Investigation of Ranking Methods Within the Military Value of Information (VoI) Problem Domain

Behrooz Etesamipour$^{(\boxtimes)}$ (ID) and Robert J. Hammell II

Department of Computer and Information Sciences, Towson University,
Towson, MD, USA
{betesamipour, rhammell}@towson.edu

Abstract. Determining the relative importance among vast amounts of individual pieces of information is a challenge in the military environment. By aggregating various military intelligence experts' knowledge, decision support tools can be created. A next step in the continuing research in this area is to investigate the use of three prominent ranking methods for aggregating opinions of military intelligence analysts with respect to the Value of Information (VoI) problem domain. This paper offers discussion about ongoing VoI research and demonstrates outcomes from a military-related experiment using Borda count, Condorcet voting, and Instant-runoff voting (IRV) methods as ranking aggregation models. These ranking methods are compared to the "ground truth" as generated by the current fuzzy-based VoI prototype system. The results by incorporating the ranking models on the experiment's data demonstrate the efficacy of these methods in aggregating Subject Matter Expert (SME) opinions and clearly demonstrate the "wisdom of the crowd" effect. Implications related to ongoing VoI research are discussed along with future research plans.

Keywords: Value of Information · Decision support · Information aggregation · Borda count · Condorcet voting · Instant-runoff voting · Rank aggregation

1 Introduction

The enormous volume of data generated everyday by computer systems and internet activities around the world cannot be easily processed and prioritized. It is an overwhelming challenge to analyze all pieces of data. The concept of Big Data Analytics introduces new challenges with its characteristics of enormous growth in data size, volume, and velocity as well as variability in data scope, data structure, data format, and data variety [1]. There are limitations in technological processing resources and human analytical expertise that do not allow the examination of all data generated every day. In a time-constraint environment such as the military, prioritizing the data can help to converge attention on the most important data first.

A "ranking" challenge in everyday human life can be defined as when someone is trying to rank a set of items based on some criterion with the goal in-mind to order those items from "best" to "worst" within some context. Essentially, the human

decision making process subconsciously reasons over multiple pieces of information and orders them in some way to achieve the "best" possible decision. Accordingly, an entirely new challenge can be identified when there are multiple individuals with multiple opinions trying to reason over multiple pieces of information.

Knowledge elicitation and aggregation from multiple individuals can sometimes provide a better outcome than the answer of any individual. This demonstrates the "wisdom of the crowd" phenomenon in which the aggregated crowd's outcome is closer to correct answer than all or most of the individual answers [2]. This work considers the matter of aggregating information by first starting with the task of ranking multiple independent judgments from multiple independent individuals. The emphasis of this paper is towards military decision making where vast amounts of data gathered from collective intelligence undertakings need to be prioritized.

Information assessment to judge and analyze the high value information, termed as Value of Information (VoI) [3], is very critical for military operations. In recent work to automate the VoI determinations, a Fuzzy Associative Memory (FAM) architecture was used to develop a decision support system for military intelligence analysts [3]. A fuzzy-based prototype system was constructed to provide VoI ratings for individual pieces of information considering the characteristics of information content, source reliability, timeliness, and mission context [3]. Later research was done that included additional knowledge elicitation efforts with subject matter experts that resulted in a complex, multi-FAM system [4].

The approach for this research is to explore the use of the ranking and voting methods of Borda count, Condorcet voting, and Instant-runoff voting (IRV) with respect to the VoI problem domain. Initially, all methods will be compared to the VoI determinations produced by the original fuzzy-based VoI system. The results from the comparisons will illustrate how well these ranking systems match the fuzzy-based approach. Using these aggregation models can also provide a way to quantitatively assess the efficacy of the recently developed VoI prototype since there are no other VoI-producing systems to compare with the fuzzy method.

The remainder of this paper is organized as follows: first, background information is presented pertaining to aggregation and some of the more popular approaches, the military value of information challenge, and a brief description of the current VoI prototype. Following that are sections that describe an experiment using Borda count, Condorcet voting, and Instant-runoff voting models and then the methodology for moving forward to accomplish the above aim of this work. Finally, conclusions and future work are discussed.

2   Background

Information judged on source reliability and content importance can be prioritized in different levels to be addressed and taken into actionable decisions on a given timeline. Information about the enemy, the battlefield environment, and the situation allow the commander and staff to develop a plan, seize and retain the initiative, build and maintain momentum, and exploit success [5]. Information aggregation and prioritization are key components in this process.

2.1 Aggregation

Based on a phenomenon called the "wisdom of crowd" [2], the aggregated rank of the crowd's choices in a voting poll has been generally identified to provide an estimate very close to the true answer; this is very helpful in many estimation tasks. This aggregated rank of the crowd's choices is usually represented in an order from the best value to the worst value. The "wisdom of the crowd" effect was first demonstrated by Francis Galton, who showed that averaging estimates from individuals regarding the weight of an ox produced a close approximation of the actual weight [6]. Consequently, many researchers believe "Crowdsourcing" [7] is driving the future of businesses by obtaining opinions on different subjects from a large and rapidly growing group of Internet users. The "wisdom of the crowd" approach is used in multiple real-world applications ranging from prediction applications aimed at markets and consumer preferences to web-based applications such as spam filtering, and others [8].

This research investigates aggregation within the military VoI problem domain by first considering aggregation related to ranking problems. The topic of aggregation and its need for, and approaches to, combining rankings from multiple sources is certainly not a new challenge; in fact, it has been studied for hundreds of years. The earliest examples of combining rankings relate to the area of "voting" and go back to the 18th century. The French mathematician and astronomer Jean-Charles de Borda in the 1700s proposed a voting method wherein voters would rank all candidates rather than selecting only one [9].

Rank aggregation has long been used in the social choice domain as well as in such fields as applied psychology, information retrieval, marketing, and others; a classification of rank aggregation methods is offered in [10]. Consistent with the overall goal of this research the focus is currently on Borda Count, Condorcet, and Instant-runoff ranking approaches. These three ranking models were chosen as representations to investigate the use of rank ordering techniques within the VoI problem domain.

2.2 Methods of Data Aggregation

Ranking is an example of obtaining assessments from a group of people to achieve one single decision on a winner. Voting on the top movie, electing the president, passing a bill, and many more examples are efforts to produce one single decision. In addition to single winner determination in balloting, voting is used to produce a ranked list. The three ranking methods compared in this paper are discussed below.

Borda Count

The Borda count model is a simple, statistical heuristic method that is widely used in voting theory [8]. The basic idea is that each item in an individual ranking is assigned points based on the position in which it is placed; then, the total points for each item is computed. The resultant totals for each item are used to sort the items and provide the aggregate ranking. One of the primary advantages of the Borda count method is that it is simple to implement and understand. Additionally, Borda count performs well with respect to rank performance measurements.

In this approach, for i items to be ranked by p participants, the total for the ith item, τ_i, can be written as:

$$\tau_i = \sum_{j=1}^{p} y_{ij}$$

where y_{ij} is the point value given to item i by participant j. The final, aggregated ranking is then found by ordering the item totals with any ties broken at random.

This traditional model is used in many applications, especially in sports. It is often used to choose the winners of sports awards such as the Heisman trophy in college football, selecting the Most Valuable Player in professional baseball, and ranking sports teams by the Associated Press and United Press International [11]. The Borda count model is considered as a baseline aggregation method for winner selection in voting. It is also used for aggregating rankings of data for decision making to produce the "wisdom of the crowd" phenomena.

The Borda count procedure is explained in [12] with an example for ranking 4 states in order from high to low by population with 5 participants. The model produces combined rankings that typically perform well relative to individual rankings.

Condorcet Method

The Condorcet method as an election method selects the winning candidate as the one who has gained the majority of the votes in an election process against all of the candidates in a head-to-head election comparison [13]. The Condorcet method requires making a pairwise comparison between every candidate. When there is single ranking item that beats every item in an election, it is called Condorcet Winner. Technically, the Condorcet winner candidate is the one candidate that wins every two-way contest against every other alternative candidate. The winning candidate should beat every other candidate in a head to head election which means the winning candidate should win a runoff election regardless of who it is competing against [14]. The Condorcet voting technique was first advocated by the 18th-century French mathematician and philosopher Marie Jean Antoine Nicolas Caritat, the Marquis de Condorcet [15].

A number of Condorcet-Compliant algorithms exist such that if there is a Condorcet winner, it would be elected; otherwise, they have different behavior. The proper Condorcet method is chosen based on how appropriate it is for a given context. A relatively new single-winner Condorcet election method called Schulze Voting or Beatpath is used in this paper and is described next. The Schulze method is recognized as a common means of solving a Condorcet's Paradox, which is a situation wherein the voters have cyclic preferences such that there is no Condorcet winner.

Condorcet/Schulze Voting (Beatpath)

As one of the Condorcet methods, Schulze Voting (or Schwartz Sequential dropping (SSD) or Beatpath) is a relatively new single-winner election method proposed by Markus Schulze in 1997 [16]. The Schulze method is an effective ranking method which many organizations and some governments are starting to apply. Debian Linux distribution has incorporated the Schulze algorithm into their constitution which can be found under the Appendix on the Debian Linux constitution [17].

While comparable to Borda count as a highly recognizable method for ranking, the Condorcet methods are more problematic to implement as they require pairwise comparisons between all candidates. Therefore, the Condorcet cycle of pairwise comparisons grows as the number of candidates grows.

In this method, after each voter ranks the candidates based on the order of preferences, a head-to-head comparison of all pairs of candidates is conducted to determine the winner of each pair. If there is one candidate that wins in all its pair comparisons, the candidate is Condorcet Winner. If there is no winner, the next step is to determine the pairwise preferences for all pair candidates in a matrix. For each head-to-head pairwise comparison, the number of voters who preferred candidate A over candidate B and vice versa is counted and noted. Once this is done, all the strongest paths for each pairwise comparison are identified; this is the most difficult and computationally intensive step. Finally, the items are ranked by their strongest path computations, producing the winner (and second place, third place, and so on). The full details of the algorithm, along with examples, can be found in [16].

Instant-runoff Voting

Similar to the Condorcet voting method, Instant-runoff voting (IRV) is a preferential ranking method which is used in single-seat elections; this method is useful when there are more than two competing candidates. Basically, voters rank the candidates or items in order of preference rather than showing support for only one candidate. There are countries that use Instant-runoff voting in their election systems such as selecting members of the Australian House of Representatives and the house of Australian state parliaments [18]. The IRV method establishes more fairness in an election when there are multiple candidates dividing votes from the more popular point of the political spectrum such that an unopposed candidate from the unpopular base might win simply by being unopposed.

In this method, once the ballots are counted for each voter's top choice, the candidate with the fewest votes will be eliminated if there is no candidate winning a simple majority (more than half) of the votes. The votes from the voters who have voted for the defeated candidate will be allocated to the total of their next preferred choice. This step will be repeated until the process produces a simple majority for one candidate; this candidate becomes the winner. At some point throughout this process, the race gets narrowed down to only two candidates; this is called an "instant runoff" competition which leads to a comparison of the top two candidates head-to-head. IRV can establish fairness and save the votes from like-minded voters supporting multiple candidates when it comes to a vote-splitting situation in an election.

In 2002, Senator John McCain from Arizona in his campaign supported instant-runoff voting and said that Instant-runoff voting "will lead to good government because voters will elect leaders who have the support of a majority." [19]. The Instant-runoff can avoid the chaos of the US 2000 presidential election and guarantee the elected candidates to have the broadest amount of support. Based on the "Record of Illinois 92nd General Assembly Bills", in 2002, Illinois Senator Barack Obama introduced SB 1789 in the Senate that created Instant-runoff voting for Congress in state primaries and local elections [20].

2.3   Rank Evaluation Metric

While none of the models discussed above require a known ground truth to produce a ranking, the existence of a ground truth will be assumed to allow the resultant rankings to be evaluated. The particular ranking metric will be discussed next.

Kendall's tau distance is an ordinal association measure between two random variables. Kendall's tau coefficient measures the rank correlation of similarities between two sets of ranks and it was developed in 1938 by Maurice Kendall [21]. This correlation coefficient based on Kendall's tau is used to compare the similarity of the derived rankings with the "ground truth" ranking. Kendall's tau is a non-parametric measure that is used to quantify the relationships between columns of ranked data. The computed value ranges from 1 to −1, inclusive [22]. A value of 1 means the rankings are identical while a value of −1 means the rankings are in the exact opposite order. Kendall's tau is calculated using the formula:

$$\frac{(C - D)}{(C + D)}$$

where C represents the number of concordant pairs in the ranking and D represents the number of discordant pairs. Concordant pairs are how many larger ranks are below a specific rank in the list; discordant pairs are how many smaller ranks are below a specific rank. This Kendall's tau value is explained in detail in [12].

Basically, Kendall tau distance measures the number of pairwise swaps needed to bring two orderings into alignment. The Kendall tau distance can be represented in a chance agreement distribution curve as shown in Fig. 1. Higher Kendall tau values indicate a greater disagreement between the ground truth and some resultant aggregated ranking. The lowest possible Tau value is 0, which indicates that the two rankings are identical. The highest possible Tau value can be calculated as:

$$Tau = n(n - 1)/2$$

where n equals the number of items being ranked. When comparing the aggregate ranking with the ground truth, Kendall tau distance can help to measure how closely rankings inferred by a ranking model match the latent ground truth. It is a way to quantify the relationships between columns of ranked data as well as the rankings provided by the participants.
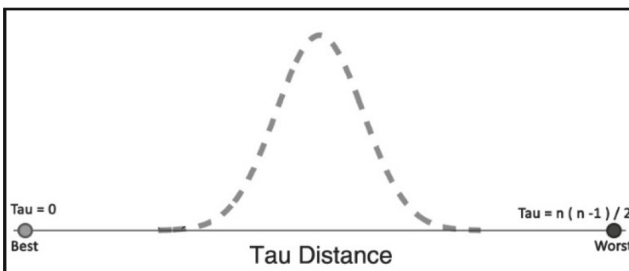

Fig. 1.   Kendall tau chance agreement distribution

2.4   Value of Information Challenge

Decision making in military environments involves massive cognitive and temporal resources to examine options and decide on the best alternative. Monitoring and analyzing real-time data to identify the most important information in the context of the mission at hand is challenging for intelligence groups. Information assessment to select and prioritize the high value information is very critical for military operations.

The Value of Information (VoI) metric is used to grade the importance of individual pieces of information. The process of making a VoI determination for a piece of information is a multi-step, human-intensive process. Intelligence analysts and collectors must make these decisions based on the characteristics of the information and also within different operational situations.

U.S. military doctrinal guidance for determining VoI is imprecise at best [5, 23]. The guidance provides two tables for judging the "reliability" and "content" of a piece of data, with each characteristic broken into six categories. Reliability relates to the information source, and is ranked from A to F (reliable, usually reliable, fairly reliable, not usually reliable, unreliable, and cannot judge). Information content is ranked from 1 to 6 (confirmed, probably true, possibly true, doubtfully true, improbable, and cannot judge).

This U.S. military doctrinal guidance does not clearly provide a method to associate these categories for information value determination. Moreover, there is no instruction on how to combine other attributes that may contribute to information value. Two other potential data characteristics include mission context (the operational tempo and decision cycle for a specific mission) and timeliness (time since a piece of information was obtained).

2.5   VoI Prototype Architecture

A prototype system using a Fuzzy Associative Memory (FAM) model has been developed to offer an effective framework for determining the VoI based on the information's content, source reliability, latency, and mission context [24]. For the prototype system, three inputs are used to make the VoI decision: source reliability, information content, and timeliness. The overall architecture of the fuzzy system is shown in Fig. 2. Instead of using one 3-dimensional FAM, two 2-dimensional FAMs were used; the reasoning behind this decision was presented in detail in [24]. The VoI metric is defined as the second FAM output and the overall system output. Note that mission context is handled by using three separate VoI FAMs. The correct VoI FAM is automatically selected based on the indicated mission context, which ranges from 'tactical' (high-tempo) to 'operational' (moderate-tempo), to 'strategic' (slow-tempo).

Fuzzy rules are used in the FAMs to capture the relationships between the input and output domains. Knowledge elicitation from military intelligence Subject Matter Experts (SMEs) was used to build the fuzzy rules [25]. More detailed descriptions of the FAMs, the fuzzy rules bases, the domain decompositions, and other implementation aspects of the prototype system can be found in [3] and [24]. The series of surveys and interviews with SMEs that were used to integrate cognitive requirements, collect functional requirements, and elicit the fuzzy rules is presented in [25].
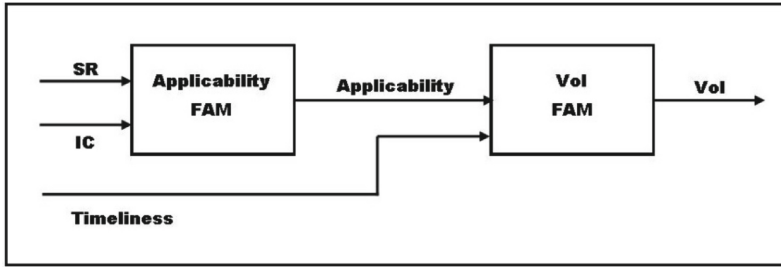
Fig. 2. VoI prototype system architecture

Note that there is no current system against which the results can be compared. As such, the system has not been tested comprehensively due to the human-centric, context-based nature of the problem and usage of the system. Approaches to validate (or partially validate) the system that do not require an extensive, expensive experiment are desired; this research seeks to assist in that effort (as explained further later).

3   Ranking Aggregation Experiment

This section describes an investigative experiment using the Borda count, Condorcet, and Instant-runoff voting methods to aggregate rankings from multiple participants. The experiment was devised not only to gather data for comparing the ranking methods to the VoI prototype system, but also to aid in understanding what data might be needed to relate the ranking models to future continued study in the VoI domain.

The VoI system details provided above in the Background section mention the use of SMEs and a knowledge acquisition process to provide a basis for constructing fuzzy rules. It should be clear that the involvement of multiple experts provides the potential for disagreement in deciding how to combine information characteristics, and multiple pieces of information, to arrive at a "value of information" determination. Another goal of this experiment was to provide first-hand familiarity regarding the efficacy of the Borda count, Condorcet, and IRV rank aggregation models with respect to their potential contribution to the VoI research.

3.1   Experiment Details and Implementation

During the summer of 2019, a team of researchers from the U.S. Army Research Laboratory (part of the U.S. Army Combat Capabilities Development Command) conducted an experiment with 34 military participants as SMEs. Each participant provided rankings for 10 different card decks, where each deck consisted of 5 or 7 cards (5 decks had 5 cards; 5 decks had 7 cards). Each card represented a piece of military information; each participant ranked each card (within each deck) based upon the attributes of source reliability, information content, and latency. An example card is depicted in Fig. 3.

Source Reliable	Information Content	Latency
B Usually Reliable	**3** Possibly True	**Somewhat Recent**

Fig. 3. Value of Information card attributes

The experiment was conducted via a computerized online interface. The resultant "Ranked Data Set" reflects the SME's determination of how each piece of information ranks with respect to information value (VoI). The highest ranking card represents the information with the greatest perceived "value" (highest VoI determination), while the lowest ranking card represents the piece of information with the lowest perceived value (lowest VoI determination). As mentioned before, each SME was charged with ranking 10 card decks in this manner. At the completion of the experiment, the Borda count, Condorcet voting, and IRV methods were used to aggregate the rankings of the 34 SMEs within each deck.

3.2 Borda Count, Condorcet, and IRV Implementation for Experiment

For the Borda count implementation, the calculations were performed in Microsoft Excel to convert the alphabetic reviewer designations to numeric rankings lists. The spreadsheet data were then imported into RStudio, an integrated development environment for the R programming language, where the aggregate Borda ranking was derived and compared to the ground truth, producing the Kendall tau distance.

For the Condorcet method, the algorithm presented in [16] was implemented in the "R" programing language and executed in RStudio. This function performs the ranking, starting with the full ballot and finding a pure Condorcet winner if one exists; otherwise, the strongest path computations are done and a winner is computed based on the Schulze algorithm as described earlier.

In the Instant-runoff voting method, a Python module was used which determines the winner card of this experiment using Instant-runoff voting rules. The algorithm executed separately for each of the 10 card deck sessions. The algorithm initially counts all the first-place choices. If there is a card in deck that has a true majority, meaning over half of the counts, the winner card with the first place is recorded. But after the initial count, if none of the cards have over 50% of participants' choices, then the card that is in last place with least number of votes is eliminated. The participants who had selected that defeated card now have their second-place cards counted. The total counts are recorded again, and this process repeats until a true majority winner card is found. Once the winner is determined, the first preferred or winner card is recorded and removed from card deck to re-run this procedure for producing a ranking of cards for second place and-so-forth.

3.3 Methodology and Results

The Borda count, Condorcet voting, and IRV methods descried above were applied to the data sets of the 10 card decks independently; again, each deck had 34 rankings – one

from each SME. For each aggregated ranking, the tau distance was computed by counting the number of bubble sort swaps required to make it match the ground truth. The ground truth was derived by computing the VoI for each card using the prototype Fuzzy system, and then ranking the cards based on this value (highest to lowest).

The result for one of the 10 card sets is visually represented in Fig. 4. The performance of the rankings inferred by all three methods as compared to the individual participant rankings is shown; note that all three ranking methods had identical performance in this case. The value of the tau distance for each participant is approximated by the person symbols. The x-axis represents the range of tau distance values that are possible. Tau distance is the number of swaps needed to make some given ranking equal to the ground truth ranking; thus, it measures the "performance" of an SME's ranking relative to the "correct" answer [21].

The light gray circle on the left side indicates the best possible ranking which is equal to the ground truth (tau distance of 0). The dark circle to the right indicates the worst possible case in which the rank order is the total reverse of the ground truth (for 5 cards that value is 10). The dotted curve depicts the chance tau distance distribution which would correspond to rankings being generated at random. The tau distance for the rankings produced by all three methods is indicated by the light-dotted circle with tau distance of 1. In this representation, the small tau distance indicates that the models produce a combined ranking that performs well relative to the individual performances, and that the aggregated rankings are very close to the ground truth.

A summary of the experimental results is shown in Table 1. The ground truth ranking for each of the 10 decks is given along with the aggregated ranking produced by each of the ranking methods. The "B" rows represent the result from the Borda count method, while "C" represents Condorcet and "I" represents IRV. The "Tau Dist" column gives the Kendall Tau Distance for the associated aggregated ranking compared to the worst case number of swaps that would be possible. The Tau distance is used to measure the ranking accuracy relative to the ground truth.
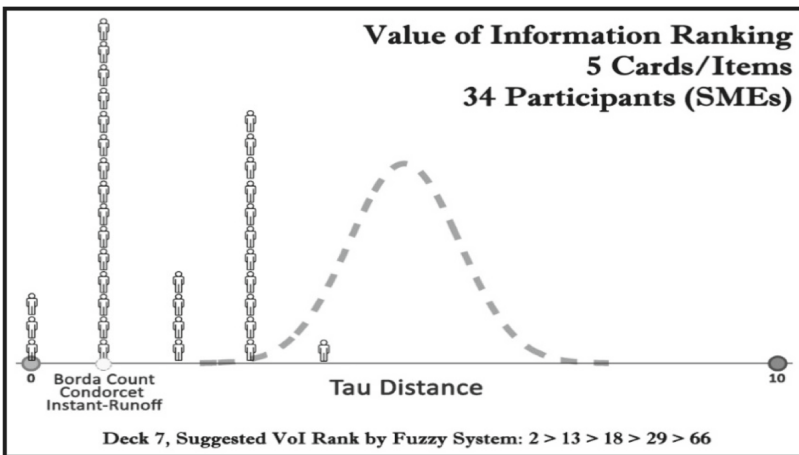

Fig. 4. Example ranking and Borda count result

The somewhat random card numbers in each deck have been mapped to an ordered set of integers reflecting the amount of cards used in the specific deck. This was done so that it is easier to understand how the aggregated rankings match the ground truth. For example, for Deck 7, the actual order of the cards for the ground truth is <2, 13, 18, 29, 66>; 2 is mapped to 1, 13 is mapped to 2, etc. So that <1, 2, 3, 4, 5> represents the ground truth order of the cards. The aggregated ranking produced by all 3 ranking models gave the actual order of <2, 18, 13, 29, 66>, which maps to <1, 3, 2, 4, 5>. It is easy to observe that the aggregated order produced by the ranking methods differs in the 2nd and 3rd cards, such that 1 swap (tau distance of 1) out of a possible 10 (worst case ranking) is required to achieve the ground truth order.

The Kendall tau distance values demonstrate that the aggregated rankings from the SMEs are relatively close to the ground truth. Note that the three methods are almost identical in their resultant aggregated rankings. The only differences are where the IRV method is 1 swap worse for Deck 8 and 1 swap better for Deck 2. Given that the Borda count model is much easier to implement, this model is deemed the "best".

Table 1. Experimental data ranking results

Deck	Ground Truth Ranking	Aggregated Ranking		Tau Dist
8	1 2 3 4 5 6 7	B	1 2 3 4 5 6 7	0/21
		C	1 2 3 4 5 6 7	0/21
		I	1 2 4 3 5 6 7	1/21
10	1 2 3 4 5 6 7	B	1 3 2 4 5 6 7	1/21
		C	1 3 2 4 5 6 7	1/21
		I	1 3 2 4 5 6 7	1/21
4	1 2 3 4 5 6 7	B	1 4 2 6 5 3 7	5/21
		C	1 4 2 6 5 3 7	5/21
		I	1 4 2 6 5 3 7	5/21
6	1 2 3 4 5 6 7	B	1 5 2 6 7 3 4	7/21
		C	1 5 2 6 7 3 4	7/21
		I	1 5 2 6 7 3 4	7/21
2	1 2 3 4 5 6 7	B	1 4 5 6 2 7 3	7/21
		C	1 4 5 6 2 7 3	7/21
		I	1 4 5 6 2 3 7	6/21
1	1 2 3 4 5	B	1 2 3 4 5	0/10
		C	1 2 3 4 5	0/10
		I	1 2 3 4 5	0/10
7	1 2 3 4 5	B	1 3 2 4 5	1/10
		C	1 3 2 4 5	1/10
		I	1 3 2 4 5	1/10
5	1 2 3 4 5	B	1 3 2 4 5	1/10
		C	1 3 2 4 5	1/10
		I	1 3 2 4 5	1/10
9	1 2 3 4 5	B	1 4 3 2 5	2/10
		C	1 4 3 2 5	2/10
		I	1 4 3 2 5	2/10
3	1 2 3 4 5	B	1 4 2 5 3	3/10
		C	1 4 2 5 3	3/10
		I	1 4 2 5 3	3/10

There are two cases where the Tau distance is 6 or 7 (out of a possible 21); the remainder of the results show much greater accuracy. These results, from Decks 2 and 6, have caused the ARL researchers to go back and look at their data and the fuzzy rules used to produce the ground truth ranking.

The experimental results demonstrate the efficacy of the ranking models in aggregating SME opinions, and highlight the "wisdom of the crowd" effect. As a corollary, it can also be said that the suggested rankings produced by the fuzzy VoI system (used as the ground truth) are reasonable given the aggregated SME rankings. As final highlight to this experiment, note that the ground truth and the aggregated rankings always agree on the most important piece of information (the first position in all rankings match) and they almost always agree on the least important piece of information.

3.4   Significance of the Experiment

While the Borda count, Condorcet, and IRV methods did not match the ground truth exactly in all instances, the methodology has shown promise to be a viable aggregation method in this particular domain. As previously mentioned, there is no current system with which to compare the VoI system results. Only a subjective validation of the efficacy of the systems by the SMEs has been possible. By providing a different approach for arriving at the basis for the rules and architectures of the systems, these methods are useful in providing the missing quantitative support of the systems.

As mentioned earlier, the fuzzy rules used in the VoI prototype systems were constructed based on knowledge elicitation processes with military intelligence SMEs. Notably, the experts did not always agree in their opinions and answers. The differences occurred in providing an applicability rating for a given SR/IC combination as well as in determining the VoI output for a given Timeliness/Applicability pattern. Further, it was not unusual for the SMEs to have varying interpretations of the linguistic terms (e.g. "fairly reliable", etc.). Based on the experimental results, it is hoped that one or more of the ranking methods can perhaps aid in better aggregating the expert's opinions in the knowledge elicitation process to create more accurate fuzzy rules. The observed differences between the fuzzy rankings and the aggregated SME rankings have already motivated a reexamination of the original fuzzy rule construction to ensure the systems are accurately representing the "ground truth".

The results herein display only the aggregated SME rankings and do not examine the spread of the 34 individual rankings (as shown in the Fig. 4 example). It is possible that extenuating circumstances could have influenced the aggregations in the instances where they did not closely match the ground truth. Factors such as the experts not fully understanding the nuances in a particular card deck (notably decks 2 and 6), or missing an understanding of the mission context in which the information would be used, or others could have come into play. Examination of variables that may have influenced the rankings is underway at the U.S. Army Research Laboratory.

Finally, while the current VoI systems assign a specific value determination to a piece or pieces of information, there may be times when military analysts simply need to rank information according to its perceived importance without the need for a numerical value. The three ranking models could certainly be useful in this regard.

4   Conclusion

In a military environment, the process of analyzing the relative value of vast amounts of varying information is currently carried out by human experts. These experts, known as military intelligence analysts, have the significant challenge of evaluating and prioritizing available information, in a time-constrained environment, to enable commanders and their staffs to make critical mission decisions. Within the domain of military intelligence analyst decision support, this work considers the matter of aggregating information. One goal of this research was to investigate the use of the Borda count, Condorcet voting, and Instant-runoff voting (IRV) aggregation models with respect to the Value of Information (VoI) problem. Another goal was to provide a way to quantitatively assess the efficacy of the recently developed VoI prototypes.

This paper presented discussion about ongoing VoI research and a preliminary experiment using the Borda count, Condorcet voting, and IRV models. The results demonstrated the usefulness of these ranking models in aggregating SME opinions and clearly highlighted the "wisdom of the crowd" effect. Additionally, this work offered some quantitative validation of the current VoI prototypes by providing results for comparison to those from the fuzzy-based systems. These efforts will help to optimize the current fuzzy rules, motivate additional knowledge elicitation efforts, and influence the development of other VoI decision support architectures altogether.

Future research will include the comparison of one or more of the models used here to the well-known Bayesian Thurstonian ranking model. As the U.S. Army continues to move toward improving its ability to create situational awareness, the ability to successfully aggregate information will be a critical factor.

References

1. Couldry, N., Powell, A.: Big data from the bottom up. Big Data Soc. **1**(2), 5 (2014). https://doi.org/10.1177/2053951714539277
2. Surowiecki, J.: The Wisdom of Crowds. Anchor, New York (2005). Reprint edition
3. Hanratty, T., Newcomb, E.A., Hammell II, R.J., Richardson, J., Mittrick, M.: A fuzzy-based approach to support decision making in complex military environments. Int. J. Intell. Inf. Technol. **12**(1), 1–30 (2016). https://doi.org/10.4018/IJIIT.2016010101
4. Hanratty, T., Heilman, E., Richardson, J., Caylor, J.: A Fuzzy-logic approach to information amalgamation. In: Proceedings of the 2017 IEEE International Conference on Fuzzy Systems, Naples, Italy (2017). https://doi.org/10.1109/FUZZ-IEEE.2017.8015667
5. US Army, US Army Field Manual (FM) 2-22.3, Human Intelligence Collection Operations, Washington, DC, Washington, DC: US Army, p. 384 (2006)
6. Galton, F.: Vox populi. Nature **75**, 450–451 (1970)
7. Estellés-Arolas, E., González-Ladrón-de-Guevara, F.: Towards an integrated crowdsourcing definition. J. Inf. Sci. **38**(2), 189–200 (2012). https://doi.org/10.1177/0165551512437638

8. Steyvers, M., Lee, M., Miller, B., Hemmer, P.: The wisdom of crowds in the recollection of order information. In: NIPS2009 Proceedings of the 22nd International Conference on Neural Information Processing Systems, Vancouver, British Columbia (2009). https://dl.acm.org/doi/10.5555/2984093.2984293

9. Borda, J.-C.: On elections by ballot. In: Histoire de l'Academie Royale des Sciences for the Year 1781 (1781)

10. Li, X., Wang, X., Xiao, G.: A comparative study of rank aggregation methods for partial and top ranked lists in genomic applications. Briefings Bioinform. **20**(1), 178–189 (2017)

11. Easley, D., Kleinberg, J.: Networks, Crowds, and Markets: Reasoning About a Highly Connected World, pp. 736–746. Cambridge University Press, Cambridge (2010). https://doi.org/10.1017/CBO9780511761942

12. Etesamipour, B., Hammell II, R.J.: Aggregating information toward military decision making. In: 17th IEEE/ACIS International Conference on Software Engineering Research, Management and Applications, Honolulu, Hawaii, (2019). https://doi.org/10.1109/SERA.2019.8886807

13. Pivato, M.: Condorcet meets Bentham. J. Math. Econ. **59**, 58–65 (2015). https://doi.org/10.1016/j.jmateco.2015.04.006

14. Newenhizen, J.V.: The Borda method is most likely to respect the Condorcet principle. Econ. Theory **2**(1), 69–83 (1992). https://doi.org/10.1007/BF01213253

15. Condorcet, M.de.: Essay on the Application of Analysis to the Probability of Majority Decisions (1785)

16. Schulze, M.: The Schulze Method of Voting. eprint arXiv:1804.02973 (2018)

17. Constitution, D.: Constitution for the Debian Project (v1.7) (2016)

18. Bennett, S., Lundie, R.: Australian Electoral Systems. Parliament of Australia Department of Parliamentary Services (2007)

19. Lavin, N., Robinson, R.: John McCain understood how ranked choice voting strengthens our democracy. FairVote (2018)

20. Richie, R., Penrose, D.: When Barack Obama Was a Leader in Seeking Fair Voting Systems. FairVote (2012)

21. Kendall, M.: A new measure of rank correlation. Biometrika **30**, 81–89 (1938)

22. Abdi, H.: The Kendall Rank Correlation Coefficient. Encyclopedia of Measurement Statistics. Sage, Thousand Oaks (2007)

23. North Atlantic Treaty Organization (NATO) Standard Agreement 2022. STANAG magazine, STANAG 2022 Intelligence Reports. (Edition 8) Annex

24. Hammell II, R.J., Hanratty, T., Heilman, E.: Capturing the value of information in complex military environments: a fuzzy-based approach. In: Proceedings of 2012 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2012), Brisbane, Australia, (2012). https://doi.org/10.1109/FUZZ-IEEE.2012.6250786

25. Hanratty, T., Heilman, E., Dumer, J., Hammell, R.J.: Knowledge elicitation to prototype the value of information. In: Proceedings of the 23rd Midwest Artificial Intelligence and Cognitive Sciences Conference (MAICS 2012), Cincinnati, OH (2012)