

“A Methodology for Simulating Net-Centric Technologies: An Operations Research Approach”

** Student Paper **

Joshua Lospinoso

Abstract

Captured enemy documents, also known as ‘CED’, are vitally important to today’s Army. With the capability to fill vital intelligence requirements, CED help Army units accomplish their missions and corroborate enemy prisoner of war interrogations. The language instant screening tool (LIST) technology being produced at the United States Military Academy provides a net-centric solution to expedite Army doctrine as outlined in FM 34-52. When modeling CED reporting procedures from FM 34-52, this study finds that LIST technologies could facilitate a net-centric enabled intelligence and information structure that is able to reduce dissemination times to a level that were previously unobtainable. The intent of this paper is to provide a methodology whereby further operations research may be conducted to support a net-centric tool suite that will ultimately benefit today’s Army.

Keywords: Captured enemy documents, latent semantic text analysis, network centric warfare, network science, language instant screening tool, discrete event simulation

Introduction

The advent of the Micro PC into mainstream computing through such companies as Sony prompted cadets¹ at the United States Military Academy to create the LIST project. LIST, which stands for Language Instant System Translation, is an attempt at facilitating the collection, reporting, interpolation, and dissemination of information contained in a captured enemy document, or CED. Since the Micro PC’s form factor is conducive to transport by personnel within a company, LIST offers a network centric tool which promises to revolutionize CED collection.

Through the use of a camera, GPS data, and a series of interoperated programs, LIST allows a soldier to take a picture of a machine printed document in any language (currently being developed in Arabic), index it for keywords through advanced latent semantic text analysis, append metadata (GPS, date-time, keywords, unit), and relay the information to higher in near real time. Latent semantic text

analysis will allow automated screening of CED, effectively pushing exploitation capabilities down to lower echelons. While this capability has the capacity to revolutionize the Army’s information structures, network centric principles indicate that organizational adaptation must accompany technological advancement.

Purpose of Research

The principles of operations research entail comparison of performance measures and values against costs; but what performance measures matter? When deciding on purchasing the LIST technology, an Army unit must evaluate more than just the ease in which LIST makes reporting CED. For without a rigorous and thorough analysis of how far the implications of LIST technology reach, decision analysis falls short.

To mitigate this potential bias, research must be conducted to evaluate not only the specifications of LIST technology (report time, processing time, battery life, etc.), but also how well LIST fits into the organizational norms, standard operating procedures, and structural considerations. For example, LIST technology promises to increase information velocity and decrease the effect of CED’s perishable nature; however, if CED begins to pile up without thorough analysis and exploitation, what has really changed?

¹ Under the direction of MAJ Ian McCulloh, Assistant Professor, Math Department

Furthermore, rigorous analysis of organizational performance permits policy recommendations beyond simple acquisition decisions as exemplified in current command and control methodologies.² As a result, it becomes increasingly evident that fruitful research will illuminate how LIST technology interacts with these methodologies.

The scope of this research project is to provide methodology and guidelines for further study. Ten active Army divisions, with varied CED processing capabilities, provide a challenge in mathematical modeling. Operations research requires close collaboration with decision makers; with so many configurations and stakeholders involved, this research project provides no directly applicable solutions. The methodology, however, promises a strong platform for further adapted studies.

Background

In support of the Office of the Secretary of Defense, Office of Force Transformation, a Network Centric Operations case study³ used inference chains from the NCO Conceptual Framework to analyze an order of magnitude increase in a team effort to demonstrate force effectiveness. In another team effort⁴, the Department of Computer Science at Texas A&M University published the results of battlefield simulations under NCO concepts in Simulating Teamwork and Information-Flow in Tactical Operations Centers using Multi-Agent Systems. By extending smaller scale distributed battlefield simulations like ModSAF and JANUS, the team created the TaskableAgents architecture to model human-decision making at the battalion and above level.

Clearly, the Army is interested in utilizing simulations to make good decisions on incredibly complex implementations. Unfortunately, the intelligence community has not benefited from a rigorous analysis of the captured enemy documents cycle. Army doctrine provides an invaluable platform for

conducting an analysis of LIST technology. The US Army Field Manual on Interrogation (FM 34-52), Chapter 4, will provide the guidelines of this study. Standard agreement (STANAG) 2084 defines a captured enemy document, or CED as “any piece of recorded information, regardless of form, obtained from the enemy, which subsequently comes into the hands of a friendly force.” CED are typically captured from enemy locations or dead, and they are rarely produced from willing sources. When enemy prisoners of war are taken, CED collected from the area are invaluable in corroborating human intelligence in the intelligence cycle. When an EPW is taken, CED physically travel with him to the appropriate echelons.

When no enemy prisoner of war (EPW) is taken and CED is collected, it is forwarded to the nearest Document Exploitation (DOCEX) center. These DOCEX centers are geographically dispersed and few in number— “[DOCEX centers] are normally organic to major [NATO] commands.” CED of this nature typically generate 525 to 5,300 sources per division per week. This astounding range and number makes thorough and timely evaluation of CED impossible.

CED are categorized into 4 compartments depending on their contents. Unless previously arranged, assets required for classification of CED are reserved at the corps level. Category A documents contain spot reportable information on enemy composition, disposition, and aid in accomplishing friendly missions. Category B documents contain secret information on cryptology and communication information. Category C documents help with general intelligence, and Category D documents are of no value.

Present Day CED Handling

Captured CED undergo accountability at the time of capture. The frontline unit responsible for acquiring the CED append a captured document tag (CDT) to the document which includes date-time groups, location, capturing unit, identity, and circumstances. Since CED are time sensitive, the value must be exploited as quickly as possible. The CED is then physically transported as quickly as

² As indicated in *Power to the Edge*, 2002

³ Gonzales et al, 2007

⁴ Zhang et al, 2001

possible to higher echelons. The following timeline is then adhered to:

1. Battalion S2 shop ensures that all untagged CED are reconciled, and performs DOCEX if possible, but places priority on expediting to brigade.
2. Brigade assigns escorts to each battalion's CED and physically transports the documents to the Division level. An IPW team can be assembled to glean whatever DOCEX information available. This team focuses on information pertinent to EPW interrogation.
3. MP escorts all CED from division to corps. A Division IPW section performs functions similar to the Brigade IPW.
4. The Corps CED team is the first to screen incoming CED for intelligence value. After batching CED into categories, the Corps CED team transports the CED to the combined-corps interrogation facility.
5. The CCIF provides a CIF DOCEX element. This element attempts to produce meaningful SALUTE reports and fill corps information requirements. Based upon the category of the CED, the CED is routed to the appropriate theatre asset.
6. Theatre Document Examiners perform final screening of CED before they are stored or disposed of.

A thorough description of all steps in the CED exploitation doctrine is available from FM 34-52. Accountability and logging of CED is a redundant and tedious process. Doctrine provides a flexible architecture for exploitation. If a battalion or brigade has the capacity to translate a screened CED, it does so and attaches a translation report. Three types of translations exist. A full translation is time-intensive and requires manpower sufficient to complete the task. It is unlikely that this kind of analysis can be performed below the corps level. Extract translations require a full translation of only a portion of a document. Summary translations still require a full read through of a document, but only require the main points to be conveyed as output. To reduce redundancy, exploitations

(translations) of CED are physically attached to the document.

Evacuation procedures for CED are conducted through the intelligence chain of command. Competence is of high importance here—a subordinate unit such as a Brigade must make an intelligent assessment of the CED based on limited exploitation capabilities and either exploit the CED at their level or decide to evacuate it to division level. Categorization is important for document triage.

Finally, when a CED is associated with an EPW, the process must be modified significantly. An EPW's cognitive capacity to retain valuable information diminish significantly with time—quickly forwarding him to the Corps level for interrogation takes priority over screening the CED associated with him.

Methodology for Analysis

Previous empirical research involving Network Centric Operational (NCO) concepts can be extended to CED processing to create an interoperable simulation package. However, the modeling of CED flow within an organization is required. With only doctrine to work with, extrapolating models for information structures within organizations as large as a theatre is impossible without empirical research. This research project, therefore, provides a methodology for interpolating this empirical research into a powerful decision making and simulation tool.

Using industry standard simulation software, mathematical statistics, and statistical regression, this research project will effectively allow decision makers to quantify aspects of information dissemination time, organizational work load, and information reliability. The first step is to define the organizational information structures. After establishing standard operating procedures (SOP's) for CED processing, the methodology requires statistical inputs. The rigorous nature of industry standard simulation software and mathematical statistics allows for virtually any random variable input. The second step is to model these random variables; such considerations include spatial separations between units, translation time, frequency of CED, categorical probabilities, time sensitivity,

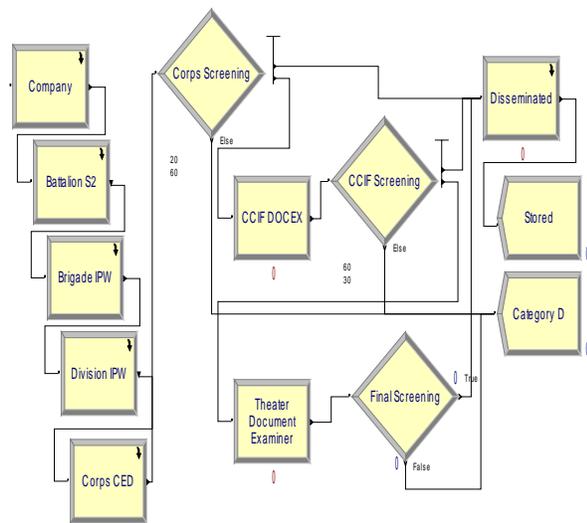
translation capabilities, translation requirements, and others. The third step involves data collection and statistical regression, which provides the capacity of natural experimentation required to produce an effective model.

With an effective model, the final step is to analytically solve for functions of these random variables within the defined organizational structure. If mathematical statistics fails to provide an analytical solution, simulation techniques can be used effectively to produce metrics which a decision maker can use to decide among alternatives.

Empirical Framework and Results

To illustrate the power of this architecture, this research project models three separate infostructure organizations: present day, present day with LIST functionality, and a proposed infostructure guided with a selective control C2 philosophy using LIST functionality.

Rockwell’s Arena Simulation package provides a powerful architecture for modeling complex organizational processes. To model the flow of CED throughout a Theater infostructure, feasible mathematical values were postulated as random variable parameters. The following flowchart shows the complexity of the Theater CED info-structure:



After running years worth of simulations in a few hours, Arena was able to determine that, based on feasible random variable parameters,

the current system performs moderately well; the average time to dissemination was 23.72 hours.

Having established a reasonable flow chart, we can estimate the simulations outcome using mathematical statistics. Since the Gamma function allows widely flexible curvature and a relatively simple composition, it is used extensively throughout the simulation. Note, however, that further empirical work could uncover the true underlying distributions of the random variables used in this simulation. The following table shows the random variables utilized in the simulation and their associated parameters:

Random Variable	α	β	Avg. Time	Var.
CED Frequency (at company level)	1	2.1	2.1	χ
Company Process Time	1	2.1	2.1	δ
Battalion Process Time	2	2.1	4.2	ϵ
Brigade Process Time	1	2.1	2.1	ϕ
Division Process Time	1	2.1	2.1	γ
Corps Process Time	2.5	2.1	5.25	η
CCIF Process Time	2	2.1	4.2	ι
Theatre Process Time	1	2.1	2.1	φ

If we assume independence among these random variables, the moment generating function technique allows us to compose summations of random variables through the product of moment generating functions. The moment generating function of a gamma-distributed random variable is the following:

$$m_g[t] = (\beta \cdot t)^{-\alpha}$$

Because we have chosen random variables with equal β parameters, composing the function for total time (κ) is trivial:

$$\kappa = \delta + \epsilon + \phi + \gamma + \eta + \iota + \varphi$$

Using the MGF technique, the following function is the MGF of κ :

$$m_{g,\kappa}[t] = (1 - \beta \cdot t)^{-\delta + \epsilon + \phi + \gamma + \eta + \iota + \varphi}$$

Recognizing that this is simply another gamma function, κ is distributed Gamma with $\alpha = \delta + \epsilon + \phi + \gamma + \eta + \iota + \varphi$ and the same β of the

other random variables. Using the parameters supplied, the PDF for total time is the following:

$$f[x] = \frac{1}{2.1^{11.5} \Gamma[11.5]} x^{10.5} e^{\frac{-x}{2.1}}$$

The expected value of κ is 24.15, validating our simulation. Since the mathematical statistics used here does not include the ability of corps and some division level assets to exploit documents, the expected value is upward biased (too large). Further empirical studies could prove to make these parameters fit present day organizations.

LIST Infusion

Since LIST technology promises to decrease the time it takes for companies and battalions to relay information to higher, it is clear that the expected values of company processing time and battalion processing time will decrease drastically. Since semantic text analysis promises to automate the process of summary translation, document triage will decrease the amount of time it takes for documents wait in queue—more Category D documents can be separated from the important CED. This indicates that the rest of the processing times will fall as well.

The following notional values are used in the simulation of List Infusion:

Random Variable	α	β	Avg. Time	Var.
CED Frequency (at company level)	1	2.1	2.1	χ
Company Process Time	.2	2.1	0.42	δ
Battalion Process Time	.5	2.1	1.05	ϵ
Brigade Process Time	.75	2.1	1.575	ϕ
Division Process Time	.75	2.1	1.575	γ
Corps Process Time	2	2.1	4.2	η
CCIF Process Time	1.5	2.1	3.15	ι
Theatre Process Time	.75	2.1	1.575	κ

Simulating the exact same flowchart over the same period of time verifies this fairly obvious hypothesis—LIST could decrease expected exploitation time of CED to 11.11

hours and decrease time sensitive CED loss to less than 1%. The analytical solution provides an expected time of 13.545 hours, verifying the simulation (again with an upward bias).

Consequently, LIST could half the CED transporting time. There are, however, limitations to this simulation. The first and most obvious limitation is that the parameters used in simulation are purely hypothetical. Empirical research and linear regression would be necessary to estimate these parameters accurately through a process called bootstrapping. Next, CED with an EPW attachment does not benefit from LIST nearly as much as CED collected from an enemy body or abandoned location. Since the CED is used only to corroborate interrogation of an EPW, it is not as useful to electronically send copies of the CED around the info-structure. It could, however, be useful for aiding interrogators at the Corps level and expedite the interrogation process.

Net-Centric Infusion

The final stage of analysis includes modifying the flowchart to reflect selective control C2 architecture. The concept here is the utilization of LIST as a small part of a greater net-centric force enabled with information age technology. In researching the Net-Centric phenomena, we see in Moxley's⁵ formula that information in a network is a function of methodology and domains:

$$M[d, b] = N_k$$

The most important concept behind this fundamental relation is that, as we have seen, network analysis is focused on empirical analysis of the physical aspects. Network science relies on a principled approach to the primordial components of networks as they manifest themselves wherever they exist, whether they are physical, social, or biological.

Moxley's Network Science Taxonomy compares our current knowledge base on these

⁵ Dr. Frederick I. Moxley, DISA Fellow/Visiting Professor and Director of Research, Network Science Center, United States Military Academy, West Point, NY.

network domains. To date, we have a higher level of physical networks such as wireless internet and telecommunications. Unfortunately, relatively little is known about the social and information domains. These latter domains entail precisely what must be fundamentally modified in the CED reporting process and the Army's infostructure in general.

The value creation process is of utmost concern in the operations research field. Through breaking down the value creation process into its individual parts, analytical methods can garner information about the contributions of each node to the overall effectiveness of an organization. With the addition of nodes into an effective organization, Metcalfe's law dictates that the value of the network increases exponentially. The following proportion illustrates this relation, where V is value in the network and n is the number of nodes connected to it:

$$V \propto n^2$$

Because CEDs represent knowledge within the network, this relation manifests direct applicability to future organizational changes to Army infostructure.

Virtual organizations have exploited this phenomenon with great results. Because the importance of physical location is diminished, virtual collaboration and virtual integration allows economies of scale to be achieved.

Army organizations can achieve full spectrum dominance through a superior information position. The challenge arises in command and control; how will an Army organization apply the relatively abstract concepts of value creation, virtual collaboration, and information superiority while conserving the redundancy, security, and interoperability that is integral in a robust Army organization?

The answer to this question is still the subject of hot debate within decision makers in the Army. Many competing philosophies exist on the subject. Dr. David S. Alberts and Richard E. Hayes have postulated on the subject extensively. Their publication, *Understanding Command and Control*, provides the basis for understanding this tradeoff. While traditional notions of Command and Control exist, they are correct in the doctrinal sense in that in order to

fully understand C2, it is necessary to overstep the boundaries of conventional thinking.

This paper provides a methodology for analyzing decisions involving Net-Centric adaptations; the analytical and simulations-based techniques covered here cannot hope to solve the complex issues involving the C2 tradeoff quagmire. Nonetheless, a fully Net-Centric CED system would be seamlessly integrated with a larger intelligence network within the Department of Defense. To eliminate separate processes which exist within the current stove-piped architecture, a solution may include the following notional architecture for CED reporting:

1. A front line company acquires CED. Using LIST technology to append metadata such as time, location, unit, and keywords (through semantic text analysis), the company intelligence NCO relays the CEDs electronically to the Battalion S2.
2. Within moments, the Battalion S2 is able to perform more processor-intensive analysis on the document in an attempt to provide DOCEX within an hour of CED acquisition. At the same time, the electronic CED is posted to a Theatre-wide server where combined-corps level assets are assigned to full or extract human translations.
3. These high echelon assets are optimally assigned to CED within hours based upon keywords gleaned by semantic text analysis.
4. Dedicated personnel from the battalion to the theatre continually monitor the Theatre-level server for potential Category A-C documents and exploit those documents, appending findings to the metadata.

Because the top-down hierarchy paradigm is completely thrown out here, competence at all levels is of utmost importance. Skilled intelligence personnel will be relied upon to make accurate and intelligent assessments CED available on the server. Through centralizing the information available to subordinate units, the theatre commander

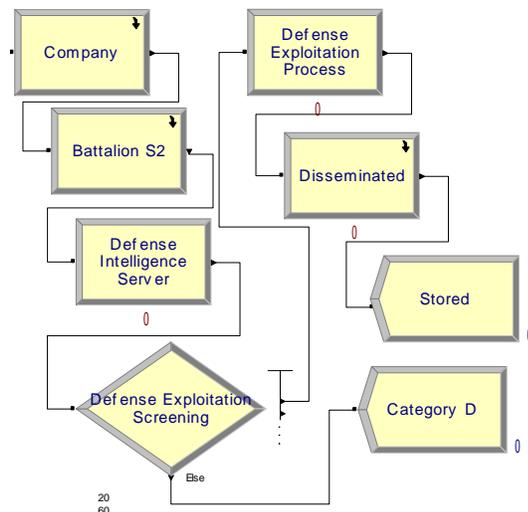
permits collaborative effort without redundancy in order to maximize efficiency. Based upon keywords, leaders at all levels have access to pertinent PIR and CED by pulling information off of the server.

Posting CED onto a server is limited only to bandwidth considerations and the processing time required by LIST. Since the previous simulation measures expected time to dissemination, this new architecture could reduce a 20 hour process into a half-hour process. Analytically, this process is akin to eliminating the brigade through theatre processing times.

Arena still provides a good tool for analyzing this architecture. Through removing the stove piping under current doctrine, a highly interconnected network of intelligence analysts can access a cornucopia of information virtually without being fettered by geographical limitations. To simulate this, the following notional values will be used:

Random Variable	α	β	Avg. Time	Var.
CED Frequency (at CO. level)	1	2.1	2.1	χ
Company Process Time	.2	2.1	0.42	δ
Battalion Process Time	.5	2.1	1.05	ϵ
Exploitation Assets Available	100 (Fixed)			
Exploitation Time	1	2.1	2.1	λ

Mapping this organization in arena is also a simplified process:



Economists strive to achieve perfectly efficient systems. When scarce resources are not allocated in the best way, the system is inefficient. This Net-Centric network represents a perfectly efficient system. Because no translator will ever be idle when an important CED needs to be exploited, the optimal configuration has been achieved. The results are impressive: given these notional inputs, CEDs are processed within 3.852 hours of being captured.

Variance and Risk

The mean time to dissemination provides a good measure of the expected efficiency of an information structure. However, a rigorous operations research analysis requires a look at variability. Variability manifests itself in two main aspects of this paper's analysis. Firstly, the notional random variable distributions have assumed a constant beta term (β), which is often referred to as the shaping parameter. It is highly unlikely that the shaping parameters are equal among these distributions. The beauty of the gamma distributed random variable is its flexibility; but how does having different shaping parameters affect analysis? For the simulations, nothing changes. The moment generating technique, however, is significantly more complicated. Instead of having a nice underlying distribution for the time to dissemination, the distribution is probably unknown. The moment generating function technique provides a solution:

$$m_g [t] = \left(\sum_{i=0}^n \beta_n \cdot t \right)^{-\sum_{i=0}^n \alpha_n}$$

By finding the first derivative with respect to t of the function above, the first moment of the function is found. Because the first moment of any function is the mean, we have found the expected value. If the random variables of interest are independent, this technique is applicable to any combination of random variable distributions with known moment generating functions. Simply multiplying these functions together will yield the moment generating function of the desired exploit time distribution.

Variation is important for the individual random variables comprising the time to dissemination, but knowing the variation of exploit time is equally important. Finding this variation is available through the simulations, or through finding the second moment of the underlying distribution (the second derivative with respect to t).

The variability of time to dissemination can tell a decision maker the risk involved with the organization under consideration. It may be desirable that it takes only 8 hours to exploit a CED, but if the standard deviation of this distribution is 10 hours, there will be some documents which could take more than 24 hours to exploit. Taking these factors into consideration will certainly be important to a decision maker, and measures of these factors are available through the methodology previously outlined.

Future Research

Because the net-centric system effectually pools the CEDs into a vast network of exploitation assets, the availability of those assets becomes more of a concern than it did in previous architectures. Because of this change, the moment generating function technique does not entail the analytical rigor required for this complex system.

Stochastic queue theory is a technique employed by businesses everywhere. A brief hypothetical situation relays the power of this technique: if the frequency of CEDs explodes into unforeseen levels, exploitation assets may be too busy to look at CEDs that have been classified via latent semantic text analysis triage as less likely to contain valid information requirements. When this occurs, time accrues for the documents in queue. When performing simulation based analysis of Net-Centric implementations, stochastic queue theory should be used to verify the model.

Conclusion

Further work is required to port the methodology contained in this paper into an applicable package for evaluating Net-Centric technologies. Empirical studies promise to calibrate the results of the simulation and analytical solution proposed, but the moment generating function technique shows us that, holding the shaping parameter β constant, reducing the mean time of any component of CED dissemination reduces total time by a factor of β . Since we can expect significant decreases in Company and Battalion relay times, even using the current doctrine outlined in FM 34-52, a unit can potentially half its total CED relay time.

The specifics of this methodology require a good deal of adaptation in order to prove useful. Nonetheless, the power of net-centric technologies has been illustrated, and a robust platform for evaluating those technologies has been presented. How organizations choose to implement net-centricity will be of paramount concern. Without organizational flexibility, the margin of power superiority that our armed forces enjoy now will certainly diminish in the future.

Author

Joshua Lospinoso is a second year cadet (sophomore) who is attending the United States Military Academy at West Point, NY. The views expressed herein are those of the author and do not purport to reflect the position of the United States Military Academy, the Department of the Army or the Department of Defense.

References

Department Of The Army. Intelligence Interrogation. 1992. FM 34-52. 25 Mar. 2007 <<http://www.fas.org/irp/doddir/army/fm34-52.pdf>>.

Alberts, D. Power to the Edge. Washington, DC: CISSP, 2002.

Gonzales, D., Johnson, M., McEver, J., Leedom, D., Kingston, G., and Tseng, M. "Network Centric Operations Case Study." Pending Distribution (2007). 25 Mar., 2007. <<http://www.act.nato.int/events/documents/nnc/stryker.pdf>>.

Moxley, F. *The Art of Network Science*. (Forthcoming) Copyright 2007.

Y. Zhang, L. He, K. Biggers, J. Yen and T. Ioerger, Simulating teamwork and information flow in tactical operations centers using multi-agent systems, *Proceedings of the tenth conference on computer generated forces and behavioral representation*, Norfolk, VA (2001), pp. 529–539.

Wackerly, D., Mendenhall, W., and Scheaffer, R. Mathematical Statistics with Applications. 6th ed. Pacific Grove, CA: Duxbury Thomson Learning, 2001.
own, UK.