bulk of the residual, the residue of capability, is devoted to the TA discipline – a discipline that has worked so well for us, yet we've got to get more out of it. We must be more imaginative and we must mechanize. In these days with the kind of peril we live in – with all kinds of synchronous crypto devices now available, people wring their hands and express great concern that pretty soon all external info is going to be denied and won't that be a shame. And we're going to be out of the TA business. And these are the same kind of people who said in 1955 "won't it be a shame when all the manual Morse is going in ten years" – I just don't believe that TA people are going to be put out of business. I think we have yet to realize our maximum capability, and it's the challenge of attempting to realize an enhanced capability that I leave with you today in this forum. I ask you to share with each other a common understanding of the importance of TA as it evolves, as we reflect upon it, as the various speakers who have been selected to talk to you provide you with their appreciation. It's going to be tough to cover your agenda, but there's no doubt in my mind that at the end of the three days there will be an appreciation – a real feel for TA as we look at it today and a challenge that must be perceived. I hope that we are energized in our thinking, and I hope we leave here with a new dedication to do something more by way of mechanizing and getting a bigger bang out of that particular discipline from which we have gotten so much pay dirt in the past and that promises to do even more for us in the future.

I welcome you all again – I'm delighted that you were able to take the time to join each other here. I think you do us all valuable service.

Again to our SCA conferees – glad to have you aboard. George, without further ado I'm going to turn the meeting over to you. Thank you – Good to see you all.

~~TOP SECRET~~

~~HANDLE VIA COMINT CHANNELS ONLY~~

Introduction to Traffic Analysis Mechanization

ROBERT S. BENJAMIN

INTRODUCTION

Our purpose this morning is to set up a logical frame of reference for the many presentations on traffic analysis mechanization that will be following in this three-day symposium. We hope to provide an overview of traffic analysis and of mechanization as it relates to traffic analysis.

Mechanization in traffic analysis goes way back, well before computers. When we first got into the cryptologic business in cryptanalysis and later in traffic analysis, in early 1943, we encountered our first pile of "IBM Runs," as they were called then, at Arlington Hall Station. When we examined them, we found that the large pile of prints consisted of headings and preambles of Japanese Army, Navy, and Air Force traffic, sorted by various parts of the message heading. Most analysts were not using the runs, but a few of us who tried to look up things found them immensely useful in net reconstruction and in analysis of routing systems.

I will be acting as an "observer" this morning, a spokesman for many of my colleagues to whom I am indebted for the ideas I will be setting forth.

THE GOAL IS INTELLIGENCE

Traffic analysis - all SIGINT, for that matter - must begin with intelligence requirements. The SIGINT process is sometimes shown as a cycle, somewhat like Sherman Kent's "Intelligence Cycle," which appeared in one of his books. In the case of SIGINT, it goes: *requirements* - leading to *collection* - to *analysis* - producing *intelligence information* - which becomes *intelligence* - leading to more or slightly different *requirements*, and so on. To put it another way, in traffic analysis, we react and respond to requirements from intelligence consumers and must never lose track of this fact. Despite the intrinsic interest of what we do, the goal is intelligence.

Recently, I was reading S.I. Hayakawa's book *Language in Thought and Action* in which he made a distinction between "maps" and "territories." His general point was that in human discourse and in semantics, humans sometimes become confused between the "real world" (or territories) and the way we represent the real world, using symbols, names, etc. (the "maps"). This has inspired the attached diagram, figure 1, as one way to visualize the traffic analysis process. In TA, we build models of other's communications, study these models, and draw inferences from them. This chart may help us to visualize the general process. In figure 1, the top half of the chart (above the line) represents the real world. The bottom half represents how we see it in SIGINT, which we have called the

"SIGINT Models." The top of our chart is the territory; the bottom half is the map. The left hand half of the chart represents the SIGINT target, and the right is communications. Consider for a moment – the target, whoever he is, whatever he is up to, exists somewhere in the world. He communicates because of *need*; his communications represent a model of his organization. We intercept his communications, reconstruct his nets, and have a model of his communications. Basing our study and observations upon these communications, we build a model of his activities and intentions and report these in SIGINT end product. The *goal* of traffic analysis – of all SIGINT, for that matter – is to produce valid intelligence about the target.

DEFINITIONS

What is traffic analysis? There have been many definitions over the years, and a rather widely used short one now current is this:

"Traffic analysis is that branch of cryptology which assembles information about communications networks by studying the elements of transmissions which are external to message texts."

One could go on to point out that the traffic analyst also may study some parts of a text itself – for example, discriminants, indicators; plaintext messages, short messages such as Air Defense "proforma" messages, and so on – and that he may make extensive use of non-

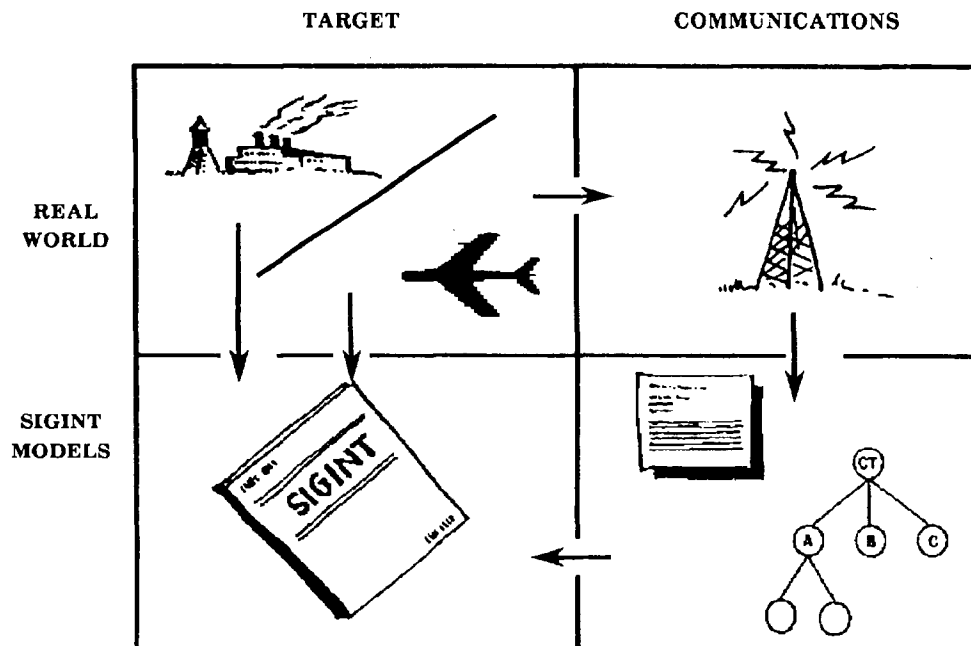


Fig. 1.

SIGINT information (collateral) to draw his conclusions. Traffic analysis has one or a combination of goals: to help SIGINT collection, to assist cryptanalysis, to improve intelligence information, and to assist the COMSEC effort.

DATA AND INFORMATION

In SIGINT, we are in the process of converting *data* ("that which is given") into information ("meaningful data"). These two terms are relative – in the total SIGINT process flow, this transformation takes place dozens and dozens of times between antenna (collection) and the point at which we issue end product.

Let us consider the traffic analytic process. As shown in the accompanying chart (see fig. 2), the analyst takes data and analyzes them. The analytic process consists essentially of making certain *observations* (counting, noting occurrences, measuring, etc.), coupled with making *inferences* based on the observations (extrapolating from the observations, drawing conclusions that are not explicit in the data, etc.).

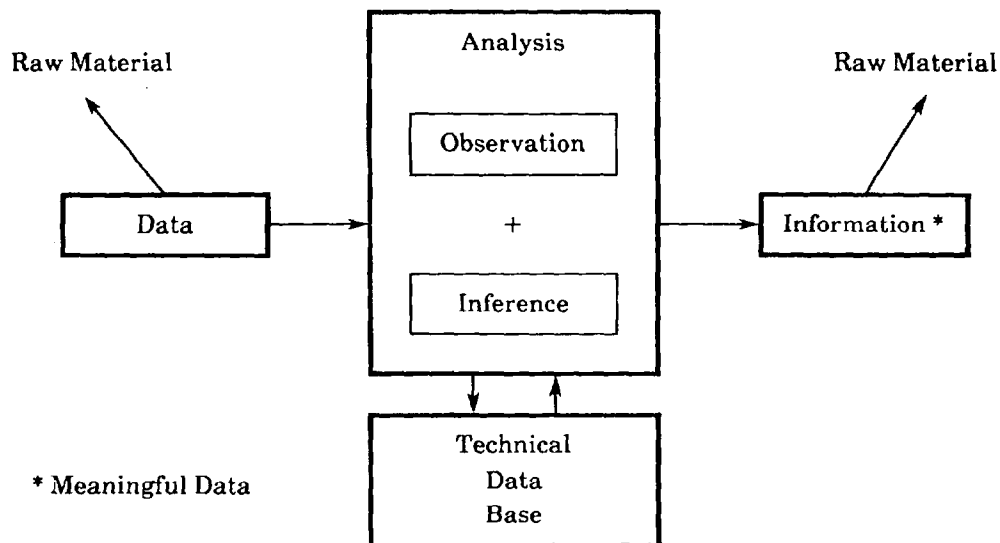


Fig. 2. The traffic analytic process

The analyst typically interacts with a technical database in this process of observing and drawing inferences. We are using the term "database" here to mean any files and records that an analyst has access to – card files, incoming hard copy of semiprocessed reports, ranging up to various computer-based files on magnetic tape, or even (in rare cases) files he can access on line with a computer.

Based on his analysis, the analyst produces information, leading to end product, which goes to consumers in the intelligence community. Much of the information may be fed back into the "technical database" to expand it and refine it.

Based on his analysis, the analyst produces information, leading to end product, which goes to consumers in the intelligence community. Much of the information may be fed back into the "technical database" to expand it and refine it.

THE TRAFFIC ANALYSIS CYCLE

Still another way of looking at traffic analysis to give us a better understanding of what happens is to consider what we shall refer to as the "traffic analysis cycle," patterned after the SIGINT cycle referred to earlier. Some persons have referred unkindly to this as the "Texan's map of traffic analysis." As you see in the accompanying chart (fig. 3), we

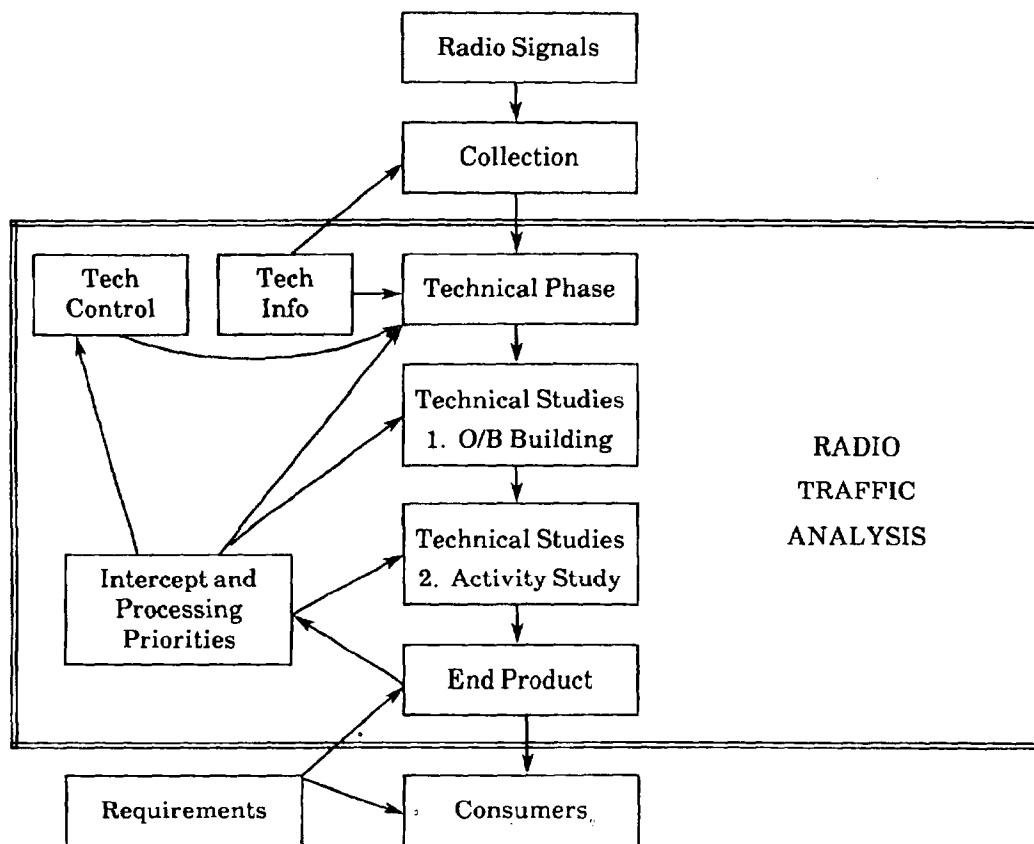


Fig. 3. The traffic analysis cycle

have set up a list of functional processes that constitute traffic analysis irrespective of echelon, things that go on between the antenna and the passing of product to consumers. In reaction to intelligence requirements, *radio signals* are collected, and the traffic analyst performs certain steps in a *technical phase* – things like traffic identification, net reconstruction, establishing net continuity, and drafting and issuing technical reports. These reports and cables provide a feedback flow of technical support to the intercept source. The next effort we will refer to as *technical studies (order of battle building)*. This consists of determining locations of radio stations and units, determining what organizations or types of units are involved, relationships between units, and the like.

Almost concurrently, the analyst carries on what we call here technical studies (activity study), looking into what is actually happening on the framework that is the order of battle. He looks for indications of things happening or about to happen (a whole talk could be given about the kinds of things that constitute SIGINT indicators); he studies activity generally; he establishes a formalized or general idea of what is normal; looks for phenomena or exceptions to the norms he has set; and he tries to interpret or explain these phenomena in terms of what is actually or probably happening to the target itself. We will return later to the idea of exceptions when we discuss mechanization. The analyst seeks to produce *end product* – I cannot emphasize enough that the goal of traffic analysis is producing SIGINT; all else is peripheral to that goal. The analyst should and often does produce product based upon his findings using non-SIGINT information as necessary, or he may furnish his findings to special research analysts, who may fuse the findings of several analysts with ELINT, cryptanalytic results, and with other non-SIGINT results to produce a comprehensive study.

Product is passed to consumers, completing the cycle. Feedback loops exist in our cycle, and our discussion is not complete unless we consider them. Requirements are passed from consumers to those in cryptology concerned with product matters. Based upon these considerations, all efforts represented by other parts of the traffic analysis cycle are constantly being readjusted – emphasized or deemphasized in terms of intercept and processing priorities. Technical control of the collection effort is being exercised in continuing response to these constantly adjusting priorities.

This, then, is the world of traffic analysis, working closely with other cryptologic specialists in the common goal of covering *data to information to intelligence*.

MECHANIZATION AND SYSTEMS

We would like to shift gears at this point and talk more about mechanization of traffic analysis, to establish a few logical bridges between the computer field and the traffic analysis field. First, a definition or so is in order. What is “mechanization”? It is generally agreed that *mechanization* means a process or system having an optimum mix of men and machines. The term *automation*, on the other hand, implies a system with no direct manual intervention, self-regulation within the system, and automatic “feedback”

loops. There are few examples of true "automation" in SIGINT (unless one takes a small piece of a system or a program); strictly speaking, by and large we are using much mechanization, little automation.

It may help to think of three types of "systems" that interact in any SIGINT process using machines (we will use the words "computer" and "machine" loosely and interchangeably). One must have *hardware systems* (the computers themselves, their tape drives, and communications linking them). There also are *software systems* (which, strictly speaking, are the programming systems that are available to a programmer when he begins to program for a particular application or problem). More important for us in this symposium, there are *cryptologic application systems* – programs and assemblages of equipments and programs that serve the traffic analyst, the special research analyst (formerly called the SIGINT research analyst), and the cryptanalyst. Consideration of these professions and of data banks that serve them leads us to a small digression in our train of thought.

OVERLAPS OF TA WITH OTHER SIGINT SPECIALTIES

Our cryptologic professions, and the techniques implied by them in SIGINT, complement each other to a remarkable extent. In fact, a person in any of our professions – i.e., Cryptanalyst, Traffic Analyst, Special Research Analyst, Signal Analyst, Collection Specialist, and so on – should know the rudiments of the other professions.

This overlap has some very interesting professional and cryptologic training applications. The reason I mention it today is that it has an important bearing on mechanization. We should not mechanize things in isolation without considering related problems. The cryptologic database interacts in dynamic ways. The need for interactions of data between large organizations – A, B, and G – and between analysts of different breeds, dictates applications systems designed to permit such interactions of the data, cross studies, and the like. When such matters are not considered, we may have two, three, four, or more databases containing essentially the same data in different, incompatible forms. The *standards* implications of what we are saying should be apparent, both at the *data element* level and at the format level. We shall have a few more words later on formats.

SOME TRAFFIC ANALYSIS MECHANIZATION IDEAS

At this point, I would like to toss out a number of ideas and terms. Most machine help to traffic analysis today still consists of *batch* processing, much like the so-called "IBM Runs" of Japanese traffic of almost thirty years ago that I mentioned earlier. We are getting more and more sophisticated, I suppose, but it is still "batch" processing – daily, weekly, monthly, etc. There is some *on-line processing* going on. Here, I mean processing where material is coming in on a signal line from the field, is going directly into a

computer without being handled manually, and some processing is being done to it before people see it. [] will be discussing some on-line applications of this type. Also, on-line can refer to on-line inquiries at a console station connected to a computer; our capabilities in that direction are growing slowly but surely. I do not mean to convey the impression that on-line processing is necessarily preferable to batch processing; each has its place and will continue to have its place, dependent upon the nature of each particular TA problem.

We referred earlier to data standards. The Department of Defense is pushing very hard in the area of data standards (and in hardware, software, and media standards, for that matter). We at NSA must consider these Defense-wide standards seriously, but because of the special problems we run into in cryptology, we are not able to use them "as is" very frequently. The comptroller and personnel organizations (who have more occasion to deal with DoD people and whose data banks interact to a greater degree) are able to use these DoD standards directly to a much greater extent. Without defining the terms rigorously, let me say that data standards are of three types of things: Data Elements (classes of data), Data Items (things that appear in these classes), and Data Codes (short abbreviated codes for either elements or items, to be used in a data bank or descriptions). Collectively, these are referred to as *Data Features*. Although some find even the idea of data standards distressing and distasteful, let me underline their importance. We simply cannot build large, useful data banks without seriously building and using standards. Dr. [] will be having some comments on the need for standards and the special problems that the overlapping and competing of national and international standards organizations create for NSA and cryptology.

In the business world, managers have found that they cannot possibly look at all the data relating to their particular concern – whether it is running a production organization, supervising a sales force, or whatever. To make it possible to control an operation effectively, a technique called *management by exception* is being applied effectively. Rather than having a huge pile of material presented to him in the morning for study, the manager using "exception" techniques will specify that he is interested, say, only in the information relating to the top three and the bottom three salesmen's records for the previous period; or he may set certain upper and lower thresholds, or "norms," and call for production and sales figures only for the departments and employees who exceeded or fell behind the norms. Computers are superbly equipped to provide reports answering such questions.

The same idea has always been used in manual traffic analysis, and we will be seeing sharply increased use of "analysis by exception" techniques by machine in the future. To use the technique, a traffic analyst studies his material, establishes norms for events and occurrences, and asks for a machine search to find deviations from these norms. Output provided him shows these deviations (or exceptions); he studies them and draws appropriate conclusions. He may have to readjust norms as time goes on, since what is "normal" changes in time. New norms would be reintroduced in the system for all subsequent processing until a new adjustment is necessary. It goes without saying that

exceptions to the norms are quite likely to contain the "goodies" that will result in end product of value to consumers – thresholds must be so designed that this will be the case.

Without getting too deeply into the subjects, I would like now to discuss *algorithmic* versus *heuristic* problem solving. An algorithm is any definite, specific procedure or formula that, when certain data are put into it, and the procedure is followed, yields certain answers that are always the same. Simple arithmetic problems, algebra, and so on, are examples. Algorithms need *not* be numerical – the machine world abounds with examples dealing with all kinds of data. Heuristic problems, on the other hand, are those that are loosely structured or are so new that approaches must be discovered. Heuristic methods may be "rule of thumb" approaches, may seem to use "intuition," and may be probabilistic (at least not deterministic); one does not necessarily get the same answer, or any answer, to a particular problem. Chess, checkers, and various board games fall into the class of heuristic problems; so does intelligence gathering, and especially so does traffic analysis. Whereas most problems of an algorithmic nature can be readily programmed to be handled by computers, heuristic approaches are much more difficult. Some progress has been made in recent years in preparing "heuristic programs" for computers, and it is hoped that there will be a spillover into our field. Some of the things that speakers will be discussing these next few days have "heuristic" features – watch for them.

THE TRAFFIC ANALYSIS MECHANIZATION PROCESS

I would like to take a look at traffic analysis mechanization in several ways. Essentially, in TA mechanization, we are manipulating *sets* of data. Today's youngsters who are learning about sets in school will probably be better traffic analysts than we were – at least they will have the benefit of some theory relating to ideas that we have been using but had not had formalized for us.

I would like to consider TA mechanization trends over the years in terms of the traditional INPUT - PROCESSING - OUTPUT breakdown. I shall mention some dates – please regard them as approximate. They are based largely on my recollections. Communications developments play a most important part throughout. What is really happening is that TA mechanization has gone global; not only are we eliminating or reducing key punching, carrying punched cards across the room, and so on, but are doing it on a worldwide scale.

INPUT

Input developments have kept pace with industry developments and have often led them. From the early 40s until 1965, TA data were punched from hard copy traffic onto punched cards or punched paper tape, introducing errors and delays. Electrically transmitted traffic and summaries were recorded on punched paper tape for later machine processing. You will be hearing about some significant improvements in the area of input

- AG-22 intercept typewriters (to be replaced in the future by the Improved AG-22 Typewriter System, or IATS), and FLEXSCOP (Flexible SIGINT Collection Position, incorporating a CP-818 computer). These devices and systems solve many but not all of our input problems.

Improved communications are giving us a path that is increasingly direct from intercept antenna to computer at NSA, with a minimum of manual intervention; at least that is the goal. CRITICOMM has been serving as an interim system, and the high-speed data link project termed STRAWHAT is now coming to fruition. Magnetic tape communications terminal devices in C Group (Projects MESHER and MUSSER) are parts of the antenna-to-computer system.

FORMATS

Trends in analytic formats have been similarly marked by distinct improvements. By formats here I mean methods of preparing and forwarding data for higher echelon processing, either machine or manual.

From the early 40s until 1951, the message heading itself was the format (in slightly edited form). Even during this period, there was some experimentation going on with direct machine input of electrically forwarded material. In 1943, a group of young GIs were sent to USM-2, Petaluma, California, for just such a project. We edited material for poking and electrical forwarding; the traffic was to be converted automatically to punched paper tape at the Arlington Hall Station end of a cross-country circuit. The tape was then to be converted to punched cards for subsequent EAM processing. It was not until later years that I discovered that errors on the cross-country circuit were so great that the material was not useful without substantial reediting and repunching; this experience accounts for much of my past skepticism about communications lashups.

In 1951, the "signalled log digest" (or LOGEX), the first true electrically forwarded analytic summary, was begun.

P.LP86-36-36
EO 8.4(d)

The concept of forwarding the results of each echelon's analysis to the next higher echelon so that analysts could start from that point was born here. The signalled log digests were really a slightly altered version of hand logs, forwarded by signal, even producing a space on the teletype print for analysts to draw a diagram, based on the "verbalized diagram" in the message. This early signalled material was neither intended nor suited for direct input into machines.

In the late 1950s, *MATSUM-type formats* were begun. These were similar to the signalled log digests but contained fielded data intended for subsequent EAM (punched card) processing. The Army Security Agency (ASA) in Europe pioneered this effort, and

NSA picked up the idea - the time and the situation were ready for the advance. Many similar formats for electrical forwarding were created and used over the years.

About 1967, *modular-type formats* began to appear. The best examples are the STRUM (Standard Technical Report Using Modules) and the FF STRUM (Fixed Field STRUM). As most of you know, the STRUM of a particular type is made up like a recipe, with whatever modules are necessary for the particular problem. I strongly believe in the modular approach and think that it is proving itself in practice. STRUMs contain some fielded information, but also have flagged and tagged information in the various modules.

Beginning in the late 1960s, *machine-generated reports* produced by field computers from AG-22 tapes started appearing. The first were the TRAPEZE project, [REDACTED]

[REDACTED] More recently, a version of the ELFAIR, a [REDACTED]
[REDACTED] has been computer generated and forwarded. There are undoubtedly others that we have not named.

P.L. 86-36
EEO 1144(c)

COMPUTERS TO THE FIELD

Ten years ago it was unthinkable to send computers to the field, but George Vergine, always about ten years ahead of his time; was suggesting sending out computers for field processing purposes. About 1964, [REDACTED] wrote a paper strongly recommending the IBM 1401 as the ideal field computer and headed up an early effort to develop programs for field use. Today we have many, many computers at our intercept sites, working with collection, doing conversion, editing, manipulating data, performing product-related tasks, and so on.

P.L. 86-36

PROCESSING

A few words about the kinds of processing we have done using computers are in order.

In the early 1940s, in fact until the mid-50s, it was mostly sort and list of message externals. I would like to reemphasize that the sort and list is still the workhorse of our business, despite all the fancy kinds of processing being done or tried.

By the mid-1950s, callsigns were being handled by computer, keys were being decrypted, and some early work on DF plotting had been done.

Programs to produce analytic *summaries* were first started in the mid-1950s. The computer began giving us totals, groupings of data, without giving us back all the input data. This was perhaps the real beginning of TA data reduction by computer.

More fancy things are being tried and have been proved workable. [REDACTED] will be telling you about the "Traffic Analysis Processing System" (TAPS), which includes some quite advanced ideas. Such things as net reconstruction, net continuities, and various recoveries are being done by computer in special instances; you will also be

hearing about some examples of *alarm systems* based on thresholds and exception reporting.

OUTPUT

The nature of output for the benefit of TA and SIGINT generally is sharply changing. Data links work both ways and provide us the capability of moving output directly to the field for their use.

We have already referred to conventional sorts and lists as typical computer "printouts." Matrix output, in which data are displayed at the intersections of row and column coordinates, has been around for about ten years now and is being used in a number of problems to advantage. Matrices are particularly good for displaying TA data summaries.

Computer graphics, which include the whole area of X-Y plotters and cathode-ray tube output devices (producing so-called "soft copy"), are undergoing a surge in popularity right now and show great promise for even further TA applications. Some such devices display only "alphanumeric" data, and others show either characters or lines, which can make possible the display of graphs, non-Roman characters, and even net diagrams.

Map overlays are an excellent example of such an application. I believe that Mr. will be showing you an example of an Airborne Radio Direction Finding (ARDF) map overlay, which illustrates both the idea of computer graphics and of exception reporting. The analyst is presented with an overlay showing *only* DF fixes that are within a radius of twenty-five miles of key South Vietnamese cities.

P.L. 86-36

The list of types of output to serve the traffic analyst could go on and on. Some examples that deserve mention are *updated databases*, the generation and *preparation of reports* for reproduction, and the *preparation of messages* for electrical forwarding.

To round out the picture, and to keep everyone honest, I would like to touch on some of the classic *problems* and advantages that people are encountering in TA mechanization.

PROBLEMS IN MECHANIZATION OF SIGINT

Some of the problems we are going to mention have changed somewhat since we first compiled this list a few years back. Some are much greater than they formerly were, and others have practically disappeared. Problems that might be cited are

- a. Objections to mechanization, as such
- b. Need for a back-up capability
- c. Problems in the data

1. Changes in the database
 2. Garbles and corruption ("garbage in, garbage out")
 3. Volumes of material (now often cited as an advantage, because of input breakthroughs)
 4. Data preparation problems
 5. Diversity of analytic problems
- d. Problems in defining logic
 - e. Problems in getting at the data in storage
 - f. People problems (carelessness and human errors in programming, data conversion, or data preparation)
 - g. Problems with "systems" (some systems are truly too large; others should be designed to interact but are not; the standards problem, etc.)
 - h. Hardware problems ("down time"; computers in the field experience environmental difficulties, etc.)
 - i. Delays in getting results (this is really a matter of relative priority of problems)

REASONS FOR THE MECHANIZATION OF SIGINT

Despite all the problems just cited, there are a number of overwhelming reasons why mechanization must be considered, and why for many problems it is already an indispensable part of the processing effort.

- a. Business vs SIGINT reasons (timeliness is more important in the case of SIGINT; cost may be an object, but the costs may be far greater not to use mechanization)
- b. Timeliness
- c. Accuracy (there is accumulating evidence that the use of AG-22s may increase the accuracy of input into our machine processes)
- d. Mechanization of target communications methods
- e. High volumes of raw data
- f. Derivation of new information from large information banks
- g. Processing of solved SIGINT systems
- h. Research into unknown and changed systems
- i. Larger look at data relationships (through information retrieval systems)

SOME CLOSING THOUGHTS ON TA MECHANIZATION

We had a serious communications gap between traffic analysts and computer people a few years back. Now, enough people have crossed the fences in each direction that things are greatly improved; there is increased understanding on both sides of the fence.

The integrity of data in machine files is a matter of concern. Somehow, we must preserve the total context of transmissions and provide an "audit trail" through machine analysis, so that we can always find out how certain conclusions were reached and what was actually transmitted in the first place. Also, the need for the simple inclusion of *validities* in machine records is a matter of concern. Things are seldom certain, but when we do not show a degree of validity, we imply that they are.

A final word on standards. I am not sure that completely "standard formats" are the right goal, although I am a staunch believer in modular formats, with many modules, appropriate for problems X, Y, and Z.

The slowness in developing workable data standards may or may not be our most serious single bottleneck in the whole area of mechanization. We must move more rapidly in standardizing data features – how we express things, irrespective of formats. Databases and retrieval systems demand this. There needs to be a more considered attempt to interweave A, B, and G Group processing systems. In fact, each office has not one but many processing systems. The key to what I am saying lies in the use of effective data standards, I believe, not in the creation of a large "system."

In conclusion, the future of TA mechanization looks bright. I can foresee the continued preparation of manually prepared modular-type formats for a long, long time in the future, but on a reduced scale. I believe that machine-produced summaries based on AG-22 output have a great future, but I would like to emphasize three things:

- a. In using AG-22s in this manner, we are essentially moving the point of analysis from one place to another; with communications improving, we can do this.
- b. We must make it possible for the analyst to get back to the total take, in some manner, if there is the need to do so.
- c. An adequate "alternate system" or path must be provided to ensure that vital information, technical feedback, and intelligence can continue to flow in the event that data links are down. The use of modular-type formats as an alternate, backup procedure for all AG-22 installations is probably the best answer to this concern; this implies the continued need for some analysts at field sites.

I hope that I have sounded neither too pessimistic nor too optimistic today. There are a lot of very real problems emerging in TA mechanization – high error rates on summaries, humidity and dust interfering with operations, machine breakdown, escalating costs, and so on, but on the whole, we are moving ahead.

I am looking forward to hearing what the other speakers have to say and to further forums of this type in the future. Thank you.

Automation of a TA Process

TIM MURPHY



What few people realize is that the workload associated with the Southeast Asian (SEA) problems did not decrease as U.S. forces withdrew from Vietnam and as B3 (previously B6) underwent successive personnel reductions. In fact, the amount of SEA communications intercepted and processed at NSA was still peaking in 1975 when the South Vietnamese government fell.

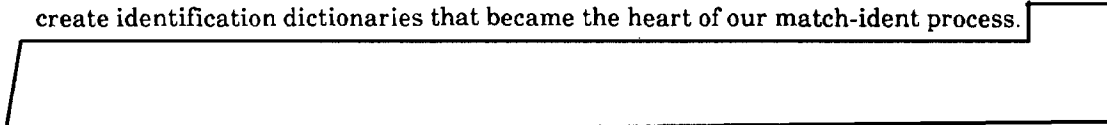


Traffic analysts have borne a proportionate share of the decrements incurred by B3, yet there has been virtually no drop in target communications. Analysts have been assigned larger case loads and have assumed responsibility for collection management; yet, in my view, the quality of analysis has improved, and the analysts, in general, are not overworked. There is little doubt that technical support to field stations has improved significantly.

The automation of a large segment of the SEA traffic-analytic processes has been the key not only to high-quality analysis but also to our ability to do the job better over the years with fewer people. Fortunately, during the early stages of U.S. involvement in Vietnam, managers with the Traffic Analytic Support Division for the Vietnamese problem took an enlightened view toward mechanization. With help from some of the Agency's leading traffic analysts and a cadre of highly qualified data-systems personnel, the first steps toward an automated traffic-analytic process were taken.



At the same time, we began a major effort to mechanize our analytic findings and create a single processing system. In addition to transmitting our analytic findings to field sites, we also used their data to create identification dictionaries that became the heart of our match-ident process.



P.L. 86-36
EO 1.4.(c)

[REDACTED]

Of equal importance to these contributions was that the methods the experts used and the approaches they took to the TA task rubbed off on many of the junior analysts, with outstanding support first from C5 and later from B42, continued this trend toward automation, always striving to relieve analysts of most of the repetitive TA tasks and to permit more time for actual analysis. Among their major contributions were

- the creation of a single analytic database – the Southeast Asian Case File (SEACF)

[REDACTED]

In addition, the availability of on-line access to our database in recent years through the COPE terminal has led to many analyst-initiated special programs that have greatly expanded the analysts' capacity for research. I should also point out that field stations tasked against SEA targets have also taken many initiatives toward automation.


SEA MACHINE DATABASES, PROCESSES, AND RESOURCES

SEACF

Perhaps the most significant single step toward reducing the workload of SEA traffic analysts was the creation of the Southeast Asian Case File (SEACF). The SEACF not only resulted in highly efficient database management, but also permitted many follow-on processes that greatly reduced the workload of SEA traffic analysis.

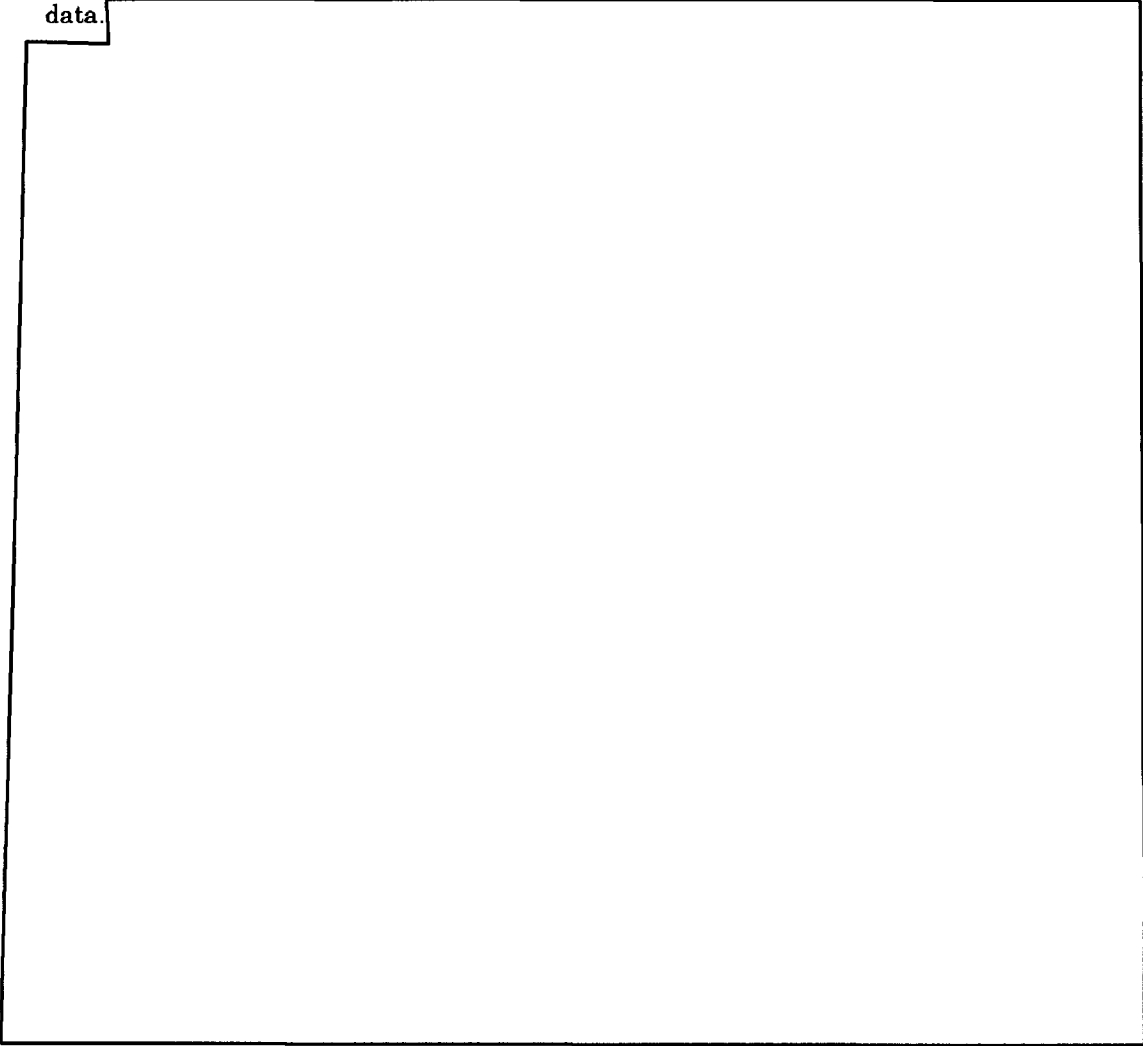
Many of the processes introduced during the early stages of the Vietnam War to mechanize the TA process required the establishment of databases to support them. As those processes expanded, the number of databases expanded to the point where analytic time saved was being spent on database maintenance. A review of those databases revealed that there were a large number of common fields of information but comparatively few unique fields. In effect, we were multiplying our file-maintenance workload and making ourselves vulnerable to contradictions between identical data items in different databases. The major databases that had to be maintained prior to the implementation of SEACF were

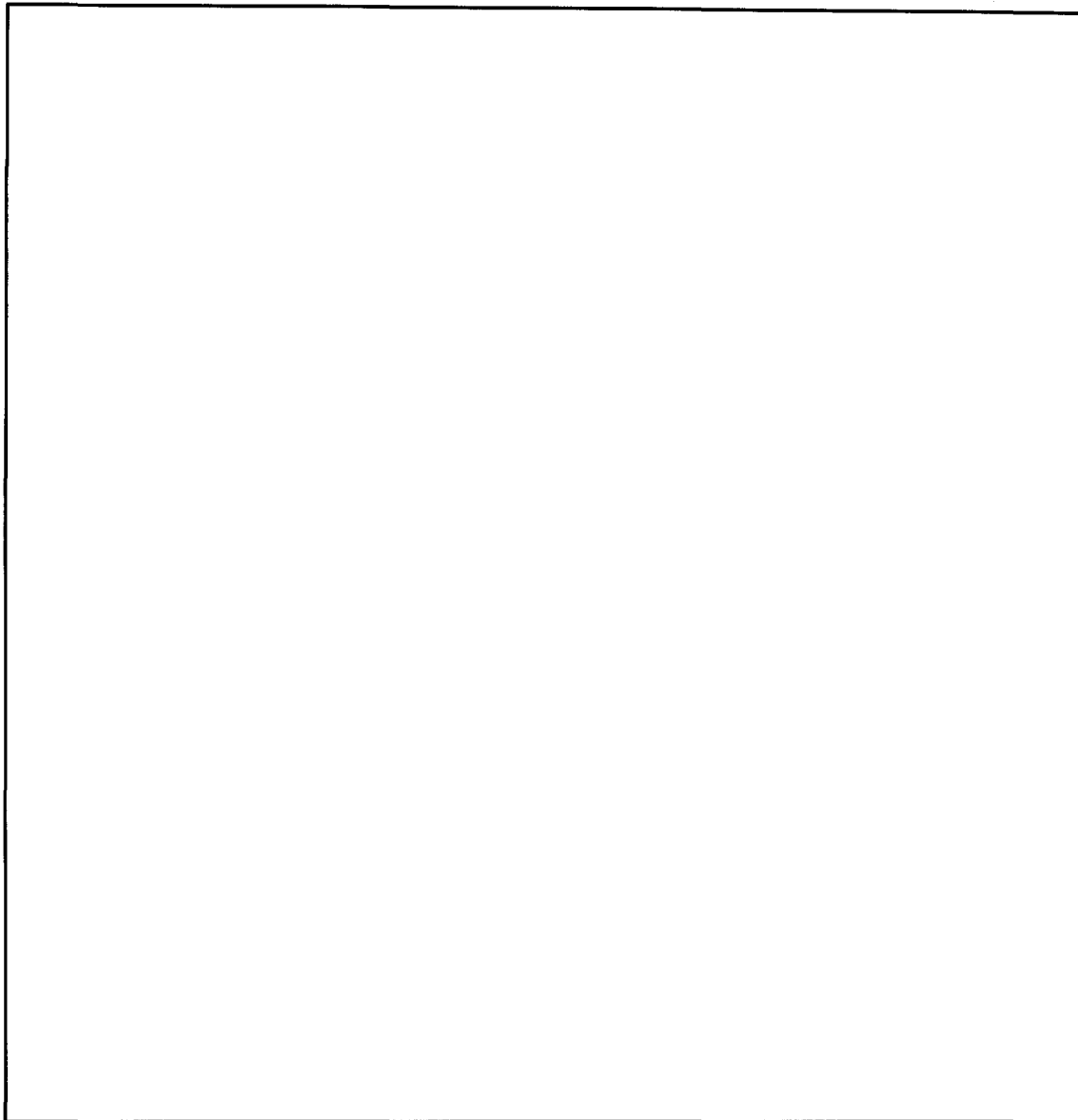
P. 8688636
EO 4.1(c)



The SEACF consolidated all these data elements into a single database, thus eliminating the requirement to maintain those multiple databases. Since its establishment, the SEACF has been expanded to include data elements that support both cryptanalytic and collection-management functions.

As now constituted, the SEACF consists of eight basic data records which analysts use to provide permanent maintenance of the communication and crypt characteristics of a given target, identification and location data on that target, and collection-management data.





It was inefficient and time-consuming for analysts to collate all these varied inputs, yet very necessary from an analytic standpoint. Hence they were eventually collated by machine as part of our daily process, and our activity database is now a composite of all primary-source technical inputs.



IATS

Another major step in automating both the traffic-analytic and the traffic-forwarding processes was the implementation of IATS. The user routine developed for SEA communications copied on IATS or AG-22 intercept positions -

P.L. 86-36
EO 1.4.(c)

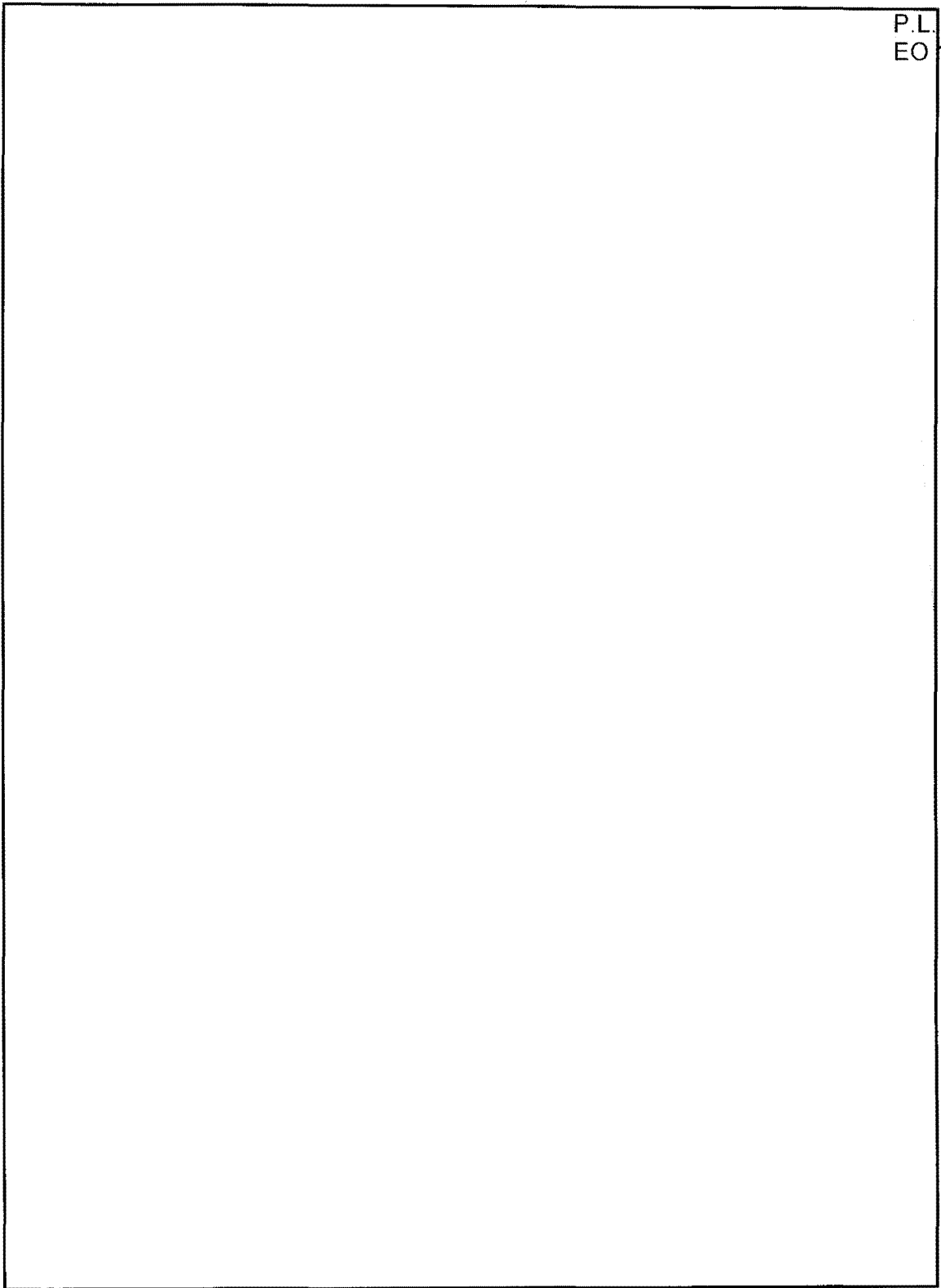
This program reduced by countless hours the task of the SEA development analyst. To a large degree, the analyst has been able to rely on the machine to isolate new continuities and then concentrate his or her efforts on identifying that continuity to region or function. The number of analysts tasked against the development problem was reduced from forty-five to ten between 1972 and the spring of 1975 with no adverse effect on the mission. Much of the credit goes to the SEADEV process.

P.L. 86-36
EO 1.4.(c)



~~SECRET SPOKE~~

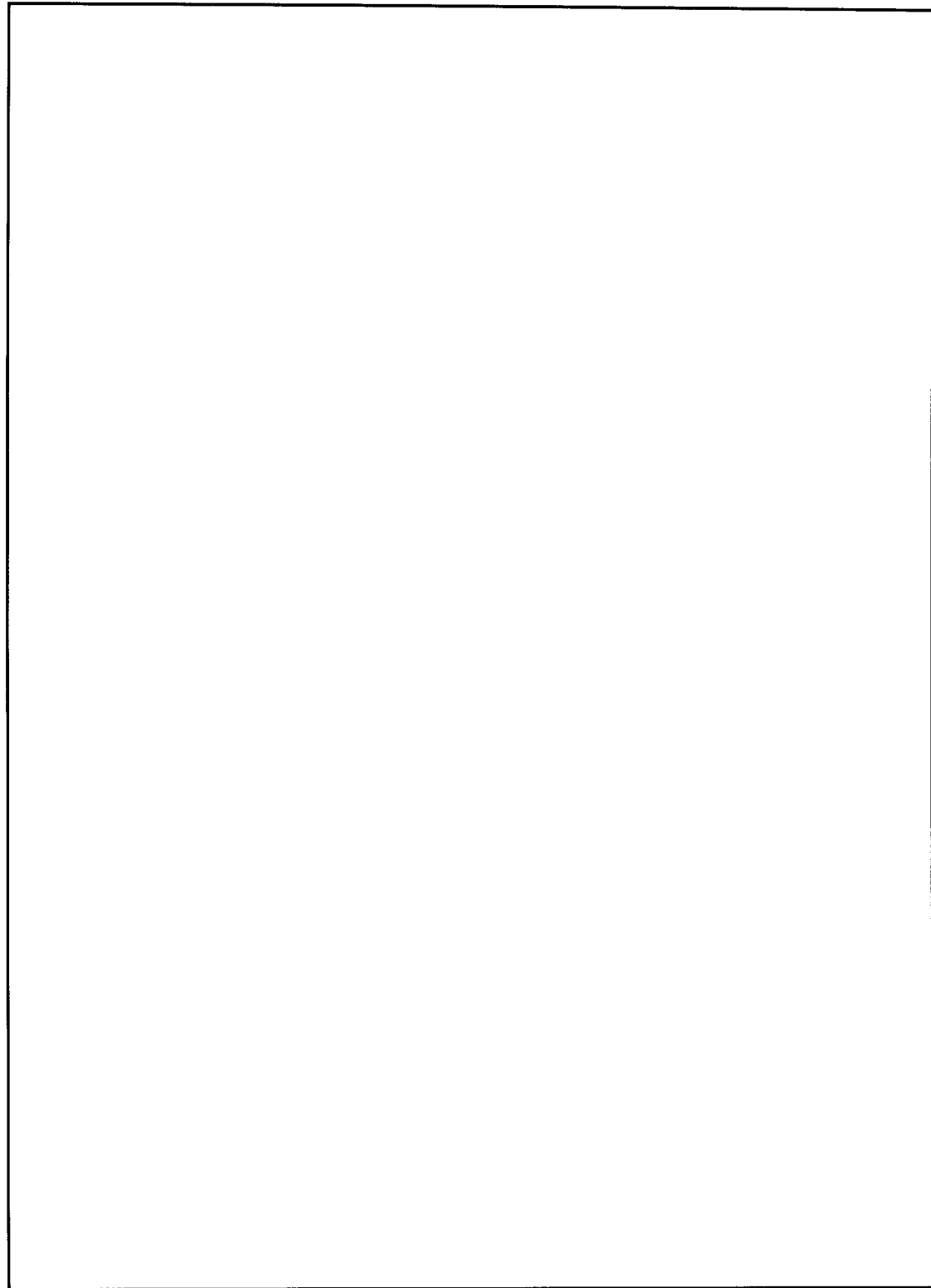
P.L. 86-36
EO 1.4.(c)



~~SECRET SPOKE~~

~~SECRET SPOKE~~

P.L. 86-36
EO 1.4.(c)



~~SECRET SPOKE~~

P.L. 86-36
EO 1.4.(c)

IMPACT OF AUTOMATION

In summary it can be said that the impact of automation on the SEA problem has significantly reduced the number of analysts required to do the job. The caliber of analysts required has increased, however, since their primary remaining function is pure analysis or, to use another term, "thinking." Many of the crutches that have kept analysts busy over the years (e.g., logging, sorting traffic) are gone.

To function effectively, SEA analysts must have an understanding of their databases and of how machines can be used to manipulate the data they contain. Imagination is currently a key asset since much of what can be imagined in terms of analytic approaches is now feasible. There is an increased demand for the traffic analyst/programmer. Knowledge of the SPECOL retrieval language is becoming a highly desirable attribute of the SEA traffic analyst. In short, the impact of TA mechanization needs "a few good analysts."

How Clean Does a Database Need to Be?



P.L. 86-36

One of the first things one learns about computers is that they require a much higher order of accuracy in the material they manipulate than do comparable "human" processes. One learns to pay an extra measure of tribute in the form of added proofreading or other forms of quality control, so that the input is "clean" enough for the computer to handle.

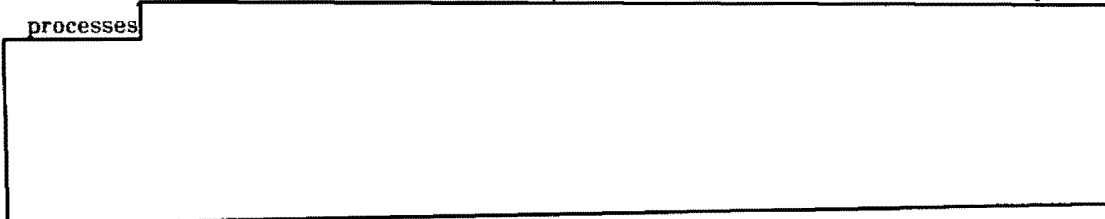
After a while, as the novelty wears off, it sometimes occurs to one that not all of the data need to be so awfully clean. If we expect to sort or retrieve on a particular field or data element, then that field or data element should be clean and garble-free; but if a neighboring item is never (well - almost never) used as a control for sorting and retrieving, then it only needs to be as garble-free as *people* need. Quite clearly, if only half of your data elements really *need* quality control, then some of that manpower now spent scrubbing each little data element might be diverted to other tasks.

It is possible to imagine categorizing data elements as "first order" if they need to be "computer clean," and as "second order" if they need to be only "people clean."

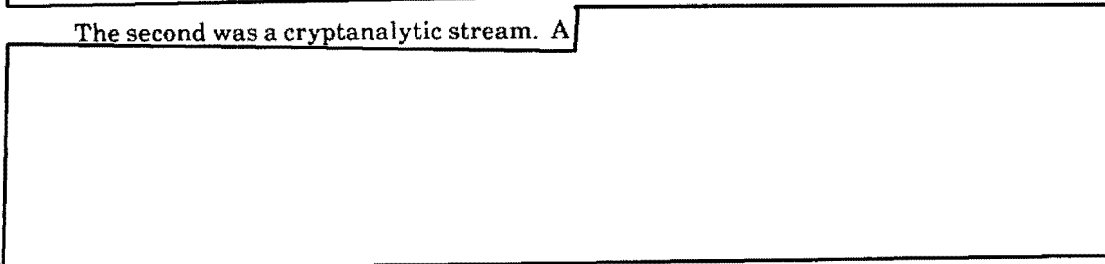
P.L. 86-36
EO 1.4.(c)


In this day of great monolithic databases, however, the use of varying quality levels can cause troubles, however laudable their manpower savings may be. A story "from life" will illustrate.

Some years ago, during the Vietnam War, we found ourselves receiving two streams of electrical material from the sites in the field, and both streams were used to feed computer processes



The second was a cryptanalytic stream. A



The specific details of these processes belong to another story (or series of stories). The point here is that there came a time when there was an operational need to identify which messages  It should have been easy. Neither system was new, and both had been working for quite a while with

~~CONFIDENTIAL~~

reasonable success. (Success is a relative term; there were always problems, sometimes earthshaking problems, but by and large, the systems did work.)

[REDACTED]

It took a while to find out why, but after a time the answer became clear. Evidently the people at the sites, knowingly or unknowingly, practiced different levels of quality control on the various data elements [REDACTED]

[REDACTED]

All of which suggests several thoughts.

Garble rates can often be determined, at least approximately, by machine. Certainly differential garble rates can be (Field 1 has more or fewer garbles per thousand than Field 2). If databases that now exist were measured to show which data elements were "cleanest" and which were "dirtiest" (perhaps arrayed in a sort of quality hierarchy),

- the unwary might be warned off using the database for sorting or controlling on the wrong (dirtiest) data elements;
- hit thresholds might have to be lowered when dealing with "dirty" elements, even at the expense of wading through more "garbage" output;
- managers might better understand the manpower costs of various control strategies; but also
- we might decide that great monolithic databases are not always the answer when one must work with a variety of data sources having widely different notions of which items are "important."

~~CONFIDENTIAL~~

The Hand Is Not Quicker Than the Eye

P.L. 86-36



Many years ago I was told that "a good analyst" does his own logging, counting, and tallying. While at first I admit I thought I was being set up to do all the menial labor (my son calls it "the donkey work"), it wasn't long before I saw that the senior people around me *did* do their own logs, and counts, and tallies – not always, but much of the time.

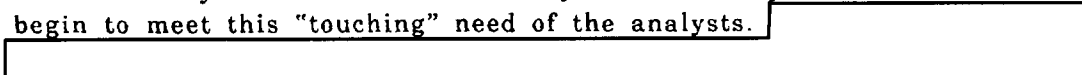
I can remember more than one callsign problem when the callsigns had been extracted (by hand) from some sort of generating matrix and inscribed (again, by hand) into pages. Often the markers of the pages would begin in an attempt to be very "random," but boredom and inattentiveness quickly set in and both processes, extraction and inscription, slowly became more orderly and regular. Near the end, as fatigue set in, the processes became virtually mechanical. Such "handwriting" patterns – top to bottom, left to right – are sometimes so strong that one can convert pages from arbitrary to true base on this feature alone. On a problem like this, one can deal with the problem on a statistical basis – one can even convince one's statistically/mathematically minded friends of the truth of one's solution – but the acid test is to sit down with pencil and cross-section paper and, by hand, duplicate the process, callsign by callsign. If your solution is right, you see it and you'll *feel* it intuitively.

In both of these processes, we gain analytic insight by doing it ourselves. Logging brings us into contact – a kind of slow-motion contact – with the material we are studying. I don't know about you, but I know that I have *discovered* more things while logging, or counting, or tallying, or some other donkey work than I have while sitting there looking at the results of the logging/counting/tallying. Especially the discoveries that were unexpected – outside the range of what I *thought* I was going to find. In the extraction/inscription kind of problem, the insight comes as my hand follows the hand of the enemy signal officer. Why did he stop just there? What made him jump over (or away from) that callsign?

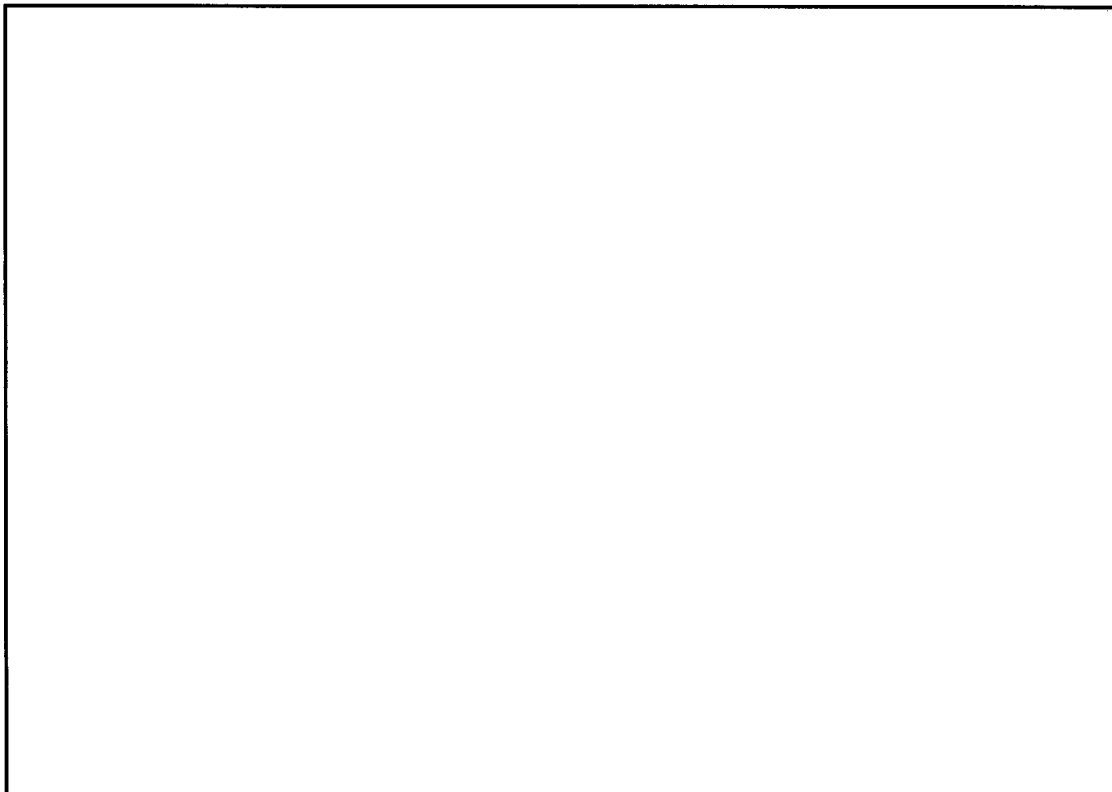
Nowadays we are told that the computers are going to do all things for us. But what happens to that intimate contact with the material when the original traffic – if you can call it that anymore – is deep inside a computer somewhere and all you've got to look at is some computer output? Well, some of us still find a way to do some hand-logging. And not just because we can't get responsive software support (that's a subject for a separate article). We *want* to log! It helps us touch the material. It's like buying a piece of land and not being satisfied with topographic charts – you have to go out and walk all over it. Then you begin to *know* the land.

P.L. 86-36
EO 1.4.(c)

I have always had the notion that someday interactive computers with *screens* might begin to meet this "touching" need of the analysts.



~~CONFIDENTIAL~~



Has anybody done this sort of thing or ever seen it done? If it works, I'd sure like to stop all this logging!

?

~~CONFIDENTIAL~~

The Lost Indicator

DR. RALPH W. JOLLENSTEN

INTRODUCTION

The purpose of this talk is to describe the type of questions that arise when a math problem is encountered in TA. Although some would argue that the problem discussed here is not strictly a TA problem, the underlying general principles are the same as those for a purely TA problem in which math is needed.

PROBLEM

A certain country has been sending military traffic in five-digit groups, using the fifth group (A5) of the message text as the indicator group according to the following rule:

Air traffic – the digits of the group sum to 23.

Army traffic – the digits of the group sum to 24.

Navy traffic – the digits of the group sum to 25.

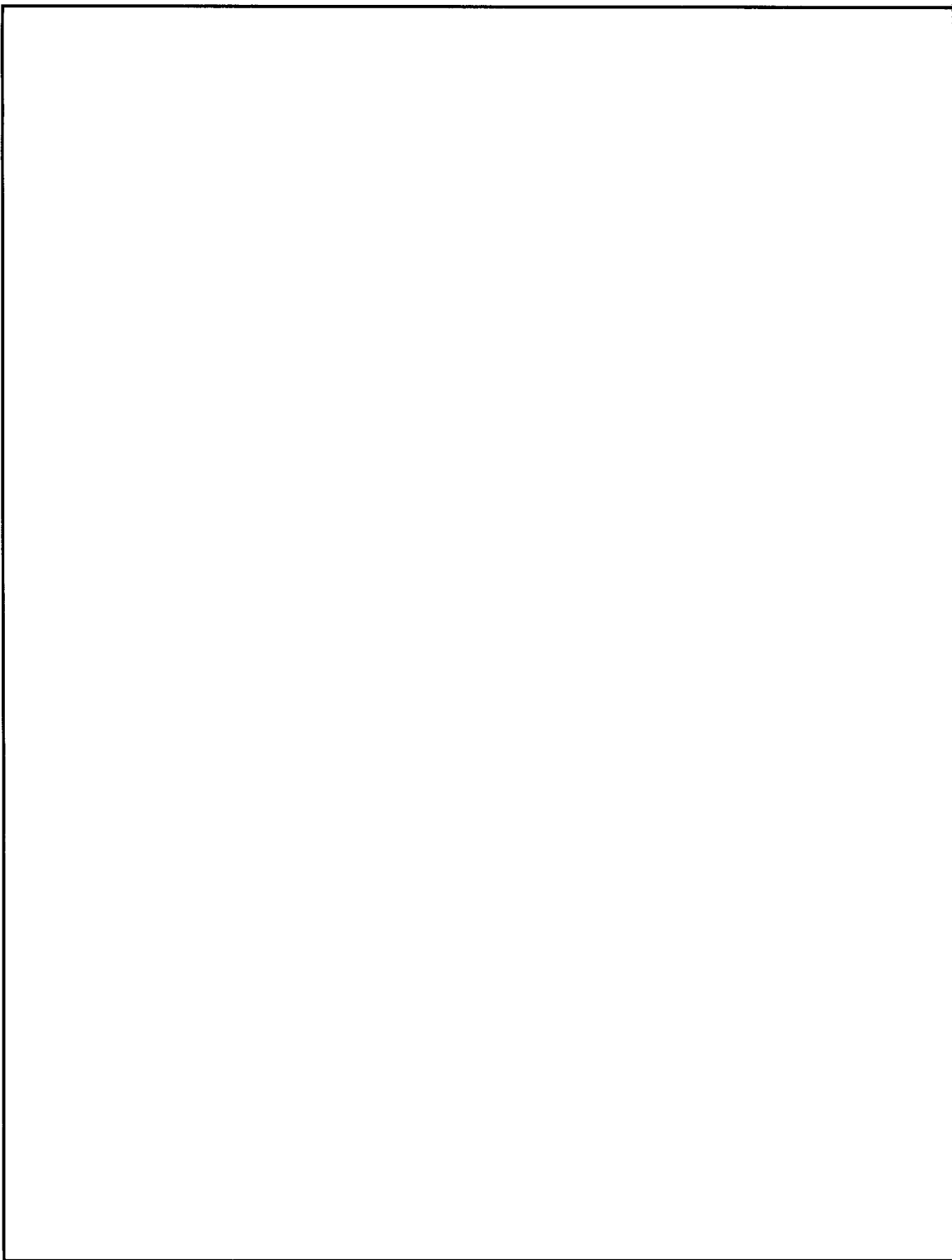
The military traffic is mixed in with other traffic having the same externals, except that the A5 position of the nonmilitary traffic exhibits no detectable bias. Obviously, the indicator characteristic is useful in differentiating between military and nonmilitary traffic.

Suddenly, on the first of the month, the indicator disappears from the A5 position of the military traffic. *What happened to the indicator group?* Did it move to another position, or has an additive been applied to it?

SOLUTION

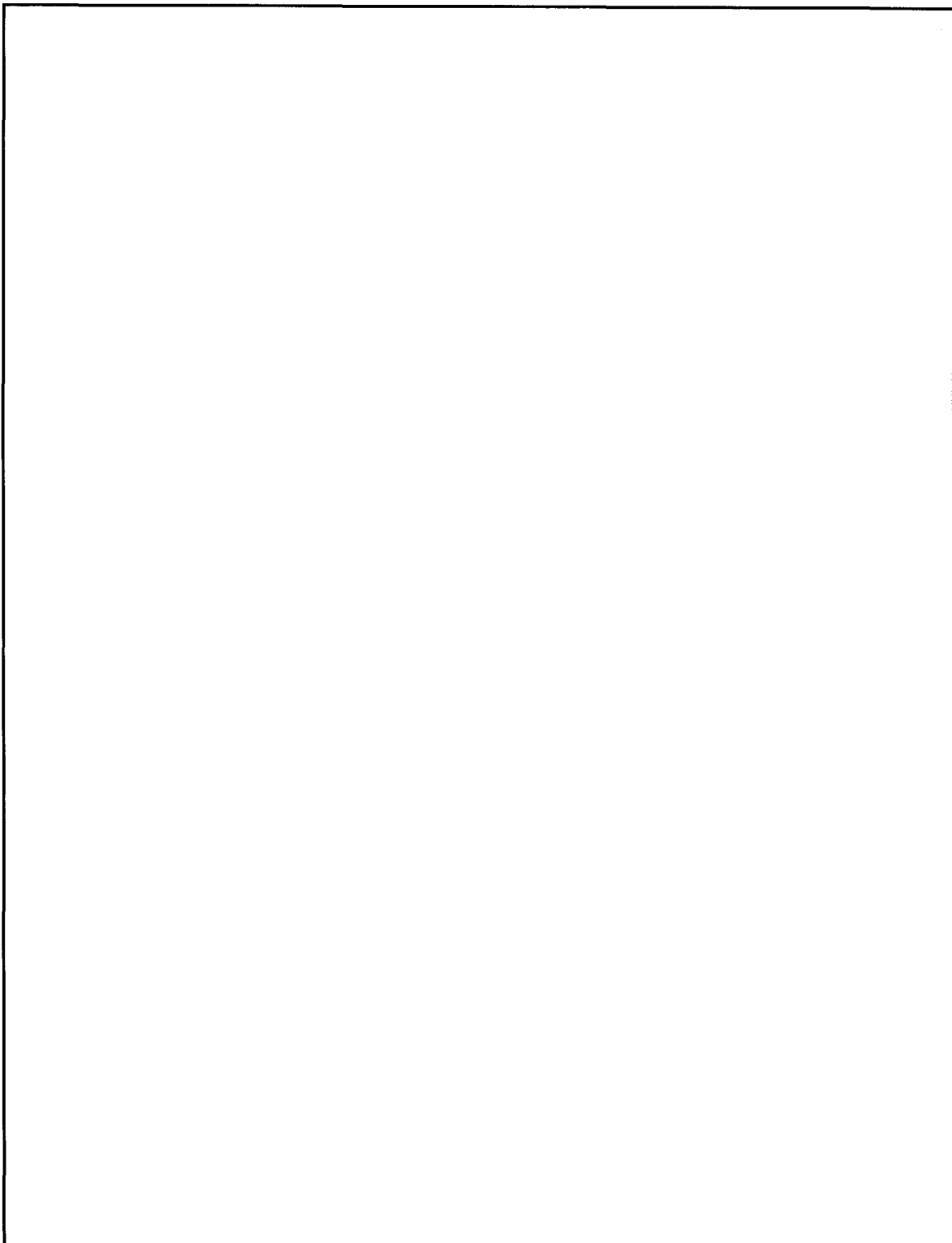
The analyst needs a method for diagnosing what has happened. To begin with, he decides to check positions A1 through A10, inclusive, to see if any of these positions exhibit any characteristics that would favor the hypothesis that the old-type indicator is present there. He has 200 messages to work with. Since he is not sure that they are all military, he doesn't know the distribution of the groups in each particular position of the messages. Some of the groups in a particular position may have causal sums while the rest do not. However, he feels sure that if the indicator is not present, the groups are distributed as if the digits of each group were equally likely. Therefore, he sets out to find or calculate the distribution of the sums of five-digit groups when each digit is equally likely.

~~SECRET SPOKE~~



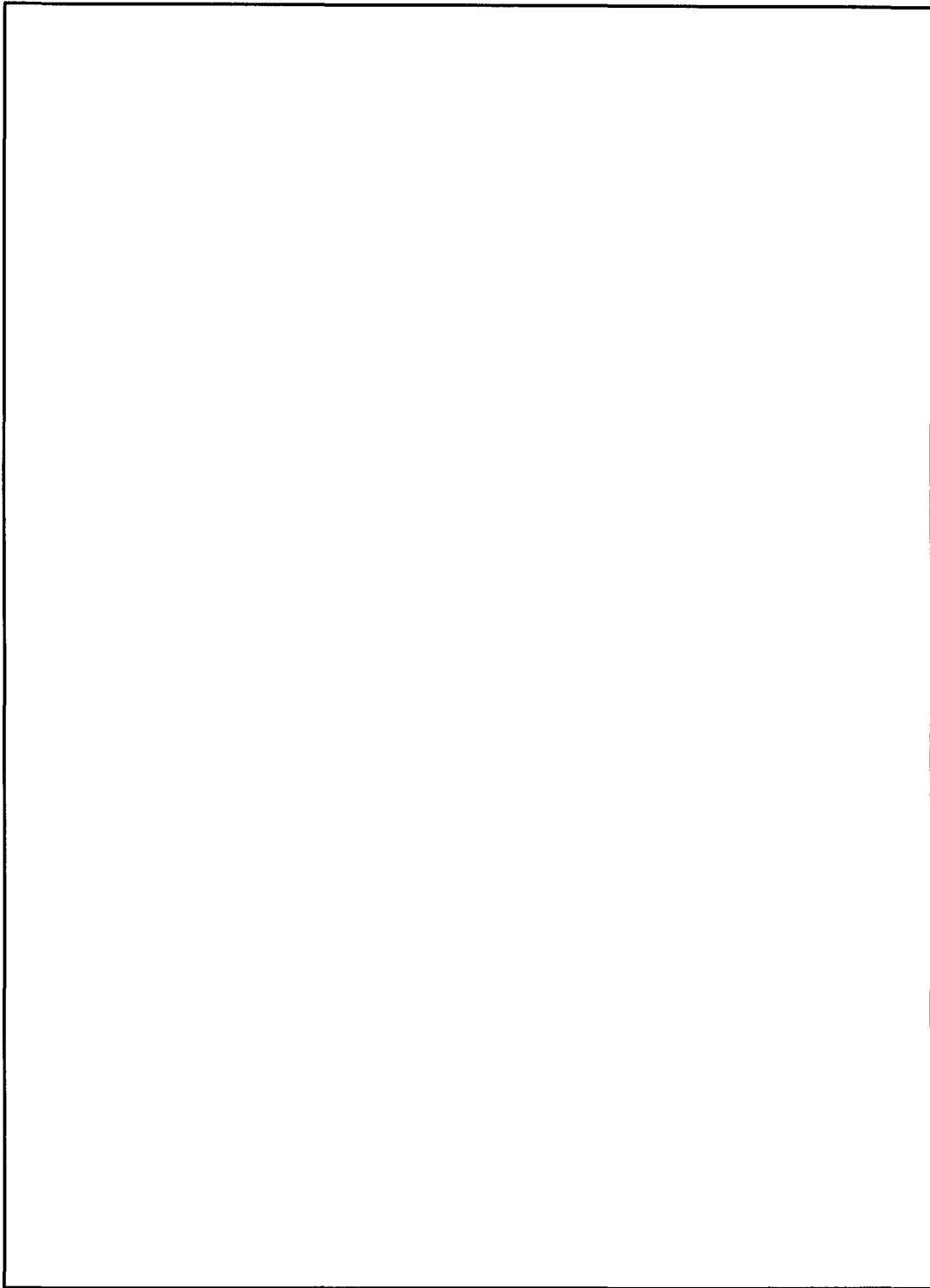
~~SECRET SPOKE~~

~~SECRET SPOKE~~



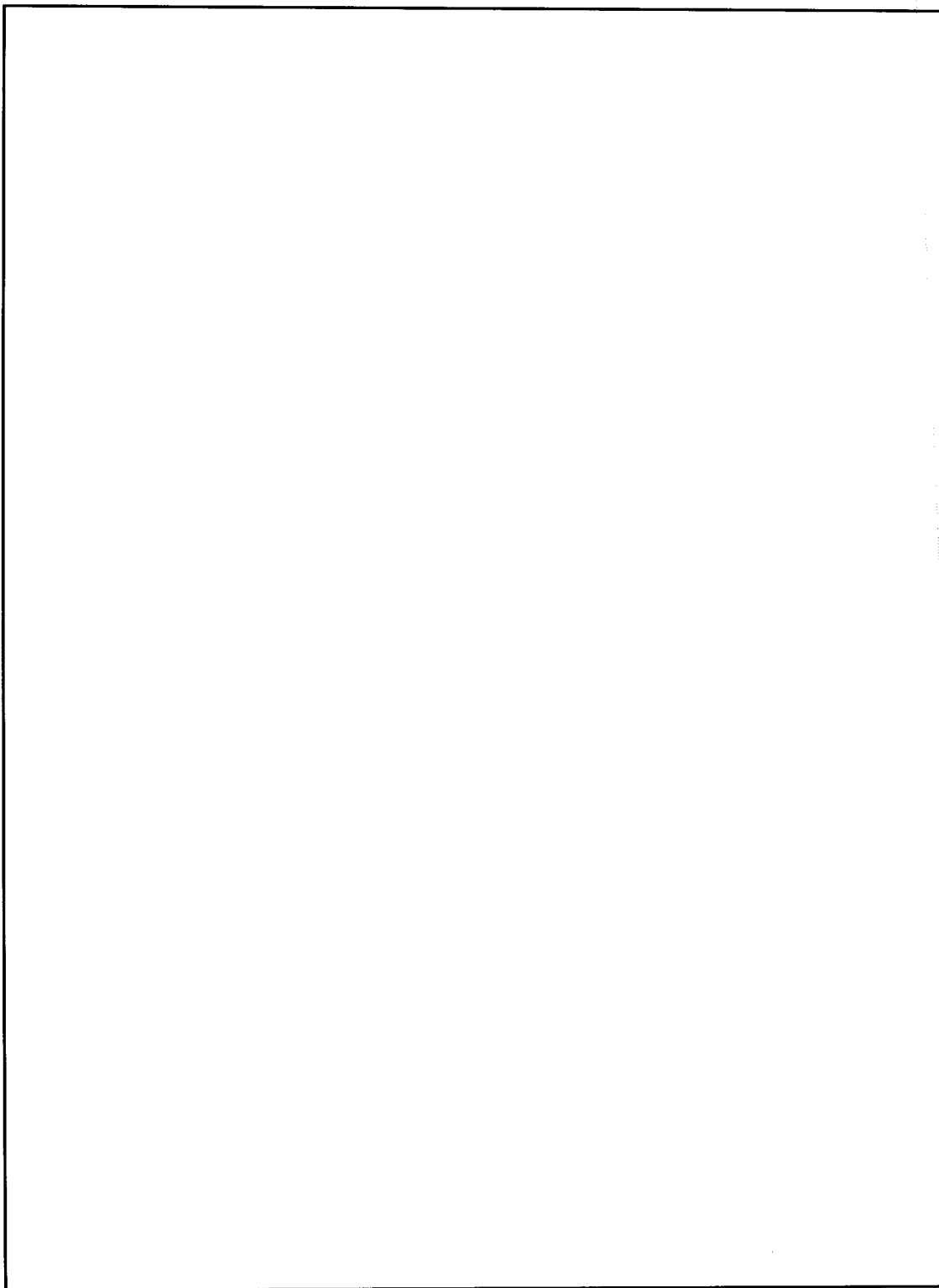
~~SECRET SPOKE~~

~~SECRET SPOKE~~



~~SECRET SPOKE~~

~~SECRET SPOKE~~



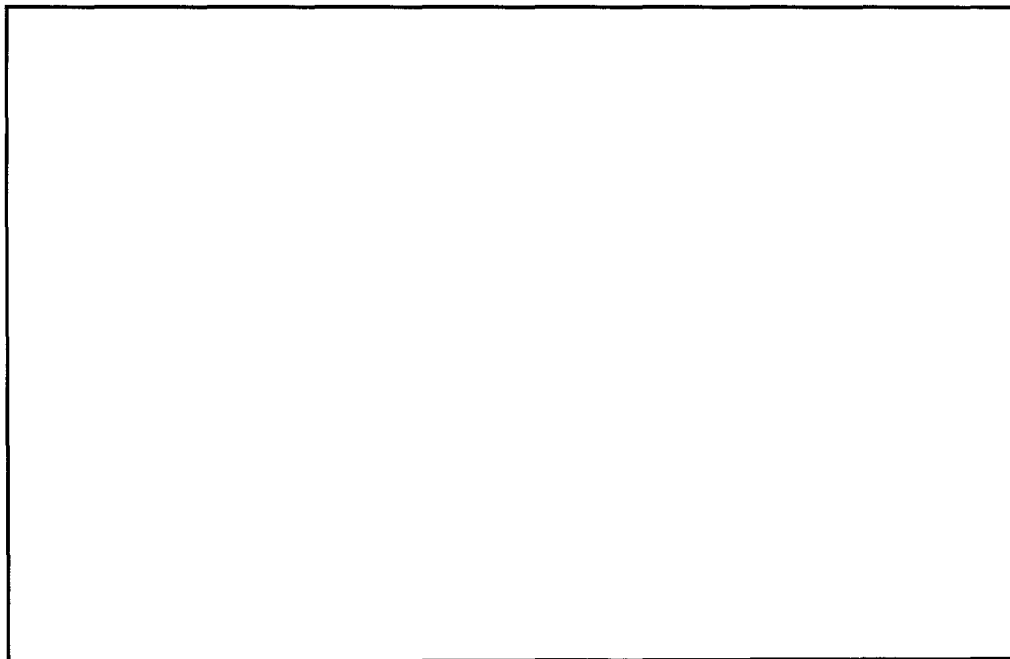
~~SECRET SPOKE~~

~~SECRET SPOKE~~



~~SECRET SPOKE~~

~~SECRET SPOKE~~



~~SECRET SPOKE~~

Applications of Set Theory to Traffic Analysis

P.L. 86-36



The purpose of this presentation is to show an approach to traffic analysis that uses set theory concepts and terminology. For this purpose, we have created a hypothetical problem designed to represent some aspects that could be reflected in a real problem. Although the following example is limited to callsign analysis, a similar approach could be used to analyze frequencies, practice traffic systems, or other data.

The hypothetical problem involves the country of Transylvania, which is divided into seven military regions labeled A through G. The callsigns used in Transylvanian communications are derived from a fifteen-page callsign system. Each page is a ten-by-ten matrix of 100 callsigns with each callsign appearing only once in the system. Further, let us assume that each military region can use callsigns from three different pages as follows:

AREA	CALLSIGN PAGES
A	01, 08, 04
B	02, 09, 05
C	03, 10, 06
D	04, 11, 07
E	05, 12, 03
F	06, 13, 01
G	07, 14, 02

A "set" is defined as a collection of well-defined objects; given any object, one can determine whether that object belongs or does not belong to a given set, i.e., whether it is an element of the set. To illustrate a callsign set, we use a Venn diagram as shown in figure 1.

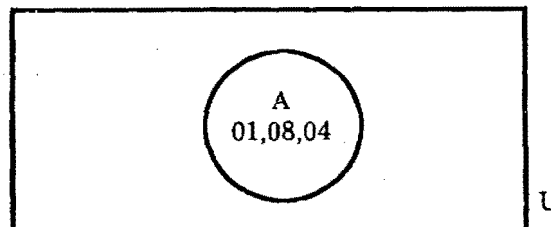


Fig. 1

The set "U," represented by the above rectangle, is the universal or complete set of callsigns in the Transylvanian system. In general, the universal set refers to the entire collection of objects being considered. The circle labeled "A" represents the set of callsigns from pages 01, 08 and 04, which are available for use on the internal communications of Military Region "A." Thus, all callsigns on pages 01, 04 and 08 are elements of the set of callsigns used by region A.

The union of two sets can be illustrated by the shaded area in figure 2.

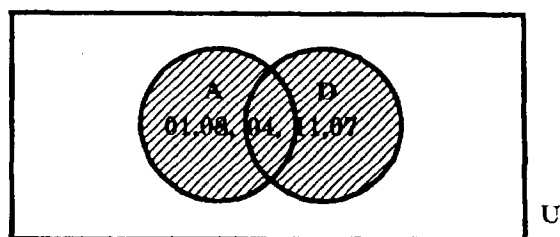


Fig. 2

We define the "union" of two sets A and D, denoted $A \cup D$, as the collection of all elements that belong to either the set A or the set D, or both. The above example represents the union of the callsign sets used by Transylvanian Regions A and D, i.e., all callsigns in pages 01, 04, 07, 08, and 11.

The "intersection" of two sets can be illustrated by the shaded area of the Venn diagram in figure 3.

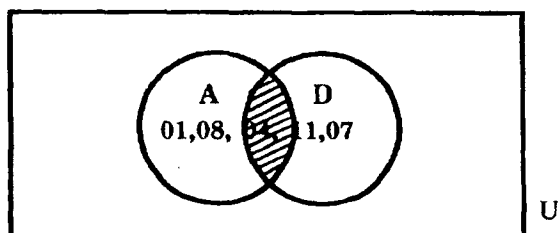
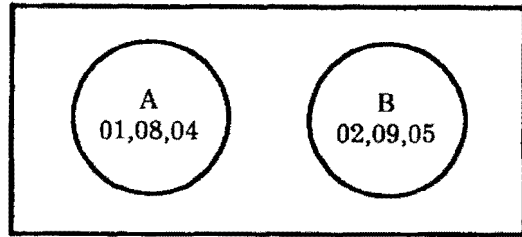


Fig. 3

We define the "intersection" of two sets A and D, denoted $A \cap D$, as the set of all elements that belong both to A and to D. In this example, only callsigns from page 04 are available for use by both Military Regions A and D. Therefore, $A \cap D$ is the set of callsigns on page 04. Suppose an unidentified radio group represents communications for elements of either Region A or D. Three possibilities exist:

- If at least some of the callsigns used are taken from either pages 01 and/or 08, then the radio group is in Region A.
- If at least some of the callsigns used are taken from either pages 11 and/or 07, then the radio group is in Region D.
- If all the callsigns used are extracted from page 04, then the radio group may be located in either Region A or D.

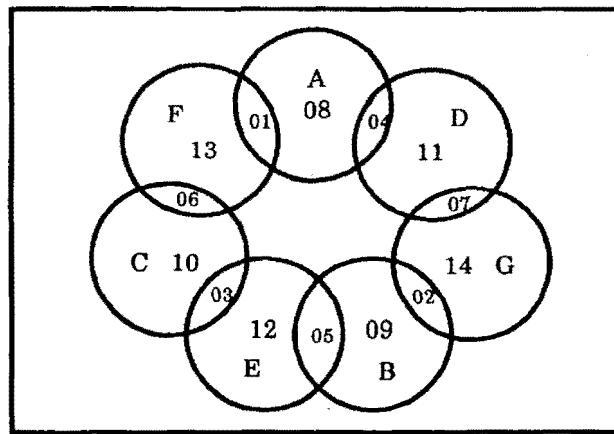
Two sets are said to be "disjoint" if they have no elements in common. For example, in figure 4 it can be seen that the communications of Region A will not use callsigns available to Region B and vice versa. This callsign set relationship gives a firm basis for the differentiation between communications of these regions.



U

Fig. 4

In order to gain more insight and understanding of the communications of an entire country, it may be helpful to represent the technical data under consideration as sets in a Venn diagram. For our example, see the Venn diagram in figure 5.



U

Fig. 5

AREA	CALLSIGN PAGES
A	01,08,04
B	02,09,05
C	03,10,06
D	04,11,07
E	05,12,03
F	06,13,01
G	07,14,02

From this Venn diagram, we can see that callsigns from pages 08 through 14 are unique to areas A through G, respectively. Consequently, these callsign pages can be used as a basis for identification of the area of communications. Since each of the callsign pages 01 through 07 is used by two different military regions, use of callsigns from these pages would reduce the problems of identification to one of two military regions. Further, callsign duplication, as shown in the above Venn diagram, may exist because of the

geographical separation of a military region by a natural barrier such as a mountain range, sea, or another country.

In conclusion, the analysis of technical data using set concepts is a systematic approach that may result in more understanding of the communications under study.

TA, CA, LOGIC, MATH: Where Do They Intersect?

P.L. 86-36



PL 886386
EO 1.44 (f)

~~TOP SECRET UMBRA~~



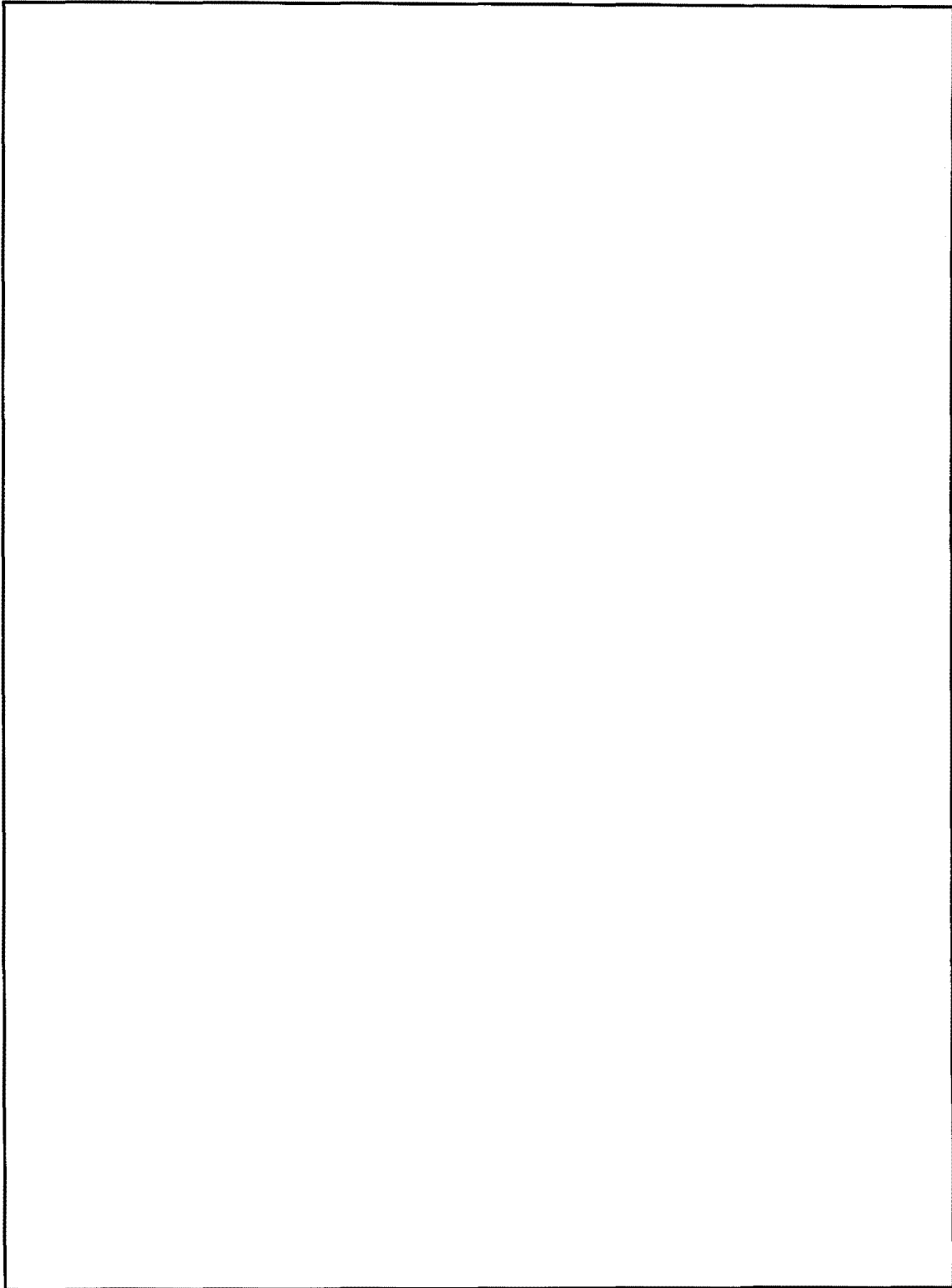
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



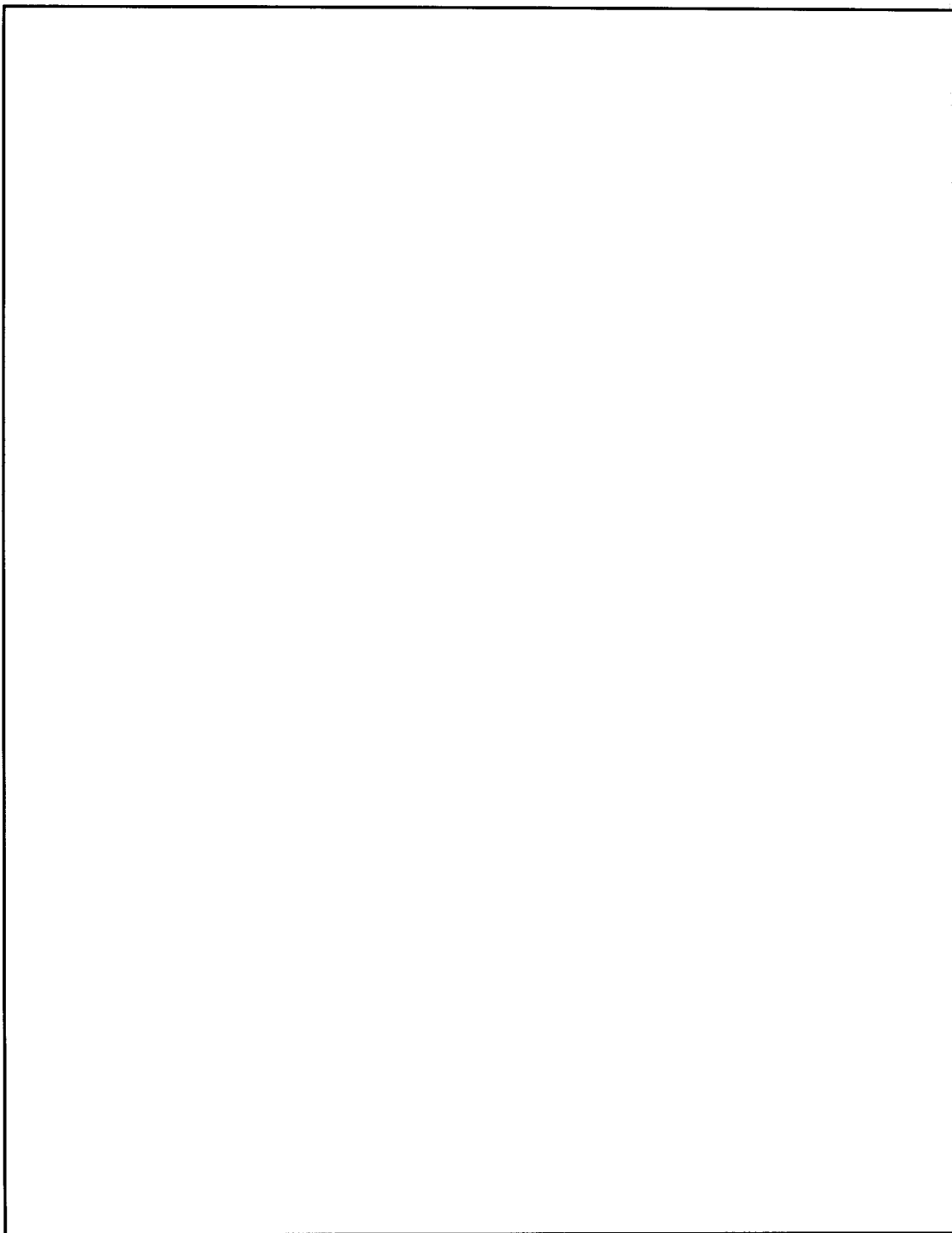
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



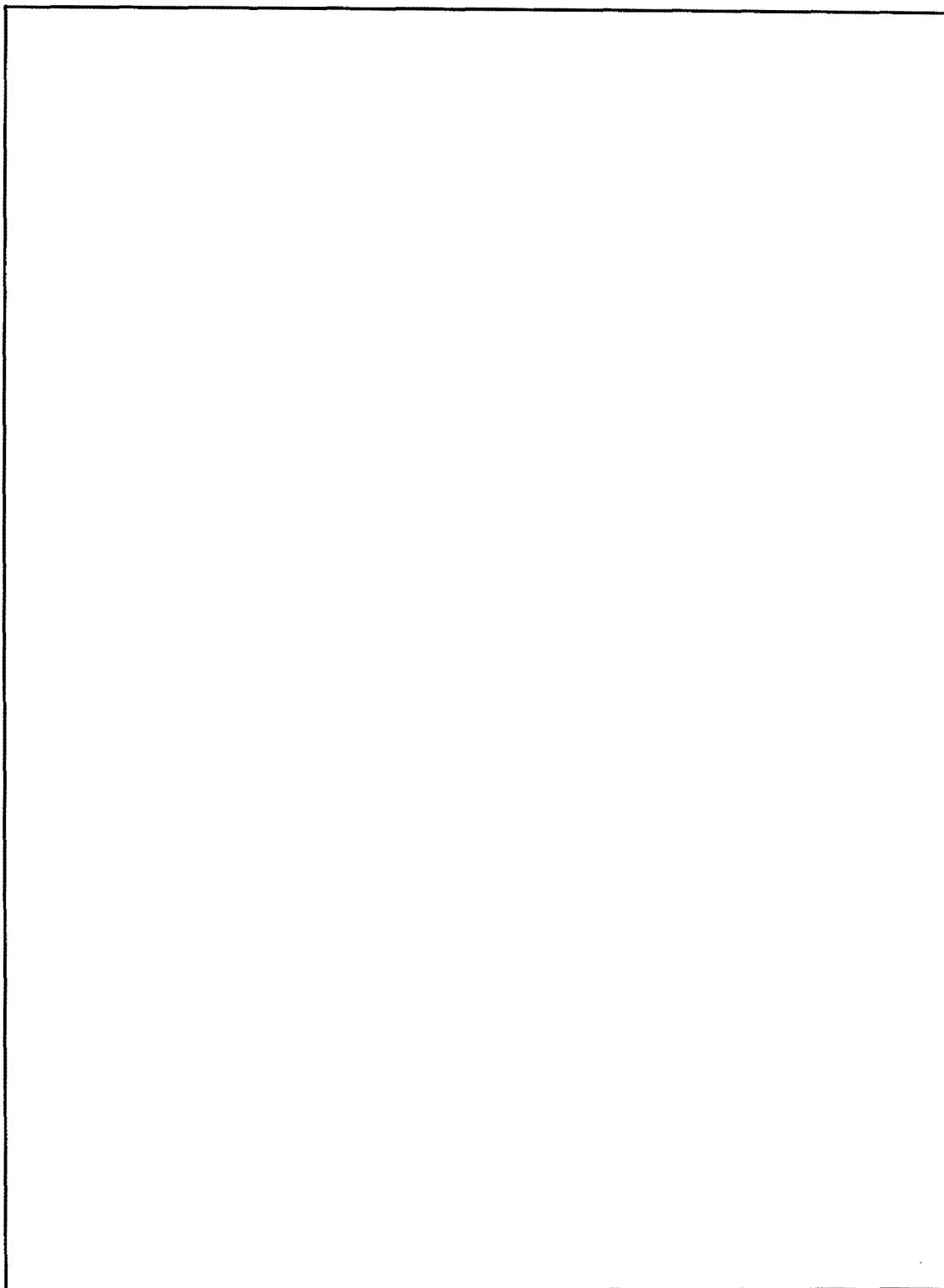
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



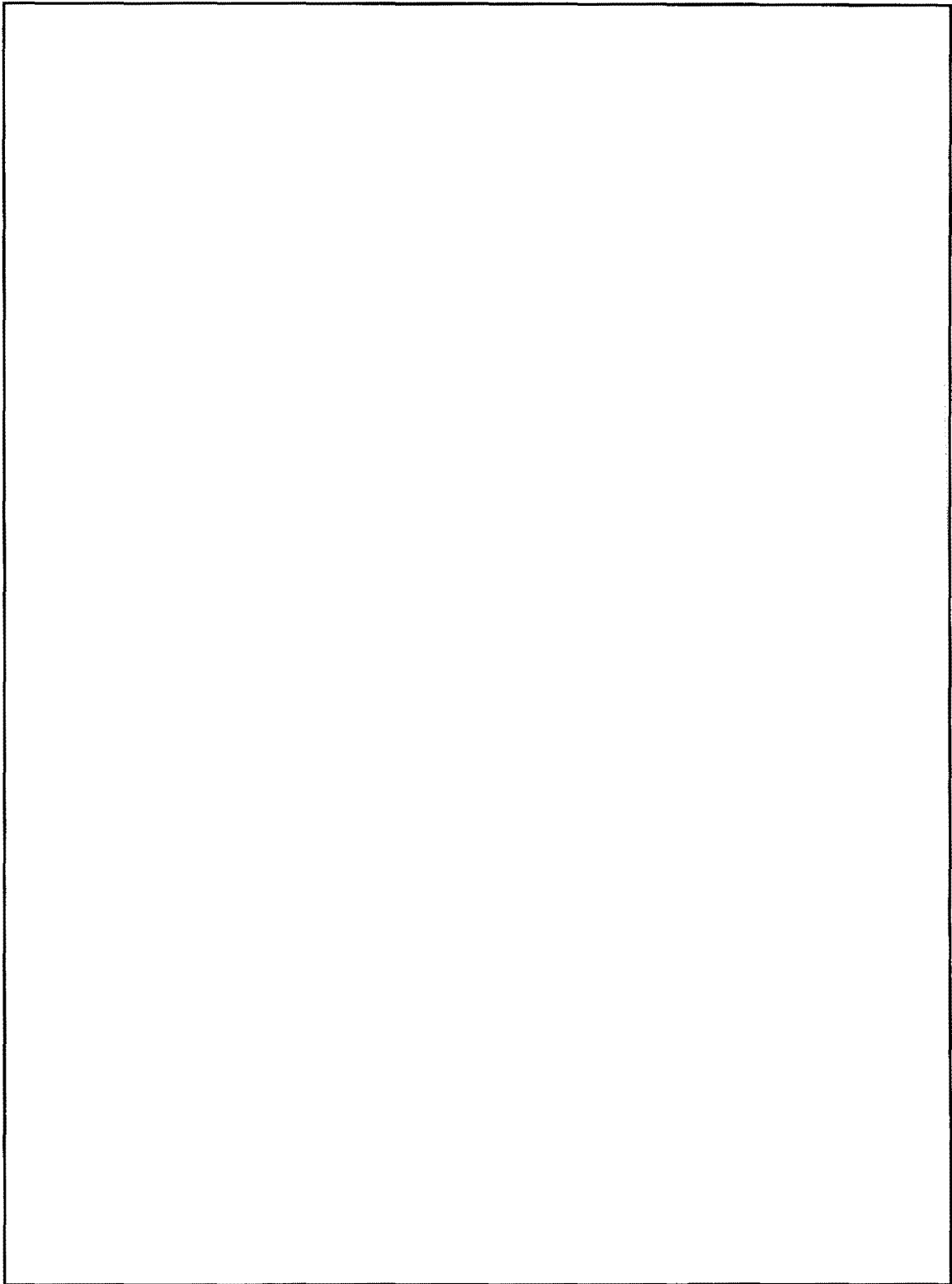
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~

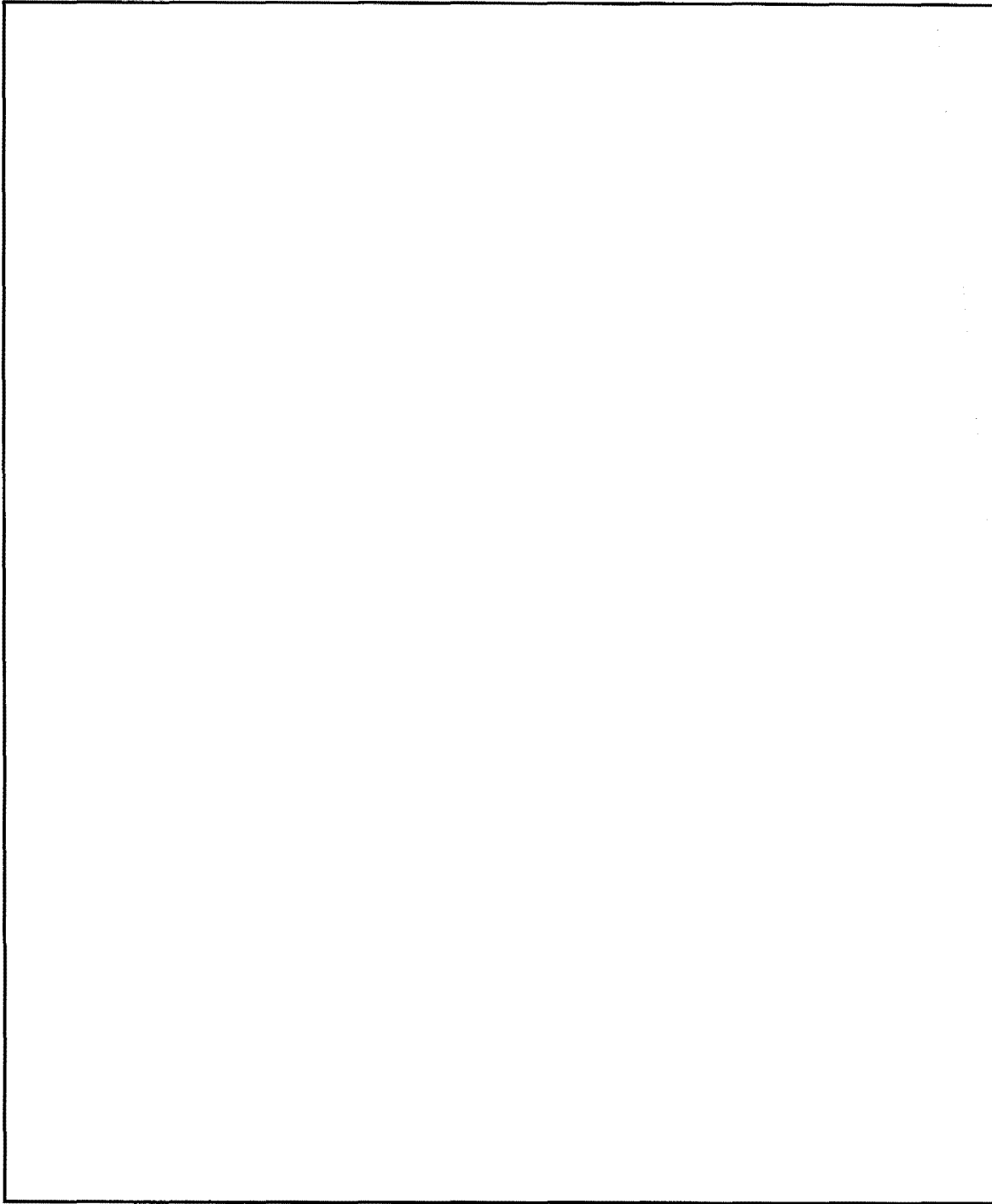


~~TOP SECRET UMBRA~~

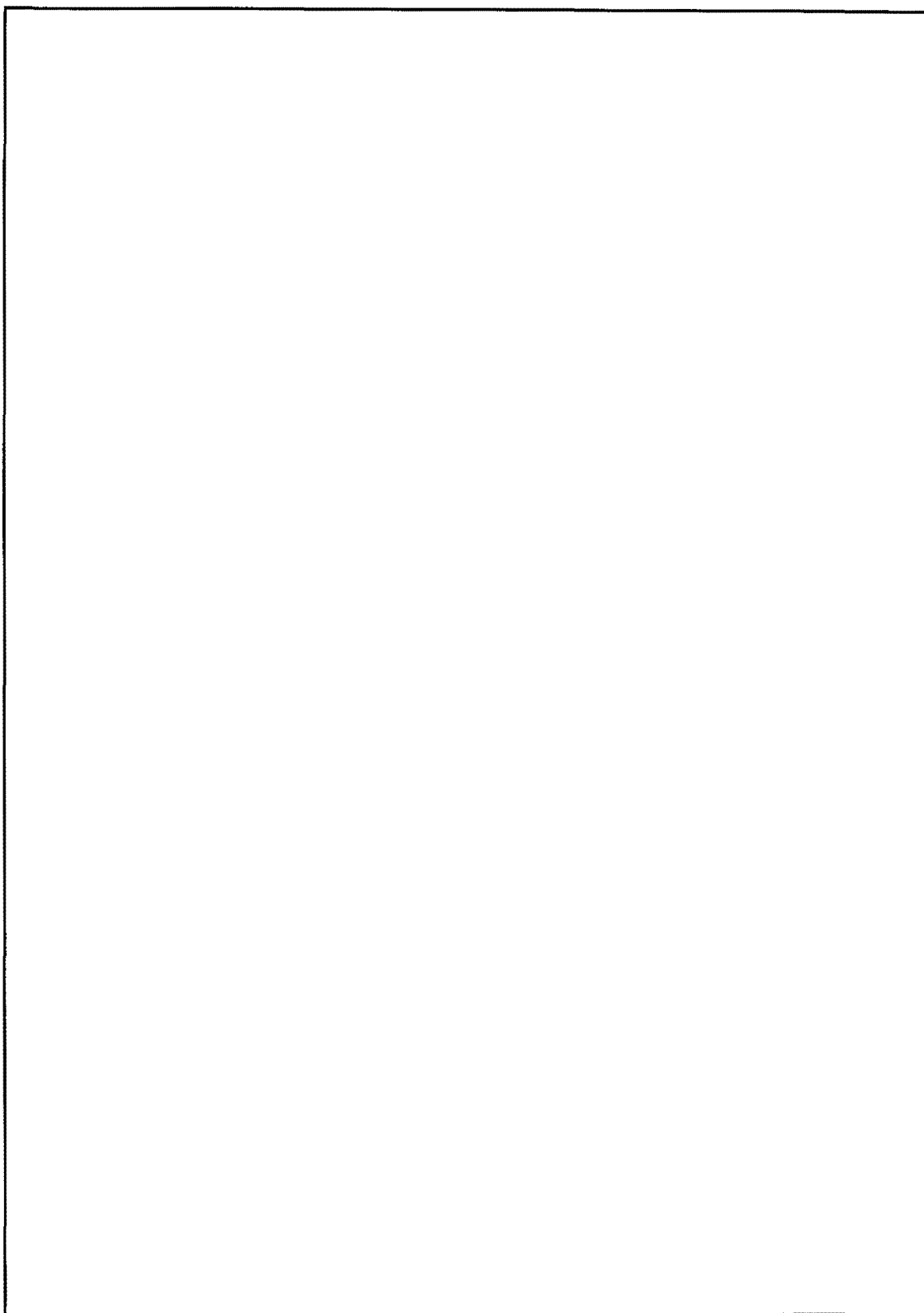
~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

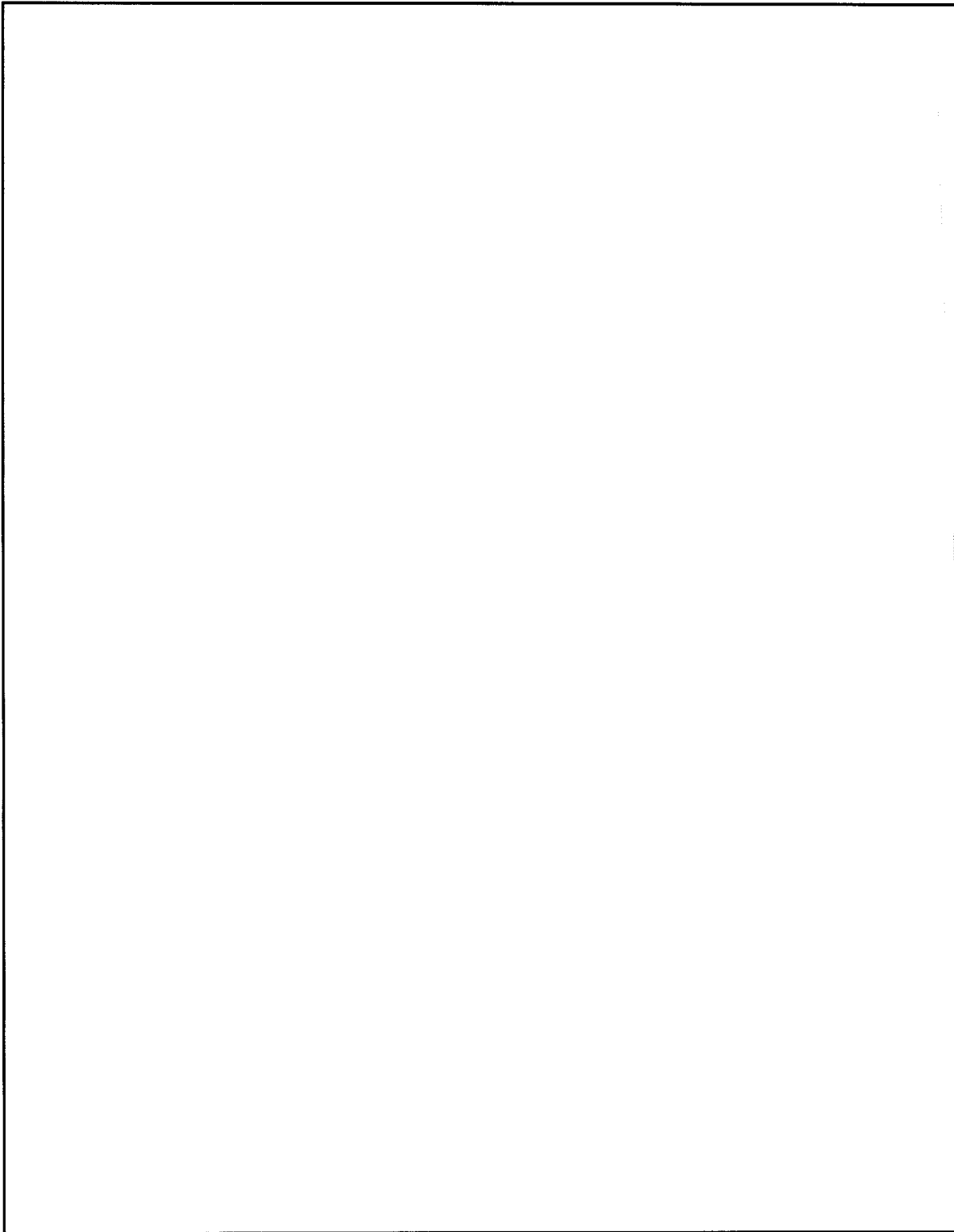


~~TOP SECRET UMBRA~~



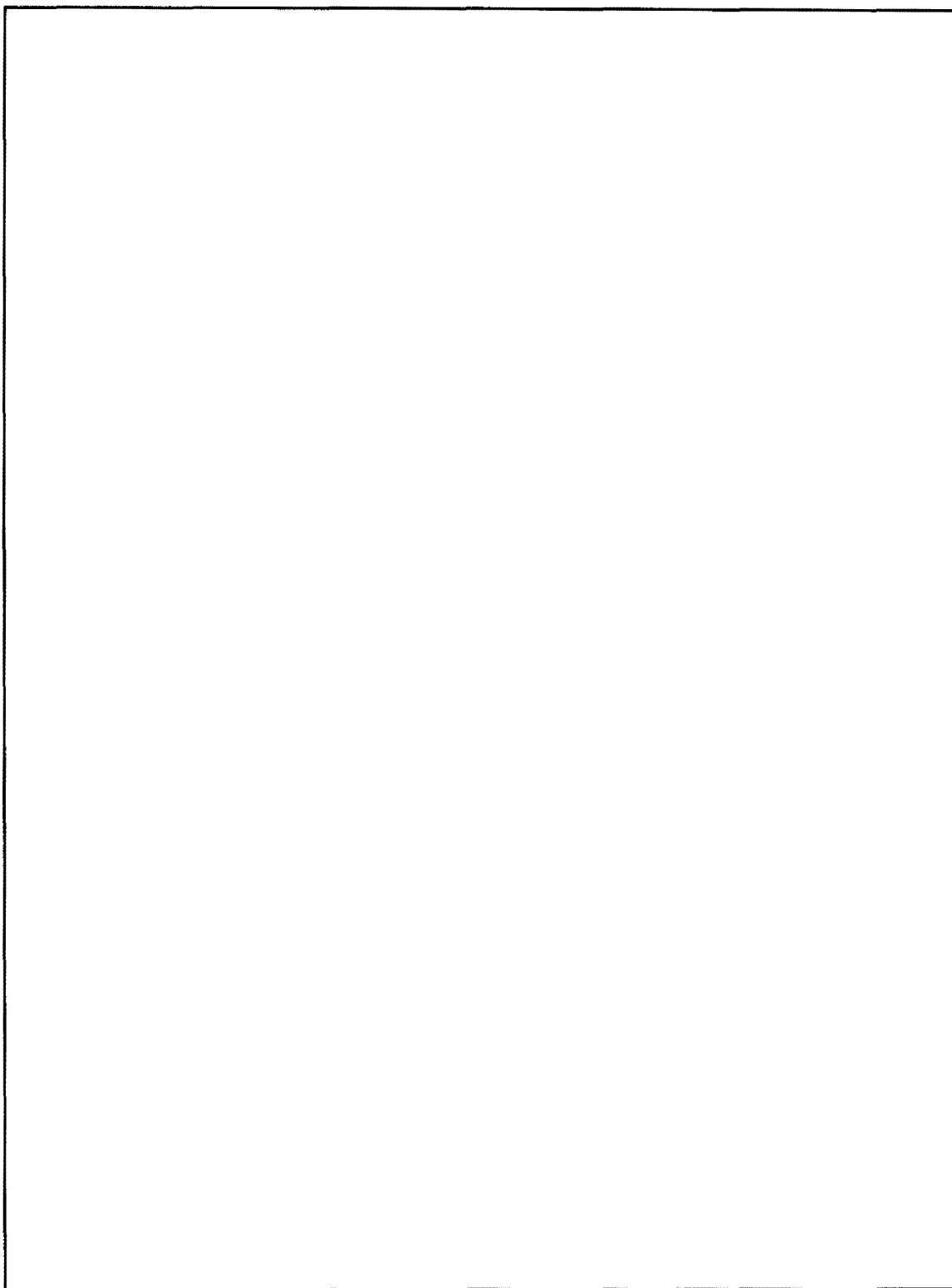
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



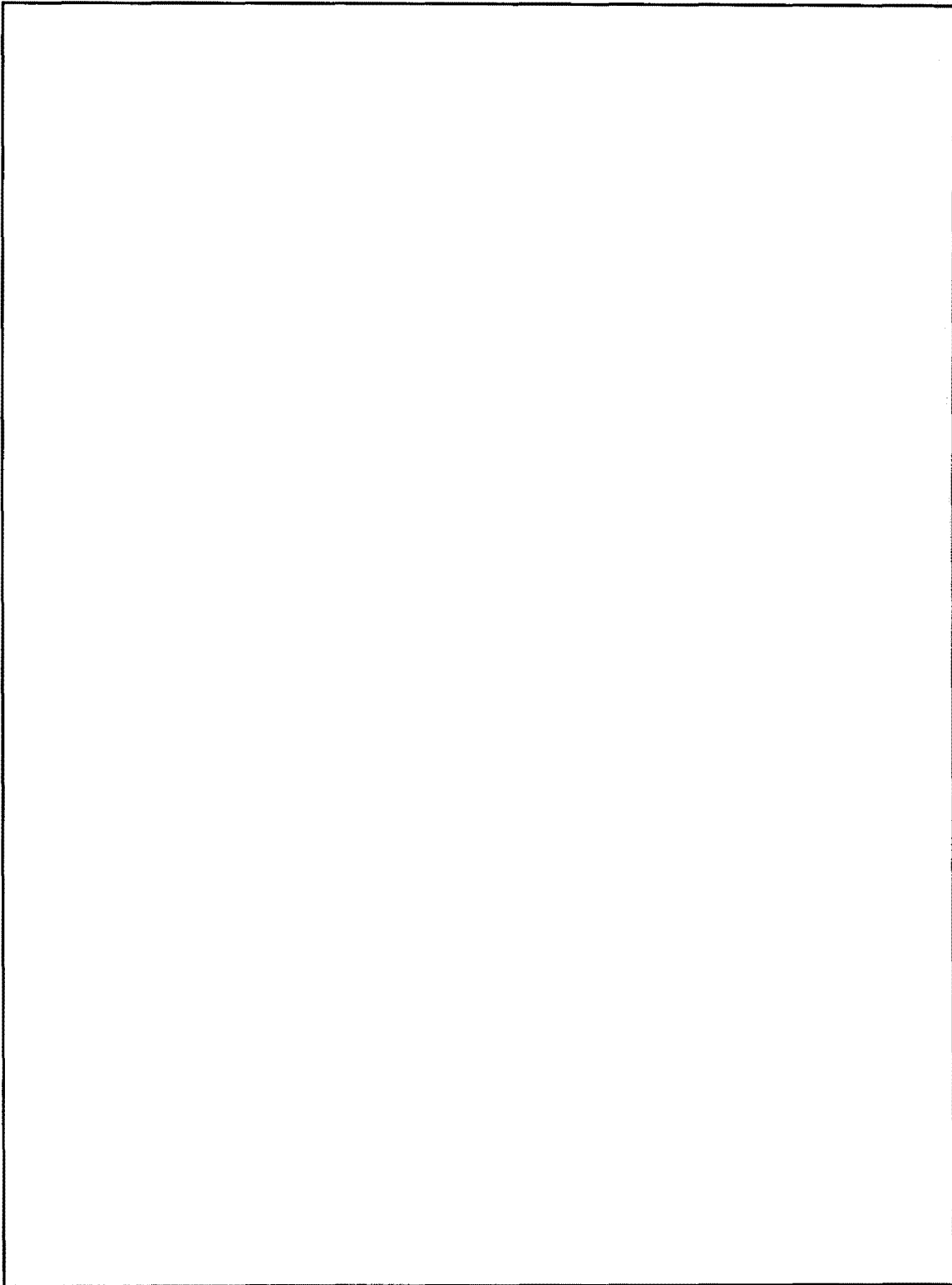
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



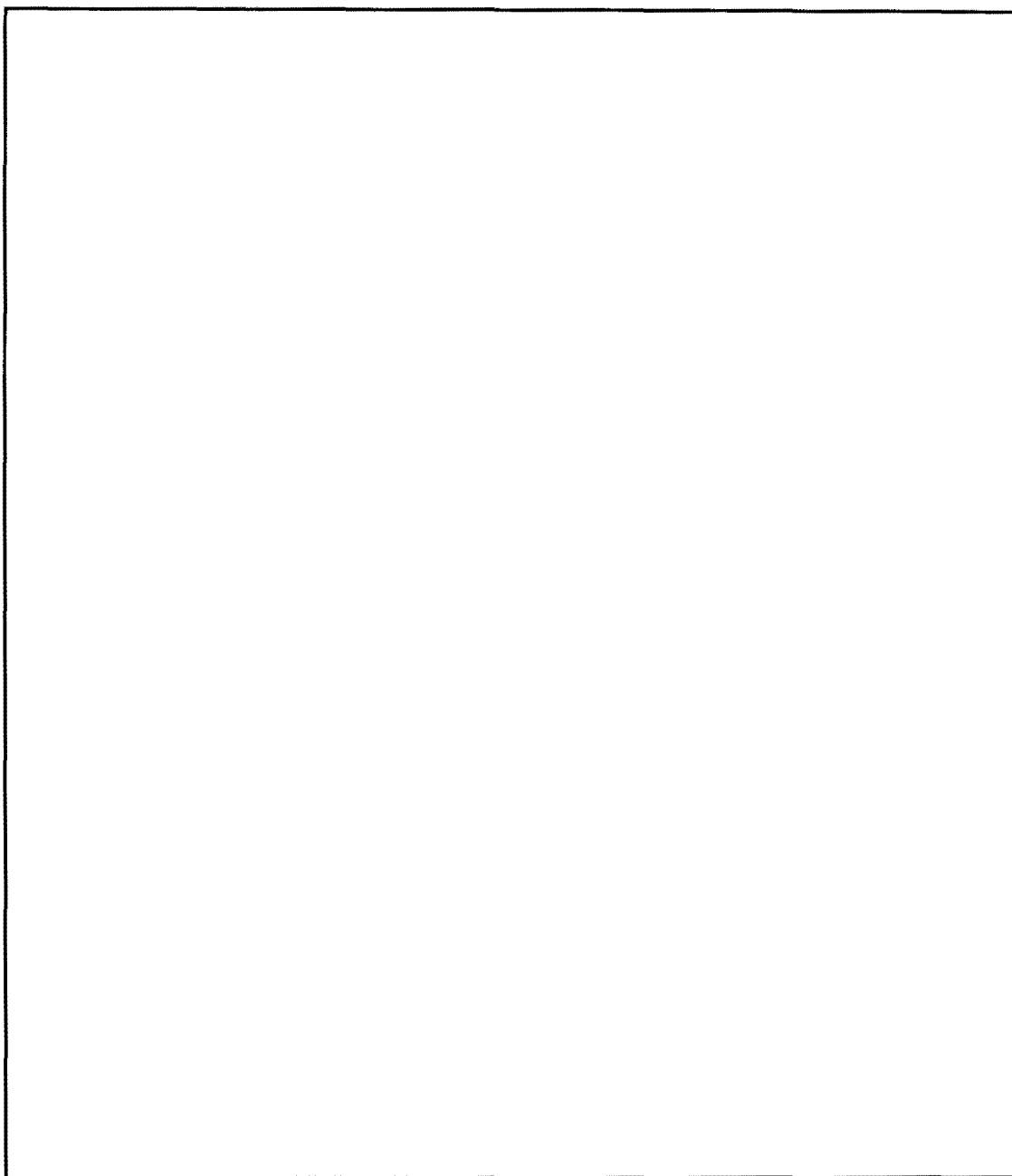
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



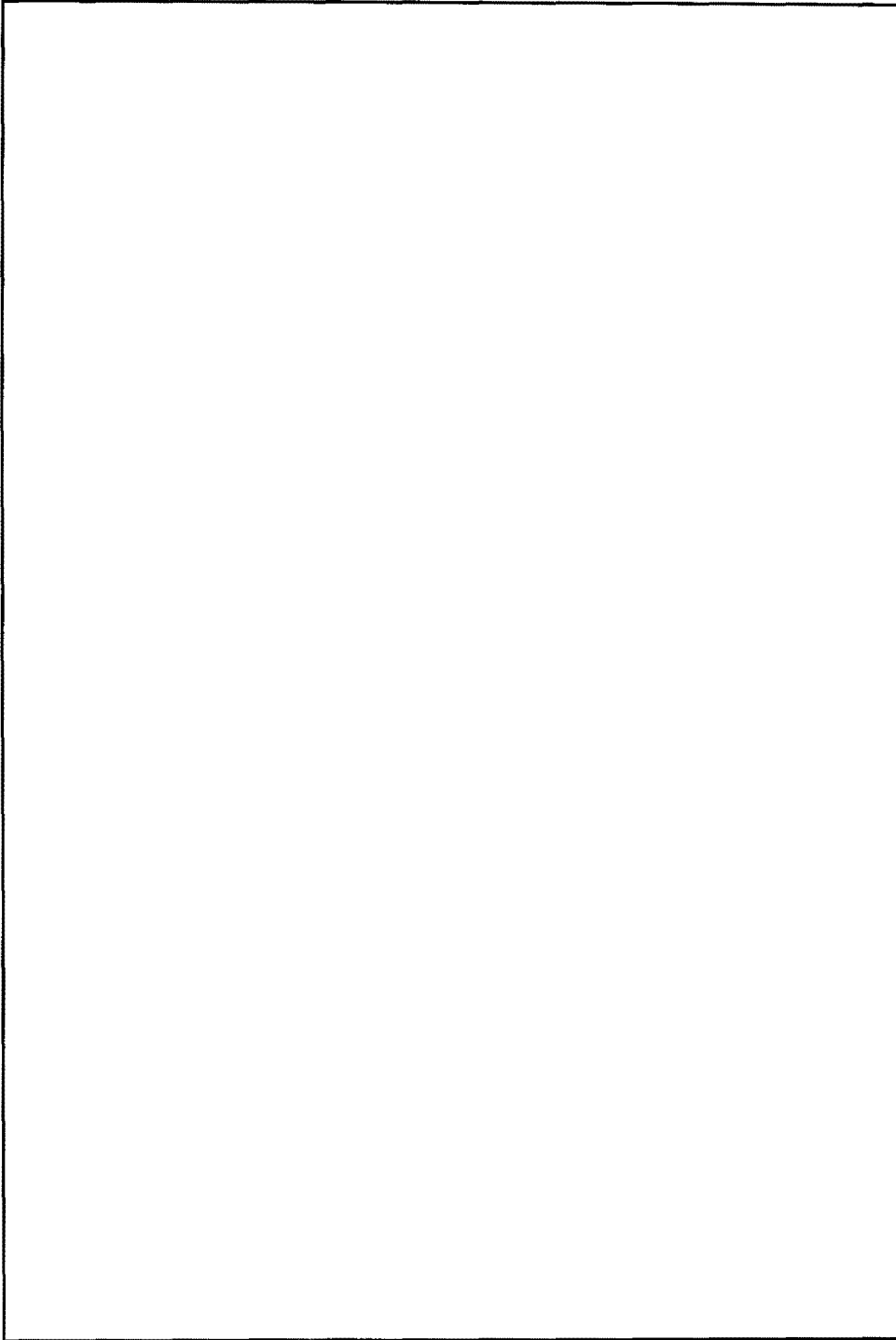
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



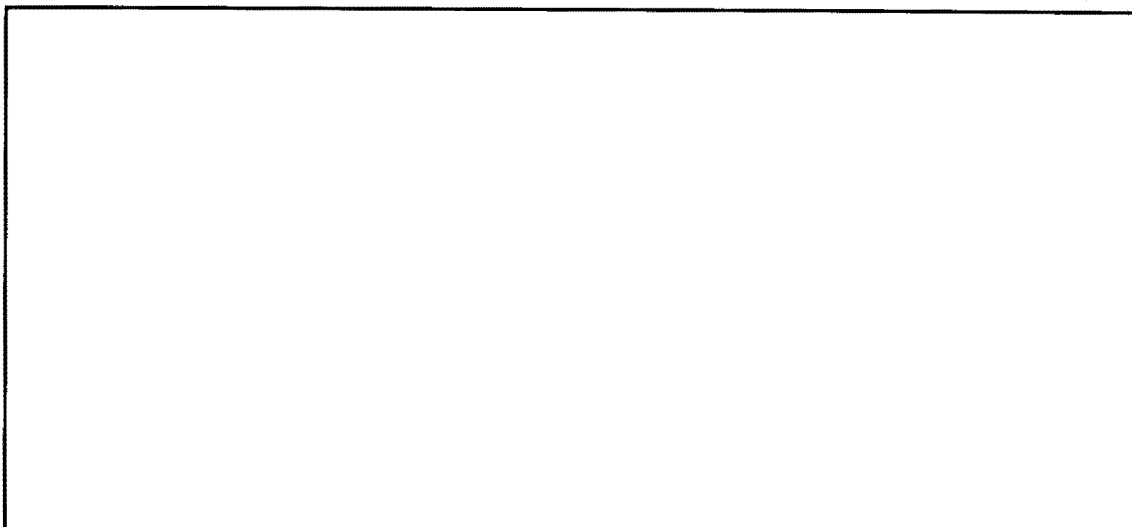
~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



~~TOP SECRET UMBRA~~

~~TOP SECRET UMBRA~~



These two examples show how TA crosses many disciplines; it is fluid, unstable in many ways and allows for wide ranges of different techniques and types of systems. Particularly, I have found math to be a very necessary tool in the diagnosis of crypto/TA problems. But perhaps we traffic analysts have neglected to record these applications of math. We have used the concepts of set theory, or VENN diagrams and the like, in research on callsign systems, procedure systems or radio frequency plans throughout all phases of TA discipline. We may, however, have called it by another name. Therefore the query: are we introducing new techniques or new names?

SUMMARY OF PANEL DISCUSSION

The panel discussion held at the close of the two-day symposium attempted to answer the question "where do we go from here?" The consensus was that traffic analysts are, generally speaking, becoming increasingly aware of a need for mathematical expertise and that mathematicians are on the threshold of discovering another field they can support. Most participants in the symposium agreed that they had arrived at a new awareness.

Ideas as to what further actions could be taken following the symposium were summarized during the panel discussion as follows:

- a. Mathematical training for traffic analysts
- b. Exposure of some mathematicians to TA problems and TA work areas
- c. Establishment of a register of mathematicians and traffic analysts and their respective skills
- d. Establishment of idea-exchange groups where problems may be described and analytic approaches recommended

~~TOP SECRET UMBRA~~

- e. Documentation of TA-math uses, ideas and developments
- f. Development of multidisciplined analysts
- g. Establishment of a focal-point (s) for TA-math ideas and problems

Let's Not Lose Our Skills

P.L. 86-36

One thing a middle-level supervisor in the Production organization realizes very quickly is that good traffic analysts are hard to find. Those traffic analysts with a skill in a specialized area such as frequency and callsign recovery are scarce. [REDACTED]

P.L. 86-36-36
EO 1.4(c)

[REDACTED] As indicated in the A/DDO memorandum, the underlying causes for this decrease in traffic analysts are the rapid change to automated methods of collecting and producing SIGINT, and the personnel limits imposed on the size of the NSA work force. Since NSA cannot hire personnel to fill shortages in critical skills, the traditional skills have been reduced to accommodate increases in linguists, signals conversion personnel, collection technicians, and data systems analysts and programmers. As a result, we are creating a static pool of traffic analysts, retarding the development of our analytic talent and altering the career-progression patterns of the traffic analytic work force. It is these effects that I wish to discuss.

The end of the Vietnam War, the subsequent tightening of purse strings and the resultant reductions in traffic analytic spaces altered the availability of traffic analysts. By limiting the hiring of new traffic analysts and not replacing those lost by attrition, the size of the analytic career field was set [REDACTED]. The immediate effects were minimal since the number of traffic analytic jobs was also decreasing, with the reduction of many of the timely requirements for information on Southeast Asia. Also helping to offset any immediate effects were the great strides made in mechanizing the traffic analytic processes during the Vietnam War. Efficiencies had been created and a degree of timeliness using methods of intelligence production never before possible had become routine.

The long-range effects probably will not be apparent until the late 1980s, but some symptoms are already beginning to appear. Our traffic analytic work force is getting old. Most of the younger analysts were hired during the 1960s and are now GG-11s or higher. Most basic traffic analytic work is now done by the military, either at the field sites or at NSA. No substantial group of young analysts at the lower grades is available for the future. The most aggressive analysts have already moved into management positions to further their careers. To aggravate what is rapidly becoming a bad situation, we have retarded the development of the younger traffic analyst. In the earlier growth days of our agency, a traffic analyst could grow in a specific target area, become recognized as an expert, and advance in grade and responsibility within his chosen career field. Today, the aggressive young analyst soon recognizes that his future is not in the technical side of the

traffic analytic business. To advance and achieve a modicum of success, he must move into management or to one of the critical-shortage skills. Consequently, we deplete our analytic talent base, and few people are left to form a nucleus for the future.

Those who are left usually have a sincere desire to remain in the technical side of the intelligence production business. Even those people are prodded by management to move into the more critical areas of data systems or linguistics. Since chances of promotion are mathematically better in these skills, many of the remaining talented young people do indeed transfer.

Those who remain face a slower career progression since the money provided for the special considerations given to the critical career areas reduced the total sum that would normally be equally divided among all those eligible for promotion. This means the traffic analyst must face stiffer competition for the promotions that are available, and ultimately his chances to achieve a position of leadership within the Agency are diminished.

As a result, probably in the near future, we are going to be faced with a severe analytic shortage similar to that which we now have with linguists. A more serious consequence will be the loss of analytic skills that can be learned only by years of experience. Specialists will be nonexistent, and major analytic recoveries will suffer. Although these problems can be alleviated to a degree by hiring from the SCAs and by programs such as the intern program, these are not immediate solutions. Unlike the data systems and, to a certain extent, the linguistic fields, our colleges and universities are not graduating many traffic analysts. It is a career field where experience is the best teacher.

To avoid future shortages, we should begin hiring some Traffic Analytic Technicians right now. These technicians could be hired out of high school at the GG-2 level and put through a program similar to that used for training linguists. Given the proper incentives, training, and experience, these people would be ready to take over the analytic work load in about ten years. If we fail to act now, we will have to react later, when our chances of success are fewer. Traffic analytic skills helped make our agency what it is today. Let's keep it that way.

P.L. 86-36

[While [] article was being prepared for publication, it was shown to [] Chief, Traffic Analysis, Office of Techniques and Standards, and he was asked if he would like to add any comments. He has submitted the following addendum. Ed.]

We could *also* hire ex-military traffic analysts, as we have in the past. This has the advantage that each "recruit" already knows what TA is, likes doing TA, and wants to make a career of it. That cannot be said of high school hires, and one must therefore expect a higher rate of "drop-out" than would apply to those already trained and experienced in TA (ex-military).

There must, of course, be some disadvantages to hiring ex-military traffic analysts. Otherwise, an agency as smart as we are would already be doing it.

LETTER TO THE EDITOR

~~(FOUO)~~ One of the advantages of being associated with a program such as the Traffic Analysis Intern program is the opportunity to watch the big picture develop. For three years I watched the traffic analysis career field being continually rescued from threats to its survival that came from whatever managerial wonders happened to be exerting influence at the time.

~~(FOUO)~~ While specific to the TA experience, these remarks point out the crucial aspects of organizational life: organization, corporations, or whatever, are really people. Hiding behind the various bureaucratic identities cannot disguise this reality. The "health" of the people affects the organization's. When people stop caring, growing, and producing, the organization ossifies and eventually fails.

~~(FOUO)~~ Personal career and organizational development are intimately linked. This agency has taken vital steps to facilitate career development, thereby assuring its future. This month's Career Development activities are a celebration of this process - in the past, present, and future - and its meaning for us and the Agency.

ROBERT HANYOK

AND FROM AN OVERSEAS READER...

P.L. 86-36

To the Editor, *Cryptolog*:

~~(S-CCO)~~ [redacted] article, "Let's Not Lose our TA Skills," (*Cryptolog*, March 1979), made an untimely appearance here [redacted]. The week before it arrived, several of the people here, due to return to NSA in the near future, had attended reassimilation briefings, during which the overstrength skills problem was not only raised but also given as a rationale for the reassignment of some to a different career field. Some of the moves were in the out-of-TA direction! It's easy enough to fall out of touch when headquarters is an ocean away, but it's even easier to be confused by the apparent contradiction between the M3 view of TA as an overage skill, and the view shared by [redacted] that TA is a field with an impending shortage crisis. Perhaps some of the fifty-odd traffic analysts now in excess could be used to head off the coming shortage. The concern about the effect of transfers out of TA and the lack of new blood is shared by many. This feeling was expressed here recently by visiting managers, who stated a need for analytic talent to work in the rapidly expanding world of multichannel communications.

P.L. 86-36-36
EO 1.4(c)

(U) What all of this may boil down to is the ever-present problem of individuals holding the title but not doing the analytic job, and the apparent inability of management to cope with this issue. Good traffic analysts are hard to find, and so may be talent in other

overstrength skills; the key may be quality vs. quantity, another problem that's always with us. If nothing else, [] article may make some managers realize that skill balance by the numbers is not an end-to-all-your-troubles elixir for reducing the work force, and that talent returning from overseas should not be regarded as a magic ingredient for such a brew.

KATHY BJORKLUND

Letters To The Editor

Last month Cryptolog printed a letter from Kathy Bjorklund in which she wondered why the view of traffic analysts as a vanishing breed, which has been expressed in Cryptolog by various people, is at such variance with the M3 view of TA personnel as an overstrength category.

To the Editor, *Cryptolog*:

(U) Since you were kind enough to ask me for a comment on Kathy Bjorklund's letter, I felt obligated to break out [] article, to which she referred. There are several hot spots in those two items and one that rises from them.

(U) First, *Cryptolog* has traditionally been an open forum, and I would not change that. But we who write for it from time to time are obligated to do some homework before we present opinions that aren't defensible. Or, maybe it's time to label fact and opinion so that readers can sort them out.

P.L. 86-36

(U) For Kathy, here are a couple of facts. While your briefing on reassimilation and career field overages was probably conducted by personnel or administrative people, they are not the ones who made the decision that TA is an overage field. As your Chief of Personnel Services, [] could have told you, M3 is part of Management Services (DDM), and it is a support or service organization that attempts to meet *requirements established by other Key Components*. In this case, it was Operations (DDO) telling M3 that there were overages in the TA field and shortages in the language field; it was DDO telling M3 to initiate the needed personnel actions, e.g., reassignments and hiring. Can you imagine the confusion if M3 went about willy-nilly hiring and reassigning people against no known requirement?

(U) Another fact is that George's article is mostly opinion. Now he has as much right as anyone else to have and express those opinions, but he knows he will get some arguments. For example, we not only lost good analysts when some TAs moved into management – we also gained some bad managers, although that's not a problem peculiar to the field of TA.

(U) But by and large, I doubt that you could find anyone who has to pick up the tab in billets or skills balances who would say we have any current or near future shortage of traffic analysts, or of TA technicians to fill the vacancies that will some day exist in the analyst ranks.

(U) Comparing real and present shortages in the language and computer arenas to "maybe" shortages ten years down the pike may not be a fair analogy. The computer and language shortfalls are there because we have added jobs or lived with vacant positions. In the field of TA that has not been, and is not now, the case.

~~CONFIDENTIAL~~

(U) Since most of our TA overages are at the technician level, I'm not sure I understand George's suggestion that we hire more technicians. But my not understanding is irrelevant - we aren't going to hire against a nonrequirement, at least not if I understand the way things work.

(U) Back to Kathy's letter for a final comment on her last statement: "...talent returning from overseas should not be regarded as a magic ingredient for such a brew." Given our selection processes, increased promotion points, and preferential treatment in assignments upon return for field people, I am a bit surprised that you believe there is an *intentional* negative attitude toward returning field people. And my opinion is that DDO, DDM, DDR, DDT and DDF would be equally surprised.

(U) Regards to you, Kathy. Congratulations to you, George. And, Dave, whenever you want an opposing view on almost any subject, please give me a call.

DAN BUCKLEY

To the Editor, *Cryptolog*:

(U) I read Kathy Bjorklund's letter with a feeling of depression - because what she says is all too true. The bodies-and-slots, or bean-counting, approach to personnel assignments is not one that is conducive to the continued development of the technical work force of the Agency.

~~(C)~~ During the skills requirement forecast of 1973, the career panels were asked various questions on personnel development covering the period FY74 through FY79. Questions such as the following were asked:

- What effect will new or emerging technology and modernization of cryptologic operations have on the skills under the purview of your career panel?
- Do you anticipate a need for developing multiskilled specialist, and if so, which skills or combinations of skills will be required?
- Will the need for specific skills (TA, CA, etc.) decline or increase?

(U) I don't know what happened to the results of this poll, since current personnel planning does not seem to reflect them, but rather continues to be based on projections of the current work force: How many people do we have in such-and-such COSC? Well, then, if we have that many, and if we are getting the work done, then that must be the right number. So let's just straightline that number for the next four fiscal years. Obviously, this approach is the basis for faulty TDs, since it makes no allowance for any shifts in requirements brought about by shifts in targets or other considerations.

~~(C)~~ Let's see how this works. At the moment Traffic Analysis is carried as an overstrength skill in A Group. The TACP [Traffic Analysis Career Panel] has two interns due to graduate this month. On the basis of their backgrounds, experience, the panel's recommendations and their own preferences, these interns should be assigned to A2. But

~~CONFIDENTIAL~~

~~CONFIDENTIAL~~

the thought of placing them in an overstrength element is enough to give the bureaucrats heartburn.

~~(C)~~ The placement of overseas returnees is similar. P41 attempts to assign personnel holding A Group overage skills to B, G, V or W; only a few of this year's returnees have been assigned to A.

~~(C)~~ In short, P41 and M3 will almost always stand in the way of any assignment to an overstrength element. I have accused P41 of approving TA intern placements using the bean-counter approach. They deny this vehemently yet state in writing, in a memo to Chief, M3:

[REDACTED]

P.L. 86-36
EO 1.44(c)

A fair share - by the numbers!

(U) [REDACTED] quoted by Ms. Bjorklund in her letter, are correct: the number of traffic analysts is dwindling. Part of this is attributable to the reasons cited above. Another rank-thinning factor is age. Almost twenty percent of the people in COSC 1411, Traffic Analyst, are over fifty years of age; less than two percent are under thirty.

(U) What is the solution? As I see it, it is twofold. An immediate measure would be some directed assignments. This would include the identification of personnel holding a given COSC in an overstrength area, but not performing that function, and making appropriate readjustments, such as transfer, retraining or reclassification. It would also include the placing of overseas returnees in areas where their skills are most needed, even where they might be a temporary overstrength condition.

P.L. 86-36

(U) For the longer term we must nurture the TA intern hire, insuring that we have at least six to ten coming in each year, and placing them in the work force where they will produce for the Agency regardless of numbers or quotas.

~~(C)~~ Let me quote from an old-time member of the TA corps:

How long does it take to build a professional traffic analyst from zero? If it takes, say, five years, then we are betting that whatever the situation is today, it will be the same five years from now. And what we are betting with is the Agency's reputation for adapting to fast-breaking changes in the world situation.

~~(C)~~ The TA intern program can and does build a professional traffic-analyst/reporter from zero with a very solid understanding of the interrelationships of the other cryptologic disciplines. The annual hiring of a few bright people - recent college graduates as well as former military analysts - should solve the problem of being able to find good traffic analytic talent in the future.

[REDACTED]
Executive, TACP

~~CONFIDENTIAL~~

~~CONFIDENTIAL~~

To The Editor, *Cryptolog*:

(U) Having had long, and often passionate, discussions with both Mr. Buckley and Mr. [] on the subject of Traffic Analysis "overages," I was surprised to see that both gentlemen missed several important points (see Letters to the Editor, *Cryptolog*, July 1979).

P.L. 86-36

(U) First, I think (please note, opinion) there are a fair number of individuals who hold TA COSCs who are not doing TA work. One example: a branch chief is a manager, not an analyst, yet the analytic COSC is often held.

(U) Second, quality is not an element included in the "bean-counter" approach; it's the number of bodies bearing the title that are counted. The fact that one quarter of the bodies may be doing three quarters of the work does not enter in.

~~(C)~~ Third, we seem desperately intent on trying to "professionalize" (read, certify) ever-increasing numbers of people in overage fields. The problem with this is twofold:

- What is the purpose of creating many more professionals than we have professional billets? What is the purpose of allowing second certifications in overage fields?

- Why are we tutoring professional aspirants? Individuals of professional quality will, like cream, rise to the top by dint of personal study and research. They will do so simply because it is in their nature to excel, not because they were pushed, prodded or carried through the test. (Please note that "top" refers to one's ability to perform, not to one's promotion record.)

~~(C-CCO)~~ I agree with [] - we are losing our TA skills (See *Cryptolog*, March 1979). It is evident in the simple fact that many case analysts do not log intercept in such a manner as to see a simple rota. It is reinforced by the fact that on our professionalization exam the net reconstruction section is pre-logged and fairly easy. The difficult section is the crypto-TA part, and few traffic analysts are charged with this type of responsibility in their daily jobs.

(U) Suggested solutions for eliminating the overage and upgrading the quality are fairly simple. Ensure that only those doing traffic analysis carry that COSC; let's start calling managers managers. Next, devise a weighting system so that a manager can state for the TD that he needs three topflight TAs, or five average, or ten below-average (or some combination thereof) to accomplish his mission. And, finally, if we feel compelled to tutor someone, let's begin by tutoring those most in need so that they may become proficient in their current jobs.

(U) Years ago achieving professionalization was akin to winning a place among the elite. Today is simply a means of proving oneself average. Elitism is not only dead, but is scorned as well.

(U) Let's put the emphasis back on the quality of our human resources.

[]

~~CONFIDENTIAL~~

~~HANDLE VIA COMINT CHANNELS ONLY~~

Simplicity in Color

C. GAROFALO

The conduct of crypto-TA studies involves the requirement to differentiate between qualitative levels or categories, and the use of a simple color scheme facilitates such differentiations. Any color scheme to be used in an analytic effort of large volume must be

- simple in nature,
- skillfully chosen, and
- systematically applied.

An analysis of the various colors available was made to determine which are practical and most suitable for the task. My experience has demonstrated that six different colors are most practical. These are black, green, red, blue, purple, and brown.

- Black – carbon (number 2) medium degree of hardness and density, suitable for both erasure* and longevity properties.
- Green – is the weakest, that is, of lightest density. All of the other colors under consideration superimpose on green quite readily.
- Red and Blue – are of equal boldness; either can be easily superimposed on both black and green.
- Purple – a bolder color that can also be produced by a superimposition of red on blue or blue on red.
- Brown – is considered to be the boldest of all colors as well as the most exclusive and conclusive.

Colors manufactured by different commercial firms vary drastically in hardness, density, and coloration; it is desirable, therefore, that having started with a particular brand to continue with that brand and not intermix brands to ensure that distinctiveness and clarity are maintained.

A color ladder may be displayed graphically as follows:

<u>Color</u>	<u>Level</u>
BROWN	5
PURPLE	4
BLUE	3
RED	2
BLACK	1
GREEN	0

* For Black erasures, a medium hard rubber erasure is adequate. For color erasures, best results are obtained with a typewriter eraser.

Another way to express the levels of the color ladder is by these definitions:

- Brown – The ultimate in degree of trueness, not to be questioned. May also represent captured, compromised information or its equivalent.
- Purple – High in degree of reliability; may be used as a substitute or companion for brown where special conditions of clarity or distinction are vital to the problem. Primarily useful as a final ordered and oriented intermediate enjoying the same general stature as brown.
- Blue – A relative base value having a significant bearing on the state of recovery.
- Red – A base of lesser value or no relativeness (completely arbitrary), a first step necessary in any endeavor.
- Black – To record or log information as it appears in its earliest or original form.
- Green – An envious color reserved completely for suspected garbles, projected or expected but unobserved values (not proven but highly suspected as being correct). Used to alter a meaning or information without obliterating the original (black) form.

Once a color scheme is established for a given problem, maintenance of color discipline is mandatory in order to achieve uninterrupted and unambiguous continuity. Discipline is also of great benefit to management in that it ensures that redistribution of analytic personnel can be effected with minimum disruption to the overall effort.

Postscript:

Colors:

Green	symbolized hope to the ancient Egyptians; in the Middle Ages was supposed to be good for the eyes; to the Mohammedan, a sacred color; in modern times has come to signify envy.
Purple	to the Tyrians and Romans, a purple robe or band of purple signified authority; became a symbol of majesty to the Romans.

The Impact of ARDF on Traffic Analysis

ALLEN L. GILBERT

The Vietnamese conflict and techniques for collection of signal intelligence developed and employed in that arena have profoundly influenced the traffic analytic approach to the Vietnamese Communist problem. One of the most effective techniques employed on a large scale in Vietnam has been Airborne Radio Direction Finding (ARDF). ARDF, in addition to revolutionizing the direct support of tactical units through timely and accurate locating of enemy units, has almost reversed the traffic analytic approach to maintaining continuity and developing new targets in some areas.

Traditionally, the traffic analyst is faced with the problem of reconstructing a communications complex through recovery of callsign and frequency systems, message externals, schedule activity and those rare compromises made by enemy communicators. This route usually requires close scrutiny and cataloging of the elements of intercept through an extended period of time, with the hope that a transmitter location will be compromised or that medium-range direction finding will suggest a location for the activity. ARDF provides a location

The availability of ARDF on target transmitters considerably shortens the period of development for new activities and provides almost instant continuity on targets effecting communications changes.

P.L. 86-36
EO 1.4.(c)

In Vietnam, the concept of ARDF tasking provides coverage in all areas of hostile troop activity.

When aircraft are deployed to a target area daily, the recovery of the signal environment in the area builds rapidly.

Certainly, all other elements of traffic analysis must then come into play to establish case notations and identifications. ARDF alone does not solve the problem, but what an advantageous beginning it provides!

Barometer - Readers' Comments

Traffic Analysts: Write it Down Now!

In approaching a traffic analytic problem for which you have responsibility, have you ever wondered how any of the major callsign systems, procedure tables, address systems, or other complex phases of transmission security were solved, or what significant characteristics contributed to their solutions? It is highly unlikely that you could find a detailed written description of the analytic effort involved in any of these solutions, so you would have to go to the individual who had been most intimately concerned with it, the individual recognized as having had the responsibility for providing the technical direction toward the solution. He is *the* person who was confronted with the problem, recognized its difficulties, determined analytic approaches, made mistakes, went up blind alleys, made critical decisions, directed other analysts, and generally supervised the entire effort. Assuming that he is available, you could ask him to describe the entire process. There is a good chance, however, that because of passage of time or lack of thorough documentation, or both, he would find it difficult to reconstruct in detail the path and the thought processes that were followed in reaching the solution. Thus, valuable expert knowledge, which might be useful in the solution of current or future problems, would have perished.

How can the loss to posterity of expert knowledge in such a situation be avoided? Very simply, the analyst who is pursuing a specific problem should *keep daily notes* on his analytic activities, his thoughts, his approaches, his mistakes, his conclusions, etc., and then, upon arriving at a solution (or failing so to arrive), *prepare a detailed written record* from these notes. The record, to be of most value, should describe the entire effort from a strictly chronological viewpoint, so that the influence of hindsight might be avoided.

If this practice had been consistently followed in the past, the Agency would now have in permanent form a wealth of analytic expertise that led to past successes but that is now unfortunately past accurate recall. The practice should be followed by every analyst who is pursuing a specific problem, whether it is large or small: he cannot predict where his efforts may lead, and he may be making an important contribution to the overall body of traffic analytic knowledge.





P.L. 86-36

How Many Angels Can Stand on the Head of a Case Notation?

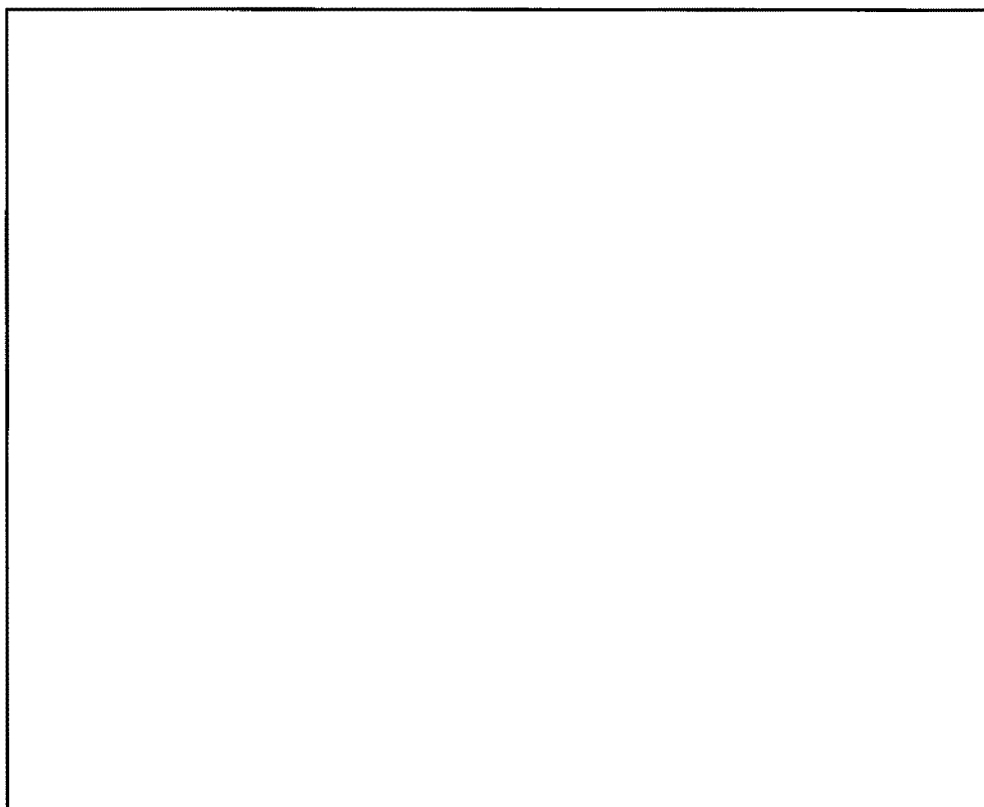
P.L. 86-36



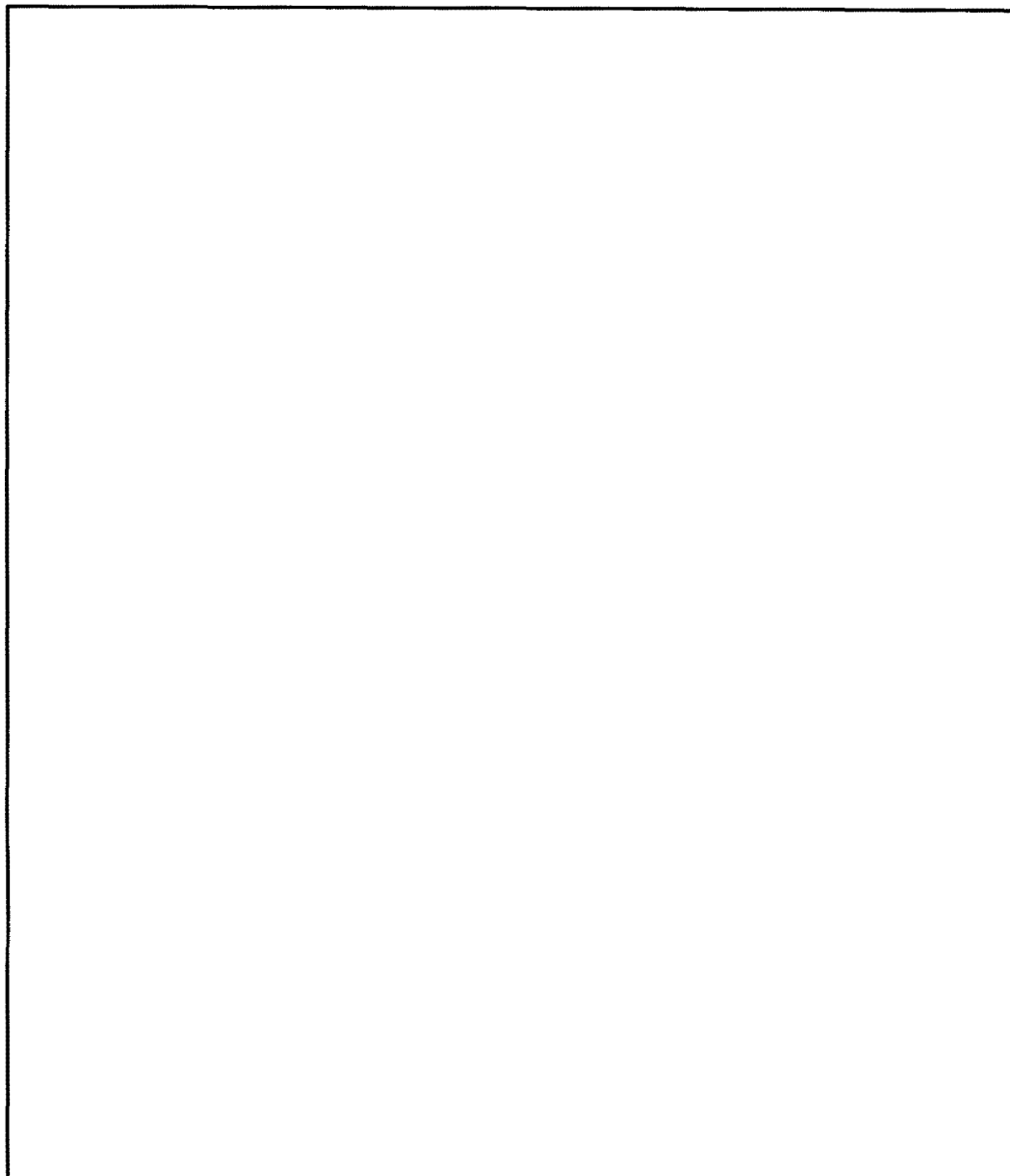
~~(S-CCO)~~ Most persons can recognize a case notation as such, even though formats differ. Few can *explain* one unless it is one worked with daily. This is because notations contain over fifty categories of information, and each of these has sub-items ranging from four or five up to about 800. In short, there must be over two thousand coded meanings. How can this happen? In addition to the unstandardized

 P. 1868636
 notations, the digits of the so-called "serial" are actually used as codes, singly or in combination, for the various categories shown below. Sometimes a given digit may have two meanings within a given notation. Also, the information selected for representation varies from target to target, from trigraphic source to trigraphic source, and even within the latter. E.O. 4.4(c)

Some of the Information Contained in Case Notations



~~SECRET~~



~~SECRET~~

~~HANDLE VIA COMINT CHANNELS ONLY~~

The 1980s and Beyond

Traffic Analysis: A Current Perspective

CW2 JAMES HOPPER, USAFS

P.E. 1868636

E.O. 4.4(c)

~~(C)~~ Traffic analysis (TA) today is beset by a variety of problems affecting the analyst, not the least of which is a lack of understanding concerning the separate functions which comprise traffic analysis. Contrary to popular definition, the traffic analysis field of today encompasses much more than the mere analysis of communications externals. Traffic analysis consists of six separate functions:

- Collection Management
- Collection Support
- Processing
- Analysis
- Reporting
- Evaluation

These are intended to be mutually supportive of each other. However, because both managers and analysts lack understanding of the interrelationship that exists between and among these functions, they all too often are performed at cross purposes, with one function inadvertently affecting the success (or failure) in a follow-on activity. In these days of meager collection and analytic resources, it is imperative that an understanding of the relationship bonding the six functions of traffic analysis be fostered at all levels to ensure the complete and proper use of these resources.

~~(C)~~ To function effectively in the traffic analysis world of today, the traffic analyst must have a rudimentary knowledge of several areas in addition to communications externals. Among these areas are

Collection Management, including the Collection Objective Performance Evaluation Systems (COPEs) and its follow-on, Collection Evaluation System (CES) statistical reporting;

Collection Support, including knowledge of working aids available to assist in collection and identification duties, as well as of target characteristics that can support both current and future collection operations;

Processing, including Automated Data Processing (ADP) routines available to assist in manipulating the intercepted data for subsequent analysis;

Analysis, including sufficient knowledge of the various traffic analysis techniques and target characteristics that develop information to satisfy consumer requirements;

~~CONFIDENTIAL~~

Reporting, including the criteria and appropriate reporting vehicle for providing information to satisfy consumer requirements;

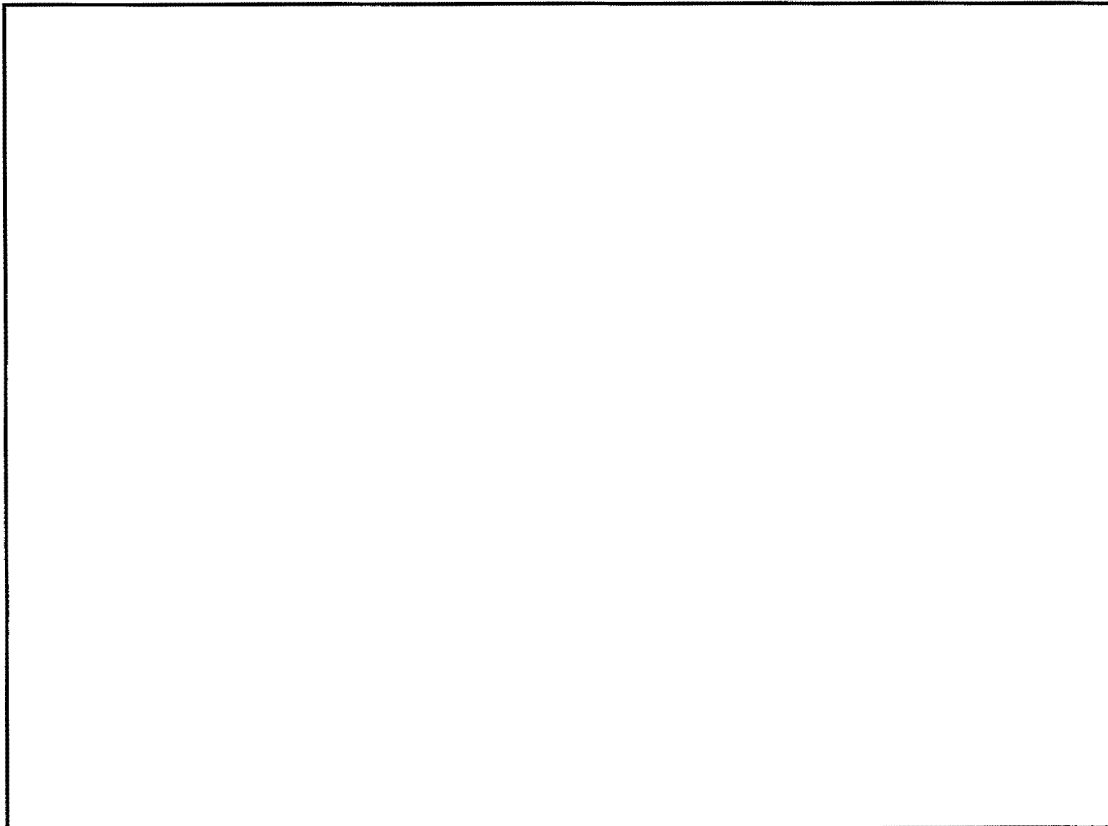
Evaluation, including the techniques available to constantly monitor and compare tasked and acquired collection so as to redirect these resources to areas of the target's communications that are susceptible to exploitation; and

Perspective, including a thorough knowledge of the target's order-of-battle (OB) structure that underlies the target's communications, as well as a thorough understanding of the target's past history from which to view changes in the target's communication habits.

A lack of knowledge in any one of these areas can ensure failure in attaining the ultimate goal of consumer satisfaction. Additionally, an incorrect managerial emphasis on any one function over another will also ensure failure in attaining this ultimate goal.

COLLECTION MANAGEMENT

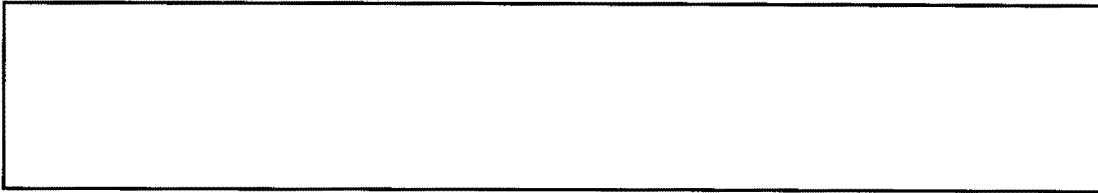
(U) The first function to be encountered by the traffic analyst is Collection Management. It is at this point that the consumer's requirement is translated into specific collection tasking to acquire the needed information.



P. P. 1868636
E.O. 4.4(c)

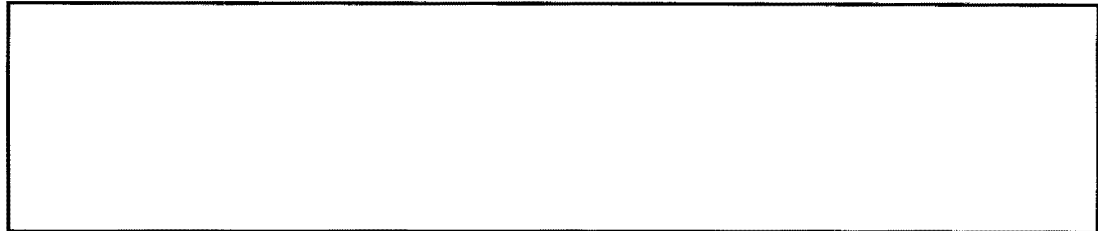
~~CONFIDENTIAL~~

~~HANDLE VIA COMINT CHANNELS ONLY~~



COLLECTION SUPPORT

Once the tasking is accurately developed and transmitted to the field site, the next function, collection support, determines the success or failure of a field element in satisfying its tasked requirements.

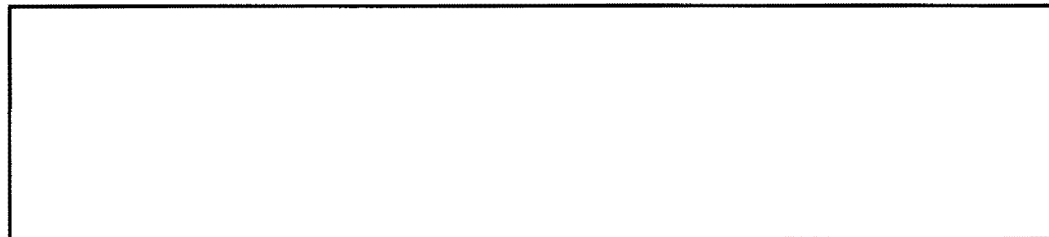


P.L. 86-36
EO 1.4.(c)

(U) At the local level, field site managers usually are faced with a decision of where to use their most talented personnel; an effective collection support element or an effective analysis and reporting effort. Unfortunately, all too often the choice is for the A & R effort because of the high visibility of analytic breakthroughs and consistent product reporting.

(U) This decision on the part of the field manager leaves the collection support effort with poorly trained analysts, generally those with the least experience in TA duties. These individuals are then placed in the fast-paced environment of the collection floor, an environment that breeds confusion and is the least likely place to learn the functions of traffic analysis.

(U) In addition to the assignment of poorly training personnel at the local level, at the national headquarters level technical support material is seldom provided in a manner that facilitates its use in the field. Most often this material is seldom provided in a bulk manner, which dictates tedious manipulation in the field to extract desired information. Yet this same manipulation is often performed with the national headquarters on a daily basis through various computer routines but is not made available to field sites. Thus, while a steady and voluminous stream of technical support material is provided to the field element, it is seldom provided so that the field can easily use it, and it is often too unwieldy to be used on the collection floor, with all its hectic activity.



Of these, the misidentification rate should be of the greatest concern since it dramatically impacts on the processing, analysis, reporting, and evaluation functions that follow the collection support function. A poor case (notation) identification rate in the collection support function means that all case identifications received for processing must be verified to ensure that the correct notation is applied, a time-consuming process for an individual already strapped for time in performing his or her other processing duties. Failure to verify the correct identification on all cases can have disastrous effects

- on processing, necessitating complete revisions of case history records;
- on analysis, which attributes incorrect technical operating characteristics to the wrong case;
- on reporting, which provides false information to the consumer; and
- on evaluation, which attributes tasking satisfaction to the wrong target.

To ensure the most effective collection support effort possible, field managers must develop an operational system that exposes all elements of the assigned analytical work force to both collection support duties and desk analysis duties. Ideally, such a system, which rotates the individuals between these two worlds, will ultimately increase the expertise levels available in both areas. Likewise, at the national level all methods of developing collection support material for field elements must be closely scrutinized to ensure that a minimum of manipulation is required on site to use the data.

PROCESSING

~~(S)~~ The next function of traffic analysis, processing, often assumes a life of its own. From the analyst's standpoint, processing (or, to use a more common term, "logging") quickly becomes associated with drudgery. The fact that processing is accomplished merely in order to organize intercepted data for subsequent analysis is quickly forgotten. Processing then becomes a mindless transfer of data from one medium, the raw intercept, to another medium, the casebook or database. Further, this attitude quickly leads to overlooked items of significance that otherwise would require immediate analytic or reporting attention.

(U) From the manager's standpoint, however, processing quickly becomes the key productivity measurement because of the statistical data available from database maintenance inputs. Since the underlying desire is for an accurate database from which to extract material for subsequent analysis, and from which to drive any automatic collection support vehicles that emanate from the database, the manager's concern becomes one of concern over database maintenance accuracy and volume. Unfortunately, this concern for accuracy normally translates into a concern over the correct formatting of the input rather than over the validity of the information itself. Likewise, the volume concern translates into a meaningless "body count" that is used to spur further productivity.

~~CONFIDENTIAL~~

P. 18686636
E.O. 4. (c)

[REDACTED] poses should be deleted from existing requirements. For example, the processing of voluminous message serialization data should not be imposed on the affected analyst if no one intends to perform a study of this aspect of the target's communications. The processing requirements should be continually geared to the voids of the target's communications that need to be exploited, and not to those characteristics that are commonplace and seldom change, as these voids are reduced through exploitation.

ANALYSIS

[REDACTED]

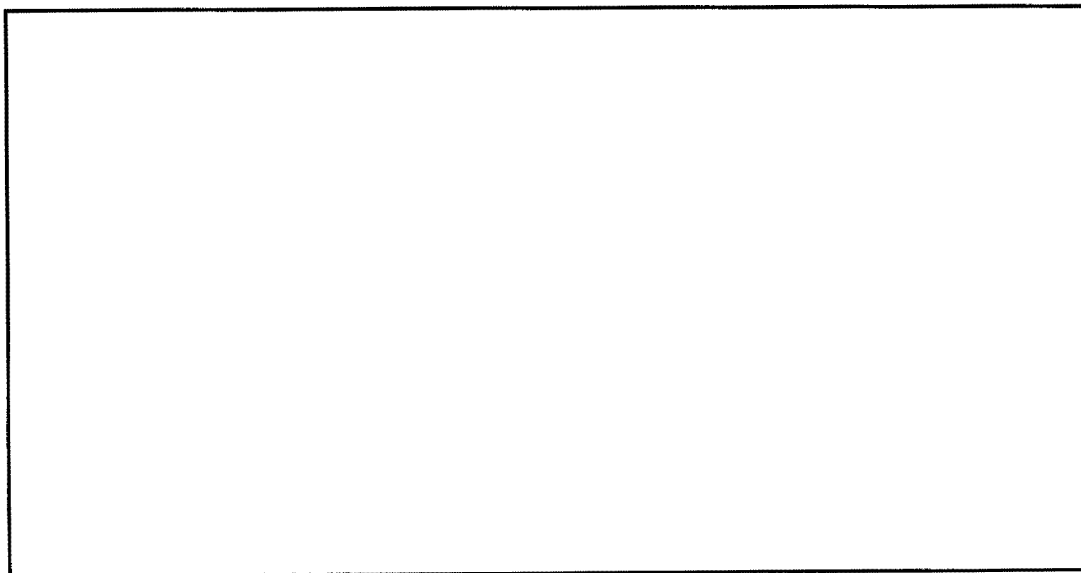
~~CONFIDENTIAL~~

[redacted] underlying military force must be the first priority of the traffic analyst. Only through such mastery can an analyst understand the communications facing him or detect changes in those communications that can be attributed to something reportable. In addition to ensuring that the analysts are knowledgeable of the structure of the military force, managers at all levels must ensure that the analytic studies to be undertaken first have the potential of either satisfying established reporting criteria or developing material that will aid collection support activities. All other studies, no matter how well-intentioned, are merely a waste of an analyst's valuable time.

P.E. 18086636
EØØ.4.(c)

REPORTING

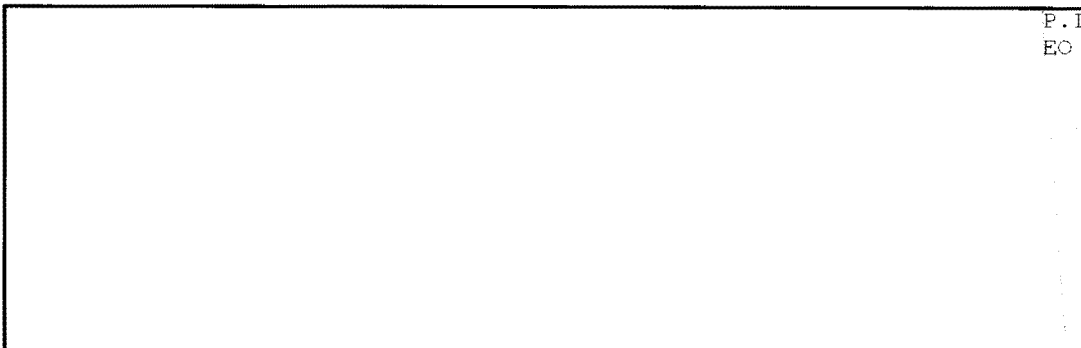
(U) The reporting function is the sum total of how well the earlier functions were conducted. If these earlier functions were conducted in a shoddy manner, they normally will culminate in an inaccurate report that bears little relationship to the consumer's initial stated requirement. Conversely, if these earlier functions were conducted in an orderly manner so that each function supported the following one, the report will probably satisfy the consumer's request. In addition to satisfying the consumer's request, the analyst must also concern himself or herself about using the proper format for the information to be used in the report. Poorly formatted reports, while factually correct, convey a sense of disorganization that can color the credibility placed on the information by the recipient.



(U) Both of these rationalizations have effectively downgraded the opportunities available to the traffic analyst to participate in the reporting function, the only function which at present is directly keyed to reporting criteria that match the consumers' requirements. By reducing the analysts' participation in this function, we also reduce their ability to ensure that the previous functions remain directed towards the ultimate

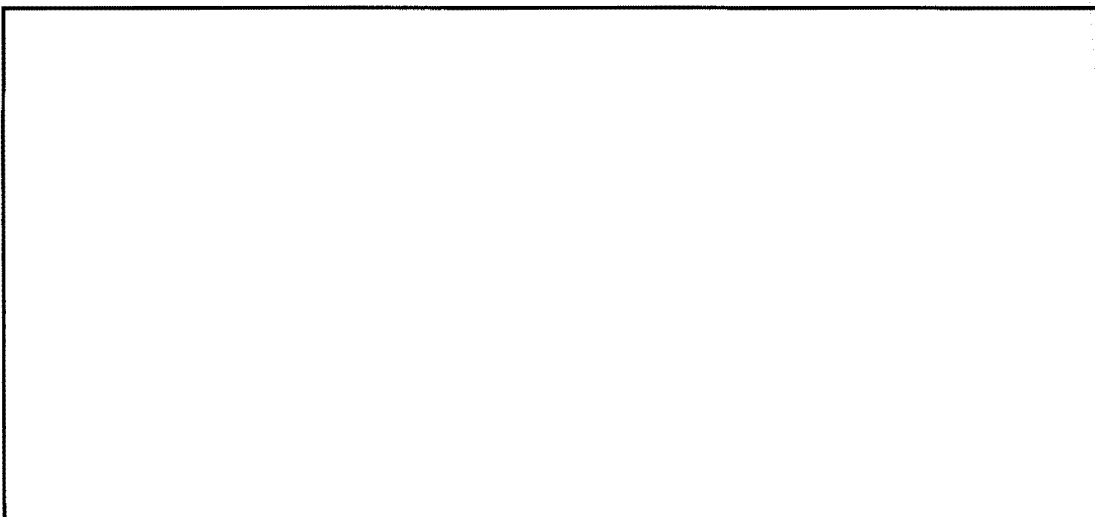
goal of consumer satisfaction. Managers at all levels must emphasize the development and maintenance of effective and imaginative reporting programs that challenge the analysts' skills to ensure that that reporting function does not become divorced from all that has preceded it.

EVALUATION



P.L. 86-36
EO 1.4.(c)

(U) Through their involvement in the evaluation process, traffic analysts are well placed to redirect collection resources to exploitable areas - or at least to areas that have not been satisfied through past intercept - of the target's communications. Unfortunately, the evaluation process is normally either not performed at all, or it keys on indicators that, when viewed alone, fail to provide an accurate reflection of the productivity of the resource.



(U) As should be apparent, the traffic analyst of today is faced with a variety of functions. However, all of these functions are tied together by a thread of continuity relating to the final goal of consumer satisfaction. If these functions are not kept in their proper relationship to each other, either through analyst neglect or improper managerial emphasis, it will be difficult - if not impossible - to achieve a cohesive operational mission

~~CONFIDENTIAL~~

that satisfies consumer requirements. Only through mutually supporting functions, as originally envisioned, can our scarce collection and analytic resources be used to their fullest extent. In other words, perhaps a return to the basics may be in order for the traffic analyst of today!

~~CONFIDENTIAL~~

~~HANDLE VIA COMINT CHANNELS ONLY~~

There's a New World Coming - Are You Ready?

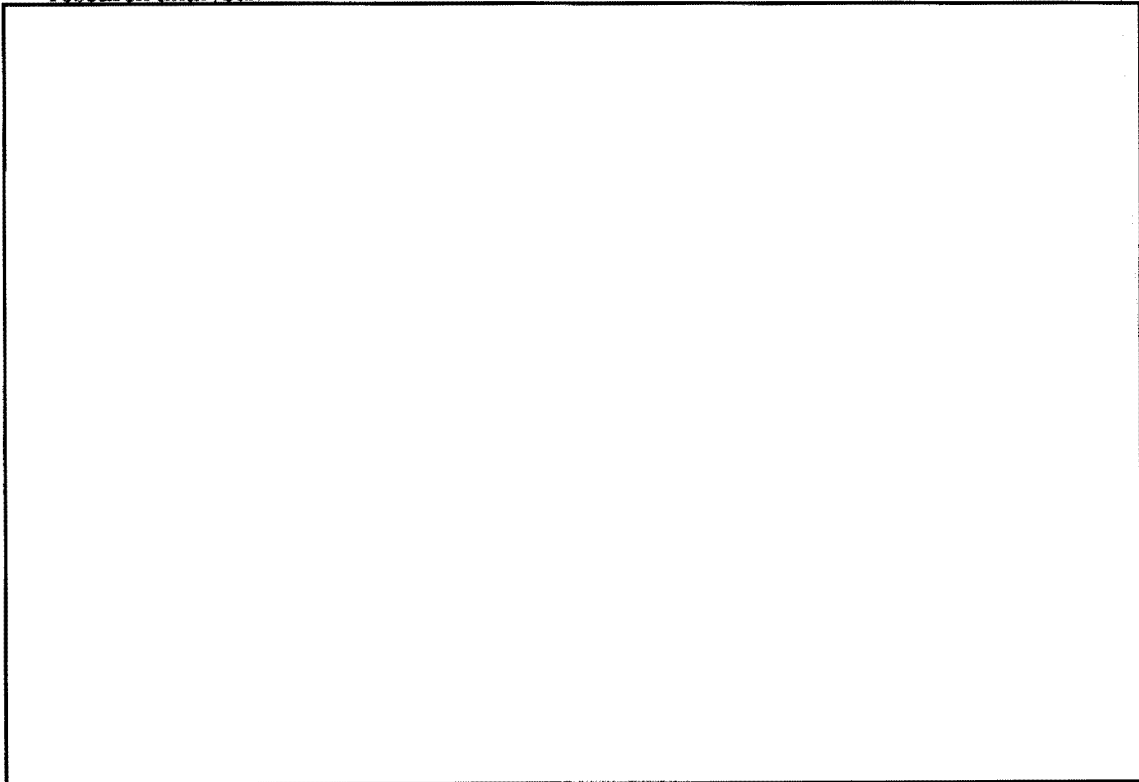
P.L. 86-36

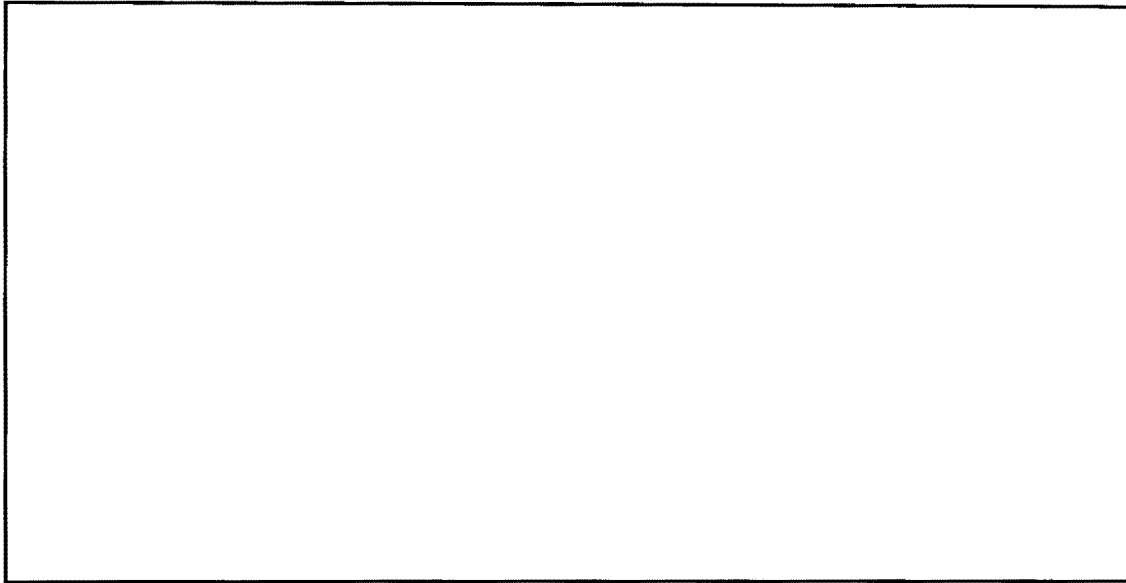


P.L. 86-36
EO 1.4.(c)

(U) Analyst response to a changing environment is an area that has piqued my interest for a number of years. This interest has become especially pronounced in the last four or five years, largely as the result of the increase in the variety and number of automated processes that have become available, my involvement in a number of long-term research projects, and the prospects of newer and more sophisticated machine-based analytic techniques. As a consequence, I have given much thought to the effects that this increased automation will have on the analyst and to what we, as an organization, can do to respond to the challenges and opportunities that will arise.

(U) If we are to accurately assess how the analyst of the future will react to the increased automation of the analytic process, it is necessary to know the analyst of today. This is logical, since the majority of tomorrow's analysts are already on board and active in the analytic field. It also becomes essential to define the functions of the analyst since all personnel who carry the title "analyst" are not analysts by the strictest definition of the word. For convenience and easy reference, I have divided these "analysts" into three categories that I've labeled, for lack of any better terms, "loggers," "case analysts" and "research analysts."





(U) The changes expected as we become more and more automated will alter the methods and procedures currently used in the analytic areas. This will have a profound effect on our operations, but to an even greater extent it will modify the lifestyle and environment of each of the three types of analysts I've defined. Each will be affected in a different way depending on the analyst's perception of analysis, his personality and his ability to adapt.

(U) Once the machine takes over the simple data manipulation and recording functions, the "logger" will become, to a great extent, obsolete. His routine will be upset and his workaday world completely disrupted. He will no longer have his logging to fill his day. He won't know what to do and he'll be completely lost. For those currently having trouble isolating the "loggers" in their outfits, they'll then be very easy to spot. Just look for those people sitting with a blank stare on their faces and for those making nervous movements at their desks as they anxiously try to find something to do. Eventually their names will appear on the list of those who are abusing their sick leave. The real "logger" will never recover from automation. He's lost his place in life. His day-to-day world will have been destroyed, and he simply won't know what to do. For many it will be too late to start over. Others will lack the initiative, while for still others, it's nothing more than a lack of talent. The world will have passed them by. They, of course, will not be to blame. It'll be the machines' fault, or management's fault, or maybe just the breaks of the game.

(U) In reality, the "logger" has always been a clerk with a professional job title and a professional paycheck. He was created by the system, and, with the advent of new automated techniques, will be destroyed by the system. Some will be salvaged. Some "loggers" will regroup, retrain and regain a place, perhaps at the "case analyst" level. Those that can't make the transition will have to be purged since the cost of automated systems will call for a decrease in personnel expenditures.

~~CONFIDENTIAL~~

(U) The "case analyst" will survive and thrive in the new automation. He really never did care for the logging and data maintenance functions associated with his job; therefore, he'll adjust. He'll reorder his priorities and use the additional time now available for the development and analysis that he never could quite get around to under the old semiautomated method of operation. His productivity will increase, and more technical data on his targets will result.

The ultimate gain will be the production of more intelligence information in satisfaction of our requirements. The "case analyst" will grow as a professional, honing his skills as he practices his trade, and in many instances will develop into a "research analyst."

(U) The "research analyst" will see few changes in his method of operation as a result of this new automation. He'll have to deal with new databases, new retrieval programs and new equipment, but his daily routine will remain pretty much the same as it is now. He will generally review the same types and amounts of material and work on the same types of projects. He should see some improvements in accuracy and completeness as a result of the elimination of the "logger" and the automation and resultant upgrade of the databases. He will be able to spend less time on the verification and data gathering phases of his assigned tasks and thereby be able to complete more assignments in a shorter time. This will allow a more efficient use of the limited number of "case analysts" and permit us to achieve maximum benefit from their talents. Automation will also improve the morale of the "research analyst" since he will be able to function almost entirely in his primary capacity.

(U) The benefit of this increased automation is readily apparent. The talent of our people can be used to the fullest, with dull, repetitive tasks reduced or eliminated. The timeliness, accuracy and quality of our product will increase. Personnel not working at their prescribed levels can be eliminated, with a resultant savings of money. All of this is, of course, predicated on manageable machine systems that will function as designed. Since our track record for the development of such systems is not impressive, let's hope that we have learned from experience, and not attempt to reap benefits before we have proven follow-on systems.

~~CONFIDENTIAL~~

~~HANDLE VIA COMINT CHANNELS ONLY~~

**In Pursuit of:
Faster Horses, Younger Women, Older Whiskey,
and More Money**

P.L. 86-36



(U) For the past two years I have been involved with a project that has given me more in terms of psychic income and pure excitement than perhaps any work-related activity that I remember. When you feel this good about something, it seems natural to want to tell everyone else and share the excitement. It's sort of a "look what I found" feeling. Of course, when you feel excited about something, it is difficult to know whether you have something worth saying and can remain objective about it. Nothing makes you feel quite as foolish as discovering the wheel only to find out that you were the only one who didn't have one all along.

(U) I've followed - from a distance - the articles, letters, and symposia decrying the diminishing number of analysts, the dilution of the career field, and the increasing work load. I really have nothing to add to the body of literature that has grown around those themes. I would like to note that some reasonably intelligent people have advanced them. Conversely, some reasonably intelligent people made the decisions that led to the described conditions.

(U) Having resisted the urge to vent my excitement on paper for this long, I thought I had it under control. Actually, I have been writing this piece all along. Part of my control mechanism was simply typing my thoughts on the screen and then hitting the delete button. That may happen to this version, and you will be spared once again. I'll tell you what "set me off" this time a bit later. First, let me tell you what I've been so enamored with.

(U) I'm a traffic analyst. Several years ago I began to work for the person I respect most in that field and share in the development of what has come to be called a Traffic Analysis Workbench System. Whatever comes of that effort, I'll always be grateful for being included.

~~(FOUO)~~ The idea is a relatively simple one: In terms of technology, the TA field is, and has always been, behind the power curve. Regardless of what high-powered machines exist, and in spite of the fact that some extremely sophisticated machine applications have been designed for analytic purposes, the analyst is still behind. There are several reasons this situation exists. However, it is primarily because analysts are directly dependent upon their machine support personnel. A few have managed (mostly out of frustration) to learn one computer system or another and support themselves. The problem with this is, if they were any good at it, they were usually lost from the field.

~~CONFIDENTIAL~~

~~(C)~~ It's time the analyst was given some help—not to catch up to technology, just to keep from getting further behind. Given the costing cycle, the procurement cycle, and the installation cycle, I'm convinced that catching up is not possible. *Not possible!* The concept of a TA workbench involves installing a terminal on the analyst's desk. Read that again! *On the analyst's desk.* Not down the hall in a "machine room," not in a corner of the basement, and not around the corner where it "won't bother anyone" — *on the analyst's desk, the one he sits at.* [REDACTED]

EP..II.. 886-336
BKD 11..41..((C))

[REDACTED] This puts the analyst in a position to access the major databases, where the daily traffic, as well as the technical working aids, resides (hides is a better word). With the terminals on their desks, analysts will have constant access to their material and perhaps approach the paperless environment.

~~(FOUO)~~ Under the umbrella called PINSETTER, we have been proceeding along a development path that will hopefully lead to the kind of help the analyst needs. Because the most precious computer resource is the programmer, the analyst must be released from depending on him for every minor need. This is true for several reasons. First, the analyst needs to be able to access his data, process that data, and change those processes without having to write memos, generate specifications, write justifications, wait for software and then participate in debugging. Second, the programmer as a resource is too valuable to be tied up with changing sort specifications every time an analyst needs a different output. Last, the plain facts are that we have a terrible time retaining good programmers. No sooner do we develop a good working relationship with a top-notch programmer than he begins to understand something about an analyst's job; then along comes a better offer and he's gone.

(U) From a machine standpoint, meeting these goals requires a system that is easy to learn, flexible, and provides a reasonable response time. By "reasonable" I don't mean instantaneous. Most analysts can live with an execution time that is not measured in nanoseconds. Most of our work has taken place on a PDP-11/70 host using UNIX as the operating system. UNIX is a high-level language that was developed by Bell Laboratories. It meets the above criteria, and it is very forgiving to a klutz at the wheel.

~~(C)~~ We have found that most of the processes that a traffic analyst needs to be able to do can be accommodated with the UNIX package. Where it was found lacking or inefficient, the solution has been provided by a unique working relationship with a small group of highly talented programmers in T333. [REDACTED]

~~CONFIDENTIAL~~

~~HANDLE VIA COMINT CHANNELS ONLY~~

(U) To digress for a moment, the realization that certain processes are simply too big for TSS applications is important to maintaining a proper perspective. This determination must be made, and large "number crunching" must be performed where the processes are most efficiently handled. However, the process can often be executed where most efficient and the results passed to where they can be best used, on the TSS. I might add, in two years of handling TA processing, it has been necessary to "send out" only one job for actual execution on a "big" machine. Of course, many of our extracts from major databases are "preprocessed," prior to transfer, to make them more TSS friendly. But I discount this, since it is largely "invisible" to the requester.

~~(FOUO)~~ If the solution were apt to provide a useful "UNIX-extension," we would request T333 help. The results have been the most rewarding part of this experience: generalized UNIX-like utilities that solve analytic processing problems. The big plus? *Anyone* who knows a little UNIX can use them. On the other hand, if the solution appeared to be problem specific, we would attack the problem with our own resources. The results of these efforts have proven equally rewarding. Based on our own experience and some operational testing in analytic elements of A3, B2, B5, G6, and G9, I'm not sure if a more effective analytic tool than UNIX could have been designed if that had been Bell's intent. This leads to a philosophical difference in user support design.

~~(FOUO)~~ There is a mask-and-menu school of thought that holds to the belief that the user should be led through the processing cycle by the software. A menu is presented with a few options to select from and a mask provided through which to make alterations. These M&Ms believe it is best to protect the user from the complexities of the system and protect the system from the klutz at the wheel. It has a place. I would look to this area for the type of handling necessary for, perhaps, TEXTA updates.

(U) Another approach is to provide the user with the modules necessary to manipulate the data, a high-level language to package the modules, and the ability to communicate with other users and peripherals such as high-quality printers: basically, a sort of Procedural Applications Language that is not unique to traffic analysis. Perhaps a Universal Procedural Applications Language approach. The user is free to design personal processes and, more importantly, *change* those processes at will. Users are not dependent upon the programmer for every minor modification, routines do not have to be recompiled after each change, and the results of the changes are immediate. I believe UNIX meets this challenge.

(U) The M&M approach keeps the analyst (or user) dependent upon the programmer for modifications, thus preserving the problem of too much demand being placed on a resource that is already overtaxed. The solution to the demand for software packages has

~~CONFIDENTIAL~~

all too often been the letting of contracting, at considerable expense, to develop processes that a few analysts, skilled in a handler like UNIX, might be able to get along without.

(U) If our own resources were concentrated in a manner conducive to the development of generalized handlers (a Universal Procedural Applications Language), and perhaps a bit of that contract money concentrated into rewarding the good programmers we have left, we might be able to come up with better analysts and better programmers. As a by-product, we might be able to handle the workload with the number of analysts we have and do a better job of it.

(U) So, what was it that set me off this time? A few days ago, while demonstrating a few system capabilities to a potential user, I was walking through the steps of a UNIX shell file (merely a collection of UNIX commands that execute sequentially and perform some process) and he asked me if I "wrote this program." The words startled me. Wrote a "program"? Me? I'm a traffic analyst, I can't "program." My rather bumbling answer was something to the effect that this is really not a program, just a collection of instructions to perform a certain process on this computer. After he left I put the shell on the screen and read it a few times. By gosh, a few years ago I would have called that mess a program myself. It "looks" like a program. It "acts" like a program. And my extemporaneous answer wasn't too bad a definition of a program.

~~(C-CCO)~~ I had to pause and reflect a bit. I put that shell together in about five minutes. What does it do?

EP..II.. 886-386
RKO 11..41.. ((C))

Based on past experience in trying to get a process to do a select of this nature, and going through the "channels" to get it, this "quickie" shell seems fairly powerful.

(U) I think I've found a faster horse, I probably couldn't keep up with younger women anyway, and I'd rather have a cold beer than older whiskey, so if anyone knows someone looking for a "programmer," I'll settle for two of four.

~~CONFIDENTIAL~~

Computerizing Traffic Analysis (U)



P.L. 86-36

This paper was presented at the November 1982 meeting of CISI.

The traffic analyst finds himself turning to data systems because he often has mountains of data to examine, because the people who receive TA results usually want their information very rapidly, and because almost all the data the traffic analyst wants to see are already inside a computer somewhere.

(U) Most traffic analysts who try to look at data systems develop a kind of schizophrenia. On the one hand, the TA data that come in today have to be processed and analyzed today because there will be another batch of data coming in tomorrow. This means that the traffic analyst has to use today's data system to handle today's data. On the other hand, it does seem to us traffic analysts that data systems people would much rather talk about tomorrow's system - the one that isn't here yet, the one that won't have all these glitches and problems that today's system has.

(U) The traffic analyst who is in the trenches on a current operational problem would easily trade all the glowing promises of some brighter tomorrow for a quick fix on some of the glitches in today's system that will keep him from bleeding to death right now. That isn't my subject today - I really want to talk about the future. But as I thought about standing up here in front of all you data systems people, I couldn't resist putting in a plug for the working traffic analyst; he needs your help, both today *and* tomorrow.

THE PROLOGUE IS PAST

(U) I might be well to begin with a little history, or at least history as I remember it. My first recollection of what we now call data systems was a lot of eighty-column cards and a card sorter. That was about thirty-five years ago. Watching those cards go through that sorter was rather hypnotic. The possibilities seemed limitless then - if we could only find a cheap and easy way to get the data onto the cards. I think the equipment was called Electronic Accounting Machines (EAM), and the people who supported the traffic analysts were called Methods Analysts (in the 1940s and early 50s).

(U) Since our data consisted of a matrix with eighty columns and many rows (one row for each card), our output consisted of that same matrix with its columns and rows transposed in some way. Later, we added the ability to look up words or strings in a dictionary and insert the result back into the matrix.

(U) Many years and computer systems later, in the mid-1960s, this was still the primary data systems support to traffic analysts: a transposed matrix (now often wider

than eighty columns) with a dictionary lookup. There were attempts to go beyond this. Most of the things we tried were made to fit one specific problem and never developed into general TA tools. We developed ponderous, monolithic record formats whose structure provided a special place for each variety of data we thought we would find in the traffic. What I remember most vividly are long, soporific meetings where all we ever seemed to talk about was what *format* the data were going to be in. We spent untold amounts of energy and resources getting all of our data *into* these unyielding, user-murky systems, and there were often little energy and resources left over to develop any user-friendly *output*.

(U) The result of this, in many areas, was that the output received by the traffic analyst was not much more than his original raw traffic, transposed both horizontally and vertically, and with some information added through dictionary lookup processes.

A LONG THIN MATRIX

The form in which the output was delivered to the analyst was often decreed by someone remote from the analyst - someone who never had to actually live with the output - and it was rarely if ever changed to fit the current needs of the local problem or individual analyst.

(U) It is still possible, even today, to see analysts sitting down with computer output and handlogging data from that computer output onto a form for their own personal use. In at least two areas, one might then see that same handwritten log being used a little later to punch cards for further computer processing!

WHAT IS THE TRAFFIC ANALYST TRYING TO DO?

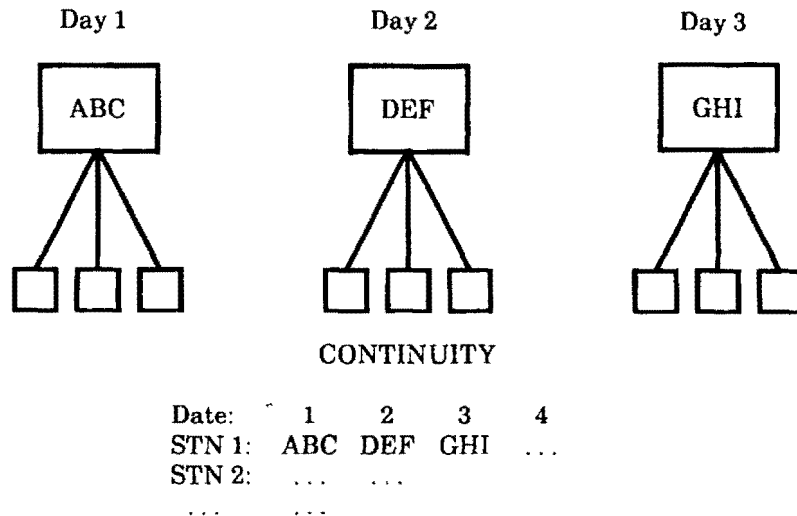
(U) The traffic analyst is trying to draw a picture of his communications target. He usually wants this picture to show how his target looks when it is operating normally. Once he knows what his target's normal behavior is, then he is in a position to detect variations, and report them to intelligence consumers.

CONTINUITY

(U) Traffic analysts are usually looking for something they call continuity. When faced with a target that has daily-changing callsigns, the traffic analyst seeks to learn which of today's callsigns matches what callsign used yesterday.

If I can say that the station that used callsign ABC on the first day is the same station that used callsign DEF on the second, then I can say that DEF (on the 2nd day) is continuity of ABC (on the 1st). On the third day, if I can say that GHI was used by that

same station, then I can add GHI (on day 3) as another link in a growing chain of continuity. Many of our TA targets do change their callsigns, frequencies, addresses, and other features on a regular basis. They do it to make collection and identification more difficult, and it is the job of the traffic analyst to defeat these changes by the development of continuity.



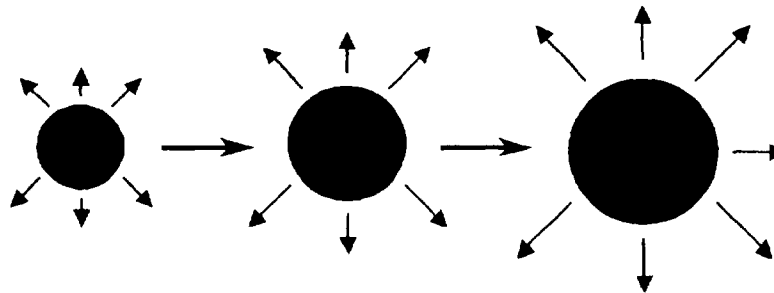
TWO KINDS OF TRAFFIC ANALYSIS

(U) There are two forms of traffic analysis on most problems: development and maintenance. To borrow an example from cryptanalysis, the attack against a cipher system often goes through two phases:

- first, diagnosing and recovering of the general cipher system, and
- second, exploiting and processing the recovered system, which often involves solving daily keys or settings.

So too, in traffic analysis, one can consider that there is a development (or recovery) phase and a maintenance (or exploitation) phase, which may or may not include product reporting. However, in traffic analysis, the two phases often occur at the same time.

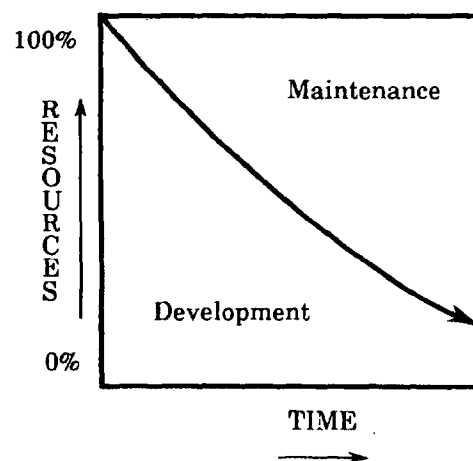
(U) In some ways, the traffic analysis process resembles a spreading oil blot. Out on the edges, new target territory is being conquered; new target communications structures are being discovered and cataloged; new methods of identifying and distinguishing various communications are being developed. But back in the central part of the oil blot, the territories previously conquered must be kept track of; the continuity of target communications structures previously recovered must be maintained.



Recovery – outer edges

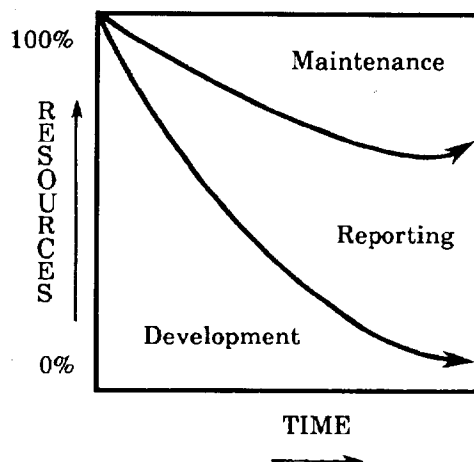
Maintenance – center

Oil Blot



(U) The more territory one conquers, the thinner the center of the oil blot becomes. The more communications structures one recovers, the more continuities there are that now must be kept track of. As the maintenance effort grows, it will use more of the available resources, draining them away from the recovery part of the effort, and at some point it will have absorbed enough of the resources so that a point of "no growth" is reached and, for all practical purposes, recovery of new structures stops. If expansion doesn't stop, the center of the oil blot will break; if development doesn't stop, the maintenance effort will fall behind and begin to lose track of continuities, which will then have to be discovered and developed all over again. This tension over resources between maintenance and development is similar to the one between software maintenance and software development.

(U) Sometimes the personality of the manager plays a part in just where this point of "no growth" takes place. Some managers are more at home in the settled, stable atmosphere of the center, where things don't change much from day to day. These



managers tend to concentrate their attention on building a smooth-running system at the center, and may put a larger proportion of their resources into that area, so that the "no growth" point is reached more quickly. Other managers thrive in the rough and tumble frontier atmosphere out on the edges of the problem, where each day is likely to bring some new and different challenge. These managers tend to concentrate their attention on the recovery effort, sometimes at the expense of the more humdrum maintenance.

REPORTING

(U) Some traffic analysis problems have a lot of potential for reporting - for providing the intelligence consumer with a blow-by-blow account of what the target is going. Targets that involve ships and aircraft often have this potential because they move around from place to place, and the analysts often find much of their time taken up with reporting which ships and aircraft were active today, in what areas and performing what missions. Where this reporting potential is high, it tends to draw off resources from both development and maintenance. Managers whose problems have a strong reporting emphasis (especially time-sensitive reporting) will generally try to pull resources from development rather than from maintenance, because losing the continuities means losing the raw material for the reporting effort. Losing the development effort is generally seen as the lesser of two evils.

(U) As an aside, I should say here that the reporting side of traffic analysis is generally well ahead of the technical side in the use of computers. Since my primary interest in this paper is the working-level traffic analyst, I will be concentrating on the technical side, and I do not propose to discuss the reporting aspects of TA except as they touch on the technical side.

(U) From the standpoint of the two kinds of traffic analysis - development and maintenance - we can express the general goals in the following ways:

TA DEVELOPMENT GOALS

(U) We rarely collect or analyze all of the communications of any given target. We are almost always working on a sample of the target. At any given time, there is some residue of the target that we do not maintain continuity on, and bits and pieces of that residue find their way into our unidentified or search pile - the file of incoming traffic that looks as if it belongs to our target but doesn't exactly fit any of our known continuities. Development TA concentrates on that pile, trying to dig out new target nets and continuities. This unidentified pile is almost like "background noise": it is always there, whether we talk about it or not. If we are still growing (if the oil blot is still expanding), then our development goal is to dig more of the target out of the unidentified pile. If we have reached the "no growth" point, then our development goal is to be able to recognize and develop any new communications that the target might put on the air - communications that ought to stand out against the "normal noise" in the unidentified pile.

TA MAINTENANCE GOALS

(U) During the maintenance phase, we want to be able to hang on to the continuity that we have already recovered. We want to do this

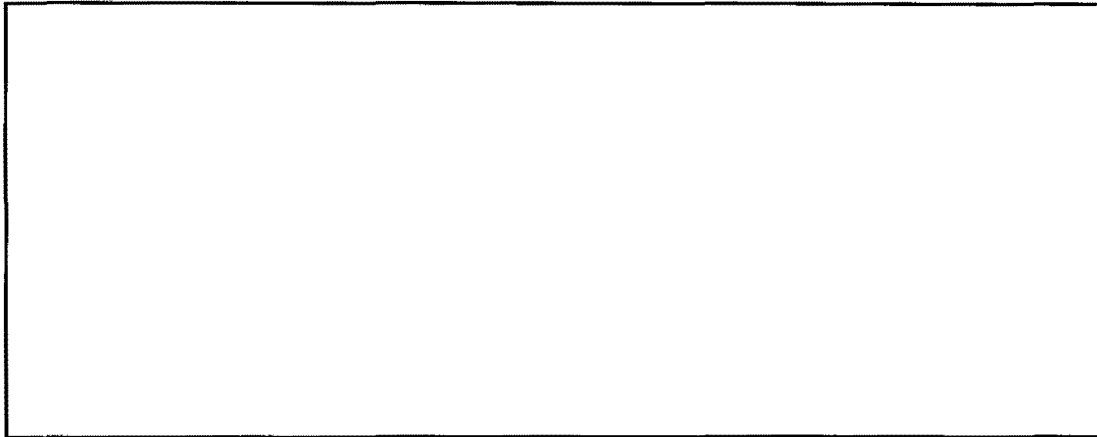
- to support whatever analysis efforts are currently engaged on the target (such as cryptanalysis, language, reporting, etc.), and
- to support whatever collection effort is working against the target.

(U) To do the first support requirement properly, we need to be able to correctly distinguish and identify each of our continuities as the traffic arrives at the point of analysis, i.e., *after* it has been collected. However, to do the second support requirement properly, we need to be able to tell the collector *before* he collects the target, what the target is going to look like and what the target is likely to do. If we can tell the collector what callsigns the target will use, what frequency he will use, what time of the day he will come up, and what he is likely to do when he does come up, then we have satisfied our goals for maintenance traffic analysis. On some problems, we are fortunate in being able to do this on a regular basis. On others, we spend the first few hours of the day (or callsign period) scurrying about to recover new keys and then, for the remainder of the day (or period), we can project what the collector will encounter in the way of callsigns, etc.

HOW CAN COMPUTER POWER BE APPLIED TO THE TA PROBLEM?

~~(S-CCO)~~ In order to consider how the power of the modern computer might be applied to traffic analysis, we need to look at the model of TA that emerges from these two phases: development TA and maintenance TA. Although I have described them as if they were distinct and separate, they really ought to be thought of as a conjugate pair, because they tend to occur together on most problems. It is also possible for certain problems to be best

described as a hybrid of these two forms: during the war in Vietnam, one out of every three pieces of intercepted traffic was unidentified, largely because of the rapidly changing nature of the target. The point I want to leave with you today is that any attempt to provide the traffic analyst, either here in this building or anywhere in the world, with a Traffic Analysis Workbench System must reckon with the fact that the problem he is working will always be some mixture of these two forms of traffic analysis. We also need to consider that a TA problem can quickly change from one form to the other.



P.E. 1868636
E.O. 4.4(c)

~~(C)~~ We decided to see if we couldn't find a way for computers to help us with the more stable maintenance problem. I remember spending several weeks laying out the logic and processes on the problem. And I remember being told, at the end of the project, that there wasn't nearly enough memory available to do what I needed.

	MAINTENANCE	DEVELOPMENT
Problem type:	slowly changing bookkeeping "Anything changed?"	rapidly changing pattern searching What's new?"
How dynamic?	slowly changing	rapidly changing
Foreknowledge:	high	low
State of solution	solved	unsolved
Control:	semiautomatic	hands-on
Interaction:	human-efficient	human-intensive
Techniques:	target specific knowledge-based?	human specific "mix-n-match"
Worst case:	"below the salt"	start from scratch

A COMPARISON OF TWO FORMS OF TA

(U) Let's look at these two forms of TA a little more closely. How do they compare when we look at them from the viewpoint of providing today's (and tomorrow's) traffic

analyst with a computer support tool kit, while using a terminal workstation in an NSA worldwide networking environment?

(U) In development TA (the garrison communications in our example), we have a bookkeeping problem,

- where the emphasis is clearly on keeping track of a lot of known continuities;
- where we expect the changes in the target characteristics to be relatively modest;
- where the technical means of keeping up with the target (i.e., callsign and frequency systems, address tables, etc.) are largely solved or understood; and
- where we have good prospects of being able to project the appearance and behavior of the target from day to day.

(U) In development TA (the training communications in our example), we have a pattern-searching problem,

- where the emphasis is on sifting through masses of low-yield ore, looking for something that forms a continuity;
- where the next success may look nothing like the last one; and
- where the chances of finding that needle in the haystack may depend as much on the personality of the searcher as on the content of the haystack.

If we can't *keep* continuities (i.e., are not *able* to), then the target stays in the development phase, no matter how much we know about it. Someone once said that TA continuities take either 95 percent of our resources or 5 percent. That number may not be right, but the idea is. Being able to keep track of the continuities is the key to whether the problem is development or maintenance in nature. A daily-changing callsign system *looks* to us as if it is rapidly changing if we haven't solved the system, but once the system is solved, we then perceive it to be slowly changing. It is a matter of viewpoint.

(U) In maintenance TA, we work largely with what the target gives us. A package of techniques to grapple with a callsign system may work well enough on a problem where the callsigns are the key to our keeping track of continuity, but may be almost useless on another problem where the callsign system isn't solved and we must rely on other things, such as serial numbers or addresses.

(U) In development TA, on the other hand, a particular technique may pull one new structure out of the search pile and then never again find anything. The development analyst may need to continually devise new attacks and new methods; to him, the search pile is a featureless mass, and it is his job to sort out the various pieces and find ways to distinguish one piece from another with some reliability.

(U) The maintenance TA problem probably needs a package that will

- look over the incoming material for the day;

- make reasonable guesses about continuities (including garbles);
- flash a warning light at the traffic analyst when things look *very* wrong or when it is confused by something; and
- provide a clean and readable summary of its results to the analyst for review.

It ought to keep up with both short-term and long-term trends, and should be especially attentive about "missing persons," portions of the target which haven't been seen for a while.

(U) The development TA problem, on the other hand, needs a tool kit that will provide the analyst with a range of diagnostic, computational, and pattern-searching techniques that can be brought to bear on the problem, in whatever mix the analyst needs at the moment.

WORST CASE

(U) I have shown what might be called the "worst case" for each of these forms of traffic analysis.

(U) In maintenance TA, one sometimes finds that a problem must somehow be worked, but that it has no real resources and not enough clout to get any. Now, in the best of all worlds, where everything is done right and for the right reasons, such problems should not exist. If a problem is worth working on at all, it is worth the resources needed to get the job done. However, in the real world, those problems that are "below the salt" will be working with whatever support they can beg, borrow, or scrounge. Proving a general package for such problems would pay for itself a hundredfold in the first few years. At the minimum, the package needs to be able to "ring an alarm bell" when the target starts to disappear, or becomes more active, or changes in some other way.

~~(S-CCO)~~ In development TA, the worst case might be the situation where nothing is known. That is not as uncommon as some people might think.

When we pull together an analysis effort for a sudden war or brushfire, the analysts are usually drawn from other problems around the building; it would be nice if they didn't have to add "learning a new system" to all the other problems they will face on the new target. Therefore, the tool kit for such situations must be quite general and all-purpose.

P. E. 1868636
E.O. 4.4(c)

TWO SYSTEMS OR ONE?

(U) What I have been describing so far may sound like two different systems, but what I am proposing is one system, with two parts. I have already said that these two phases or aspects of TA occur together, and I should add that on more than one problem, they are frequently done by the same people. New continuities are recovered by the development

~~SECRET~~

TA process, and then handed over to the maintenance TA effort to be kept track of. Information is often derived by the maintenance effort that will help the development effort. What the traffic analyst needs is one system that has enough flexibility for him to move whichever way his TA problem takes him. It would also be useful if the language we use is one that isn't going to change every few years because some equipment in the basement is being upgraded.

PINSETTER

(U) Several years ago, we began to work on the concept of a Traffic Analysis Workbench System, with the covername PINSETTER. Some of what I have described here comes out of that experience. PINSETTER has been described elsewhere, so I will not spend time on it here. However, I will share with you some of my personal conclusions about PINSETTER, especially those which seem to be pertinent to the future.

(U) There are aspects of traffic analysis that resemble word processing, and a good screen editor seems to go a long way toward putting the analyst in contact with his traffic, letting him rearrange it and touch up the rough edges and garbles the way he (the owner) wants them. It lets him look at the data *before* he decides what processing to apply to them. It also puts him in a good position to generate reports about his problem, especially the technical reports with technical data embedded in narrative text.

(U) A good tool kit, similar to UNIX and the PINSETTER extensions, is invaluable in providing the traffic analyst with the ability to tailor-make his own flexible processes for large scale manipulation of his traffic.

(U) Many of the practical results of PINSETTER results that found their way into daily applications on specific targets were not limited to traffic analysis. It became a regular occurrence to hear people from other cryptologic disciplines tell us that much of the UNIX/PINSETTER package for traffic analysts was what they needed, too.

PROBLEMS THAT NEED SOLVING

(U) Among the many problems that need to be solved, I would like to mention two. Both of these are areas that are critical to the future TA Workbench System.

ARCHIVES

(U) Some of our continuities form chains that stretch back to the end of World War II. One of the things that data systems people don't like to hear is that we need storage for data whose lifetime must be measured in years, and perhaps decades. Some years ago, there came a time when all of our incoming data went solely into the computers in the basement. It was the culmination of the dreams of a number of people: to take the raw traffic away from the analyst! I don't challenge that decision. It is history. But I must say

~~SECRET~~

that on many TA problems around the Agency, there are no good records on our *known continuities* from that date forward, unless there were analysts still keeping some sort of hand records. The philosophy on most computer hosts is that any records not accessed within some period (usually a year or less) are taken off the system.

(U) Even if the data are put onto tape, the medium will deteriorate. Once on tape, the data are "out of sight and out of mind." The software understands that data will sooner or later disappear or be "improved." Nevertheless, the analysts on that problem are still responsible for that period of time, and may still have to field questions about their targets for that time period. So far, we have dodged this bullet, but sooner or later we will have to face the need for long-term archives.

INFREQUENT USE OF PROCESSES

(U) The second problem involves the question of software that is only infrequently used. For example, suppose that one of our larger targets has a major communications change every five or six years. The effect of this change is so great that it interrupts intelligence reporting on that target until the new communications structures are understood and recovered. Each time the change occurs, an intensive effort is therefore mounted to recover our continuities in the shortest possible time.

(U) In the old days, when the special effort was over, everything was bundled up and packed away for retrieval when the next change came along. But how do we handle this now that we have modern data systems support? After five or six years, how much of the software is still useful? Chances are that the database has been changed, as well as the host on which it resides.

(U) Another example might be the diagnostic techniques to attack a particular kind of callsign system. Once the system in question is solved, how should we preserve the software so that it doesn't need to be reinvented the next time such a system is encountered? Suppose we don't find a similar system for five, or six, or even ten years?

CONCLUSION

(U) I don't offer either my observations or my experiences as criticisms, but rather as areas of traffic analysis support which need to be solved. I have tried to avoid mentioning specific hardware or software, except as examples. A man named Bob Biles taught me long ago that users should never tell computer people what equipment to use.

(U) Perhaps traffic analysis has lagged behind other cryptologic disciplines in making full use of modern data systems. But that is changing, thanks to the patience, ingenuity, and hard work of many of you here today. I still keep a supply of pencils around, and I still have a pencil sharpener on my desk - but I have noticed that I don't really use them very much anymore.

Traffic Analysis: Specialty Without Portfolio



P.L. 86-36

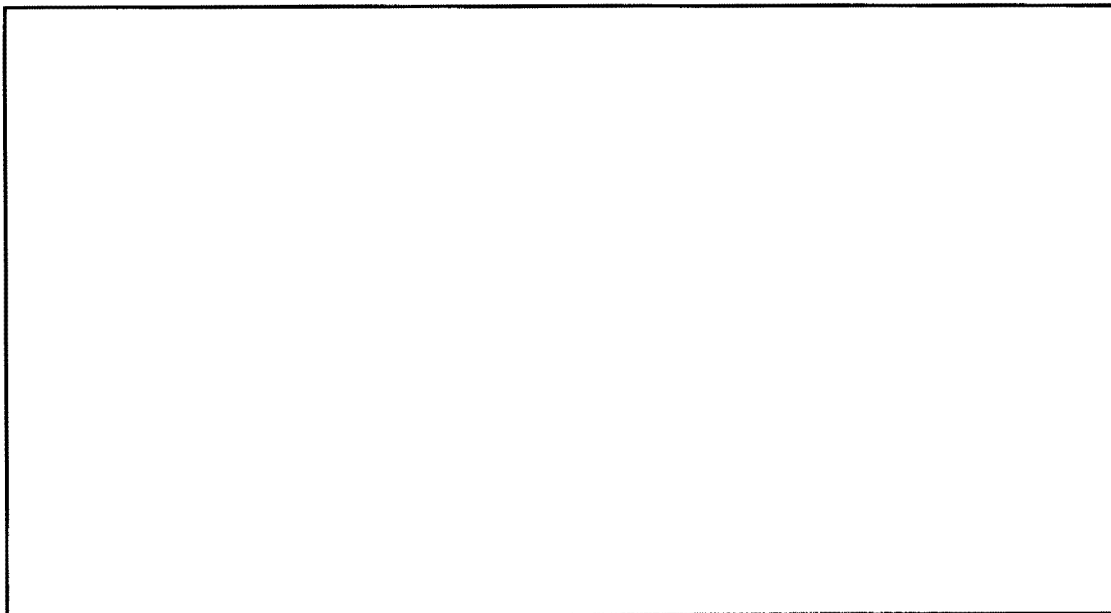
Editor's Note: This is an address given by [redacted] the A2 Senior Analyst, at the 1980 symposium of the Communications Analysis Association; it is based on his extensive experience on the Soviet military problem.

In addition to cryptanalysis, which I do not intend to deal with in this discussion, communications intelligence has always involved two primary, overlapping but distinct, tasks with respect to target analysis:

- ordering and understanding target communications structure and procedures—commonly referred to as *traffic analysis*, and
- describing target organizations, activities, intentions and trends – known as *intelligence analysis* or *special research analysis*.

The major overlaps of the two disciplines are in their contribution to target organization, or order of battle, and command and control.

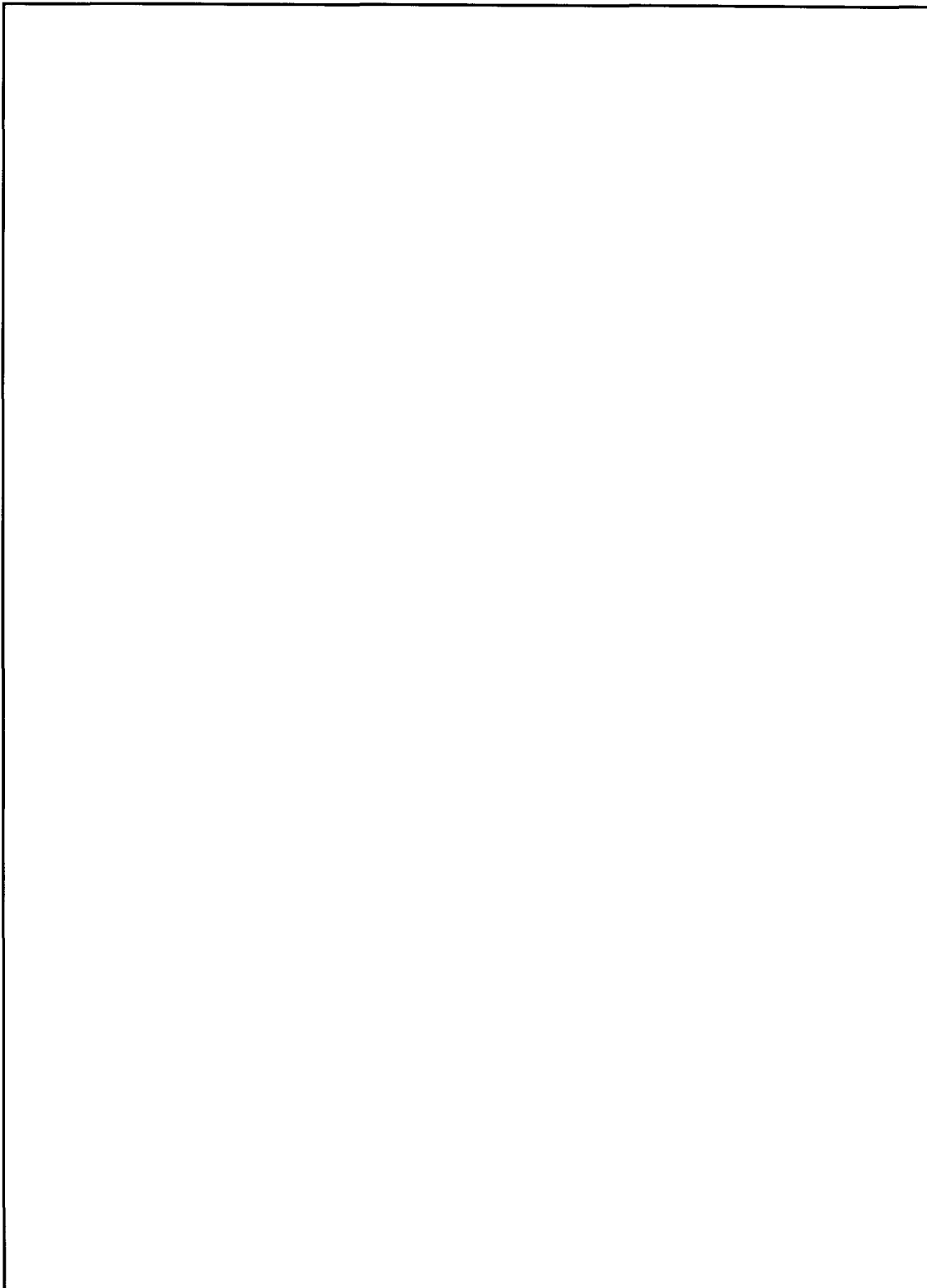
Of these two skills, only the latter one, intelligence analysis, has been effectively accommodated in the NSA career structure. Traffic analysis, in terms of its application, is largely regarded as a skill to employ if time and resources permit, which means that it is basically nonfunctioning in many important areas. In career terms, it has been essentially inaccessible to representative numbers of people, being an apprentice slot that must be abdicated by those with more than modest ambitions.



P. 1868636
E.O. 4.4(c)

~~SECRET~~

P.L. 86-36
EO 1.4.(c)

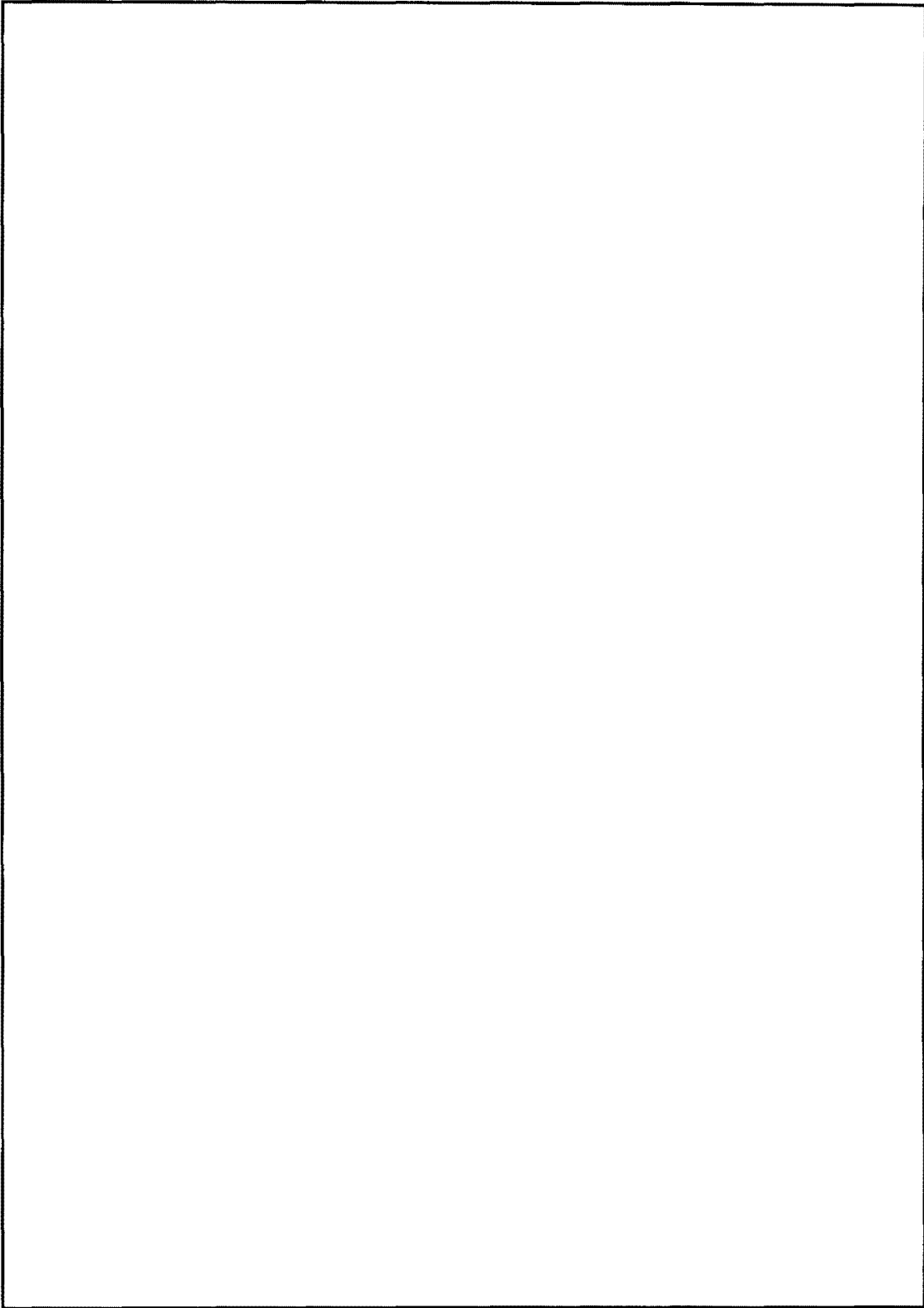


~~SECRET~~

~~HANDLE VIA COMINT CHANNELS ONLY~~

~~SECRET~~

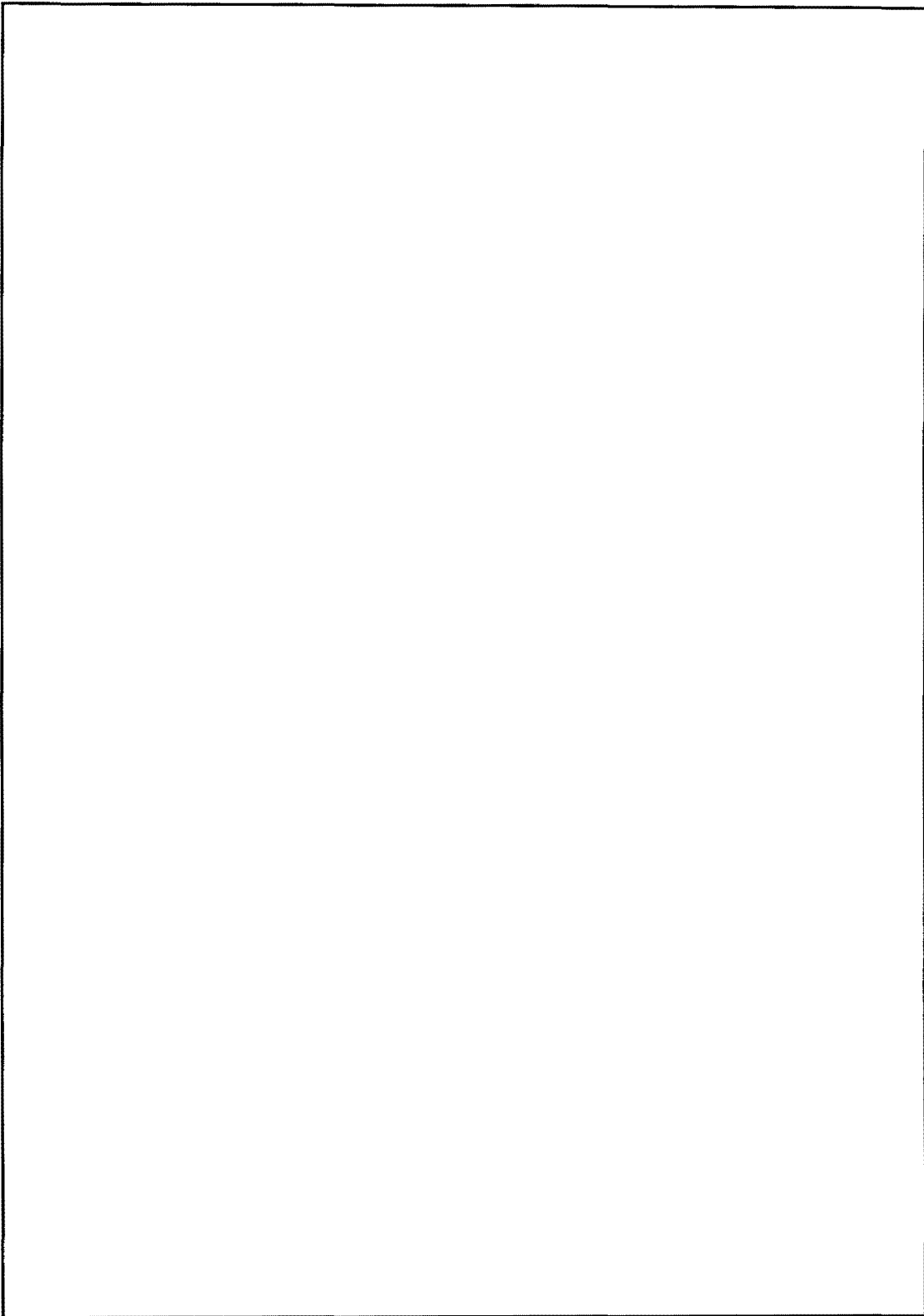
P.L. 86-36
EO 1.4.(c)



~~SECRET~~

~~SECRET~~

P.L. 86-36
EO 1.4.(c)

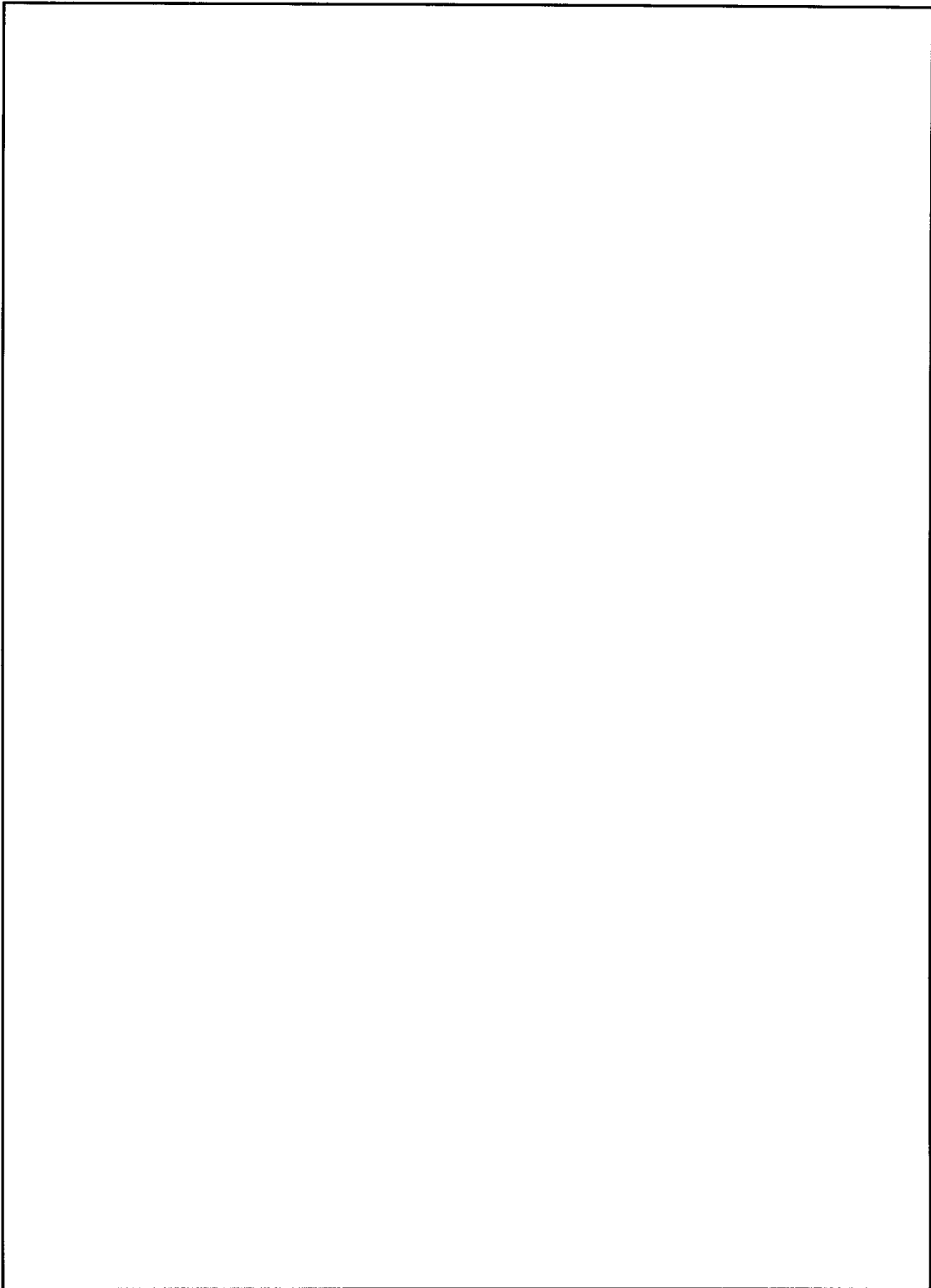


~~SECRET~~

~~HANDLE VIA COMINT CHANNELS ONLY~~

~~SECRET~~

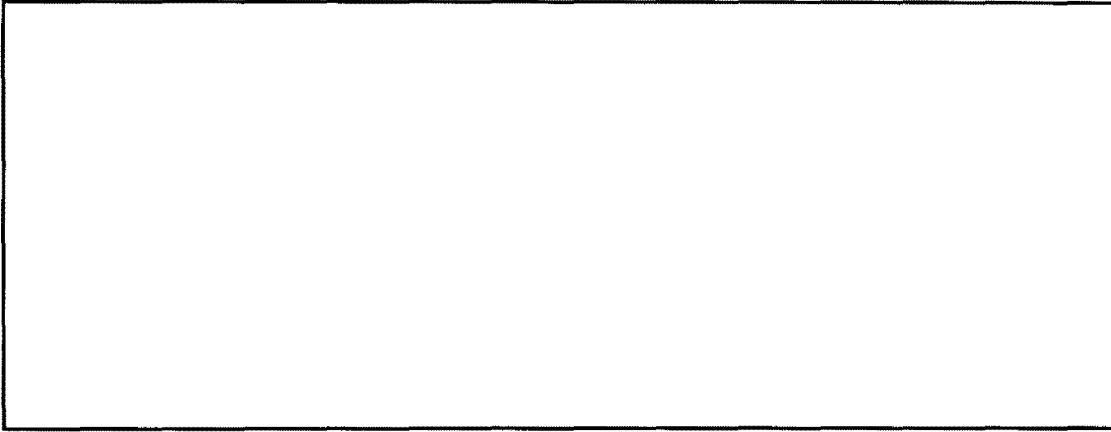
P.L. 86-36
EO 1.4.(c)



~~SECRET~~

~~HANDLE VIA COMINT CHANNELS ONLY~~

~~SECRET~~



P.L. 86-36
EO 1.4.(c)

~~SECRET~~

~~HANDLE VIA COMINT CHANNELS ONLY~~

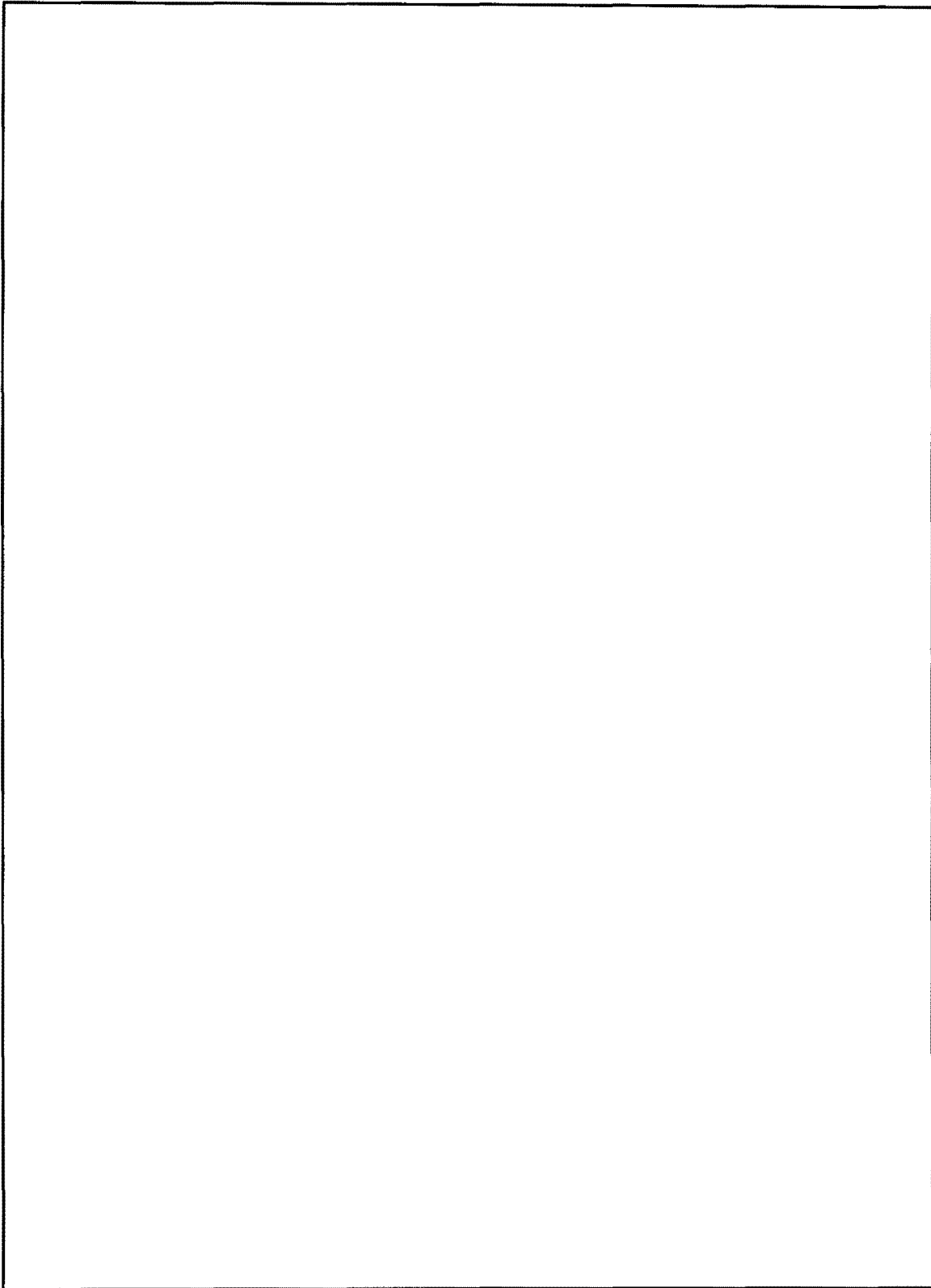
~~SECRET SPOKE~~



~~SECRET SPOKE~~

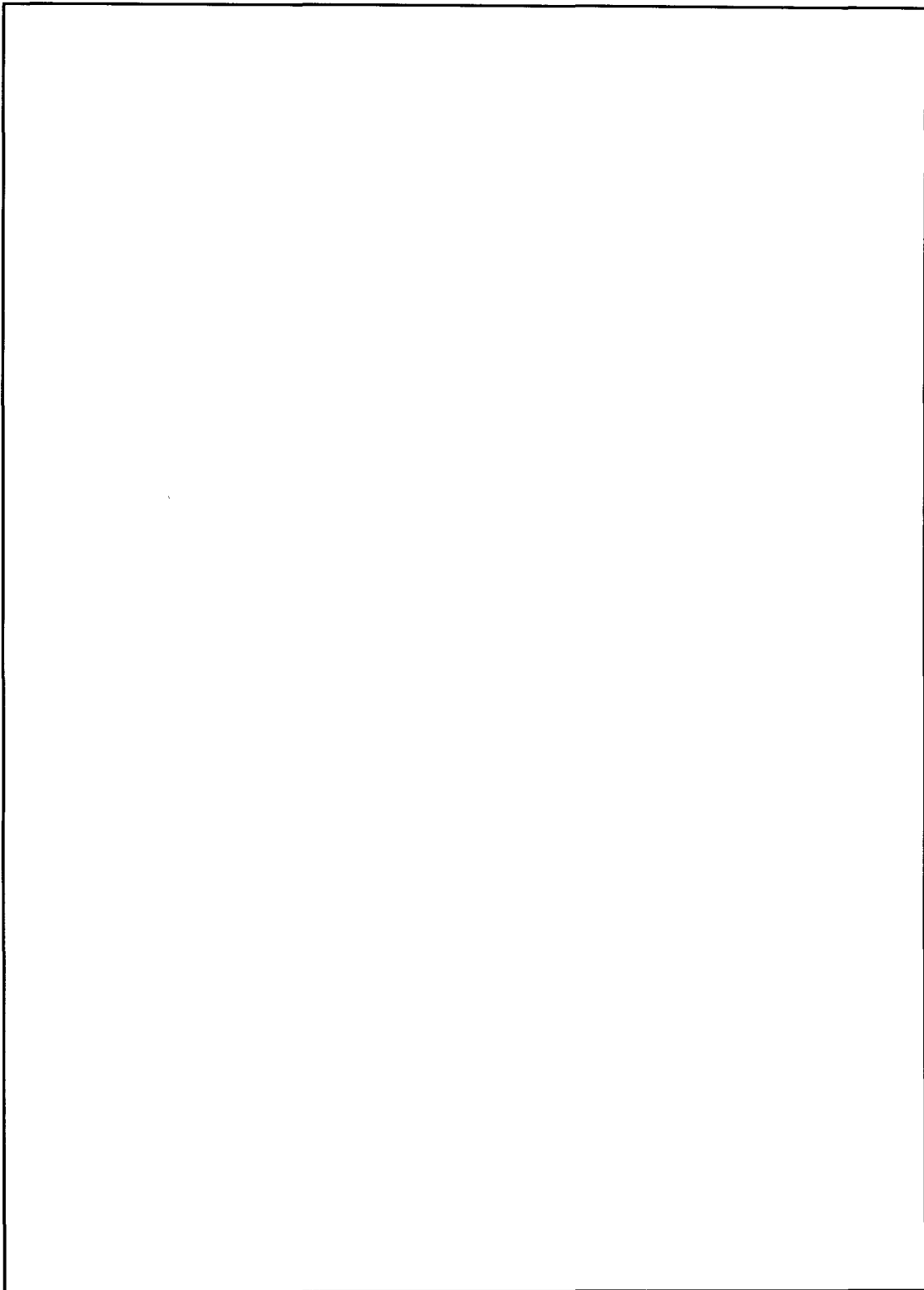
~~SECRET SPOKE~~

PELL 886386
EEO 144(c)
EEO 144(d)

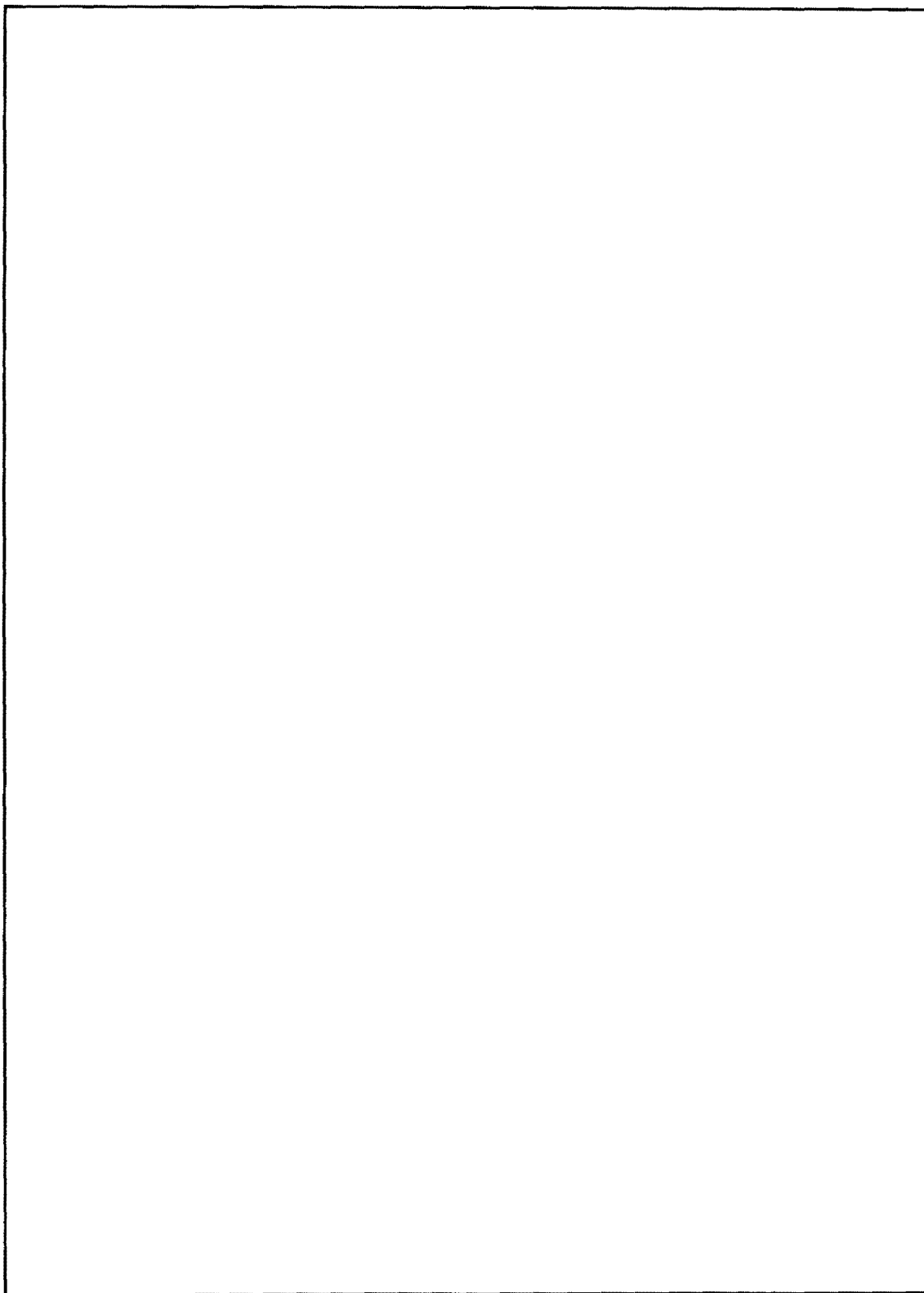


~~SECRET SPOKE~~

PELL 886386
EEO 144(c)

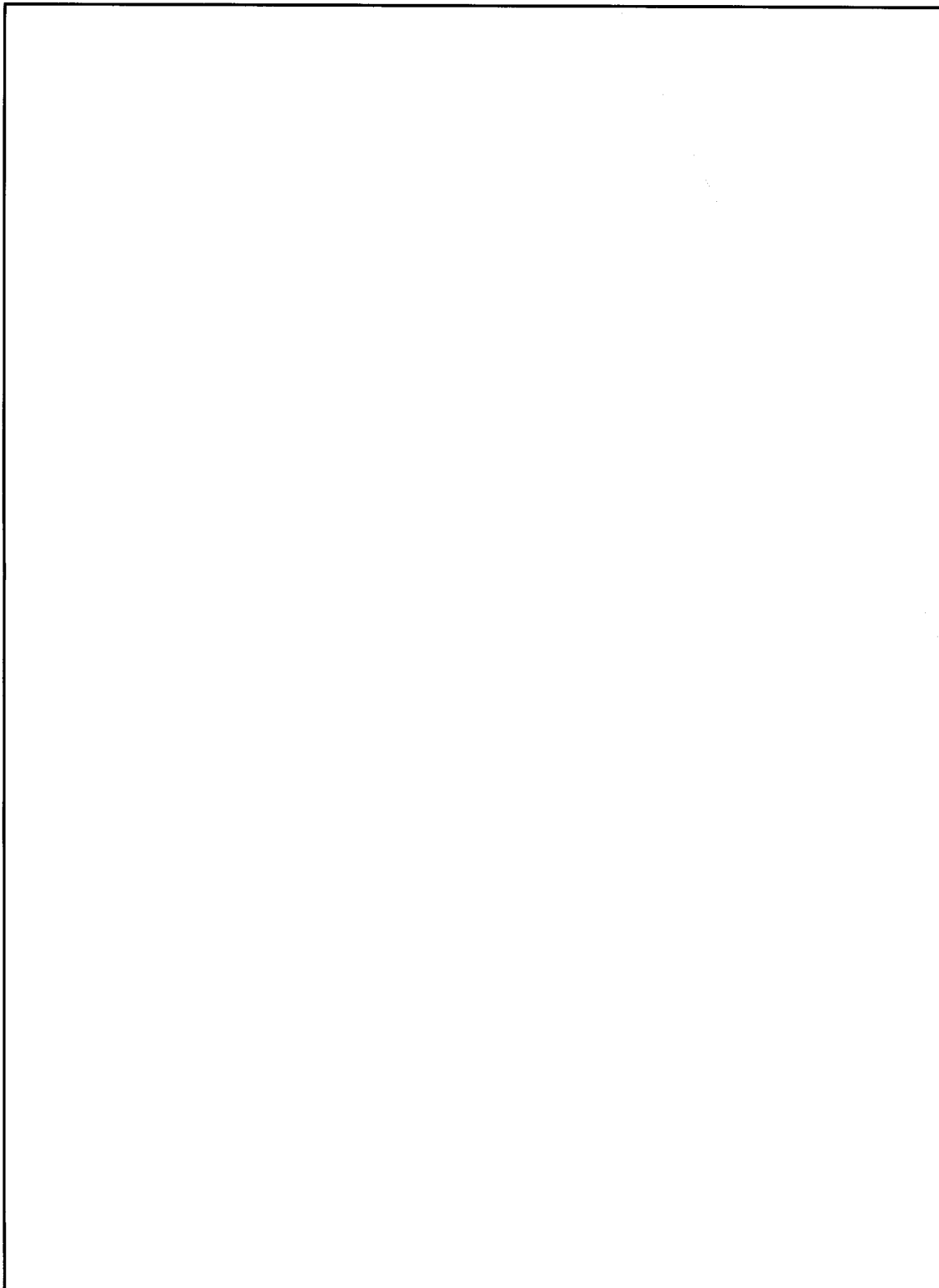


~~SECRET SPOKE~~



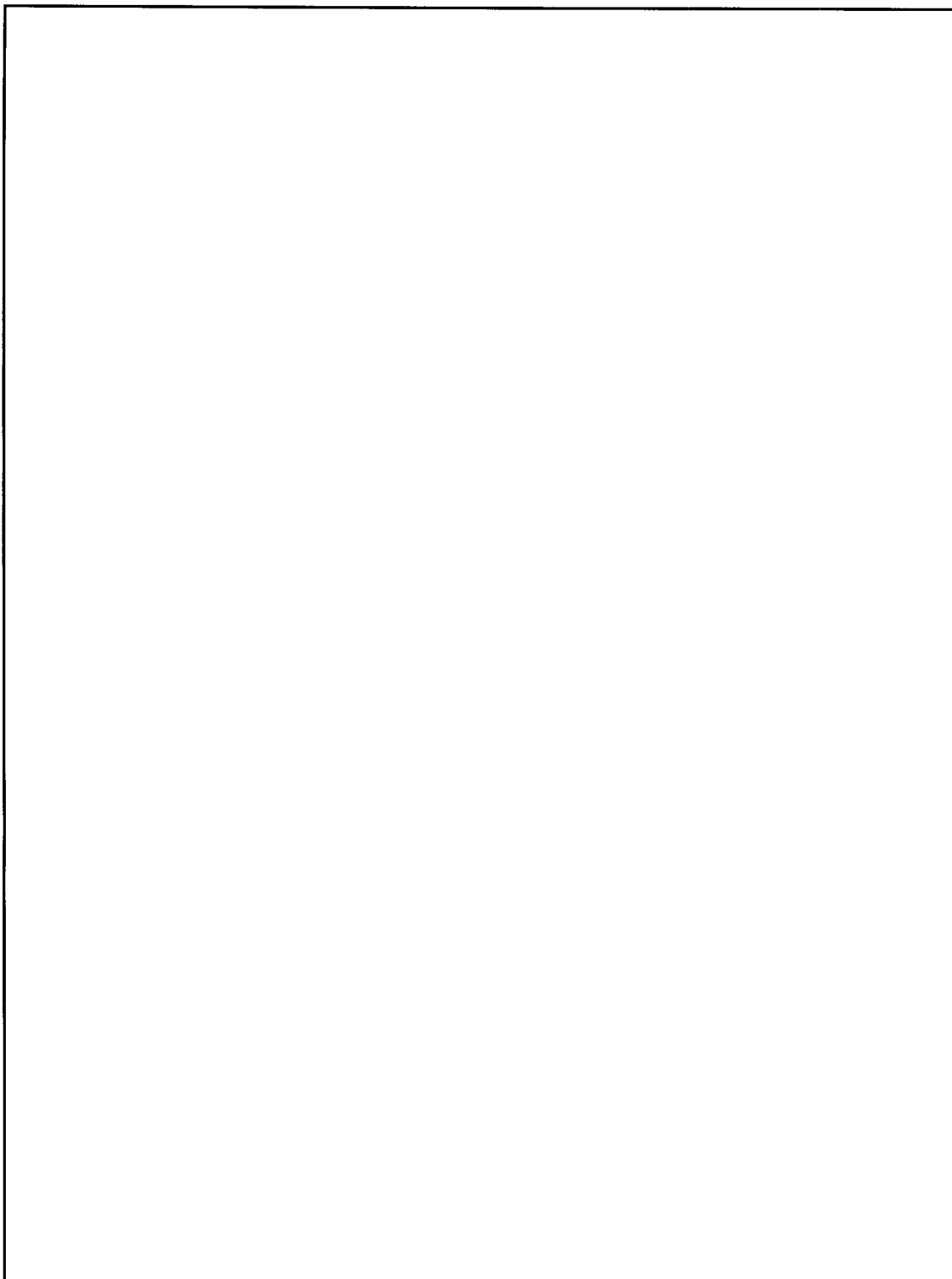
~~SECRET SPOKE~~

~~SECRET SPOKE~~



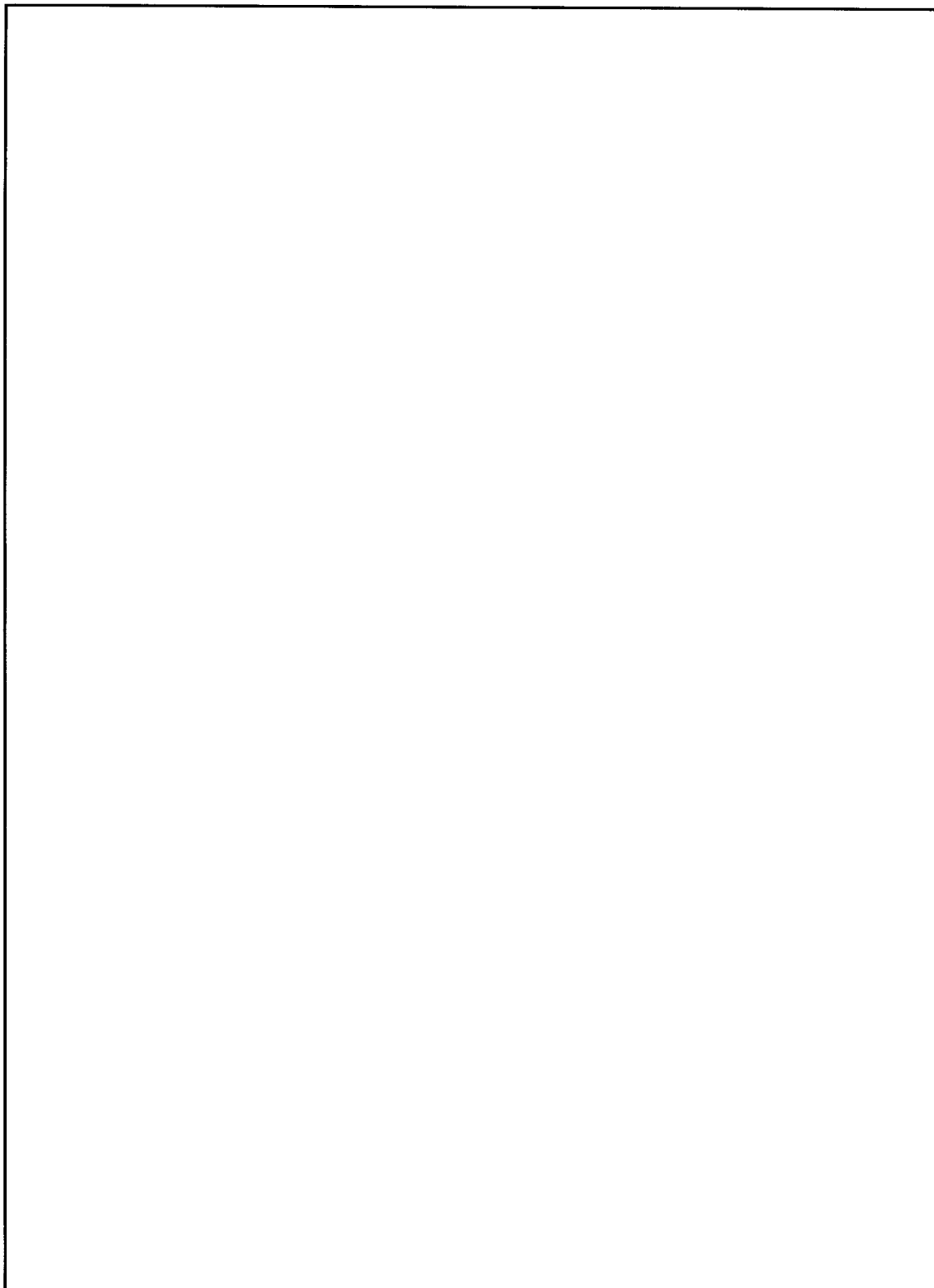
~~SECRET SPOKE~~

~~SECRET SPOKE~~



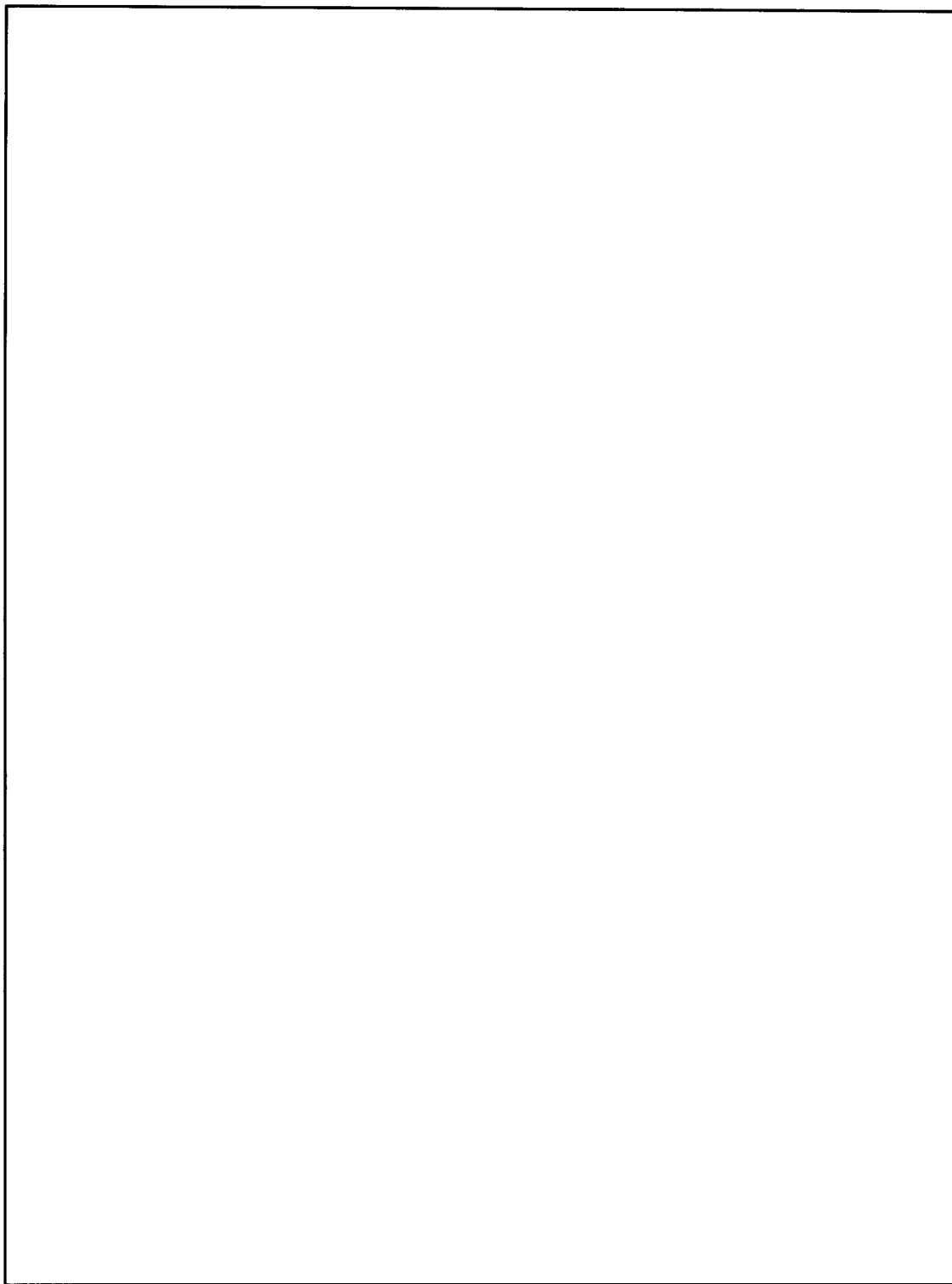
~~SECRET SPOKE~~

~~SECRET SPOKE~~

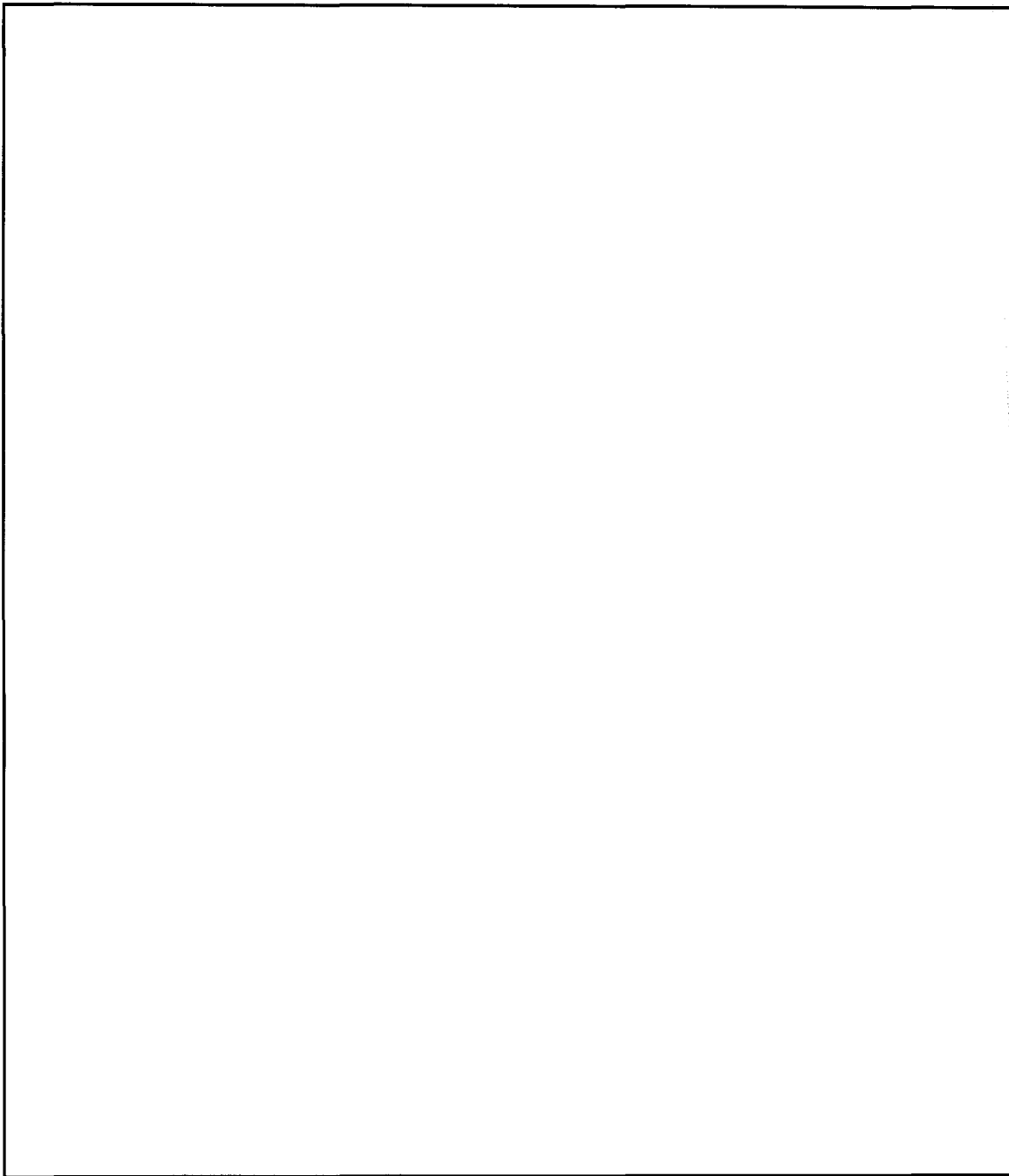


~~SECRET SPOKE~~

~~SECRET SPOKE~~



~~SECRET SPOKE~~



~~SECRET SPOKE~~



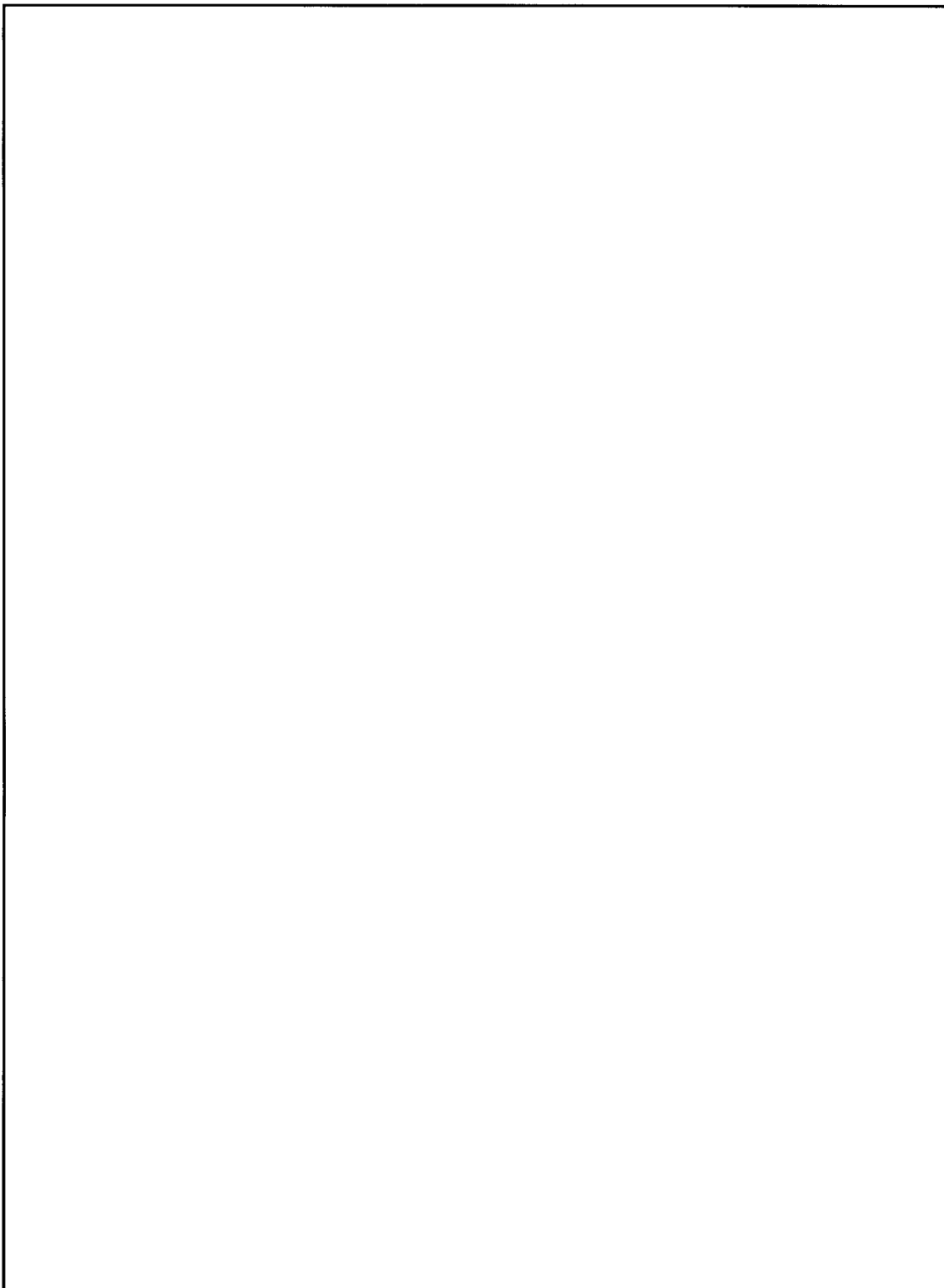
~~SECRET SPOKE~~

~~SECRET SPOKE~~



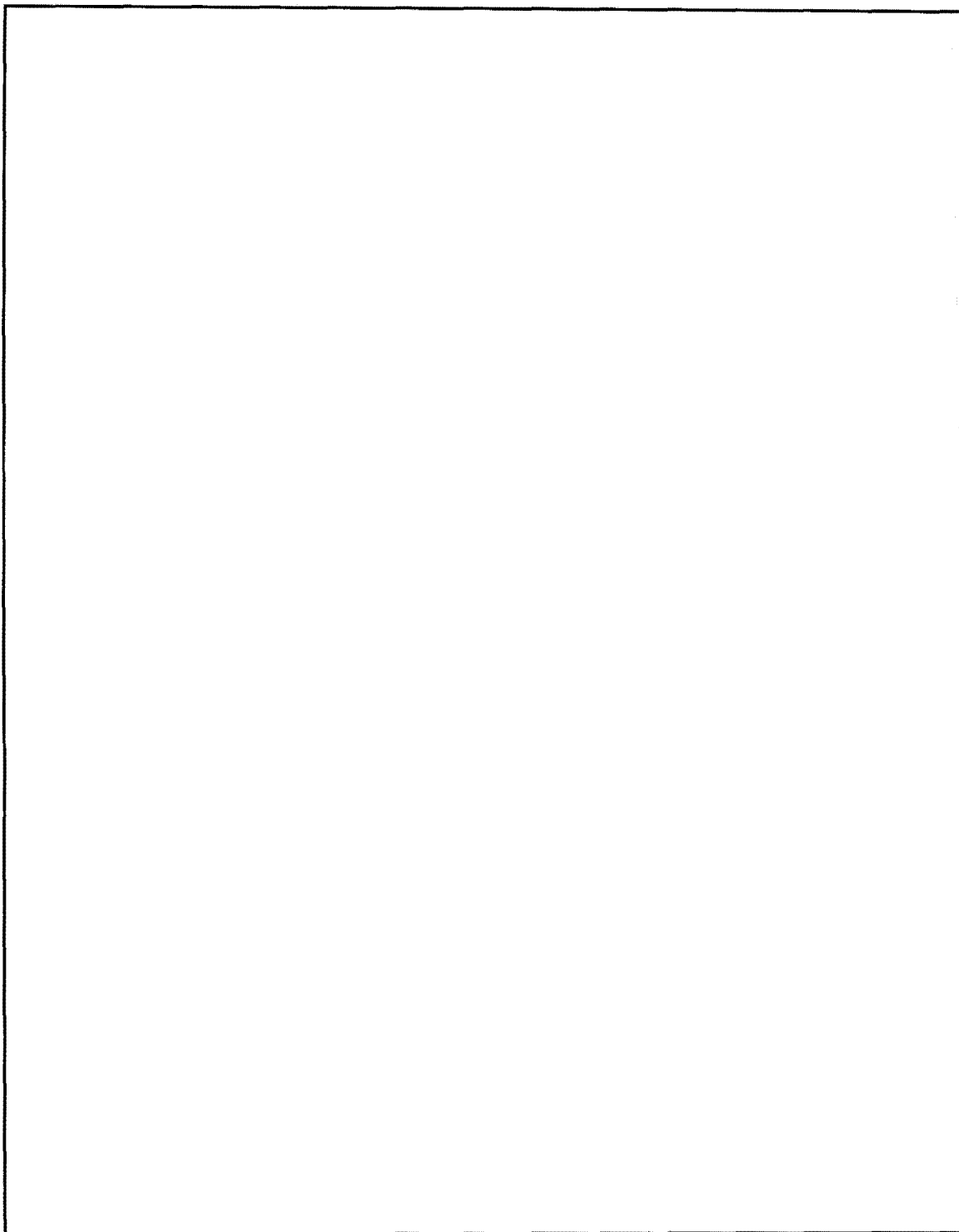
~~SECRET SPOKE~~

~~SECRET SPOKE~~



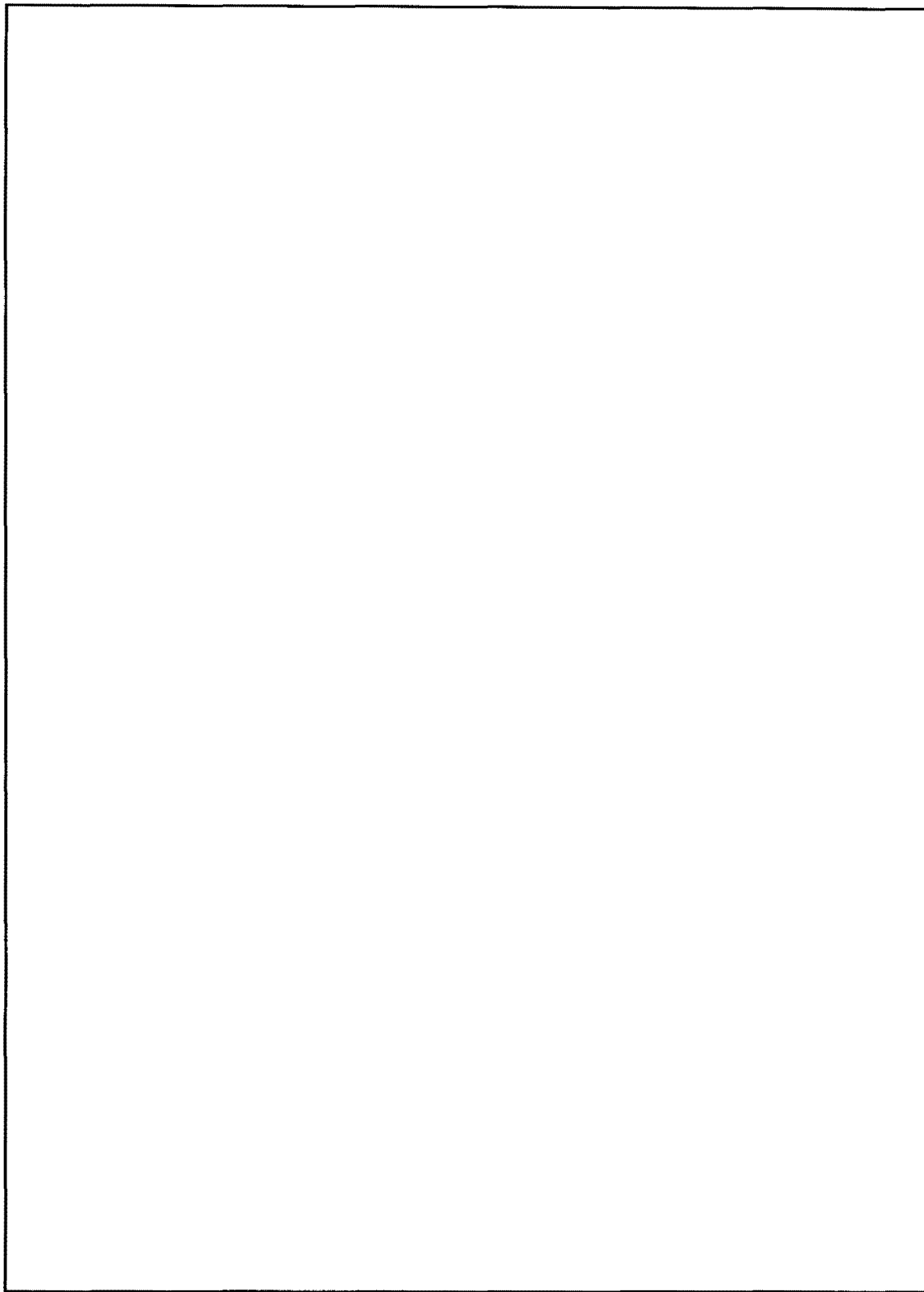
~~SECRET SPOKE~~

~~SECRET SPOKE~~



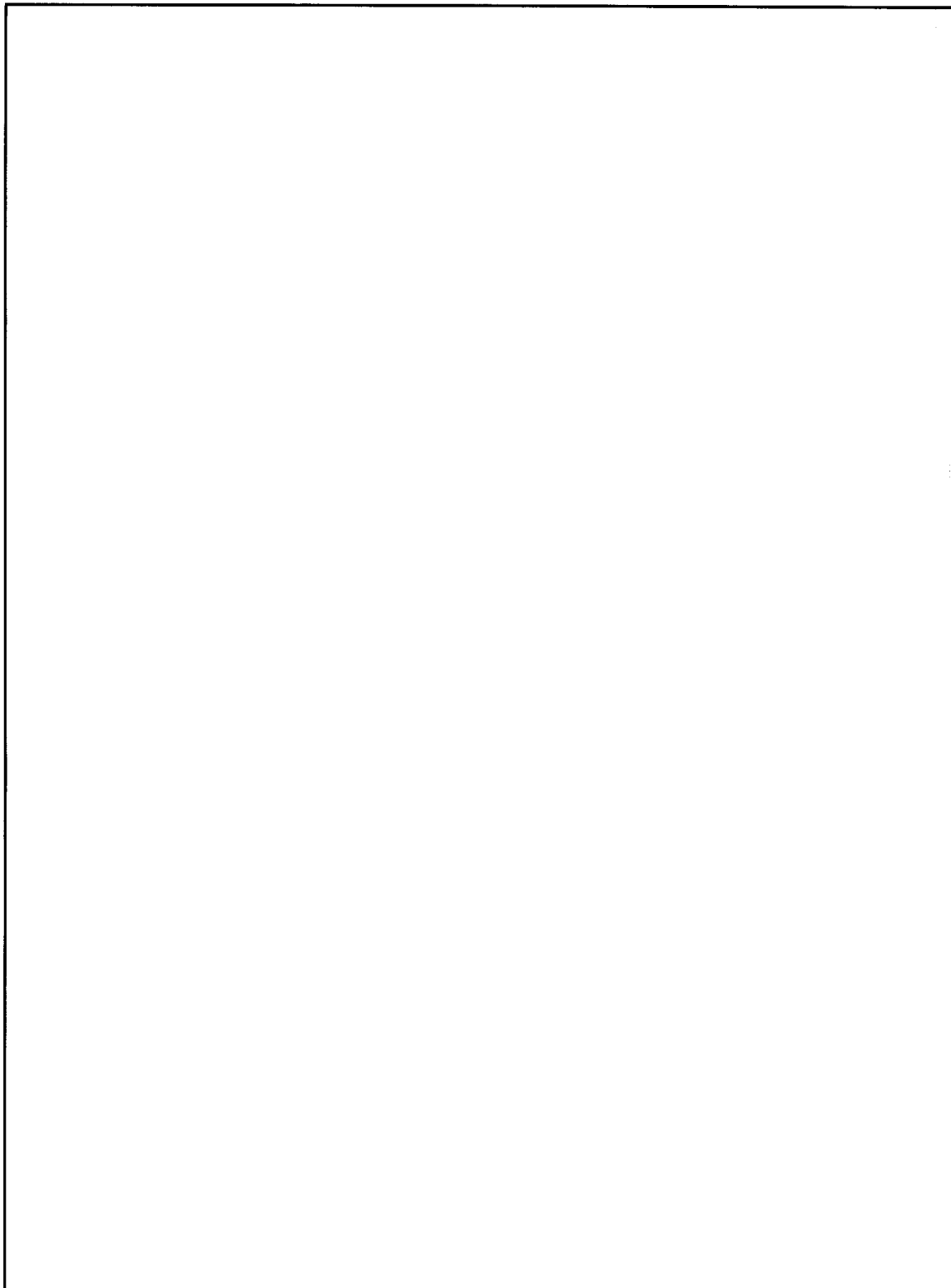
~~SECRET SPOKE~~

~~SECRET SPOKE~~



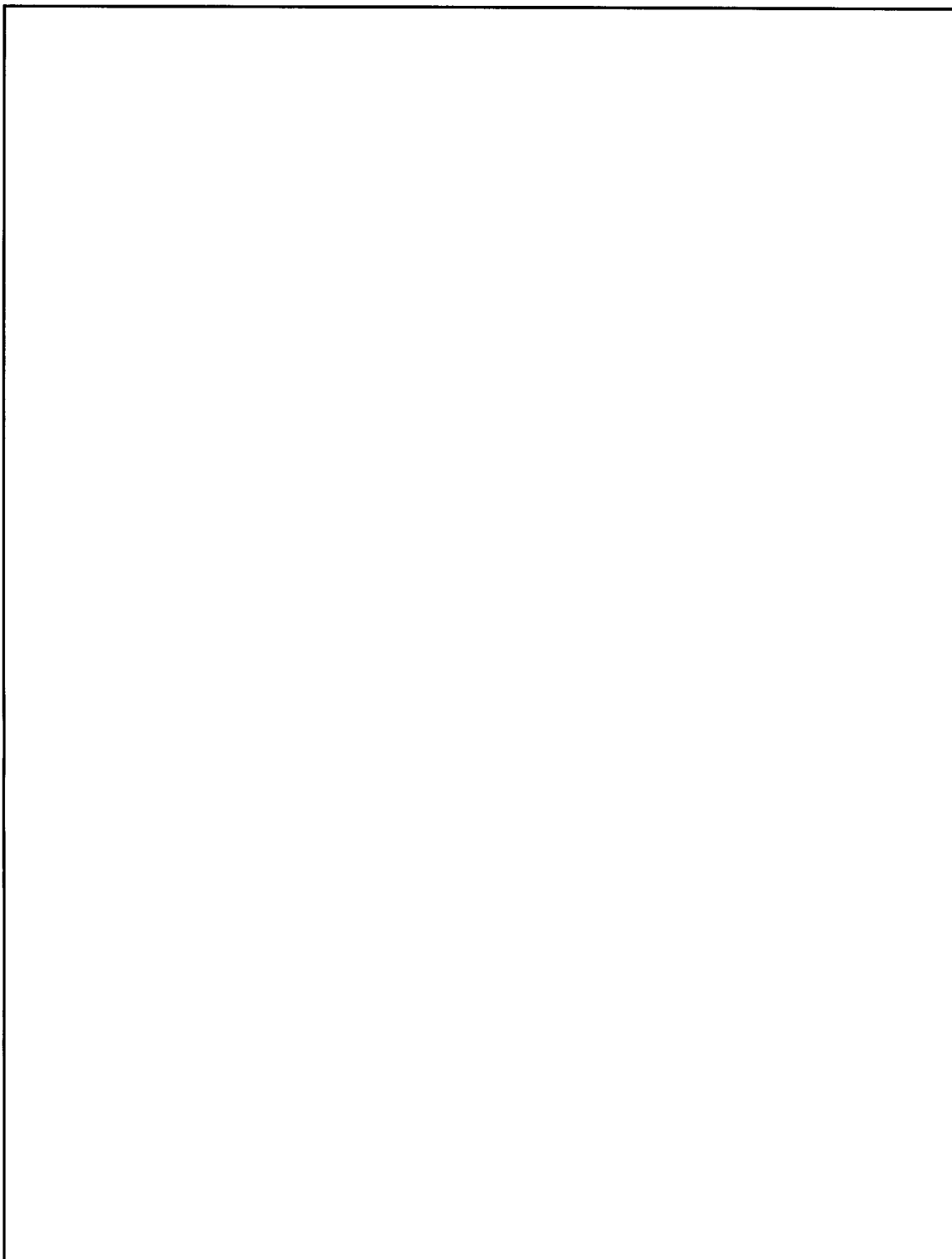
~~SECRET SPOKE~~

~~SECRET SPOKE~~



~~SECRET SPOKE~~

~~SECRET SPOKE~~

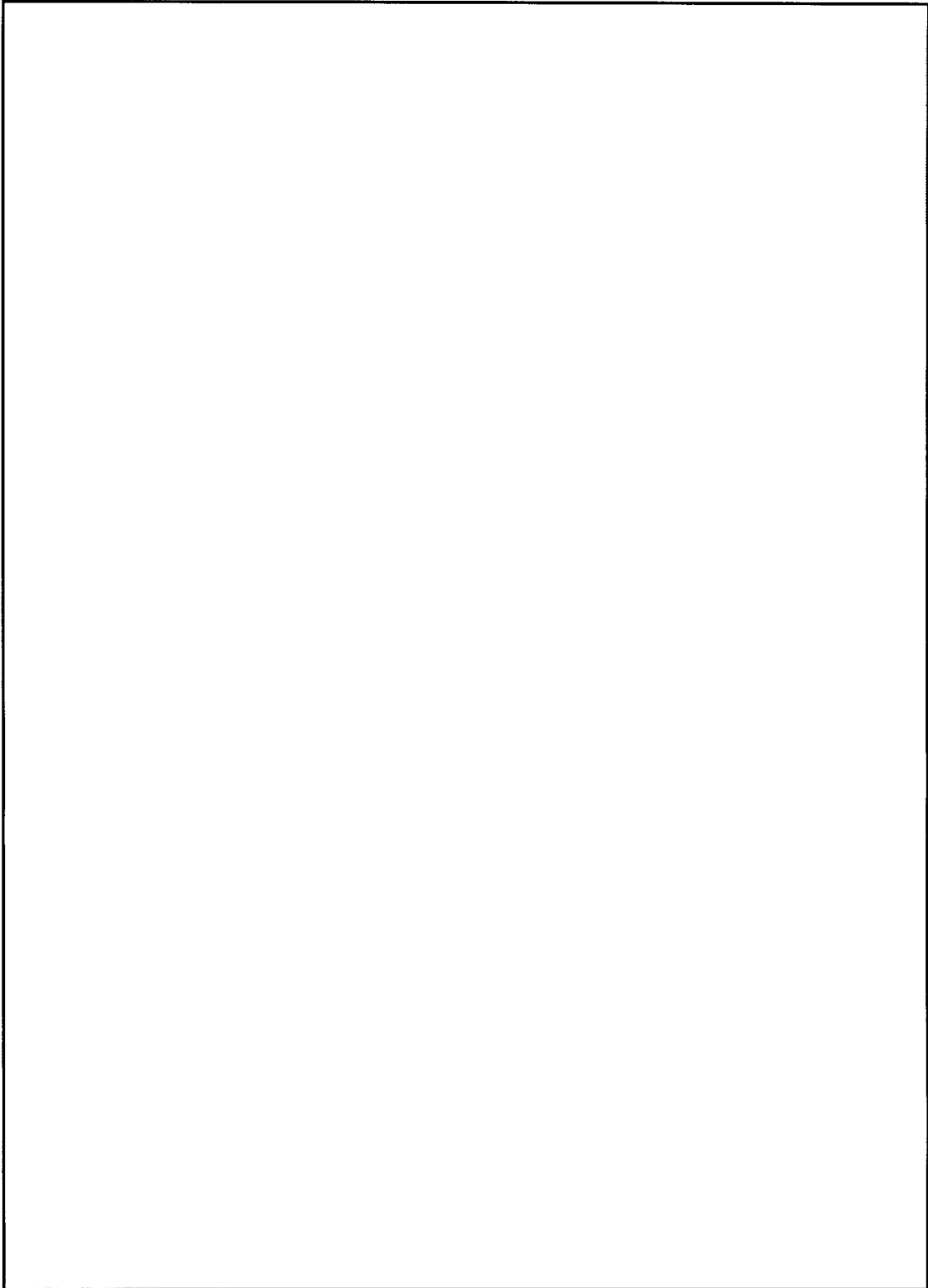


L. 86-36-36
O E (t) 4 (c)
O E (t) 4 (d)

~~SECRET SPOKE~~

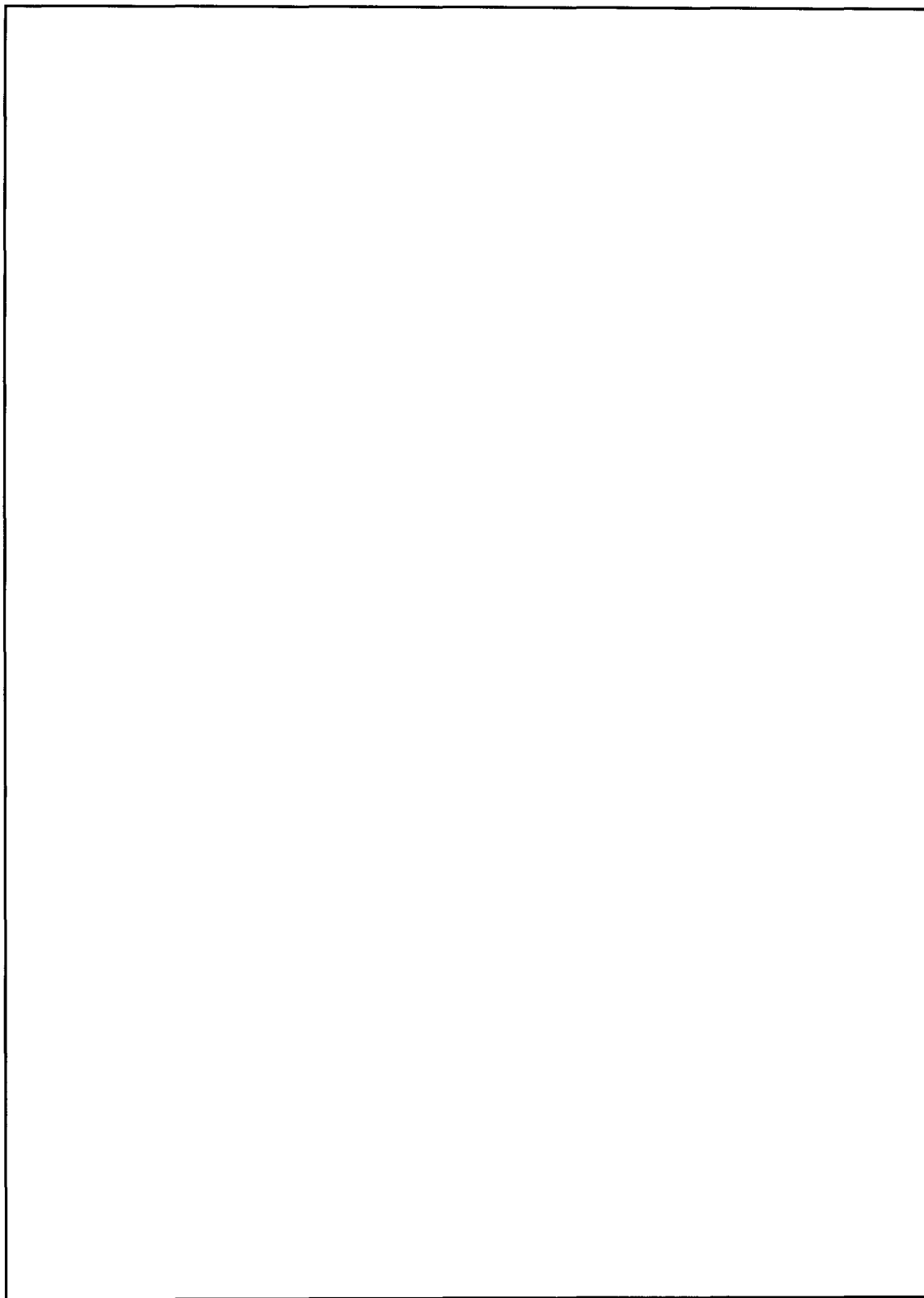
P.L. 68636
EO 1.4(c)
EO 1.4(d)

~~SECRET SPOKE~~



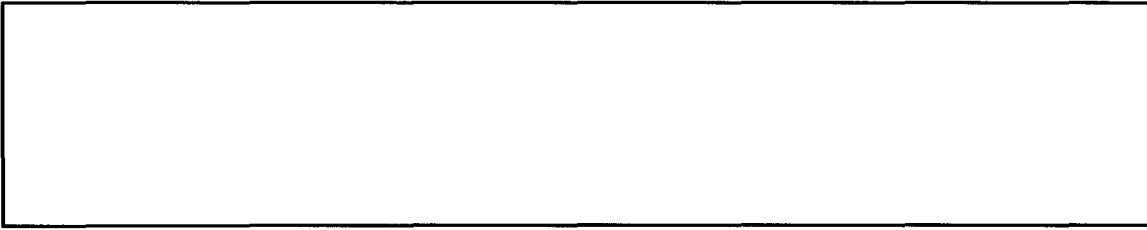
~~SECRET SPOKE~~

~~SECRET SPOKE~~



~~SECRET SPOKE~~

~~SECRET SPOKE~~



~~SECRET SPOKE~~

The Future of Traffic Analysis

P.L. 86-36



It is a pleasure to be here and to be asked to think about the future of TA. That's a subject I've been thinking about, in one way or another, for most of my adult life.


I would like to begin by making some predictions about TA twenty years from today. I found a crystal ball and was going to set it up here on the podium for these predictions. However, the guy that owned it couldn't figure out why I wanted it, so I decided it wasn't worth the trouble.

Here are the predictions.

- *Twenty years from today, people outside the field will be telling us that TA has no future.*

Back in 1946 when I first started, people were quick to tell me that TA was a wartime skill, that it did a fair job during the war, but that it wouldn't be around very long in peacetime. In fact, I can remember not really wanting to be in the field of TA. I was convinced that CA was more interesting, and for a while I did my best to get out of TA so I could work in CA. (Later on, when I got my hands on some actual raw traffic and began to work with collectors, I found that TA was a lot more interesting than I thought.) And people have been predicting the demise of TA ever since. By now it's an old song, but I even hear it today. So it doesn't take much nerve to predict that we will still be hearing that song twenty years down the road.

- *The highest graded individual in the TA field will be a supergrade, or whatever they are calling them twenty years from now.*

This is another easy prediction.  got to be a supergrade as a Traffic Analyst, and so did I. I see no reason why other Traffic Analysts can't do the same. If anything, it ought to be a little easier for you, because the Agency is now more attentive to the technical track than it has been in times past.

- *Most of the daily front page news in the intelligence community will be provided by TA.*

This is another easy one. We do it now. We have been doing it. It seems an easy projection to say that we will continue. We will still be relying on TA for much of our Information [sic] & Warning. There may be some new high-tech wrinkles in the I&W world twenty years from now, but TA has a good track record for reliability and consistency in this area. It is so simple for TA to "reach out and touch" targets to see if

they are normal. Even if some good gadgetry comes along, the result will be some sort of partnership between the new techniques and the old reliable methods of TA.

- *We will still rely on TA for collection steerage.*

It is funny how often this is overlooked in day-to-day operations. And yet some TA shops could almost pay their own salaries in the savings from more efficient use of collection resources. Signals are getting more complex. Some of today's targets have a lot of money to spend and are modernizing faster than our collection system. All of this will require even *more* efficiency for collection steerage, and that's where TA has to come in.



FPL 886386
EEO 1144(c)

This hasn't changed for years, and probably won't over the next twenty. TA people are still "batching" traffic into homogeneous groups to prepare it for cryptanalytic attack. Ask your favorite cryptpies if they would like this to continue. Once cryptanalysts experience good TA support, they become believers; they will ask for it again and again, because they know it helps them to do their own job well.

- *On Tuesday, the 24th of May, in the year 2005, someone will be heard to say, "My problem isn't a standard TA problem – we don't use callsigns."*

I think I first heard this about thirty-five years ago when I started working on a radioprinter problem. I really don't think there is a *standard* TA problem. What we have is one or two problems that are so large, we tend to think of them as standard. But they aren't. Traffic analysts have a large bag of tools, built up over the years. We keep adding to that bag of tools, but we don't throw old tools away. Lately, we've been added some computer tricks to our tool kit, and over the next twenty years, I suspect that other new tools will be developed. We need every tool we can get.

Art Levenson used to say that we were a bit like a man that has lost his wallet on a very dark street. We don't look for it where we think we lost it – in the dark. We look for it under the street light where we can see. We have to work with what the target gives us. And we use whatever tools are appropriate for the features the target gives us.

This leads me to my last prediction, one I hope does *not* come true:

- *Sometime during the next twenty years, your target, or the target you are most interested in, will have a major communications change.*

Maybe the target will find out what we use to exploit his communications, how we do our little magic tricks to produce TA results. Maybe they will have a spy, or maybe some public figure will slip and say too much, or maybe someone will go out and write another

dumb book. Whatever the reason, the target will change its system, perhaps at great cost and inconvenience.

First there will be panic and lot of running around. Then, after things settle down a bit, the traffic analysts will begin to take a look at what the target is giving us to work with. Then they will dip into that big old bag of TA tools and start building another processing system, a system that will feed the collectors, and the cryptanalysts, and the intelligence community. As they recover the new signal plan, what the traffic analysts find will be new, and maybe some of the tools they use will be new. But it will still be TA.

Valedictory of a Traffic Analyst

JOSEPH STARR

P.L. 86-36

(U) On August 1, 1985, I retired from the National Security Agency after almost forty years of cryptologic work, which included at one time or another most of the things NSA does in the fields of collection, analysis, and SIGINT reporting. I enjoyed every minute of it. I honestly don't believe I could have chosen a more enjoyable or rewarding career. If I had been born rich, I would have done traffic analysis for nothing; if I had been independently wealthy, I would have paid the Agency to let me do it. Hang with me for a few minutes of serious discussion about machine processing problems, and I'll share with you a few observations, comments, and recommendations that may contribute to your enjoyment of the profession and perhaps help you to avoid some of the pitfalls that beset the paths I wandered in.

(U) The most important question confronting the profession today is why, from a traffic analyst's viewpoint, data processing systems have not yet produced the Utopia we were led (or stampeded) to expect back in the early 1960s.

(U) In the January-March 1980 *Cryptolog*, [] correctly sorts working analysts into three categories: loggers, case analysts and research analysts. He then predicts that automated processing techniques will permit us to eliminate the loggers, and assesses that the research analyst is the one with the bright future because he is concerned with the "why of analysis...driven by his desire to explain." But every analyst should be saying why, and George's prediction that the logger will become obsolete just hasn't happened, although it should have. Indeed, we now have a new generation of loggers who have even less understanding of traffic analysis than those who preceded them.

(U) After George discusses how analysts are going to have to adapt in order to be productive in a modern (machine-oriented) world, he puts his finger squarely on the major problem area, "manageable machine systems that will function as designed." And he goes on to point out, "... our track record for the development of such systems is not impressive."

(U) In the December 1981 *Cryptolog*, [] published an article on machine processing that everyone who claims to be a traffic analyst should have read. Part of what Dale addressed was the problem with keeping competent programmers on the job, and how the traffic analyst could be freed from total dependence on the programmer.

(U) Traffic analysis is basically a very simple art if you are able to keep your mind loose so that you can recognize and exploit whatever the target gives you to work with. This is a fairly easy exercise intellectually, but it is much more difficult to structure your thought processes in order to design a machine processing system that will give you essentially the same versatility. Unfortunately, we have some managers and analysts who have produced processing systems and tried to make the target fit the processing

scheme, and I include in this category most of those procedures that were designed to replace a hand log kept by the analysts with a similar log prepared by the machine. The results were predictable. What we have now are logs prepared by machine that in many instances are inferior to those the analyst prepared by hand. And, without the benefit that accrued from working the traffic by hand, today's analyst has even less understanding of the data and may in fact be at a total loss to explain observed anomalies.

~~(C)~~ We can reverse this trend. We need to build on the good decisions we have made with regard to automating data processing for traffic analysis, and we must be willing to discard the bad decisions. What we really need to do now is to get on with developing an expert system that will process automatically all incoming traffic, compare what is observed to a dictionary of stated norms, and send appropriate alert messages to analysts and reporters on trouble spots.

(U) None of this should be difficult. Building and testing dictionaries will be tedious, but in the process of doing this, the analyst will be compelled to learn a great deal more about his cases than he now knows, and supervisors will be compelled to work with analysts to a much greater extent than is taking place now. When an expert-type system is in place and functional, one of the benefits will be that time will become available for a different and much more meaningful kind of on-the-job training.

~~(S-CCO)~~ We began exploring how to develop an expert system for processing
 just before I retired. I expect initial development to be slow, painful, and frustrating, but I do expect the system to "function as designed." What we must have to make it work is a sufficiently confident estimate of our analytic judgment and capabilities to tell the machine how to do *all* the dog work for us and then let it do the job. There can be no doubt as to our ability to make the right technical and reporting decisions based on the alert messages because each analyst must understand his responsibilities well enough to have described his case, net, and network norms to his analytic dictionaries.

(U) We will see vastly improved databases as a side benefit of this kind of processing. If the input data aren't good, the expert system is going to make noise; it will not permit any analyst to do a poor job on editing input data without sounding off.

(U) I am not persuaded that an expert system can be developed and implemented by evolution; probably the best (and certainly the quickest) way to effect such radical changes in the way traffic analysis is accomplished would be to select the right analysts and programmers, charge them with developing an expert system to do traffic analysis on a specific target, and set a deadline of perhaps a year hence for the initial job to be finished.

There should be no restrictions as to how the task is to be accomplished, and there should be no arbitrary limitations as to the techniques to be used (e.g., storage of norms for ordinary comparisons should not be made excessively difficult by the intricacies of an existing database).

(U) There is no question that the Agency will develop an expert system at some point in the future, if not for normal traffic analytic functions, then to provide rapid data evaluations for tactical support of military forces. I regret that I will not be here to see it.

(U) Here is the advice I promised earlier.

(U) Learn to be persuasive as well as informative in any presentation you make. The best ideas you have aren't worth much if you can't sell them. You cannot depend on decision makers to steer the right course because it is (or appears to you to be) the logical thing to do.

~~(FOUO)~~ Reporting the results of your analysis is part of your job. Learn what the intelligence requirements are for the target you are assigned and relate them directly and specifically to how you work your traffic. If you don't understand what it is that you are supposed to be getting out of the traffic and don't try to find out, you are still a logger, and that's all you're ever going to be.

(U) Avoid meetings. I am convinced that many people at the Agency really and truly believe that they are making progress as long as they are talking with each other regardless of whether anything is actually being done. I have attended some meetings where the only concrete thing accomplished was an agreement to meet again at some specified date in the future to discuss the same subject. And so help me, they all left happy.

(U) Don't take sinful pride in your own words. If the staff officer who reviews your message or report wants to change your words for his, don't be offended. Staff officers do serve a useful function. If you have treated them with respect in the past, they will often help you get out of trouble.

(U) I finished this article just as my NSA career came to an end. Reading it over, I find a line or two that may be mildly objectionable to one person or another. It was not my intention to offend anyone by my choice of words or by my comments, and if I have given offense, I apologize. This doesn't sound like me, but it does sound nice, and I think I'll leave it right there. Goodbye, good luck, God bless you all.

~~(S-CCO)~~ Upon joining the U.S. Navy (WAVES) in February 1943, Mrs. Filby trained in aerology at Lakehurst Naval Air Station. In the summer of 1944 she was stationed at the Naval Communications Annex (NCA), where she accepted an appointment as a civilian in May 1946 and served as an analyst, translator, reporter, and staff product and technical editor in the weather branch until 1958. This period included a tour with the [REDACTED] (1953-55). From 1958 until 1964, Mrs. Filby served as an analyst, cryptolinguist, and reporter in Soviet naval and merchant shipping problems. This period included assignment to the Military Cryptanalytics course (taught by Lambros Callimahos), a year-long Advanced Intensive Russian course, and the pilot SIGINT reporting course at the National Cryptologic School (NCS). She was instructor and developer of the NCS SIGINT reporting course (1964-83), organizer of a two-day seminar on SIGINT reporting including Second Parties (1978), and designer of the first SIGINT reporting course for the field (1982). Mrs. Filby was chief of the Intelligence Skills Division, Intelligence and Analysis Department (E4) at the National Cryptologic School (1983-86). She continued teaching and presented the field SIGINT reporting course in [REDACTED]

[REDACTED] In 1986 she was assigned as special assistant to the Dean, E4. She studied the status of SIGINT reporting throughout the Agency; developed a course on the National SIGINT Operations Center; and developed and managed seminars on SIGINT Users, Support to Military Operations, and Current Issues in SIGINT Reporting. In July 1991, Mrs. Filby was assigned to the Center for Cryptologic History. She has been president of the Crypto-Linguistic Association (1975-1976) and has received several awards: NCS Teacher of the Year (1970); the Meritorious Civilian Service Award (1972); and the CIA Sherman Kent Award for the most outstanding contribution to the literature of intelligence (first NSA winner, 1983).

FPIL 886386

EO 1144 (C)

EO 1144 (C)

SOURCES

[] "How Many Angels Can Stand on the Head of a Case Notation?" *KEYWORD*, November 1972, 4.

Benjamin, Robert S. "Some Thoughts Concerning Traffic Analysis Mechanization." *NSA Technical Journal*, 1966, 97-101.

[] "Introduction to Traffic Analysis Mechanization." Traffic Analysis Mechanization Forum, February 1971, Ref. S # 194,544, 4-22.

[] "Applications of Set Theory to Traffic Analysis." Traffic Analysis-Mathematics Symposium, May 1973, Ref. S # 208,891, 202-205.

Bjorklund, Kathy. Letter on reassimilation of traffic analysts. *Cryptolog*, June 1979, 17.

[] Letter on personnel assignments. *Cryptolog*, July 1979, 14-15.

Buckley, Dan. Letter on TA overages. *Cryptolog*, July 1979, 14.

[] "Traffic Analysis or [] Data Transmission Systems." *COMMAND*, October 1968.

[] "Area Studies and Their Place in Traffic Analysis." *COMMAND*, October 1968, 14-16.

Delaney, James D. []
[] *NSA Technical Journal*, Fall 1971, 91-106.

PPIL 886-386
EOD 1144(0)

[]
COMMAND, March 1971, 31-52.

[] "TA, CA, LOGIC, MATH - Where Do They Intersect? (TACALOMA)." Traffic Analysis-Mathematics Symposium, May 1973, Ref. S # 208, 891, 226-246.

[] "Simplicity in Color." *Cryptolog*, February 1982, 24-25. (with title "Gary's Colors," *Cryptolog*, September 1974; *COMMAND*, October 1971.)

Gilbert, Allen L. "The Impact of ARDF on Traffic Analysis." *DRAGON SEEDS*, December 1971, 7-8.

[] "Maybe It's Related to the Phases of the Moon." *DRAGON SEEDS*, June 1972, 5-12.

Hanyok, Robert J. Letter on TA career field. *Cryptolog*, May, 1985, 10.

Hooper, James P. "Traffic Analysis: A Current Perspective." *Cryptolog*, June 1984, 1-4.

[redacted] ~~(S-CCO)~~ *Cryptologic Quarterly*, Spring 1991,
87-105

[redacted] "Traffic Analysis: Write it Down Now." *COMMAND*, February
1974, 31.

P.L. 86-36

Jollensten, Ralph E. "The Lost Indicator." Traffic Analysis-Mathematics Symposium,
May 1973, Ref. S # 208,991, 106-116.

[redacted] "Traffic Analysis: Specialty Without Portfolio." *Cryptolog*, April-
June 1981, 15-18.

Mason, Fred. "Cleaning the Augean Stables or How Much TA Can a Computer Do?" *NSA
Technical Journal*, Summer 1968, 119-127.

_____. Letter on data queries. *COMMAND*, March 1969, 25.

_____. Letter on human confirmation of computer data. *COMMAND*, August
1970, 41.

Morrison, John E. Opening address, Traffic Analysis Mechanization Forum. February
1971, Ref. S # 194,544, 1-2.

Murphy, Tim. "Automation of a TA Process." *Cryptolog*, October 1975, 33-35.

[redacted] "Let's Not Lose Our TA Skills." *Cryptolog*, March 1979, 11-12.

_____. "There's a New World Coming - Are You Ready?" *Cryptolog*, January-
March 1980, 1-2.

[redacted] "Traffic Analysis of the Future." *COMMAND*, October, 1968, 12-
13.

_____. Letter on use of computers. *COMMAND*, August 1970, 40.

[redacted] "The Reality of Communications Changes." *DRAGON SEEDS*, June 1972, 13-
16.

[redacted] "In Pursuit of Faster Horses, Younger Women, Old Whiskey, and More
Money." *Cryptolog*, December 1981, 16-18.

[redacted] "Establishing Communications Norms." *COMMAND*, August 1969, 18-
22.

Shaw, R. H. "What Good is Traffic Analysis?" *NSA Technical Journal*, Fall 1979, 68-73;
from P1 Informal No. 10.

[redacted] Letter on TA overage. *Cryptolog*, August-September 1979, 17-18.

Starr, Joseph. "Valedictory of a Traffic Analyst." *Cryptolog*, June-August 1985, 9-10.

Stoffel, Wayne E. "Chatter Patterns: A Last Resort." *NSA Technical Journal*, October
1957, 63-75.

~~SECRET~~

[REDACTED] "A Note About NR's." *COMMAND*, August 1969, 11-12.
[REDACTED] "Recovery of a Vietnamese Communist Callsign System."
DRAGON SEEDS, December 1971, 5-6.

REF ID: A66386

[REDACTED] "A Note About Organizing T/A Problems." *COMMAND*, March 1974.

[REDACTED] "How Clean Does a Data Base Need to Be?" *Cryptolog*, January 1975.

[REDACTED] "Computerizing Traffic Analysis." *Cryptolog*, May 1983, 36-43.
[REDACTED] "The Hand Is Not Quicker than the Eye." *Cryptolog*, March, 1978, 9.

[REDACTED] *Cryptolog*, September 1978, 9.

REF ID: A66386
REF ID: A66386

[REDACTED] "True Base: Two Tales." *Cryptolog*, May 1982, 15-19.

[REDACTED] "The Future of Traffic Analysis." *Cryptolog*, May 1983, 9-10.

~~SECRET~~

ADDITIONAL ARTICLES ON TRAFFIC ANALYSIS IN NSA SOURCES

[] "Model 204 - The Analyst's Computer." *NSA Technical Journal*, Summer 1976, 65-75.

Benjamin, Robert S. "Overview of Mathematics Support to Traffic Analysis." Traffic Analysis-Mathematics Symposium, May 1973, Ref. S # 208,891, 3-17.

[] "The Origination and Evolution of Radio Traffic Analysis, the World War I Era." *Cryptologic Quarterly*, Spring 1987.

[] "The Origination and Evolution of Radio Traffic Analysis, the Period between the Wars." *Cryptologic Quarterly*, Fall-Winter 1987-88.

[] "The Origination and Evolution of Radio Traffic Analysis, World War II." *Cryptologic Quarterly*, Winter 1989.

Cahill, Robert. "...Be Considered a Single Field?" *KEYWORD*, July 1973, 5.

Chiles, Gloria. "TA Draw." *COMMAND*, March 1971, 23-26.

Callimahos, Lambros D. "Introduction to Traffic Analysis." *NSA Technical Journal*, April 1958, 1-11.

P.L. 86-36
EO 1.4.(c)

[] "The Cryptography of [] Traffic Analysis-Mathematics Symposium, May 1973, Ref. S # 208,891, 59-71.

[] "The Basic Probability Concepts and Their Applications to Traffic Analysis." unpublished article in TA library collection.

Dale, D. E. Letter on intelligence reporting. *COMMAND*, March 1971, 53.

[] "Computer Aided Traffic Analysis in [] Division." *Cryptolog*, October-November 1986, 1-8.

P.L. 86-36
EO 1.4.(c)
EO 1.4.(d)

Donym, Sue. "Second Sighting." *Cryptolog*, January 1979, 12.

[] "Squares." Traffic Analysis-Mathematics Symposium, May 1973, Ref. S # 208,891, 93-104.

Gilbert, Allen L. "Mechanization for T/A Development." *DRAGON SEEDS*, September 1972, 14-15.

[] "How the Traffic Analyst Can Help Others." unpublished article in TA library collection.

"The GOLD NUGGET Award for Excellence in Traffic Analysis." *Cryptologic Quarterly*, Winter 1989, xv-xviii.

[] "Is There a Doctor in the House?" *Cryptolog*, September 1977, 9-10.

[] "Can Traffic Analysis and Special Research....." *KEYWORD*, July 1973, 4.

~~SECRET~~

[redacted] Jackson, William J. "TA Management Information System." Traffic Analysis Mechanization Forum, February 1971, Ref. S # 194,544, 213-224.

Jackson, William J. "TDB: The TEXTA Data Base." *Cryptolog*, August 1974, 4.

[redacted] "An Approach to Callsign Analysis." *Cryptolog*, December 1974, 7.

[redacted] "Net Reconstruction - A Basic Step in Traffic Analysis." *NSA Technical Journal*, July 1958, 31-45.

[redacted] "Collection-Support TA Is Not for Everyone." *Cryptolog*, February 1978, 7-8.

[redacted] "Go-Go Traffic Analysis." *COMMAND*, October 1968, 10-11.

[redacted] "AG-22: Where Do We Go Now?" *DRAGON SEEDS*, December 1972, 22-23.

[redacted] "Use of Displays in Traffic Analysis." Traffic Analysis Mechanization Forum, February 1971, Ref. S # 194,544, 72-82.

[redacted] "Collection Data Services Support to TA." Traffic Analysis Mechanization Forum, February 1971, Ref. S # 194,544, 40-46.

[redacted] "The Desk Analyst's Math." Traffic Analysis-Mathematics Symposium, May 1973, Ref. S # 208,891, 49-58.

[redacted] "TA Continuity." *COMMAND*, March 1969, 14-24.

[redacted] "Where Are You Going and How Do You Know?" *COMMAND*, August 1970, 35-39.

[redacted] Letter on intelligence in TA technical data, *COMMAND*, August 1970, 42.

[redacted] "Normal Reporting." *COMMAND*, March 1971, 28-30.

[redacted] Letter responding to comment on reporting. *COMMAND*, March 1971, 53.

[redacted] "Some Notes on TA Management." *COMMAND*, October 1971, 32-35.

[redacted] "After Continuity, What?" Traffic Analysis-Mathematics Symposium, May 1973, Ref. S # 208,891, 18-21.

[redacted] "Traffic Analysis of the Future." *Cryptolog*, June 1979, 10.

[redacted] TEXTA: What Is It? Where Is It Going?" *Cryptolog*, December 1981, 19-25.

Taylor, James W. "Traffic Analysis of the Present." *COMMAND*, August 1969, 13-15.

~~SECRET~~~~HANDLE VIA COMINT CHANNELS ONLY~~

_____. "Traffic Analysis Processing System (TAPS)." Traffic Analysis Mechanization Forum, February 1971. Ref 194,544, 235-238.

Tiren, David J. "T/A Math Symposium Reviewed." *DRAGON SEEDS*, December 1972, 34-39.

"NSA Traffic Analysis Overseas." unpublished article in TA library collection.

"Practical Traffic Analysis from the Viewpoint of the Theory of Knowledge." manuscript in TA library collection.