Exploring the Analytical Processes of Intelligence Analysts

George Chin Jr., Olga A. Kuchar, and Katherine E. Wolf
Pacific Northwest National Laboratory
P.O. Box 999, Richland WA 99352 USA
{George.Chin, Olga.Kuchar, Katherine.Wolf}@pnl.gov

ABSTRACT
We present an observational case study in which we investigate and analyze the analytical processes of intelligence analysts. Participating analysts in the study carry out two scenarios where they organize and triage information, conduct intelligence analysis, report results, and collaborate with one another. Through a combination of scenario-based analysis, artifact analysis, role-playing, interviews, and participant observations, we explore the space and boundaries in which intelligence analysts work and operate. We also assess the implications of our findings on the use and application of key information technologies.

Author Keywords
Intelligence analysis, homeland security, national security, participatory design, work practices, work-oriented design, artifact analysis, participant observation, collaboration.

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION

Intelligence analysts (IAs) work in a very demanding environment. They receive massive amounts of complex information that comes from various sources and are tasked to make sense of the data and discern critical patterns and anomalies. They conduct analysis under demanding time constraints and strong political pressures, and often the lives of others depend on their ability to make accurate predictions and assessments. Yet, while the intelligence community has been steadily improving its ability to collect information, its ability to analyze information has not progressed as significantly [1].

Since the September 11, 2001 terrorist attacks in New York and Washington, DC, renewed interest has emerged in characterizing and analyzing the work of IAs in the context of developing advanced information technologies and tools to improve intelligence analysis [2, 3, 4, 5]. Many of these research and development efforts have been or are being funded and spearheaded through US government programs such as the Department of Homeland Security (DHS) National Visual Analytics Center (NVAC) and the Advanced Research Development Agency (ARDA) Novel Intelligence for Massive Data (NIMD).

As computational scientists at the Pacific Northwest National Laboratory (PNNL), we are also interested in the research, development, and deployment of information technologies to support intelligence analysis. To create and evolve such technologies, however, requires a deep understanding of the analytical processes that IAs carry out. How do they organize their information for analysis? What computational tools do they apply? How do they collaborate with others? What are their analysis products?

The objective of our study is to capture, examine, and understand the analytical processes and work practices of IAs. Restricted from accessing and using authentic intelligence sources and data, we developed two scenarios with fictional but representative material and data. We then observed IAs as they worked through the two scenarios and developed hypotheses and conclusions. We applied various observational and analysis methods to capture and examine the analysts' work. Furthermore, we evaluated the applicability and usefulness of specific information technologies that appeared relevant to this work.

RESEARCH METHODS AND SETTING

At PNNL, we organized an analyst workshop to elucidate, discuss, and share the general work and analysis activities of IAs. PNNL was establishing a homeland security program and committed research funding to approximately a dozen computational sciences research projects. Most project members had very little exposure to the work of IAs. The analyst workshop was intended to provide project teams their first glimpse of the work activities that IAs perform and allow them to begin exploring what kinds of computational tools would be useful in supporting intelligence analysis. For security reasons, IAs were not permitted to discuss existing analysis projects or show the physical confines in which they operate. Given this overall context, we were driven to contrive our own intelligence analysis scenarios and to rely on self-reporting methods to gather details and requirements.

Five IAs participated in the analyst workshop. Their normal job functions covered a wide spectrum of intelligence areas

CHI 2009, April 4–9, 2009, Boston, Massachusetts, USA.

including cyber-security, threat analysis, critical infrastructure protection, counterintelligence, and nuclear non-proliferation. Most of the analysts had prior intelligence community or armed services work experience. The IAs knew one another and some had worked together on past assignments. Most worked for the same department.

A sixth analyst developed the material that appeared in the two separate scenarios. This analyst's participation was vital in developing material that was reasonable and consistent with what one might find in authentic intelligence cases. All information contained in the scenarios was either fictional or publicly available.

The objective of the first scenario was to understand and explore how IAs conduct intelligence gathering and analysis. The objective of the second scenario was to examine how IAs collaborate in real-time as they jointly carry out intelligence analysis.

In conducting this study, we drew from approaches and methods that have been applied in work practice studies and in the work-oriented design of information systems [6, 7]. Specifically, we utilized a workshop format [8] to bring designers, developers, and users together, and developed and interacted with scenarios [9] to detail and evaluate human processes and practices. For scenario 1, we conducted artifact analysis [10, 11] to analyze and compare the work products of an intelligence analysis process. For scenario 2, we engaged IAs in role-playing or theater [12, 13] as participants collaborated, acted out roles, and engaged in specific behaviors as they carried out the scenario. We also conducted participant observations [14] as we observed the team of analysts working together through the hypothetical scenario. Analysis methods were applied using a "participatory design" approach [15, 16], where we consciously encouraged and engaged users in the analysis and design of their own work practices.

The study we present is qualitative and designed to capture, elucidate, and understand the work practices of IAs. It is intended to be descriptive. Like other qualitative studies, it may be used to inform the designs of computer tools and systems, but not to directly specify them.

Scenario 1

The first scenario centered on a domestic terrorism case involving a fictional extremist group in the Northwest. Each participant was asked to play the role of a FBI Intelligence Operations Specialist assigned to a Crisis Action Team. The intelligence information in the case was supposedly collected by a previous specialist. Documents included background information on the group, witness statements, press articles, intercepted electronic files, working notes from the previous specialist, a listing of online associates of group members, and floor plans from the White House. Also included was a set of spreadsheets containing the names of foreign terrorists and computer hackers, chemicals one should never mix, and cheese market information.

IAs were instructed to develop a report assessing the general threat of the extremist group and potential ways it might attack the US government or population. We did not ask IAs to report their finding in a prescribed format, so that we could observe how they would naturally describe, substantiate, and connect their hypotheses and claims.

All documents were provided to participants in electronic form. Participants were given two weeks to analyze the case, after which, they would present their findings to an audience during a workshop. Participants were asked to work alone. Furthermore, they were informed that no set or correct solution existed. Some information in the case was deliberately ambiguous, conflicting, and/or irrelevant. During the course of the two weeks, participants spent a total of four to eight hours working on the case.

We asked the IAs to apply typical analysis tools in the manner they normally would on real cases. Thus, employed analysis tools and methods were representative for our set of participants. Most of the IAs were moderate computer users. The analysis process they performed and tools and data sources they utilized were often not mandated at any level. From our experiences in working with analysts from other intelligence agencies, we found that the tools and methods the IAs employed in our study were common across the wider intelligence community.

At the workshop, each participant presented his or her findings and conclusions. Furthermore, each was asked to walk through the analysis process he or she conducted. After the individual presentations, participants organized into a panel and took questions from the audience.

Scenario 2

In the second scenario, participants worked together as an investigative team. The team consisted of a case chief overseeing the work of four detectives. The case chief was given a briefing document containing background information on a gang known as the Gregorian Brotherhood. Each of the detectives received a spreadsheet report that listed different sets of crimes. The crimes occurred in four named districts over different periods of time. The objective of the investigative team was to discern patterns of involvement of Gregorian Brotherhood members in the crimes occurring in the four districts. The details of the second scenario were provided to participants at the time the scenario was initiated.

As the team conducted analysis, new information would spontaneously arrive. The team was to integrate the new information into their working analysis. This scenario better reflects the nature of how information is received in the intelligence community. Rather than receiving bulks of information all at once at the beginning of an investigation, data tends to dribble in over time in disparate pieces.

The team was instructed to collaborate on the analysis of the case over the course of an hour. IAs had two

whiteboards, markers, pens, paper, writing pads, rulers, and sticky notes among other office supplies at their disposal. At the end of the time period, the case chief was to report their findings.

Like the first scenario, Scenario 2 did not have a set or correct solution. The case also had ambiguous, conflicting, and irrelevant information. IAs conducted Scenario 2 in an auditorium where their actions were observed.

ANALYSIS RESULTS

In this section, we present some of the important concepts and themes that emerged from our observational study as evidenced in analysis products and artifacts and interview responses produced by participants.

Intelligence Analysis Strategies

One area we wished to examine in the study was whether IAs prescribed to and followed specific intelligence analysis strategies. In other studies, various researchers [1, 17, 18, 19] have extensively described and examined many of the approaches and strategies that IAs apply in their practice. In our study, we wished to identify which strategies were most prominent, the details and nuances of how they are applied, and how they might affect, add to, or color analysts' views and philosophies.

One common intelligence analysis strategy mentioned by our IAs is the analysis of competing hypotheses. Under this strategy, an analyst lays out all possible hypotheses and then maps each piece of data to each hypothesis to assess whether the data supports, counter-indicates, or is irrelevant to the hypothesis. The hypothesis that is best supported by the data is considered the most credible.

To illustrate the competing hypotheses approach, one IA in our study presented three competing hypotheses of how the extremist group in Scenario 1 might attack US citizens or the US government. As shown in Figure 1, each hypothesis is annotated with evidence that support it. In addition, an IA would typically also list evidence that refutes each hypothesis. Identifying the most valid or likely hypothesis amounts to determining which hypothesis is collectively best supported and least refuted by the gathered evidence.

In contrast to the competing hypothesis approach, another IA argued that intelligence analysis should not begin with any preconceived notions or hypotheses. This analyst warned against the tendency of "taking one's favorite hypothesis and making the data fit it," and suggested letting "the data prove itself and suggest itself." Elaborating on this view, the analyst continued,

> We don't have any assumptions, meaning we haven't drawn any preliminary conclusions on what we're going to find. We do have assumptions about how things work, like hackers, scientists, and such, but we didn't draw any conclusions from the assumptions.

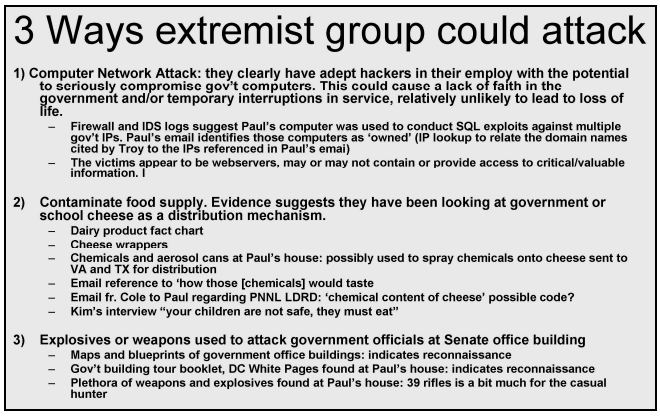

Figure 1: Competing hypotheses with evidence.

A third IA promoted a strategy of investigating the four specific areas of access, intent, motivation, and capability. As the analyst described,

> We know going in, that a threat entity of any sort, whether it being an individual or group, will have or be working towards either access, from intent, for a motivation, and with some type of technical capability… What technical capability does the adversary have, why are they doing it, what might they do with it, and where can they access to make it happen? Those are the things I start an analysis with. What does the information present me in those four categories?

Intelligence analysis is often seen as indefinite and self-perpetuating as described by the following analyst,

> Once you go through the information you have, then you know what you don't have, and issue requests out to the community that supports you and try to get additional resources to support the information that you do have.

Thus, IAs are always moving to new queries and new lines of investigation. For example, in Scenario 1, several IAs noticed that a suspect was using a company email address after he left the company. Different analysts wondered why the suspect still had an email account, how and why the suspect left the company, what kind of work did the suspect perform, with whom did the suspect correspond with via email and physical mail, and with whom did the suspect have personal relationships? As information and evidence is collected, many new questions and lines of inquiry arise. The goal is to eventually reduce the number of outstanding questions as answers converge to form a coherent story.

Intelligence analysis is always conducted within a specific period of time. To IAs, an analysis is never fully complete but rather is valid as far as the current evidence shows and time of analysis allows. As described by an analyst,

> You never have enough time to finish an analysis, because you're never finished. Every time I get a call from someone who is waiting for a report,
>
> *When are you going to be done with your analysis?*
>
> *Well, when do you want it by?*
>
> *By Thursday.*

Then I'll be done by Thursday.

And you'll get what I've got by Thursday. And then on Friday, you'll get the essential piece of info that you really need.

The IAs in our study expressed that they would often abandon a systematic approach to satisfy time constraints. Under time pressures, IAs may mentally conduct many aspects of an analysis without adhering to a particular investigative path or documenting intermediate findings.

Information Collection and Triage

The first step in the intelligence analysis process is information gathering. For IAs, physical files and folders continue to play prominent roles in the collection and organization of information. For Scenario 1, despite receiving all case information in electronic form, each of the IAs printed out hardcopies of the individual documents. One analyst sat on the office floor and laid out the hardcopies in a circle around her such that she could see the documents all at once. She labeled each of the documents to better track them and piled them according to their types (e.g., firewall logs, email, interviews). Within a pile, the IA made smaller piles perpendicular to one another to further organize documents into subtopics. She then manually drew a graph that showed the relationships among documents. Later, the IA redrew the document relationship graph in Microsoft PowerPoint as shown in Figure 2.

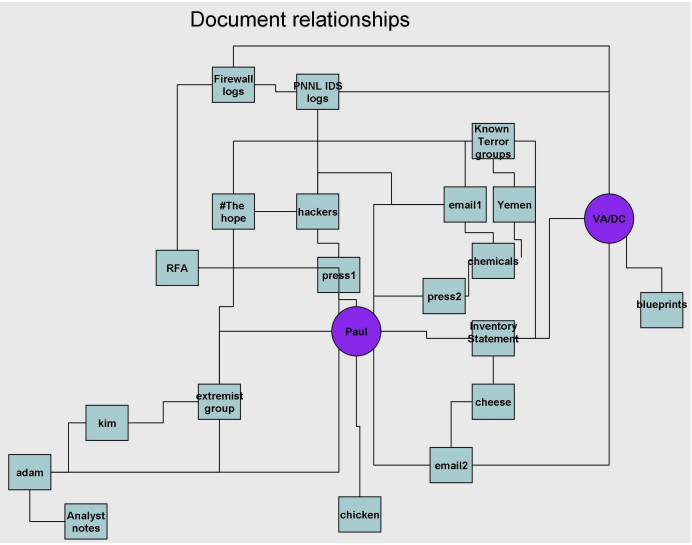

Figure 2: Relationships among documents.

Another IA made five hardcopies of every document and placed them into different folders representing different types of relationships. She duplicated this process on the computer as well, where she placed electronic copies of documents into electronic folders.

A third IA read through all the documents and then marked each one with a number identifying its relevance and importance. He then physically ordered the documents in a pile from most to least important and relevant. The analyst called this ordering process "triage." Once ordered, the analyst would extract key facts from one document at a time and insert them into a spreadsheet.

As described, the IAs participating in our study always began their Scenario 1 investigations with their evidence in physical form even though the materials they received were all initially provided in electronic form. In our follow-up interviews with analysts, we found that all of them at one point had conducted intelligence analysis solely using physical files and folders without computer support. Out of comfort and familiarity, the use of traditional media continued to be a part of their analysis processes.

Data or fact extraction was a common procedure for all IAs, but the analysts highlighted and extracted facts in different ways. All five IAs highlighted key facts and details in the physical document with highlighter pens. Each IA would then replicate and collect those key facts into common representations or analysis artifacts such as graphs and spreadsheets. Computer applications like spreadsheets and drawing packages provide analysts basic capabilities for managing, indexing, and linking information together, which would be much more cumbersome to accomplish using physical files and folders. Fact extraction often marks the point where information and evidence migrates from its physical form on paper to a digital form on computers.

In their investigations, the IAs would typically store and lock physical case folders and files in filing cabinets as raw evidence much like police detectives might stow away physical evidence from a crime scene. The analysts believed that saving the raw evidence was necessary should the analyst need to re-examine it in the future.

Identifying Patterns

Once the IAs in our study collected and filtered evidence from case materials, they turned to examining the data to identify relevant patterns and trends. In describing how IAs look for patterns, one analyst explained,

> We're trying to find some type of order in the information we were given. We found they were in different time frames. We're trying to see where there are common borders. We're looking for some type of pattern. Start looking at the times, dates, what activity occurred.

In conducting intelligence analysis, IAs search for key relationships or connections among facts and evidence. Relationships are established by identifying concepts that seem similar or orthogonal, or to naturally aggregate. IAs will look for concept details that are similar or the same in different places or contexts, or concepts that seem to co-exist in time, space, and/or other dimensions.

The analyst heavily relies on and exercises his or her personal knowledge when defining and establishing relationships. For example, in the context of identifying geospatial patterns, one analyst described how assumptions and background knowledge impacts his analysis,

> Criminals generally conduct the crimes in their general region or location where they live. (This is) an assumption. Is there a pattern of distance between all the events that took place and some type of central location that would dictate, maybe that's where we need to start looking? And then you just search. It's a query you would just run to see who in this area might fit that profile. It would be something in the back of my mind rattling as I go through the data.

In another example from Scenario 1, an IA with a background in stenography and encryption closely examined an intercepted email containing a long sequence of numbers. As shown in Figure 3, the analyst combed through the message and partitioned numbers into sets of threes. The analyst then scrutinized the number sets in hope of discerning some hidden message. In this particular case, no patterns were found.

000 477 470 343 570 043 157 735 351 157 331 130 557 334 700 130 570 040 377 503
374 350 040 333 750 300 435 333 704 114 333 450 750 034 530 750 373 357 570 503
733 571 435 747 501 317 073 357 145 433 374 000 507 075 550 445 411 734 343 430
047 070 413 014 503 374 577 353 050 053 441 053 150 300 505 030 374 755 350 305
037 701 073 543 017 335 341 444 137 304 503 047 334 100 305 030 747 577 457 354
444 007 415 209 142 223 020 914 222 326 677 652 311 563 319 533 143 231 861 962
041 081 017 119 210 454 3

Figure 3: Looking for numeric patterns in an email message.

Information Analysis Tools

For Scenario 1, IAs were free to apply any analysis tools to which they normally would have access. We found that our IAs did not typically apply specialized computational tools in their analysis, but rather relied on basic applications.

Several IAs captured facts and relationships by simply sketching them on paper (see Figure 4). Another analyst elected to draw facts and relationships using Microsoft PowerPoint. As shown in Figure 5, this analyst identified relationships and patterns among people and specific topics or themes. She then redrew the graph restricting the nodes to known terrorists and terrorist groups (see Figure 6). In addition, she drew a timeline highlighting the dates of specific events as found in the documents (see Figure 7).

Another IA considered the spreadsheet to be the "Swiss army knife" for intelligence analysis. As he read through documents, he would identify and cut and paste key facts from the documents into a spreadsheet. During this process, categories for the key facts would naturally fall out, which the analyst would then apply as columns of a table. Should broader themes in the data emerge, the IA would move the newly themed data into a new table or spreadsheet.

Figure 8 shows two of the tables the IA constructed for Scenario 1. During the course of analysis, the IA would highlight and un-highlight different cells of the spreadsheet as he found the contained information to be more or less relevant. He would also completely delete information once he was convinced the data was extraneous to the investigation. When new significant discoveries or themes emerge during analysis, the analyst would sometimes save a new version of the entire spreadsheet file. For Scenario 1, the analyst saved eighteen different versions of his analysis spreadsheet, which captured his overall progression of analysis and work. The analyst would return to previous versions of his analysis to review the path or progress of his investigation or to reclaim information that previously had been deemed irrelevant and deleted, but then found to have greater significance due to new incoming information.

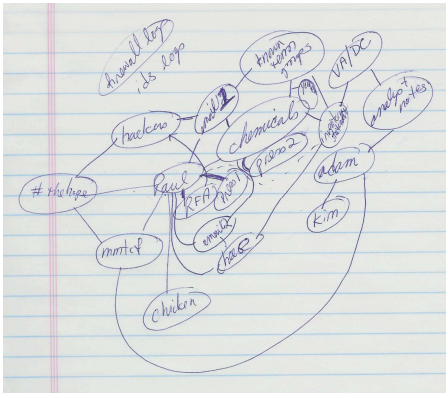

Figure 4: Graph containing facts and relationships.

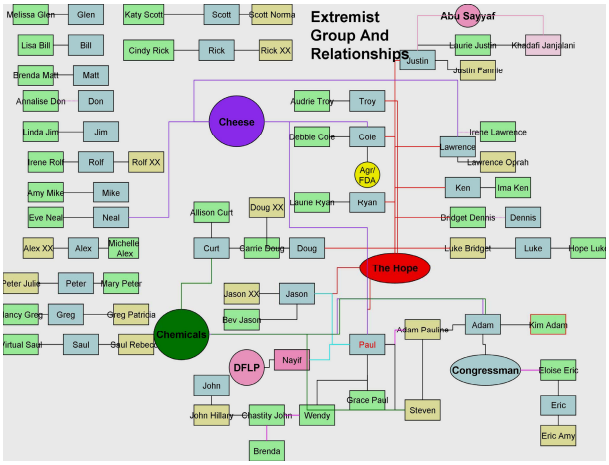

Figure 5: Graph containing people and topics.

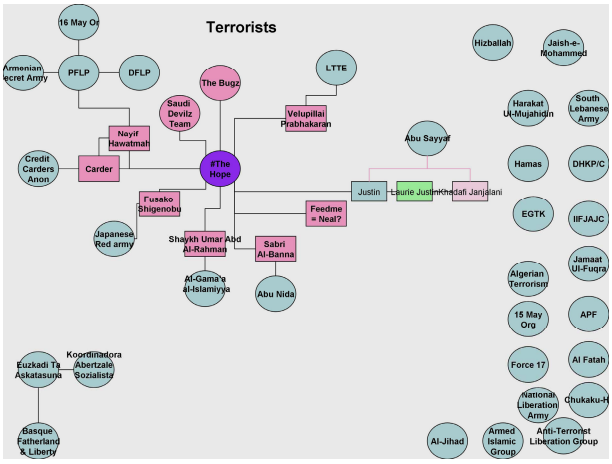

Figure 6: Graph containing terrorists and topics.

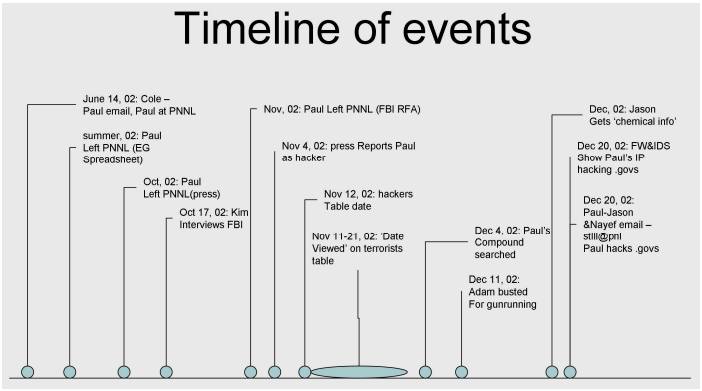

Figure 7: Timeline of events.

As we have shown, both graphs and spreadsheets may be used to collect and derive relational information. Given time, IAs are able to craft the use of everyday computer tools and representations to fulfill their analysis needs.

In Scenario 2, we found that IAs would use common, everyday physical tools and representations to orient and view data from different perspectives in hopes of discerning patterns of behaviors and activities. For example, one of the analysts tore a calendar page from her personal notebook organizer and circled the dates on which crimes occurred (see Figure 9). A pattern she found was that no crimes occurred on Tuesdays and Thursdays. Two other analysts plotted the locations of crimes on paper maps that had been provided. They annotated certain locations with letters to indicate the crimes that were committed at those locations.

Figure 9: Annotated calendar.

Continuing with the calendar, two IAs marked all the crimes on the timeline of Figure 10. The analysts noticed that the crimes seemed to move from the Prospect District to the Pine Hill District and then to both the Mississauge and Weston Districts. As shown in Figure 11, the two IAs then listed the crimes under the districts in which they occurred and noticed a pattern that the crimes became progressively more violent over time.

IAs often annotated and encoded information on documents using colored highlighters and pens. As shown in Figure 12, an analyst coded different crimes on a police report using different colors and symbols. She noted the days of the week that crimes occurred to examine whether weekly patterns existed. She also highlighted critical details and discrepancies in blue.

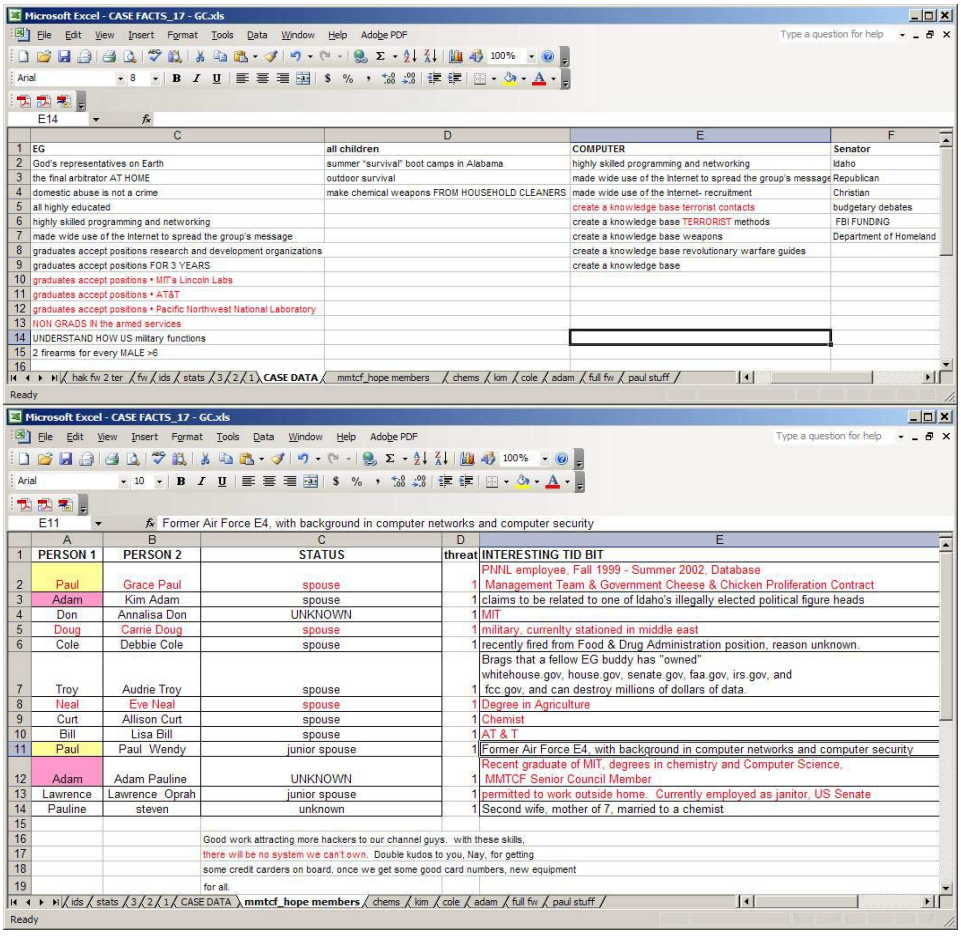

Figure 8: Intelligence analysis using spreadsheets.

Figure 13 shows a paper map annotated by IAs. Dots on the map indicate locations of crimes, dots annotated with the letter "R" indicate locations of rapes, circled blue dots indicate locations of murders, and the single blue circle indicate the location of a combined rape and murder.

Evidence and Credibility

IAs speak in terms of "facts." In the context of intelligence analysis, facts are not necessarily concrete truths. Rather, as one analyst describes, "When I say facts, it doesn't mean it's true, it's simply the evidence we have."

IAs in our study view intelligence information through different prisms. One analyst, who "was

paid to be paranoid," had a general mistrust of information until that information could be corroborated by other sources. Another analyst had the following pragmatic view,

> I like to think information is valid, unless I have reason to think it's otherwise. You have to work with what you're given. If the data conflicts itself, then obviously your confidence in the data degrades. But that can be a significant finding, when you find something invalid.

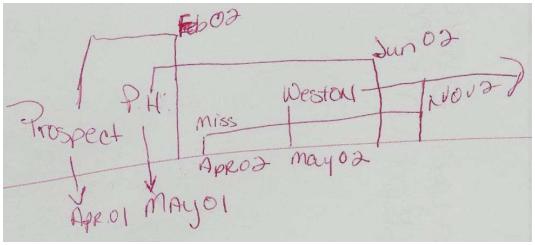

Figure 10: Crime timeline.

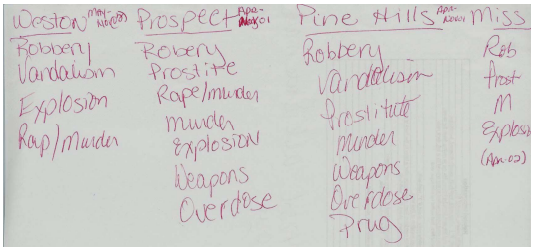

Figure 11: Crimes by district.

The topics and relationships that emerge from the documents have variable levels of credibility and validity as determined by the individual IAs. The level of credibility and validity is often judged by the credibility of the information source, the circumstances under which it was collected, and corroboration by other sources. In some cases, information may conflict, which degrades the credibility and validity of that information.

Data sources are also judged to have different levels of credibility. For example, IAs generally consider information found on Websites to have low credibility. The exception would be Websites operated by known, reputable organizations with credible references and publications. In assessing reports and interviews received from voluntary witnesses and external parties, IAs need to consider whether a report is intended to "inform" or "influence." Information gathered through technical means such as networks and firewalls are generally viewed as credible, but even so, such data may still be deceived or compromised such as through "spoofed" IP addresses. Information without attribution, where the source and circumstances of the data are unknown, is considered highly suspect.

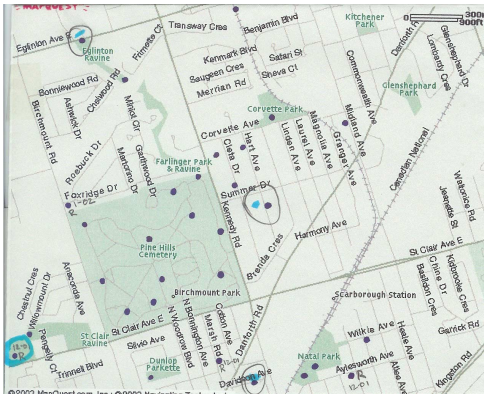

Figure 13: Annotated paper map.

The corroboration of information is an ongoing activity during analysis. As described by an analyst,

> There is a certain amount of protection we have to give to every source of information and so that it continues and doesn't dry up. At the same time, we're constantly checking and double-checking the veracity of each source we use, whether it be a technical source or some means of gathering some type of signal, or whether it is a human source. We're constantly checking through other sources of information to add credence to what that person or technology is giving us to provide a higher-level of trust in that information.

IAs still must deal with information that may be ambiguous, incomplete, and/or conflicting. Imperfect information has undesirable consequences as one analyst notes,

> Everything I do has to be weighed against something measurable to say there is more veracity to this piece than that, but even at that point, I have to understand that as an analyst, I may prove that wrong. I may prove that that's not valid. On occasion, I have conducted analysis and filed reports, but then found out that a source gave information just for money and not having any true facts. When this happens, the whole analysis becomes invalid.

An important step in the analysis process is to eliminate data and information that are deemed irrelevant to narrow the scope of the investigation and concentrate on the more essential details and facts. Recalling the people and topic graphs of Figures 5 and 6, a number of relationships existed outside the larger graphs. Here, the

Crime	Location	District	Date/Time	Victim/Accused	Related Information
Robbery	St. Clair Ravine	Pine Hills	December 1, 2001	Harriet Lodgings	19-year-old-female victim. Purse snatched.
Over-dose	Greenlaw Ave.	Prospect	December 10, 2001	Velva Wright	Librarian found dead near the university. She was walking home from work. Murder occurred around 11 p.m.
Robbery	Aylesworth Ave.	Pine Hills	December 11, 2001	Jane Morena	Domestic -- stolen ring, cash, and DVD player.
Weapons trafficking	Gilbert Ave.	Prospect	December 25, 2001	Hector Tongilski	
Robbery	Marsh Rd.	Pine Hills	December 27, 2001	Ed Kletzman	House robbery. Walls defaced with colorful paint.
Robbery	Foxridge Dr.	Pine Hills	January 2, 2002	Pepper West	18-year-old-female victim. Mugged. Wanted to kidnap her but she got away.
Over-dose	Railroad	Pine Hills	January 9, 2002	Marilynn Raffety	University student found dead after leaving her study group for the evening. Murder occurred around 1 a.m.
Weapons trafficking	Kennedy Rd. and Cemetary	Pine Hills	January 24, 2002	Carl Youngblood	
Murder	Hope St. and Day Ave.	Prospect	February 11, 2002	Mark Garrett	Black man gunned down by unknown person driving a blue Oldsmobile. Car never found.
Vandalism	Pine Hills Cemetery	Pine Hills	February 14, 2002	none	Cemetary headstones. Crushed.
Robbery	Corvette Park	Pine Hills	February 15, 2002	Caroline Montvale	33-year-old victim. Purse snatching. Nothing of value was in purse.
Drug trafficking	McRoberts Ave and Cemetary	Prospect	February 26, 2002	Calvin Bass	
Murder	Eglington Ravine	Pine Hills	March 13, 2002	Antonio Gandee	Hispanic male that was assulted, beaten, and killed with a knife to the liver.
Vandalism	Pine Hills Cemetery	Pine Hills	March 17, 2002	none	Cemetary headstones. Crushed.
Drug trafficking	Farlinger Park and Ravine, near cemetary	Pine Hills	March 28, 2002	Fred Meyer	
Robbery	Mississauga Hospital - Building C	Mississauga	April 12, 2002	none	Bank robbery during the middle of the night. Unknown items were taken from safety deposit boxes, including cash.
Vandalism	Pine Hills Cemetery	Pine Hills	April 22, 2002	none	Cemetary headstones. Painted and defaced.
Robbery	Hurontario St.	Mississauga	April 27, 2002	Jack Cox	2 males. Stolen ATM card.
Vandalism	Pine Hills Cemetery	Pine Hills	April 29, 2002	none	Cemetary headstones. Painted and defaced.
Vandalism	Pine Hills Cemetery	Pine Hills	May 8, 2002	none	Cemetary headstones. Painted and defaced.

Figure 12: Codes and annotations on police report.

analyst set people and relationships aside to concentrate on those that were most critical and intertwined in the case.

When writing reports, IAs also convey the credibility and confidence they have in their findings and conclusions through the language they use. As one analyst describes,

> An analyst will generally not make a statement that "this is exactly what happened." He will use words to moderate. "There are indicators." "There are some indications." The strength is often couched in the language in which we write things.

Another analyst confirms this view,

> Part of passing information on to other people in the intelligence community is to make sure you put in the appropriate qualifiers. Language may include "a credible source," "a very credible source," "a not so credible source," so when the reader reads the piece of intelligence you create, he has a sense of how credible the overall document is by the use of these words.

IAs themselves are also data sources when they publish and disseminate intelligence reports. Over time and through experiences, IAs evolve their own views on the credibility of other analysts. As one analyst describes,

> As you do analysis work, the longer you're in the field, you know who you can trust, you know about the circular reports, you know who plagiarizes.

A circular report is a reported analysis that is not directly based on collected facts and information, but rather on the analysis results and conclusions of others.

In most cases, IAs are generally viewed as credible sources because analysts want to develop strong reputations. As one analyst relates,

> One of the most closely guarded things is an analyst's reputation. A lot of this goes back to that. "Can this person be trusted?" And so, we have a lot to invest in making sure that information is clean.

The credibility of information, evidence, and people are subjective and will depend on an analyst's experiences, exposure, and general outlook. For example, an analyst that is generally skeptical of all evidence is going to judge credibility much more harshly than the analyst that accepts all evidence to be true until it is conflicted.

IAs may also consider different details and evidence to be relevant in a case. For Scenario 1, for instance, two IAs believed the cheese information was irrelevant while the other three considered it important. In addition, one analyst believed the firewall information was irrelevant, while the others did not. The judgments and conclusions IAs reach depend largely on their personal and professional experiences. Regarding the cheese information, for example, an analyst might have particular knowledge on the ingredients of cheese or how it is manufactured, and this specific knowledge would then lead the analyst in a particular analysis direction.

Collaboration

The lack of collaboration among intelligence agencies has been noted as one of the factors that contributed to the inability of the United States to thwart the 9/11 terrorist attacks [20]. The intelligence community has long had a competitive culture and environment in which individual agencies sought to be the first to identify and extinguish national security threats. As an analyst laments,

> I like to refer to the intelligence community as independent organizations who take turns outside of a room where they go inside one at a time, grab a handful of puzzle pieces, walk out the other door, and never coordinate with one another to figure out what the heck the puzzle is suppose to be, and then have to report what this puzzle is all about.

For our IAs, collaboration was considered more important in the sharing of data and resources than in the sharing of analysis results and findings. As one analyst mentions,

> What I will not trust and put into my analysis is somebody else's analysis. I need to know the source of the information and build on that so that I can put my level of trust in it and then it's my name at stake when I provide an answer… I won't trust their analysis until I look at the source of the information, and it will be, "Do I agree with the conclusions that they came to based on the facts and the evidence?"

Yet, the IAs still view collaborative analysis as a useful and necessary approach, where they would work together on analysis rather than simply passing on results and conclusions. As described by one of the analysts,

> How you look at the data, how you twist the data and the assumptions you make, can lead you different ways. No matter how many different analysts you have, you'll probably have some differences in analysis. It's when all of our analysts get together and work out the differences and challenge each other with facts that we get to a better and more prominent answer.

A second analyst had a similar view,

> One of the things that is most beneficial is when you have a group, like the five of us here, with different backgrounds, and different assumptions built-in, attacking the same problem, and the value is combining that information, merging it, and deconflicting it. That's when you get the best information.

For Scenario 2, a desirable outcome of the collaboration among IAs was greater attention to resolving discrepancies in the data. The team of IAs spent significant effort and time in examining and resolving information that seemed ambiguous, conflicting, or illogical. For instance, the IAs noticed that the police reports indicated that a number of the ATM robberies occurred at the Pine Hill Cemetery. The IAs found this information to be strange since ATM machines are normally not found at cemetery sites. The analysts concluded that the cemetery listed in the police reports refer not specifically to the cemetery grounds but to the larger neighborhood in which the cemetery resides.

In another example, a discrepancy existed in the police report where one entry listed the crime as rape and murder, but connected material described the crime as a burglary.

The IAs concluded that the entry was incorrect and they simply ignored the entry in their analysis because the information conflicted and was not verifiable.

IMPLICATIONS FOR INFORMATION TECHNOLOGIES

In this qualitative study, we strive to understand and elucidate aspects of the work and practices of IAs so as to find opportunities for tool and technology development. We would like to identify information technology capabilities that would be useful to IAs and able to enhance their investigative and analytical abilities. The design of specific tools and systems should follow from this qualitative study.

As previously described, IAs will often abandon a systematic analysis approach for the sake of time. Without a systematic approach, however, IAs run the risk of missing important details and/or critical steps in their analyses. One solution might be to auto-generate a set of standard analysis perspectives given a set of facts and relationships. The analysts, for examples, often mapped data onto timelines, geospatial maps, and organizational structures and hierarchies. A computer system that could generate a set of standard views from a data set would provide the analyst a way to systematically explore and investigate the data along specific themes or patterns. Furthermore, since the data is pre-wired to be displayed in ways that analysts naturally slice and manipulate data, such systems would likely improve the speed and efficiency of IAs and their analyses.

Workflow management systems such as Kepler [21] and Taverna [22] may also be useful to organize and automate the analysis process. The dynamic, ad hoc manner in which intelligence analysis is conducted, however, may be too spontaneous to be captured and automated in workflows.

Pens and highlighters provide IAs natural and familiar tools for encoding and annotating information. In general, computer tools do not afford the same naturalness, dexterity, and ease-of-use as physical writing tools. To support free-form encoding and annotation, sketching tools [23] may be useful in allowing IAs to quickly capture concepts and attach them to documents and data. Optical character recognition tools may also be valuable for converting sketched information into text and graphical forms that would be more amenable to future editing. For collaborative encoding and annotating, shared window tools (e.g., Microsoft Live Meeting [24]) may provide collaborative analysis support to IAs separated by distance.

A benefit of laying out physical information on desks and floors is the large amount of real estate one can garner. Recent studies on multiple-screen displays [25] suggest that the virtual real estate on computer systems may be expanded to better accommodate users by adding more monitors. Such multi-screen systems may also be of use to analysts in providing larger collaborative spaces.

Intelligence data is typically multivariate. In intelligence analysis, however, IAs often view data along just one or two dimensions such as time and space, and then integrate those views largely in their heads. Research efforts to develop multivariate visualizations for the intelligence community [1] may prove to be valuable if they are able to evolve visualizations that are intuitive, accurate, and conform to the analytical views and models held by IAs.

Collecting evidence to confirm facts and support conclusions is an important process for IAs. Evidence tools may be applied to collect and graphically link evidence to facts and conclusions. More sophisticated evidence tools such as DECIDE [26] may also be useful for automatically generating confidence scores provided that IAs find such tools accurate and trustworthy.

When IAs create graphs by drawing facts and relationships, they conduct *link analysis* [27]. In our study, IAs created link analysis graphs using basic drawing tools (e.g., pen and paper, Microsoft PowerPoint), but a variety of link analysis tools are commercially available including i2 Analyst's Notebook [28] and Visual Analytics VisuaLinks [29]. The benefit of using these commercial tools for link analysis is that they provide high-level capabilities for storing, managing, editing, and querying link analysis graphs and data import features that support the ingestion of data from many different kinds of data sources.

During the study, IAs frequently described how much they rely on history when conducting predictive analysis. They predict the future by mapping emerging information and facts onto those of past events and historical situations. To support these kinds of comparisons, case management tools for collecting, managing, and querying past cases would be of benefit. Research graph systems, such as the Scenario and Knowledge Framework for Analytical Modeling system [30], provide more sophisticated pattern-matching capabilities that may allow IAs to locate past cases more effectively, accurately, and efficiently.

CONCLUSION

In this paper, we presented an observational study that characterized and analyzed the analytical processes of IAs. We elaborated various qualities of analysts' work such as what investigative methodologies do they apply, how do they collect and triage information, how do they identify patterns and trends, what physical and computational tools do they apply, how do they work with hypotheses and evidence, and how do they collaborate on analysis. Furthermore, we discussed the relevance and application of specific information technologies that may support various aspects of intelligence analysis.

ACKNOWLEDGMENTS

The research described in this paper was conducted under the Laboratory-Directed Research and Development (LDRD) Program at the Pacific Northwest National Laboratory, a multiprogram national laboratory operated by

Battelle for the U.S. Department of Energy under Contract DE-AC06-76RL01830.

REFERENCES

1. Thomas, J.J. and Cook, K.A. (2005). *Illuminating the Path: The Research and Development Agenda for Visual Analytics*, IEEE Computer Society, Los Alamitos, CA.
2. Card, S.K. (2005). The science of analytical reasoning. In Thomas, J.J. and Cook, K.A. (Eds.), *Illuminating the Path: The Research and Development Agenda for Visual Analytics*, IEEE Computer Society, Los Alamitos, CA, pp. 33-68.
3. Chen, H., Wang, F. Y., and Zeng, D. (2004). Intelligence and security informatics for homeland security: information, communication, and transportation. *IEEE Transactions on Intelligent Transportation Systems*, *5*(4), pp. 329–341.
4. Wong, P.C, Rose, S.J., Chin, G. Jr., Frincke, D.A., May, R. Posse, C., Sanfilippo, A., and Thomas, J. (2006). Walking the path: A new journey to explore and discover through visual analytics. *Information Visualization,* 5(4), pp. 237-249.
5. Yen, J. (Ed.) (2004). Special issue: Emerging technologies for homeland security. *CACM*, 47(3).
6. Ehn, P. (1988). *Work-Oriented Design of Computer Artifacts,* Arbetslivescentrum, Stockholm.
7. Greenbaum, J and Kyng, M. (Eds.) (1991). *Design at Work: Cooperative Design of Computer Systems*, Lawrence Erlbaum Associates, Hillsdale, NJ.
8. Kyng, M. (1995). Creating contexts for design. In Carroll, J.M. (Ed.), *Scenario-Based Design: Envisioning Work and Technology in System Development*, John Wiley & Sons, New York, pp. 85-107.
9. Carroll, J.M. (Ed.) (1995). *Scenario-Based Design: Envisioning Work and Technology in System Development*, John Wiley & Sons, New York.
10. Carroll, J.M. and Campbell, R.L. (1989). Artifacts as psychological theories: The case of human-computer interaction. *Behaviour and Information Technology*, 8, pp. 247-256.
11. Monk, A.F. and Wright, P.C. (1991). Claims, observations, and inventions: Analyzing the artifact. *SIGCHI Bulletin*, 23(1), pp. 52-54.
12. Buchenau, M. and Suri, J.F. (2000). Experience prototyping. *Proc. DIS 2000,* ACM Press, pp. 424-433.
13. Seland, G. (2006). System designer assessments of role play as a design method: A qualitative study. *Proc. NordiCHI 2006,* ACM Press, pp. 222-231.
14. Lincoln, Y.S. and Guba, E.G. (1985). *Naturalistic Inquiry*, Sage Publications, Newbury Park, CA.
15. Floyd, C., Mehl, W.M., Reisin, F.M., Schmidt, G., and Wolf, G. (1989). Out of Scandinavia: Alternative approaches to software design and system development. *Human Computer Interaction*, 4(4), pp. 253-350.
16. Muller, M.J. and Kuhn S. (Eds.) (1993). Special issue on participatory design. *CACM*, 36(4).
17. Dulles, A.W. (2006). *The Craft of Intelligence: America's Legendary Spymaster on the Fundamentals of Intelligence Gathering for a Free World*, Lyons Press, Guilford, CT.
18. George, R.Z. and Bruce, J.B. (2006). *Analyzing Intelligence: Origins, Obstacles, and Innovations,* Georgetown University Press, Washington, DC.
19. Heuer, R.J. (1999). *Psychology of Intelligence Analysis*, US Government Printing Office, Washington, DC.
20. National Commission on Terrorist Attacks upon the United States. (2004). *The 9/11 Commission Report: Final Report of the National Commission on Terrorist Attacks Upon the United States*, W.W. Norton, New York.
21. Ludäscher, B., Altintas, I., Berkley, C., Higgins, D., Jaeger-Frank, E., Jones, M., Lee, E., Tao, J., and Zhao, Y. (2005). *Concurrency and Computation: Practice and Experience*, 18(10).
22. Oinn, T., Greenwood, M., Addis, M., Alpdemir, M., Ferris, J., Glover, K., Goble, C., Goderis, A., Hull, D., Marvin, D., Li, P., Lord, P., Pocock, M.R., Senger, M., Stevens, R., Wipat, A., and Wroe, C. (2005). Taverna: Lessons in creating a workflow environment for the life sciences. *Concurrency and Computation: Practice and Experience*, 18(10), pp. 1067-1100.
23. Landay, J.A. and Myers, B.A. (2001). Sketching interfaces: Toward more human interface design. *IEEE Computer*, 34(3), pp. 56-64.
24. Microsoft Office Live Meeting. http://www.microsoft.com/uc/livemeeting.
25. Colvin, J., Tobler, N., and Anderson, J.A. (2004). Productivity and multi-screen computer displays. *Rocky Mountain Communication Review*, 2(1), pp. 31-53.
26. Cluxton, D. and Eick, S.G. (2005). DECIDE™ hypothesis visualization tool. In *Proc. Intl. Conf. on Intelligence Analysis 2005*.
27. Sparrow, M. (1991). The application of network analysis to criminal intelligence: An assessment of the prospects. *Social Networks*, 13, pp. 251-274.
28. I2 Analyst's Notebook. http://www.i2inc.com/products/analysts_notebook.
29. Visual Analytics VisuaLinks. http://www.visualanalytics.com/products/visuaLinks.
30. Chin, G. Jr., Kuchar, O.A., Whitney, P.D., Powers, M.E., and Johnson, K.E. (2004). Graph-based comparisons of scenarios in intelligence analysis. In *Proc. IEEE SMC 2004,* IEEE Press, pp. 3175-3180.