DOCID: 3838665

Net Reconstruction—A Basic Step in Traffic Analysis

BY [(b)(3)-P.L. 86-36]

~~Confidential~~

An explanation, for those unfamiliar with the subject, of the nature of net reconstruction and the techniques normally used to accomplish it.

Communications intelligence is possible because any organization, particularly a military organization in battle, must carry on rapid communications in order to achieve planned, coordinated, and effective action. More often than not, radio is used; since radio is susceptible to interception, disguises become necessary to conceal the content of the message traffic and the structure of the organization.

Traffic analysis utilizes all aspects of the intercepted traffic to penetrate these disguises, to attempt reconstruction of the communications network, and to ascertain the system of signal operations in use. Moreover, when the message portion of the traffic is so disguised as to deny the intelligence of its contents to the intercepting party for an interminable period of time, the role of traffic analysis becomes even more important. Indeed, traffic analysis becomes the only source of communications intelligence and, in some cases, may be the only source of any intelligence.

This is not to restrict the importance of traffic analysis solely to traffic with unreadable message cryptosystems. When cryptanalysis is "yielding," the traffic analyst uses the results of his study to guide intercept activities, to fill the intelligence gaps left by the cryptanalytic results, and to give assistance to the cryptanalyst. Such assistance includes searching for *isologs*—messages encrypted by different methods but having the same plain text so that a knowledge of one system furnishes a crib for the other—but consists principally in ascertaining who originated a message, "where" it originated, and "who" is receiving it "where." Without this information even a complete decrypt would be of limited value to an intelligence expert.

Reconstruction of the enemy communications system may be considered as a basic step in traffic analysis. Specifically, such reconstruction will yield valuable order-of-battle intelligence, exposing the disposition, intent, and capabilities of the hostile forces. Of course, to

Approved for Release by NSA on 06-16-2008, FOIA Case # 51505

DOCID: 3838665
(b) (3)-P.L. 86-36

achieve these results, *communications continuity*[1] must be maintained at all times, but for the purpose of this discussion, let us assume that this has been done. Net reconstruction can be illustrated best by a purely fanciful example; therefore, let us, the Romans, project the campaigns of Hannibal in Italy during the Second Punic Wars into the Twentieth Century and attempt to reconstruct and identify his communications.

A body of traffic, intercepted from a number of intercept sites south of Rome, has been tentatively isolated and identified as the communications serving the Carthaginian Military Expeditionary Force in Italy (CMEFI). This general identification has been made by associating the characteristics of the CMEFI with those of known Carthaginian military communications previously intercepted by means of general search missions. However, beyond this tentative identification, little is known, from the stand-point of COMINT, of the CMEFI order-of-battle, the logistic structure supporting it, or the intentions of Hannibal in the campaigns sure to follow.

It has been ascertained that this body of traffic is composed of the following:

Low-level radiotelephone (Voice): Exploitable cryptanalytically or sent in the clear, this voice traffic contains basic tactical intelligence concerning the activities of Carthaginian infantry units from platoon to regiment size.

Voice/Morse: Exploitable cryptanalytically or sent in the clear, this voice/Morse traffic consists of voice transmissions interspersed with messages sent by means of manually keyed Morse. This traffic contains tactical intelligence concerning Carthaginian artillery units from division down.

Morse: This traffic is, as yet, unexploitable cryptanalytically and no plain text has been intercepted. It is believed to contain intelligence strategic in nature.

No other type of signal communications, such as radioprinter, has been intercepted or is known to exist.

Let us concern ourselves with Morse communications. On the basis of what is known about Carthaginian low-level radiotelephone and voice/Morse, we may assume, by the process of elimination, that Morse communications serve the following:

[1] In traffic analysis, communications continuity may be defined as the equation of callsigns, frequencies, schedules, and other variable elements assigned to a given radio station, link, or net before and after a change. For a detailed explanation see [redacted] "Chatter Patterns," *NSA Technical Journal*, Vol II, No. 4, pp. 63 and 64.

P.L. 86-36

(1) Carthage to the Carthaginian Supreme HQ in Italy.

(2) Carthaginian Supreme HQ in Italy to armies and infantry divisions.

(3) Infantry divisions down to regiments.

(4) Communications among or between infantry divisions.

(5) Carthaginian Supreme HQ in Italy (or Army HQ) to artillery divisions.

(6) Communications among or between infantry and artillery divisions.

(7) Communications serving the logistic command, possibly from Carthaginian Supreme HQ in Italy to regional logistic commands and/or from such regional commands to armies, infantry and/or artillery divisions within such regions.

To attack this Morse phase of the problem, the following external features of CMEFI traffic are studied:

a. Circular Messages: Messages destined for two or more different recipients, encrypted in a key which they hold in common. Usually, circular messages pass from superior to subordinate.

b. Isologs: Cryptograms in which the plain text is identical or nearly identical with that of a message encrypted in another system, key, code, etc.

c. Schedules: Times fixed beforehand during which links, groups, or nets operate. These may be constant or may be set by one-time prearrangement, the time of each period being announced at the end of the one before.

d. CQ Schedule: The method by which a control transmits a message to all outstations of a radio net simultaneously at a predetermined time. During a CQ schedule, outstations answer, request servicing, and receipt for the transmitted message or messages.

e. Station Serial Numbers (NR's): Reference numbers assigned by a station in serial order to all messages that it transmits.

f. Call-up Orders: A fixed sequence of sets of signals. These sets of signals are employed by a radio station to establish contact with other stations and to prepare for the transmission of traffic.

g. Relay Messages: Relays, for the purpose of this discussion, are messages requiring more than one transmission to effect proper delivery, but carrying the *original* encipherment in *all* transmissions and usually bearing a message NR from the NR series established between the originator and the addressee.

h. Procedure Signs: One or more letters or characters, or both, used to facilitate communication by conveying, in a condensed standard form, certain frequently used orders, instructions, requests,

DOCID: 3838665
(b)(3)-P.L. 86-36

and information related to communications.[2] Procedure signs are normally found in chatter between operators and are not considered as messages.

i. Callsigns: Groups of letters, numbers, or both, serving as a means of identification for a telecommunications station (or stations) when stations are establishing and maintaining contact with each other.

It is observed that a station (Station A), located by means of direction finding (D/F) in the Florence area, sends circular messages to four outstations twice daily on CQ schedules. This is ascertained by means of schedule analysis, call-ups, and message preambles.

A schedule analysis reveals that Station A maintains strict point-to-point contact with each of its four outstations in a definite call-up order as follows:

	OUTSTATIONS				
Station A to	B	C	D	E	CQ Schedule
	0600	0700	0800	0900	1000
	1200	1300	1400	1500	1600
	1800	1900	2000	2100	——

During these CQ schedules at 1000 and 1600, the message traffic bears the peculiar A1 (or first textual) group of AAAAA and uses a unique NR pattern (2000 series), whereas point-to-point communications utilize a "one-up" (1,2,3,4,··· etc) daily NR series to each outstation from Control. For example: the callsign for Station A on 31 October has been identified as PXLA. The callsigns for the outstations on 31 October have been identified as:

Outstation B: BPMC

Outstation C: DGRD

Outstation D: SNTC

Outstation E: MCPL

Thus, for these two CQ schedules on 31 October, the call-up is,

BPMC DGRD SNTC MCPL DE PXLA.

A circular message passed during one of these two schedules bears the typical preamble,

[2] See L. D. Callimahos, "Introduction to Traffic Analysis," *NSA Technical Journal*, Vol. III, No. 2, p. 6.

NR 2314 GP 150 A2B 31M10 14M30 $\overline{\text{BT}}$,

which means, "This is the 314th circular message (sent this month); it contains 150 groups, has a *Routine* (A2B) precedence, and was filed for transmission at my message center on 31 October at 1430.".

(*Note*: $\overline{\text{BT}}$ means "break" or end-of-preamble and the M is used as a separator).

After receipt and acknowledgment of this message, each of the four outstations reencrypts and transmits it to *its* outstations, maintaining the original NR in addition to its own NR series. Each of these four outstations, in communicating with its subordinate stations, becomes in effect a control of its own net (and as part of that net has a different callsign). As control it may choose to retransmit the circular message received either to each of its subordinate stations separately, or, by means of a separate schedule, to all of its outstations simultaneously. Let us assume the Standing Signal Instructions call for a point-to-point transmission. The reencrypted and retransmitted message would bear the typical preamble, NR 32/2314 GP 150 A2B 31M10 16M15 HAN $\overline{\text{BT}}$, which means, "This is the 32nd message sent to this particular station on this day. It was received as a circular message NR 2314 (the 314th sent this month); it contains 150 groups, has a *Routine* (A2B) precedence, was received at my message center on 31 October at 1615, having been relayed from Carthaginian Supreme HQ in Italy (as expressed by the trigraph HAN)."

Previous analysis has revealed that Stations B and C have three outstations each, Station D has six, and Station E has two. Thus, the close association and relative subordination among these five nets, which previously seemed unrelated, may now be assumed. As illustrated in the example above, this assumption is based on:

(1) The usage of circular messages, such as NR 2314 from Station A retransmitted eventually to Outstation 1;

(2) the incidence of isologs (i. e., NR 2314 is the same message text as NR 32/2314 although the latter has been re-encrypted, and retransmitted to Outstation 1);

(3) analysis of the schedules used by Station A (the schedules were intermeshed and no conflicts were evident);

(4) the time differences (1430 from Station A and 1615 from the recipient to his outstation);

(5) the use of NR's: a four-digit NR for circular messages by Station A, a one-up NR for point-to-point transmissions, and a split NR, 32/2314, for the retransmitted circular message;

(6) the peculiar A1 group AAAAA.

DOCID: 3838665

These phenomena have enabled us to reconstruct a part of the CMEFI communications which may be presented visually as follows:

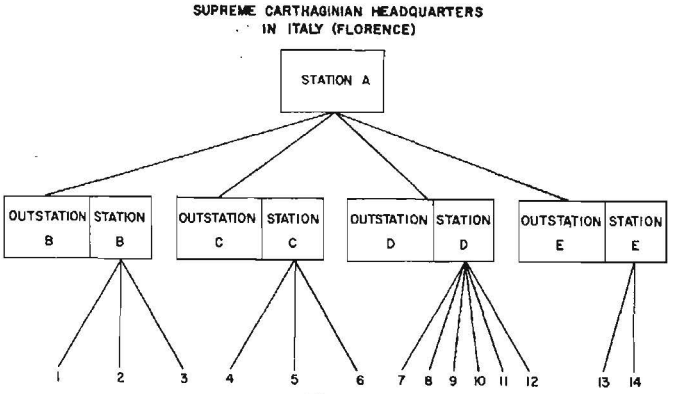

Fig. 1.

If Station A is identified as Carthaginian Supreme HQ in Italy, it may be assumed that Stations B, C, D, and E are the next lower echelon, or armies, and the 14 outstations of these four stations might, in fact, equate to divisions. At this point, a check of non-COMINT collateral sources such as prisoner-of-war interrogations, or agent and defector reports reveals that there are indeed four armies in Italy—the Numidian, Lusitanian, Mauretanian, and Cisalpine—with a total of 14 infantry divisions.

Attempts are now made to isolate and identify communications between the Carthaginian Supreme HQ in Italy and (1) artillery divisions, (2) possible logistic commands, and (3) Carthage itself. Again, a check of collateral sources reveals the possible existence of five Carthaginian artillery divisions and four regional logistic commands. Our next step is to look for controls colocated with Station A.

For this purpose, Special Identification Techniques (SIT) are used. These consist of Direction Finding (D/F), Radio Fingerprinting (RFP), and Morse Operator Analysis (MOA). D/F, or *radiogoniometry*, is the process of locating a radio transmitter by employing special receiving equipment, including directional antennae, which can determine the direction from which a signal emanates. RFP is the technique and process employed to codify and identify the unique characteristics of an individual radio transmitter by the study of the oscillograms of its signals. MOA consists of cataloging and identifying manual Morse operators by their individual sending characteristics.

RFP analysis does not reveal any similarity between transmitters used by a station X (with five outstations), a station Y (with four outstations) and Station A. D/F "shots" taken on these transmitters place them in the same general area, but the terrain does not allow of exact pinpointing. On the other hand, MOA continued over a period of time, does reveal that one radio operator has worked and is still working on all three.

Field intercept operators assigned to copy Station A communications note a Station A operator's "fist" peculiarities; other field intercept operators assigned to copy Station X and Y communications have observed the same "fist" peculiarities over a period of time and, indeed, intensive field analysis suggested the possible colocation of all three stations.

Analysis of Station X and Station Y traffic reveals circular message NR's in the 3000 and 4000 series respectively. A study of the circular message NR's of *all* CMEFI communications shows that Station A, Carthaginian Supreme HQ in Italy, and these suspected colocated stations are the *only* stations utilizing four-digit circular message NR series.

The combination of these facts makes it possible to assume that Station X with five outstations may be Carthaginian Supreme HQ in Italy communicating with artillery divisions, and Station Y with four outstations may be the same HQ communicating with the four regional logistic commands.

In addition, the link between Carthage (the city) and Italy is found, and is seen to be using a unique cryptosystem and a higher frequency range. Moreover, it employs automatic Morse devices for transmission, whereas all such communications in Italy are limited to hand Morse.

Thus, we have reconstructed a larger part of the CMEFI communications and identified their components. A diagram of them is shown in Fig. 2.

By means of traffic analysis and D/F, we have further ascertained that Outstation B, for example, is colocated with Station B. However, RFP has determined that Outstation B and Station B are using different transmitters, and MOA, confirmed by intercept operator notes, reveals the employment of different radio operators for each. Since the call-signs used by each station are separate and distinct, we may assume a certain compartmentation of communications functions at the army level (and presumably at subordinate levels as well).

So far, lateral communications between armies, for example, have not been found. We might assume that such communications are handled by means of relays through Control Station A. But analysis

(b)(3)-P.L. 86-36

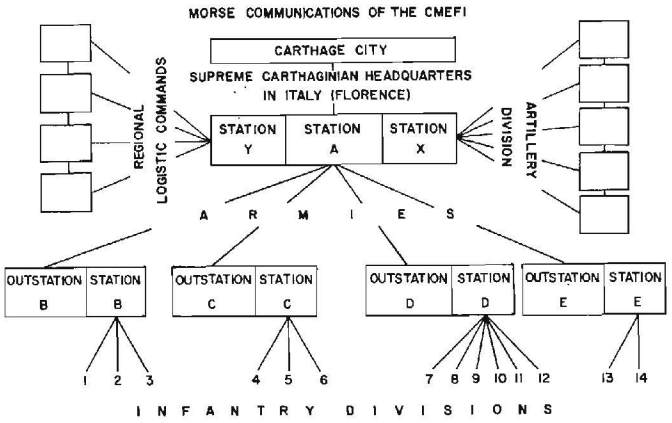

Fig. 2.

does not confirm this assumption. Therefore we reexamine previously unidentified CMEFI traffic for such lateral communications, expecting to find still a third colocated station (lateral Station B, for example), in communication with other armies and perhaps with other units in support of armies, so that the structure would look somewhat as follows:

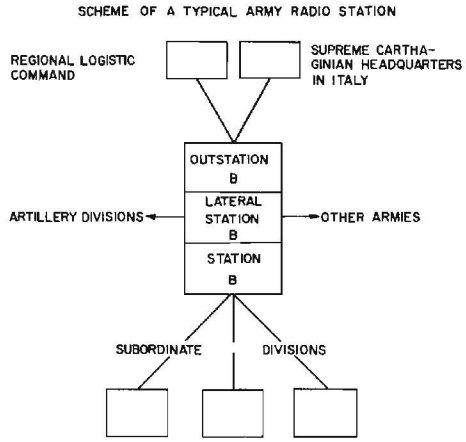

Fig. 3.

Some nets seem to fit the pattern but cannot be confirmed definitely at this time. We redirect our efforts accordingly.

It now becomes necessary to attempt the numerical identification of the infantry and artillery divisions. For this task many traffic-analytic techniques are used, but two in particular prove fruitful. These are *analysis of the call-up order* and *analysis of relays*.

Call-up Order: The call-up order for Station A has been discussed previously. Analysis confirms the fact that Stations B, C, D, and E likewise maintain strict point-to-point contact with a definite call-up order. For example, schedule analysis of the Station D radio net reveals the following:

Station D to outstations	9	7	12	11	8	10
	0445	0515	0545	0615	0645	0715
	0845	0915	0945	1015	1045	1115
	1245	1315	1345	1415	1445	1515
	1645	1715	1745	1815	1845	1915

The call-up order is thus established for Station D as being: 9—7—12—11—8—10.

Data obtained by means of prisoner-of-war interrogation, and agent and defector reports have indicated the existence of six divisions within the Lusitanian Army, numerically identified as the 19th, 22nd, 24th, 31st, 33rd, and 40th Divisions. There is a suspicion, therefore, that Station D may be the Lusitanian Army and a hope prevails that the call-up order may reflect the numerical designations within that Army. However, this information must be confirmed or rejected definitely by a more thorough analysis before it can be used. We turn then to a study of the relay system used by the Carthaginian armies.

Relays: The prime factor necessitating the use of relay procedures within the CMEFI seems to be one of time—either the urgency of the message, or a difficulty in transmission. Relays, apparently, are necessary to expedite important, high-precedence communications. As a consequence, such relayed messages merely reflect normal communications between previously established correspondents, and are sent by an alternate route merely as a matter of efficiency. Obviously, an originator may establish a relay route only to the extent that he is informed of the communications activities and regulations of intermediate stations with which he is in contact.

Consistent behavior of CMEFI communications further reveals that such relays are almost always transmitted in such a way that two "legs" (i. e., transmission paths) are all that is required for the delivery of the message. No instances of relays using more than two trans-

DOCID: 3838665 (b)(3)-P.L. 86-36

missions have been observed, although such a situation might exist in theory. Below is a diagram of a typical relay:

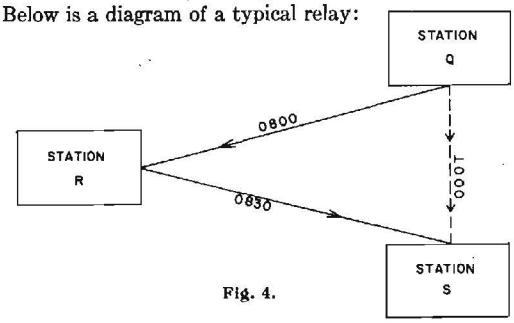

Fig. 4.

It is 0730 and Q has a message with an *Urgent* precedence for S. Scanning his schedule listing, he observes that he has an 0800 schedule with R and a 1000 schedule with S. However, R has an 0830 schedule with S. The sender, Q, transmits his message to R at 0800 with instructions to relay it to S, who receives the message soon after 0830 instead of 1000.

In order to avoid any misunderstanding as to the originator and addressee, the sender further adds a series of indicators in the preamble of the relayed messages as follows:

(Station R) (Station Q)

KARK DE NTPC

NR 14 GP 75 A1A 4M11 07M30 ZQQJ/S9B $\overline{\text{BT}}$,

which means "This is the 14th message sent this day from Q to S, containing 75 groups, of *Urgent* precedence. It was filed for transmission at Q's message center on 4 November at 0730. The originator of this message is ZQQJ and the addressee is S9B". (Note: The originator has used the NR series he would normally use to transmit to the ultimate addressee.) Station R will relay this message as follows:

(Station S) (Station R)

LZPU DE TRIS

NR 2/14 GP 75 A1A 4M11 07M30 ZQQJ/S9B $\overline{\text{BT}}$,

which means "This is the second message sent from R to S and the 14th sent from Q to S on this date, etc., as above." (Note: All Morse stations in the CMEFI network use separate callsigns for transmitting and receiving.)

The encoded station indicators (such as the originator ZQQJ and the addressee S9B above), though changing periodically every ten days, may be broken out by accumulating and studying them in relation to the suspected echelon (army, division, etc.) of the originating station and of the ultimate recipient, and by applying the suspected call-up order to ascertain the numerical designation (see page 33, *Call-up Order*). For example: Station Q is believed to be an outstation of Station D (Lusitanian Army) and may be assumed to be an infantry division. It has been the second outstation called by Station D in the call-up order (again, see page 33, *Call-up Order*). If, as suspected, the call-up order reflects the numerical designations within the Lusitanian Army, Station Q may be the 22nd Infantry Division; "Q" in the station indicator ZQQJ may equate to "2"; and "Z--J" may equate to "infantry division". Thus, for this specific period, the following may be reconstructed:

P = 1	U = 6		7P-	=	armies
Q = 2	V = 7		Z--J	=	infantry division
R = 3	W = 8		--9B	=	artillery division
S = 4	X = 9		G-	=	logistic command
T = 5	Y = Ø	and	HAN	=	Carthaginian Supreme HQ in Italy

Thus Station Q, the originator, may be identified as the 22nd Infantry Division, and Station S as the 4th Artillery Division.

The basic net reconstruction of the CMEFI (as of 4 November) has been accomplished, lateral communications confirmed, numerical designations and order-of-battle identifications of stations established, and the relationship of logistic commands and artillery divisions to armies and infantry divisions uncovered. These are shown in Figs. 5, 6, and 7.

It now becomes necessary to establish norms of activity and communications patterns within the CMEFI network in order to understand the normal working activity and organization and be able to recognize departures from it. This is accomplished by intensively studying various specific characteristics such as:

(1) *NR.* The volume of activity of each station is observed by the number of messages sent and received.

(2) *Radio Frequencies.* It is quite likely that a definite radio frequency usage allocation to major military elements has been made by the CMEFI Signal Officer in order to allow for maximum frequency utilization, without interference, among radio stations. Within these

DOCID: 3838665 (b)(3)-P.L. 86-36

SCHEME OF THE CARTHAGINIAN COMMUNICATIONS IN ITALY

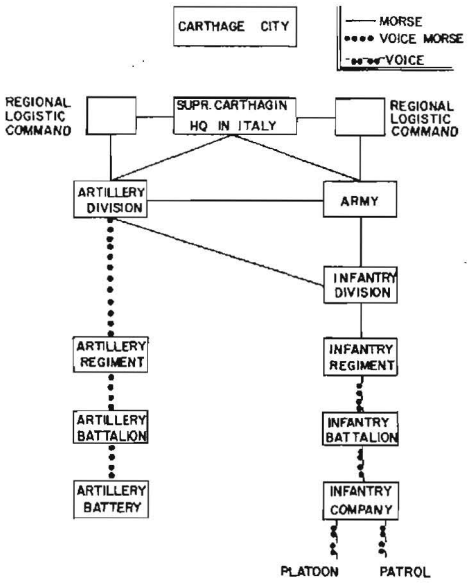

Fig. 5.

MAINLINE MORSE CMEFI COMMUNICATIONS

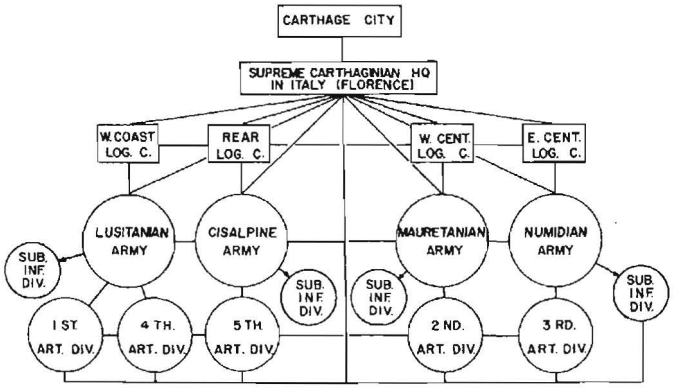

Fig. 6.

MORSE COMMUNICATIONS OF A TYPICAL ARMY

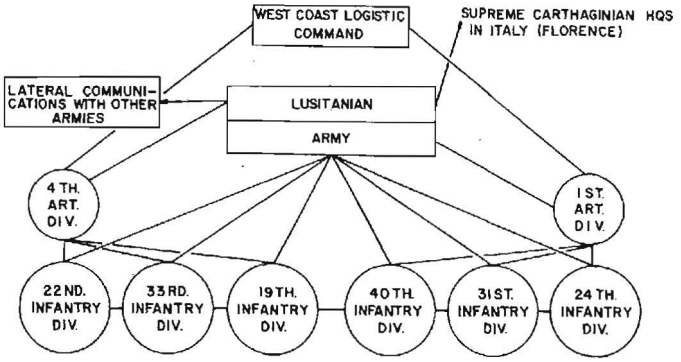

Fig. 7.

assigned bands, such major military elements may further allocate specific frequencies to subordinate units.

(3) *Callsigns.* The use of a large number of callsigns in a communications network, such as exists when callsigns change daily, necessitates a definite system to prevent confusion and misunderstanding. The reconstruction of such a system (including the methods of allocating callsigns) becomes a major task and, if successful, provides the traffic analyst with a powerful tool for maintaining continuity and identifying the users.

(4) *Procedure Signs.* The analysis of procedure signs yields much valuable technical information. Even in the most secure and disciplined radio network, procedure signs are sometimes found which reveal schedules in use or to be used, existence of messages not intercepted, identities of units or radio operators, and other information relating to continuous operation between stations. More often than not, such procedure signs are disguised and require solution. Since messages are not transmitted during every schedule, analysis of procedure signs becomes, at times, a major source of material for traffic analysis.

As the months slip by, the relationship of logistic commands to armies, to infantry and artillery divisions, and among the various divisions themselves, becomes clear. For example, the possible movement of a 23rd Infantry Division from Lusitania to reinforce Hannibal might be detected when it enters Italy for the first time. It will communicate first with Carthaginian Supreme HQ in Italy, then with the Rear Logistic Command, and will finally appear in the Lusitanian Army net

DOCID: 3838665

(Station D net), being called in the call-up order after the 22nd Division and before the 24th Division.

As the 23rd Infantry Division moves toward the front, its lateral communications with adjoining infantry divisions will indicate its relative position in the line (to be confirmed by D/F). Communications with the Rear Logistic Command will cease and contact with the West Coast Logistic Command will be observed.

As Hannibal prepares for the spring offensive, the volume of traffic among stations of the Lusitanian Army begins to rise and indicates that the front occupied by that Army may be the point of thrust. There is an increase of communications among the Lusitanian Army HQ, its infantry divisions, the supporting artillery divisions, and the West Coast Logistic Command; this reveals the number of units to be engaged in the thrust and the number of units comprising their support. Artillery preparations and increased local contacts by small units and patrols begin to be reflected in low-level radiotelephone and voice/Morse communications.

Thus, the Roman COMINT effort has fulfilled its mission, that of furnishing strategic (and tactical) intelligence; it has indicated the possible point of thrust for the coming offensive, the units to be engaged, and the relative functions of each. This fulfillment of mission has been accomplished, in large part, by applying the principles of traffic analysis to the intercepted traffic and reconstructing the Carthaginian communications system.

The die is cast, as Hannibal, unknowingly, has lost the element of surprise. It now remains for the Roman soldier to take over the defense of the Republic.

Non - Responsive