A METHODOLOGY FOR DETERMINING CRITICAL DECISION POINTS THROUGH ANALYSIS OF WARGAME DATA

A Thesis
Presented to
The Academic Faculty

by

Mahmoud A. Abdelaal

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of School of Aerospace Engineering
Georgia Institute of Technology
August 2016

A METHODOLOGY FOR DETERMINING CRITICAL DECISION POINTS THROUGH ANALYSIS OF WARGAME DATA

Approved by:

Professor Dimitri Mavris, Advisor
School of Aerospace Engineering
Georgia Institute of Technology

Professor Daniel Schrage
School of Aerospace Engineering
Georgia Institute of Technology

Professor Graeme J Kennedy
School of Aerospace Engineering
Georgia Institute of Technology

Dr. K. Daniel Cooksey
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Jean Charles Domerçant
Georgia Tech Research Institute
Georgia Institute of Technology

Date Approved: July 06, 2016

For Amber and Riley

ACKNOWLEDGEMENTS

I would like to thank Professor Dimitri Mavris for providing guidance and mentorship throughout my experience in graduate study. Your technical and academic guidance have been crucial to the development of this thesis and my competence in performing research. Additionally, your level professionalism something I one day strive to attain. I would also like to thank Dr. Charles Domerçant. As a fellow Veteran and accomplished researcher, your influence and counsel were essential during this process. Thanks to my other committee members as well, Professor Daniel Schrage, and Professor Graeme Kennedy and Dr. Daniel Cooksey, for your involvement and advice on my formulation and thesis. Your collective input elevated my work substantially.

I would also like to thank my friends and family for supporting me through this arduous and fulfilling journey from beginning to end. To my extended family, your encouragement through the years has been uplifting, allowing me to fight another day. To the many friends I have made while here at Georgia Tech, Thank you. Your support and comradery made the most grueling parts of this process achievable.

Lastly, I would like to thank my Wife, Amber Abdelaal. Your unwavering support and steadfast devotion were essential to completing this thesis. I could not have done this without you. I can't do much without you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

xiii

LIST OF SYMBOLS AND ABBREVIATIONS

Agent based models (ABM)

Artificial Intelligence (AI)

Balanced Scorecard (BSC)

Capability, Awareness and Position (CAP)

Clustering Analysis for Decision-point REcognition (CADRE)

Command, Control, Computers, Communications, Intelligence, Surveillance, and Reconnaissance (C4ISR)

Courses of action (COAs)

Critical Decision Points (CDP)

Dynamic Adaptive Planning (DAP)

Dynamic Time Warping (DTW)

Hierarchical Agglomerative Clustering (HAC)

Human Intelligence (HUMINT)

Joint and Coalition Operational Analysis (JCOA)

Joint Operation Planning Process (JOPP)

Line of Sight (LOS)

Line of Influence (LOI)

Longest Common Subsequence (LCSS)

Map Aware Non-uniform Automata (MANA)

Naval Simulation System (NSS)

Network (Net)

Network Centric Operations (NCO)

Network Centric Warfare (NCW)

Number of clusters (NC)

Number of nearest neighbors (k)

Number of nearest neighbors outside of cluster (q)

Operations Research (OR)

Probability of kill (P_{kill})

Reinforcement Learning (RL)

Root-mean-square standard deviation (RMSSTD)

Rules of Engagement (ROE)

Tactical Warfare Model (TACWAR)

Tactics, Techniques and Procedures (TTP)

Universal Joint Task List (UJTL)

SUMMARY

Utilizing the latest in technology, today's military is engaged in complex conflicts and operations across the globe. These disparate operations can occur simultaneously and within the same theater. The ability to handle the spectrum of operations in a single engagement or in multiple locations and engagements, requires an extremely capable force. However, difficulties in planning due to the blurring of lines connecting cause and effect on the modern battlefield inhibit the most efficient use of this capability.

Joint operational planning processes are often characterized as data, labor and time intensive, and courses of action (COA) must be planned and executed within the enemy's decision-making cycle while allowing planning time for lower command echelons. This represents a computational burden that scales dramatically with increasing numbers of systems and actors within the modern battlespace. Furthermore, the most time consuming part of the planning process is COA analysis and wargaming. This highlights the research objective developing a methodology to aid military planners by utilizing a new process for analyzing and evaluating COA alternatives.

This methodology will take shape from the merits of dynamic adaptive planning where, both reactive and anticipatory methods of plan adaptation are utilized to maintain operational effectiveness. This is done by analyzing simulated wargame data to identify leading indicators and metrics for monitoring the battlespace. By analyzing the effect of the interaction of the terrain, systems, and actors within the battlespace, it was hypothesized

that leading indicators and metrics associated with a specific battlefield configuration will allow the identification of battlefield heuristics.

The desire to obtain actionable knowledge from within the simulation data presents challenges specific to the problem. Wargaming different scenarios will result in a heterogeneous database to analyze because variations in plans that will be simulated will have an effect on the protraction or brevity of the simulated operation. Furthermore, this heterogeneous data is temporal in nature, where the metrics are time series associated with the occurrences on the simulated battlefield. Using these data constraints to develop the methodology, Hierarchical Agglomerative Clustering, using the Dynamic Time Warping algorithm for similarity measurement were chosen. Moreover, a novel method for cluster validation was created to establish relative value between different linkage algorithms using the similarity height of Dendrogram and the statistical significance of the within cluster outcomes.

A series of experiments were planned to test the hypotheses developed in this thesis. The first experiment was performed to test the effect that the number of clusters have on the rankings of the linkage methods provided through utilization of the HSig Index. The purpose of the next experiment was to verify the hypothesis referring to the need of CAP data in addition to traditional MOEs and MOPs for identification of CDPs. Additionally, in extracting well known chess heuristics from within the data and CDP analyses will showcase the tractability of the overall hypothesis of the ability to identify battlefield heuristics from leading indicators and metrics associated with a specific battlefield configuration.

As the number of pieces, board size, or number of actors in a scenario increases, so does the branching factor. The final experiment will be conducted to investigate what effect scaling the problem will have on the methodology by applying the methodology to two reduced datasets. For combinatorically complex problems it may not be feasible to have an exhaustive data set. This experiment is done to show the effect that complexity will have on both the methodology due to the size and nature of the data. This serves to inform the level of investment in data prior to performing analysis.

With the relevant experimentation compete in Experiments 1, 2, and 3, the methodology was utilized in an agent based model of a ground combat scenario between blue and red agents from the blue agent perspective. This was done in order to illustrate the merit of methodology on shifting from chess as a wargaming simulation to a full scale agent based modeling and simulation wargaming suite. Data specific to the simulated battlefield was utilized to affect the operation of the agents as they attempt to meet objectives. Additionally, COA components were developed as variations of routes taken and sequences of engagements with red agents. There were a total of 600 COAs simulated, collecting Capability and Position data for each simulation. The CADRe methodology was performed on each metric and identification of COAs and CDPs were attempted.

The primary contribution of this thesis is offering of enhanced COA analyses of wargame data allowing the identification of decision points and heuristics. This was motivated by the desire to provide a capability to recognize patterns in the evaluation metrics successful COAs. This was shown successful utilizing the methodology, CADRE, here. Resulting from the formulation and the experiments, this work has aided in the

creation of a cluster validation index that considers inner cluster similarity and the statistical significance of the outcome of scenarios. In addition, this work has further effectuated breadth of analysis by the determination of the effect of scaling of complexity in the number of elements and of the underlying wargame on the selection and performance of linkage algorithms, HSig Index, identification of decision points and heuristics. The motivated the next contribution, a process to analyze the data at differing resolutions, called HSig Optimization.

An investigation into the effect of form and complexity of the data was performed. This was motivated by the experimentation performed to test the methodology, and was found to vary quality of the output of this process, COAs and CDPs. The chief forcing function behind this effect was dependent on the factors surrounding the data retrieval process.

Next, an investigation into the effect of metric sensitivity was performed utilizing the CADRE methodology. This investigation yielded an identification of such sensitivity in the metric and are presented in this work. Therefore, as an additional contribution, this methodology enables the analysis of the sensitivity of performance metrics to the battlespace and associated decision space.

A baseline analysis for which to compare the results of the CADRE methodology was identified. It was shown that insights obtained utilizing the CADRE methodology would not have been identified through the baseline analysis case alone. There would be no other way to obtain the cluster information from the baseline case as well. It was shown that the ability to use cluster specific information on traditional scatter plots of MOPs

enable the CADRE information to be viewed from multiple dimensions. Thus, the CADRE

methodology can be utilized to further enhance traditional decision-making paradigms or

as a standalone paradigm.

CHAPTER I

INTRODUCTION

Utilizing the latest in technology, today's military is engaged in complex conflicts and operations across the globe. These engagements can range from humanitarian assistance, as seen in the Nepali earthquake relief effort, to combat operations, as seen in Iraq and Afghanistan. These disparate operations can occur simultaneously and within the same theater [1]. Additionally, mission creep can cause one type of mission or operation to turn into another. An extremely capable force has the ability to handle the spectrum of operations in multiple locations and engagements, or in a single engagement. Hence, utilizing the Global Firepower Index, the United States military ranks as the most powerful military in the world [2]. However, difficulties in planning due to the blurring of lines connecting cause and effect on the modern battlefield inhibit the most efficient use of this capability [3].

The difficulties associated with modern Network Centric Operations (NCO) extend beyond the disparity of potential operations. The utilization of coalitions which brings together heterogeneous command structures and equipment intensify these difficulties. This can be seen when observing the makeup of the coalition and assets used in Operation Odyssey Dawn. In order to implement a no-fly zone in Libya in 2011 the U.S. and coalition partners France, U.K., Italy, and Canada employed fourteen types of aircraft and five types

of ships and submarines. The operation lasted nine days with an estimated cost of $373.6B [4]. The obvious high cost of performing this operation would have been for not if improper planning lead to mission failure. Additionally, the lack of availability of resources used in this operation were not able to participate in other operations would have also been wasteful given mission failure. This articulates the 'high-stakes' nature of operational planning and the first order effects of failure with higher order effects influencing the socio-economic and political domains.

Given the challenges associated with modern warfare and the U.S. military's capability, the obvious challenge of employing capability presents itself. Mission failure or accomplishment can be traced back to how capability is utilized through planning. The remainder of this chapter will include an expansion of the challenges in military planning on the modern battlefield. Furthermore, planning paradigms are discussed and contrasted with respect to the effects of modern combat. The chapter closes on the formulation of an overall research objective.

1.1   Technology & the Modern Battlefield

Throughout history technology has been a potent enabler to capability on the battlefield. As technology matures, so does the practice of warfare because the nature of warfare changes as a result of new technology. The implementation of technology influences changes in tactics, techniques and procedures (TTP). This is due to the creation

of one or more friendly capabilities and/or cancellation of one or more enemy capabilities [5, 6]. This occurrence can be seen throughout history. An example is the transition from the Bronze to the Iron Age. Armies embracing iron technology were able to equip more soldiers with swords as a result of the capability of smelting a more readily available iron ore. This introduced a capability mismatch, forcing lower technology armies to modify TTPs, effectively limiting or canceling one or more capabilities while high-tech armies had to modify TTPs to introduce one or more capabilities as a result of technology infusion.

Today's battlefield is influenced by improvements in air power, imagery, electronic, and information technology warfare which have magnified effectiveness of combat power elements [7-9]. Specifically, these improvements have heightened the level of battlespace awareness and maneuvering capability, effectively increasing the information flow to make decisions and capability to act accordingly [8-10]. This increases the speed of combat. Additionally, the proliferation of networked systems and Command, Control, Communications, Computers, and Intelligence (C4I) technology, further enables capability but increases complexity, which is the diversity and number of components, interactions, and resulting behaviors [8, 10, 11].

The use of networked and C4I systems by battle-commanders is coined Net-Centric Warfare (NCW) [9, 12]. As classic performance improving technology, where systems are made physically better, is a capability multiplier, so is NCW. With it comes the promise of further increases in speed of combat, in addition to offering battle-commanders more

operational flexibility. But the NCO paradigm and the operational flexibility has made planning more difficult as well. With too much operational flexibility delegated to subordinate units, the commander runs the risk of losing control. However, with too little operational flexibility, commanders lack operational flexibility, and can potentially inhibit timely action [9].

The mechanism that allows NCOs to enable overall capability is illustrated in Figure 1. It is based on employing an information-based strategy to facilitate a competitive advantage [12]. According to Mitchell and Alberts et al. [9, 12], capability is enhanced through a 'NCW Value Chain' that begins with enabling information structure and leads to combat power. The enabling information structure facilitates awareness generation through fusion and management of sensor data. By generating increased and shared awareness, the enterprise can make resource allocation decisions that maximize value with respect to the overall enterprise as opposed to increasing value locally. The heightened awareness singularly, and by the whole, allows for processes that exploit awareness through the creation of a new virtual decision-making hierarchy. Time and space are collapsed allowing geographically dispersed teams to contribute to a common goal yielding a measurable benefit of increased tempo, responsiveness, and profit gained at a reduced cost and level of risk [12]. However, the current military planning paradigm can limit the effectiveness of NCO.

```
┌─────────────────────────────────┐
│      Information Structure       │
└─────────────────────────────────┘
                ↓
┌─────────────────────────────────┐
│         Sensor Netting          │
│           Data Fusion           │
│      Information Management      │
└─────────────────────────────────┘
                ↓
┌─────────────────────────────────┐
│    Vastly Improved Awareness    │
│         Shared Awarness         │
└─────────────────────────────────┘
                ↓
┌─────────────────────────────────┐
│      Virtual Collaboration      │
│      Virtual Organizations      │
│   Substitution of Information for│
│        People and Material      │
│        Self Synchronization     │
└─────────────────────────────────┘
                ↓
┌─────────────────────────────────┐
│         Increased Tempo         │
│     Increased Responsiveness    │
│           Lower Risks           │
│           Lower Costs           │
│          Higher Profits         │
└─────────────────────────────────┘
```
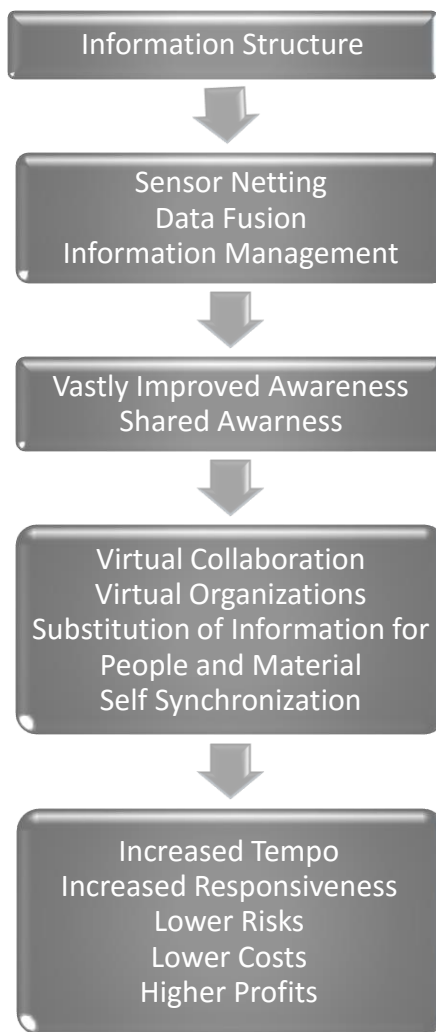
Figure 1. The Network Centric Enterprise and influence on organization through improved awareness and collaboration increasing the effects of capability[9, 12].

1.2   Challenges in Military Planning

Military planning is a constrained and multi-dimensional decision-making process in which planners must reconcile battlespace parameters and objectives [9, 13].

5

Specifically planners must reconcile ends, means, ways and risks in order to achieve a desired military end state [13]. This defines the process of defining a Course of Action (COA), a method in which to achieve some sort of military objective. The crux of this problem is in the reconciliation of the aspects of the battlespace parameters planners must resolve. These aspects are further explained and expended upon next.

The battlespace increases the dimensionality of the problem as it is the physical, information, cognitive and social domain where operations take place [9]. As introduced in the previous section, the impact of technology to the battlefield magnifies capability and increases the speed of combat, reducing the allowable time for action. Furthermore, action must be planned and executed within the enemy's decision-making cycle while allowing planning time for lower command echelons [13, 14]. Operations ranging from Peace Keeping to Power Projection rely on subsets of operational elements for accomplishment, adding to problem dimensionality as seen in Figure 2. These effects are magnified as the military planning process is often characterized as data, labor and time intensive [15]. Therefore, although the increased proliferation of NCO and technologies enabling NCO effectively increase capability, they compress action timelines. This compression exacerbates an already data, labor and time intensive process.

Figure 2. Operation types and required elements adding dimensionality of military planning as operation types range from peace keeping to power projection [1].

1.2.1 Uncertainty on the Battlefield

In addition to the dimensionality and time constraints imposed on planners, the battlefield is rife with uncertainty. Clausewitzian views on uncertainty are centered on the enemy's physical and strategic state and utilization of the "laws of probability" to decipher the enemy's intent. Another view on uncertainty considers imperfect knowledge. Clausewitz states "The only situation a commander can know fully is his own…" and knowledge about his enemy is only known through imperfect intelligence [16]. Additional consideration of uncertainty is included when considering aspects of general strategy, such as the element of surprise and the cunning of an enemy [16].

While many of Clausewitz's theories are timeless, traditional views of uncertainty are more holistic, where it is considered that a commander does not even know his own

7

situation fully [17]. Uncertainty represents a lack of knowledge and extends throughout the battlefield state and elements and into potential inferences from the state [18]. Although cases can be made to the contrary, generally, uncertainty can be attributed to the probabilities associated with detection technology. These technology solutions can only see what is physically observable, which excludes enemy intent [17]. Most uncertainty present on the battlefield can be reduced by gathering more information or generating more knowledge. This is classified as Epistemic Uncertainty [19]. However, Aleatory Uncertainty cannot be reduced and is the product of natural variability [20]. The potential for C4I technology is not expected to fully eliminate uncertainty [17]. Additionally, factors on the battlefield including time can affect the levels of uncertainty which commanders must reconcile.

Information on battlespace conditions obtained by the commanders is subject to uncertainty depending on the goodness of the information sources and the relative age of the information [21]. Additionally, the types of information retrieved may be subject to information atrophy, where the lifespan of the information can be affected by its criticality or the nature subject under consideration. For example, consider the lifespan of information of the location of an enemy mobile offensive system and the location of a bridge. The location of the bridge is time insensitive whereas the location of the enemy system is considered time sensitive by the nature of the system.

Information collection elements ranging from sensors to Human Intelligence (HUMINT) are associated with different levels of confidence respectively [22]. Source uncertainty is a byproduct of the level of reliability of the information collection element. However, this uncertainty of information gained from sensors is generally known as a function of system performance. Content uncertainty can vary depending on the source of HUMINT and, when dealing with multiple reports, leads to correlation uncertainty [22].

The preceding taxonomies of battlefield uncertainty are a first step in uncertainty reduction and management. These types of uncertainty are magnified by a complex operational environment and Complex Adaptive Systems seen on the modern battlefield [13, 23]. As a result, they must be reconciled prior to or during planning, further complicating the process of determining a proper course of action.

1.2.2 Impact of Ineffective Planning

Regardless of the specific mission, engagement, time allotted or information available, military personnel must develop an appropriate plan prior to commencing operations. Failure to do so can lead to mission failure or mission accomplishment incurring heavy losses. Losses can degrade the ability to complete future missions in the case of lost equipment or friendly casualties. These types of losses can prevent the accomplishment of a subset of mission types and objectives, and can lead to a general lack of agility and battlespace awareness resulting in increased confusion [24, 25]. However,

losses with respect to the local civilian population degrade the combat effectiveness by enabling a loss in the battle of narratives, which is the battle for how the civilian population and international community perceives a situation. Additionally, with the spread of social media use, the effects of civilian population loss are not limited to the specific locale of occurrence. These effects can be distributed to multiple locations and theaters of operation.

Despite unparalleled capability, operational effectiveness in Afghanistan and Iraq has been criticized as a result of difficulties in planning [25-27]. Rates of fratricide during the Gulf War and again in the Global War on Terror are greater than previous conflicts [28, 29]. Additional failure modes associated with planning are inadequate knowledge, disconnects in planning assumptions among stakeholders, political pressures and pressures from rapidly emerging situations. Specifically, Flynn and the Joint and Coalition Operational Analysis (JCOA) synthesized findings from studies covering the spectrum of military operations between 2003 and 2012. The result was a designation of eleven themes or lessons to consider for the future where the path forward to improve planning was given for a majority of the themes [25].

1.2.3 Challenges in Joint Operational Planning

Joint operational planning processes are often characterized as data, labor and time intensive [15]. Courses of action (COAs) must be planned and executed within the enemy's decision-making cycle while allowing planning time for lower command echelons

[13, 14]. This represents a computational burden that scales dramatically with increasing numbers of systems and actors within the modern battlespace [30]. For example, Operation Dragon Strike in Afghanistan required planners to assign spatially and temporally, a coalition of U.S. and Afghani troops, and U.S. State Department personnel to operations in Arghandab, Zhari and Panjwai Districts. The coalition included troop numbers exceeding eight thousand, utilizing a combination of ground forces, assault helicopters and artillery [30, 31]. Operations like this require planning that must be done dynamically in a rapidly changing environment while accounting for uncertainties discussed previously and becomes more constrained during rapidly developing situations [14].

According the U.S. Joint Chiefs of Staff, analysis and evaluation of potential COAs is the most time consuming part of the planning process for the following reasons:

- Analysis and evaluation is often done manually [13]

- New scenarios for computer aided analysis are time consuming to create, analyze and train users [32]

Limitations of the method of analysis and evaluation of COAs are due to the following reasons:

- There is a large design space of alternatives to analyze

- Time constraints and the urgency of military operations prevent more than a few alternatives from being created

In constrained situations, only the most probable and the most dangerous enemy actions are considered with severe constraints enabling only the most dangerous enemy actions or reactions considered. In order to match the unparalleled capability of the U.S. Military, the planning timeline must keep pace with the rate of changing battlespace conditions.

1.3   Research Objective

The task of planning military operations is difficult given the scale of U.S. military capability, the spectrum of potential operations and the levels of uncertainty associated with modern warfare. By embracing technology, the U.S. military has effectively multiplied existing capabilities. However, they have also introduced complexity into the process of identifying how to employ capability. Identified in the previous section is a bottleneck in the planning process, which is analysis and evaluation of potential solutions to a military problem or COAs. This fact and the factors in the battlespace discussed previously, result in the following research objective of this Thesis:

Research Objective: Develop a methodology to aid military planners by utilizing a new process for analyzing and evaluating COA alternatives.

This methodology must provide military planners with a means to evaluate the impact of changes in the selection, sequencing, or timing of decisions and events within the battlespace on operational effectiveness as well as other measures of performance. Additional derived objectives to provide this capability include:

- To create a large number of COAs

- To assess a large number of COA alternatives without excessive computational burden

- To identify critical decision points

- To account for uncertainties

Chapter Two presents a literature review of both state of the practice and state of the art in operational planning as well as other relevant work addressing the challenges of Joint Operational Planning. This will enable a benchmark from which the proposed methodology can be contrasted in addition to the identification of gaps in capability. Chapter Three will describe the technical approach proposed to satisfy the stated research objective by defining research questions and corresponding hypotheses. Chapter Four will describe the application problem and experimental plan to test the stated hypotheses.

Chapter Five will summarize the results from experimentation. Chapter Six will present salient conclusions as well as define the contributions of this thesis and future work identified.

CHAPTER II

PROBLEM FORMULATION

In order to identify a methodology to aid in the analysis and evaluation of COA alternatives, the planning process must be characterized. Initially, generic planning formulations are investigated. Afterwards, the Joint Operational Planning Process (JOPP) is evaluated and the factors required for COA creation are determined. Next, the identification of success measures is undertaken and, methods that will aid commanders in the analysis and evaluation steps of planning are discussed and appraised. Finally, gaps in the state of the art, with respect to the research objective are identified.

2.1 Planning Paradigms

For many domains, it is desirable to construct a detailed plan prior to execution. This is facilitated by the expectation of anticipated future through the use of trend analysis of empirical data, probability functions, or through experts [33-36]. As a result, traditional decision-making is based on the creation of static plans. Under the assumption that the future can be predicted, decision makers seek to identify an optimal plan for a single most likely future. This is referred to as a static optimal plan and is visualized in Figure 3 [37]. However, in regard to the battlefield, one cannot assume to predict the future perfectly, other methods must be used.

15
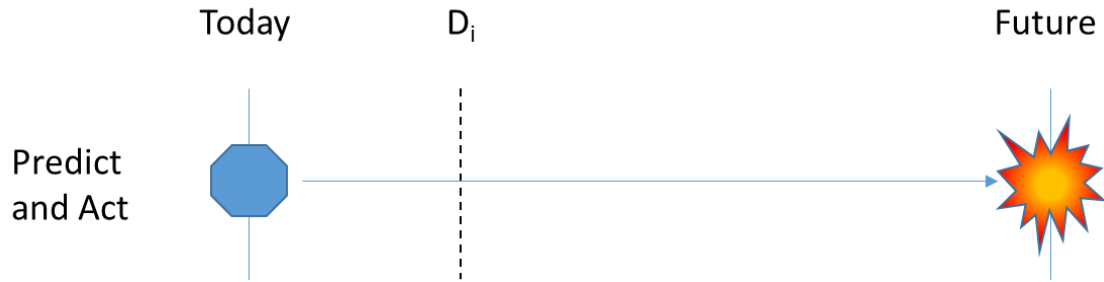
Figure 3. Static Optimal Planning and Decision Making utilizing one decision point for one possible predicted future state based on information available today [38].

Based on the level of uncertainty present in the battlefield, optimal results are not attainable. For this, a static robust plan has been proposed by several authors to achieve the best alternative based on a range of possible futures [39-42]. In these solutions, uncertainty is mitigated by identifying a subset of possible outcomes and choosing an optimal compromise across the range of possible futures. This type of planning assumes a known likelihood of possible outcomes or futures and the associated plans are robust with respect to that subset of possible outcomes [42]. This type of planning is illustrated in Figure 4. While this assumption is perfectly acceptable for a vast range of decision-making problems, the multitude of actors and systems on a Modern Battlefield and associated interactions render static plans fragile and are considered highly idealized [37].

16
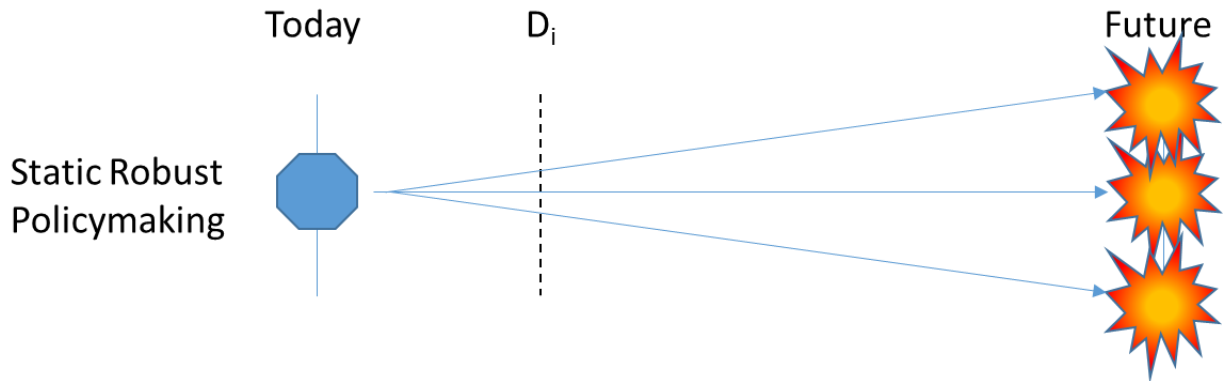
Figure 4. Static Robust Planning and Decision Making using one decision point for a range of possible predicted future states based on information available today [38].

Static plans, both optimal and robust, are vulnerable to changes in the nature of the environment [41]. Assumptions about the battlespace can be rendered obsolete due to a change in enemy situation. Changes in policy like restricted airspace, and new or altered Rules of Engagement can greatly affect the capability and battlespace and render assumptions invalid. Unexpected events, like asymmetric attacks causing a loss of resources, can have a great effect on battlefield conditions and capability reducing the set of possible actions available to the decision maker. By considering these aspects of battlespace variability, it is assumed that the future cannot be predicted.

According to Walker et al. [43], literature offers three general and overlapping methods of resolving uncertainties like the ones observed on the battlefield. These methods and their descriptions follow:

17

- Resistance is a method of hedging against uncertainty by planning for the worst case scenario

- Resilience is a method which ensures rapid recovery from the effects of uncertainty

- Adaptation is a method in which it is assumed that the plan will change based on the uncertainty, while still in pursuit of the original goals

Successful planning on the battlefield requires other methods of dealing with the inherent uncertainty rather than hedging against it as in static robust planning. Adaptive and dynamic adaptive planning approaches attempt to embrace uncertainty. Walker et al. [43] provides a state of the art review of adaptive approaches for policymaking in different domains.

Adaptive methods are separated according to two general schools of thought; adaption based on the anticipation of negative effects are one, and adaption based on the reaction to negative effects are the other [43-47]. Adaptive planning relies on an intrinsic knowledge of vulnerabilities to facilitate the anticipation of failure modes. For reactive adaptive planning, intrinsic knowledge of strengths is necessary to facilitate reaction to failure modes.
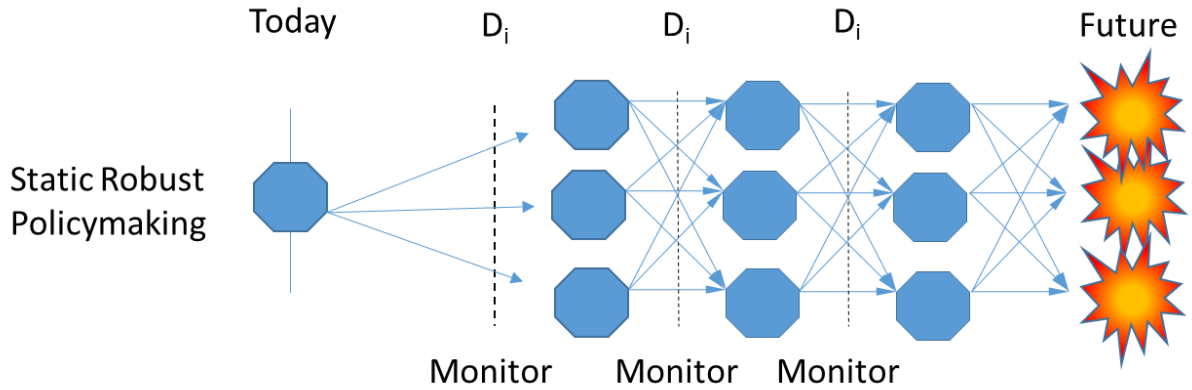
Figure 5. Dynamic Adaptive Planning and Decision Making using multiple decision points for a range of possible predicted future states using refreshed information.

Reactive or anticipatory methods require the assumption that failure modes associated with uncertainty are either not catastrophic, or knowledge is sufficient such that anticipation is possible. Neither of these assumptions are explicitly true but utilizing Dynamic Adaptive Planning leverages the strengths of both assumptions by employing both adaptive and anticipatory policies to planning [43, 48, 49]. However, the adaption of either adaption or anticipation to the modern battlefield is problematic due to the multitude of actors, systems and associated interactions between them. The problem regarding the complexity stated presents a challenge in both the anticipation of unknown events and effects, and complete performance knowledge of a Network Centric Force.

2.2   Joint Operations Planning Process

Decision making in a joint setting is done utilizing the Joint Operation Planning Process (JOPP). This is a method to initiate, develop and approve a course of action (COA) that will appropriately solve a problem; This problem is the achievement of a desired military end state and is outlined in Joint Publication 5-0 [13]. This process is illustrated below in Figure 6 and begins with Planning Initiation which is the receipt of instruction by a higher command for deployment. The next step is Mission Analysis, where the recipient is to analyze the instruction given by higher headquarters. Once these two steps are complete, a COA will be developed, analyzed, compared and approved by the commander.

Joint Operation Planning Process

Step 1	Planning Initiation
Step 2	Mission Analysis
Step 3	Course of Action (COA) Development
Step 4	COA Analysis and Wargaming
Step 5	COA Comparison
Step 6	COA Approval
Step 7	Plan or Order Development

Figure 6. The Joint Operation Planning Process beginning with planning initiation and spans the progression arriving at a plan that is disseminated to lower echelons [13].

Based on the research objective, the emphasis of the subsections will include steps three and four. The descriptions that follow regarding steps three and four will be based on the corresponding joint publication for joint operational planning [13]. This section will end with limitations and difficulties of the JOPP.

2.2.1 Course of Action Development

COA development is initiated after an understanding of the higher order is reached. This step requires inputs of staff estimates, operational approach and desired information

seen in Table 1. Staff estimates are battlespace specifics used to understand the situation and conduct a functional analysis of the COA. The operational approach contains the Commander's approach to solve the given problem. The Commander's Critical Information Requirement (CCIR) is elements of information that the commander identifies as being critical to timely decision-making. With the input of these elements, COA alternatives can be developed.

Table 1. COA development inputs.

COA Inputs

Mission Statement

Staff Estimates

Operational Approach

Critical Information

The COA alternatives must contain what is to be done during the operation, size of forces and time force capabilities must be utilized. In short, the development of a COA answers the following questions:

- What type of forces will execute the tasks

- What are the tasks

- Where will the tasks occur

- When will the tasks begin

- How the commander should provide direction to enable components to accomplish tactical actions

- Why will each force conduct their tasks

- What Intelligence, Surveillance and Reconnaissance (ISR) support is needed

The COA is the product of multiple elements that includes utilization of multiple assets, which are directed at multiple objectives, which negotiate multiple axes and include multiple operations. The objective to be satisfied by the COA can be arrived at by any combination of generated effects either direct or indirect. Figure 7 shows a notional COA sketch, where even in this simplified example, there are eighteen potential COAs. This sketch comes from the joint publication expanding on the operations process. The units,

represented in green have three avenues of attack (in blue) towards three objectives (in red), and the total combination of alternatives scales exponentially with the introduction of more systems or players [50]. Additionally, options, activities during a COA that when executed will enable achievement of the objective, must also be generated, allowing the commander to rapidly transition due to changing conditions on the battlefield. Next, the COA alternatives must be tested for validity, analyzed, wargamed and compared for final approval.
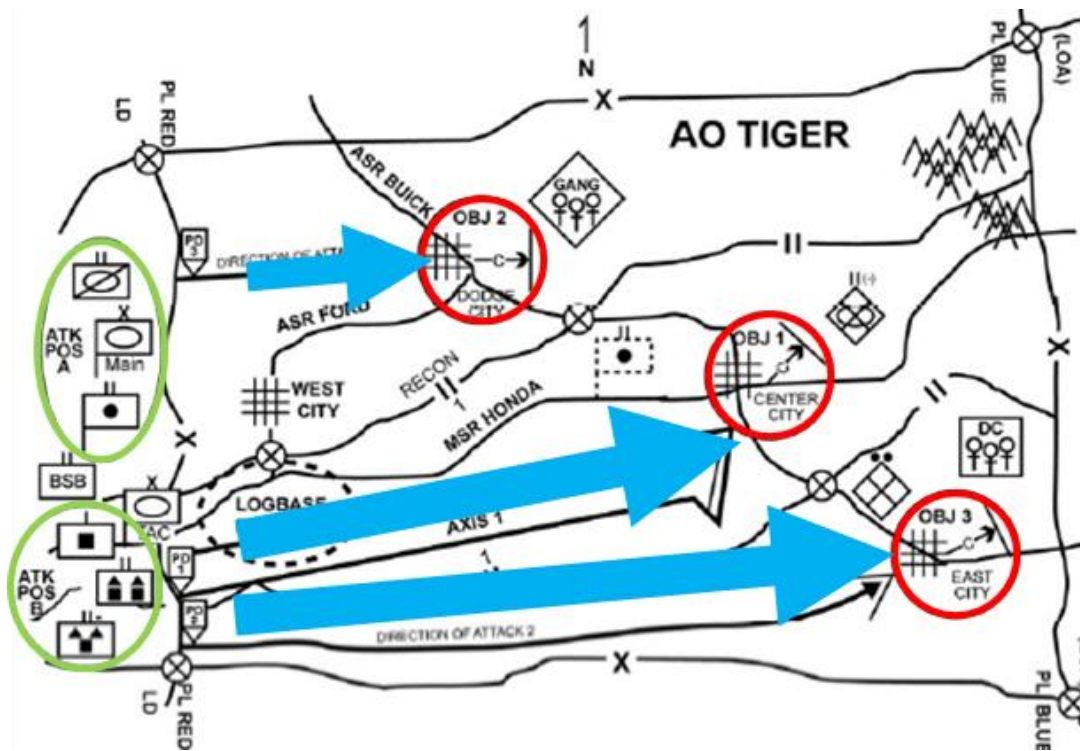
Figure 7. Notional COA sketch showing a set of units in green, which can attack three targets in red, utilizing three axis in blue yielding 18 possible attack combinations[50].

24

Validation of a COA requires assessment of suitability with respect to the requirements listed in Figure 8. Adequacy considers whether or not a COA accomplishes the mission while meeting conditions for the desired end state. Feasibility is achieved, if the COA can be executed with the forces, support and technology available. Acceptability considers compliance with Rules of Engagement (ROE) and whether or not the plan contains unacceptable risks. Distinguishability ensures that each COA is sufficiently different from the next. This is required because of the manual nature of this problem and the JOPP, where creation of similar COAs is seen as wasteful. Finally, consideration of completeness is based on incorporation of objectives, desired effects, tasks to be performed, forces required, concepts for deployment, time estimates for objective achievement and success criteria definition.
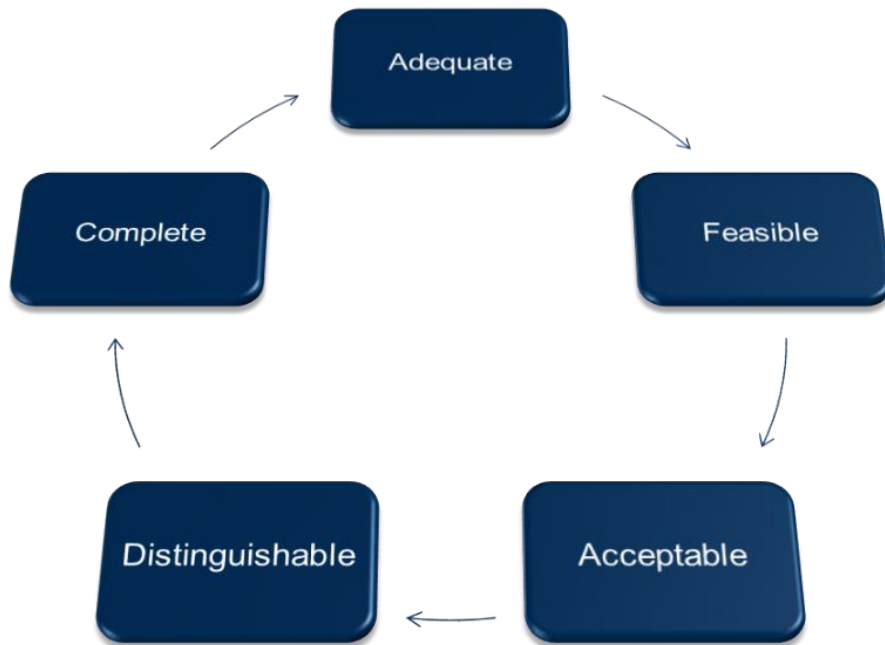
Figure 8. COA validation requirements defined in the JOPP but are not all viable when considering modifications to the JOPP or other planning processes.

2.2.2 Course of Action Wargaming and Analysis

The next step in JOPP is Step Four, COA Wargaming and Analysis. COA analysis is the process of examining each potential COA for validity and determining details of each COA for comparison, therefore the primary input is the set of COA alternatives. The COAs

are validity tested more rigorously than in the development step. Details include but are not limited to advantages, disadvantages, feasibility and acceptability. Output is a wargamed COA and Critical Event/Decision Point where Critical Events are essential and mission specific tasks and Decision Points are decisions required by the commander to ensure successful execution of resources. COAs from the development step are infused with more detail during the wargaming step, therefore, the primary output of Step Four is a set of wargamed COAs. New COAs can also be developed in Step Four as well.

The primary means of performing COA analysis is wargaming. During analysis, the commander and battle-staff attempt to visualize the flow of the operation considering all relevant aspects of the operational environment/battlespace through simulation. This facilitates a common understanding of COAs which allows for advantages and disadvantages of each COA, which will be used for comparison in a later step. Analysis tools include maps, sketches, worksheets and computer aided simulation tools. The primary methods of wargaming are Analog and Computer.

Analog wargaming is a traditional method of wargaming and has been used for centuries for training, tactics analysis and mission preparation [51]. This type of wargaming is done utilizing maps of the battlefield and pieces as abstractions of soldiers, units and equipment involved in the battle. It is still utilized by the military today. In a similar fashion, live action wargaming is conducted with "boots on the ground" running specific scenarios. These types of wargaming are much more resource intensive, as it

requires more time, logistical planning and incurs a higher cost. An example of this is Millennium Challenge in 2002. This was one of the biggest wargames of all time, taking two years to plan, including 13,000 troops and costing $250M [52].

Computer wargaming can either by computer aided or computer assisted. Computerized wargaming has roots dating back to the 1960's and has evolved much since then. Exercise Internal Look 2003 utilized CENTCOM's Command Deployable Headquarters, CDHQ, to test the command, control and communications of CENTCOM, prior to the war in Iraq. The footprint of CDHQ is the size of a football field, supporting hundreds of troops, including 100 servers and more than 400 workstations [53].

Regardless of the wargaming capability available, the commander will select COAs to be wargamed as well as wargaming guidance and evaluation criteria. COAs from the development step are infused with more detail during the wargaming step. Enemy COAs are developed considering factors that could potentially influence the achievement of the desired end state. The adversary COA includes capability and intent, and actions by other actors, both friendly and hostile. The direct output of the analysis and wargaming step is the probable effect that each opposing COA has on the friendly COA. This facilitates determination of acceptability and feasibility.

2.2.3 Limitations of the JOPP

With reduced decision cycle time resulting from a dynamic battlespace, a method to create COAs rapidly is needed. The planning process is severely time constrained. As a result, only two to three COAs are developed in the process. Additionally, all feasible adversary COAs should be considered during friendly COA development further enhancing time sensitivity and in time constrained situations, only the most probable and most dangerous adversary COAs are considered. In the most severe time constraints, only the most dangerous enemy COA is considered. In addition to time constraints, there are manpower constraints to developing COAs as well. Developing COAs in parallel requires the same personnel structure on each development team. Furthermore, by pursuing parallel development, it is likely that non-distinct COAs will be developed. As a result, sequential COA development is recommended but is severely limiting with respect to a large decision space.

Other actors in the battlefield may also make operations more difficult, as a result, the planners must also consider this as well. For example, the presence of civilians or non-combatants in or near the battlespace will complicate the operation adding another dimension to consider which can potentially limit feasible solutions that will satisfy the military objective.

According to the U.S. Joint Chiefs of Staff COA Analysis & Wargaming is the most time consuming part of the planning process [13]. Analysis is often done manually and new operational scenarios are time consuming to create. Live exercises take time to organize, coordinate, play out, and post-process. Hurdles in COA Analysis & Wargaming include:

- Large design space of alternatives to analyze

- Time constraints and urgency of military operations prevent more than a few alternatives from being created

- Emerging situations influence the rate of change of the battlespace

To match planning capability to operational capability, the planning timeline must keep pace with the rate of changing battlespace conditions.

2.3 Measuring Battlefield Success

Evaluation criteria are considered and selected during wargaming for use in the comparison step to identify the best COA. They are used to assess the relative effectiveness and efficiency. Criteria are the aspects of the situation that are critical to mission accomplishment which change from mission to mission. If they are not explicitly stated,

they are derived implicitly from the commander's intent statement. Criteria should be objective and interpretation should be constant when considering different COAs and evaluation. A potential list includes [13]:

- Adherence to commander's guidance and intent

- Mission accomplishment quantification

- Cost

- Doctrinal fundamentals with respect to the type of operation

- Residual risk

- Implicit significant factors of the operation (need for speed, security…)

- Factors relating to specific staff functions

- Elements of operational design

- Political constraints, flexibility, simplicity, surprise, speed, mass, sustainability, C2, infrastructure survivability…

This list of evaluation criteria overlaps with the validation measures seen in Figure 8, where the list of validation measures appears to be more general and easily applicable to a greater range of operational situations. Therefore, adequacy, feasibility, acceptability,

31

distinguishability and completeness are criteria used to evaluate COAs. However, a subset of these criteria will be used in this formulation. Additionally, these criteria must be measured somehow. This will be discussed in the following section.

2.3.1 Measures of Effectiveness & Performance

Success on the battlefield is measured using Measurements of Effectiveness and Measures of Performance. A Measure of Performance (MOP) is a measureable indicator that assesses friendly ability to create a desired outcome in the battlespace [54]. MOPs are used to track completed tasks and are often discrete measures of some battlefield occurrence. An example of an MOP in a combat scenario is number of enemy aircraft neutralized. This example is a discrete measure that generates a desired outcome. A Measure of Effectiveness (MOE) is a measure used to assess changes in the battlespace, with respect to a desired end state [54]. An example MOE is a reduction in enemy air capability. This measure can be influenced by a number of MOPs. For example, the attainment of a reduction of enemy air capability can be had by neutralizing enemy aircraft, as in the MOP example above, or by destroying a runway.

MOEs are calculated using battlespace measures (environment), while MOPs are a list of completed tasks. According to the Joint Staff, "…trying to build a linkage between MOP and MOE is a proven waste of time for staffs." [54]. However, sequencing of tasks (MOP) affects the MOE. While evaluating the military outcome and battlespace

parameters is necessary, doing so in isolation is not sufficient to define the degree of success [55, 56].

Perry et al. highlight the inadequacy of traditional MOEs and MOPs to account for the effects of information and decision-making in regard to NCW [57]. Through the use of two vignettes, an MOE and MOP were defined for both. The MOE is survivability while the MOP is a series of equations that measure knowledge for the first vignette which is protection against a ballistic missile attack. For the second vignette, the acquisition of a target, the MOE is the probability of kill (P_{kill}) and the MOP is time on target which is time that the defending asset has to acquire the target [58].

The effort to adapt current MOEs and MOPs by Perry et al. is derived utilizing two predefined vignettes making them idealized [57]. Additionally, this method is based on the utilization of tuning parameters, where prior knowledge of the parameter values would remove the necessity of the model [57, 59] and the parameters themselves would be utilized to facilitate the measuring of performance and effectiveness.

2.3.2 Contrasting the Results

There are various decision support tools available to military planners that aid in the decision-making process. A Synchronization Matrix is a decision-making tool and a method to record the results of wargaming. With it comes the organization of decision points, potential evaluation criteria, CCIRs, COA adjustments and branches and sequels.

It helps the staff synchronize the COA across time and space with respect to adversary COAs and helps identify cross-component support resource requirements. A completed Synchronization Matrix yields the basic structure of a COA and enables orders development [13, 60].

With the completion of the Synchronization Matrix, each COA is wargamed with the results and action, reaction and friendly counteraction documented in a COA Wargame Worksheet. This, in addition to the Synchronization Matrix, enables the completion of the COA Comparison and Decision Matrix. In this matrix the weightings for each of the commander's evaluation criteria are documented. In addition, the advantages and disadvantages are documented as well, allowing the weightings of evaluation criteria, advantages and disadvantages to identify an advantageous COA to pursue [13, 60].

2.4   *Methods of Wargaming and Analysis*

As stated previously in section 2.2.2, analog wargaming is the oldest method of wargaming [51]. Whether it is using maps and acetate overlays or by conducting live exercises, these types of wargaming become intractable when the number of players and associated interactions increase as they do in the modern battlefield. Computer wargaming, on the other hand, scales well to the size and density of the modern battlefield but still includes limitations.

34

2.4.1 Deterministic Wargames

The approach to computer wargaming is varied with many implicit assumptions. As a way to leverage computational power, deterministic models assume perfect knowledge on the battlefield. The foundation on which deterministic computer wargaming models are built are Lanchester's equations.

Lanchester's equations describe attrition of a force-on-force engagement. Each side degrades the other at a rate proportional to the remaining size multiplied by the firing rate [59]. This provides the basis for modern, deterministic air and ground combat simulations. An example of this is Tactical Warfare Model (TACWAR), which utilizes Lanchester's equations for its attrition model. It provides a computerized deterministic representation of tactical warfare at the theater level and allows the measures of kill ratios, killer-victim scores and Forward Line of Own Troops movement rates to include controlled area [61].

Models like TACWAR and others using Lanchester's equations require an equation for every interaction between factors in the process that are modeled [59]. As the complexity of the model increases, by introducing more actors and equipment and engagements, the ability to model interactions becomes limited due to the combinatorial nature of military planning. The aggregation of effects assumption, implied when using deterministic models, may ignore local behavior effects [59]. Additionally, the implied

35

regularity assumption states that small changes in inputs will have proportional changes in the outputs which may not be true in a complex system.

2.4.2 Stochastic Wargames

Stochastic Wargames are built on stochastic models, which are similar to deterministic models but have one or more deterministic equations replaced with formulations that include some sort of uncertainty [59].

Naval Simulation System (NSS) is a computer-based framework for Command, Control, Computers, Communications, Intelligence, Surveillance, and Reconnaissance (C4ISR) centric multi-warfare simulation of naval theatre operations [59, 62, 63]. In this framework, combat is represented as a chain of independent events or as sets of basic interaction equations [59]. Additionally, it includes agent based and discrete event simulation models and has the ability to receive information which makes it useful for analyzing ongoing operations course of actions [62].

OneSAF is another example model which is both entity based and stochastic [59]. It is based on physical attrition models which favor number of assets as opposed to arrangement of assets [59]. According to Cares, "attrition obscures the importance of sensing or the sequencing of attacks" [59]. Therefore, stochastic force-on-force wargames lack the ability to model element interactions at the troop level with respect to attrition and

36

the effects on the network from losing nodes cannot be identified without entity level simulation.

Similar to the deterministic wargame, stochastic wargames have implicit assumptions of perfect knowledge, aggregation of effects and implied regularity but extend further. An additional assumption is that, due to uncertainty of input variables, it is acceptable to utilize random variables. Furthermore, interactions between variables is assumed to be uncertain. Therefore, the variables are also assumed to be independent random variables. This negates the existence of causality in the models [59].

2.4.3 Limitations of Conventional Wargaming

While facilitating a greater study of the decision space, conventional wargaming carries limitations that prevent efficient use of capability due to planning. Both deterministic and stochastic models assume perfect information is available to the commander, which is not true given the number of elements available to a commander and their associated dispersion on the battlefield. Furthermore, the requirement for an equation for every interaction in the process being modeled makes thorough analysis cumbersome [59]. As the complexity of the model increases, the ability to model interactions becomes limited due to the combinatorial nature of military planning. The aggregation of effects assumption, implied when using may additionally ignore local behavior effects [59]. The implied regularity assumption states that small changes in inputs will have proportional

changes in the outputs which may not be true in a complex system. Stochastic wargames have implicit assumptions that enable utilization of random variables which negate the existence of causality in the models [59]. Thus, utilization of conventional wargaming methods prevent the understanding of interactions among the many battlespace actors. This ultimately prevents the understanding of causality on the battlefield which exposes the friendly force to unknown dangers while preventing the friendly force from seizing causality associated opportunities.

2.5   Gaps in the Literature

General decision methods are either reactive or anticipatory methods and require the assumption that failure modes associated with uncertainty are either not catastrophic, or knowledge is sufficient such that anticipation is possible. Neither of these assumptions are explicitly true: However, utilizing Dynamic Adaptive Planning leverages the strengths of both assumptions by employing both adaptive and anticipatory policies to planning. However, adapting either anticipation or reactivity presents a problem due to the multitude of actors, systems and associated interactions between them present on the modern battlefield. It reduces the ability to anticipate unknown events and effects, and to have complete performance knowledge of a Network Centric Force on a battlefield exhibiting complexity.

The reduced decision cycle time resulting from a dynamic battlespace results in a severely constrained planning process. As a result, only two to three COAs are developed in the process and in the most severe time constraints, only the most dangerous enemy COA is considered. Additionally, there are manpower limitations to developing COAs. Parallel COA development requires the same personnel make up on each development team and further increases the likelihood that non-distinct COAs will be developed. Therefore, sequential COA development is recommended but is severely limiting with respect to a large decision space. Complexity in the form of entity type adds to the difficulty of planning. Decision makers must not only consider friendly and adversary elements but also non-combatants, non-government organizations and non-state adversaries. This complicates the operation adding another dimension to consider which can potentially limit feasible solutions that will satisfy the military objective but will complicate the sequence of operations.

Evaluation of the myriad of potential situations that can be encountered on the modern battlefield is problematic as well. Utilization of conventional wargaming methods prevent the understanding of interactions among the many battlespace actors. This ultimately prevents the understanding of causality on the battlefield which exposes the friendly force to unknown dangers while preventing the friendly force from seizing causality associated opportunities.

Utilizing MOPs and MOEs is helpful to identify adequate performance of a mission toward a stated goal. While evaluating the military outcome and battlespace parameters is necessary, doing so in isolation, which is the prescribed method, is not sufficient to define the degree of success. Analysis of event sequencing is necessary for a more holistic review of the battlefield and situation. Additionally, the Synchronization Matrix, COA Wargame Worksheet and the COA Comparison and Decision Matrix weightings for each of the commander's evaluation criteria facilitate a highly subjective analysis of COAs using MOEs and MOPs which may be constraining in the analysis of new and developing situations.

Wargaming and analysis is often done manually preventing rapid evaluation. One reason is that the design space of alternatives is extremely vast. Adding to the burden, time constraints and urgency of military operations prevent more than a few alternatives from being created. To match planning capability to operational capability, the planning timeline must keep pace with the rate of changing battlespace conditions.

CHAPTER III

METHODOLOGY FORMULATION

In Chapter Two generic decision-making paradigms as well as the JOPP were discussed. In doing so, the benchmark process was defined. Utilizing this benchmark and maintaining the research objective, this chapter will serve to identify areas requiring resolution to meet the stated research objective of developing a methodology to aid military planners by utilizing a new process for analyzing and evaluating COA alternatives while facing the challenges associated with NCW and the modern battlefield. Such resolution will be arrived at by identification of research questions that require resolution as well as a combination of responses and hypothesis to research questions.

In the first section, a general methodology will be defined that will enable the development of a methodology aiding the analysis and evaluation of COAs. The remainder of the chapter will complete the steps in this general methodology and will culminate in the definition of the specific methodology by chapters' end.

3.1   General Methodology for Formulation

The stated objective requires the use of a decision-making methodology for accomplishment. In pursuit of creating a specialized methodology to aid in COA development and analysis, a more general decision-making methodology is sought to

augment the development. While the proposed methodology will aid decision-making in a Net-Centric battlefield, it will not serve as a decision-making methodology in and of itself. Through review of general decision-making methods, the Generic Top-down Decision Making Process by Mavris et al. developed for Integrated Product/Process Development (IPPD) [64]. The steps shown on the left side of Figure 9 will be adhered to and provide structure to the development of the methodology to allow completion of the highlighted steps of the JOPP on the right hand side of the figure as well.

The first two steps in the general process have been completed in Chapters One and Two. The next three steps will be expanded upon in the following sections and will aid in the development, analysis, wargaming and comparison of COAs. The next section of this chapter will continue the Generic Top-down Decision Making Process.
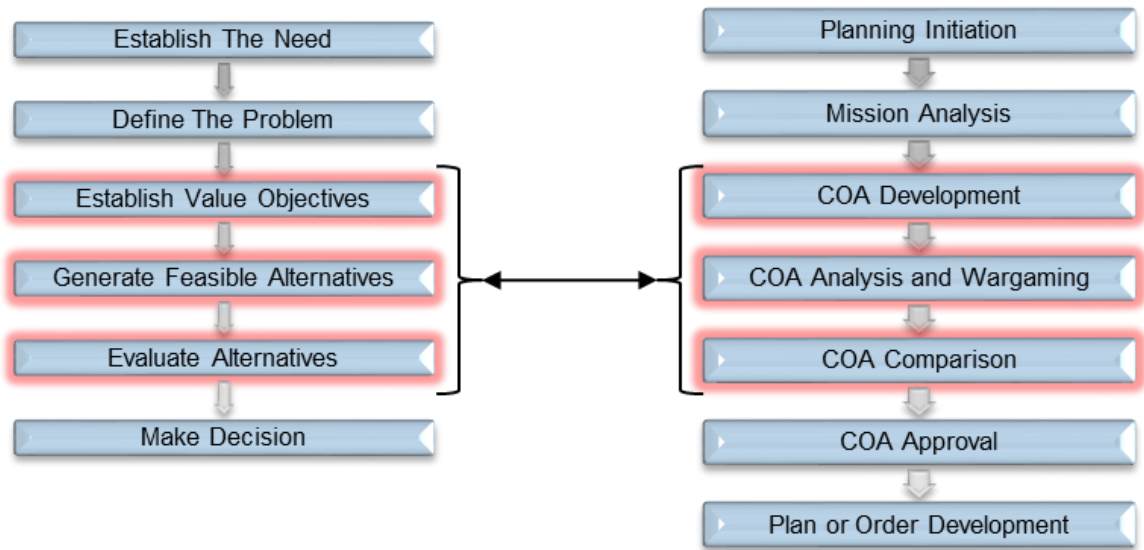
Figure 9.   Generic Top-down Decision Making Process and likeness to the JOPP highlighting which JOPP steps will be modified and how [13, 64].

42

3.2 Establishing Value Objectives

Today's fast-paced battlefield is influenced by improvements in air power, imagery, electronic and information technology warfare have magnified effectiveness of combat power elements [7-9]. Specifically, these improvements have improved the level of battlespace awareness and maneuvering capability, effectively increasing the information flow to make decisions and capability to act accordingly [8-10]. The proliferation of networked systems and C4I technology and adaption of Network Centric Operations further enables capability but increases complexity, which is the diversity and number of components, interactions and resulting behaviors [8, 10, 11]. Based on the level of uncertainty present in the battlefield and the complexity associated difficulties brought on by modern technology, factors that neutralize these effects must be defined.

Operational freedom enabled by NCW is a double edge sword. If too much operational flexibility, enabled by NCW, is delegated to subordinate units, the commander loses control. However, with too little operational flexibility, timely action can be inhibited [9]. As a result, flexibility alone is not sufficient to check the liability of the modern battlefield. The realization of the enhancements offered by NCW must be aided by agility in command, decision-making and action in order to minimize the realization of difficulties brought on by the use of C4I technology in an uncertain battlespace [3].

Generally, when speaking of agility, physical agility is considered. Physical agility describes the capability of an asset to maneuver quickly. However, the Joint Chiefs describe agility as the ability to adapt to changing situations [1]. Weill utilizes a definition of Strategic Agility which relies on the capability of the organization [65]. Dekker defines elements of agility that affect different aspects of the mission to include tactical, organizational, deployment, sustainment, acquisition and conceptual agility with associated metrics [66]. However, Moffat describes characteristics of agility that can be applied generally to each type of mission which is preferred to the situational definitions offered [3].

In Figure 10 the characteristics of Agility are displayed. Utilizing Moffat's definitions, the characteristics follow [3]:

- Robustness is the ability to maintain effectiveness across a wide range of operating conditions

- Resilience is the ability to recover from an encounter with a destabilizing force

- Responsiveness is the ability to react to changes in the environment quickly

- Flexibility refers to the ability to maintain a multitude of options to accomplish a task and to transition easily from one option to another

- Innovation is the ability to find new approaches to both old and new problems

44

- Adaptation refers to the ability to change the organization and inherent processes of an organization

Figure 10. Characteristics of organizational and strategic agility required to hedge the influences of a network-centric operational environment.

These characteristics are overlapping ways to limit the effects of a Net-Centric operating environment and are additional to the three methods of resolving uncertainty on the battlefield [43].

Based on the level of uncertainty present in the battlefield, strategies that foster agility must be utilized. Static plans, discussed in Chapter One, achieve the best alternative based on a most likely or a range of possible futures [39-42]. In these solutions, uncertainty is mitigated by identifying a subset of possible outcomes and choosing an optimal compromise across the range of possible futures in the case of robust plans. As seen in Chapter Two, this process fails when time is so constrained that only the worst case scenario is planned for. This case is representative of a resistance strategy. Resistance is a method of hedging against uncertainty by planning for the worst case scenario.

With these assumptions, static plans are shown to lack agility. In the best case, they can be robust; but in the worst case, they adhere to a resistive strategy. The multitude of actors and systems on a Modern Battlefield and associated interactions render static plans fragile and are considered highly idealized [37].

Dynamic Adaptive methods based on adaption motivated by the anticipation of negative effects and are adaption based on the reaction to negative effects, which offers increased agility over static or adaptive plans [43-47]. Anticipatory adaptive planning requires an intrinsic knowledge of vulnerabilities in order to facilitate anticipation of failure

modes. Therefore, the state of the system in question is monitored for specific preconceived vulnerabilities. When these vulnerabilities are encountered, they are resolved using a predetermined plan.

For reactive adaptive planning, intrinsic knowledge of strengths and how the strengths are changing with time, is necessary to facilitate reaction to failure modes. This indicates that the reaction to a stimulus (an identified failure mode) is not preconceived and is a function of time and state. Additionally, it is expected in this paradigm that both the stimulus and the reaction to regain stability will not be predictable.

Reactive or anticipatory methods require the assumption that failure modes associated with uncertainty are either not catastrophic, or knowledge is sufficient such that anticipation is possible. In modern military operations, these assumptions are not valid. Neither of these assumptions are explicitly true in modern military operations but utilizing Dynamic Adaptive Planning (DAP) leverages the strengths of both assumptions by employing both adaptive and anticipatory policies to planning [43, 48, 49]. Applying DAP and JOPP to modern problems produces Robustness, Responsiveness and Adaptation but not Resilience, Flexibility and Innovation. These characteristics will be resolved in other ways.

DAP enables attainment of the previously discussed characteristics, however, adapting these schools of thought to the modern battlefield is problematic due to the

multitude of actors, systems and associated interactions between them. It presents challenges in both the anticipation of unknown events and effects, and complete performance knowledge of a Network Centric Force. Action motivated by anticipation and/or adaption must be stimulated by the realization of changing conditions. This realization can be arrived at by the recognizing of indicators in the battlespace. These challenges motivate the following research question:

Research Question 1.1: What indicators will allow battle commanders to identify situations that require action through anticipation and/or adaption?

As discussed in Chapter Two, success on the battlefield is measured using Measurements of Effectiveness and Measures of Performance. MOEs are used to assess changes in the battlespace, with respect to a desired end state [54]. This is a measure of the response of the environment and is therefore necessary for both anticipation and adaption needed for agility. A Measure of Performance (MOP) is a measure that assesses friendly ability to create a desired effect in the battlespace [54]. MOPs are used to track completed tasks and are often binary in value.

While the utilization of MOEs and MOPs is limiting towards developing a holistic degree of battlefield success, it is necessary and must be updated to reflect changes in the battlefield as a result of NCW. Adapting current MOEs and MOPs has been attempted by Perry et al, however, the measures defined are derived utilizing two predefined vignettes

which specifically embody NCW and NCO [57]. Additionally, this method is based on the utilization of tuning parameters, where prior knowledge of the parameter values would remove the necessity of the model [57, 59].

Through discussion of MOEs and MOPs, their importance and necessity is signified. It is of paramount importance to measure effectiveness and performance. However, the objective of deciphering situations requiring action cannot be met through these measures alone. According to Klein et al. and the U.S. Joint Chiefs of Staff, this can be facilitated by utilizing Critical Decision Points (CDP), which are points in time where alternative COAs are available yielding different probabilities of success[54, 67]. Observations of MOEs, MOPs and CDPs motivate the following research question:

Research Question 1.2: What battlespace factors must be analyzed to allow identification of Critical Decision Points?

Utilization of situation specific MOEs and MOPs is suggestive of a Balanced Scorecard (BSC) approach to measuring success and has proven to be limiting in today's dynamic battlespace [68]. A BSC approach was defined by the financial industry in order to facilitate analysis on the profitability of an institution. It calls for the analysis of traditional financial factors (profit) and also non-traditional non-financial factors (customer satisfaction) to develop an understanding on how financial factors are affected. The tabulation of all factors makes up the balanced scorecard.

According to O'Donnell, a BSC approach relies on developing the MOEs and MOPs based on a finite set of perspectives, which prevents seizing the potential of a dynamic environment [68]. Using each factor type as an example, the MOPs are the financial factors and are directly measureable while MOEs are the non-financial factors and are aggregated and deduced from other sources or other aspects of the simulation. The conclusion of his work applied to joint operational planning is to utilize systemic knowledge which focuses not only on the metrics associated with outcomes, but also data associated with all development stages. Similarly Portfolio-Analysis tools by Davis et al., all rely on similar score card methods, however these do not allow the identification of favorable sequencing and critical decision points [69]. Furthermore, Agoglia notes that traditional use of MOEs and MOPs measure outputs, while measurements of outcomes could be more telling of the progression of a military effort[70]. In identifying CDPs linked to the progression of success or failure of an operation or conflict, the near term outcomes must be measured continuously to offer a clear indication of the profitability of a sequence of actions.

Battle planning is a resource allocation problem where the commander's resources include personnel and equipment. Through utilization of the resources, some sort of capability is offered. Capability is traditionally expressed in terms of MOPs and MOEs. Additionally, information is transferred to the Commander from the battlefield in the following ways:

- Communications Systems E.g. Radio and SIPRNet

- Sensors e.g. RADAR, UAS and Satellite

Utilizing these resources, the commander attempts to increase/sustain the level of Situational Awareness. The planning process results in an assignment of personnel and equipment to targets. The Commander seeks advantage by utilizing Position to magnify combat power. Using the observations presented here on the specifics of battle planning, the following hypothesis were formulated:

Hypothesis 1.1: Utilizing Capability, Awareness and Position metrics in addition to MOEs to assess the battlespace will allow the enhanced identification of COAs as compared to traditional Capability metrics and measured by the identification conditions signifying a critical decision point.

This hypothesis will resolve both research questions 1.1 and 1.2. Utilization of Capability, Awareness and Position (CAP), identification of CDPs is possible, which fosters both anticipation and adaptation enabling agility.

3.3   Generating Feasible Alternatives

With a hypothesized way to measure value, the fitness of plans related to the research objective requires the generation of alternatives. This will facilitate the retrieval of data that can be used to determine agility. This mirrors Step Three in the JOPP in Figure

6 and is necessary for the following wargaming step to be completed. This necessity motivates the next research question:

Research Question 1.3: What method or methods should be utilized to generate a sufficient number of COAs that will facilitate analysis and generation of CDPs?

Based on Joint Operation Publication guidance, COAs are generated using force ratios, troop-to-task analysis and Universal Joint Task List (UJTL) [13, 71]. A commander will review the mission and apply the UJTL to identify tasks that will accomplish the mission. Next, the commander will perform troop-to-task analysis and force ratios in order to ensure adequate resources are allocated to each task. This procedure relies heavily on knowledge of the UJTL and relevant experience to arrive at a COA. Furthermore, the output of this process is intended only to be a first round approximation needing further analysis. This exemplifies the need for more efficient COA Development, Wargaming and Analysis aided by computers.

Schlabach et al. utilize a similar process of COA Development called Fox Ga. This process leverages computational efficiency utilizing a genetic algorithm to generate COAs for a fixed scenario with a mechanized infantry brigade [72]. Each COA and relevant situational factors are codified into a chromosome, creating an initial population of feasible solutions. Gonzales utilizes an air battle planning module in his Contingency Theater Automated Planning System (CTAPS), where a human planner selects from six types of

aircraft missions and is then aided in assigning details to the plan and call signs, IFF codes and tanker aircraft are assigned to a mission [73].

An application used in Game Artificial Intelligence, is surveyed by Browne et.al [74]. Here, Monte Carlo Tree Search (MCTS) is utilized to find optimal decisions in games such as Go and Chess. This method identifies these decisions by performing random sampling of the decision space, and then building a search tree with decision points and associated outcomes of streams of decisions.

Through crossover and mutation using Fox GA, a feasible plan can be invalidated. Additionally, different feasible plans can include different sequences and lengths of sequences. Utilizing crossover and mutation operations will drive the sequences towards a similar set and/or length which may not be ideal. Furthermore, the strength of these methods rely on the creation of an objective function which is based on the preferential rankings of MOEs and MOPs, which will not aid in CDP identification. Therefore, a Genetic Algorithm formulation of analysis is not preferred for this problem. However, the utilization of a sort of random sampling in all of the methods is appealing as seen in MCTS. In the scenario where there is trivial cost of performing simulation, random sampling of the decision space will offer the identification of decision points as well as common sequences yielding improved probabilities of success.

Response to Research Question 1.3: While it is noted that in each of the above methods, there is an issue with plan sufficiency, it is determined that performing allocation analysis utilizing a random sampling of the decision space will yield conditions signifying CDPs and associated success probabilities.

3.4   Evaluating Alternatives

Using the financial industry as an analogy, an operational planner would seek a portfolio of options to hedge against risk due to market uncertainty. Analogously, conventional planning guidance has planners choose a portfolio of options that will accomplish a task that hedges against battlefield uncertainty. However, the high frequency trader seeks out market efficiency measures to determine buy or sell points to generate advantage from volatility at small time sequences, capitalizing vs. hedging against uncertainty [75]. Utilizing the high frequency trading paradigm as motivation, this section will expand on methods to wargame and create data that can be used to generate advantage from volatility at smaller time sequences.

3.4.1 Wargaming

Since World War II and the growth of Operations Research (OR), methods of discovery to aid decision-making and general performance have been sought after. The search for heuristics that provide "rules of thumb" or guidelines to assist the decision maker can facilitate the ease of which this performance can be gained [76]. One of the earliest

examples of this was taught by Sun Tzu, where the attainment of the high ground was deemed preferential. More modern examples of military OR analyses in use today aid commanders in deployment of forces. For example, the Joint Publications set force ratios for deployment specific to the expectation of an insurgency or conventional force [77, 78]. This thesis aims to continue the pursuit of battlefield heuristics to aid the decision maker. The desire to establish cause and effect relationships on the battlefield motivates the desire for data to analyze. Due to the scale of modern engagements, physical wargaming to achieve this data is intractable. Additionally, this method will not reveal the nuances present in NCO. Therefore, computer wargaming through M&S will aid in this retrieval of wargame data to facilitate analysis.

Manual and analog wargames have been used for centuries for training, tactics analysis and mission preparation. These wargames utilize maps of the battlefield and pieces representing soldiers/units in battle. It is the most flexible method to wargame but offers the least information by way of element interactions [51]. Live action wargaming is conducted with "boots on the ground" running specific scenarios. For example, Millennium Challenge 2002 was one of the biggest wargames of all time, taking two years to plan, including 13,000 troops and costing $250M [79].

Digital wargaming has roots dating back to the 1960's and has evolved much since then. Although, the scale and cost of some digital wargaming efforts is large, the investment can be preferable to conducting full scale live action wargaming. For example,

Exercise Internal Look 2003 utilized CENTCOM's Command Deployable Headquarters (CDHQ), to test the command, control and communications of CENTCOM, prior to the war in Iraq [80]. It spanned the size of a football field, supporting hundreds of troops, including 100 servers and more than 400 workstations. However, once the initial investment is made, many COAs can be evaluated with little effort. Therefore, the criteria for COA development defined in the previous paragraph may be relaxed. Specifically, there is no need to ensure that each COA generated is distinguishable from others because digital wargaming makes the evaluation of each COA trivial with respect to manually performing this step as in the current paradigm.

Other computerized wargaming models include general deterministic and stochastic models. Lanchester's equations describe attrition of a force-on-force engagement, where each side degrades the other at a rate proportional to the remaining size multiplied by the firing rate [59]. These equations provide basis for modern, deterministic air and ground combat simulations. For example, Tactical Warfare Model (TACWAR), utilizes Lanchester's equations for its attrition model. It provides a computerized deterministic representation of tactical warfare at the theater level [61]. This allows the measures of kill ratios, killer-victim scores and Forward Line of Own Troops movement rates to include controlled area. Deterministic models require an equation for every interaction in the process being modeled [59]. As the complexity of the model increases, the ability to model interactions becomes limited due to the combinatorial nature of military

planning. Specifically, the aggregation of effects assumption may ignore local behavior effects and the regularity assumption states that small changes in inputs will have proportional changes in the outputs [59].

Stochastic models used for computer wargaming include Naval Simulation System (NSS)and others. NSS is a computer-based framework for C4ISR-centric multi-warfare simulation of naval theatre operations [59, 62, 81]. It represents combat as a chain of independent events or as sets of basic interaction equations [81]. Other stochastic large scale force implementations include EADSIM and JWARS [82, 83]. While pursuing a truer depiction of the battlefield by including uncertainty, these methods still favor physical attrition models which favor number of assets as opposed to arrangement of assets [81]. According to Brutzman et al., "Attrition obscures the importance of sensing or the sequencing of attacks" [81]. Therefore, stochastic force-on-force wargames lack the ability to model element interactions at the troop level with respect to attrition.

As discussed in the Response to Research Question 1.3: While it is noted that in each of the above methods, there is an issue with plan sufficiency, it is determined that performing allocation analysis utilizing UJTL will be utilized to perform assignment of troops to task. However, this process must be matched to the method of wargaming and the stated goals. While the methods discussed above allow enhanced analysis over large scale computer and live wargaming, they lack the ability to investigate element interactions present in a modern battlefield.

57

3.4.1.1 Agent Based Wargames

Due to the complexity of actors, systems and the interactions in the battlespace, in addition to the adaptiveness of the adversary, the modern battlefield is represented well as a complex adaptive system (CAS)[84]. Cil et al. characterize CAS according to the non-linear interactions of the elements within the battlespace in addition to the following characteristics[84]:

- Aggregation, allowing groups to form

- Flows, allowing the transfer of information and resources

- Diversity, allowing a variation of behavior in agent behavior

    Moffat adds the following to the description of CAS [3]:

- Decentralized control where behavior is emergent through the effects of local interaction and co-evolution

- Self organization of a system to approach a steady state

- Non-equilibrium order potentially preventing stasis of system resources

- Adaptation of behavior

- Collectivist dynamics that influence the element behaviors

As Moffat notes, a military force operating in a modern battlespace must endure the aspects of a CAS. Specifically, forces on a modern battlefield are composed of many heterogeneous systems and agents which interact non-linearly. The effects of NCO aim to compress the command structure affording lower command elements more decision-making authority effectively decentralizing control. NCO also effects self-organization and adaption through a continual feedback between agents in the command structure. These effects are motivated through the nature of NCO where decision rights, information dissemination and interactions are altered by the nature of the information age [3].

Agent based models (ABM) have been utilized to study CAS and are well suited due to the adaptability of ABM's. By attempting to simulate a real world system using adaptive agents, an artificial world is created allowing real world systems, made up of elements and real agents to be approximated by autonomous agents and modeled elements [84]. Through the use of both a top-down and bottom-up modeling strategy, we can capture the effects of element interactions from the bottom up and facilitate knowledge discovery from the top down [85]. However, there is still a problem with identifying how to apply this type of modeling with respect to the problem at hand, which motivates the next research question.

Research Question 1.4: In what way must alternatives be generated and passed given the selection of Agent Based Wargaming?

Response to Research Question 1.4: The utilization of UJTL will be done randomly choosing from a feasible subset of tasks that make up a complete COA. By randomly selecting from the UJTL for the first element of a COA, a feasible subset of the entire task list can be defined for the second element. This process continues until all required elements of the COA are defined. Therefore, for the required number of elements in a potential COA, an element is selected randomly from the feasible subset of the UJTL. An example of how this is implemented is shown in Figure 11. In this case, three units can take three axes to three targets. In this implementation, there are 27 combinations of COAs that can be generated. Two examples of these are also shown in the figure. They comprise the assignment of Unit, Axis and Target so that each target is assigned. Ensuring each target is assigned is the criteria, in this example, ensuring completeness of the plan.

COA Matrix of Alternatives		
Units		
Unit 1	Unit 2	Unit3
Axis		
Axis 1	Axis 2	Axis 3
Targets		
Target 1	Target 2	Target 3

$$COA_1 \begin{cases} unit_1, axis_1, target_1 \\ unit_2, axis_2, target_2 \\ unit_3, axis_3, target_3 \end{cases}$$

$$COA_2 \begin{cases} unit_1, axis_2, target_2 \\ unit_2, axis_1, target_1 \\ unit_3, axis_3, target_3 \end{cases}$$
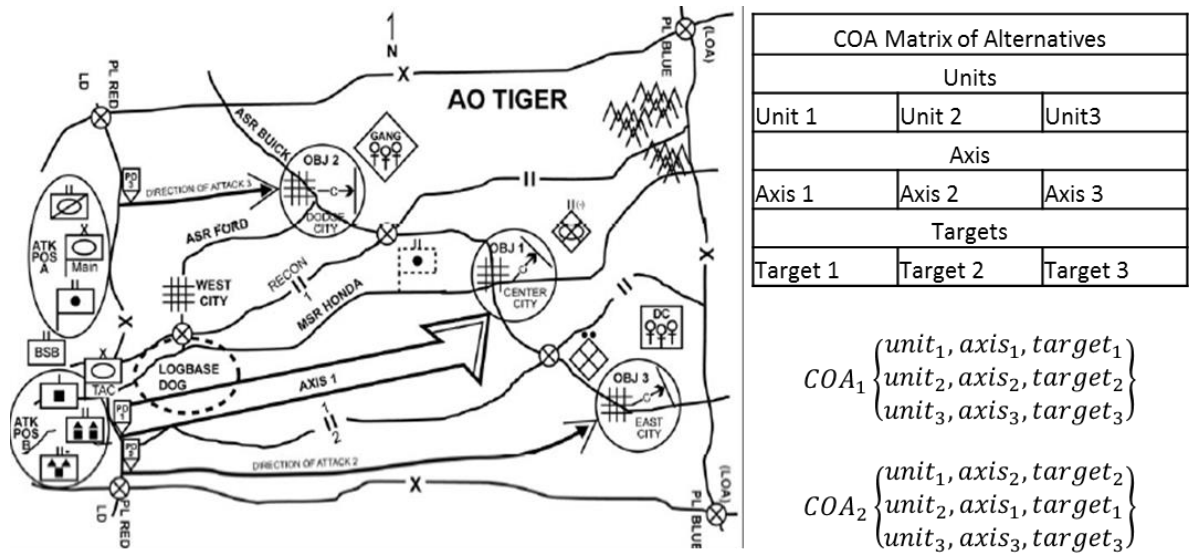
Figure 11. COA generation example using a notional area of operations and a COA matrix of alternatives showing two potential COAs out of all possible combinatorics.

3.4.2 Analysis of Alternatives

According to Davis et al., exploratory analysis over sensitivity analysis is essential for pursuing strategies that leverage agile operations, specifically with respect to flexibility, adaption and robustness [69]. This is approached via Pareto analysis, where frontier and near frontier solutions, based on score-card are considered [69, 86].

PAT-MD by Dreyer et al. uses inputs of various options e.g. particular interceptors, radars, and battle-management systems to be used for missile defense, as well as capability and risk investments, cost data and control parameters to obtain rankings, effectiveness and

cost comparisons [86]. This method maps traditional metrics to outcomes to generate an overall score.

While these methods extend the applicability of traditional MOEs and MOPs, they still cannot account for the effects of information and decision-making in regard to NCW. This realization leads to the next Hypothesis:

Hypothesis 1.2: Analyzing CAP sequence data in combination with MOEs and MOPs will allow the identification of CDPs

Rather than treatment of MOEs and MOPs alone, this thesis, instead, seeks to utilize the High-Frequency Trader's philosophy, by analyzing the elements of CAP in Performance. The ends to which this philosophy will be utilized highlights the overarching Hypothesis of this thesis:

Hypothesis 1.0: Utilizing simulation data, battle planners can extract COAs that offer an increased probability of success over the average of many simulations

By utilizing Hypothesis 1.1 and 1.2, planners will be presented with a stream of data that adequately depicts the battlefield, enabling goal oriented planning. By first utilizing the metrics to assess the battlespace, specifically enabling analysis of the time series data; CDPs can be defined giving way to a selection of COA that outperforms other COAs in the analysis pool. However, this brings about new factors as artifacts of the M&S

methods which require reconciliation. These artifacts include the form and volume of the incoming data streams.

Simulation of combat operations using traditional MOPs, MOEs and CAP generates multidimensional data streams for each combination of adversary and friendly COAs. As more scenarios are considered, the volume of CAP data increases, accumulation of CAP data creates a temporal database. Different scenarios will have data streams of differing lengths. Furthermore, these data streams are of non-uniform length and sequential/temporal in nature since the method of wargaming uses agents to step through the engagement in time-steps. For example, two identical friendly COAs wargamed against two different enemy COAs can have different outcomes and last different lengths of time, resulting in two non-uniform data streams. As a summary, the characteristics of the data are listed in Table 2.

Table 2.   Data characteristics of simulation state data providing complexity and impracticality when utilizing traditional methods of data analysis.

Data Characteristics
Multidimensional
Sequential/Temporal
Non-Uniform Stream Length

Considering these data characteristics, a method of analysis must be selected that allows consideration of CDP. Furthermore, it is the objective of this thesis to distill CDPs through analysis of CAP data. Therefore, methods in Data Mining will be investigated to reach that end and motivate the following research question.

Research Question 1.5: What is a useful method for analyzing simulation sequence data which can be of different lengths and multidimensional?

3.4.2.1 Data Mining

Antunes et al. define temporal Data is an ordered sequence of events and methods to analyze these data streams must be able to illicit hidden relationships from this order [87]. In this case, this temporal data is a set of continuous real-valued data and is therefore referred to as a time series [87]. Generally, applications that deal with time series serve to either predict or diagnose [87]. In addition to data related characteristics, our stated goals will inform the methodology as well. Next, data mining methods are categorized and discussed [88, 89].

Data Mining methods enabling analysis of time series data are shown in Antunes et al. and in Fu's survey on time series data mining for engineering applications in AI [87, 90]. Specifically, Association Rules, Classification and Clustering methods are surveyed in both sources. Association Rules identify the conditional probability of achieving some response given the current Conditions[89]. While beneficial in use to the ends of learning

player tendencies, Rushing et al. observe that the perception of the conditions may not be all encompassing, suggesting the presence of hidden variables [91]. This allows inconsistency in the decisions as well as prevents understanding of causality which is desired for the identification of CDPs [91].

Classification is a supervised method of machine learning. In this method, pre-defined classes, or types of conditions in this case, are used to identify and arrange the raw data into these types. Alpaydin uses credit scoring as an example, where credit worthiness was identified from the data [89]. In this case, utilizing bank data consisting of savings, income and past loan history, classification was done separating the data into two classes: high-risk and low-risk customers. In classification, it is possible to learn the bounds and limits separating the classes, in order to identify those conditions in future data. While extremely powerful in application, this method relies explicitly in predefining classes. Applied to battlefield wargaming, decision makers are presented with unique and unprecedented situations preventing the outright definitions of classes. Additionally, applied to games, supervised methods perform poorly for games which they were not defined [92]. Therefore, changing battlefield conditions that affect outcomes is the same as changing rules of the game. As a result, methods that do not utilize prior assumptions are desired.

Unsupervised learning methods make no assumptions about the character of the data. Methods of unsupervised learning include Reinforcement Learning (RL) and

Clustering [89].   Classes and relationships are discovered as a result of analysis. Reinforcement learning is a method for learning sequential decision-making strategy in the analysis of cause and effect relationships. It is powerful in situations where adversary intent is unknown or not easily implemented. RL methods identifies good policies, by considering the sequence of actions, rather than the actions themselves, and the associated outcome in the form of reward [89]. Once enough sequences are analyzed, there is sufficient learning to facilitate decision-making. Madeira et al. utilize RL to inform an Artificial Intelligence (AI) agent to control the French army in a simulated battle with the Russian army in 1812 with positive results; however, learning facilitated by RL required training the AI for 10,000 games [93]. While this level of training doesn't negate the benefits of RL, the need for training data imposes the same problems of applying supervised methods to games for which they were not defined or trained for. This negates the benefit being pursued in the formulation of this methodology.

3.4.2.2 Clustering

Clustering is similar to Classification but does not require the predefined classes. Elsing et al. defines these methods as ones which identify hidden relationships within the data, where elements within a cluster are similar to co-cluster elements and are different from out of cluster elements [94]. Most applications of clustering are used to illicit user behavior from games to aid game development and/or AI development [95-100]. However, different clustering methods exist which are aimed at solving different problems.

66

Bauckhage et al. identify four types of clustering algorithms: Hierarchical, Centroid Distribution-based and density Clustering [100], while Liao and Berkhin characterize static clustering methods into five major categories: partitioning methods, hierarchical methods, density-based methods, grid-based methods and model-based methods [101, 102].

By applying constraints with regard to the temporal nature of the data, Liao identified only three of these methods remain applicable to the problem: Partitioning methods, Model-based methods, and Hierarchical methods [94, 101]. This taxonomy will be utilized for Clustering method selection with methods described and surveyed below.

Model-based methods are utilized for data with high dimension assume that the observed data entries are the product of probabilistic models [103]. Similar to RL model-based methods require the use of training data and become invalid when the battlefield has changed [93, 101].

Partitioning methods include k-means, k-medoids and fuzzy c-means, where these procedures only work with time series of equal length [101, 102]. However, with modification, Liao et al. utilized clustering battle outcome prediction using raw data that was transformed into five indicators of battle state time sequences, which were then used to attempt prediction for different number period ahead for the indicated state [104]. While this method saw success in application, prediction of outcomes is not the objective of this thesis. This work aims to recognize, within the battlespace, conditions signifying CDPs in

order to facilitate COA development and analysis. However, using elements of Liao et al.'s work, it is likely that identification of related COAs will lead to identification of CDPs.

Hierarchical clustering methods are separated into two types: agglomerative which is a bottom-up strategy, and divisive which is a top-down strategy [89, 101, 105]. Generally, agglomerative methods are preferred to divisive because divisive methods are much more computationally complex; agglomerative methods are $O(N^2)$ while divisive methods are $O(2^{N-1} - 1)$ for N data points [106]. Agglomerative methods begin with each data element existing in a separate cluster, and are merged iteratively to the cluster with the closest proximity [106]. This process ends when there are no remaining data elements to place and the relative orientation and similarity is known among all data elements. This form of clustering group instances that are most like other instances within the cluster and most unlike instances in other clusters [89]. The output of this process is a hierarchy of the data and a method to visualize the hierarchy and structure of how the data is related through the use of a Dendrogram [106]. This will facilitate the identification and analysis of hidden patterns within the data and the identification of CDPs visually from the Dendrogram, fostering an improved understanding of battlefield variables which will aid in the research objective [102].

Response to RQ 1.5: Hierarchical Agglomerative Clustering best suits the need to analyze data to the end of identifying CDPs for the following reasons:

- It has the ability to analyze temporal data streams of different lengths

- It does not require training data or preconception of classes expected within the data

- It is less computationally complex over divisive methods

- The output of this process is a hierarchy of the data and a method to visualize the hierarchy and structure of how the data is related

As previously stated, HAC is an agglomerative method of grouping or clustering elements in a dataset. These elements are defined in a similarity matrix that define the distance between the elements. Elements that are close together are similar where elements that are relatively far from one another, are dissimilar. The process is illustrated below in Figure 12, where a notional data set is shown. Below, the letters 'a' through 'f' represent data elements and are shown to have some sort of relationship as seen in the graph. In this case, elements a and b are one unit apart. As are elements d and e. Element c is approximately 1.4 units away from element a while element f is two units away from element d. Utilizing these distances, a hierarchy of the data can be made.
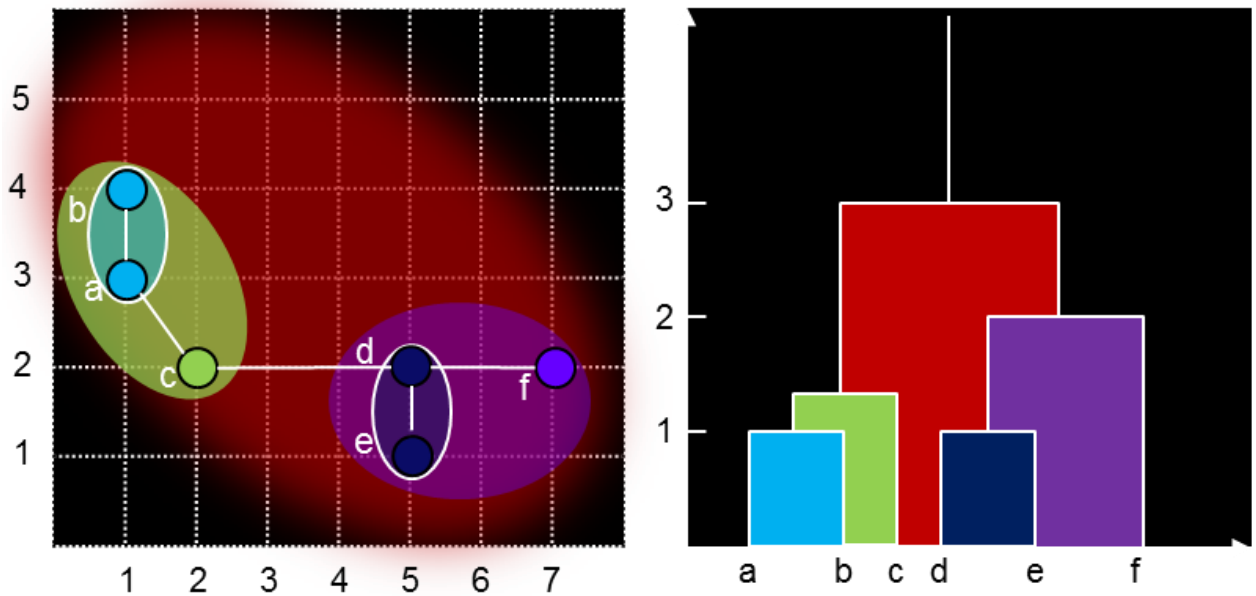
Figure 12. An example of Hierarchical Agglomerative Clustering on a notional data set illustrating the operations required to identify a hierarch from the data set.

The way that a hierarchy like the one pictured in Figure 12, is through the analysis of the relative distances of each element of the dataset. In the example above, the Euclidean distance between elements is used, as illustrated in Figure 13, a similarity matrix can be created by determining the pairwise Euclidean distance between elements. This similarity matrix represents the distance from each element of the dataset to every other element.

	a	b	c	d	e	f
a	0	1	1.4	4.1	4.4	6.1
b	1	0	2.2	4.4	5	6.3
c	1.4	2.2	0	3	3.1	5
d	4.1	4.4	3	0	1	2
e	4.4	5	3.1	1	0	2.2
f	6.1	6.3	5	2	2.2	0

Figure 13. Similarity matrix associated with the notional data set illustrating pairwise distances that are used to create a hierarchy through HAC.

Utilizing the Single linkage algorithm, a method that utilizes the two closest elements in separate clusters to determine the distance between two clusters, HAC is performed. On the first iteration of this method, the Similarity matrix is operated on to identify the closest pairs of elements. In the case above, seen in Figure 12, and Figure 13, these elements are a and b, as well as elements d and e. These four elements are joined in two sets with the height at which they join equating to the distance between those points, one. In this case, there were two sets of clusters that were joined in the first iteration. This

71

is because both sets of elements have identical similarity or distance. If only one set of elements was the closest or most similar set in the data set, then the first iteration would only form one cluster. Likewise, if there was a tie in similarity among many clusters, then there would be appropriately many clusters formed in the first iteration. Next each of the two newly formed clusters gains another element. a and b join c and d and e join f. According to the Single linkage algorithm, the heights that these clusters are joined are at 1.4 and 2 respectively. This is due to the fact that in the Single linkage, the closest points in two clusters are identified as the inter-cluster distance. For a and b joining c, the closest members between a, b, and c are elements a and c with a distance between them of 1.4. Similarly, elements d and e join f according to the closest elements. This is d joining f with a value of two, therefore on the Dendrogram, the cluster including d and e, joins f at a value of two, signifying the distance from cluster d-e to f. On the final iteration, the two larger clusters of a-b-c and d-e-f are joined. Per the Single linkage algorithm, the two closest points or elements between these clusters represent each cluster respectively. These elements are elements c and d, which are 3 units apart. This can be seen in the join height of the larger clusters in Figure 12. The HAC algorithm ends when all elements are joined in a cluster. The output of this process is a Dendrogram and the hierarchy of all of the data offered in the Dendrogram.

3.4.2.2.1 Intra-Cluster Distance Measure

In order for Hierarchical Agglomerative Clustering (HAC) to be effective, proper selection of a distance measure must be made. The first of two distance measure types discussed will the intra-cluster distance measure. This distance is the distance between two points. Generally, Euclidean distance, which is a case of Minkowski distance, is used [89]. Manhattan distance is another case of Minkowski distance [89]. The form of each distance is shown in Table 3 and are generally referred to as shape based similarity measures [106]. According to Aghabozorgi et al. the measures in Table 3 are best suited for static data which is not utilized in this research [105]. Additionally, the data application presented in this thesis utilizes data sets that are non-uniform in structure which Aghabozorgi et al identify as another constraint to using these shape based similarity measures [105]. However, two methods of measuring distance between data points as seen in this thesis are identified: Dynamic Time Warping (DTW) and Longest Common Subsequence (LCSS), which will be discussed in the following paragraphs [105, 107, 108].

Table 3. Between point distance measures utilized in clustering all based on the Minkowski distance are operations on the spatial separation between points x and y [106].

Measure	Form
Minkowski Distance	$D_{ij} = \left(\sum_{i=1}^{d} \left\| x_{ij} - x_{ji} \right\|^{1/n} \right)^{n}$
Euclidean Distance	$D_{ij} = \left(\sum_{i=1}^{d} \left\| x_{ij} - x_{ji} \right\|^{1/2} \right)^{2}$
Manhattan Distance	$D_{ij} = \sum_{i=1}^{d} \left\| x_{ij} - x_{ji} \right\|$

Dynamic Time Warping is a method that employs an elastic distortion of time to enable comparison of time series of unequal lengths. Senin and Müller illustrate the portions of the Classic DTW method that are shown in this thesis [109, 110]. First two time series, X and Y, where $X := (x_1, x_2, \dots, x_n)$ and $Y := (y_1, y_2, \dots, y_m)$ are chosen for comparison and the pairwise distances are calculated forming an NxM matrix. The time series are collapsed into a single dimension, a comparison value, however, the dimensionality of the data set remains, as each comparison represents a single dimension.

74

Therefore, if there are 500 elements or simulation data elements, there can be 500 dimensions of the similarity pairwise distance matrix.

Next, the minimum cost path between X and Y is found based on constraints that enforce boundary conditions of the sequences ensuring that the distance between the first and last points in each sequence are the first and last points in the cost path, monotonicity in cost path which ensuring time ordering of cost path points, and step size limits which ensure connectedness of the cost path barring large shifts in time [109, 110]. The form of the cost function can be seen in Equation (1) and utilizes the Euclidian distance between points x_{nl} and y_{ml}. For large matrices, there will be many feasible cost paths to consider and can be computationally burdensome. However Dynamic Programing utilizing the DTW distance function shown in Equations (1) and (2) aids in the efficient calculation of pairwise distances of sequences [109, 110]. After the distances are calculated, the data can then be clustered using HAC.

$$c_p(X, Y) = \sum_{l=1}^{L} c\left(x_{nl}, y_{ml}\right) \qquad (1)$$

$$DTW(X, Y) = c_{p^*}(X, Y) = min\{c_p(X, Y), p \in P^{NxM}\} \qquad (2)$$

Longest Common Subsequence (LCSS) model is similar to edit distance where sequences are allowed to stretch allowing comparison, while some elements remain unmatched [108]. LCSS identifies the longest string or substring present in two sequences under consideration. Similarity is treated as the ratio of common substring length to average of each string length.

As an example to illustrate the utility of each method, two non-uniform vectors are defined below:

$$Let \; x = \; [1, 1, 2, 3, 2, 0]$$

$$and \; y = \; [0, 1, 1, 2, 3, 2, 1]$$

The distance between the two vectors according to DTW is calculated below in Figure 14. The distance must be converted to similarity to compare with LCSS. This is done by converting the non-zero entries in the distance vector to binary values, where zero

76

represents complete similarity and one represents complete dissimilarity. This modified

vector is shown as $\widetilde{D}_{xy}$.

Figure 14. Dynamic Time Warping calculation using dynamic and programing finds the optimum path which is then translated to a similarity from a distance vector calculation.

Using the same vectors, x and y, the LCSS identification and the associated

similarity measures are defined below.

$$Let\ x = [1, 1, 2, 3, 2, 0]$$

$$and\ y = [0, 1, 1, 2, 3, 2, 1]$$

$$\therefore LCSS = [1, 1, 2, 3, 2]$$

$$S_{xyDTW} = 1 - \frac{\sum_{k=1}^{n} \tilde{D}_{xy_k}}{n} = 0.714$$

$$S_{xyLCSS} = \frac{|LCSS|}{min(m,n)} = 0.8333$$

While the previous example was shown to illustrate the inner workings of each method, it is not representative of the expected data set. The expectation is that time series vectors with differing lengths will be obtained through simulation. Therefore, with similarity definitions in place, it is now possible to determine canonical examples to define the strengths of each method to aid in final selection that mirror the expected data type. Four cases were developed to assess the performance. The first case, Case zero, was made to illustrate identical sets of telemetry data. Vectors defined below as X0 and X1:

$$X0 = sin(x)$$

$$X1 = sin(x)$$

$$where\ x = [0: 0.1: 2\pi]$$

In this case, both similarity measures are identical and equate to one. Next the remaining three cases with descriptions will be shown.

Case one employs two vectors with similar telemetry data. In this case, X1 remains unchanged while X2 is modified by changing the first five vector entries to zero, and remaining vector entries are the same as X1 as seen in Figure 15. Case two modifies the data in X3 with respect to X1. The end result is a vector X3 where there is a shift in phase and translated preceded by a run of zeros which can be seen in Figure 16. Case three is the most modified where the data sets appear different. This case has vector X4 linearly increase to a value of one, then continues with the X3 formulation. This case can be seen in Figure 17.
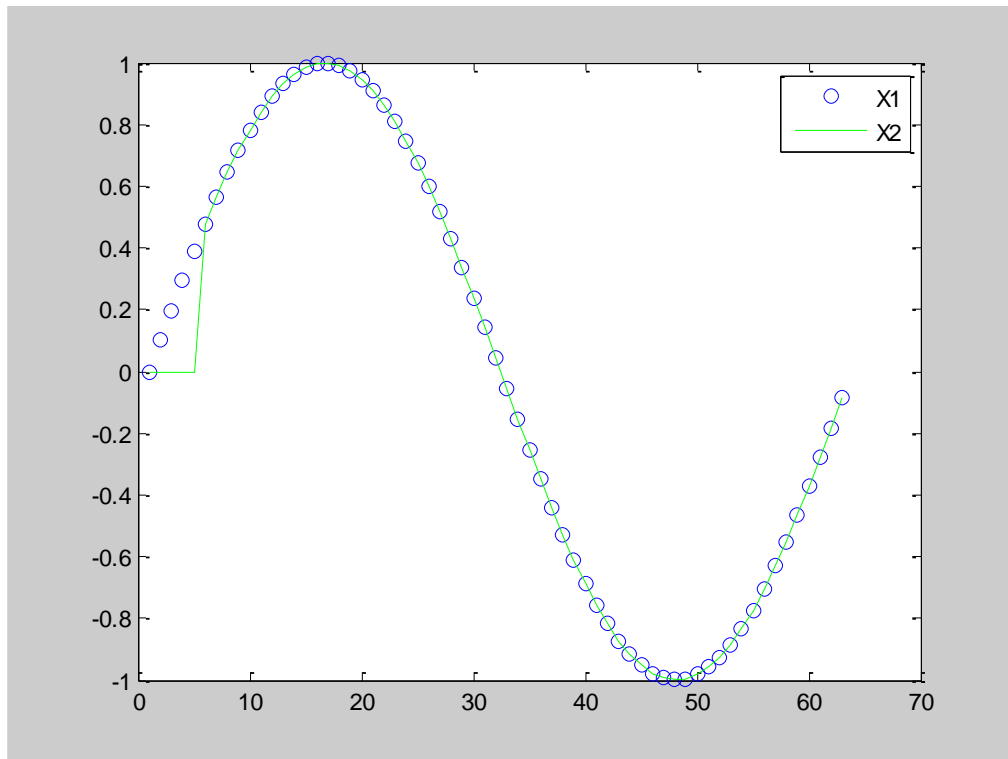
Figure 15. Case one comparison of two trajectories where X2 is modified by replacing the first five values with zero to illustrate how a slight variation can affect the distance.
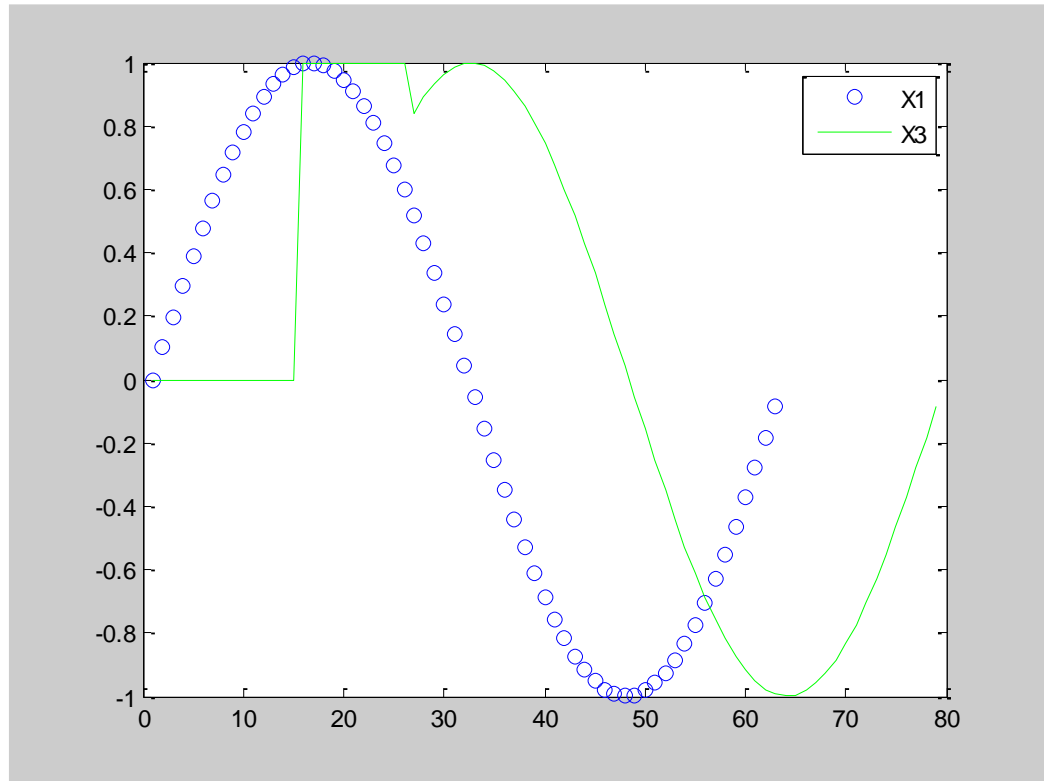
Figure 16. Case two comparison where X3 is modified by replacing values in two sequences, then translating the entire data sequence by twenty values to illustrate how a more substantial modification affects the measure of distance.
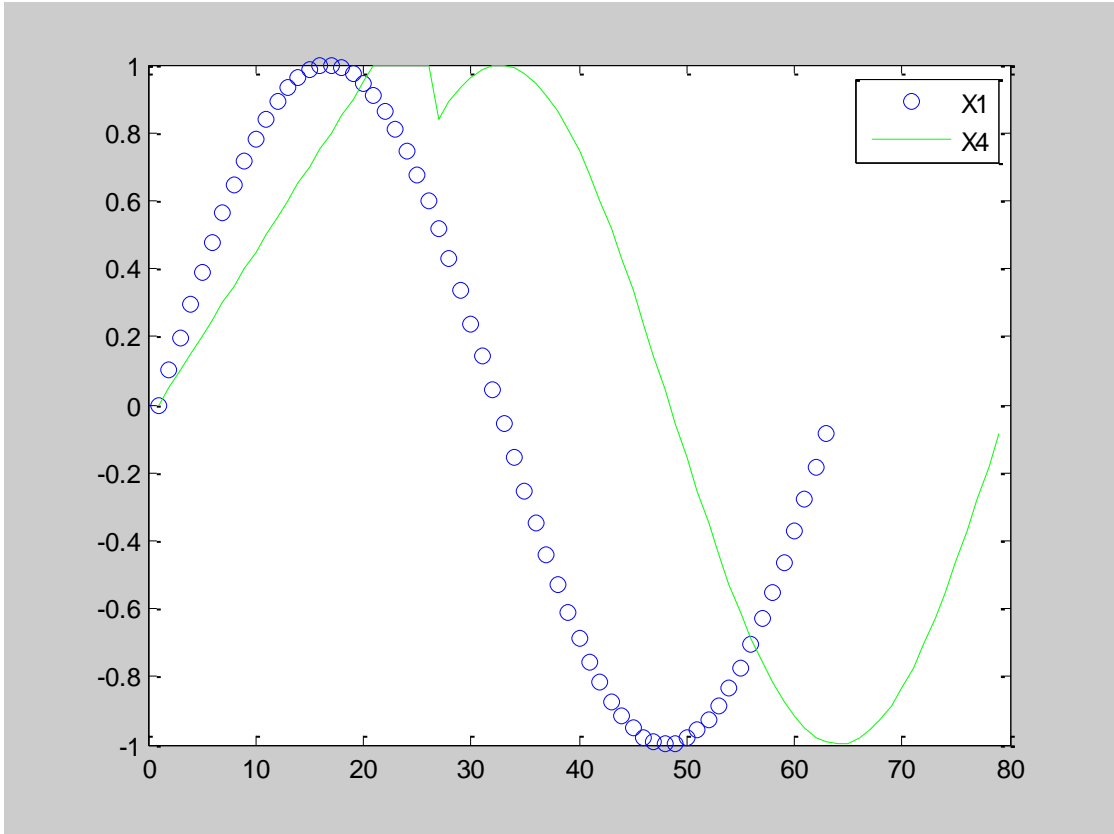
Figure 17. Case three comparison where X4 is translated and contains replaced values, but, also includes a unique sequence not present in X1 to illustrate the effect on distance.

With a visual description of each case, the similarity measures shown in Table 4 have increased utility. As shown in the graphs and table, Dynamic time warping performs best on time shifted series but appears to have a diverging similarity measure while a strong similarity is still observable, while the LCSS measure appears overstate the similarity as can be seen in Figure 17. Observe that the total proportion of the green trend that deviates

81

from the Sin wave is approximately 30 units out of the total 80 units. This represents a 62.5% deviation from the original Sin wave. While neither method identifies the stated deviation from the previous sentence Table 4 shows that DTW furnishes the more conservative estimate.

Table 4. Similarity comparison of LCCS and DTW showing the core detriment of LCSS and the lack of granularity in assessing the similarity of the last two cases.

Case	Similarity LCSS	Similarity DTW
Case 0	1	1
Case 1	0.9365	0.9403
Case 2	0.7606	0.7738
Case 3	0.7606	0.5625

3.4.2.2.2 Linkage Algorithms

With intra-cluster distance measures defined, it is necessary to define an inter-cluster distance measure. This requires defining a spatial distance measure, such as Euclidean, and also a method of measurement which is known as a Linkage Algorithm. According to Murtagh et al., linkage algorithms can be classified into two groups: graphical methods that utilize the actual data points for measure, and geometric methods that identify the cluster center [111]. These methods can be seen in Table 5 and Figure 18. The graphical methods use different philosophies in determining the relative distance

between clusters. Single linkage utilizes the closest points between two clusters to determine the distance between clusters. Complete linkage uses the distance between the two furthest elements to determine the cluster distance. Average uses the unweighted average of all of the distances of all of the elements in each cluster. The Geometric linkage, a graphical method utilizes the calculation of a geometric center for each cluster and identifies the inter-cluster distance as the distance between the two cluster-specific geometric centers.

The differences presented in the previous paragraph extend into and influence the performance of each method leading to identifiable characteristics associated with each method. According to Jain et al., Complete and Single linkage are the most popular, but have different and conflicting strengths [112]. These differences will be elucidated in the coming paragraphs.
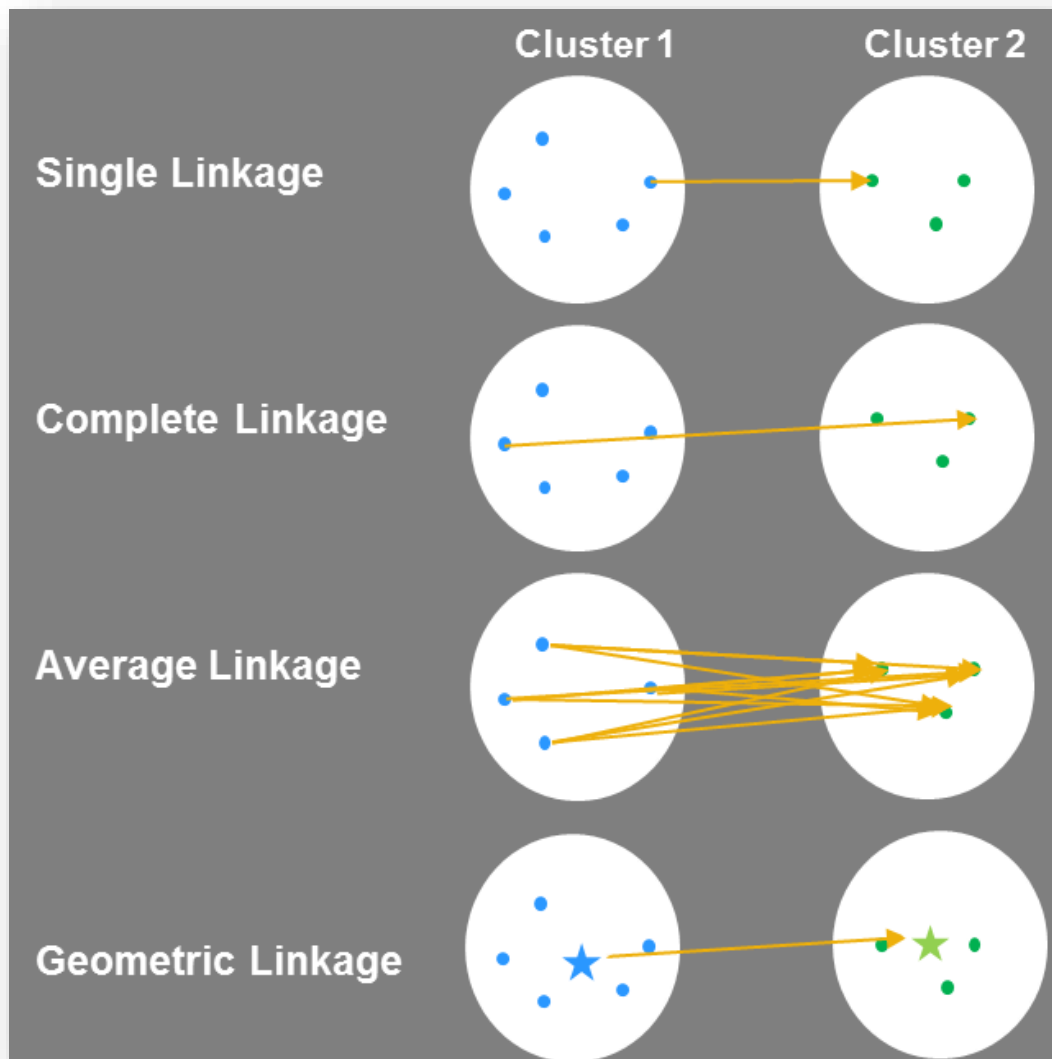
Figure 18. Graph and geometric linkage examples illustrating the mechanism behind inter-cluster distance evaluation for both geometric and graph methods.

While the Single linkage algorithm is popular, a detriment in some cases is the fact that this method forms chained clusters. Conversely, Complete linkage algorithm forms tightly bound clusters [112]. Ward's method creates well defined clusters [113]. UPGMA outperforms single link on standardized data and ensures inversions in the hierarchy do not occur [114, 115]. Weighted Average (WPGMA) performs similarly to UPGMA but employs a priori weightings [116]. Centroid and median linkage methods do not preserve monotonicity, which implies that it is possible to merge latter clusters with former clusters, affecting the hierarchy [117]. Each algorithm has desirable attributes, but making the final selection on linkage algorithm for this problem troublesome which leads the next research question below.

Table 5. Linkage algorithms utilized in this work spanning both graph and geometric measures of similarity and inter-cluster distance.

Linkage Algorithm	Method
Single	Graph
Complete	Graph
Group Average (UPGMA)	Graph
McQuitty's Method (WPGMA)	Graph
Median (WPGMC)	Geometric
Centroid (UPGMC)	Geometric
Ward (Minimum Variance)	Geometric

Research Question 1.6: What Linkage algorithm should be used based on the nature and

state of the data?

The choice in linkage algorithm affects the hierarchies obtained from the data sets. In a sense, using an incompatible linkage algorithm for clustering can distort the data, rendering the output of the process invalid. However, utilizing suitable algorithms for the data can yield a hierarchy of simulations and COAs, from which a subset can be found that can increase the probability of success. A synthetic data set was created to illustrate this example and is illustrated in Equation (3). After the creation of the synthetic data set, HAC was performed using two linkage algorithms to illustrate the peculiarity possible from mismatched data and algorithms. Observe in Figure 19 that by using different algorithms, the final hierarchies are different, which can be identified by the red line highlighting the location of the highest value in X. The inner squared distance linkage places that data point within a cluster while the single linkage algorithm has the same data point in a cluster all to itself. This is particularly problematic because of the goal of using clustering, which is illustrated previously and is presented below in the following Hypothesis:

Hypothesis 2: It is possible to select a linkage algorithm that yields groupings of simulations that improve the probability of success, over other linkage algorithms.

The implications of mismatching a linkage algorithm and a specific dataset can be illustrated with the example below utilizing Equation (3), a synthetic data-set. In this example, fifty uniformly distributed samples between one and fifty were drawn and stored in a vector X. This vector was considered a data-set and was subjected to HAC via two linkage algorithms, Inner squared linkage and Single linkage. The in observing the vector,

87

it is seen that the highest value present is fifty, which is expected. However, the linkage

algorithms place that most extreme value differently. Utilizing the inner squared distance,

this value is placed in a cluster with similar magnitude neighbors, while the Single linkage

treats it as an outlier.

$$X \sim U(1,50)$$

(3)

$$where\ X = [x_1, x_2, x_{3,\dots} x_{48}, x_{49}, x_{50}]$$

Drawing from the parallels of this example to the intended use cases of this thesis,

the importance of matching a proper linkage algorithm to the data sets presented here is of

utmost importance. If true hierarchies cannot be identified, then the identification of COA

types with CDP's cannot be identified either. Therefore, there is a need to match a linkage

algorithm to the data that will be experienced in this thesis.
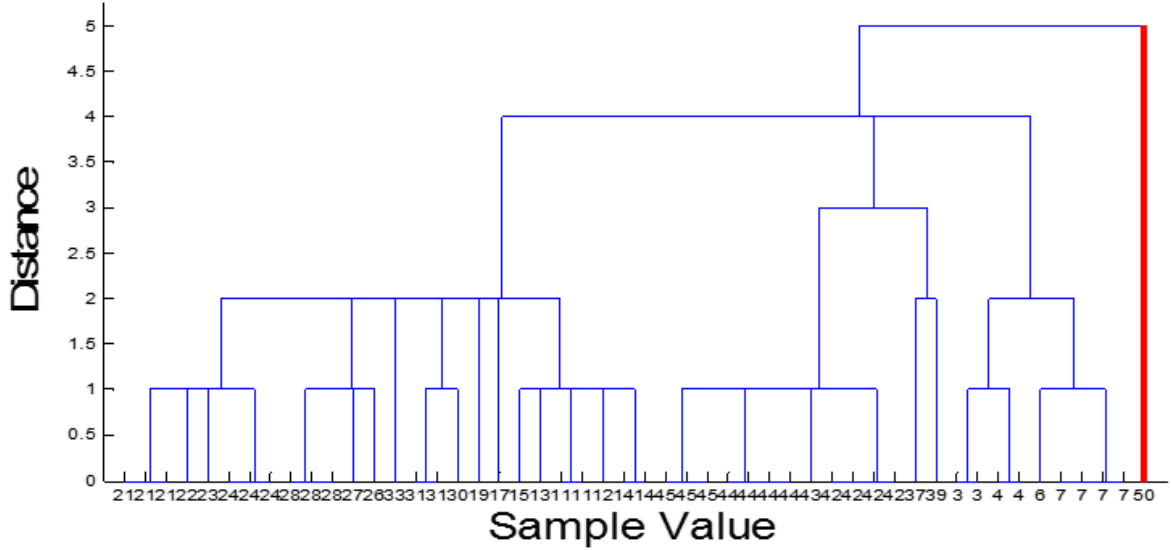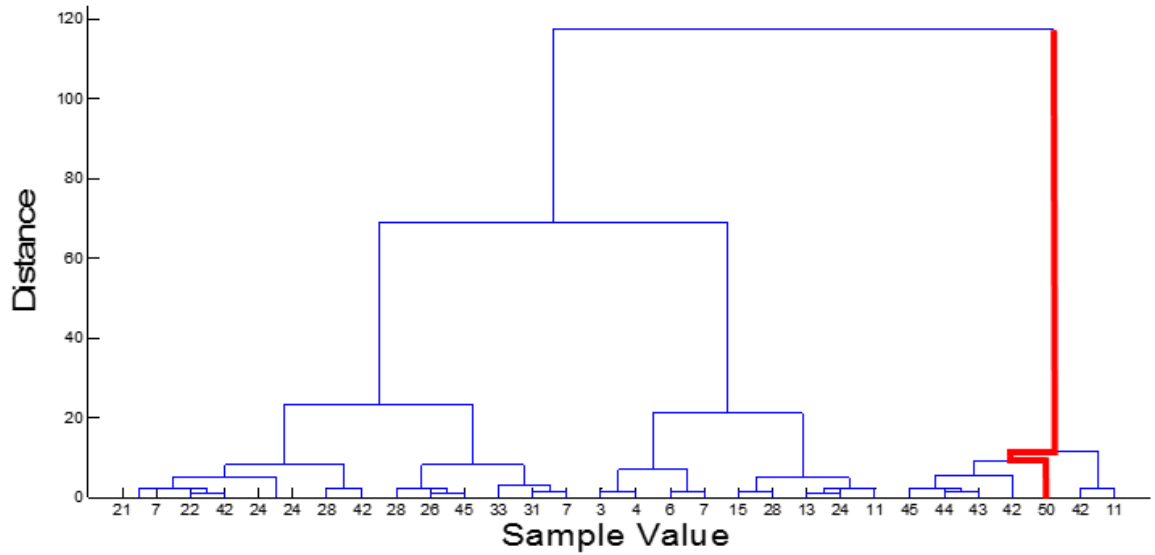
Figure 19. HAC of synthetic data utilizing two different linkage algorithms can have a drastic effect of clustering an element or treating it as an outlier.

3.4.2.2.3 Cluster Validation

In order to compare and contrast these methods there needs to be a way to assess behavior and performance, but first, there must be a method to conveniently and generally transition between linkage methods. The Lance-Williams dissimilarity update formula shown in Equation (4) with coefficient values for each linkage type shown in Figure 20 is the method to do so and will be employed here[118]. Although this facilitates the comparison of linkage clustering methods, the method of comparison is yet undefined. This motivates the next two research questions where the initial question will aim to resolve how to transition between linkage algorithms and the latter question will seek to identify validation metrics to compare linkage algorithms.

$$d_{k(ij)} = \alpha_i d_{ki} + \alpha_j d_{kj+} \beta d_{ij} + \gamma \left| d_{ki} - d_{kj} \right|$$

(4)

Hierarchical Clustering Methods	Lance and Williams dissimilarity update formula	Coordinates of center of cluster, which agglomerates clusters i and j	Dissimilarity between cluster centers g_i and g_j
Single link	$\alpha_i = 0.5$ $\beta = 0$ $\gamma = -0.5$		
Complete link	$\alpha_i = 0.5$ $\beta = 0$ $\gamma = 0.5$		
Average link	$\alpha_i = \dfrac{\lvert i \rvert}{\lvert i \rvert + \lvert j \rvert}$ $\beta = 0$ $\gamma = 0$		
McQuity's link	$\alpha_i = 0.5$ $\beta = 0$ $\gamma = 0$		
Median link	$\alpha_i = 0.5$ $\beta = -0.250$ $\gamma = 0$	$g = \dfrac{g_i + g_j}{2}$	$\left\lVert g_i - g_j \right\rVert^2$
Centroid link	$\alpha_i = \dfrac{\lvert i \rvert}{\lvert i \rvert + \lvert j \rvert}$ $\beta = \dfrac{\lvert i \rvert}{(\lvert i \rvert + \lvert j \rvert)^2}$ $\gamma = 0$	$g = \dfrac{\lvert i \rvert g_i + \lvert j \rvert g_j}{\lvert i \rvert + \lvert j \rvert}$	$\left\lVert g_i - g_j \right\rVert^2$
Ward's link	$\alpha_i = \dfrac{\lvert i \rvert + \lvert k \rvert}{\lvert i \rvert + \lvert j \rvert + \lvert k \rvert}$ $\beta = \dfrac{\lvert k \rvert}{\lvert i \rvert + \lvert j \rvert + \lvert k \rvert}$ $\gamma = 0$	$g = \dfrac{\lvert i \rvert g_i + \lvert j \rvert g_j}{\lvert i \rvert + \lvert j \rvert}$	$\dfrac{\lvert i \rvert \lvert j \rvert}{\lvert i \rvert + \lvert j \rvert} \left\lVert g_i - g_j \right\rVert^2$

Figure 20. Lance and Williams recurrence formula values[111].

Research Question 1.7: What methods exist to compare outputs of the clustering process?

In order to compare linkage algorithms validation measures must be defined. Gan has separated the approaches with regard to Figure 21 [119]. The primary limitation on index selection is based on whether the desired criteria is external, internal or relative. For the application presented in this thesis, external criteria are not possible as there is no pre-specified structure in which the data exists. In Gan's formulation, Internal and Relative criteria are applicable [119]. Internal and Relative measures of validity evaluate the clustering performance based on inherent features of the data. While a variety of indices, measures and associated inherent features are defined in Gan's work, strengths and weaknesses associated with each measure, and rules of thumb for application are not specified.
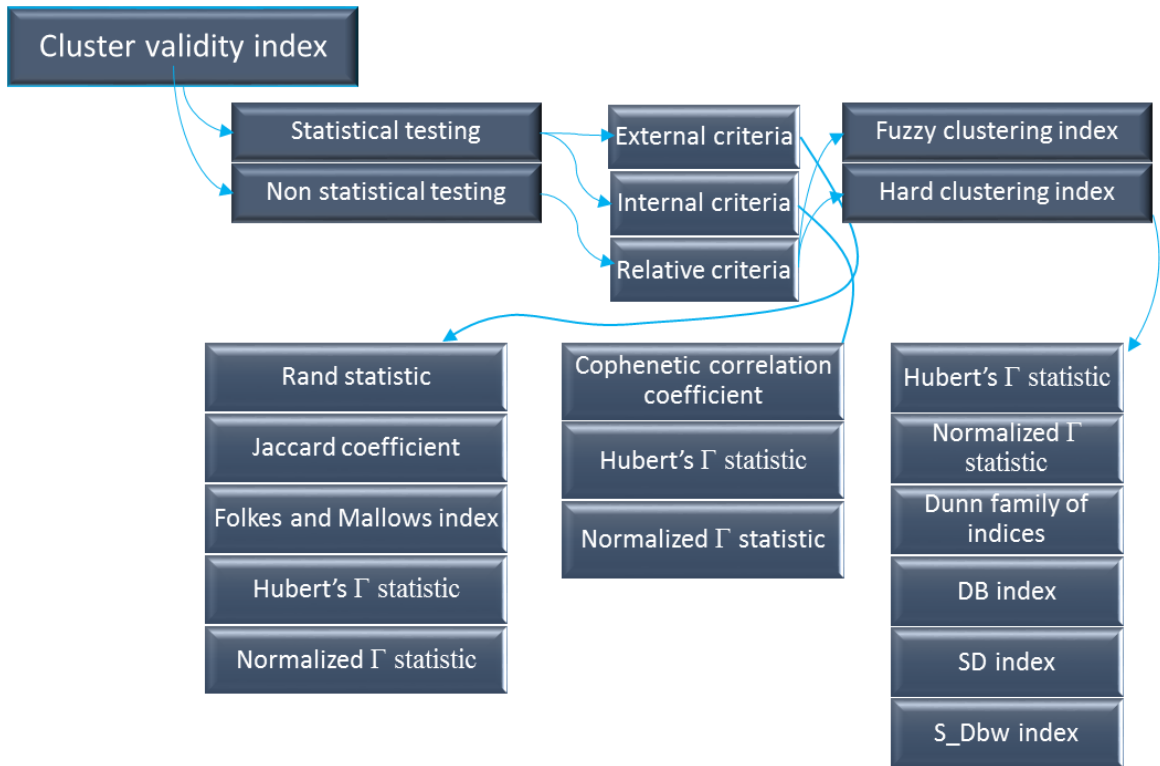
Figure 21. Cluster Validity Index selection process flow showing a reduced subset of indices that can be utilized regardless of desired criteria preference [119].

Aggarwal and Liu separate validation indices into either External or Internal, where Gan's Internal and Relative indices are merged into one internal index list [120-122]. Additionally, this work identifies rules of thumb for index employment based on the nature of each index and the sensitivity of the indices to monotonicity, noise, density, sub clusters, skewed distributions and arbitrary shapes present in the data [120-122]. These data characteristics will be discussed in the following paragraphs.

Monotonicity of internal validation indices is based on the number of actual clusters present in the data set [120]. If an index either increased or decreased monotonically based on the linkage algorithms output of number of clusters, such a measure can be seen as troublesome. For example using the Root-mean-square std. dev. (RMSSTD), the number of clusters influences the index value as can be seen in Equation (5), where NC is the number of clusters and is in the denominator of the equation [120]. This is what drives the index value down as the number of clusters increases. Therefore, *ceteris paribusa*, a linkage algorithm that tends towards splitting clusters will outperform a conservatively clustering algorithm, and the margin of performance will continue to increase for an increased number of perceived clusters. For example, in Figure 22, there appears to be five clusters. If a conservative linkage algorithm identified five clusters, and another algorithm split each cluster in two, the algorithm seeing ten clusters may have a higher performing index as a result of incorrectly identifying ten clusters as opposed to five.

$$RMSSTD = \sqrt{SSE/P(n-NC)} \qquad\qquad (5)$$
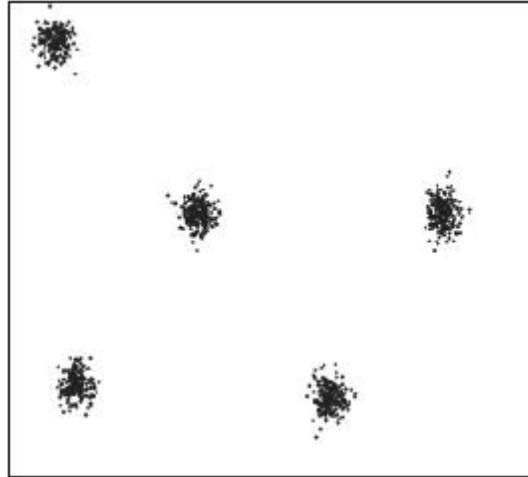
Figure 22. Monotonicity example data showing that increasing the number of perceived clusters in a data set can affect the way indices and linkage algorithms are regarded.

According to Chapman et al. the presence of noise or the effects of noise on data can affect the linkage algorithms and indices as well [120]. As seen in Figure 23, the presence of noise serves to decrease the density of the data and also introduces the presence of outliers within the data. However, this phenomenon can be seen in non-noisy data as well. Non-linearity and complexity can project the same behaviors within the data. Whether or not the actual data is noisy, the presence of outliers in the data can be troublesome for linkage algorithms. For example, the Calinski-Harabasz index is sensitive to noisy data and for an undistorted clustering will be undervalued if noise is introduced, making the measure unstable [120].
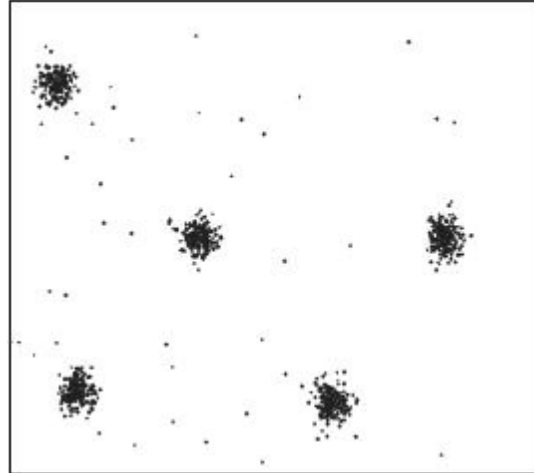
Figure 23. Noisy example data illustrating the potential effect of loosely assembled clusters on the evaluation of indices and subsequent linkage algorithms assessed.

In a similar way to monotonicity and noise, the size, density and presence of sub-clusters complicates the clustering process and affects the way in which to measure the performance of linkage algorithms. This is in a large part due to the methods in which validation indices evaluate the algorithms. The presence of heterogeneous clusters, with different sizes, shapes and densities, is troublesome for a generic calculation that is better suited for a homogeneous cluster set. More robust measures are needed for these types of data particularly present in scenarios where elements of the data are new or poorly understood and the potential for any type of data exists.
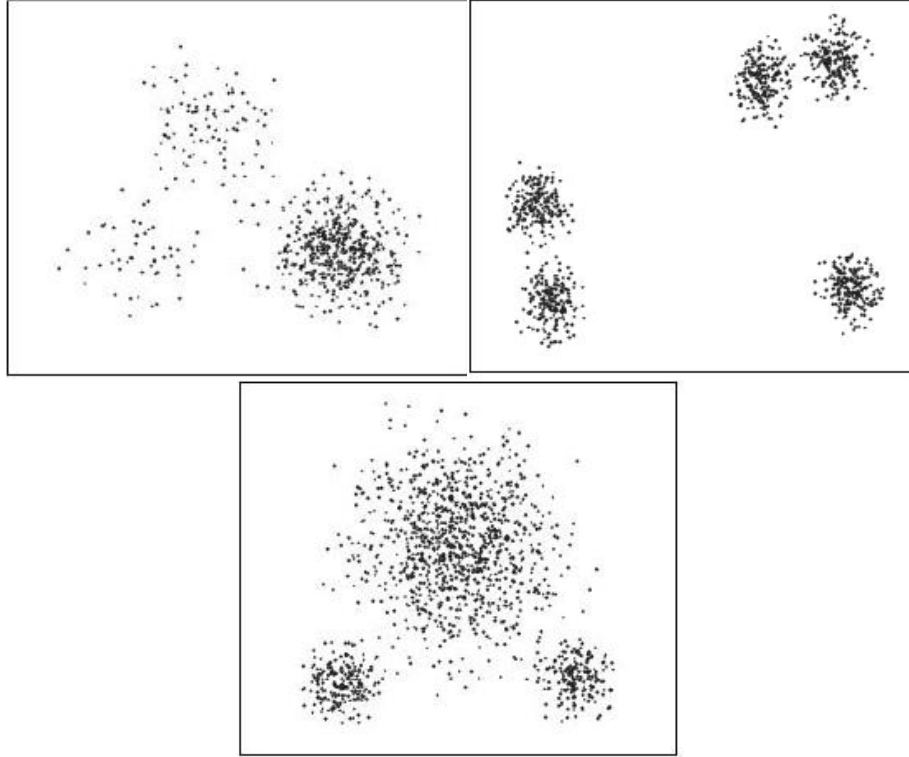
Figure 24. Skewed, variable density and sub-cluster data also affect the evaluation of linkage algorithms.

Data of arbitrary shapes influences the index measures in a similar way to skewness, density and presence of sub-clusters. This can be seen in Figure 25. In data sets like these, there is an obvious behavior and pattern to the data that may or may not be spherical as in the other examples above. Additionally, these types of data sets, can have multiple and competing behaviors giving rise to the different shapes found. Furthermore, these data set clusters may not be convex either. Xiong et al. illustrated the classifications of the data and applied to twelve validation indices through the use of six synthetic data sets [120]. The results are shown in Table 6, where the clustering validation index based

97

on nearest neighbor (CVNN) appears to have adequate performance across all of the data

sets [120].
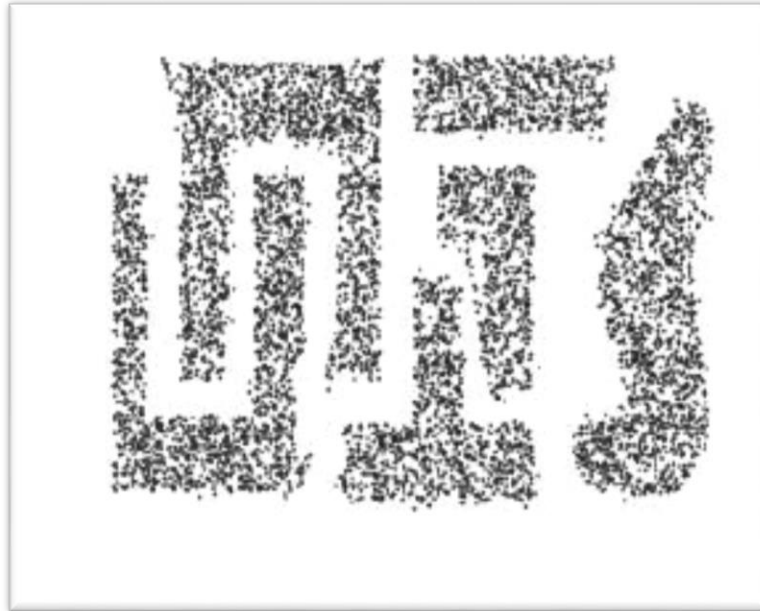
Figure 25. Arbitrarily shaped data.

Table 6. Validation index expected performance across data types where green is meets or exceeds expectations, red does not meet expectations and yellow is untested.

Measure	Monotonicity	Noise	Density	SubIcluster	Skew Distribution	Arbitrary Shape
RMSSTD	red	yellow	yellow	yellow	yellow	yellow
RS	red	yellow	yellow	yellow	yellow	yellow
Γ	red	yellow	yellow	yellow	yellow	yellow
CH	green	red	green	green	green	red
I	green	green	red	green	green	red
D	green	red	green	red	green	red
S	green	green	green	red	green	red
DB**	green	green	green	red	green	red
SD	green	green	green	red	green	red
S_Dbw	green	green	green	green	green	red
XB**	green	green	green	red	green	red
CVNN	green	green	green	green	green	green

CVNN is calculated using intercluster separation and intercluster compactness by using dynamic geometric information of the cluster. This is facilitated by the utilization of the maximum value obtained by utilizing nearest neighbors for each cluster. This calculation enabled by the nearest neighbors ensures that there is a unique measure for each cluster that best describes its nature.

Rules of thumb for index employment based on the nature of each index and the sensitivity of the indices to monotonicity, noise, density, sub-clusters, skewed distributions and arbitrary shapes present in the data exist and have been shown to aid in selection of

clustering parameters [120-122]. However, in highly dimensional datasets, visually identifying sparsity, compactness, sub-clusters, or noise in order to select a proper validation index is not possible. Time series data, regardless of the dimensionality, has the same effect. This effect is further magnified if the series are of different lengths. Specifically, sparsity, compactness and sub-clustering are not applicable when inspecting the trajectories of time series data.

With no obvious index to select we look to some of the methods previously discussed to identify any trends. Below in Figure 26 we see that there are multiple methods that can be used to identify an optimum number of clusters. Using the UPGMA linkage algorithm, HAC was performed. Next, for all possible values of cluster numbers, metrics were calculated and plotted and can be seen in Figure 26. The performance of each of the methods can be seen and appear to incline towards different optima. Specifically, S_Dbw and Average Cluster Sum of Squares, while unequal, seem to be conservative measures of the number of clusters, where Average Silhouette Width and the clustering validation index based on nearest neighbors (CVNN) tend towards a more extreme interpretation of the Dendrogram [120, 123, 124]. Based on the available data and the performance of these methods there does not appear to be a most suitable method to utilize. Therefore, we look to the primary motivation behind attempting to classify the data, which is to identify classes of games that can lead to an increased probability of success.
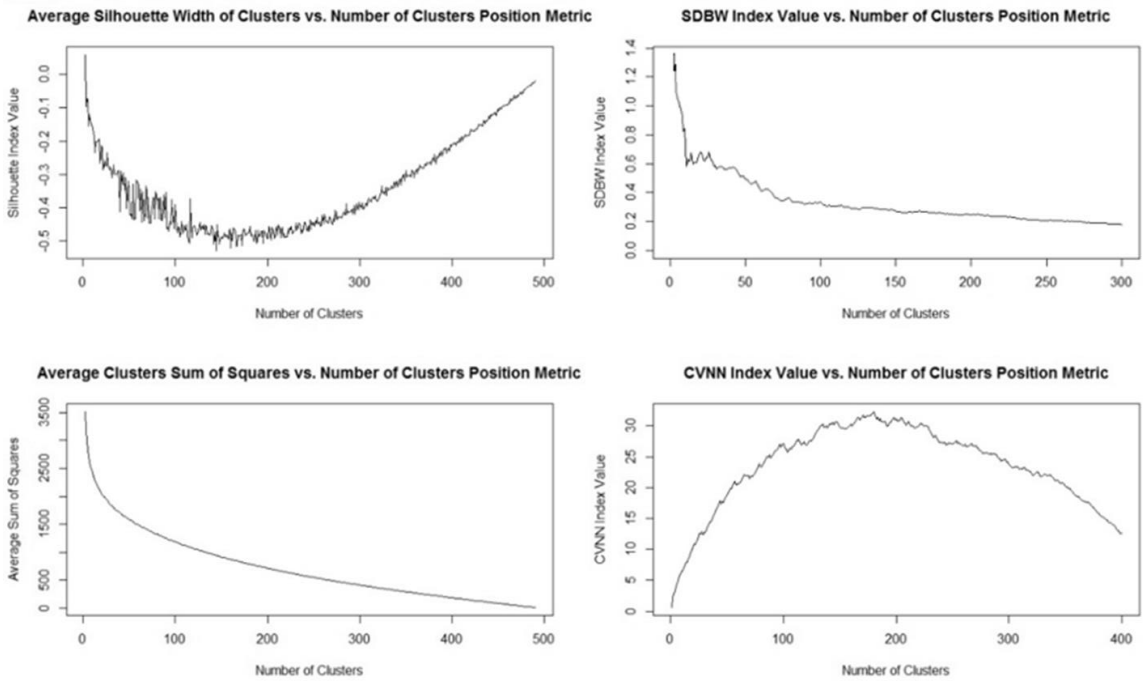
Figure 26. Performance of validation indices vs. number of clusters is varied depending on the index, yielding contradictory insights and showing no clear optimum.

By comparing the formulation of the validation indices shown above to our primary objective, it is hopeful that the index will lead to the best possible configuration of both linkage algorithm and cluster parameters to identify regions of the decision space that lead to an increased probability of success. With that in mind, inter-cluster compactness and intra-cluster separation were chosen to begin the investigation with a preference towards maximizing the former and minimizing the latter. This preference sees the ideal case as one with very compact clusters which are very far from other clusters. Thus, the elements of a cluster are very similar to one another and very different from the elements of other clusters. These measures were chosen because they represent the majority of internal validation indices as most of the internal indices are manipulations of both inter-cluster compactness and intracluster separation [120]. By creating hierarchies using HAC for all of the linkage algorithms listed below in Figure 27 the average distance within clusters and between clusters were calculated by sweeping through the number of clusters. The result of which can be seen below.

Performance of Lingkage Algorithms
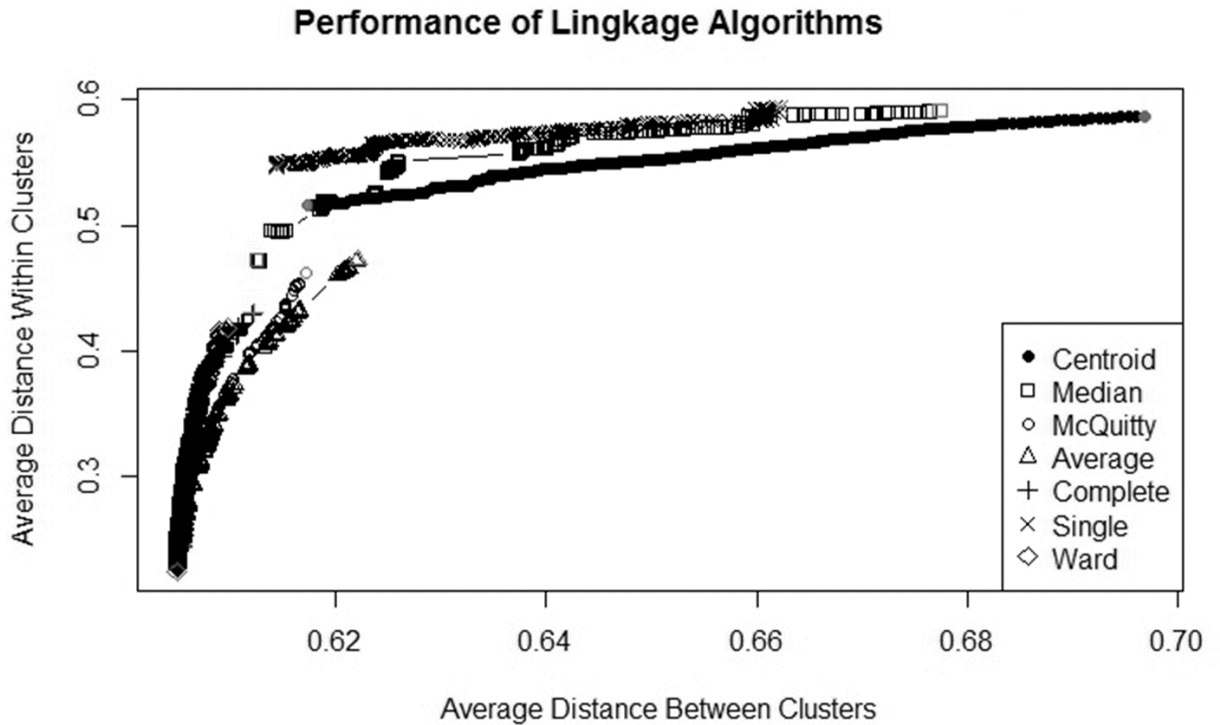
Figure 27. Performance of linkage algorithms using within and between cluster distance without a well-defined optimum value of number of clusters or linkage algorithm.

The primary observation to be had is that this procedure formed a non-convex Pareto frontier. Next, it seems that the entire frontier is made up of subsets of the linkage algorithms. Specifically, the Centroid Algorithm makes up the frontier from the upper right hand side of the plot to approximately the middle of the graph, while the Average Linkage Algorithm makes up the remainder with a small portion of Complete Linkage Algorithm points on the lowest left hand side of the plot.

A few general points can be taken from these behaviors. First, it seems that the Centroid Linkage Algorithm maximizes the intra-cluster separation best but does so at the cost of inter-cluster separation. Next, the Average Linkage Algorithm appears to be a compromise between the two objectives. Lastly, the Complete Linkage Algorithm behaves conversely to the Centroid Algorithm by minimizing the inter-cluster separation at the cost of intra-cluster separation. With a solution for preferences of each extreme and for a compromised solution, it is evident that a final selection of a linkage algorithm must draw from more information present in the data than only the distance within and between clusters. Furthermore, a known preference towards inter-cluster proximity or intra-cluster distance does not reveal the number of clusters to analyze. In an effort to facilitate the analysis of COAs using unsupervised learning, a method encapsulating pertinent, data-set specific information to make cluster parameter selection is needed.

As stated in the previously, a method that utilizes problem specific information to find the optimum number of clusters and linkage algorithm is sought after. While the benefit of a generic method to identify clustering parameters is simplicity, different analysis and datasets are subject to different levels of hazard associated with misclassification. A characteristic found in this dataset is a conditional outcome in addition to the metric information chronicled from each run. While this formulation will be utilized on ABMs, it is implied that this methodology can be applied to any dataset that has metric time series information coupled with a condition or outcome.

104

Internal validation indices commonly operate on measures of inter-cluster compactness and intra-cluster separation [120]. As a result, nuances present in different datasets have no bearing on the performance of the linkage algorithms and associated index measures making recommendations of the process generic to all datasets. It is the aim of this step in the formulation to determine a way to identify which resolution to view the Dendrogram giving the most significant view of the hierarchy so that groups of runs that have increased or decreased win probability can be identified. The change in probability due to clustering will be seen in the absolute case and will be referred to as polarity. While it is more advantageous to seek clusters where the win probability is increased, knowledge of areas of decreased win probability are also advantageous in a strategic setting as well. A reference frame centered on polarity will aid in the remainder of the formulation where we attempt to find groups of each type, win or loss and test the magnitude of their homogeneity. To do this first, we investigate how clustering can be shown to increase the polarity of a class of run type.

The proposition above can be accepted or rejected based on the following points. If there is no clustering assignment that will increase the polarity, then groups will be comprised of a mixture of wins and losses equivalent to the total win probability. This point can be shown using hypothesis testing and for the outcome types appropriate to this problem, win or lose. Therefore, the binomial distribution will be used to identify the significance of the polarity of groups.

105

In a simple experiment, observed in Figure 28, one thousand binomially distributed samples were drawn, simulating an initial dataset and can be seen in the top of the figure where black represents a win and grey represents a loss. Next 100 samples were drawn from the initial dataset and a binomial test was performed on the group with the results on the right hand side of Figure 28. As can be seen, the results of the 10 samples do not give enough evidence to reject the null hypothesis for a rejection region of 0.05. However, a limitation of hypothesis testing is that it only tells the probability of seeing more extreme behavior. For the problem set provided, it is the probability of seeing more polar behavior. It does not give the actual win probability, but gives the probability that the win or loss probability is higher or lower respectively. While lacking, this is still a very appealing aspect to utilize. Therefore, if a similar foundation is set for the analysis of clusters of simulation data, p-values less than 0.05 will give sufficient evidence to reject the null hypothesis, suggesting a significant classification that can be used to increase probability of success.
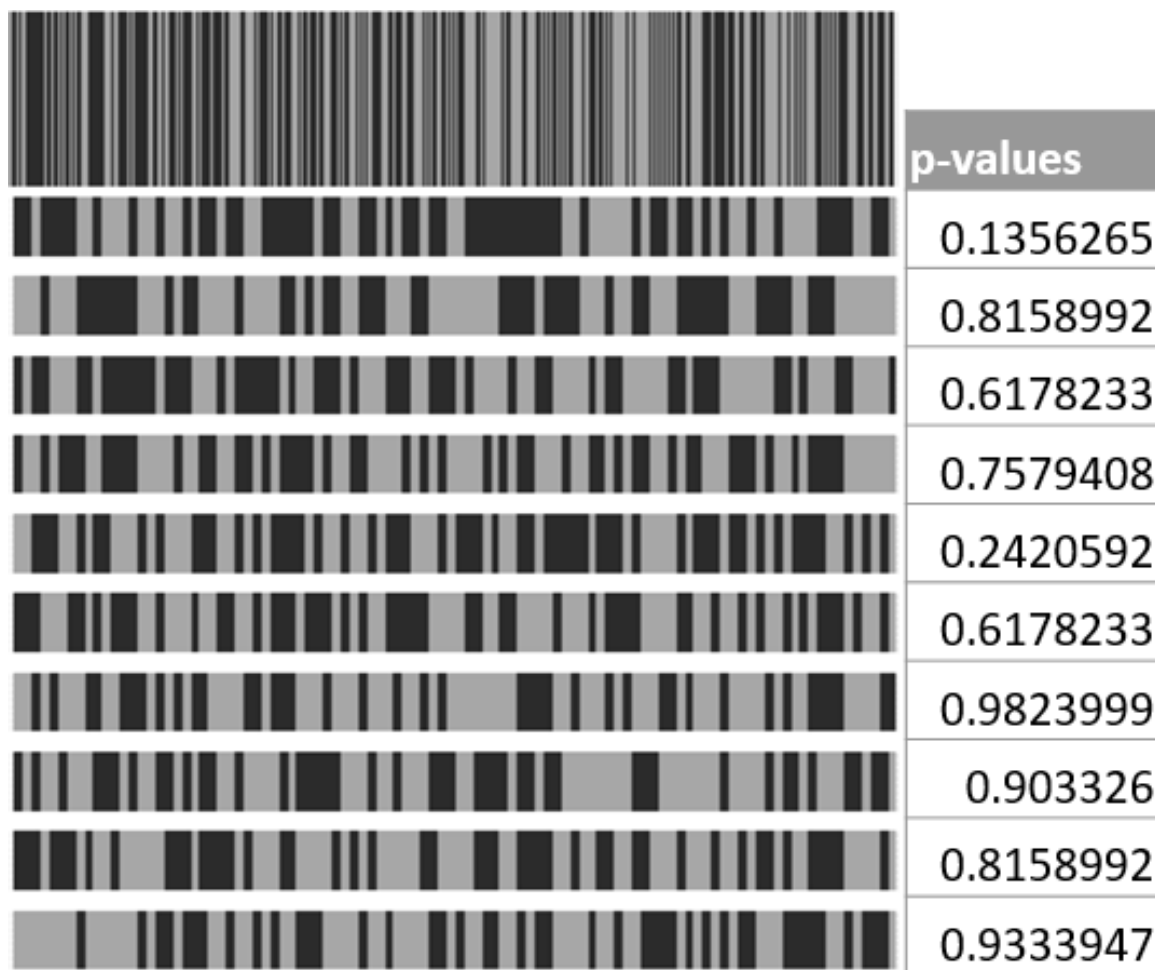
Figure 28. Testing of random samples from a binomial distribution does not give p-values low enough to reject the null hypothesis of 50% win/loss probability.

The dataset utilized to form this methodology is from the random tournaments between two computer chess engines. The environment chosen for this simulation is WinBoard version 4.8 [125] and the engines chosen are two instances of Poseidon version

1.18 [126] with brain mode set to random. The configuration specified was set and played 1500 times, where each engine played for either white or black pieces until one of a series of outcomes is reached. The following chess-specific rules were enforced for each game:

- The "Fifty-Move-Rule" is in place to prevent a game from going on indefinitely. It comes into effect when either player makes fifty moves without moving a pawn or making a capture.

- A Checkmate is a board configuration where one player's king is in position to be captured without a legal move available to avoid the capture.

- The "Insufficient Mating Material" rule defines when a draw occurs and results from neither player having adequate pieces to deliver a checkmate based on known chess configurations.

- The "Repetition Draw" results from the same relative position occurring three times during game play. In this case, the same relative position is where there are identical legal moves and geometry.

- A Stalemate occurs when neither player has a legal move and is not in check.

- A "Trivial Draw" occurs when certain board configurations are present and is an approximation of the likelihood of a draw.

Figure 29 captures the end results of 1,500 simulated chess games. While the majority of outcomes are draws, we argue that they are an artifact of the game and the

random moves of the engines. However, a preference can be had regarding the ending

configuration of a game. To make a translation from chess to a military endeavors, a draw

where less troops and equipment are compromised is preferred to one where considerable

losses are incurred. Using this perspective, a preference was assigned to the games

depending on the available material left at the end of the game and in doing so, only 143

games of the initial 1,500 were irreconcilable. The remaining games were evenly split

between white win and black win with 680 out of the remaining 1,357 games for black and

677 wins for white. This aided in maintaining the perception of a fair game with

approximately equal probability of win for either side.
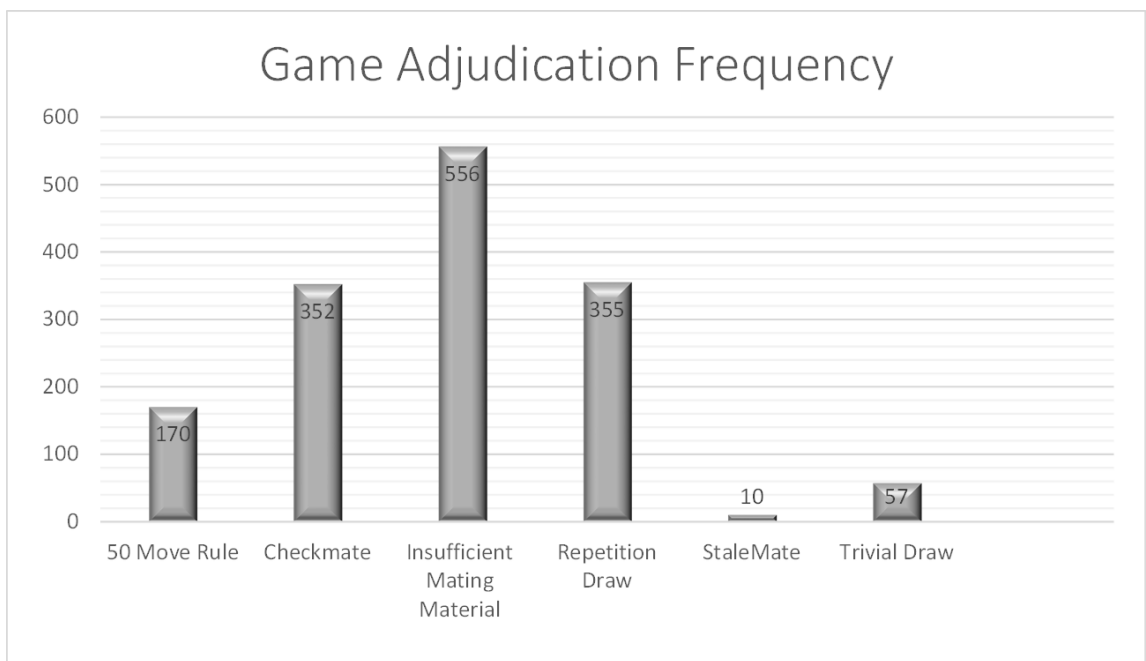
Figure 29. Game adjudication frequencies showing actual stalemate, the strongest form of stalemate in chess, as least frequent with other imposed draws as much more frequent.

109

Using DTW on the metric data for the 1,357 games, a distance matrix was made by performing pairwise comparisons between all of the games where the entry matrix was the DTW distance. From this distance matrix, a linkage algorithm was used to create a Dendrogram. By utilizing the Dendrogram, an assignment of cases to communities or clusters can be seen by imagining a cut made at a certain height of the Dendrogram. This represents viewing the dataset at different resolutions.

By viewing the Dendrogram at the highest level or the greatest Dendrogram height, all cases are in one community and inferences can be made on the community as a whole. By visualizing the Dendrogram at lower levels or lower Dendrogram heights, the single community of all the data breaks up into sub-communities of the data. These sub-communities are comprised of elements that are similar to their neighbors but dissimilar to elements in other communities. As lower cuts across the Dendrogram are made, clusters contain fewer elements as cluster outliers are removed from clusters, forming single element clusters, or singletons but become more homogeneous in the process. The level at which a cut is made on the Dendrogram means the lower limit of similarity is at the cut level. Therefore, non-singleton clusters that remain after making a cut at a certain Dendrogram height are as similar as or more similar than the cut level. This behavior increases as cuts are made further down the Dendrogram, however, a limit is reached when every case in the Dendrogram is a singleton, which can be seen in Figure 30.

110
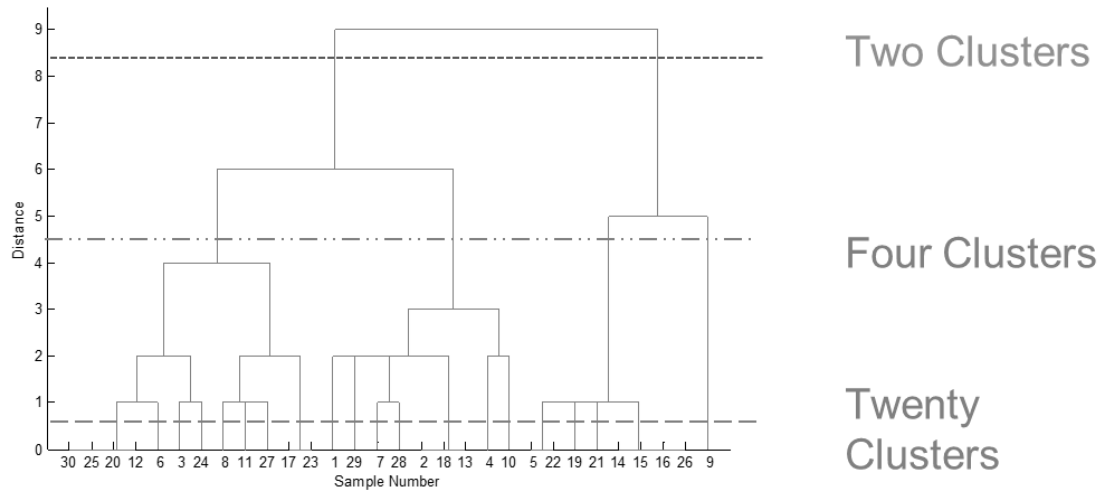
Figure 30 . Illustration of dendrogram and similarity showing the result of looking lower on the dendrogram impacting the number of perceived clusters with greater similarity.

The p-value exhibits opposite behavior to relative distance or similarity as cuts are made further down the Dendrogram. This is due to the reduction in cluster size. With all things held equal, as the cluster size is reduced, the p-value increases. This is because of the reduced statistical significance of smaller clusters. There is an increased probability of witnessing more extreme behavior when a cluster size is sufficiently small. Therefore, a measure to identify the relative goodness of a linkage algorithm, as well as the optimum number of clusters which to view and analyze the Dendrogram should include both the similarity and the average p-value associated with a particular cut. This index called the Height Significance Index (HSig Index) and is shown below in Equations (6), (7), and (8).

111

$$\text{HSigI} = \hat{\alpha} + \hat{H} \tag{6}$$

$$\hat{\alpha} = \frac{\vec{\alpha}}{\|\vec{\alpha}\|}, \quad \text{and} \quad \hat{H} = \frac{\vec{H}}{\|\vec{H}\|} \tag{7}$$

$$\vec{\alpha} = \frac{1}{n} \sum_{j=1}^{m} \sum_{i=1}^{NC} \left[p_{val_i} \cdot |c_i| \right]^j \tag{8}$$

The HSig Index is reached by the addition of both the magnitude of the average p-value across the cut of the Dendrogram and the magnitude of the height of the cut which can be seen in Equation (6). The magnitudes of the average p-value and height vectors are shown in Equation (7). The average p-value equation is shown in Equation (8), where the equation describes the double summation of the p-value and the cardinality of the current cluster. First, the average weighted p-value is calculated, where the weight is the proportion of the size of cluster I, written as the cardinality of c_i, to the size of the entire dataset, n, for a given cluster. This is repeated for all clusters, i, at the given cut height, j, and then for all cut heights, forming a vector of average weighted p-values of length m. For each cut height, the HSig Index value represents one data point. Across all cut heights, HSig Index is a vector, and plotted against the number of clusters, a unimodal behavior can

112

be seen as in Figure 31 due to the competing behaviors of the fall of significance and rise of similarity as the Dendrogram is cut further down.
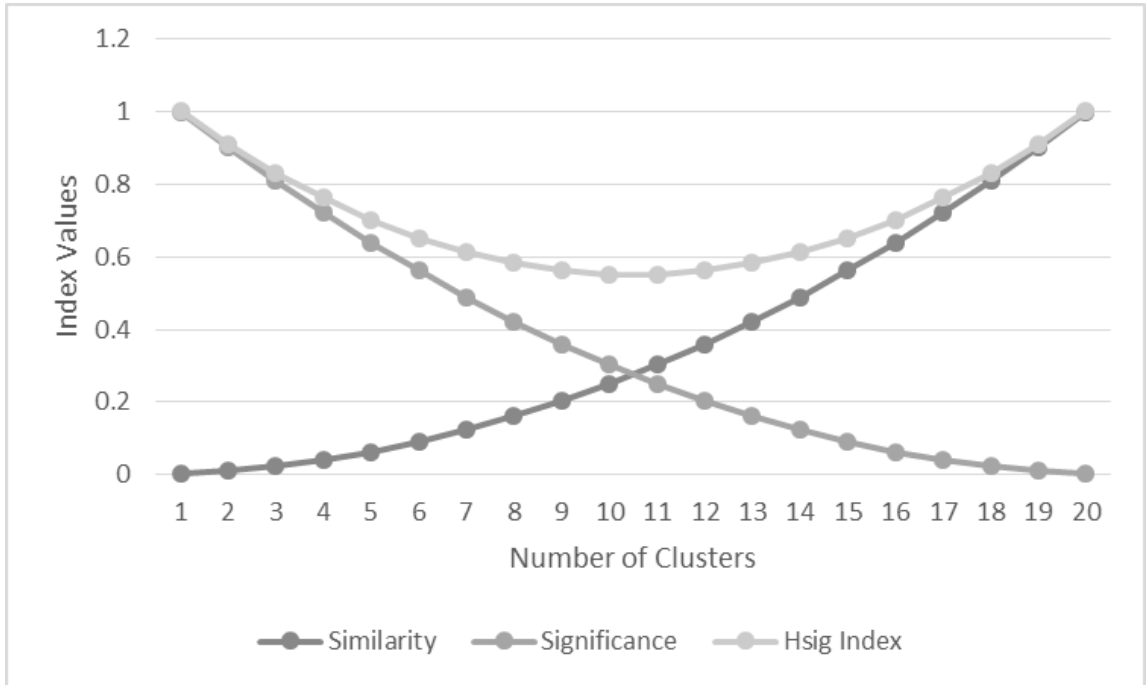
Figure 31. Behavior of Significance, Similarity and HSig Index forming a quadratic relationship where a minimum value of the HSig Index presents the best compromise of Similarity and Significance attainable at a range of heights.

The conflicting trends seen in the significance and the similarity result in unimodal behavior driven by the size of the clusters and the size of the data as well. Similarly, to the effects of cluster size on significance as discussed in the previous paragraphs, this phenomenon can be mirrored when considering a reduced dataset. With a sparse sampling of the decision space, the probability of witnessing more extreme behavior increases as the size of the data and the scale of enumeration is reduced. Additionally, as the richness of

113

the data is reduced, the quality of the hierarchy is reduced as well. This motivates the next hypothesis:

Hypothesis 3: The size of the data will impact the quality and significance of clusters found

Put simply, if there is insufficient data, the hierarchy formed will not result in clusters where the elements have sufficient similarly. Additionally, the clusters will not display significance with respect to the outcomes.

In utilizing the HSig Index, performing a single cut or calculation of the optimal number of clusters may not yield clusters that are beneficial for further analysis. This is because, at an optimum number of clusters, there will be clusters that can outperform the average significance value across all clusters, as well as clusters that cannot outperform the average, as the clusters found by any linkage algorithm will vary in significance. If additional cuts were performed on the data set, by viewing the cluster as a data set and considering a cut within the cluster, the effect would be an increase in intra-cluster density as shown in Figure 32. By requiring a higher density within a cluster, the net effect is a cluster that has elements that are more similar to each other.

114
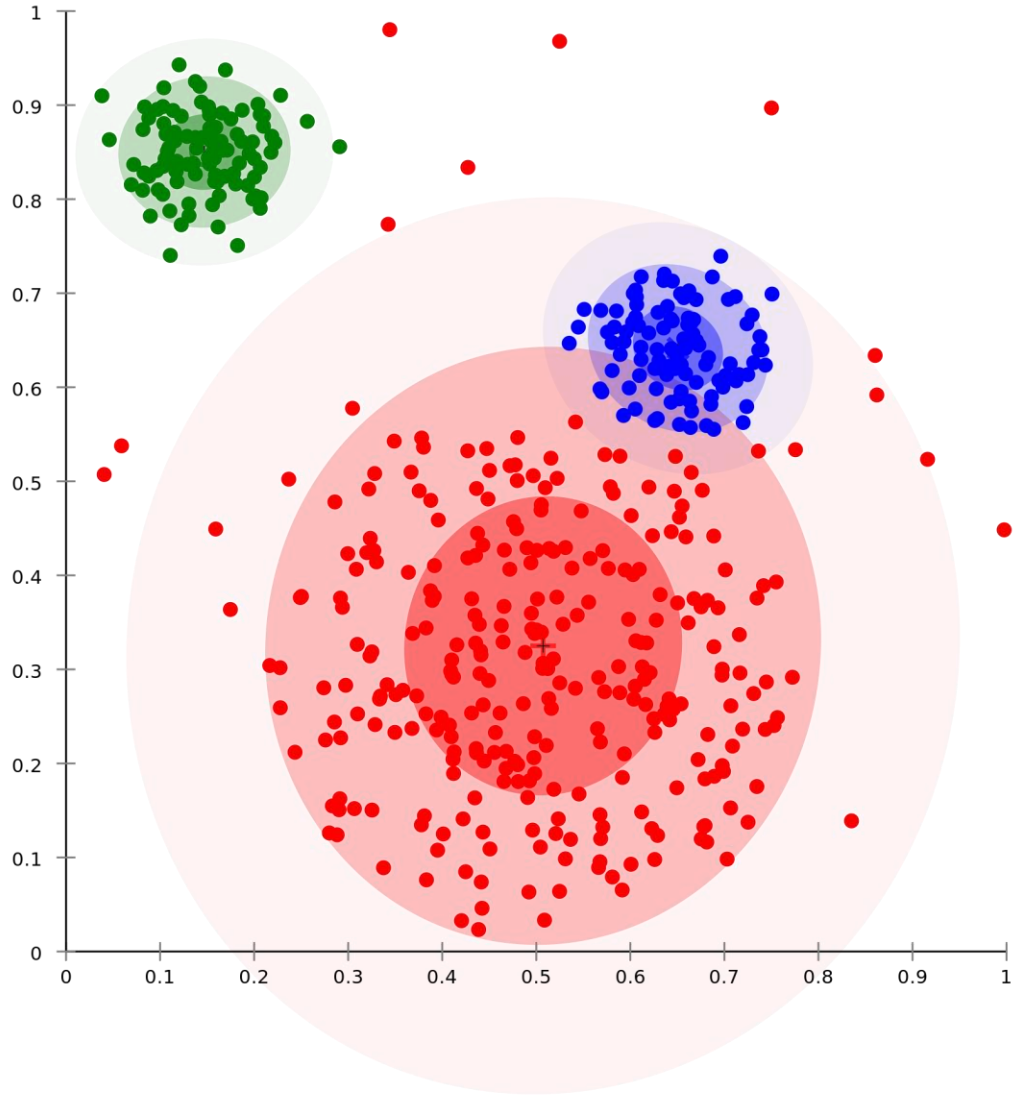
Figure 32. 2-D representation of cluster data showing the effect of expanding clusters effectively requiring higher density from sub-clusters by shedding outliers[127].

While viewing denser clusters enables identifying more similar elements, the benefit is easily seen when looking at the win percentages found in Figure 33. In this figure, an initial cut separating the data set into two clusters can be seen in the top plot. Viewing the data set at this level does not illuminate any increase in win probability. However, by making a cut lower down, making the clusters denser and changing the perceived structure, a cluster rises that appears to increase the probability of success.
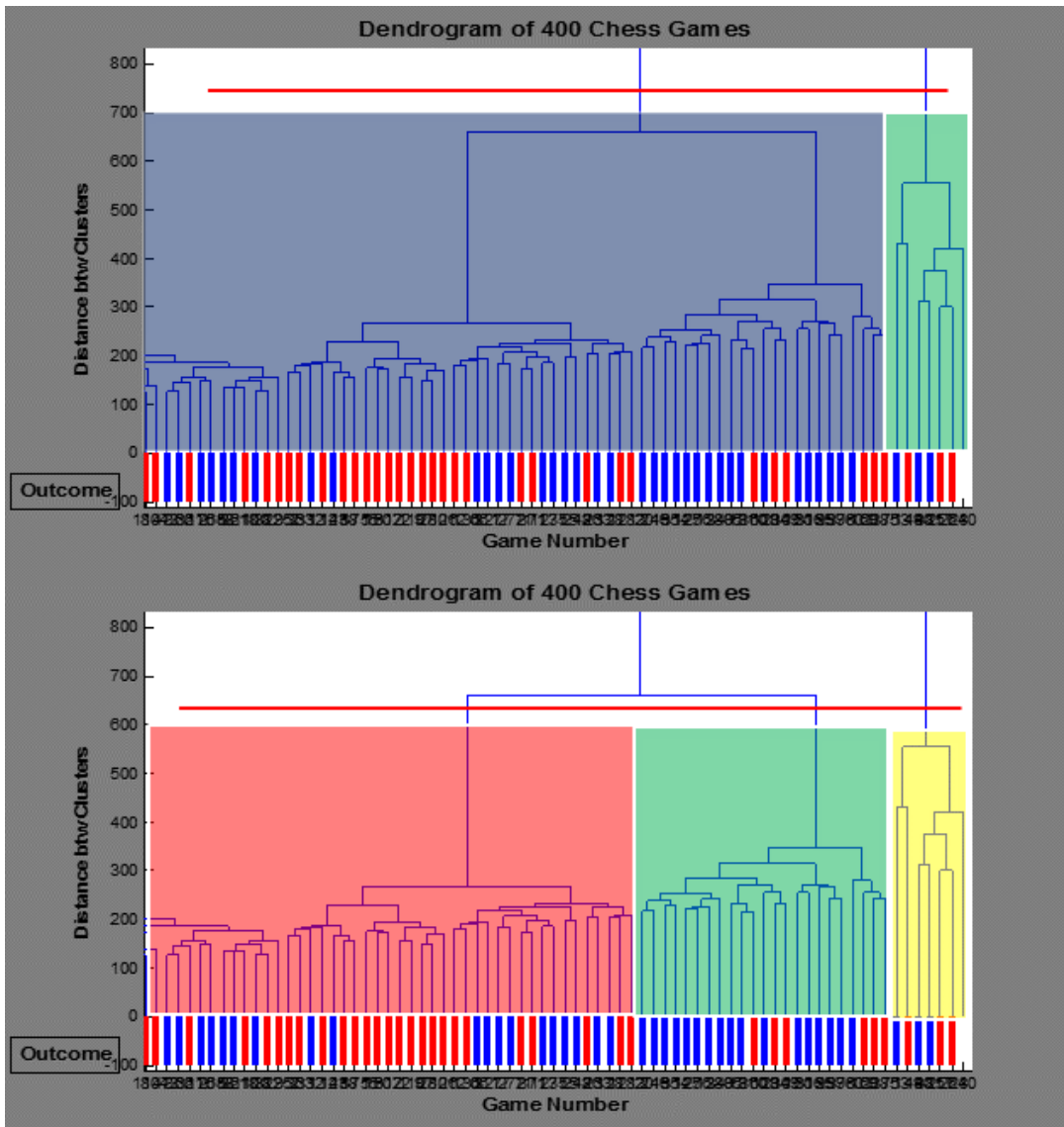
Figure 33. Dendrograms illustrating the effect of viewing deeper into the dendrogram on the ability to find clusters of higher density enabling the identification of clusters with an increased probability of success as in the green cluster in the lower plot.

117

Therefore, it is necessary to enumerate and expand all potential decision points within the hierarchy, only stopping when all nodes are bound. In the previous example, the hierarchy was not identified by making the first cut of the Dendrogram. However, this method will enable a multi-resolution view of the data set allowing the analysis of classes of COA that range in similarity, yet meet the significance constraints. This process ends with nodes pruned through assessment of optimality or through dissatisfaction of a constraint threshold. A visual example is presented below in Figure 34.
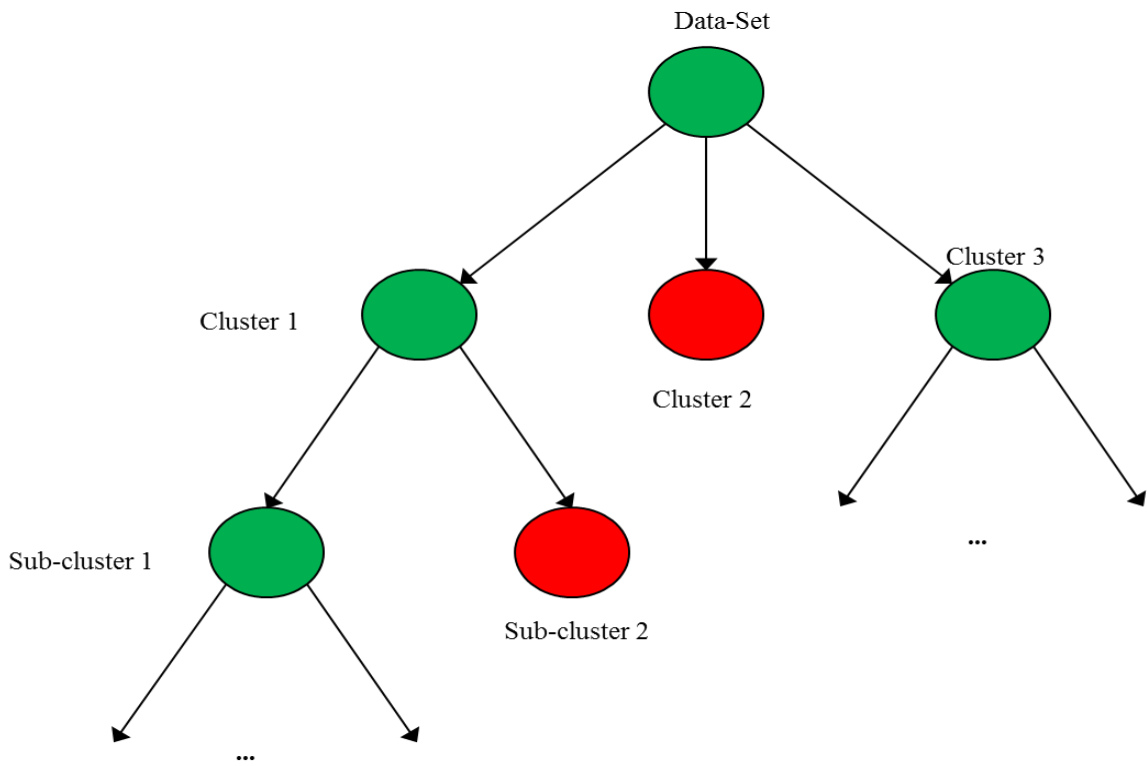
Figure 34. Bounding visualization using HSig Optimality.

Bounding the solution space by optimality results when each node is considered as an isolated dataset. Then the HSig Index is calculated, sweeping through the range of possible number of clusters, for each cluster or node in question. When the result of this process is an absence of nodes that outperform their predecessors, the previous node is known to be bound by optimality.

The other way in which a node can be bounded is through the use of significance threshold. This threshold for which to reject the null hypothesis must be set prior to analysis. This is usually set to 5% or 10%. In any case, if the clusters analyzed do not meet or exceed the significance value set, the clusters analyzed will be discarded from the analysis in the administration of bounding by threshold. However, this is necessary for bounding but not sufficient. The insignificant clusters will be subjected to HSig Optimization as well. This process is illustrated below in Figure 35.
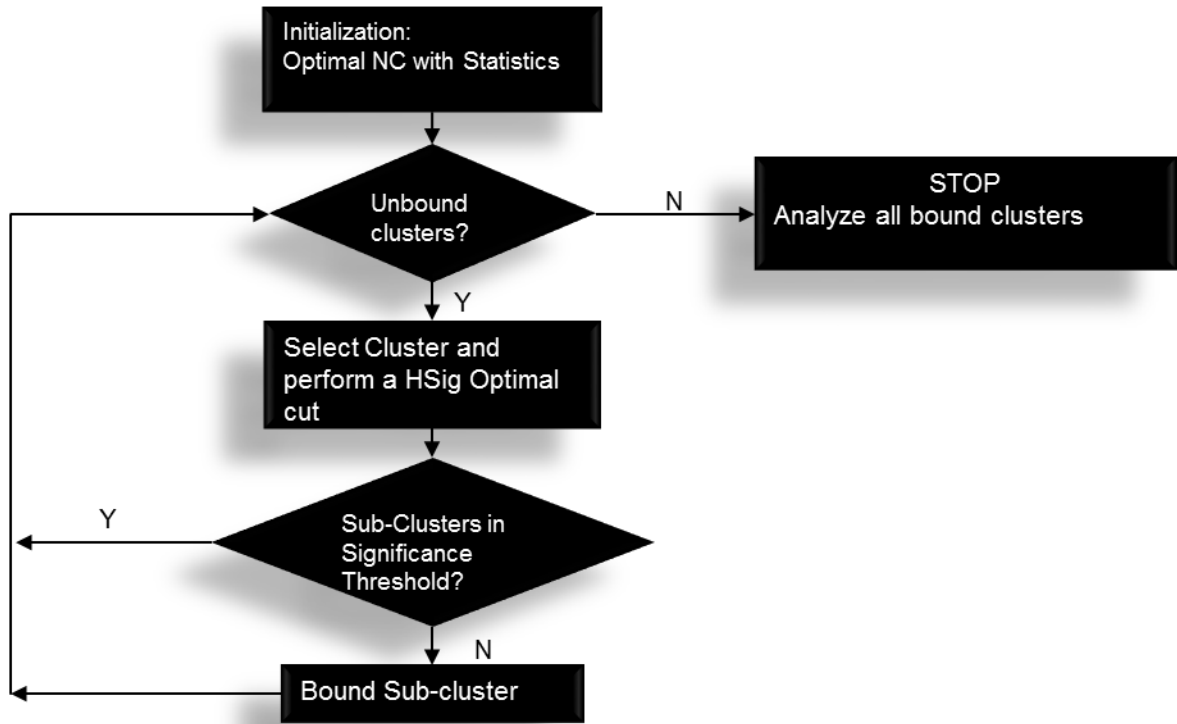
Figure 35. HSig Optimality algorithm process flow.

3.5   Formulated Methodology

As discussed in section 3.1, this formulation followed a general methodology born from a generic decision making process, the Generic Top-down Decision Making Process by Mavris et al. developed for Integrated Product/Process Development (IPPD) [64]. This process was utilized to fashion the aspects required of COA Generation, Analysis and Wargaming of time-series simulation data, in addition to final selection of a reduced set of

COAs and is illustrated below in Figure 36. The summarized steps follow however, the description of each step is expanded in the following paragraphs.

- Identify the metrics utilized for monitoring the battlespace
    o In Chess this is Position
    o In Combat modeling this is Position, Awareness and Capability
- Generate alternative COAs
    o Use random selection of COA elements to facilitate a random search of the decision space
- Analyze the alternatives generated
    o Obtain the pairwise distance between each COA using Dynamic Time Warping
    o Utilize each linkage algorithm to get a hierarchy for each linkage
    o Calculate the HSig Index for a sweep of 'Number of Clusters'
        ▪ For each possible number of clusters identify the height and average significance of all clusters at the corresponding level
    o Identify the appropriate linkage algorithm by finding the linkage algorithm yielding the minimum value of the HSig Index
    o Using the linkage algorithm and optimum number of clusters perform HSig Optimization
        ▪ Remove clusters not satisfying significance threshold
        ▪ Break clusters into sub-clusters using HSig Index
            • Bound clusters by significance threshold or optimality
    o Analyze clusters obtained through last step
        ▪ Analyze trajectories for trends and insights
        ▪ Identify equitable sequences of decisions

Initial provisions are made for establishing a method to determine value generically. With this methodology formulation in mind, these values are determined to be Capability, a traditional performance measure, in addition to the proposed metrics of Position and Awareness as well. Evaluating Position will enable the identification of prosperous plans from the perspective of the time varying physical location and the impact on overall performance. Measuring Awareness will allow the identification of the impact of decision elements on overall performance. In formulating the methodology, it was hypothesized that the observation of the battlespace will allow the enhanced identification of COAs as compared to traditional Capability metrics. Additionally, this observation is considered to enable the identification of conditions signifying a critical decision point.
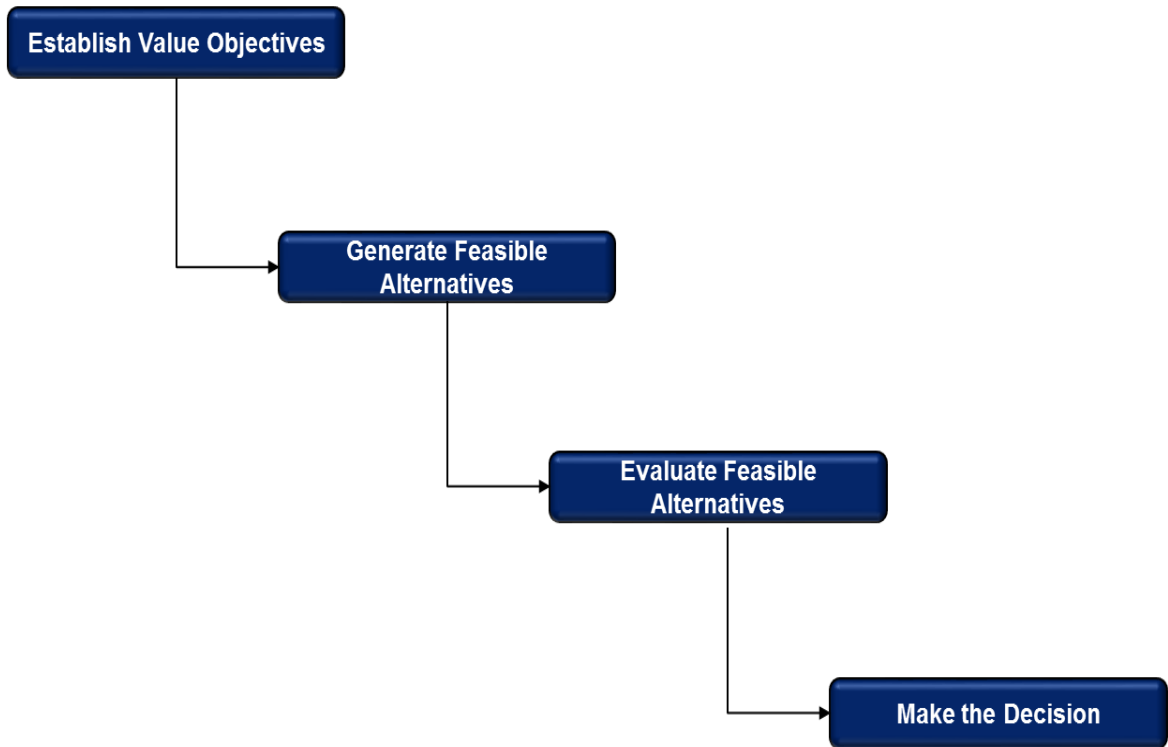
Figure 36.   General methodology for decision making used to illustrate flow of methodology formulation and as a framework for discussing the implementation here.

With value metrics defined, evaluation of the decision space, according to a battle planer, must be performed. However, a method to select portions of the decision space to analyze is needed. Presented in the section 3.3, this selection is motivated by MCTS methods used in game AI. Like in MCTS, identifying decisions by performing random sampling of the decision space, and then building a search tree with decision points and associated outcomes of streams of decisions is desired. However, this methodology needs a repository from which to generate random samples. This repository is the UJTL, and complete plans/COAs will be made by randomly selecting plan elements from a feasible

subset of the entire task-list. The number of times this is repeated represents the level to which the decision space is sampled. Each feasible COA represents one point in the decision space requiring evaluation, the majority of this chapter and overall effort.

Evaluating each COA generated requires a multi-step approach and is shown in the CADRE Methodology in Figure 37. First a method to simulate each COA must be identified. Wargaming is this method, and specifically, an ABM implementation of wargaming will be utilized. In performing simulation and collecting the value metrics defined earlier, a data stream capturing the implications and causality of choosing specific COA elements is obtained. However, as the complexity of the missions intended grows, so does the combinations of possible COAs and the required simulation to enable capture of decision implications. This makes the analysis of the data streams much too difficult to do manually because to determine decision implications, pairwise comparisons must be made to see how one decision differs from another. Therefore, the methodology requires assistance to identify patterns in this vast dataset. Machine Learning enables this identification.

Within Machine Learning, HAC was identified as being most suitable given the form of the data most recently defined in the previous paragraph. Additionally, DTW was selected as a method to allow the similarity to be calculated for many data elements by reducing each comparison to a single dimension, while maintaining the salient information

present in each simulation or game. However, there are many linkage algorithms to utilize in this form of clustering. This is why the next step of the algorithm is to identify the most suitable linkage algorithm. The most suitable algorithms will be so due to the degree to which it forms clusters containing the most similar data streams that lead to skewed outcomes for the cluster. In this best case, these clusters will be near homogenous from the perspective of loss or win, which is defined in this work as polarity. Therefore, to meet this objective, the methodology calls for the utilization of the HSig Index to assess the relative performance with respect to finding polarized clusters.

The HSig Index allows the identification of the relative performance of linkage algorithms as well as identifying the number of clusters giving the most polarized and crisp definition of the set of clusters for a given cut height. This index operates on three factors, Similarity, Polarity and p-value. The Similarity is assessed based on the cut height and signifies how closely elements within the cluster are joined, or more simply put, how dense a cluster is. As lower cuts are made on the Dendrogram, there is an implicit density constraint imposed on each cluster. This phenomenon is balanced against the Polarity and p-value associated with each cluster through hypothesis testing because, unchecked the density would grow infinitely as the cluster sizes approach one. The p-value is influenced by the Polarity and cluster/sample size.

Using this information, in addition to the success probability of the entire data set, a significance of the combination of Polarity and cluster size can be determined. In extreme cases, highly polar and large clusters have a very low p-value and are highly desirable. Additionally, very small clusters approaching size one will have a p-value equal to one. This p-value translates to the probability of witnessing more extreme, or more polar behavior then was seen in the cluster in question. This calculation is performed for each cluster as seen from performing a cut at some similarity, or height on the Dendrogram. The weighted p-value is determined for each cluster and is averaged for each cluster at a given cut. The product of identifying the HSig Index at each cut height will be a quadratic relationship showing an identifiable minimum. At this minimum, the optimum HSig Index and optimum number of clusters is identified. Performed on all linkage algorithms, the most ideal algorithm can be denied as the one attaining the smallest HSig Index value and analysis may commence utilizing the optimum number of clusters for that linkage algorithm given the optimum HSig Index value.

The first pass of analysis will require an evaluation of the significance of the clusters found at the optimum number of clusters cut height. This evaluation requires the definition of a significance threshold for which to allow evaluation of clusters. Typically, this value is less than 10% probability of witnessing more polar behavior in the data set. Therefore, after setting this threshold, the only clusters that will be analyzed in the first pass are the ones that are within the defined threshold. The clusters not satisfying the

126

significance constraint will be passed through HSig Optimization to identify if there is a cut height where they will satisfy the significance threshold, effectively shedding outliers in the cluster until sufficient polarity and density is attained. Additionally, clusters satisfying the significance threshold will be subjected to HSig Optimization as well, essentially forming an array of clusters to continue analysis. This process will continue until there are no clusters that satisfy the significance threshold.

With the array of satisfactory clusters in hand, the analysis will commence, where visual analysis will be performed on the metric time series to establish what the most salient features are in each cluster and if the win/loss can be attributed to those salient features. Assuming there was a sufficient investment Generating Feasible Alternatives, and Wargaming portions of the methodology, there will be classes of decisions yielding different probabilities of success that can be compared and contrasted. The differences in the classes will attribute the outcomes of each. For a set of cluster that are similar, but yield different probabilities of success, a CDP can be found. In this case, cluster features which are the difference in COA elements, that diverge between clusters will signify the actual CDP. It will show the branch in planning that leads to eventual win or loss. For the remaining data, there will be a reduced set of COAs to facilitate further analysis. In this case, winning COAs can be analyzed further for pursuit and losing COAs can be used to assess vulnerabilities, or to avoid outright. In any case, the output of this methodology is

a reduced set of COAs to analyze that was obtained from a rigorous search of the decision

space as well as CDPs.
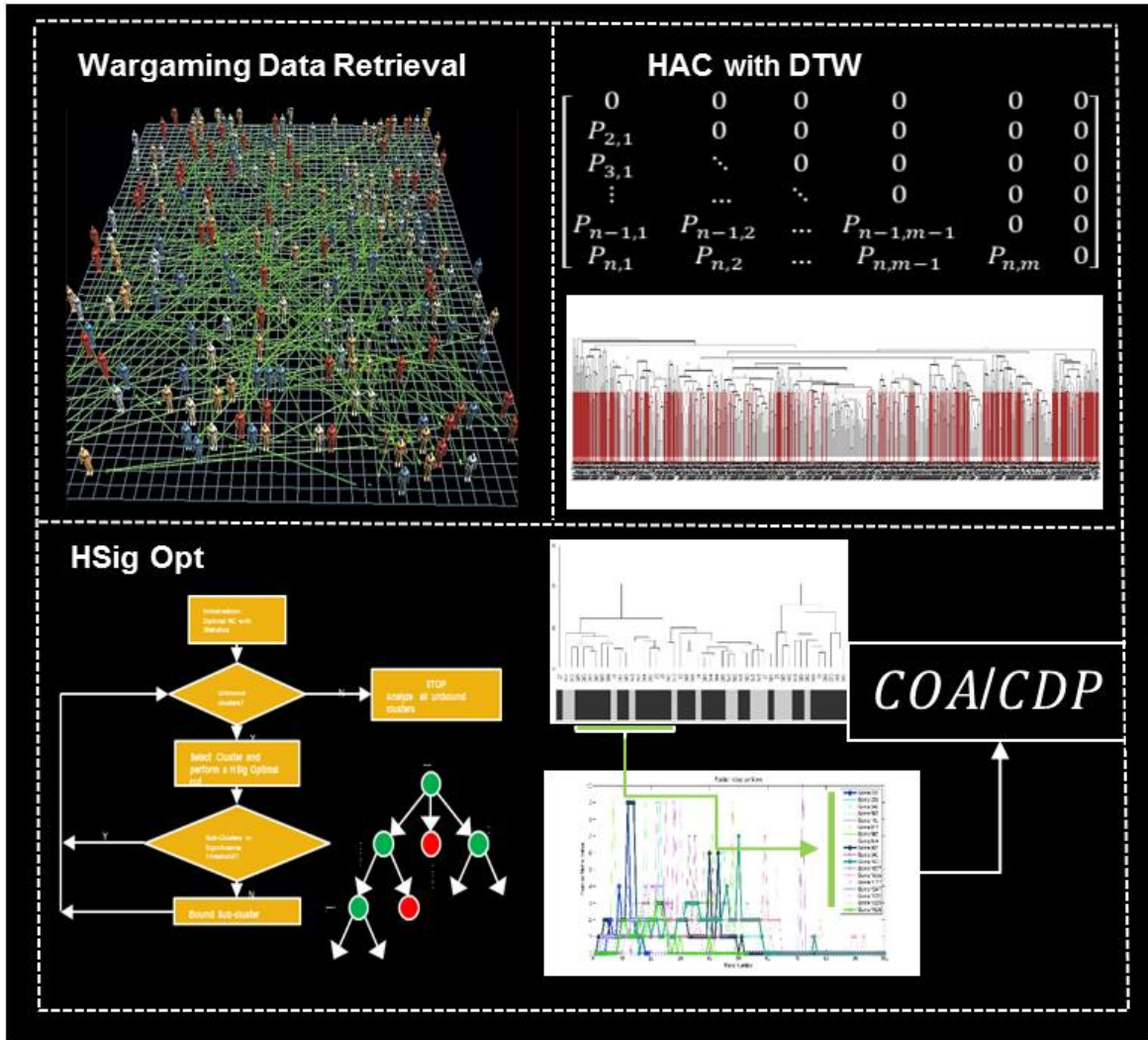
Figure 37. CADRE methodology including steps for data retrieval, comparison, analysis
and the determination of candidate COAs and CDPs to facilitate further analysis.

CHAPTER IV

EXPERIMENTAL PLAN

In order to evaluate the logic and assumptions used to formulate this thesis experimentation must be conducted. An experimental apparatus is required to test the hypothesis formulated in this thesis. To serve the needs of experimentation, three M&S suites are utilized to obtain wargame data. As a canonical problem the formulation of a reduced simple chess engine and full size simple chess engine is developed. Next a full scale agent based wargame suite, Map Aware Non-Uniform Automata (MANA), is described and will be utilized as a real world example problem showing the methodology.

The experimentation aims to both complete the methodology and to test hypothesis about the scalability of the methodology. Two sets of experiments have been formalized and carried out to achieve this end. The first is to fill the capability gap needed to identify a suitable linkage algorithm. This is then tested in the face of increased scale. The next set of experiments are for the applicability of selecting CDP's using HAC, and the behavior in the face of scaling as well. The methodology will then be utilized in a full scale agent based wargame, where the applicability will be illustrated and the generation of battlefield "rules of thumb" will be demonstrated.

4.1 Modeling & Simulation for Data Retrieval

In the pursuit of identifying a methodology to allow computerized COA analysis and wargaming, a modeling and simulation environment must be acquired. Due to the complexity of actors, systems and the interactions in the battlespace, in addition to the adaptiveness of the adversary, the modern battlefield is represented well as a complex adaptive system (CAS) [85].

Agent based models (ABM) have been utilized to study CAS and are well suited to the task due to the adaptability of ABM's. Attempting to simulate a real world system using intelligent agents enables the creation of an artificial world comprised of real world systems. The dichotomy of real systems in an artificial world allows the analysis of complex systems which can be approximated by autonomous agents and modeled elements [84]. While ABMs are well suited to simulate combat operations, applying ABMs presents several challenges. The ABMs in use within and across different military services can be of varying fidelity, complexity, and availability. While the primary goal of this research is to develop a methodology to aid in COA development for military operations, it may be necessary to utilize a canonical problem at first for methodology formulation. This canonical problem should have enough complexity as to avoid being trivial in comparison, yet should be a more constrained and well-understood simulation to facilitate hypothesis testing and verification.

4.1.1 Chess as a Canonical Problem

With the previous descriptors of ABM, computer chess can be thought of as an agent based formulation of a wargame. Similar to an ABM, the board represents a grid where action is localized and bounded. There are a range of chess pieces that have specific behaviors as dictated by the rules of chess and the logic of the different chess engines. Each type of chess piece can be thought of as a different agent type, with each chess engine dictating different and complex behaviors. Additionally, these chess pieces move and interact with other pieces in the environment with the goal of physically overtaking an adversary's pieces. These movements and interactions can be recorded using appropriate movement metrics.

While there may be a smaller and more constrained number of different pieces and movement patterns when compared to actual or simulated combat operations, the true complexity of the game of chess is revealed in the large branching factor of possible moves as the game progresses from opening to close [13]. The game of chess, when used as a canonical problem presents an adequate test-bed for creating the COA development methodology because it is a well-studied problem with references on good potential COAs, tactics and strategies. Specifically, inherent in the formulation of a chess engine is the accomplishment of the elements of the methodology relating to generating alternatives. For example, a COA as translated to the chess analogy is simply the list of move sequences from opening move to checkmate for a given game or simulation. Given this sequence,

the state of the board for each move can be recreated allowing calculation of the metrics of interest in addition to the overall outcome of the COA. The end result of a chess game is a multi-dimensional vector consisting of move sequences, and the associated outcome for one game. By repeating this process, a database consisting of multi-dimensional vectors is created which facilitates the analysis methods presented here. Equation (9), below, is the method used to identify the Position value of a player's pieces from one turn to the next for either side. In Equation (1) m is the number of either player's pieces remaining and n is the number of legal moves available to the player's piece in question. Simply put, this calculation is the sum of values associated with the enemy pieces that are in striking range of each opposing side's piece. In this case, if a piece can attack multiple enemy pieces by performing one legal chess move, it will be the sum of all of the values in Table 7 of the associated enemy pieces. This is then repeated for all of the remaining chess pieces and the total is summed over all pieces for each side per turn. High position values are reached when multiple friendly pieces have multiple opportunities to attack.

$$\text{Position Value} = \sum_{i=1}^{m} \sum_{j=1}^{n} \text{PieceValue}_{ij} \tag{9}$$

132

Table 7. Chess pieces and values utilized to calculate the position metric.

Symbol	♟	♞	♝	♜	♛	♚
Piece	Pawn	Knight	Bishop	Rook	Queen	King
Value	1	3	3	5	9	10

The dataset under consideration is from the random tournaments between two computer chess engines. The environment chosen for this simulation is WinBoard version 4.8 [35] and the engines chosen are two instances of Poseidon version 1.18 [36]. There are a range of pieces that have specific behaviors as dictated by chess rules and logic of the engines motivating agent behavior. Each piece type can be thought of as different agent types. Additionally, these pieces interact with other pieces in the environment as they move and alter the values of the CAP metrics as well as by physically overtaking adversary pieces. Similar to an agent based model, the board represents a grid where action is localized and bounded. However, the complexity in the diversity of actors and expected interactions is considered much less in chess due to the lower branching factor [128]. Although, due to the back and forth nature of chess, the expected combinatorial vastness

of the decision space is expected to be much more complex than planning conventional warfare.

Through monitoring the decision sequences (COAs) and the Position metric expressed and defined for chess above, the Methodology may be applied. As in the previous chapter, the main objective is to use Dynamic Time Warping to find similar COAs by comparing the Position metric time series of each simulation. Below in Figure 38, two example COAs are provided. They represent the time varying nature of the time series they represent and influence. The decisions made can impact the brevity or protraction of a combat maneuver, or chess game in this example. This further reinforces the selection of DTW to determine the similarity of runs so that they may be clustered.
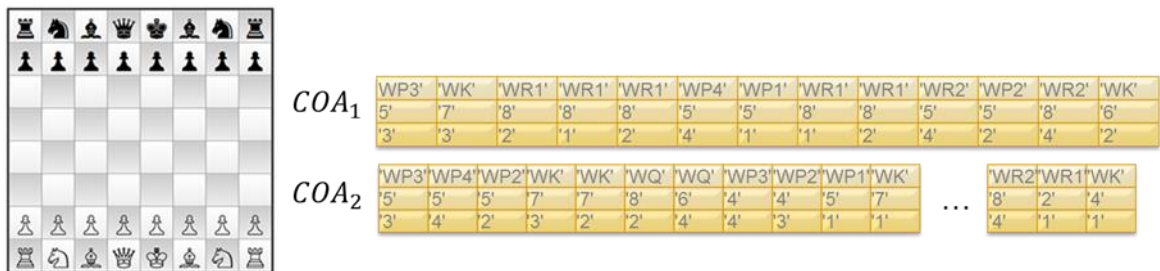
Figure 38. Example COAs using Chess to show a COA in Chess is a list of moves from beginning to end game that can be evaluated through analyzing the board configuration.

While the sequences themselves represent COAs, the metric time series will be monitored and analyzed to allow advantageous COAs to be identified. Using Equation (9),

134

the value associated with each move is calculated and stored. The time series are of great importance when comparing runs. By plotting the Position metric time series, we may see that each game has a 'signature'. As in Figure 39, there appears to be a behavior among the games shown. This behavior is in the similarity of moves, however is the product of comparing the actual trajectories of each game to others in the data set. Each game is replaced by a vector and is compared to other games using DTW to arrive at a similarity matrix as sown in   Figure 13.
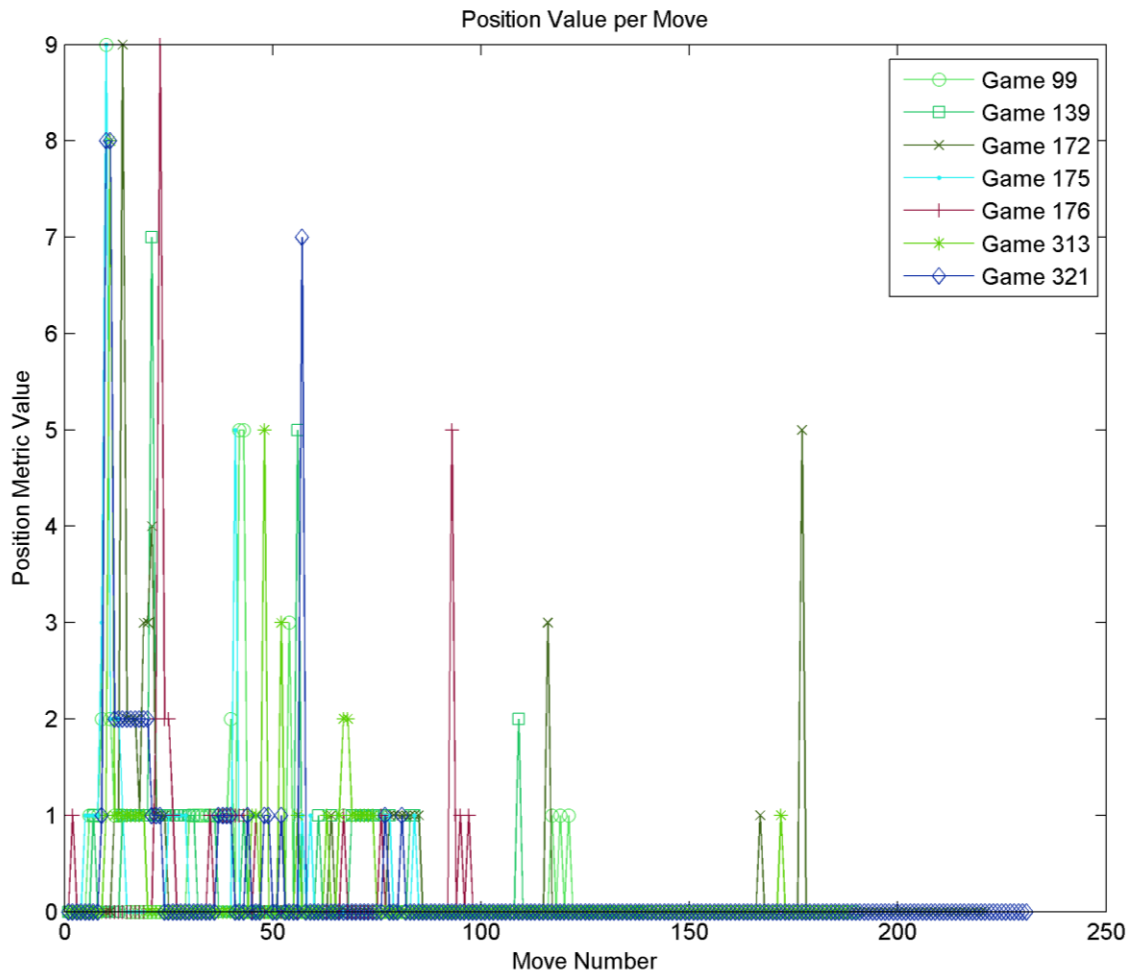
Figure 39. A series of chess games position metric time series illustrating an observable behavior in attack sequences among games shown.

For the engines described, the behavior and sets of possible interactions are much less than what is expected in a true agent based formulation. For example, in chess, there are approximately 30 moves possible at every turn which remains somewhat constant until

the endgame where there are just a few pieces left [129]. This is in contrast to an agent based combat simulation which can have thousands of agents acting almost continuously each time step. However, in a full board chess formulation, the decision space is combinatorial as well as, the number of possible moves in a total game is approximately 10^{120} from initial move to checkmate [129].

4.1.2 Map Aware Non-uniform Automata (MANA)

Map Aware Non-uniform Automata (MANA) is an agent based model and has been identified as complimentary to the objective of this thesis. As an ABM, strengths are met as a result of a trade off with detail. Under this paradigm, details that are not essential to determining behavior in a scenario are omitted in favor of decreased computational complexity facilitating a more encompassing exploration of the battlespace. Additionally, the relatively simple interactions defined between elements in an ABM allow the observation of complex macro level behaviors due to the non-linearity present in CAS. As researchers wanted to analyze the effects of awareness on the battlefield and net-centricity, the predecessors to MANA were developed in place of existing physics based models. As a result, MANA has been used for numerous military OR studies to include modeling of civil violence, maritime surveillance, studies of NCW and counter-IED measures, and other investigations of modern warfare as a complex adaptive system.

MANA is an agent-based combat model and is specifically characterized as an agent based, distillation intended to model the "essence" of a particular problem [130]. In this way, MANA is primarily used to create a bottom-up abstraction of a scenario to capture the global behavior. Its power lies in the emphasis on leaving out detailed physical attributes of the battlefield entities in place of a parametric characterization. With this parametrization, users can easily create many types of entities with different performance characteristics easily and effectively. For example, the difference between a mobile ground unit and one on foot will include less armor and relative capable speed for the ground unit on foot as compared to the mobile unit which has the cover and of armored vehicles [130]. Additionally, these diverse battlefield elements can be easily applied to different battlefields or locations by implementing different maps in the form of bitmap images as seen in Figure 40. In this way, the makeup of a force developed parametrically, can be assessed in various environments. Some of the advantages and disadvantages are enumerated below [131].
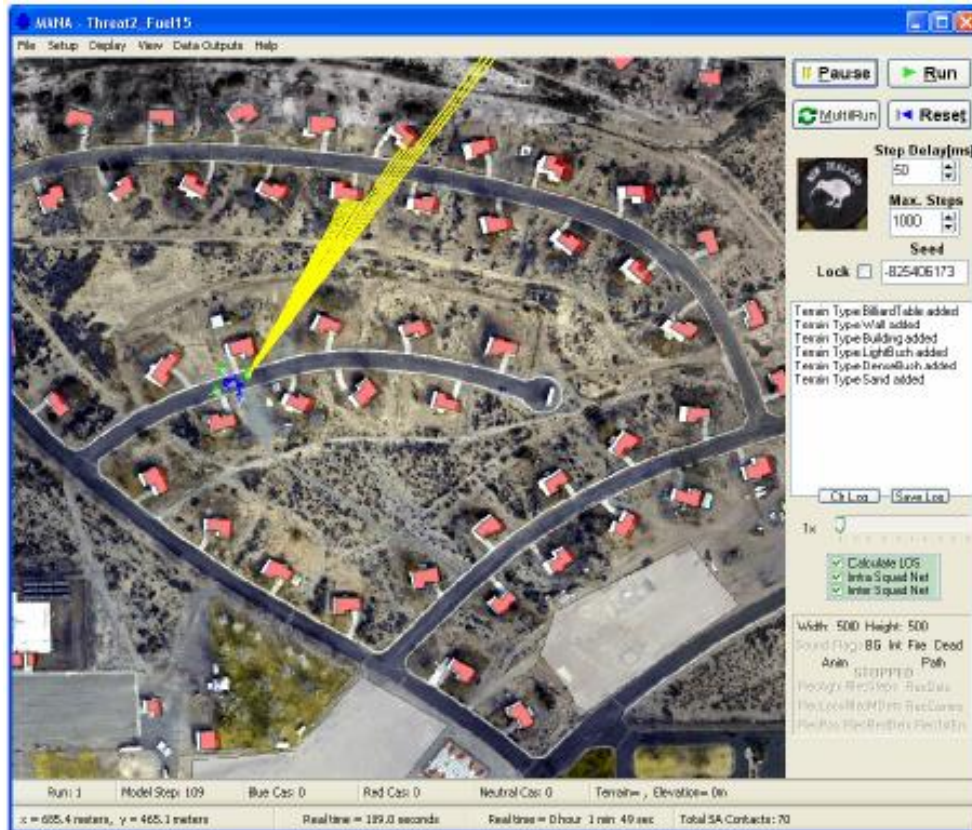
Figure 40. Background image in MANA showing real world terrain use [131].

Advantages:

- Still in development

- Large OR user community

- Vector based movement

- Parametric agent definition

139

- Import/modify battlefield data (elevation and terrain)

- Enables large range of scenarios with short set-up time

- Includes non-attrition-based combat model

Disadvantages:

- Doesn't cover the spectrum of military behaviors (can't represent actual fighting formations)

- Terrain/battlefield data is draining on computer memory

- Entities don't always make the "correct" decisions

Due to the ability to represent spectrum of capability in a range of operating environments and scenarios simultaneously, MANA appears complimentary to the stated research objective and the desire to perform analysis on wargaming data to facilitate decision-making. However, the published shortcomings of the model may inhibit aspects of the analysis. Specifically, if the military behaviors are too simple, requiring abstraction to gain rules of thumb, action sequences may not be legible. The coherence in which entities act will also affect the analysis. Specifically, care must be taken to acquire sufficient run data to ensure, agent behavior does not dominate the lessons learned through analysis.

4.2 Experimentation

This section will expand on the experimentation required to complete the methodology and to test the hypothesis posed in this thesis. The first experiment, Experiment 1, will be performed to test Hypothesis 2 which states that it is possible to select a linkage algorithm that creates a hierarchy offering an increased within cluster probability of success by utilizing the Height Significance Index. The next experiment, Experiment 2, will be conducted to investigate efficacy of utilizing HSig Optimization to select a COA that offers a significant increase in the probability of success. This experiment will be conducted to test the validity of hypothesis 1.0, 1.1 and 1.2 stating, COA identification is possible through the use of CAP metric sequence data in addition to traditional MOEs and MOPs to identify CDPs. Experiment 3 will be conducted to test Hypothesis 3, which states the utility of the analysis is heavily influenced by the depth and breadth of the data analyzed.

4.2.1 Experiment 1

Experiment 1 will be performed to test Hypothesis 2 which refers to the possibility of select a linkage algorithm that creates a hierarchy offering an increased within cluster probability of success by utilizing the Height Significance Index. As in the development of the HSig Index, this experiment will utilize statistical hypothesis testing to prove or

disprove Hypothesis 2. Additionally, this experiment will establish a baseline for the methodology as well as to determine the correct linkage algorithm. The primary steps involved in this experiment are enumerated in the list below:

- A data set will be created

- Linkage algorithms will be utilized to generate hierarchies

- HSig Index will be utilized to assess each linkage algorithm

- For the optimum linkage algorithm, the optimum number of clusters will be found

With respect to the Generic Top-down Decision Making Process, these steps can be seen in Figure 41.
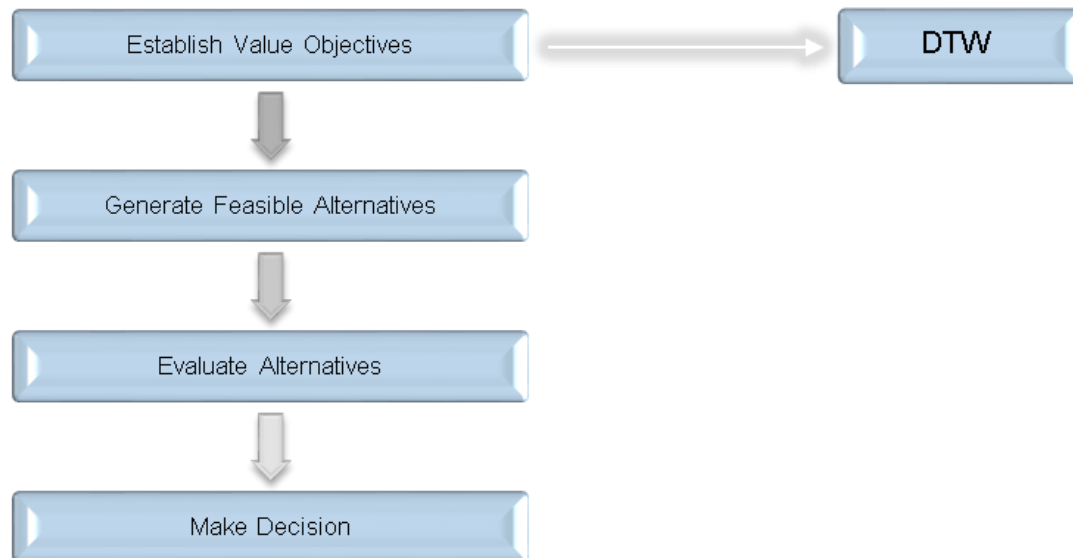
Figure 41. Experiment 1.1 decision support process and flow of work highlighting the selection of dynamic time warping for evaluating value.

142

The first step according to the process is the generation of alternatives and is illustrated in Figure 42. The HSig Index will be utilized to measure relative goodness of each of the linkage algorithms. For this experiment, the Poseidon chess engine with brain mode set to random will be utilized on WinBoard 4.8 in order to create a wargaming database of alternatives for analysis. Through this database, similarities in game play will be identified using the Dynamic Time Warping algorithm. Next each linkage algorithm will be applied to the proximity matrix to arrive at corresponding set of clustered data alternatives. After generation of the alternatives, evaluation will commence by analysis of the HSig Index values of each linkage algorithm selection for a range of number of clusters. By analyzing the plots of this behavior, the relative performance of each linkage algorithm will quantitatively assessed, in addition to the optimum number of clusters being defined.
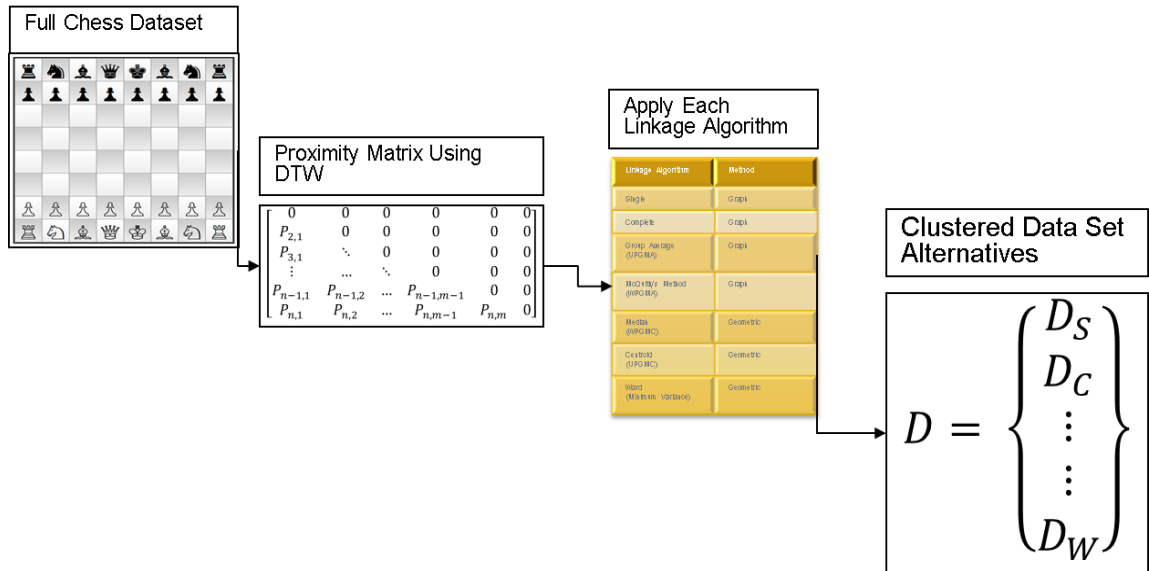
Figure 42. The Process of Generating Alternatives for Experiment 1.1 beginning with simulation using Chess and resulting in a set of clustered data alternatives.

After HSig Index resolution, the alternatives that were created previously will be the evaluated. Final selection will be based on the optimum value HSig Index as seen in Figure 43, where the minimum value obtained from utilizing all represented linkage algorithms will be identified as the best in class measure with respect to the data set on hand.
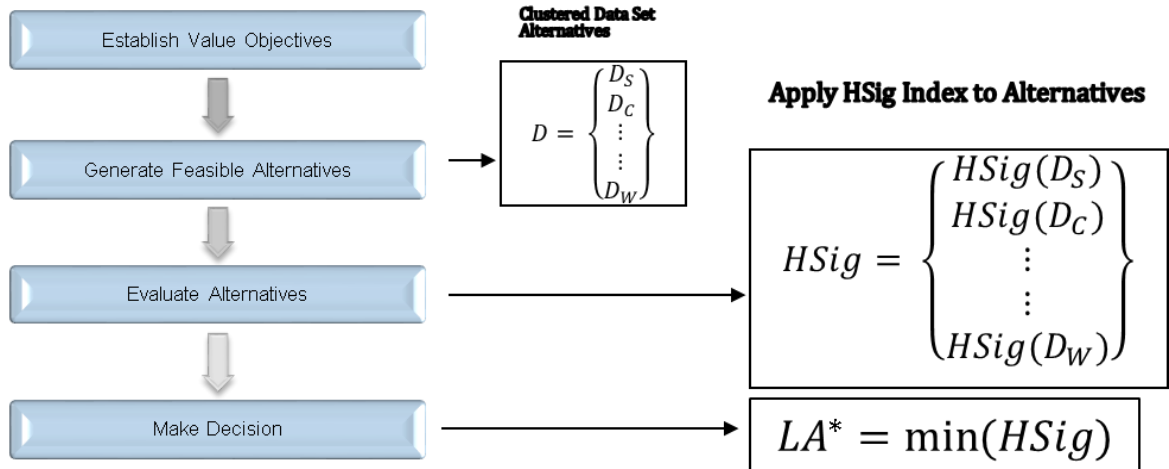
Figure 43. Process of making the final selection of linkage algorithm utilizing the HSig Index and where the best performing algorithm has the minimum HSig Index value.

4.2.2 Experiment 2

With the Linkage Algorithm selected, CDP's must be identified from within the Dendrogram. Using the chess M&S suite of WinBoard and Poseidon, data will be obtained and clustered facilitating CDP analysis. Experiment 2 will be conducted to test the validity of hypothesis 1.0, 1.1 and 1.2. It states that COA identification is possible through the use of CAP metric sequence data in addition to traditional MOEs and MOPs to identify CDPs. Also stated is that CDP identification from move or action sequence alone is possible using a wargame data resulting from a relatively low branching factor process.

Through analyzing the Dendrogram, areas of increased win density can be identified as can be seen below in Figure 46. Further analysis of those areas and the relative distance or similarity can aid in identifying a common node with increased win density and regions of decreased win density. As a result of the formulation of the HSig Index, a single cut or calculation of the optimal number of clusters is necessary but not sufficient to give the best chances of arriving at a best in class policy for COA adoption. This is because, at an optimum number of clusters, there will be clusters that can outperform the average significance value across all clusters, as well as clusters that cannot outperform the average, as the clusters found by any linkage algorithm will vary in significance. Therefore, it is necessary to enumerate and expand all potential decision points within the hierarchy, only stopping when all nodes are bound, a process that ends with nodes pruned through assessment of optimality or through dissatisfaction of a constraint threshold. A visual example is presented below in Figure 44.
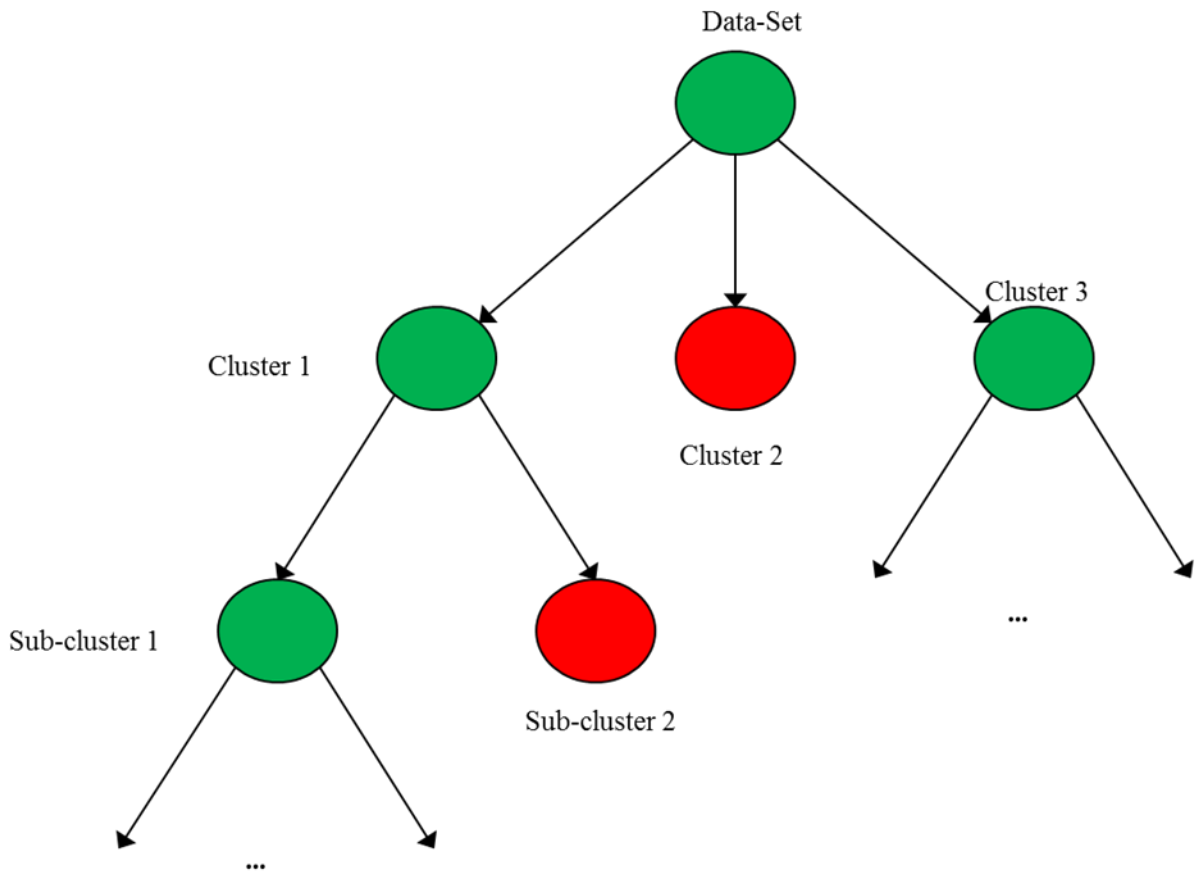
Figure 44. Bounding HSig Optimality visualization motivated by the necessity to expand all clusters and sub-clusters within the hierarchy bounding along the way.

Bounding the solution space with regard to optimality will result when each node is treated as an isolated dataset and the HSig Index is calculated for a range of number of clusters with respect to the cluster or node in question. When the result of this process is an absence of nodes that outperform their predecessors, the previous node is known to be bound by optimality. Additionally, bounding the solution space can also come by way of

147

significance threshold.   Prior to analysis, there must be a set significance threshold for

which to reject the null hypothesis.   This is usually set to 5% or 10%.   In either case, if no

clusters remain that meet or exceed these thresholds, then the remaining nodes are seen to

be bound by threshold.   This process is illustrated below in Figure 45.
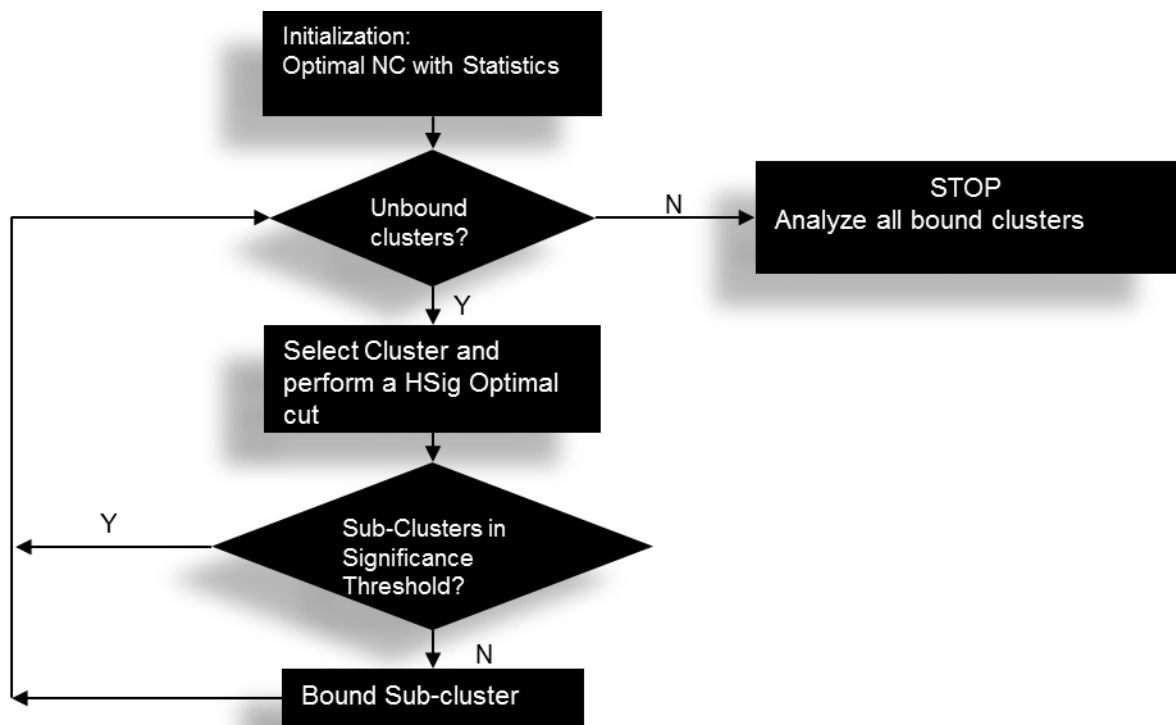
Figure 45. HSig Optimality algorithm flow chart, initiation and bounding criteria.

This motivates the use of HSig Optimization, where the most significant clusters are analyzed further to uncover an equitable COA. Identifying through the hierarchy, where the differences lie between an equitable and unequitable COA will reveal a node defined as critical decision point.
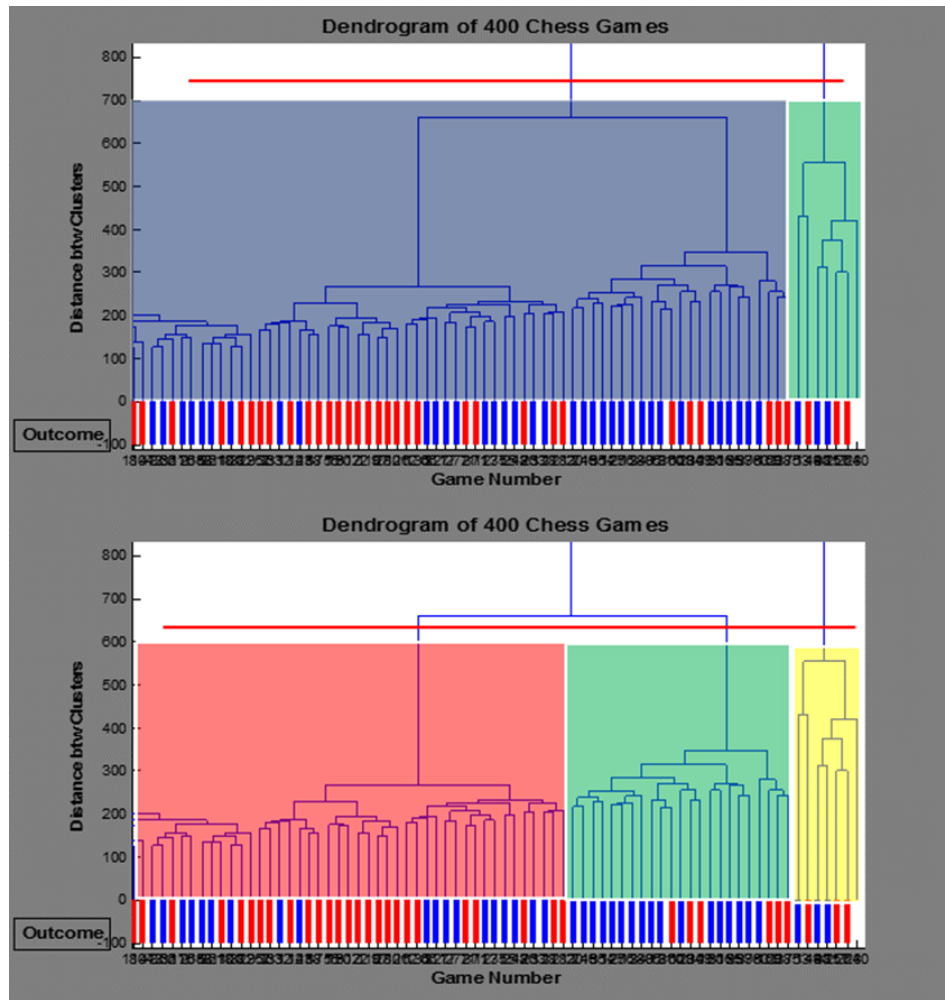
Figure 46. Chess move sequence dendrogram analysis for CDP identification illustration.

149

In order to verify the hypothesis, qualitative analysis of the data and Dendrogram must be performed to identify the CDPs. Additionally, identification of well-known chess heuristics from within the data and CDP analysis will serve as an indication of verification. Specifically, by identifying desirable COAs associated with CDPs, it is expected that this will show the behavior associated with position desirability preferences seen in the opening of a chess match as seen in Figure 47. In this figure, the desirability of board positions is plotted on a notional chess board. This plot illustrates the desire for center board control heuristic [129].
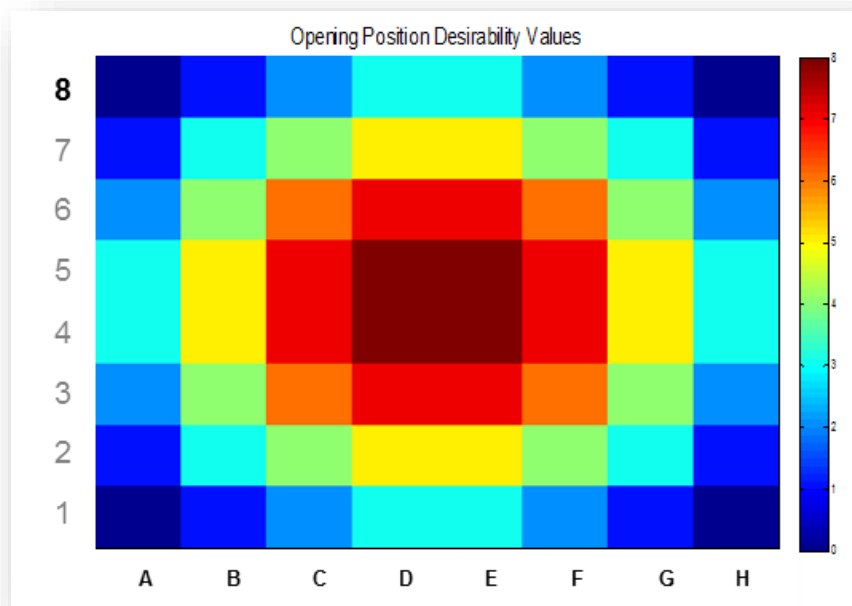
Figure 47. Chess board position desirability values showing center of board importance.

4.2.3 Experiment 3

The next experiment will be conducted to investigate what effect scaling the problem will have on the methodology. As the decision space grows due to an increased number of actors and actions to preside over, as must the investment in simulation. In cases where this investment is not possible, or where it is limited from the start Hypothesis 3 states that this will negatively influence the quality of the results. In this experiment, fractions of the initial data set are analyzed using HSig Optimization and the significant clusters are to be shown. This experiment will test data sets of 400 and 800 simulations to determine the effect on significance. These results will then be comparted to the results in the Experiment 2. The form of this experiment will parallel that of Experiment 2 as well and is depicted in Figure 48.
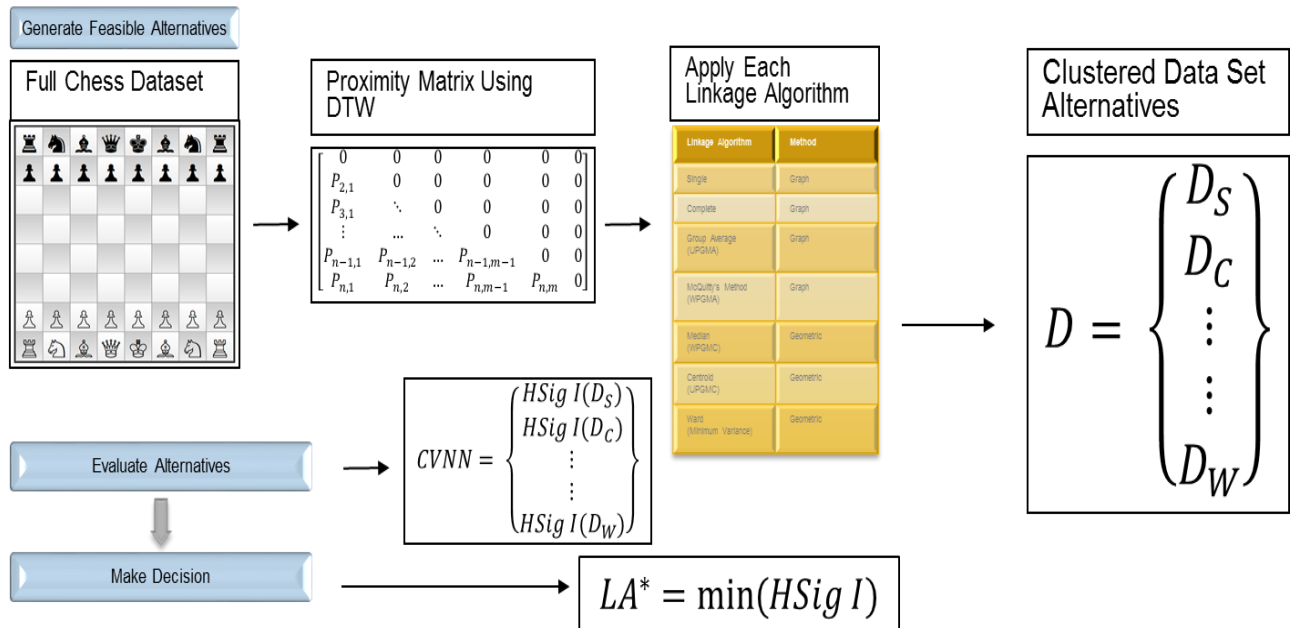
Figure 48. Experiment 1.2 process flow utilizing reduced portions of the chess data set to illustrate the effect of scaling on the methodology and subsequent decision output.

4.2.4 Methodology

Clustering Analysis for Decision-point REcognition (CADRE) is the methodology produced in this thesis and can be seen in Figure 49. The first element includes a method in which to obtain wargame data. Presented here are the two methods utilized in this thesis: Chess and MANA. Next is the utilization of HAC with DTW to determine similarity and proximity of each wargamed COA. Implicit in this step is the process of identifying an adequate linkage algorithm to perform the analysis and the optimum number of cluster

from which to analyze the clusters and the data. To provide this information, the HSig Index will be utilized. The final step as shown in Experiment 2 is the decision point analysis. This analysis will be done by analyzing the group of significant clusters to determine similarities and differences in winning and losing policies. Additionally, the heat map trajectories are analyzed to identify trends giving way to "rules of thumb" in the metric data.
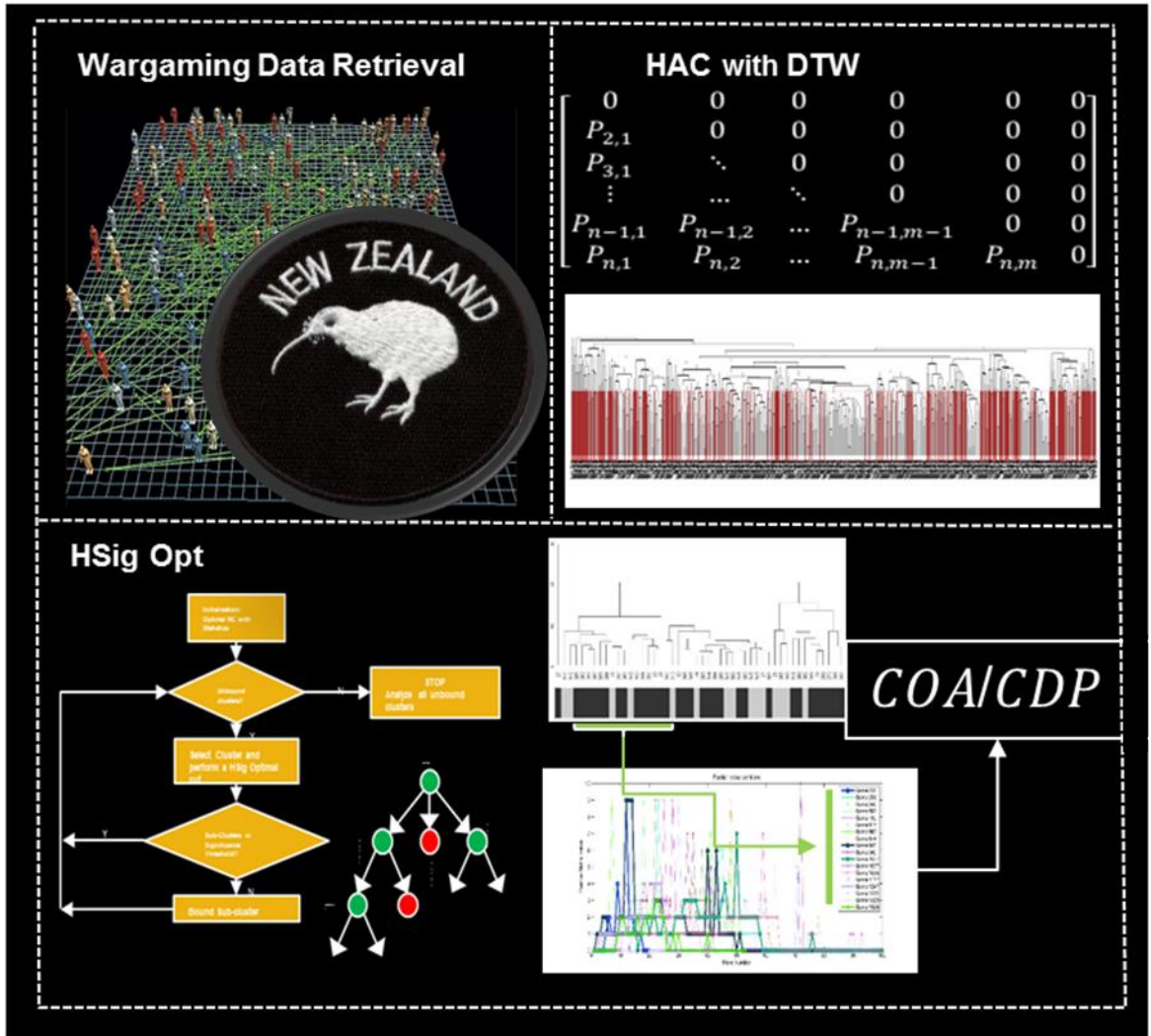
Figure 49. CADRE process visualization utilizing MANA as the method of wargaming and PAC as the data retrieved, leading to the eventual selection of COAs and CDPs.

4.3 Application Problem

Using MANA to wargame and CADRE to evaluate the data the effects of scale will be further evaluated. One vignette will be defined and CADRE will be applied enabling

decision point analysis to be conducted to identify trends based on favorable nodes associated with the battlespace.

- E.g. Portions of the engagement, where C/A/P is more important

- Portions where combinations of CAP data are important

- Situations, where the intersection of heightened CAP is important

Additionally, battlefield heuristics will be defined based on the decision point analysis. This will be attempted in a similar fashion to the attempt in experiment 2.1 where a similar position value map can be generated showing valuable areas for acquisition. This will be repeated taking awareness and capability into account to determine the coherence with the position value map.

CHAPTER V

RESULTS

5.1 Experiment 1

In this experiment, the HSig Index was calculated for a range found by sweeping through the possible range of number of clusters. The dataset used was the position data for WinBoard and Poseidon chess engines defined in the Modeling and Simulation section. Finally, the data was compared using the Dynamic Time Warping algorithm available in R [132]. Finally, clustering was performed using the Centroid, Complete, McQuitty (WPGMA), Median, Single, Ward, and Average (UPGMA) linkage algorithms. The results can be seen below in Figure 50 and Figure 51.

156
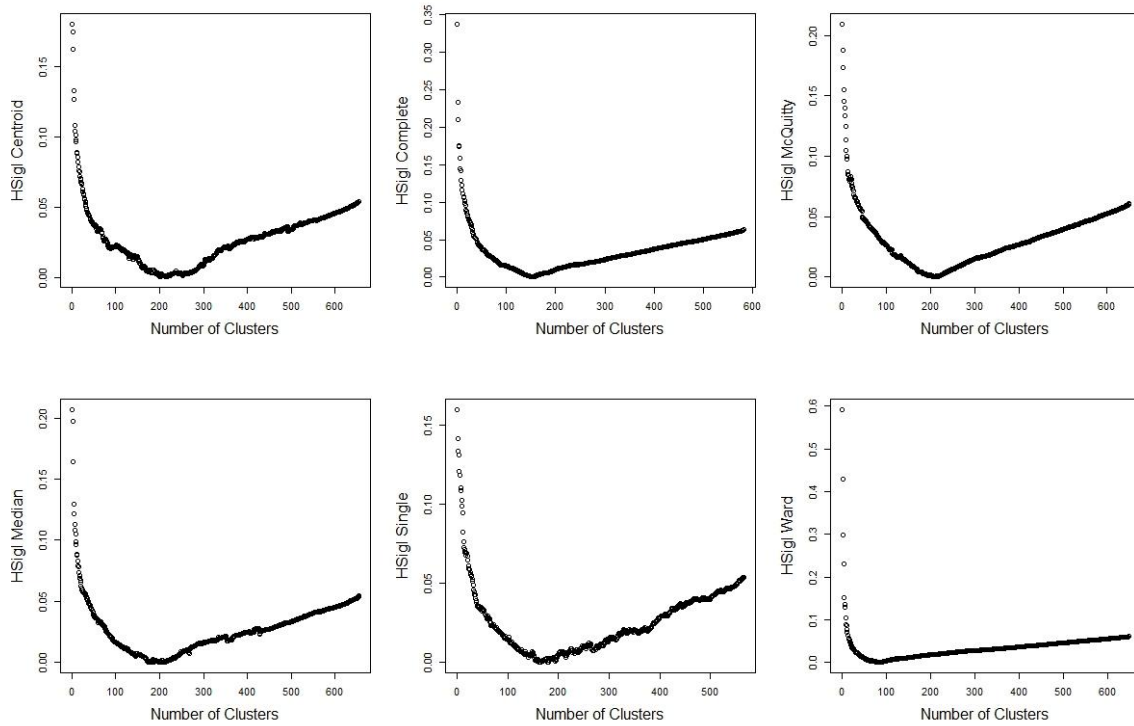
Figure 50. HSig Index Values of Linkage Algorithms ranging in perceived number of clusters illustrating algorithm performance and dynamics.
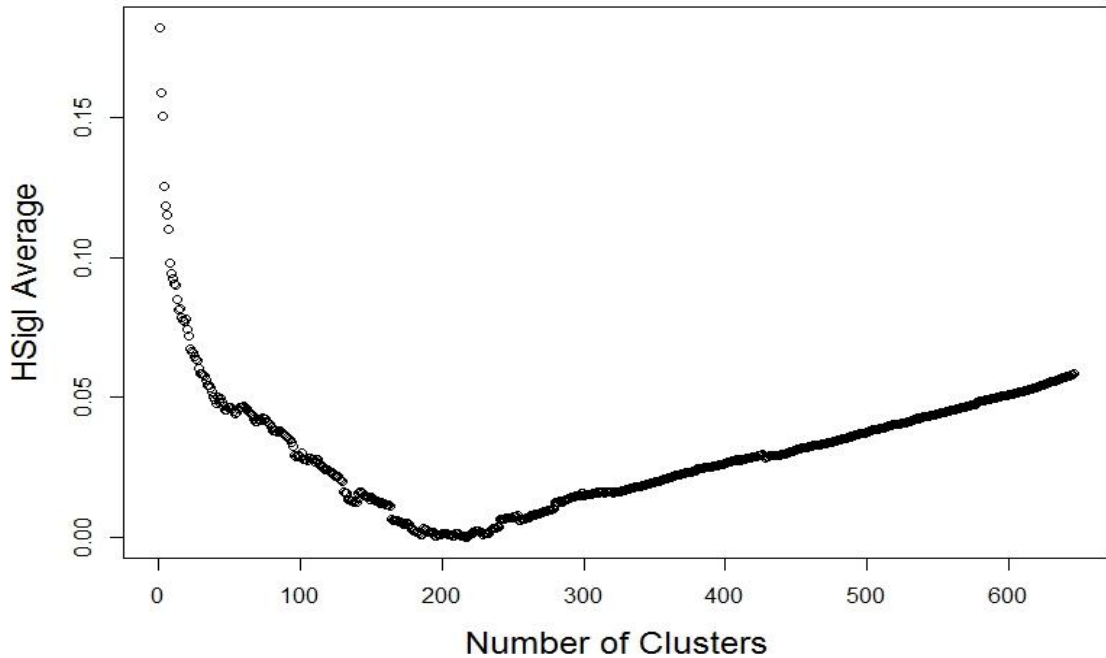
Figure 51. HSig Index Plot for Average (UPGMA) Linkage showing the HSig Index value of the optimum linkage and the associated number of clusters.

The behaviors seen in Figure 50 and Figure 51 reinforce some of the expected behaviors of the linkage algorithms. For example, Ward's method minimizes the value of HSig Index with the least number of clusters of all linkage algorithms. This can be attributed to the fact that Ward's linkage algorithm creates well defined clusters [113]. Next is the Complete Linkage Algorithm, which sees the minimum index value at the second lowest value for number of clusters. This is attributed to the Complete Linkage Algorithm's formation of tightly bound clusters [112]. However, the lowest index values

were seen by the Average Linkage Algorithm, and the Single Linkage Algorithm. Although, the Single Linkage Algorithm arrives at its minimum earlier in the plot, the Average Linkage Algorithm arrives at a lower minimum overall. This reinforces the expected performance of the two algorithms as UPGMA is expected to outperform Single Linkage Algorithm [114, 115]. This also reinforces some of the behavior seen in Figure 27, as the Average Linkage Algorithm represents the most compromising solution present on the Pareto frontier. The Single Linkage Algorithm is the second-best solution according to the HSig Index. The Single Linkage Algorithm, though fails to show up on the Pareto frontier. The remaining results can be seen in Table 8.

Table 8. Linkage algorithm performance unitizing Chess data and the relative optimality and specific number of clusters assessed by each resulting linkage algorithm.

LINKAGE ALGORITHM	OPTIMUM NUMBER OF CLUSTERS	OPTIMALITY RANK
AVERAGE	217	1
CENTROID	164	2
MEDIAN	182	3
COMPLETE	215	4
MCQUITTY	153	5
WARD	83	6
SINGLE	210	7

5.2 Experiment 2

After splitting the Dendrogram into 217 clusters, the statistics are collected and tabulated for each cluster. This includes the p-value, number of successes, number of trials and polarity for each cluster. For this portion of the analysis, a significance threshold of 10% was utilized. Therefore, clusters with p-values higher than this threshold are ignored. After performing HSig optimization, there remain 10 clusters and 71 games needing further analysis. However, there will be several clusters analyzed in the following paragraphs that do not meet the threshold while still enabling the identification of interesting behaviors. The first cluster to be analyzed is cluster number whose Position metric time series and Dendrogram with Win Density can be seen Figure 52 and Figure 53 respectively.

A general observation of this cluster is that there are multiple instances where a high valued enemy piece is within striking range. Specifically, Games 2, 223 and 503 make several advances on high valued pieces in the first twenty moves. The remaining games do so in the midgame similarly, though appear to be phase shifted. The strategy of this cluster appears to be to quickly seize attacking potential in the game opening in order to make attempts on high valued enemy pieces. As can be seen in Table 9, while there are only two losses in eight trials for this cluster, the associated p-value is not low enough to allow the rejection of the null hypothesis. Therefore, pursuing similar games as in cluster two without more data will not guarantee a result of a greater than 50% of success.

160
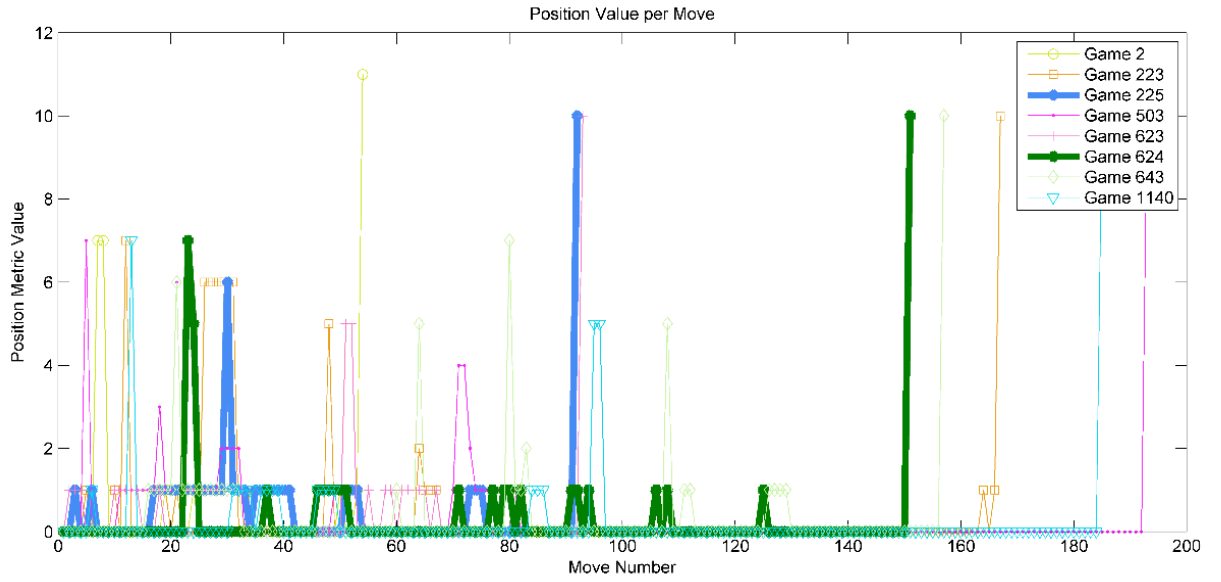
Figure 52. Cluster two position metric time series high level attack frequency.
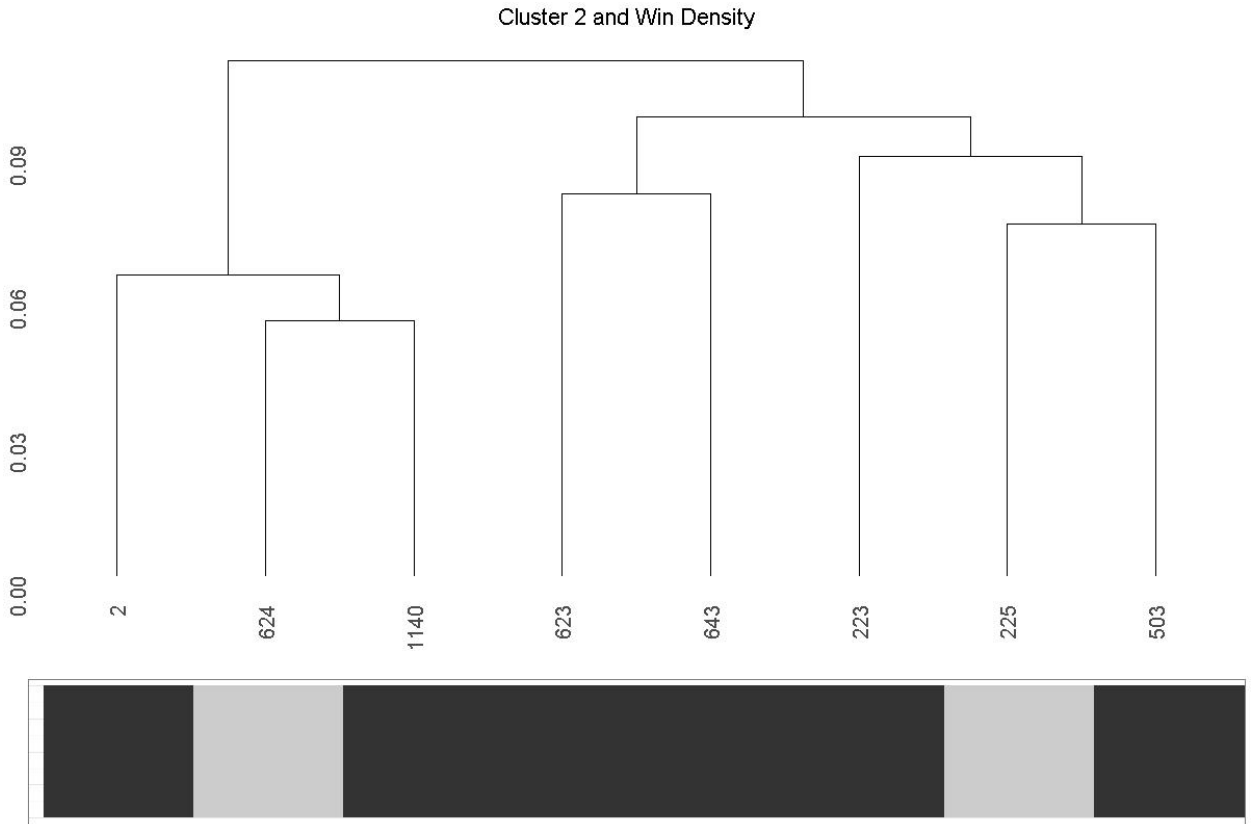
Figure 53. Cluster two dendrogram and win density.

Table 9. Cluster two statistics with insufficient evidence to reject the null hypothesis.

Cluster Number	2
p-value	0.144531
Number of trials	8
Number of successes	6
Polarity	0.75
Win Percentage	0.75

While we inspected a cluster that appeared to offer an increased win percentage in Cluster number 2, inspection of clusters with very low win percentages offer great utility as well. For example, Cluster 111 offers some insight as a losing cluster. Although, the p-value in Table 10 is not significant enough to formally reject the null hypothesis, investigation of the Position metric time series in Figure 54 shows that in all of the losing cases, no advances on the enemy pieces were made throughout the game. The exception is Game 212, where an advance on an enemy pawn was made in game opening, and the outcome was white win. While there was an advantageous outcome for Game 212, as seen in the win density of Figure 55, it is not expected that a single pawn attack in a game can yield victory with any likely interval. This will be shown as significant clusters exhibiting significant and increased polarity exhibit much more complex behavior. However, regardless of the polarity and significance of this cluster, this case is trivial. One would accept this as a losing strategy regardless of the magnitude of the evidence presented.
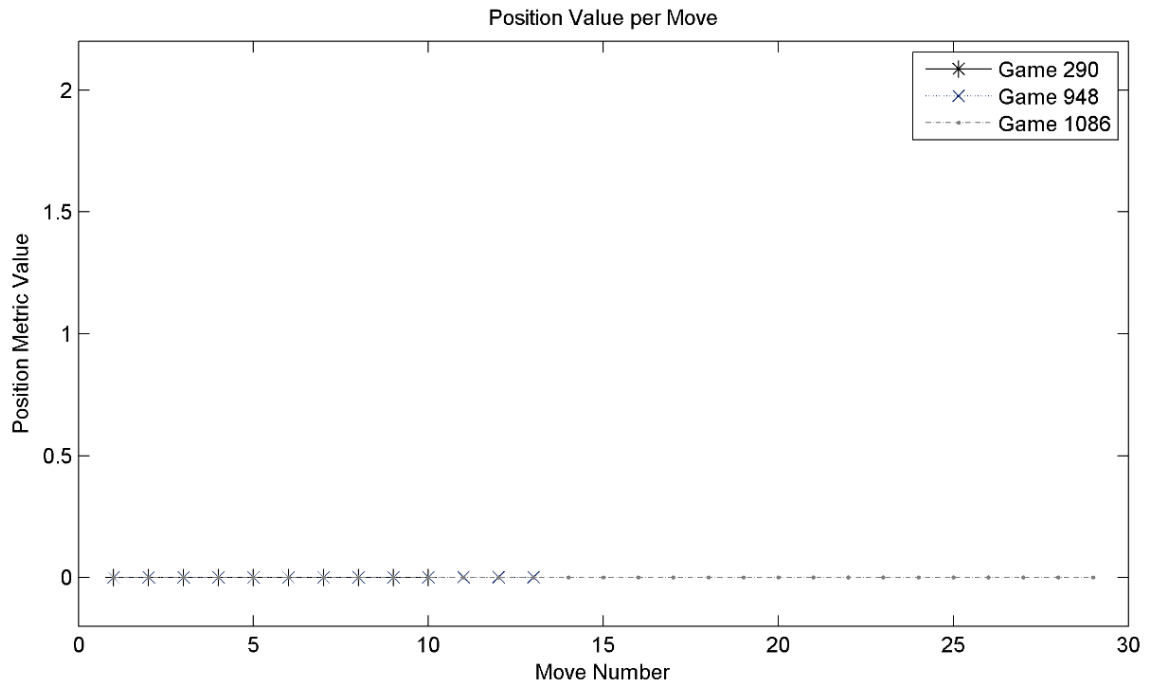
Figure 54. Cluster 111 Position metric time series for perceived identical cases where each Game retains zero Position metric value but are of differing game lengths.
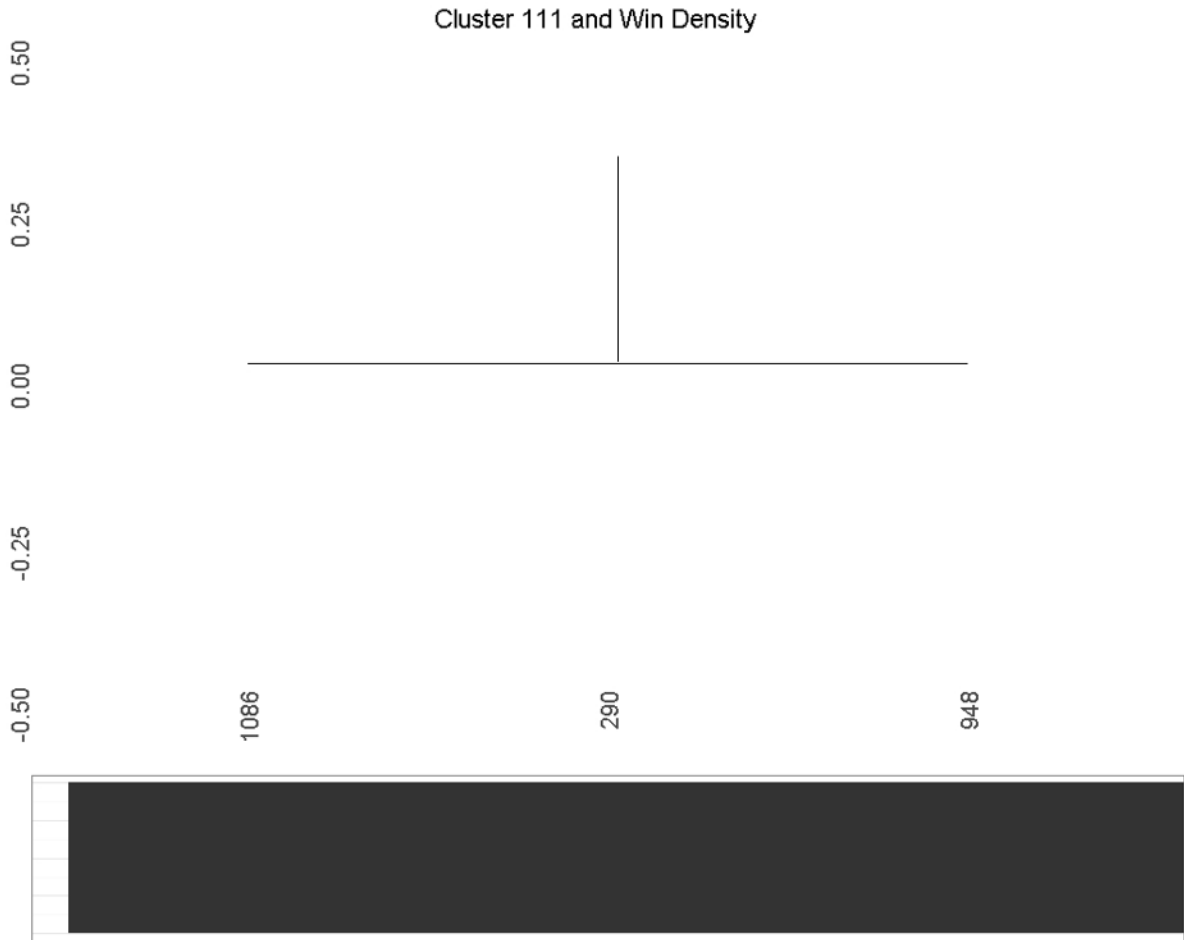
Figure 55. Cluster 111 dendrogram and win density identical cases.

Table 10. Cluster 111 statistics with insufficient evidence to reject null hypothesis.

Cluster Number	111
p-value	0.125
Number of trials	3
Number of successes	3
Polarity	1
Win Percentage	0

165

A compliment the case seen in Figure 54 and Figure 55 is Cluster 120 in Figure 56 and Figure 57. While this example is not significant with a p-value in excess of 12%, it serves to show a lack of actionable knowledge gain when extreme cases are similar and are clustered together. In this cluster, there is not much to be seen throughout opening and mid-course. However, these games achieve similar profiles throughout game play motivating the formation of the cluster. Additionally, the most salient aspect of these games is their achievement of Checkmate. This can be observed as the last Position metric value rises to at least ten negating the possibility of the capture of any other piece. While these games are similar, and their outcomes desirable, there is not much knowledge obtainable other than the importance of achieving Checkmate as early as possible.
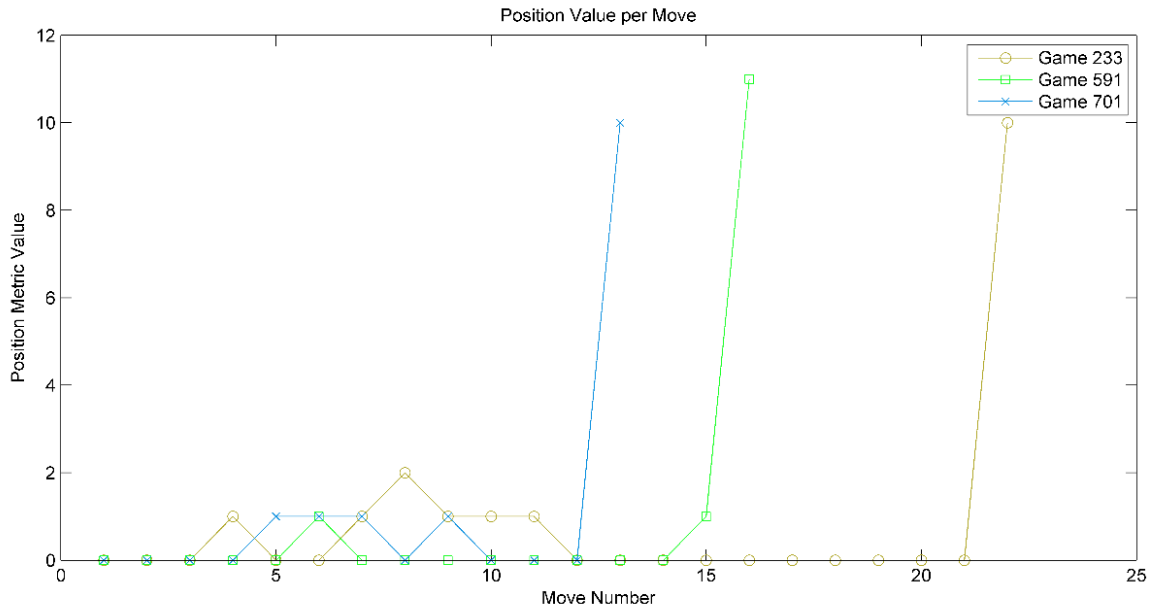
Figure 56. Cluster 120 Position metric time series showing similar game evolution and similar attempts on the enemy King in the last move.
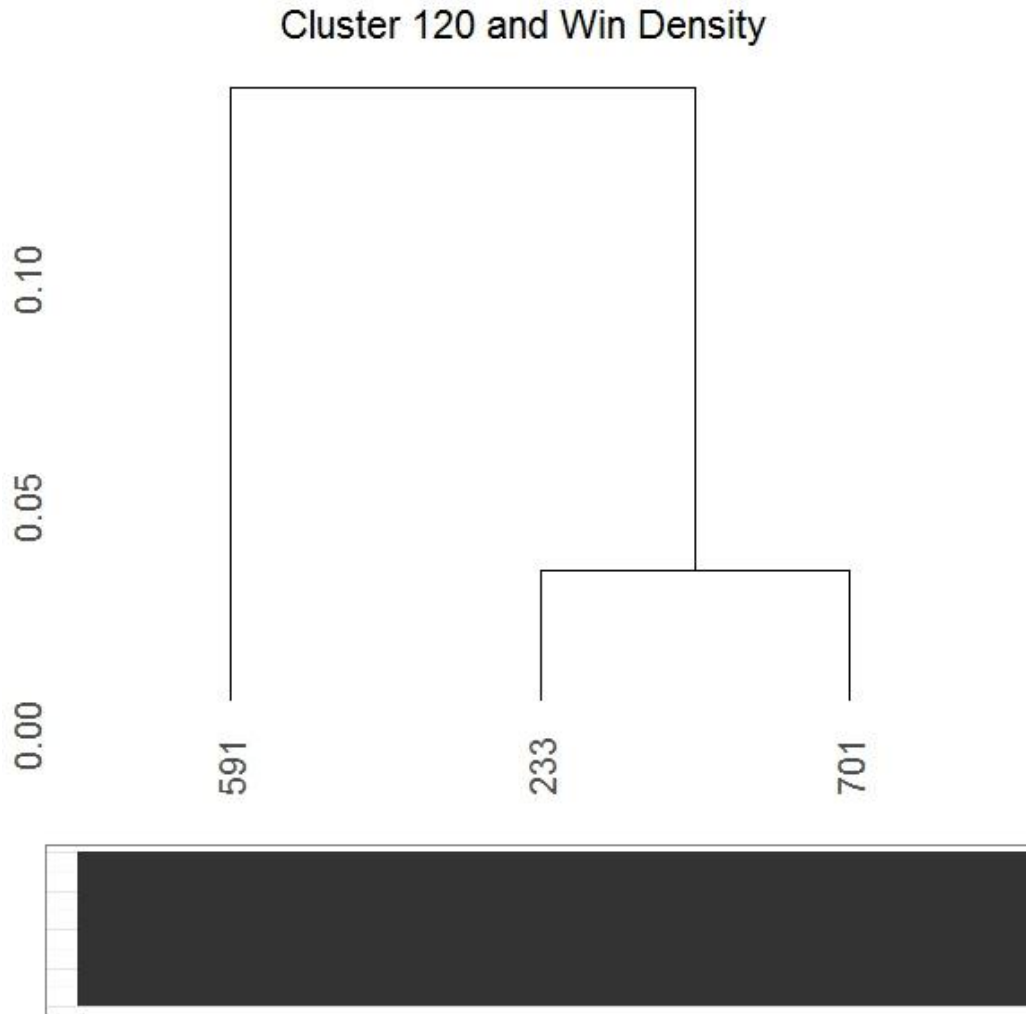
Figure 57. Cluster 120 dendrogram and win density.

Contrary to what was seen in Clusters 2, 111, and 120 the remaining clusters will adhere to the significance threshold with the exception of Clusters 22 and 159, which are slightly out of the 10% threshold at 10.9%. The initial identified clusters can be seen in Table 11 and are the product from a first pass of the HSig Optimization. They include a

range of acceptable p-values for a range of sample sizes, or cluster sizes. The following

analysis will investigate the overall behavior of or trends seen the clusters, and will avoid

making inferences about single games alone. Doing so would conflict with the overall

hypothesis and could potentially lead to accepting unsubstantiated trends through making

Type I or Type II errors. The remaining analysis will follow the order of significance

beginning first with Cluster 119.

Table 11. Identified clusters statistics within threshold value.

Cluster Number	119	175	22	159
p-value	0.048	0.062	0.11	0.11
Number of trials	18	7	6	6
Number of successes	13	6	5	5
Polarity	0.72	0.85	0.83	0.83
Cluster Win Sampling	0.72	0.85	0.83	0.16

Upon first look of Cluster 119, general observations from Table 11 are that it is the

largest intact cluster with the lowest p-value among remaining clusters, but it is the least

polarized. After analyzing the cluster further, the number of successes in this case is wins.

The proportion of wins to losses is approximately 72%, seen in Figure 59, however, the

limits of this method and with statistical hypothesis testing in general, is that we are only

to reject the null hypothesis. Therefore, Cluster 119 exhibits greater than 50% chances of

success, but the exact probability is unknown. Looking to Figure 58, the trajectories of the

Position metric can be seen. General trends appear to be early and frequent attacks on high

value enemy pieces. By looking at Game 865, this trend can be seen. In this game, at approximately move seven, a move is made which puts the enemy queen in striking range. This appears to be a spike in the Position metric. While, a value of nine of the Position metric is possible by being in attack range of multiple enemy pieces, this spike is attributed to the enemy queen. This is because if the value were due to multiple pieces, the enemy removing one of the threatened pieces would result in a partial drop in the position value because other pieces are still in striking range as one side is allowed one move at a time. This scenario can be seen in multiple cases, where the position value is constant for multiple moves. A true ambiguity exists in identifying what happened after a spike in position as in Game 865. It is unknown whether, the drop is due to a successful attack on the enemy queen, or a successful evasion by the enemy queen. In either case, the queen is no longer in striking range. One way to attempt to identify what happened is by searching for other spikes in the time series for that game. However, as promotion is permitted in the chess engine rules, additional spikes to that level of position value, can be due to re-emergence of the enemy queen on the board. This further reinforces the position to only extract cluster specific behavior as opposed to game specific behavior. For this cluster, the trend appears to be maneuver early to attack high valued targets, while simultaneously maneuvering to make multiple lower level attacks.

170
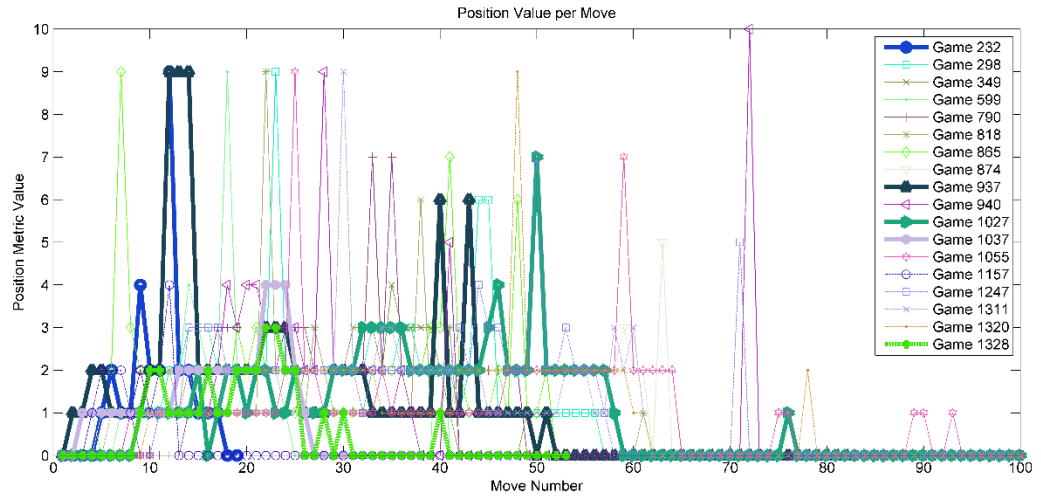
Figure 58. Cluster 119 Position metric time series showing similarity attack evolution and a high and sustained attack frequency in cluster.
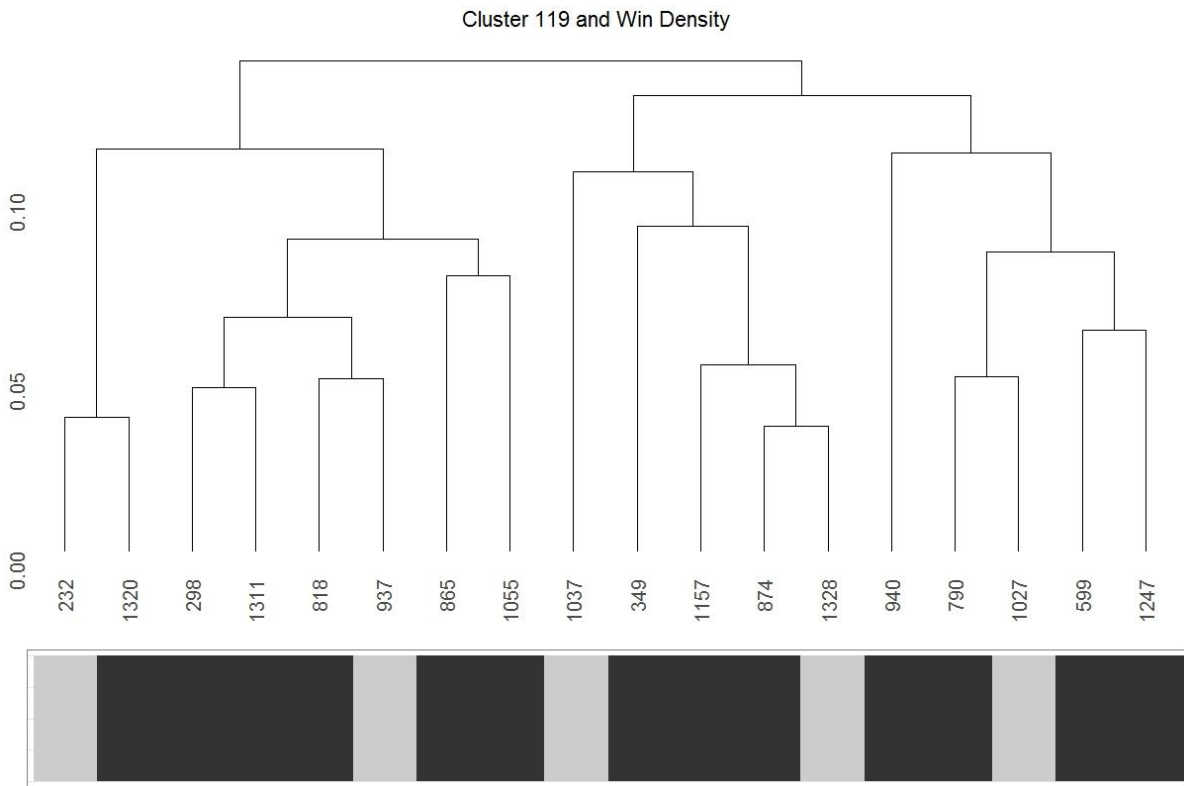
Figure 59. Cluster 119 dendrogram and win density.

Another dimension to analyze the cluster of games is in the frequency that positions of the eight by eight board are utilized. In game opening, it is well known, that occupation of the centerboard is desirable [129]. The preference for specific board positions can be seen below in Figure 60. However, a slight difference was found when observing the games of twelve Chess Grand Masters [133]. Instead of occupation of the center, these players show a desire to form a chevron in the center with an inclination towards the king's side. The preferences of the Grand Masters can be seen in Figure 61, where for each player,

172

the average moves played to positions per game are plotted when playing white on the left and when playing black on the right. Similarly, for Cluster 119 the total number of times a position was played, for all games in the cluster, is plotted in Figure 62. The result mimics the behavior seen in the opening position desirability and the analysis of twelve grandmaster games. In the Cluster 119 plot, there seems to be an affinity for the center of the board as in Figure 60, but there also appear to be an affinity for the kingside as in Figure 61. However, this kingside preference appears much more drastic and the centerboard preference appears distorted with greater variance in preference of the center. An inference from the greater inclination towards the kingside is that it offers the king more protection to direct an assault in that direction. The increased distortion is likely a product of the randomness of moves selected by the chess engines. The analysis of the Chess Master games is the product of a decisive strategy and, therefore, form more crisply observable behaviors. The product of randomness appears to give to the distortion of the appearance of board position preferences. Regardless, there still appears some promising behavior.
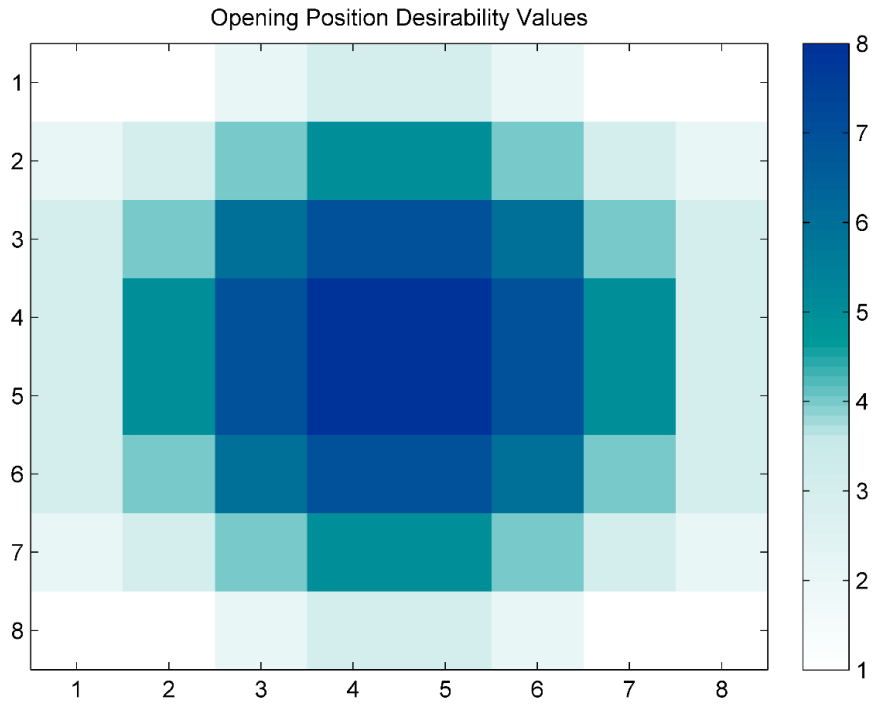
Figure 60. Opening and mid-game board position desirability.
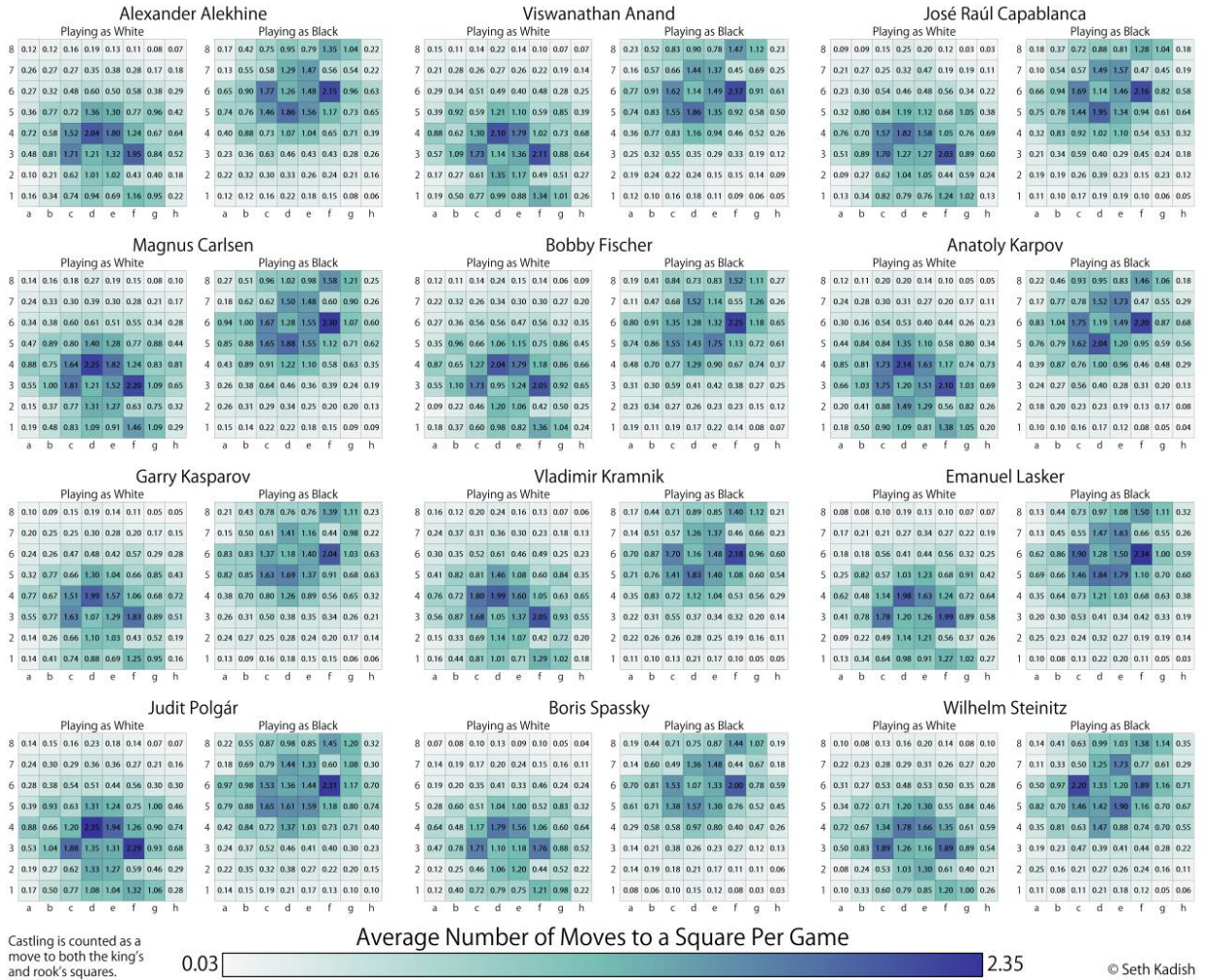
Figure 61. Utilization of squares by chess masters showing a tendency towards center board and to the King side for game play using either black or white side [133].
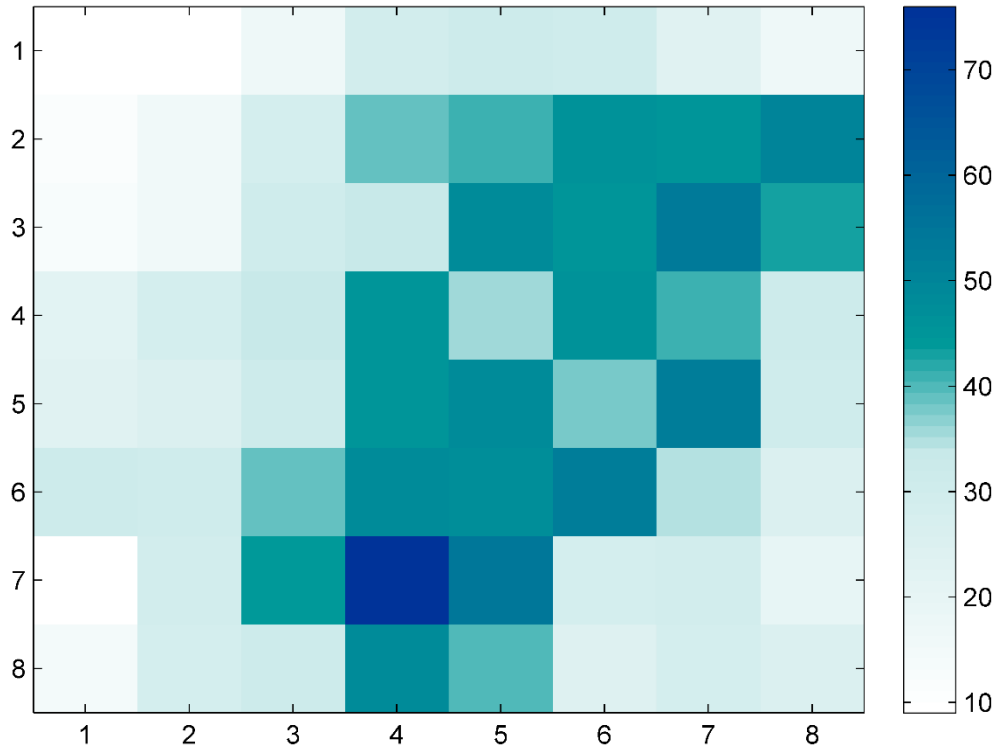
Figure 62. Cluster 119 Position histogram favoring center and king-side.

While Cluster 119 exhibits actionable behavior to increase the probability of success, there may be lower level behaviors identifiable, as it is the largest intact cluster. HSig Optimization was performed on this cluster, as the HSig Index was calculated for a range of potential clusters. This can be seen below in Figure 63. The optimum number of sub-clusters for Cluster 119 is eight clusters. This point sees the lowest value of the HSig Index. However, after looking at the statistics in Table 12, there are no sub-clusters with low enough p-values to merit further investigation. Therefore, Cluster 119 is optimally bound with no benefit from further decomposition.

HSig Index for Cluster 119
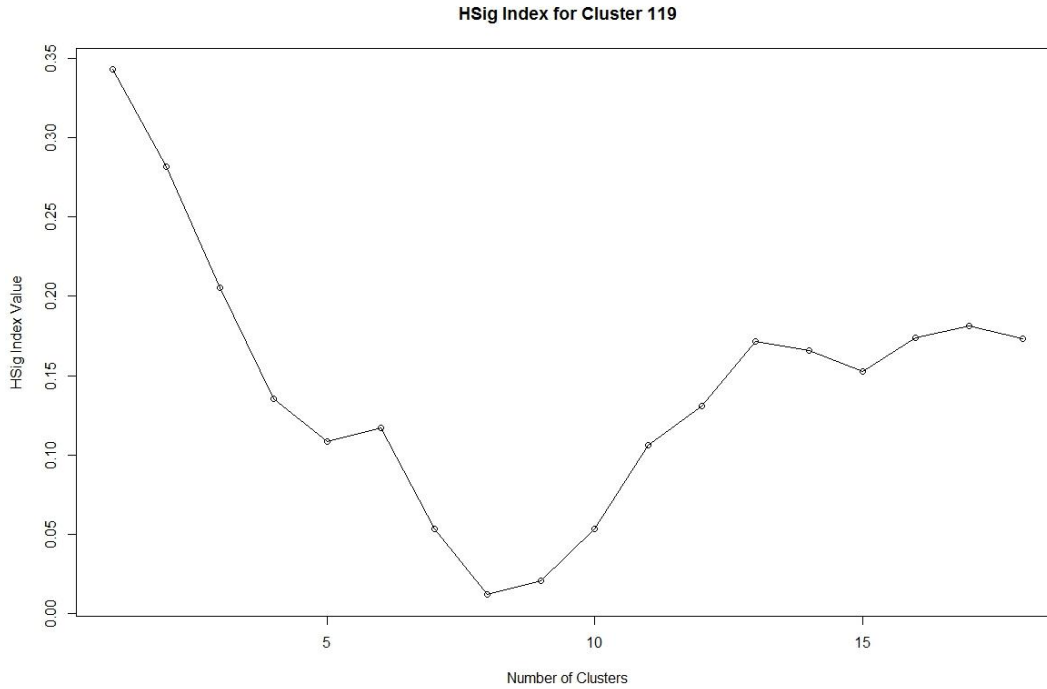
Figure 63. HSig Index Values for Cluster 119 showing a new achievable optimum.

Table 12. Cluster 119 Sub-cluster statistics showing a lack of statistical significance.

Sub-cluster Number	1	2	3	4	5	6	7	8
p-value	0.75	0.31	0.5	0.31	0.25	0.5	0.5	0.5
Number of trials	2	4	1	4	2	3	1	1
Number of successes	1	3	1	3	2	2	1	1
Polarity	0.5	0.75	1	0.75	1	0.67	1	1

The next cluster to be investigated is Cluster 175. This cluster has the second lowest p-value and the highest win percentage among all initial clusters. Upon first look at Figure 64 and Figure 65 similarities can be seen in the time series, however, it seems as though

177

this cluster can be separated into two clusters. While this is possible to do visually, HSig Optimization will not allow it in this case because, separating this cluster will result in an increase in the p-value. In any case, this cluster exhibits a few trends that reinforce the significance of this cluster as one which increases the probability of success. Observing the total number of moves in all games indicates that these games ended relatively quickly, with the longest game lasting 36 moves. Additionally, none of the games end with a high position metric value, indicating a win by forcing attrition of the opponent. This is reinforcing as most of the games make early, frequent and sustained attacks of enemy targets. Specifically, Game 658 exhibits this trend in the fourth move, where a position metric value of seven is attained. This value is not through the attack of one piece as no piece value is seven. The only combination of targets that can give this value includes Pawns, Knights and Bishops, whose values are one, three and three respectively. Therefore, the minimum number of pieces in attack range for move four of Game 658 is three. A similar trend can be seen in Games 1166 and 1251, where at different points of the game, the same value is reached, however, in each case, the values are increased by one during the next move. This can only be due to the addition of an enemy Pawn to the pieces in attack range.
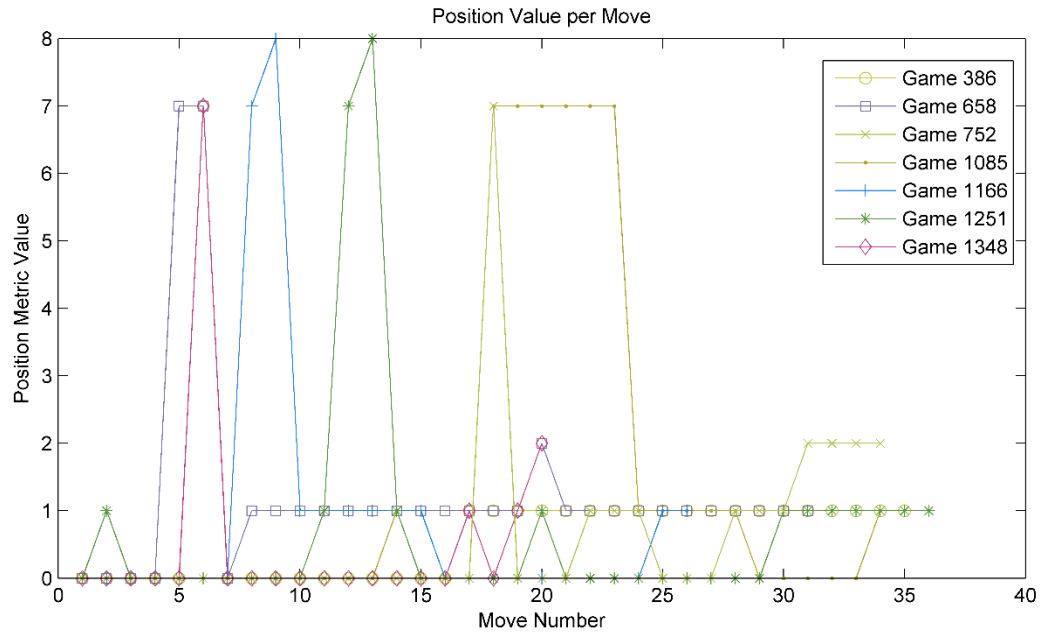
Figure 64. Cluster 175 Position metric time series showing similarity of attack sequence.
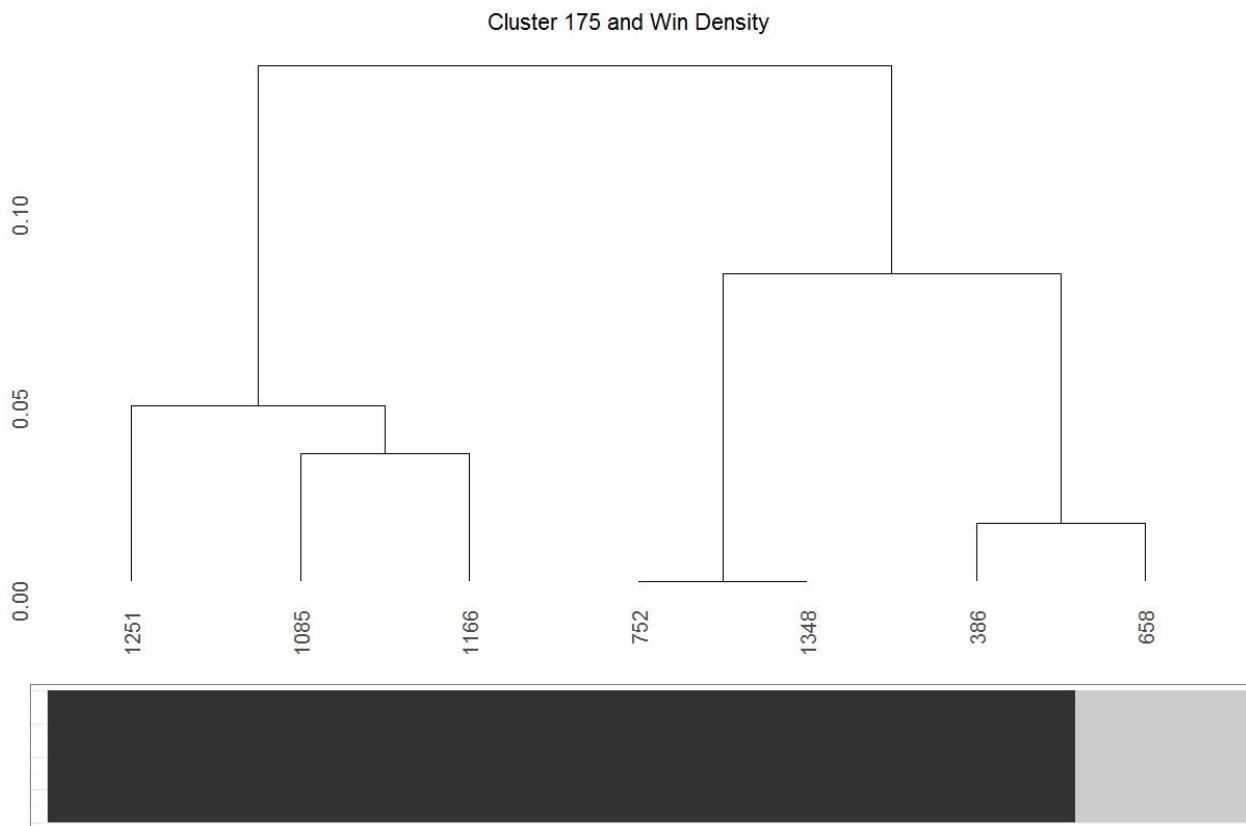
Figure 65. Cluster 175 dendrogram and win density.

Other games of interest include Games 752 and 1348. The distance between these two games, according to DTW, is zero as seen in Figure 65. Therefore, they are treated as identical games. However, looking to Figure 64 there are some differences. First, Game 1348 is in attack range by move five but it takes Game 752 eighteen moves to attack. However, they attack the same value of targets, and these attacks are both their first attacks of the game. Similar differences are seen in the games on further analysis, but each game reaches the same attack potential reaching identical values of the position metric. Another

peculiarity of these games can be seen by observing the final move, move 20 of Game 1348. In this game, the final value is due to the attack of two pawns. This occurs in Game 752 as well, but it occurs later in the game at move 31 and is sustained for four moves until the game ends. This is an artifact the DTW measurement. In these cases, time is warped allowing one point of Game 1348 to equal four points of game 752. It is due to the time warping that the games, while out of phase, can be seen as identical. This behavior, while peculiar, is a testament to the effectiveness of DTW on this type of analysis. By looking at the Position metric time series, it can be seen that these games are out of phase, but are identical in every other way.

While this cluster is appealing due to the observed increased win probability, the significance is at the edge of acceptability. Additionally, analysis of the Position Histogram in Figure 66 does not depict strong behavior. While a middle board Kingside propensity can be seen, the trend is week. This is most likely do to the relatively short duration of the games in a somewhat small cluster. In this case there is not enough data to identify strongly a board occupation strategy.

181
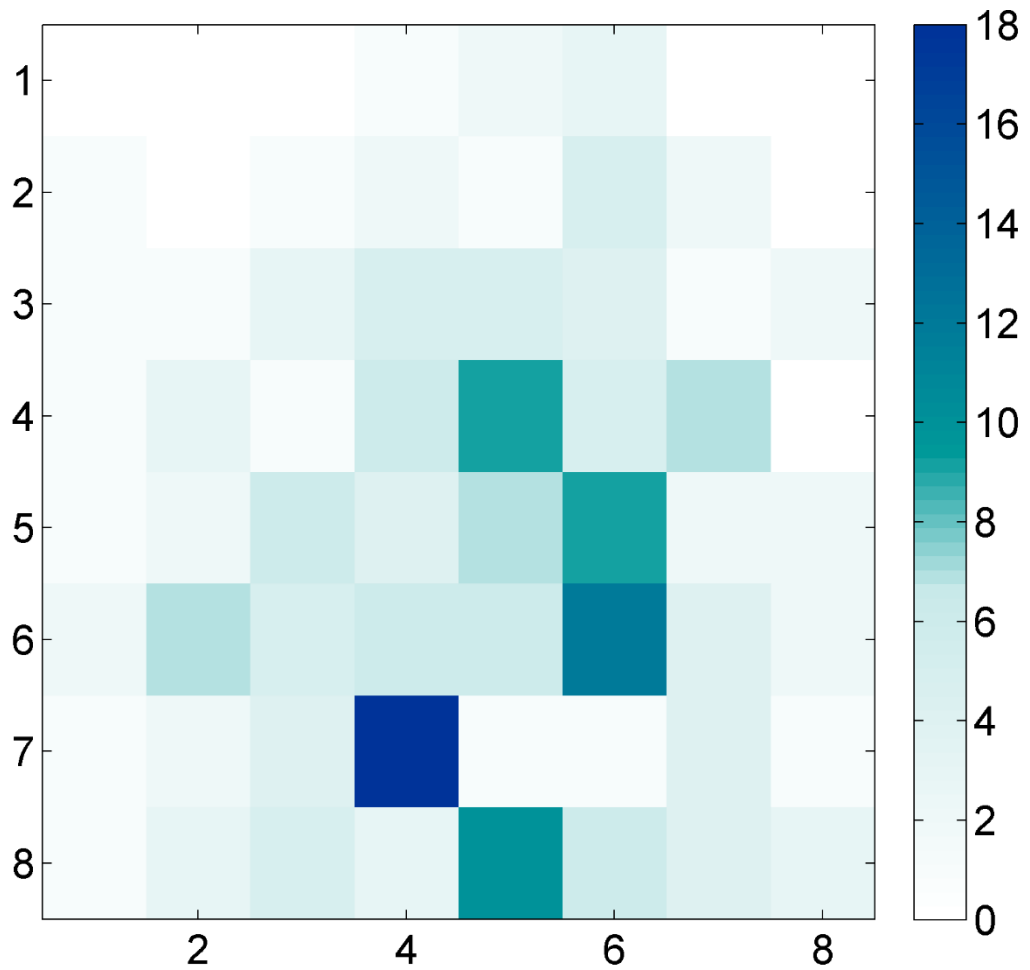
Figure 66. Cluster 175 Position histogram showing slight king-side center preference.

The next cluster to be analyzed is Cluster 22. Similarly, to Cluster 175, this cluster is at the edge of acceptability with regard to the p-value. However, both the Position metric time series and the Position histogram plots provide clues as to why we are able to reject the null hypothesis in this cluster. By observing Figure 67, it can be seen that none of the

182

games exceeded 35 moves. Additionally, none of the games resulted in a Position metric exceeding a value of three. Therefore, it can be deduced that this clusters components are games that resulted in a draw. However, this cluster offers better than 50% chances of success because of the reduction in enemy material. While, no game exceeds a Position metric value of three, none of the games go back to zero after first contact. Therefore, through a sustained attack on low level pieces, a draw is forced, however a win is declared because at the time of games end, there are fewer enemy pieces on the board, then friendly pieces.
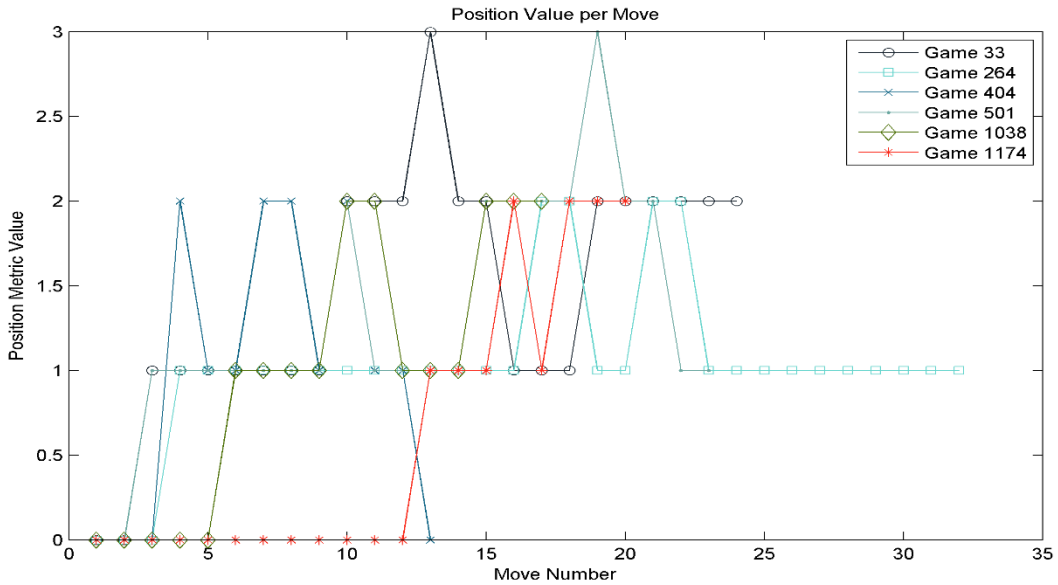
Figure 67. Cluster 22 Position metric time series showing similarity in game development and evolution maintaining non-zero metric values in the endgame.

Additional trends can be seen by observing the trajectories of the Position metric time series throughout the game as well as the Dendrogram with win density in Figure 68. Similarly, to Cluster 175, this cluster has games whose distance is zero but are not visually identical. Observe in Games 1038 and 1174 the time to first contact is different by seven moves. Additionally, Game 1038 remains at a value of one for more moves than Game 1174 before the next attack, where Game 1038 stays at that value for longer. However, as stated previously, this is an artifact of the DTW process, but the measurement and contrast of the two games still gives an indication of the benefit of the identical tempo found in both games.
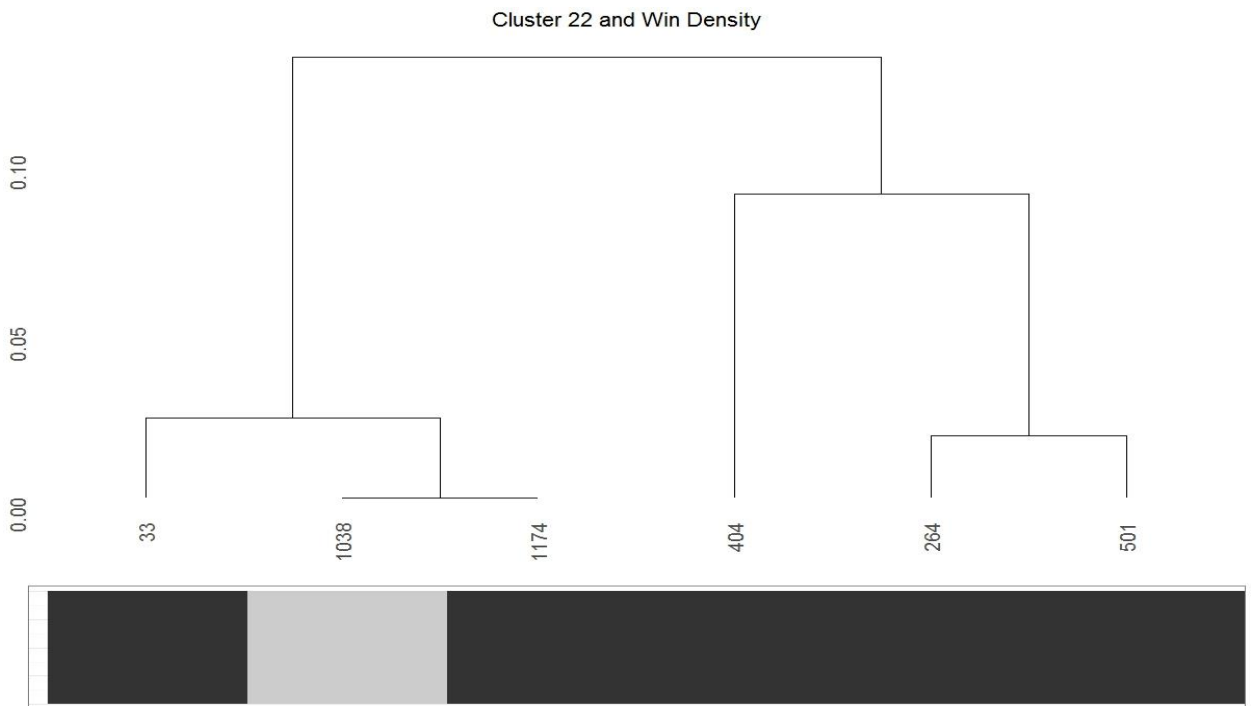
Figure 68. Cluster 22 dendrogram and win density.

184

Further analysis of the cluster with respect to the position histogram reinforces the trends seen in the Position metric time series as seen below in Figure 69. Generally, there appears to be a preference towards the center of the board, with a bias towards the Queen side generally. However, there is a slight deviation from the trend in very heavily played position on the King side. This appears to be advantageous board positioning on the center board Queen side with provisions for King defense towards the rear King side seen in the sixth row and sixth column of the board. A peculiarity of note can be seen in row two and column six of the board. This position is played approximately 5 times by all of the games in Cluster 22 and, in theory, should constitute a higher value for the Position metric as, this board position is next to the enemy King and Bishop initial locations. However, in none of the games does the Position value exceed three. Therefore, in this situation, the histogram is not telling of the game specific potential to attack high valued enemy targets. Conversely, it is preferential to occupy board positions in this area due to the potential for piece promotion, increasing the relative capability by promoting a pawn to a more mobile piece.
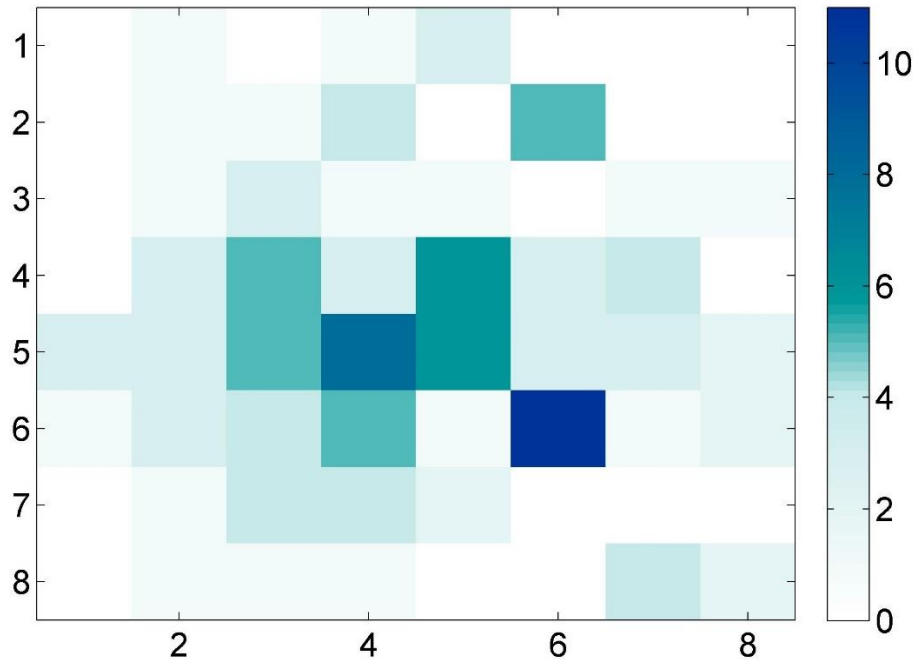
Figure 69. Cluster 22 Position histogram with strong center board preference.

Cluster 22 is the final cluster observed after the first pass of the HSig Optimization. For each other cluster defined at the initial cut, HSig Optimization is performed by treating each cluster as its own data set, clustering and sweeping through the range of possible numbers of sub-clusters within the original cluster. In doing so, only one cluster was left unbound after the process. Meaning, based on the data set provided, no additional information can be gleaned from the other clusters. The remaining cluster was Cluster 1 and has an optimum HSig Index value at six clusters. The statistics associated with this cluster and sub-clusters can be seen below in Table 13. The only sub-cluster that is bordering on the insignificant is Sub-cluster 20. However, the other five sub-clusters offer

significant evidence to reject the Null Hypothesis and will be investigated further in the following paragraphs.

Table 13. Significant sub-clusters of Cluster 22 and statistics.

Sub-cluster Number	69	124	13	56	87	20
p-value	0.007	0.031	0.032	0.062	0.062	0.109
Number of trials	7	5	11	4	4	6
Number of successes	7	5	9	4	4	5
Polarity	1	1	0.818	1	1	0.833
Win Percentage	0	1	0.818	1	0	0.167

The first sub-cluster to analyze is Sub-cluster 69. In this sub-cluster, there are seven games grouped together that all correspond to losses. As a result of the size of this sub-cluster as well as the polarity of the outcomes give way to the overwhelming significance of the cluster to allow rejection of the null. Of all clusters analyzed, this cluster offers the most compelling evidence to signify polarity. While this cluster represents an inclination to lose games, it is technically an increase in the probability of success, as the way this formulation is posed, any change from the 50% probability of success is treated the same. This treatment is in the Polarity and only when calculating the Win Percentage does the magnitude of Polarity get the direction of Win Percentage.

An investigation into the trends seen in the Position metric time series in Figure 70 gives some indication of why this represents a negative impact on the Win Probability. Games 99 and 139 are the only games that do not see high Position values late in the game.

Specifically, all of the other games in the sub-cluster achieve higher than a value of three after move fifty in the game. This indicates that the enemy is able to retain use of their high value pieces, or are able to force a promotion after friendly capture of enemy pieces. However, a majority of the games are seen to get into attack range of high value enemy pieces, a trend that has led to an increased win probability in other clusters.

A potential source of this conflicting behavior can be seen through analysis of the Dendrogram in Figure 71. The result of Games 99 and 139 varying from the remaining games in the sub-cluster can be seen as, these two games are a kind of outlier within this sub-cluster. Furthermore, none of the games within this sub-cluster are as similar as other games analyzed so far, a behavior that can also be seen in Figure 72, where there appears to be some tendency to occupy the center of the board, but a stronger tendency to occupy the rear King side. While the second of the two tendencies appears foolhardy compared to established board play heuristics, the presence of limited center of board tendency calls into question the entire sub-cluster behavior and lessons learned. Therefore, in this cluster, with profound statistical evidence, there appears to be a within cluster incoherence preventing the learning of any heuristics. The reason for this disconnect is due to the HSig Index itself. There is no weighting between the significance of the outcomes or the height or similarity between elements of a cluster. In the example of Sub-cluster 69, the significance was a much higher driver for cluster formation than the relative similarity

between the elements within the cluster. This example makes the case for identifying a threshold value where clusters will not be formed based on the within cluster similarity.
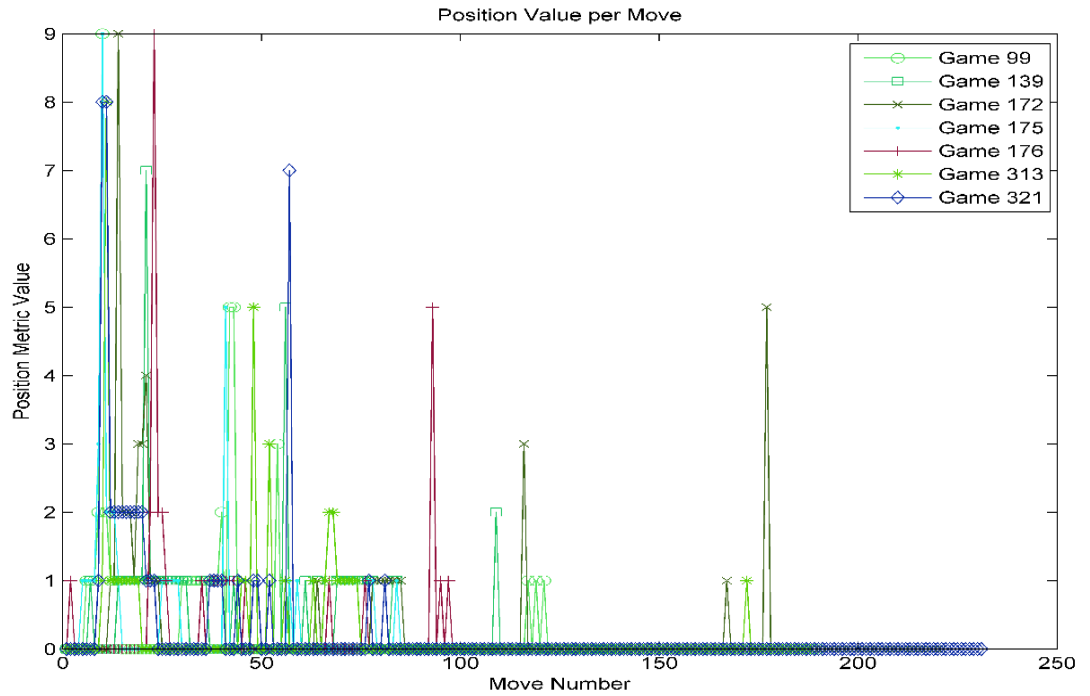
Figure 70. Cluster 1 Sub-cluster 69 Position metric time series showing early high value piece attacks.
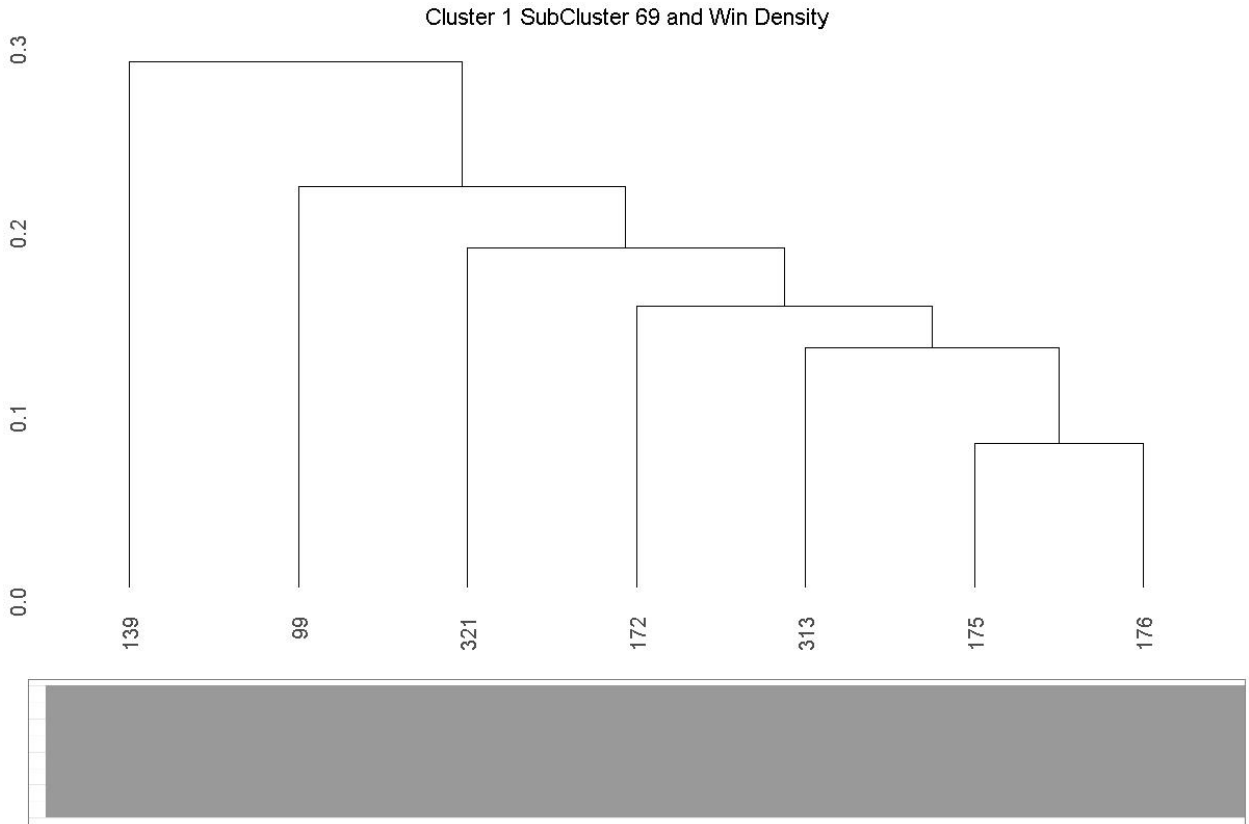
Figure 71. Cluster 1 Sub-cluster 69 dendrogram and win density showing cost of high value attacks.
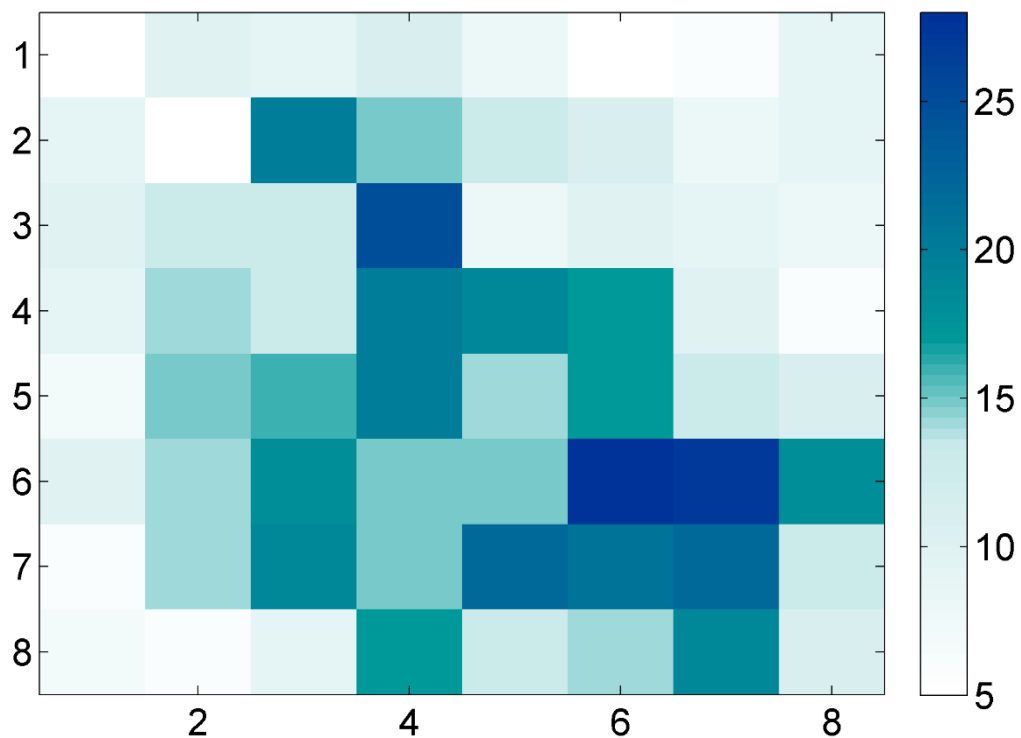
Figure 72. Cluster 1 Sub-cluster 69 Position histogram showing slight center board preference.

Sub-cluster 124 on first glance appears similar to Sub-cluster 69, however, under increased scrutiny, there appears to be more inter cluster coherence as seen in Figure 73 and Figure 74. This cluster exhibits similar within cluster game similarity as identified in the Position metric time series as well as through the similarity seen in the Dendrogram in Figure 74. Additionally, in opposition to Sub-cluster 124, after move 60 there does not appear to be any attacks made on high value enemy pieces. Therefore, it can be inferred

that there are no more enemy pieces remaining for most of these games. This inference can be made because a majority of the game in this cluster continued for another 60 – 80 moves. In that span, if there were other high valued enemy pieces, it is expected that by chance, they would have arrived within attack range. Furthermore, all games reach a Positon value of at least seven within the first twenty moves and some do so repeatedly until approximately move 60 where high value piece activity drops. Therefore, it can be inferred that all high valued enemy pieces except the Queen and King are removed from the board in the game mid-course and that similarly to other clusters analyzed, the win in this cluster came from a draw where the white pieces had more pieces left after the last move.
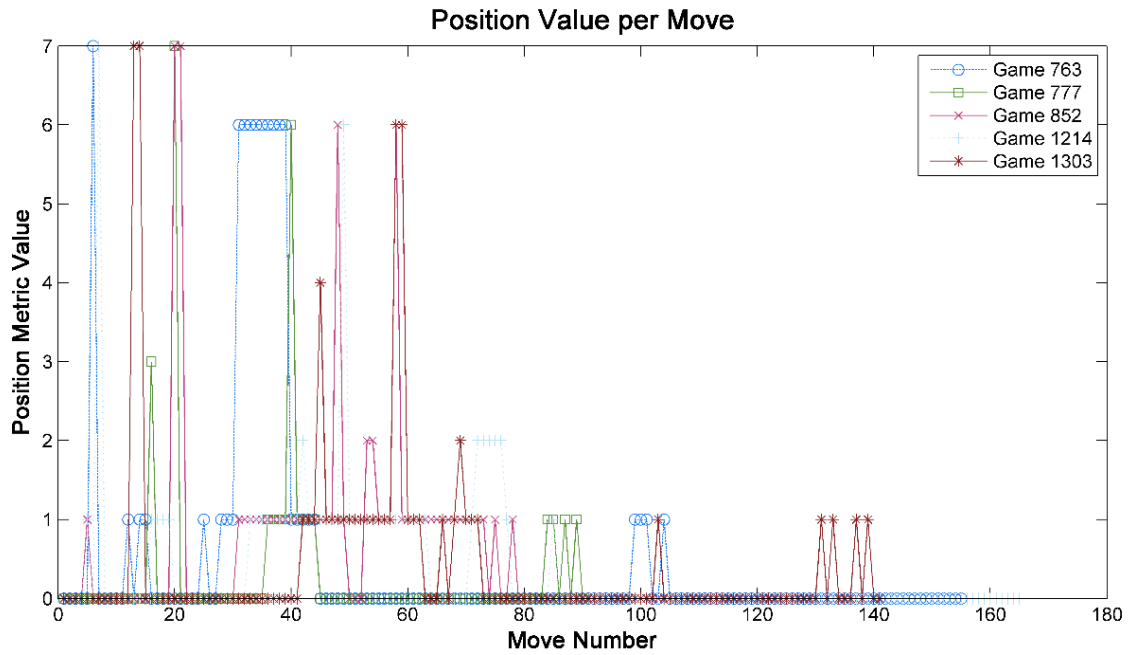
Figure 73. Cluster 1 sub-cluster 124 Position metric time series showing the signature of game development and attack frequency oscillating between attack and zero metric value.

Figure 74. Cluster 1 Sub-cluster 124 dendrogram and win density.

This sub-cluster appears to have a more defined structure to board positon preference than the previous sub-cluster although, the structure is not crisply defined. In other clusters and sub-clusters, there appeared a strong tendency towards the center of the board while this example occupies what seems to be an inverse chevron with a slight tendency towards the King side which can be seen in Figure 75. The most played position on this board with respect to the games in this sub-cluster is the position in front of the

194

white King. This can be due to the King moving frequently to that position or due to other pieces moving into that position serving to protect a static King. In either case, this serves to make the white King a more difficult target as in either case, there is an obstacle to surmount in order to force a Mate. In one case, it is hitting a moving target and in the other, it is the surrounding pieces moving into protection range of the King. The King side trend appears to fade as we move up to the forth row of the board although, it is still a relatively strong behavior signifying King defense farther up the board.
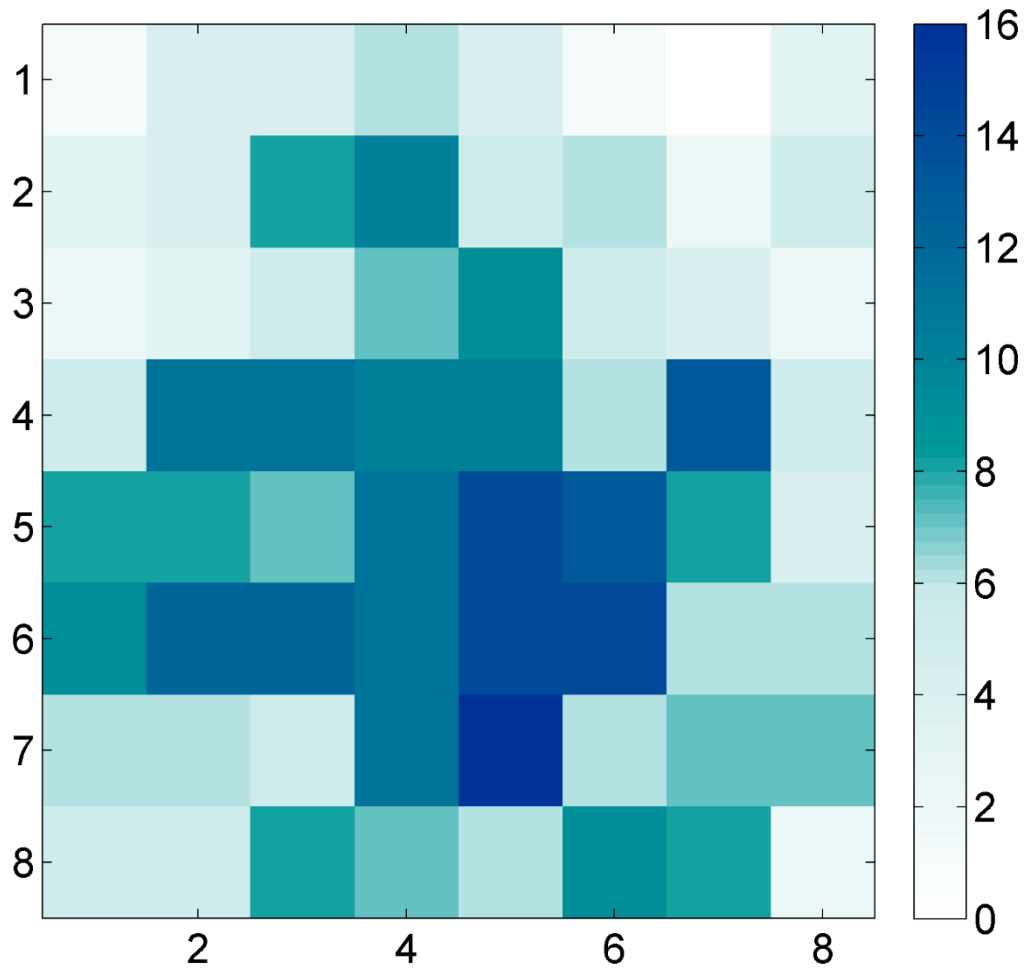
Figure 75. Cluster 1 Sub-cluster 124 Position histogram showing center board preference.

Sub-cluster 13 appears to be the most aggressive cluster of all clusters analyzed. By observing the Position metric values in Figure 76, it can be seen that a majority of the games reach Position values of 8 frequently throughout the game with some doing so in the game opening. Specifically, Game 44 reaches that height first and is followed by the

others. With the exception of Game 54, the other games see no high value piece attacks after approximately move 40. However, Game 54 sees a Positon value of 11 spike in one move. Therefore, it can be determined that this was the result of a single move where a white piece came into attack range of a King and a Pawn, a Queen and two Pawns, or a larger combination of lower value pieces. With the exception of Game 54, the other games in this sub-cluster appear to show more similarity.
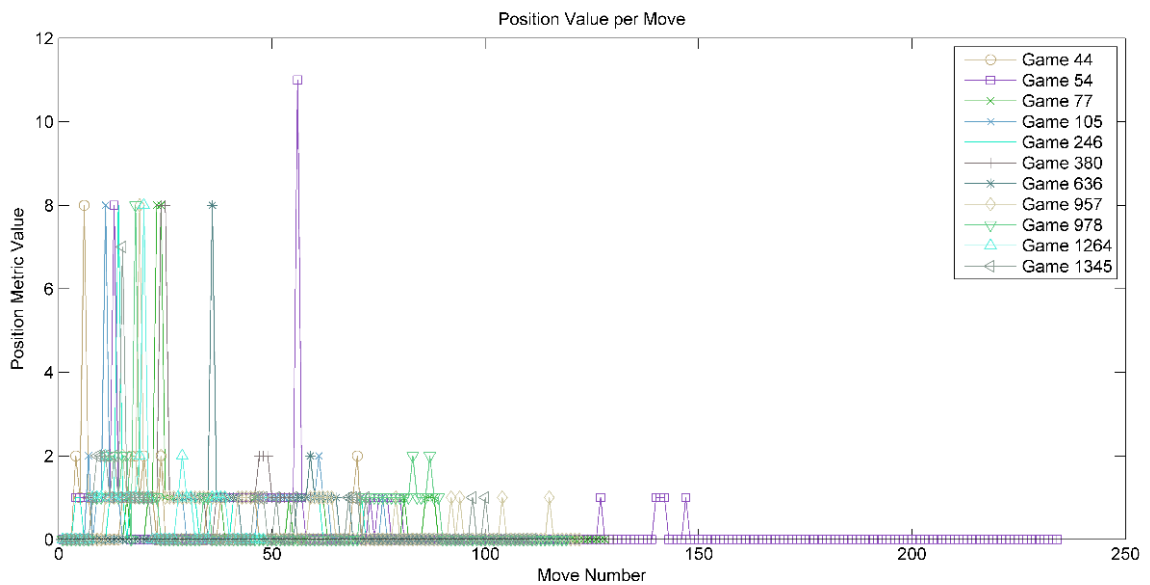
Figure 76. Cluster 1 sub-cluster 13 Position metric time series showing similar attack behavior and outlier Games exhibiting unique behavior among cluster neighbors.

By looking to Figure 77 the entire cluster appears to show a strong similarity because the maximum height joining cluster elements is 0.06. Specifically, DTW finds Game 44 0.06 units away from the remaining games in the sub-cluster. Additionally, Game

197

54 with the uncharacteristic spike in Position value more similar than Game 44 to the remaining games in the sub-cluster. Additionally, with two losses present in this cluster it still offers significant evidence to reject the null hypothesis with a p-value of 0.03.
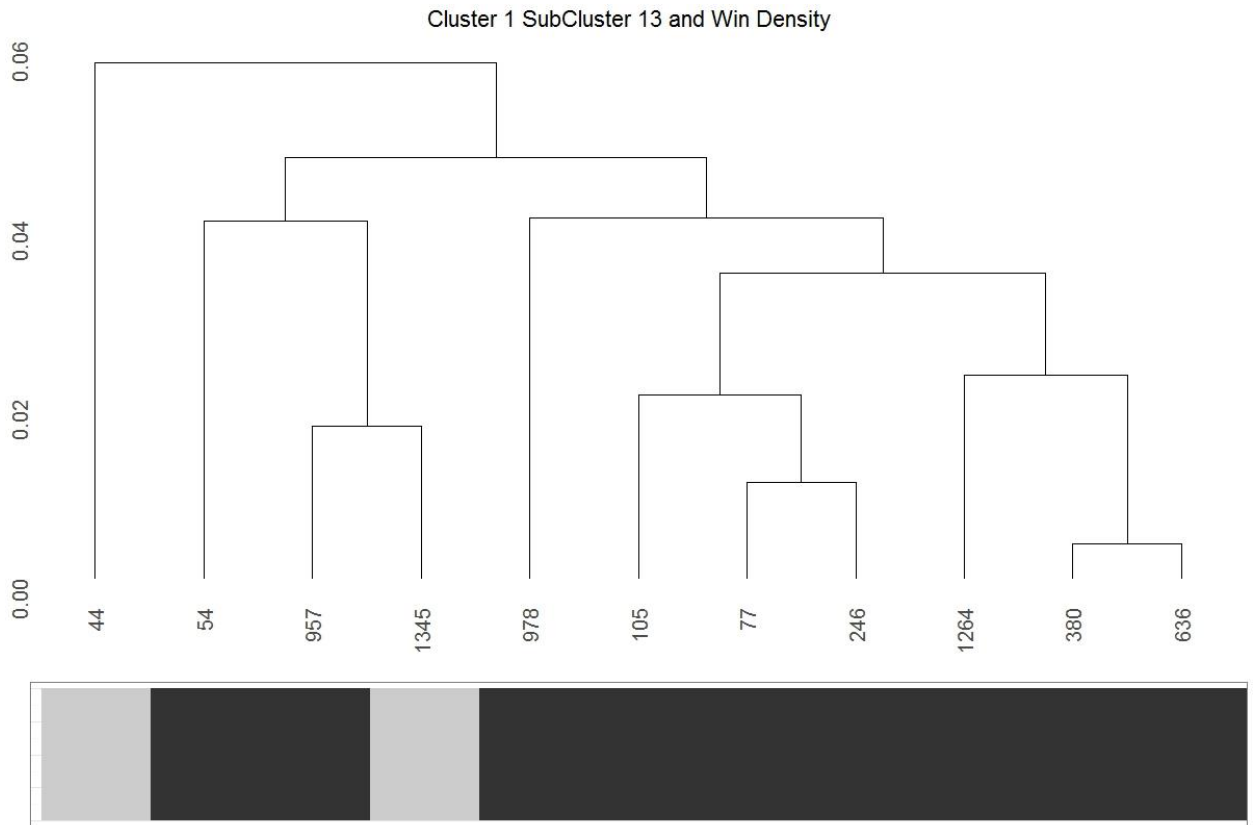
Figure 77. Cluster 1 Sub-cluster 13 dendrogram and win density.

Further analysis with regard to the positions occupied on the board as this sub-cluster of games was played indicates a somewhat compelling illustration of strategy as there appears to be some center board tendency found in Figure 78. With the exception of

the position in front of the Queen and the position of row seven and column seven, the frequency of positions played was most in the middle of the board. However, the center of the board is not crisply defined as in other clusters but there does appear to be a King side preference. The lack of definition is due to the range of the length of games found in this sub-cluster. The longest game exceeded 200 moves while the shortest game was less than 60 moves. This disparity in addition to the number of games that extend past any major Position values serve to muddle the observable trend seen in Figure 78.
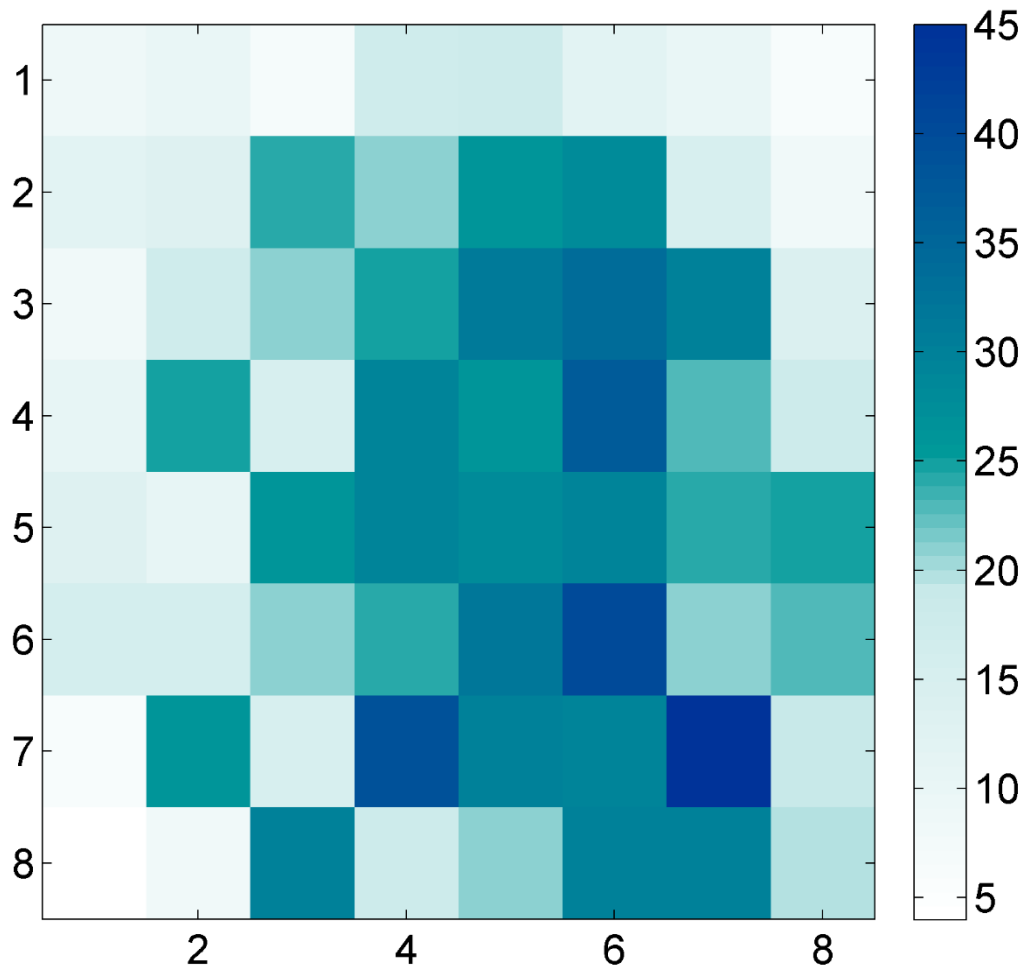
Figure 78. Cluster 1 Sub-cluster 13 Position histogram showing center board preference.

Next Sub-cluster 56 will be analyzed. This cluster, like Sub-cluster 13 has components that are similar with the most distant tree height of approximately 0.05 as seen below in Figure 80. This trend can be further demonstrated by the time series values seen

in Figure 79 below as the games appear to have a similar tempo making similar moves to achieve the same Position metric values at similar intervals. However, similarly to other wining clusters, this sub-cluster appears to reduce the amount of higher value enemy pieces early in the game. Specifically, the early attacks appear to be on the mid-valued pieces such as the Knight, Bishop and Rook. The maximum value of Position attained is seven, however there are more frequent values of six. These two trends are attained by a combination of Pawn, Knights Bishops and Rooks, whose numbers on the board exceed those of the Queen and King. This is a known strategic heuristic which relies on the development of pieces and the board before utilizing the queen [134]. While there are cases of specific attacks and games by chess masters, the conventional recommendation is to preserve high value pieces for the end game and not to squander them during opening and midgame if possible.

201
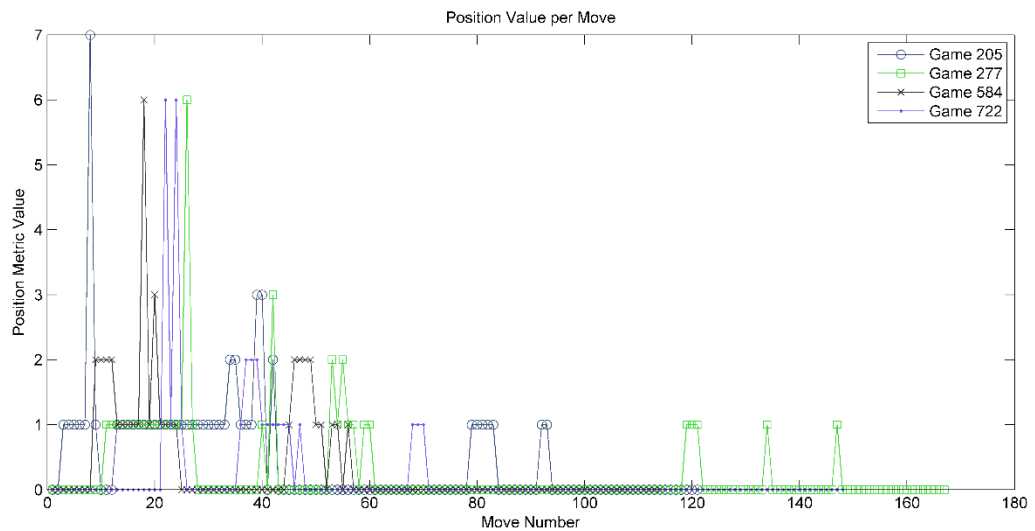
Figure 79. Cluster 1 Sub-cluster 56 Position metric time series showing similar tempo short games.
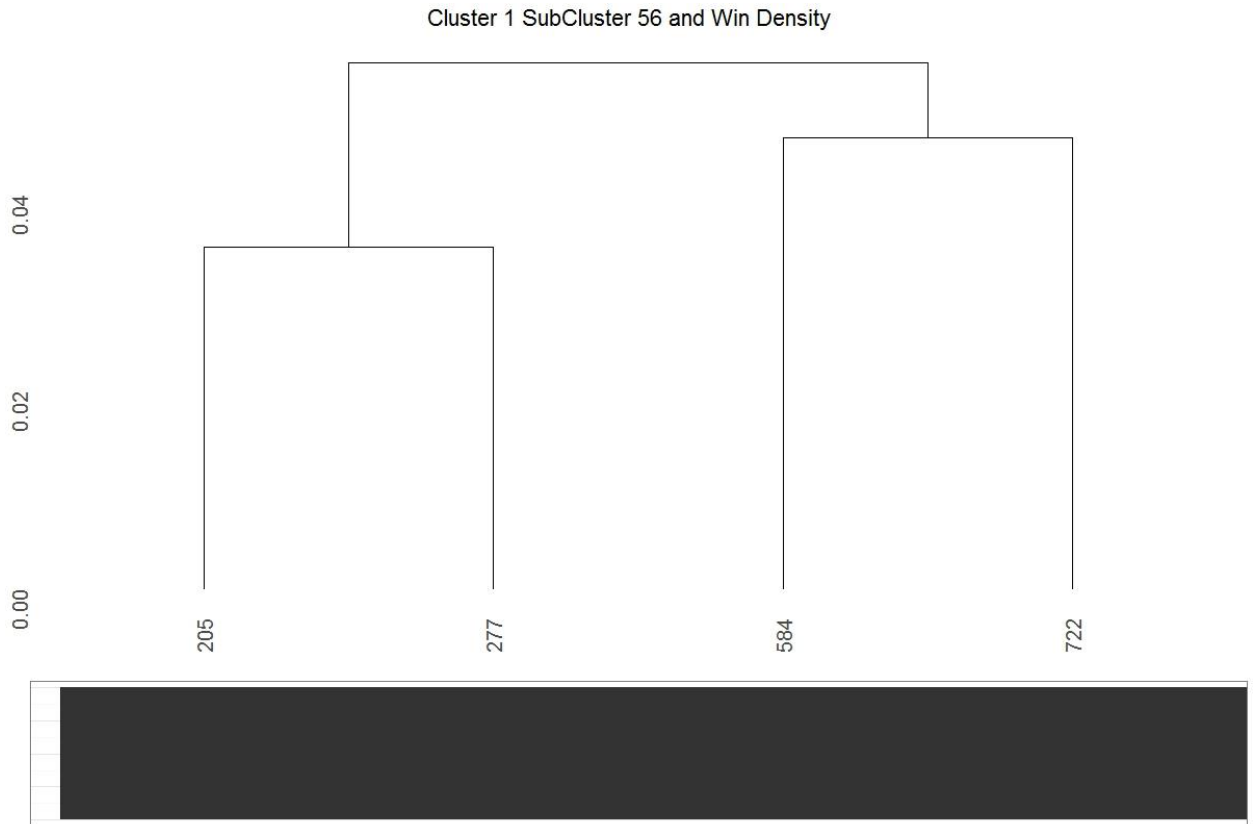
Figure 80. Cluster 1 Sub-cluster 56 Dendrogram and win density showing strong similarity.

In analyzing the board position history of Sub-cluster 56 in Figure 81, we can see what appears to be two boundaries on the board where a majority of the white moves are. The first appears to be center-board with an inclination to rear and King side. However, the second boundary appears to be on the black Queen side. Two inferences can be made

202

from this occurrence. First it seems as though the game play occurred near the black Queen side, therefore, it is that side that was under frequent attack. Secondly, the entire first row was utilized by the games in Sub-cluster 56. Specifically, the spot of row one column 3 was utilized approximately ten times. While this is a risky area to play due to highly mobile pieces existing in the area, it offers a chance for white to promote pawns and regain capability if high value pieces were lost in the game opening and mid-course.
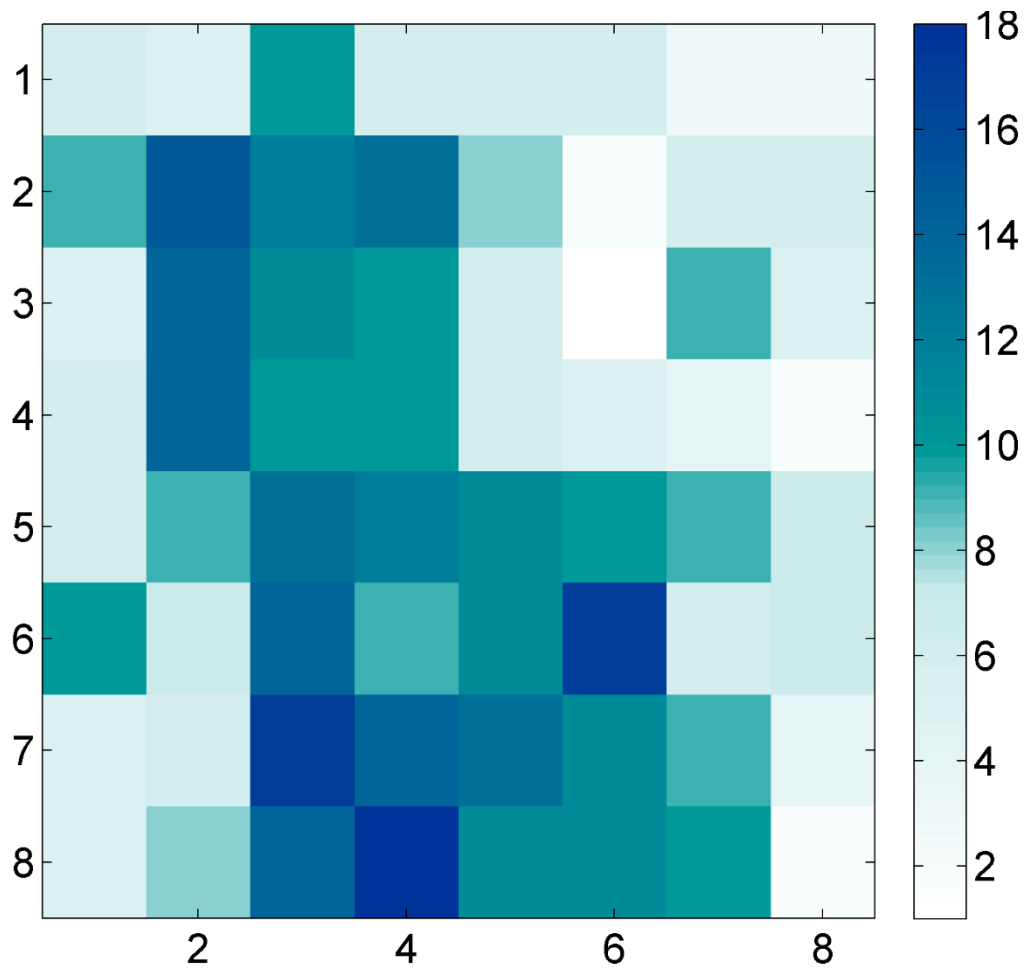
Figure 81. Cluster 1 Sub-cluster 56 Position histogram showing center board preference.

The next sub-cluster is similar to Sub-cluster 69 insofar as they both represent losing policies and are not strikingly similar as can be seen in the Position time series in Figure 85 and from the similarity measure in Figure 86. However, there appear to be trends in game play. The most salient observation can be seen in Game 156. Although, multiple attempts were made at high value enemy pieces early in the midgame, there were still relatively high Position metric values attained close to move number eighty. This represents a failure to complete the attack on multiple enemy pieces, which ultimately resulted in the loss for that game. Games 134 and 232 were able to make early attempts at high value enemy pieces in the opening of the game signifying a failure to properly develop the board, ultimately resulting in the loss of pieces to the risky early attacks. The final game to consider is Game 425 which is an outlier as can be seen in the Dendrogram below. This game fails to acquire positions of value until approximately move 30, but is able to sustain an attack on a pawn for five moves until a move is made that increases the value to eight. However, the game ends abruptly on the next move where the value only drops by a value of one. This game is confirmed to be a Checkmate for black. In summary, although statistically significant, this cluster does not contain elements similar enough to draw any cluster specific conclusions.
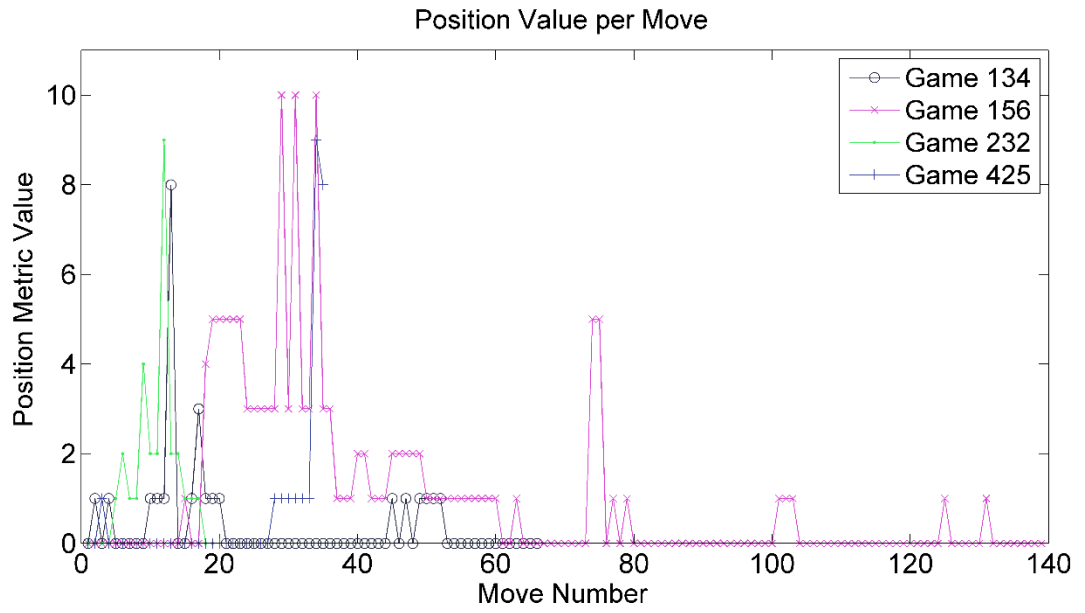
Figure 82. Cluster 1 Sub-cluster 87 Position metric time series for short games.

Figure 83. Cluster 1 Sub-cluster 87 dendrogram and win density with weak similarity.

In analyzing the board position frequency plot a peculiarity is found. As the Dendrogram and Position time series indicate a lack of similarity, it is expected that the position histogram would exhibit more randomness than the other histograms. However, this is not entirely the case as Figure 84 is investigated below. There appears to be asymmetric game play on the board favoring the King's side. Additionally, there appears to be no center board preference and atypically, the most played position is in row two and

column six which is directly next to the enemy King's original location. This appears to reinforce the overwhelming loss of material of the white side as a result of failure to develop the board.
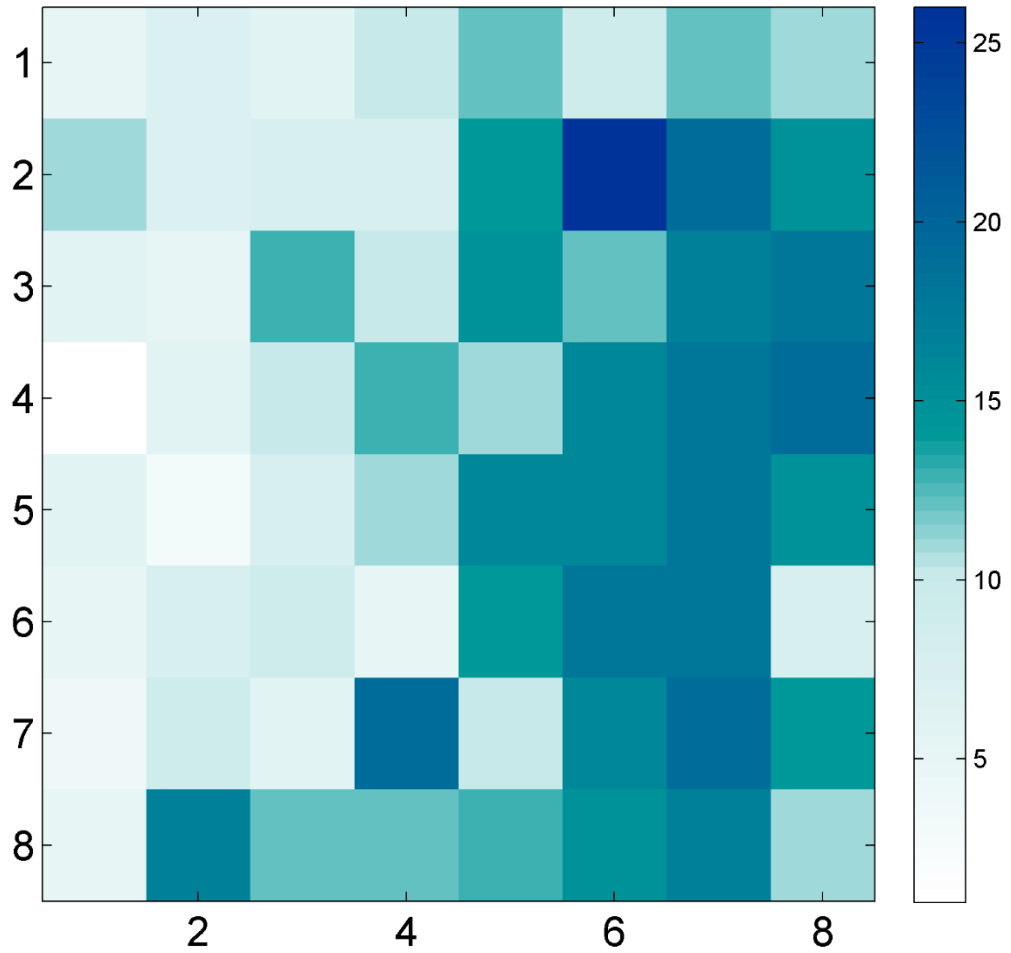
Figure 84. Cluster 1 Sub-cluster 87 Position histogram without center board preference.

The final sub-cluster to be analyzed is number 20. This sub-cluster, similar to the previous cluster, is a losing policy. However, upon viewing Figure 85 and Figure 86 it can be seen that this cluster is a bit less significant but includes elements that are much closer in similarity. Specifically, most of the games plotted below, attain a Position value of at least seven within the range of game opening. However, this trend continues far into the game mid-course indicating a failure to effectively act on an achieved strike position. Another attribute of this cluster enabling the losing trend is the off-center game play displayed in Figure 87. In addition to playing off-center the majority of the Sub-cluster's moves are on the King side but to the rear of the board with the most commonly played position is the original location of the King's Knight.
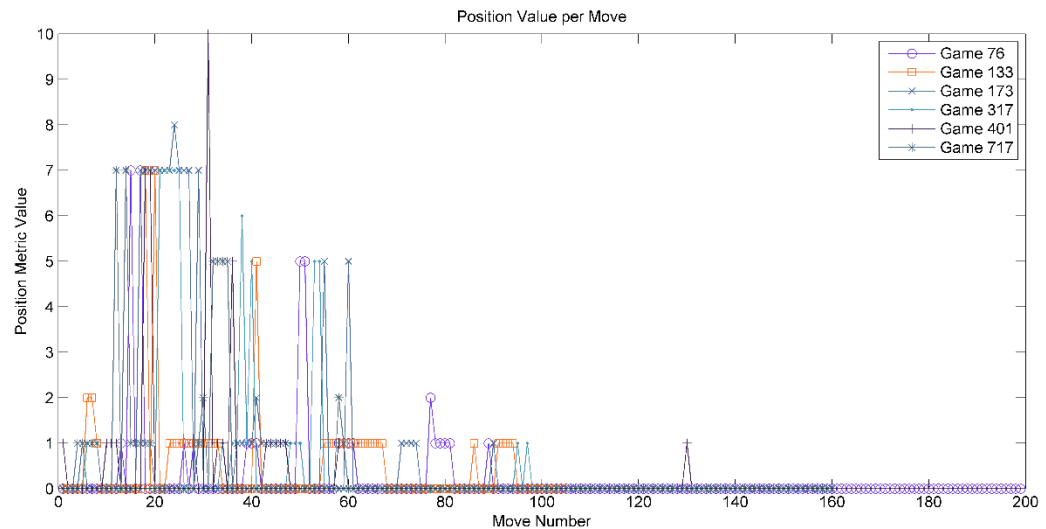
Figure 85. Cluster 1 Sub-cluster 20 Position metric time series showing cluster preference to achieve high attack tempo and maintain into mid-course.

208

Figure 86. Cluster 1 Sub-cluster 20 Dendrogram and win density.
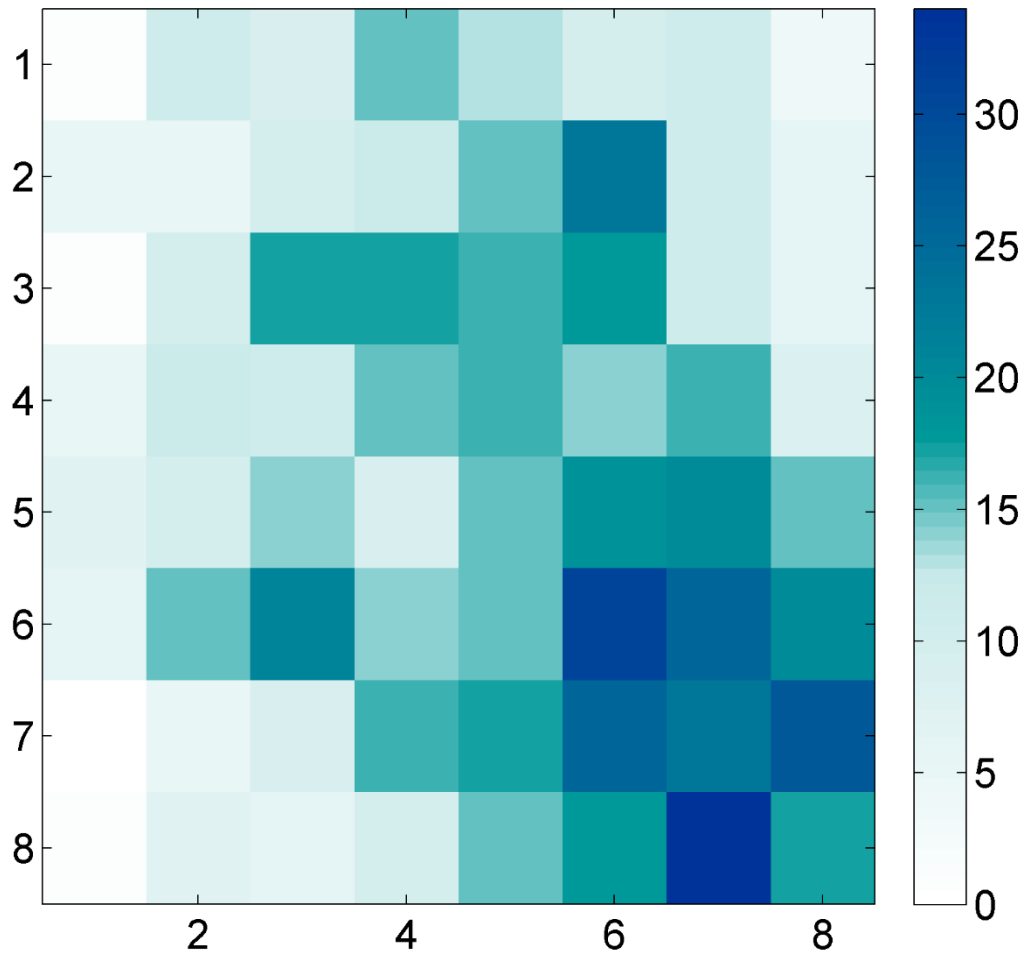
Figure 87. Cluster 1 Sub-cluster 20 Position histogram showing lack of center board preference.

5.3 Experiment 3

As was seen in the culmination of the previous experiment, the results and specifically the clusters formed showed a spectrum of significance and closeness yielding

different levels of applicability to further analysis. In all cases where either the significance was insufficient, or the similarity within the clusters was inadequate, the only resolution is to attempt to extract more data by performing more random runs. However, there were still actionable results found from the original dataset. Conversely, there still is the question as to what happens to the analysis and how it is degraded by considering a reduced data set. The practical reason for why this would happen is if there is a sufficiently complex situation that must be performed to gain understanding. If there is substantial cost in running this simulation, there may be an upper limit on the initial investment to M&S. Therefore, in cases where an exhaustive or substantial search of the decision space is not feasible, the effect on the analysis must be investigated. Furthermore, the effects on HSig index and the Optimization process due to variations in the availability of data must be known.

The in initiation of experiment two begins with the reduction of the total data source. Initially, only 400 of the total 1300 cases were selected. In the next case 800 out of the total 1300 cases were selected and further analyzed. In both cases, this was done by performing random sampling and collecting random samples of the initial runs. This was then processed for analysis by pruning the rows and columns of the initial proximity matrix found by performing pairwise Dynamic Time Warping. By identifying two random samples of indices, two vectors consisting of the indices of the original data set were found enabling the creation of the two reduced datasets. Afterwards the HSig Index was

211

calculated for a range of possible number of clusters assuming utilizing the Average Linkage Algorithm was still feasible for both data sets. The investigation paralleling the last experiment follows.

The relative quality of the clustering can be assessed by comparing the similarity heights to the left of the figures below. This is telling of how similar the elements of each cluster are. If two elements are not relatively similar, but are the two most similar cases present, they will be clustered together. Given a reduced dataset, it is likely this will be the case, as clustering on a sparse dataset will result in a best grouping structure given the data. Therefore, measuring the effect of reducing the size of the dataset will be done using the Dendrogram height as well as the significance values.

The HSig Index appears below in Figure 88 and still displays unimodal behavior allowing the identification of a minimum and an optimum number of clusters. The Dendrogram with win density is shown below in Figure 89, however, does not indicate clusters regions of increased win density. Additionally, when performing HSig Optimization, no significant clusters are found unless the significance constraint is relaxed. Relaxing this constraint allows one cluster that has a p-value of more than 0.1 with the remaining statistics shown in Table 14. Furthermore, assessing he Dendrogram of this cluster, we find that the cases represented are much more dissimilar than the ones seen

resulting from the complete dataset, impeding the extraction of any trends from the data of this cluster.
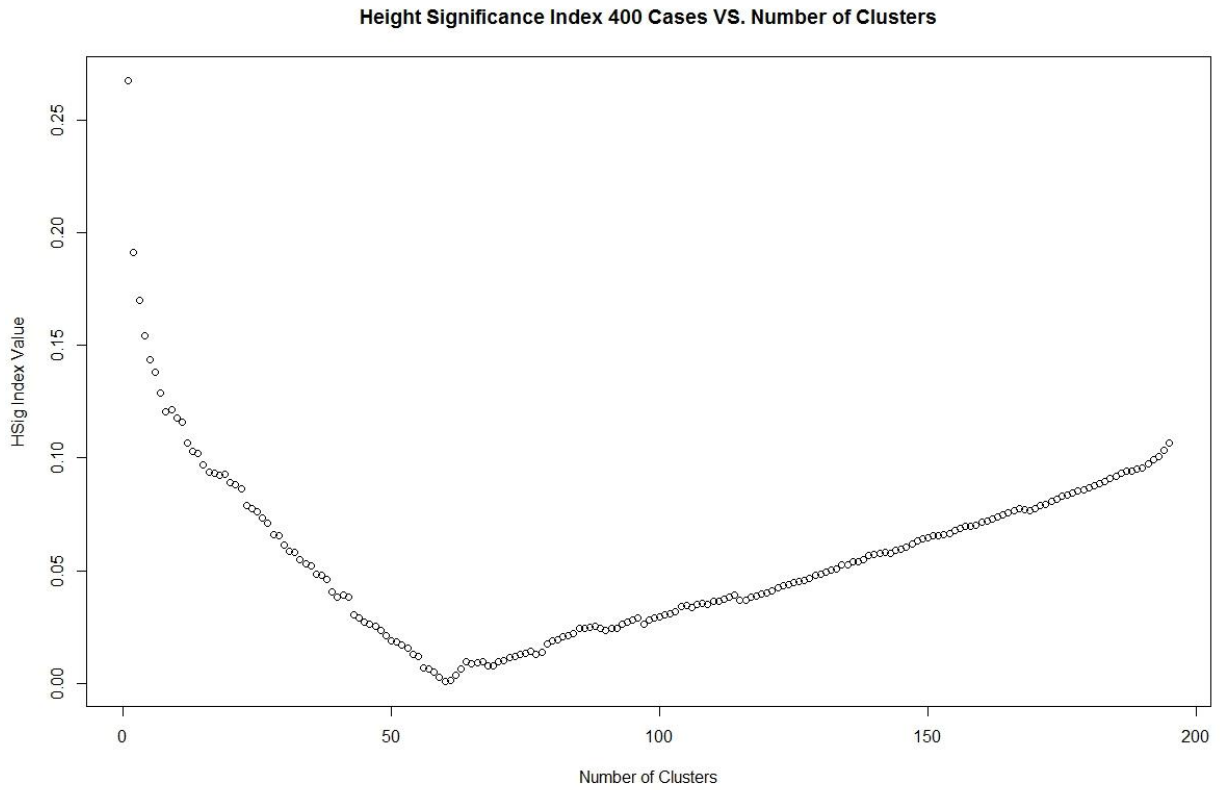
Figure 88. HSig Index values for 400 case reduced dataset showing clear minimum.

Figure 89. 400 Case Dendrogram with win density showing regions of increased win density and regions of decreased win density.

Table 14. Significant Cluster from 400 case dataset.

Cluster Number	14
p-value	0.105
Number of trials	16
Number of successes	11
Polarity	0.6875

214
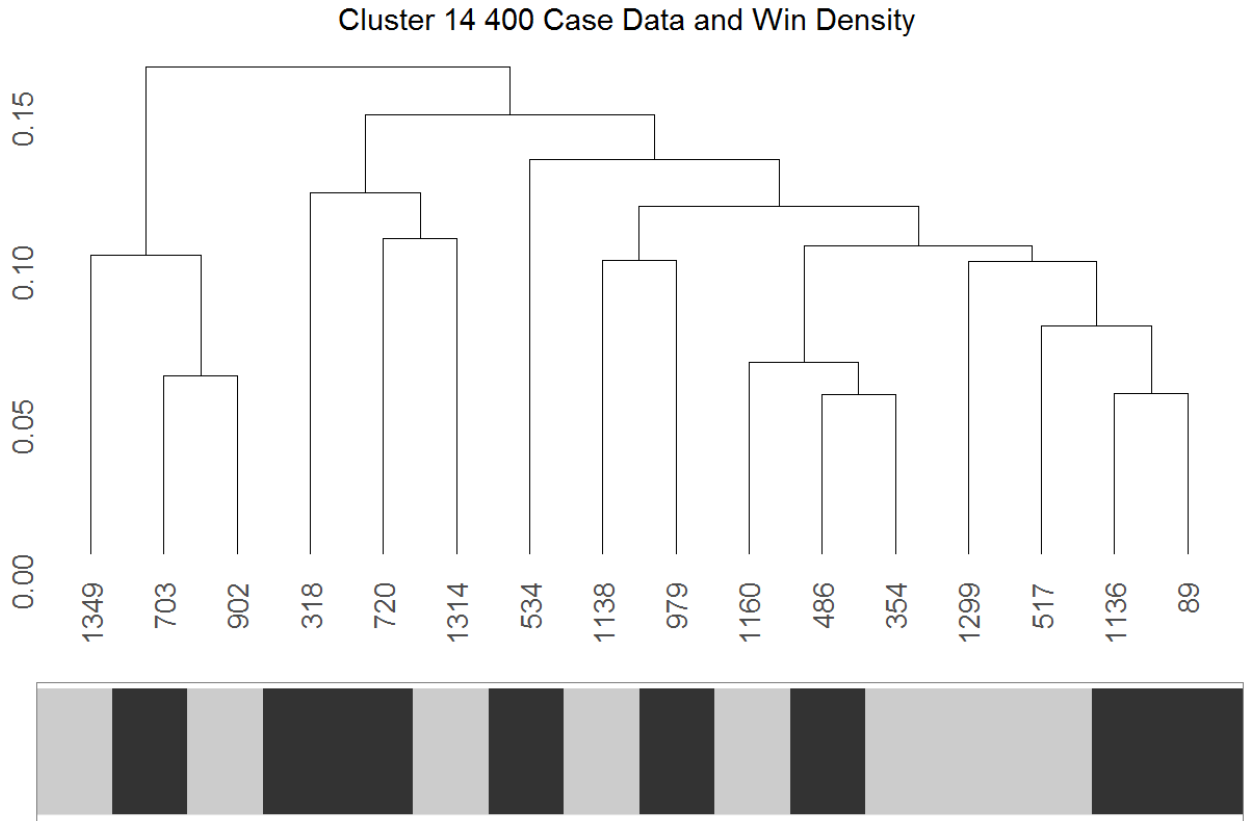
Figure 90. Cluster 14 of 400 cases and win density.

Similarly, to the 400 case example above, this process was performed on a data set twice the size of the initial reduced size. It too displayed unimodal behavior resulting in the successful identification of an optimum number of clusters and can be seen in Figure 91. This example, however, did not require relaxation of the significance constraints in order to identify clusters, revealing four significant clusters shown below in Table 15. However, an observation comparing this to the case of full data in the last experiment, is

the size of the clusters was greater in the full data set. The analysis of the Dendrogram
next will allow a comparison to the quality of the clusters.

Height Significance Index 800 Cases VS. Number of Clusters
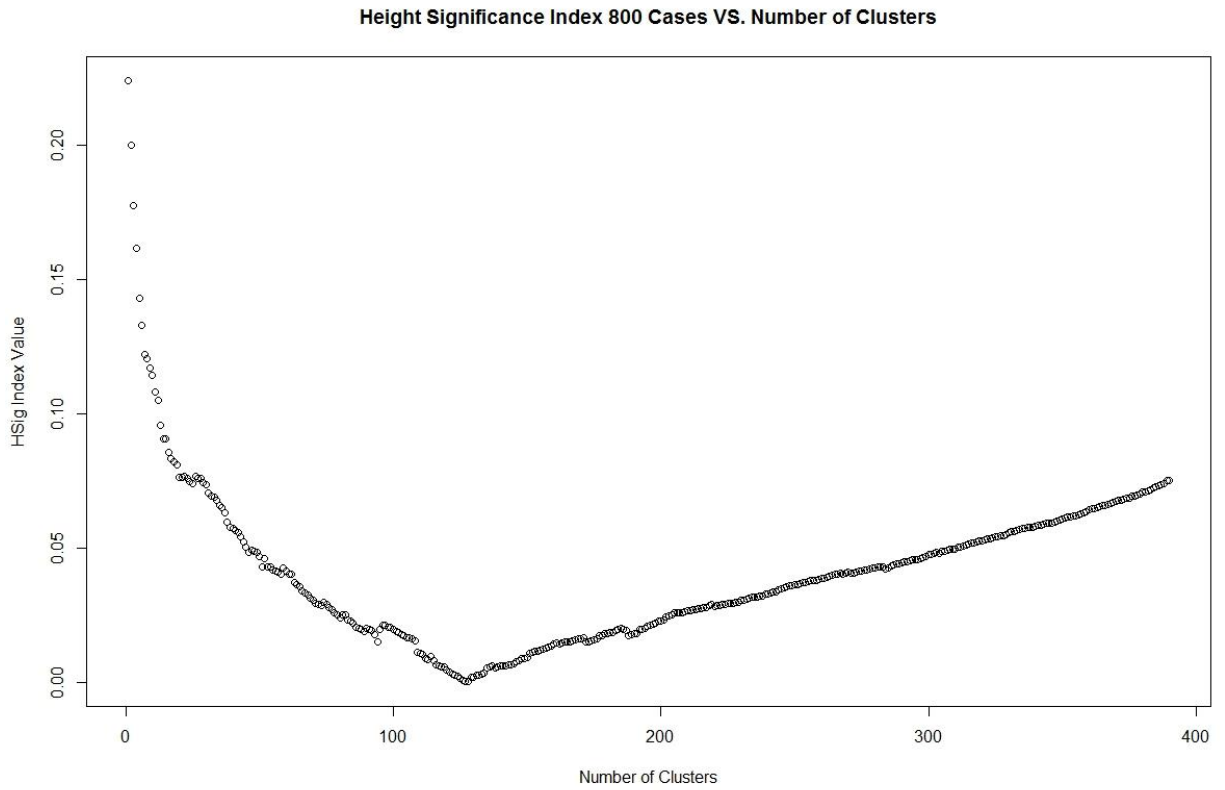
Figure 91. HSig Index values for 800 cases.

216

Table 15. Significant clusters from 800 case dataset.

Cluster Number	47	20	32	63
p-value	0.048	0.062	0.062	0.062
Number of trials	18	4	7	4
Number of successes	13	4	6	4
Polarity	0.722	1	0.857	1

The analysis of the Dendrograms associated with the significant clusters of the 800 case

data reveals only one cluster that appears to have adequate inner cluster similarity. Cluster

47, below in Figure 92 is both significant and shows similarity warranting further analysis.

However, the remaining clusters of Table 15, Figure 93, Figure 94, and Figure 95

demonstrate substantial inner cluster separation reducing the quality of the analysis. While

four clusters appeared to have significance, only one cluster was able to be carried forward

in the analysis. This is in contrast to the eight clusters of Experiment 2 that had less than

a 0.15 Dendrogram height.
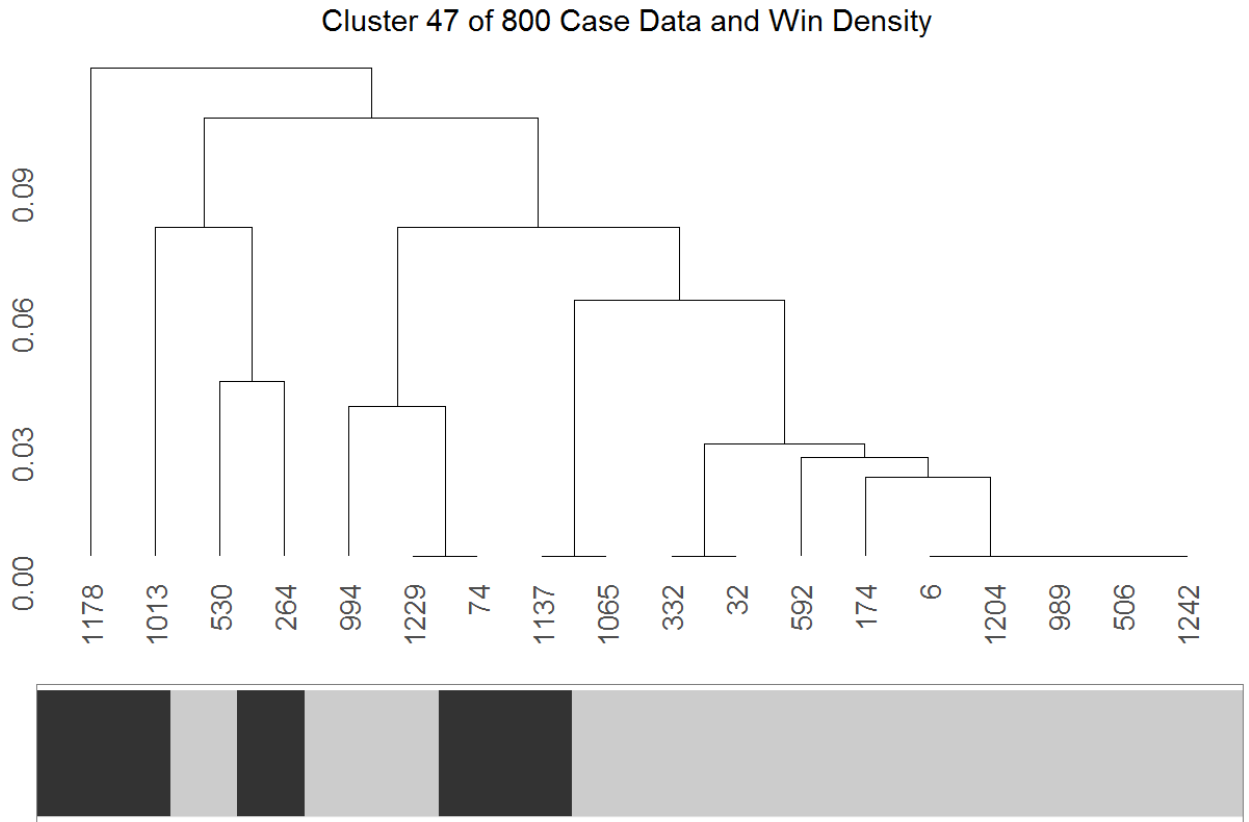
Figure 92. Cluster 47 of 800 case data and win density.
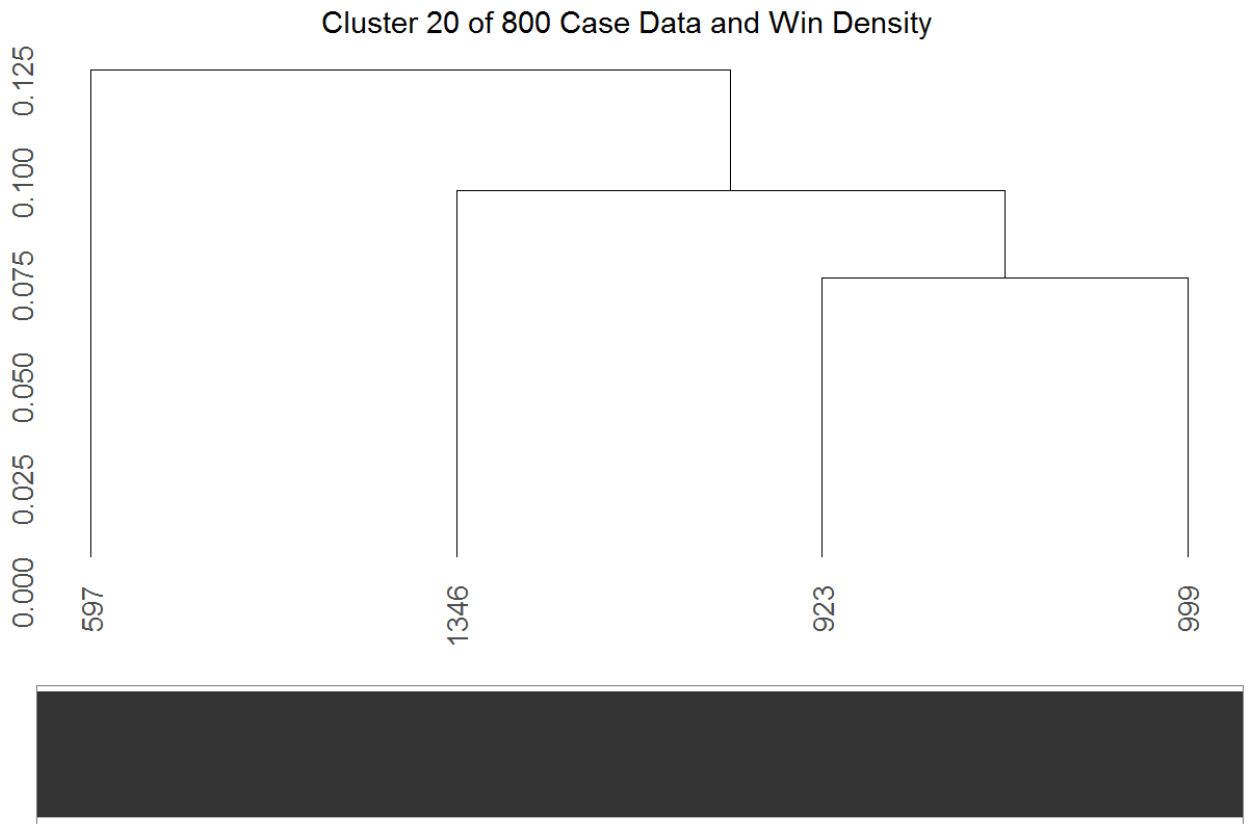
Figure 93. Cluster 20 of 800 case data and win density.
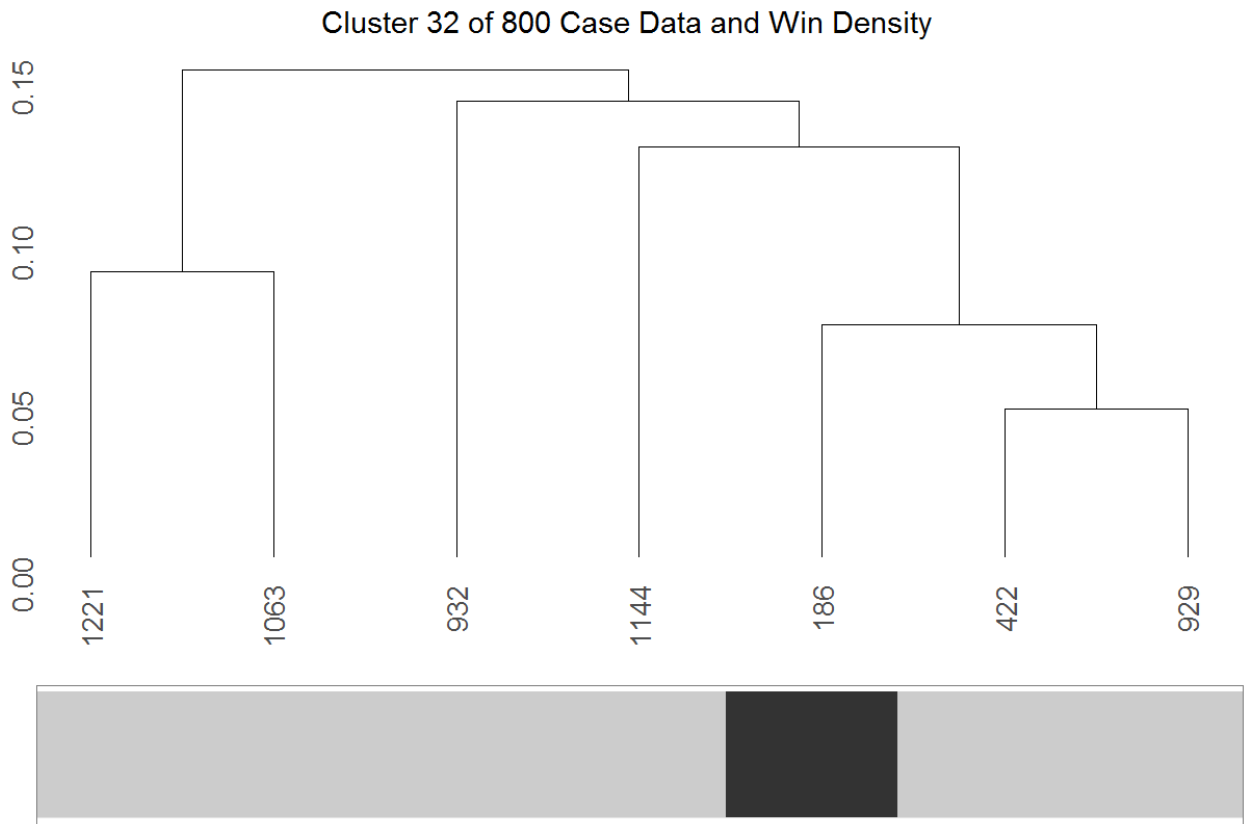
Figure 94. Cluster 32 of 800 case data and win density.
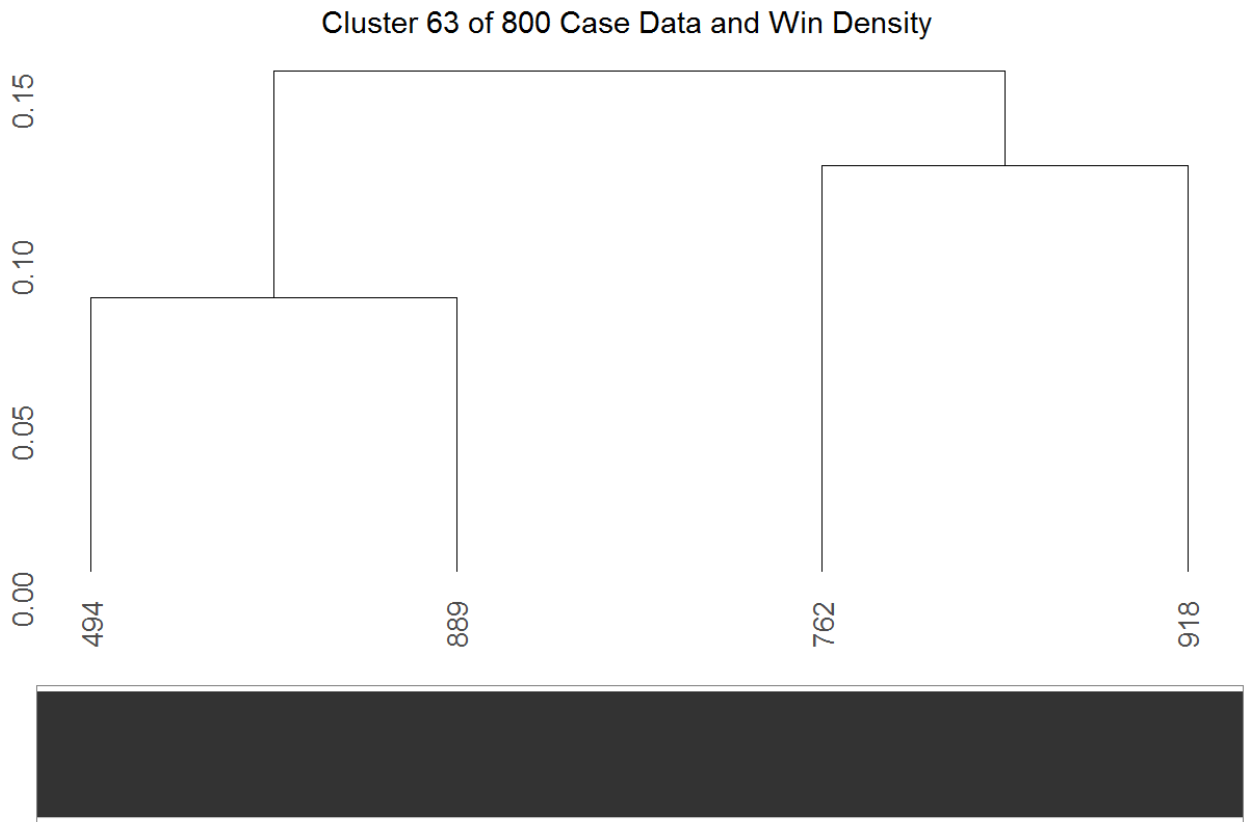
Figure 95. Cluster 63 of 800 case data and win density.

5.4 Application Problem

This is the final experiment and it will serve to close the loop on the methodology. The primary objective of this experiment is to create a ground combat scenario that can by simulated via an agent based simulation. Specifically, MANA will be utilized to run the defined simulation varying the list of potential COA elements randomly. Through extraction of simulation time series data of the metrics defined previously for a combat

scenario, Position, Awareness and Capability, a set of three time series data sets will be created to perform the remainder of the methodology.

Each of the datasets will be transformed into a similarity matrix utilizing dynamic time warping at which point the most applicable linkage algorithm will be selected based on the results of the HSig Index for a variation of potential number of clusters. After the optimum linkage algorithm and corresponding number of clusters is defended, HSig Optimization will commence, at which point significant clusters will be selected for further analysis. Finally, the Candidate COA will be selected and be contrasted from the COAs in other clusters to define the Critical decision point.

This experiment relies on the simulation of ground combat subject to operational and environmental constraints. Specifically, the scenario is meant to illustrate geo-political as drivers to military maneuvers. Additionally, terrain elements and elevation will be utilized to influence the rate of change of the simulated forces in the simulations and is defined in the next paragraph.

The country of Suisun Valley is unsurpassed in the quality wine grapes and artisan wine produced. The neighboring and larger country of Napa Valley, is second to Suisun in quality, however, it is the leading exporter of grapes and wine to the region. In order to control the entire market, Napa has invaded Suisun commandeered its wineries and has

taken the capitol of Suisun, the Suisun Valley Wine Coop. Napa is currently occupying the region of Suisun with a substantial presence in three locations preventing unencumbered access to the wineries of Suisun. The first location is represented by the largest red star in Figure 96 and is the location of the Suisun Valley Wine Coop which is being utilized by Napa as the headquarters. The next location is represented by a single red star in Figure 96 represents a road guard, protecting the enemy positon from southern attacks, and by preventing access to the Coop via easy going terrain. The next enemy position is represented by a cluster of three red stars and represents a roving patrol that traverses the route intersecting each of the stars. This position prevents access to the Coop from the East and North via easy going terrain.

A neighboring ally of Suisun, Sacramento has staged a forward operating base North West of Suisun, represented by a large blue star in Figure 96 and has commenced planning for an invasion in order to repel Napa forces from the area. Given a starting point and targets defined previously, the combatant commanders of Sacramento must utilize their resources, applying them towards targets of varying sequence to identify the greatest chance of success.
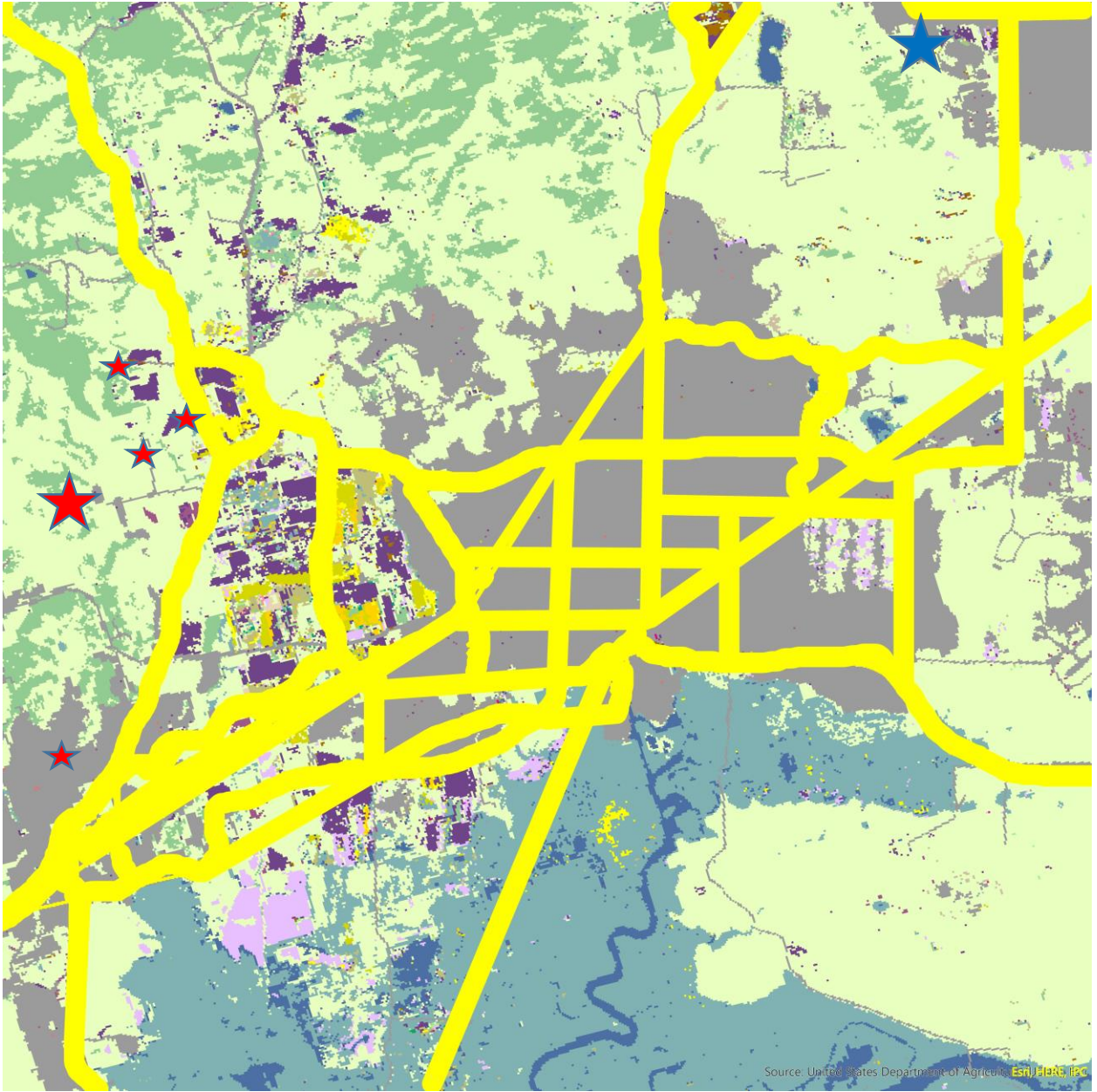
Figure 96. Suisun terrain map showing terrain features, farm land types and targets utilized[135].
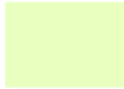
	Going	Cover	Concealment
	1	0	0
	0.75	0.1	0.3
	0.2	0.3	0.9
	0.1	0	0
	0.4	0	0

Figure 97. Terrain influence on agent movement through simulated battlespace.

In addition to the operational and geo-political constraints imposed as defined in the previous paragraph, there environmental constraints as well. The area for which this operation will take place has unique attributes that make this problem distinct. In order to capture the environmental factors, location specific data must be utilized to illustrate the effects. This was accomplished by utilizing ArcGis, a geographic information system that utilizes datasets provided by the U.S. Geological Survey [135]. With this capability, the terrain data specific to the Suisun and Napa Valleys, was added the maps of the region allowing a more realistic arena for the agent based simulations to operate. Specifically, for

reach type of terrain featured in Figure 96, there will be a different set of conditions for travel by the agents, which are depicted in Figure 97. The three influences on the propagation of the agents through the battlespace are Going, Cover and Concealment. Going defines how the Agent's speed is affected by the terrain. Cover defines the degree to which the environment can protect an agent from being shot. An example of Cover is being shielded by a wall from enemy fire. Concealment is the degree to which the terrain shields the agent from being seen.

The specific implementations of the color scheme and terrain features are described according to the scenario shown above. The yellow and blue areas which represent major roads and lowlands, shown in Figure 96 will allow unencumbered travel for the agents. The light green, light vegetation, will offer slight impedance, but will enable more concealment from enemy agents. The dark green, representing mixed woodland, will have more impedance and more concealment over the light green. The grey, representing urban areas will offer slight impedance and no concealment. Most importantly, the purple areas represent vineyards, are designated 'no-go' zones per the rules of engagement, and will offer no concealment and a substantial impedance for agents traversing. Additionally, elevation information was utilized for the defined area affecting line of sight for the agents, both friendly and enemy and is shown in Figure 98. For this specific scenario, friendly agents will have perfect information of other friendly agents, however, if line of sight between an enemy and friendly agent is broken by elevation features, neither will have

226

information about the other. The remaining simulation specific details are stored in XML

files attached in Appendix B.

Figure 98. Suisun elevation data where lighter features are at increased elevation levels
over dark regions[135].

The features discussed in the previous paragraph will affect the performance of a specific plan. Additionally, the relative performance will also be identified as more plans are assessed. Specifically, each COA will vary the approach to different targets as well as vary the sequence of which targets are attacked. However, the defending force attributes will remain static, as will the enemy force. The implementation of this approach is done in three phases. First, the routes to targets as well as target sequencing is created. Next MANA is utilized running the predefined scenario with the unique target and sequencing data for the blue team. Finally, the data from each simulation is collected.

In order to assess the effects of terrain and sequencing on the overall performance, a plan must be made. This plan is a COA and in this situation, it is made up of the routes to the sequence of targets. A series of twelve routes to the three targets was made. Then, for each of the twelve base paths, they were duplicated and injected with noise until there remained 600 paths or COAs which can be seen below in Figure 99. Each of the COAs were inputs to an XML file read by MANA and simulated, capturing time series data. Specifically, the positions traversed by the agents and the time and location of blue casualties were captured for further analysis. Each of the simulations halted once the objective was reached, neutralizing all enemy targets, or when all blue agents were killed. The definition of this scenario as well as the implementation of MANA yields the same form of the data analyzed in the previous two experiments. Therefore, the CADRE methodology will be utilized to identify CDPs given this scenario.

228
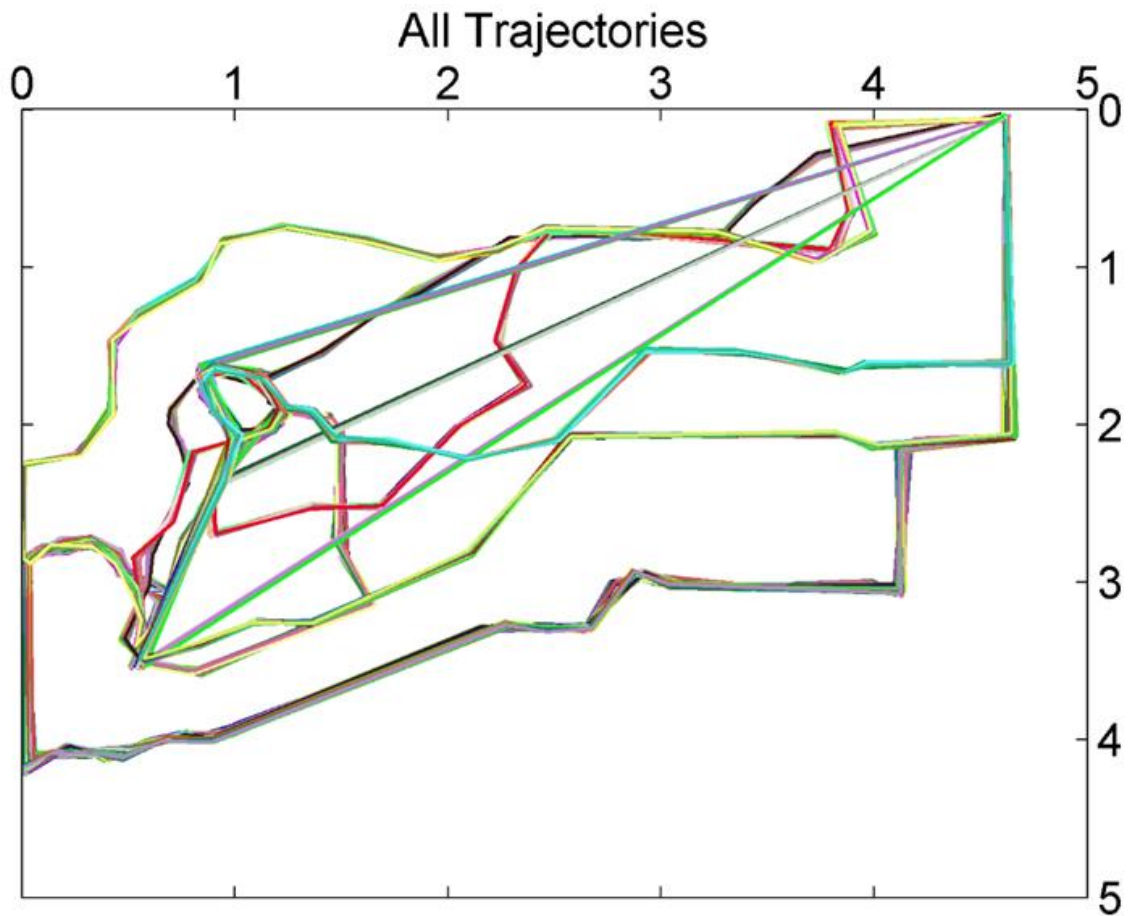
Figure 99. All trajectories traversed by agents.

5.4.1 CADRE Methodology Application

The first portion of the methodology is to identify which linkage algorithm allows the most significant analysis of like COAs. As a result, the HSIG Index was utilized for all linkage algorithms in order to identify a 'best-in-class' with respect to the given data.

The first dataset analyzed is the Capability dataset, which is the time series of capability degradation of the blue agent team due to attrition forced by the red agents.

After analyzing the data with the different linkage algorithms, the Median linkage algorithm was identified as offering the most statistically significant clusters. Utilizing this linkage algorithm yields 35 as the optimum number of clusters, four of which are significant, warranting further application of HSig Optimization and are presented below in Table 16 and the remaining clusters are in Appendix A. While the polarity is not high, the significance is noteworthy and an analysis of the formed clusters will follow to assess the applicability to determining an optimal COA and identifying a CDP.

Table 16. Significant cluster capability data.

Cluster Number	15	6	20	3
p-value	0.000919	0.006642	0.044766	0.050568
Number of trials	135	48	35	84
Number of successes	86	33	23	50
Polarity	0.637037	0.6875	0.657143	0.595238

The first of the clusters selected for analysis is cluster is Cluster 15, which is the most statistically significant cluster among all shown in Table 16. The first observation of the Dendrogram associated with this cluster in Figure 100 is that the polarity is inclined

towards losses. Additionally, there is a large variation in the similarity heights within the cluster weakening the inferences possible in this cluster. Observation of the Cluster 15 Position trajectories shown in Figure 101, shows a sampling of all possible combinations attack plans generates. Therefore, although the cluster is significant, there is not enough information to enable decision-making. As a depth first approach, this cluster was assessed for bounding through HSig Optimization and was found to have two significant sub-clusters which are shown in Table 17.
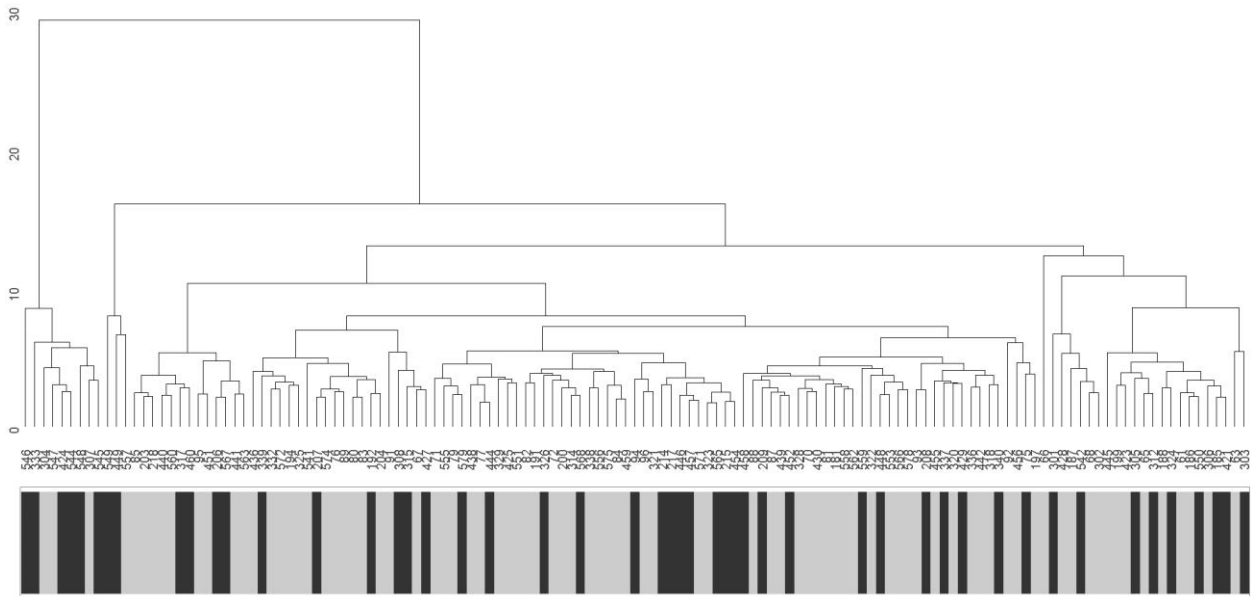
Figure 100. Cluster 15 Capability dendrogram and win density.
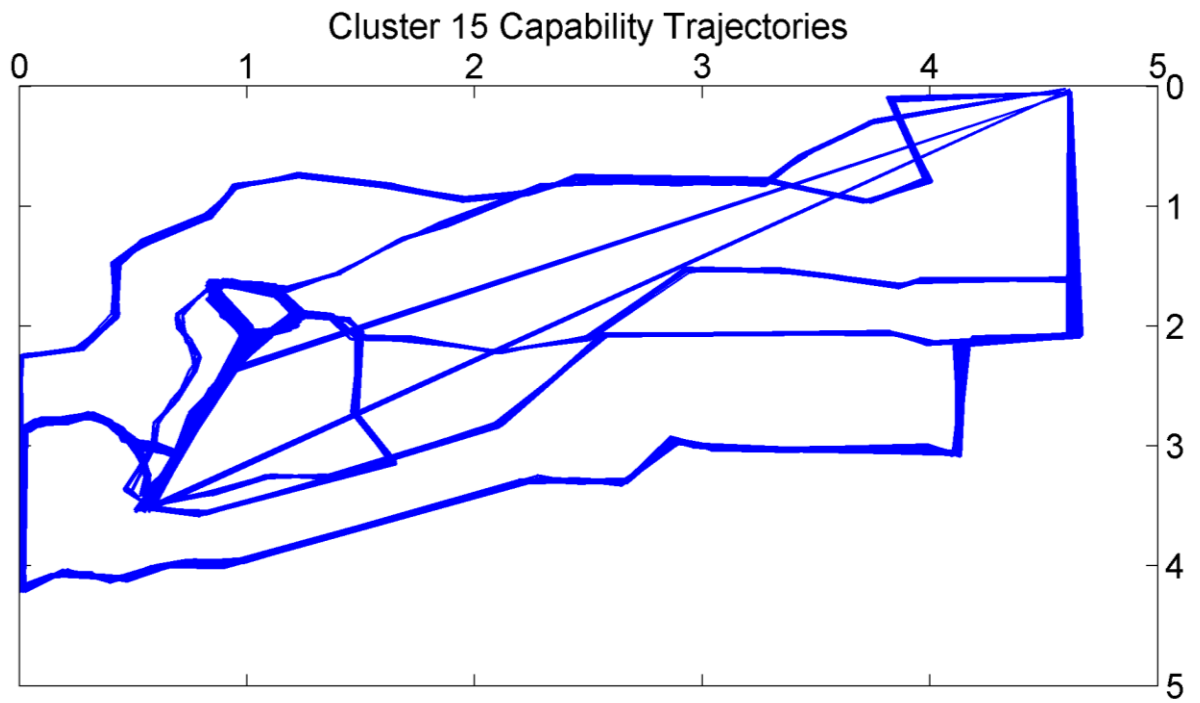
Figure 101. Cluster 15 Capability trajectories.

Table 17. Cluster 15 Sub-cluster statistics.

Sub-cluster Number	9	7
p-value	0.019287	0.031784
Number of trials	12	19
Number of successes	10	14
Polarity	0.833333	0.736842

The first of Clsuter 15's sub-clsuters to be analyzed is Sub-cluster 9 pictured in Figure 102. Well within significance bounds, the members of the clsuter are collected and the trajectory is plotted below. This plot contrary to Figure 101, exhigbits some behavior that provides some actionable informtion, however does not indicate a preference for a particular COA as a majority of the COAs are still represented in this plot. This cluster does seem to offer more information than the Clsuter 15's other Sub-cluster, Sub-cluster 7 pictured in Figure 103. Converseley, one can see that in both plots, there is a high density of trajectories int eh region joining all of the enemy targets. Therefore, it appears that, all of the COAs represented in Cluster 15 and its Sub-Clusters have a common problem area.
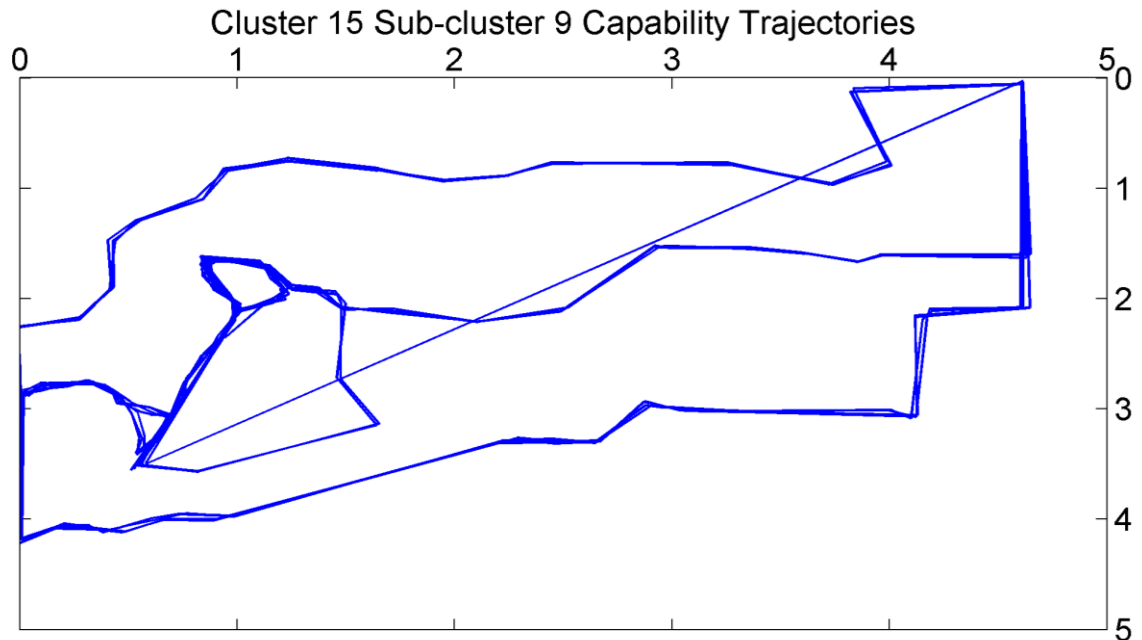
Figure 102. Cluster 15 Sub-cluster 9 Capability trajectories representing a subset of all trajectories and a potential down selection of potential COAs.
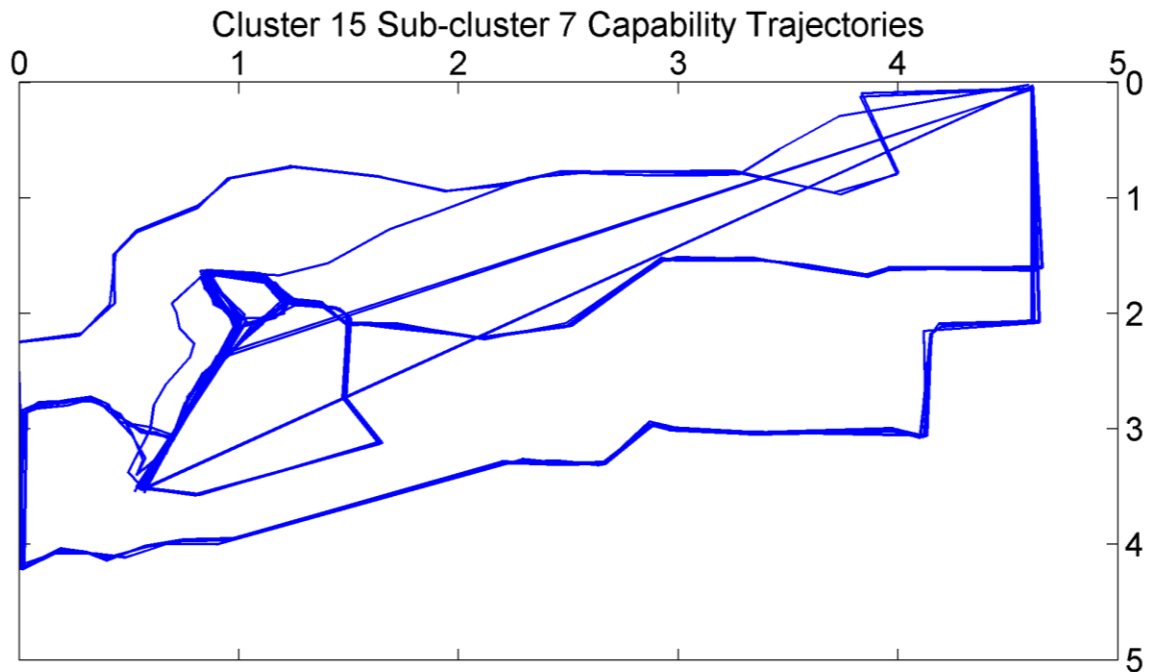
Figure 103. Cluster 15 Sub-cluster 7 Capability trajectories representing a larger subset of all trajectories indicating a diluted representation of a set of candidate COAs.

The next Cluster to be analyzed is Cluster 6 shwon in Figure 104. While this cluster is inclined towards an increased win density contrary to Cluster 15, the within cluster similarity is poor. Just as in Cluster 15, this cluster was operated on using HSig Optimization resulting in only one significant Sub-cluster.

The Statistics of Sub-cluster 2 of Cluster 6 is shown below in Table 18. Within the bounds imposed on significance, it appears to be inclined towards increased win density. However, analysis of the within cluster trajectories shown in Figure 105 are not telling. An interestin observation to notice in this plot of the trajectories, is the heaviness surrounding the roving patrol. In previous plots, all paths connecting the enemy positions

234

were heavily represented while this plot appears to represent the roving patrol moreso. However, it can be concluded that utilization of the Capability metric for this scenario and the HSig Index, and Optimization, was not successful in determining a set of candidate COAs or CDP for which to continue analysis.
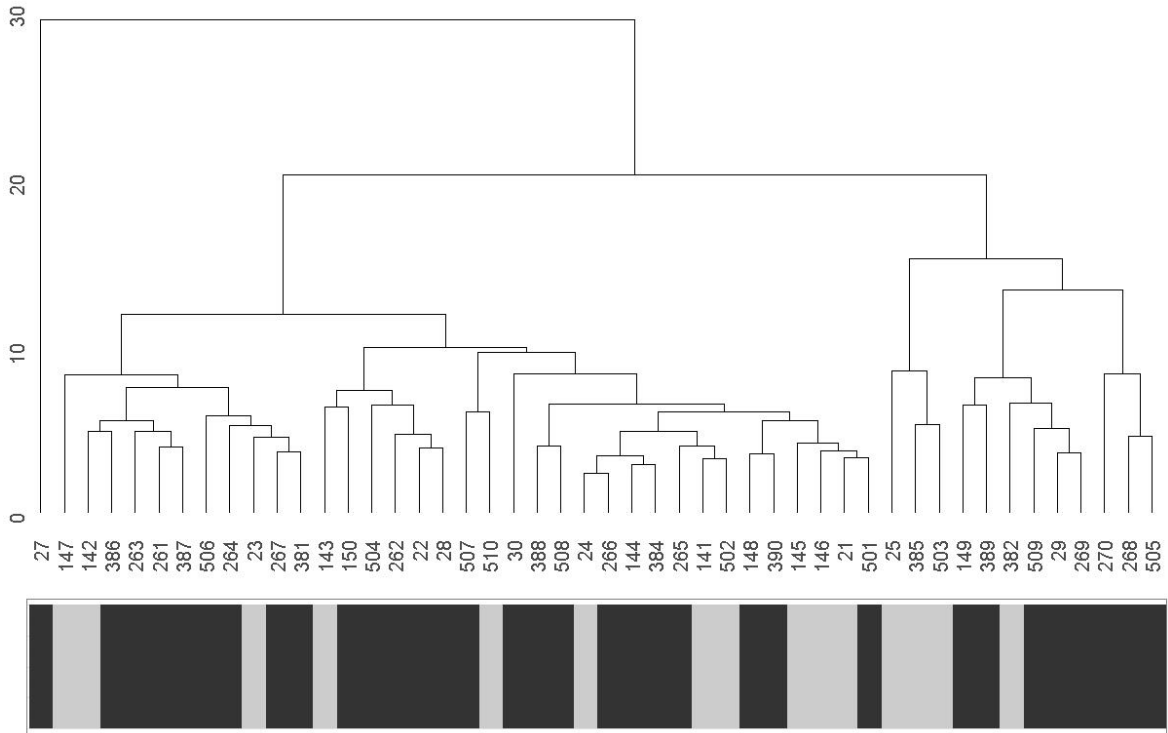
Figure 104. Cluster 6 Capability dendrogram and win density.

Table 18. Cluster 6 Sub-cluster 2 Capability statistics.

Sub-cluster Number	2
p-value	0.0625
Number of trials	4
Number of successes	4
Polarity	1
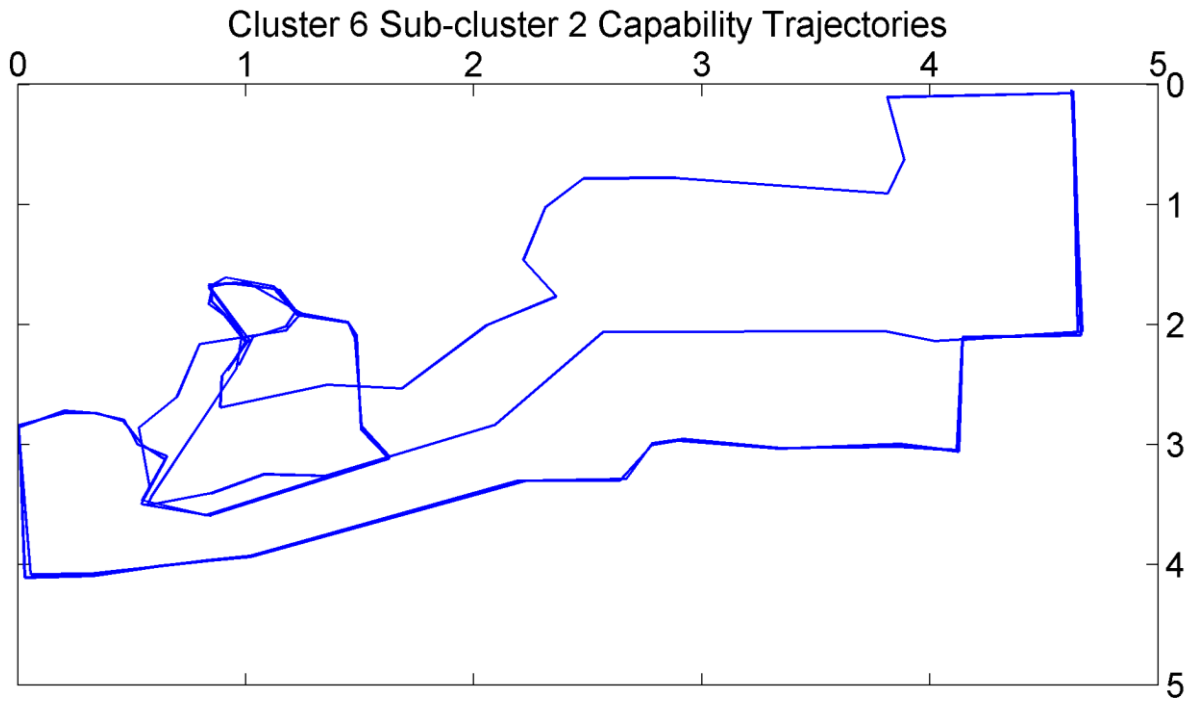
Figure 105. Cluster 6 Sub-cluster 2 Capability trajectories representing a reduced subset of total trajectories showing a more vivid representation of comparable COAs.

Next this methodology will be performed on the Position metric in another attempt to determine advantageous COAs and define CDPs. The simulation background is defined as 5000 by 5000 pixels which can be represented by coordinates. The Position time series presented here are the actual coordinates traversed by the agents during each simulation. These metrics are collected and translated to a set of linear indices, at which point, they are utilized to find the pairwise distance between all simulations utilizing the DTW algorithm. This pairwise distance matrix is fed into the HSig Index and Optimization process. As a result of the HSig Index analysis, it was determined that the Single linkage algorithm was the 'best-in-class' for this particular dataset, polarizing the clusters most significantly among the linkage algorithms utilized. With twelve clusters as the optimum number of clusters, six remain that are within the significance bounds set forth and are shown below in Table 19. The remaining insignificant clusters are attached in Appendix A.

Table 19. Significant clusters Position data showing compelling evidence to reject the null hypothesis of a 50% standard probability of success.

Cluster Number	6	2	4	5	12	11
p-value	1.13E-12	2.10E-09	1.19E-05	1.19E-05	0.001301	0.01642
Number of trials	50	50	50	50	50	50
Number of successes	48	45	40	40	36	33
Polarity	0.96	0.9	0.8	0.8	0.72	0.66

The first cluster that will be analyzed is Cluster 6. By observing the p-value and polarity in Table 19, one can identify the appeal to this clustering as it very strongly indicates paths and COAs that are optimal. Although the relative similarity height appears extreme in Figure 106 and Figure 107, a look to shows a convergence to one established path. Furthermore, the great similarity height is a result of the dimensionality of the board which includes 25 million positions. The route depicted is one where the blue force avoids the highways, cutting across the foothills and directly attacking the Suisun Valley Wine Coop. Afterwards, the blue force neutralizes the roving patrol utilizing Suisun Valley Rd, then takes a concealed route to attack the road guard in the south.
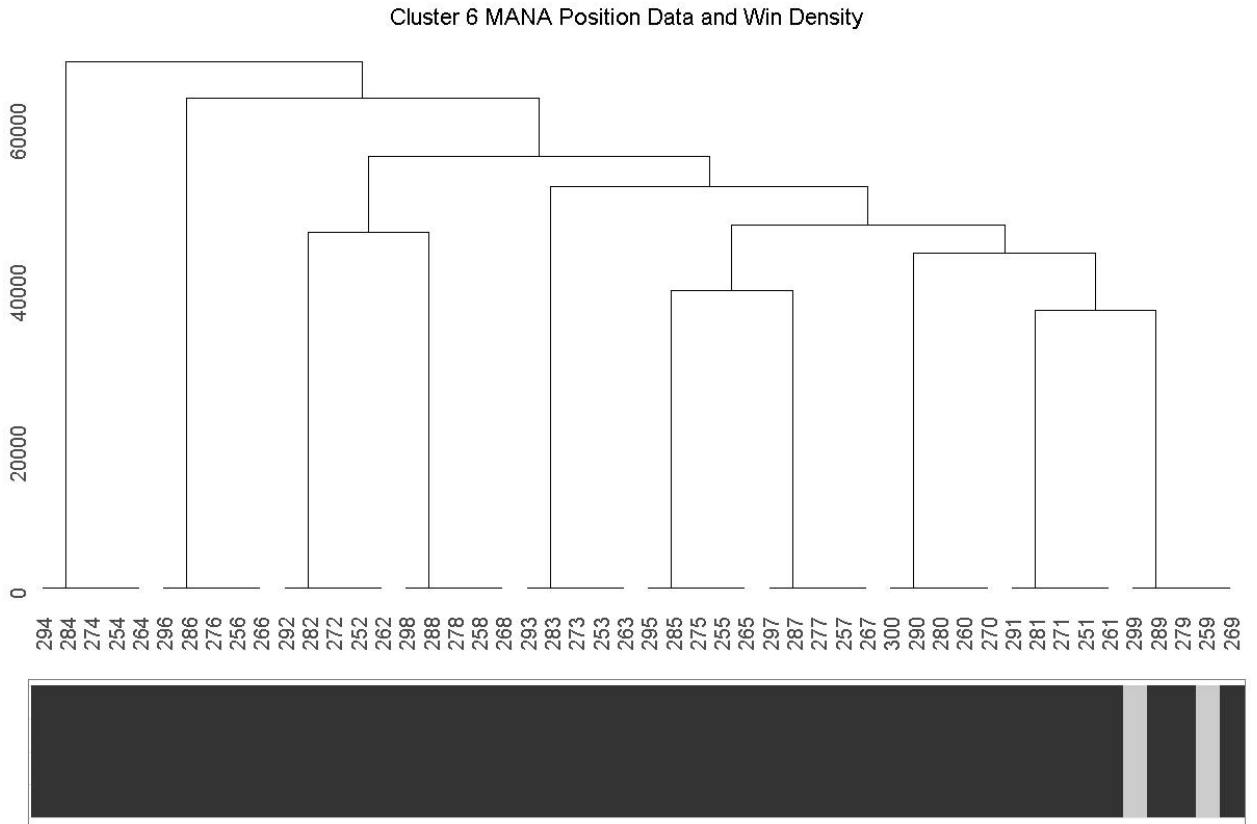
Figure 106. Cluster 6 Position dendrogram and win density.
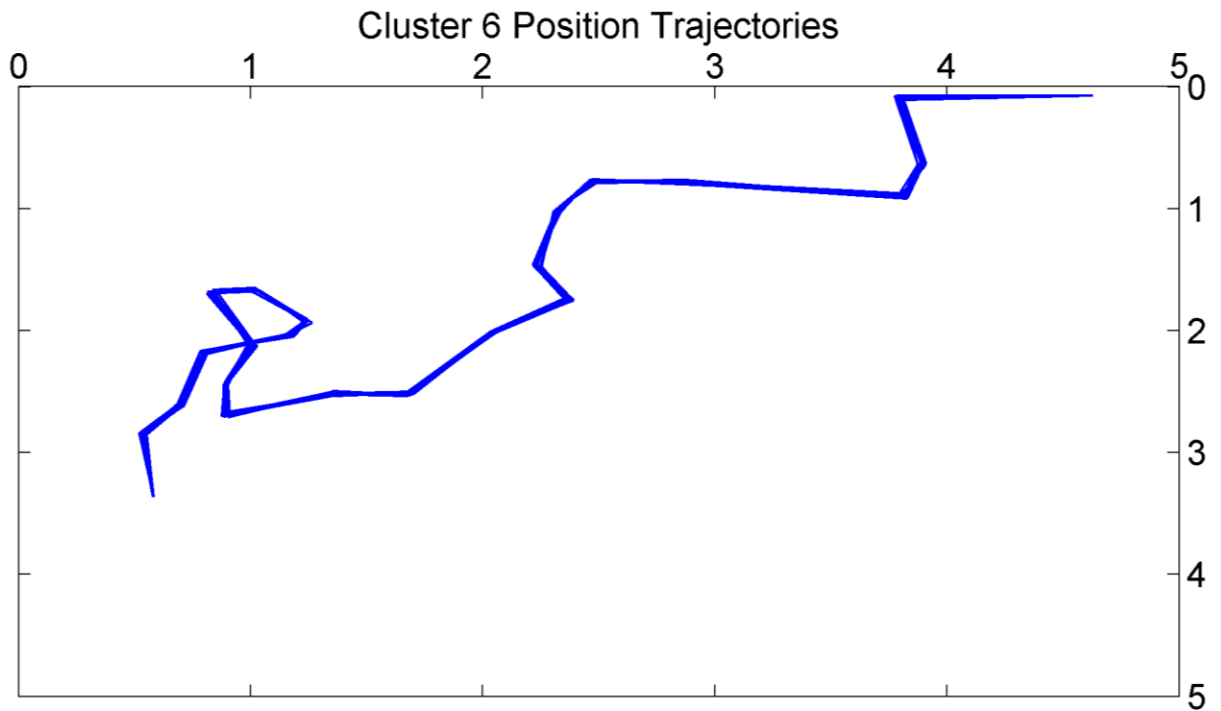
Figure 107. Cluster 6 Position trajectories representing a northern mountain route with southern roving patrol attack followed by headquarters and road guard attack.

The next cluster to be analyzed is Cluster 2, which is almost as noteworthy as the previously discussed cluster, but differs in the direction of the polarity as can be seen in Figure 108. This cluster represents a potential for a catastrophic conclusion where the blue force takes the path of least resistance, the interstate, until passing Suisun Valley Rd, then cutting though the foothills to attack the road guard, then the roving patrol and then finally the Coop. This description can be visualized below in Figure 109. The result of this COA appears to be quite poor, however, the significance testing and HSig Index is not able to

quantify the actual probability of some event occurring, but whether or not witnessing a more extreme event likely. Therefore, while it is appealing to assign a probability of success to this COA, or others presented here, it should not be done unless a particular COA is simulated enough times to create a stable set of statistics.
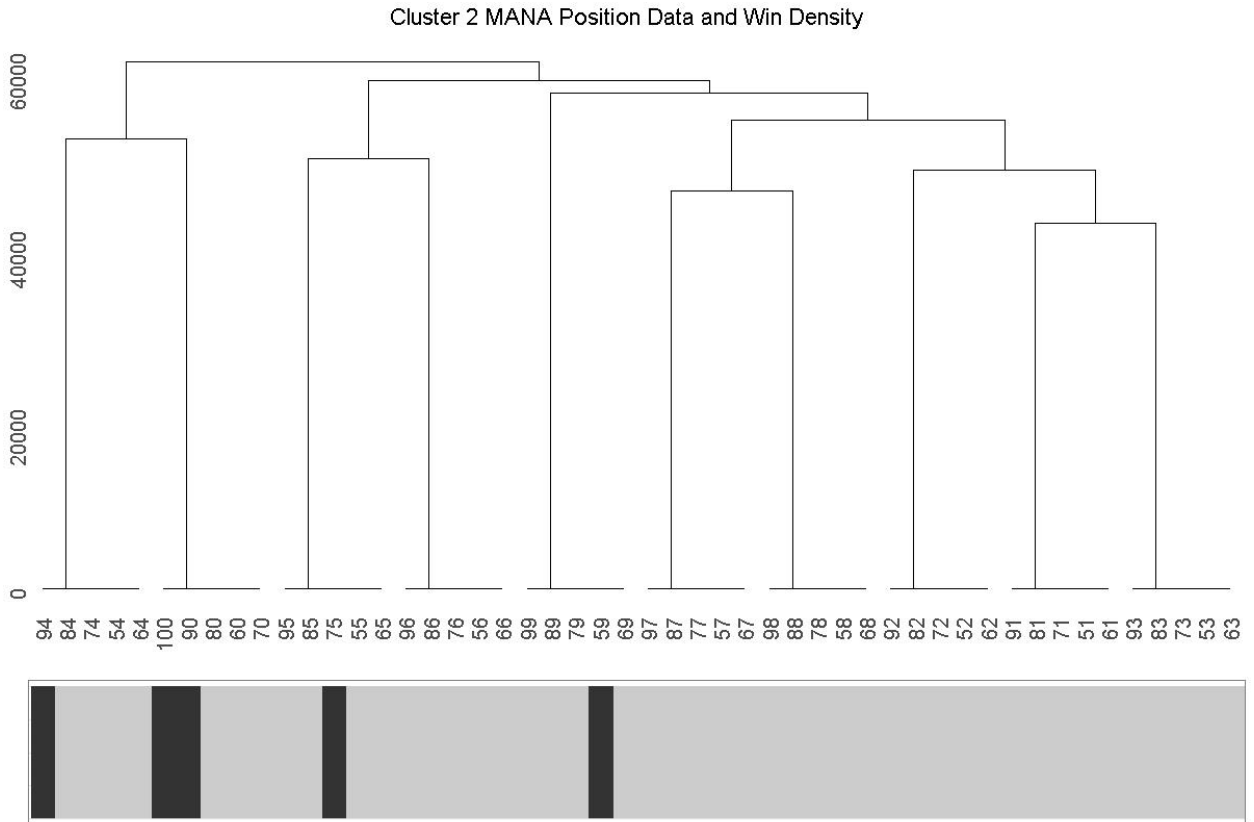
Figure 108. Cluster 2 Position dendrogram and win density.
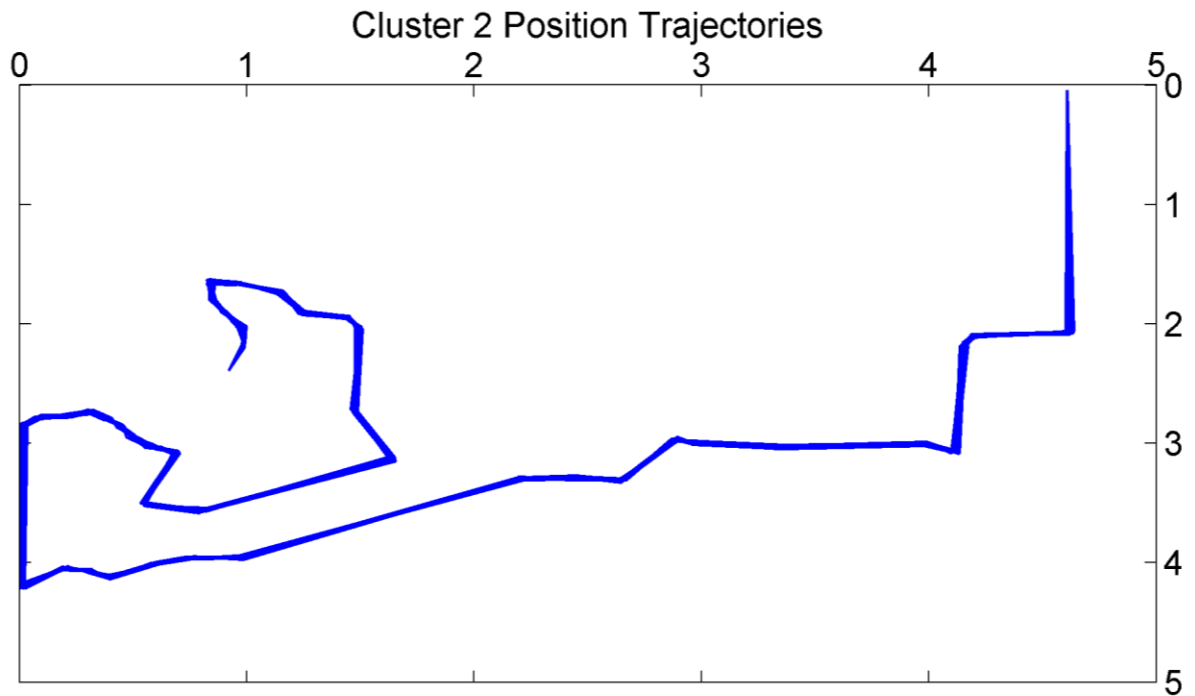
Figure 109. Cluster 2 Position trajectories representing a road-going route giving way to a mountain pass route to road guard attack followed by the roving patrol and headquarters.

The next cluster to be analyzed is Cluster 4 and is shown below in Figure 110. Similar to Cluster 2, this cluster is interesting because of its significance and sub-par probability of success. It is similar to Cluster 4 in that it utilizes the interstate before utilizing the foothills in the Southwest. However, the difference is in the sequencing, where, Cluster 4 utilizes Suisun Valley Rd and engages the road guard to the south immediately before engaging the red headquarters and finally engaging the roving patrol. It is possible that the change in sequencing serves as an improvement over the sequence

seen in Cluster 2, however, more simulation is needed to determine the likelihood that changing sequence would increase the benefit of one COA over another.
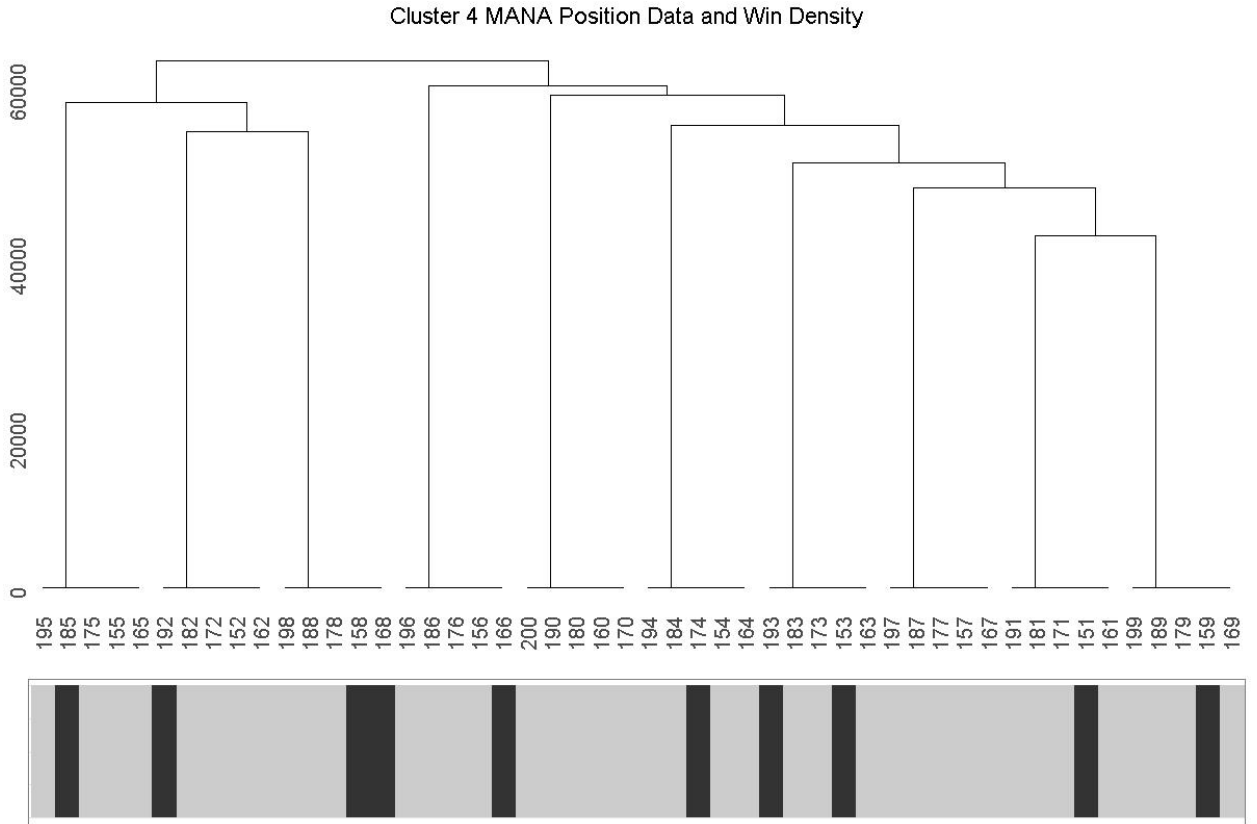
Figure 110. Cluster 4 Position dendrogram and win density.
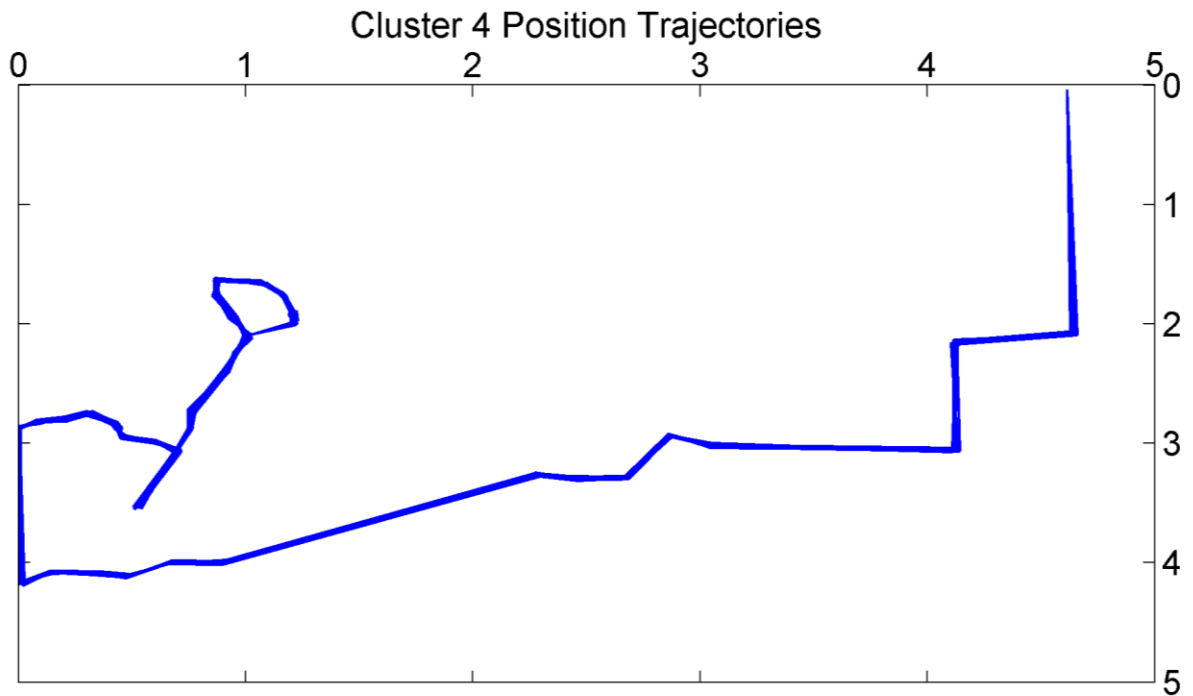
Figure 111. Cluster 4 Position trajectories representing a mountain pass attack on the road guard to the south, then enemy headquarters and finally roving patrol attack.

The next cluster, Cluster 5 exceeds 50% probability of success and is an example of another concealed foothill route which can be seen by observing Figure 112 and Figure 113. However, this cluster is different from Cluster 6 in the sequencing. Cluster 5 engages the roving patrol first before returning to concealed terrain then attacking the red headquarters from the rear and then maintaining concealed travel to flank the road guard in the south. Recall, in Cluster 6, the blue team attacked the headquarters first, taking the road to the roving patrol, then a concealed route to the road guard. While this is a slight variation in route, the sequence is a most obvious departure, and while there are no claims

244

on any change to the probability of success, the change can be seen in the significance level

of the win conditions associated with either of het clusters discussed.
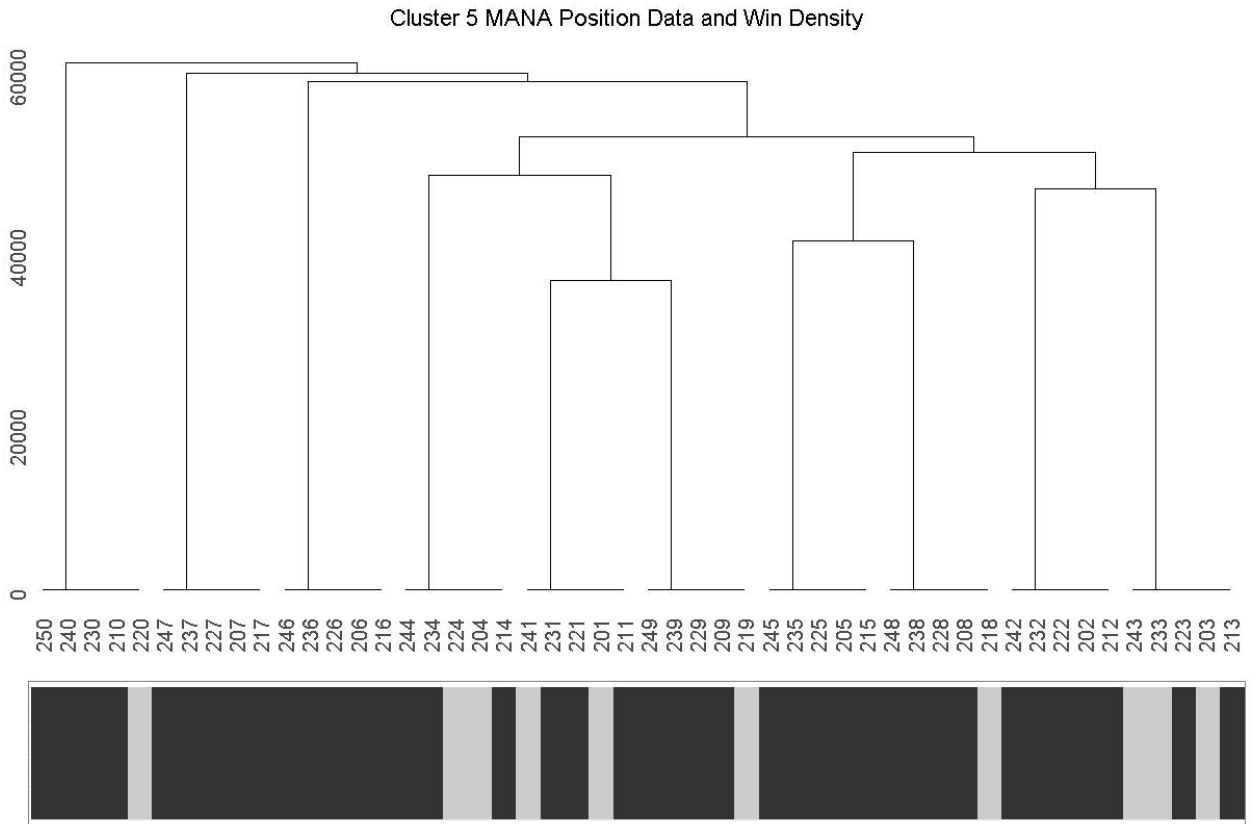
Figure 112. Cluster 5 Position dendrogram and win density.
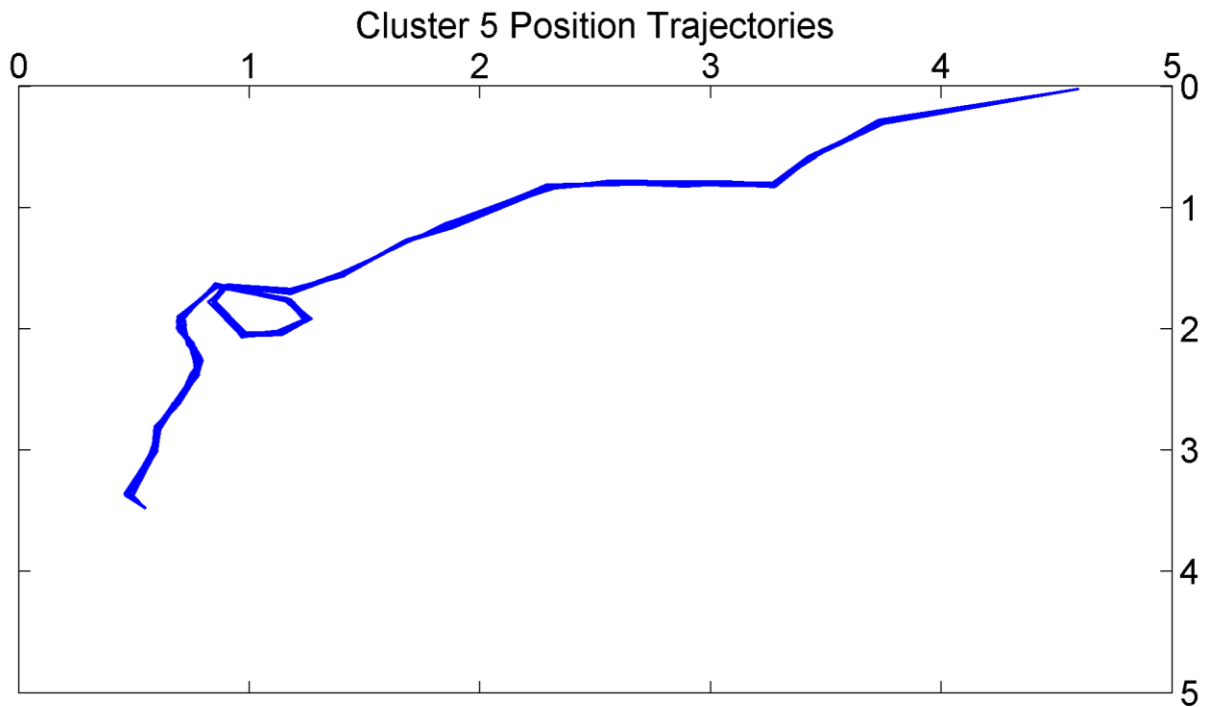
Figure 113. Cluster 5 Position trajectories representing a direct route to the roving patrol, followed by an attack on enemy headquarters and road guard from concealment.

The next cluster to be analyzed is Cluster 12. This cluster represents a policy that has lower than 50% probability of success as can be seen in the win density of Figure 114. The nature of het path can be identified by observing the trajectories in Figure 115. This path represents another variation of a concealed path. The salient feature is after initially engaging the roving patrol, the blue force attacks headquarters and the road guard using an unconcealed Suisun Valley Rd. The result of which is the poorest performing concealed route yet.
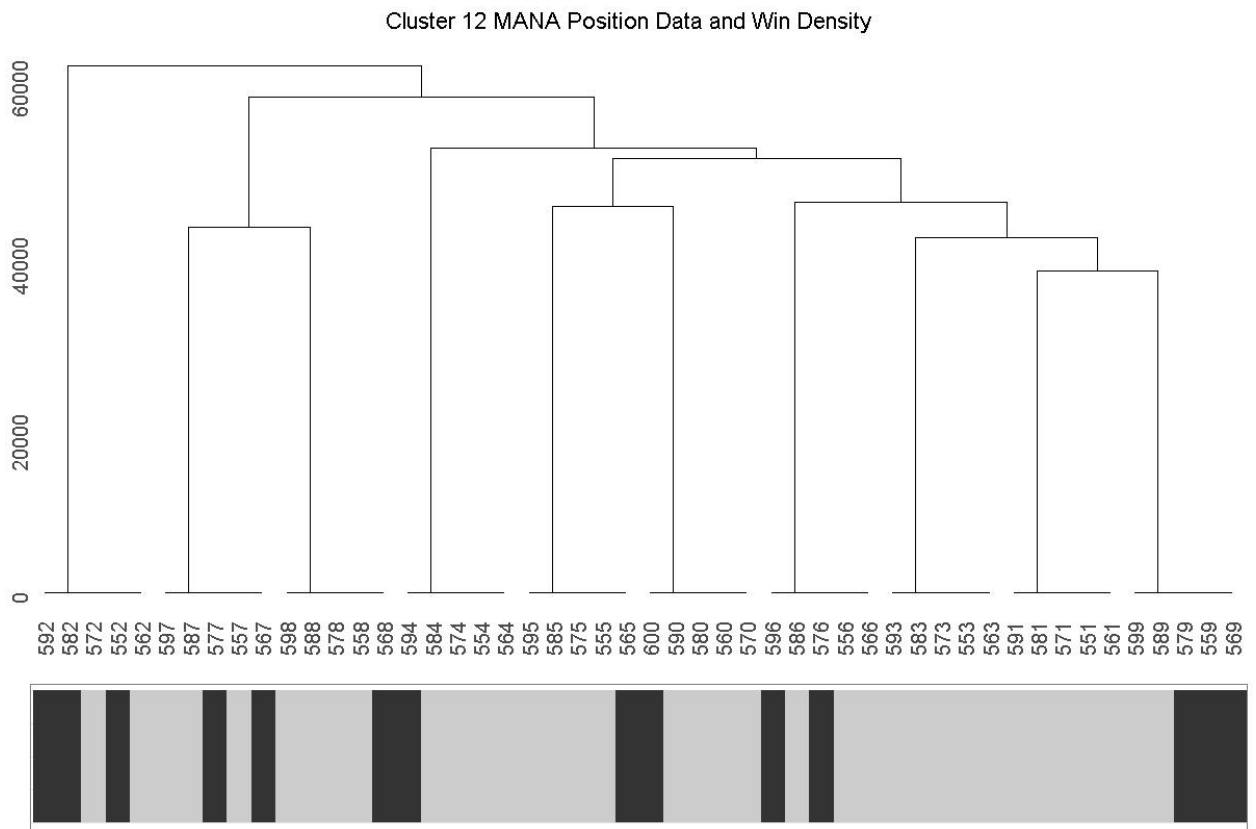
Figure 114. Cluster 12 Position dendrogram and win density.
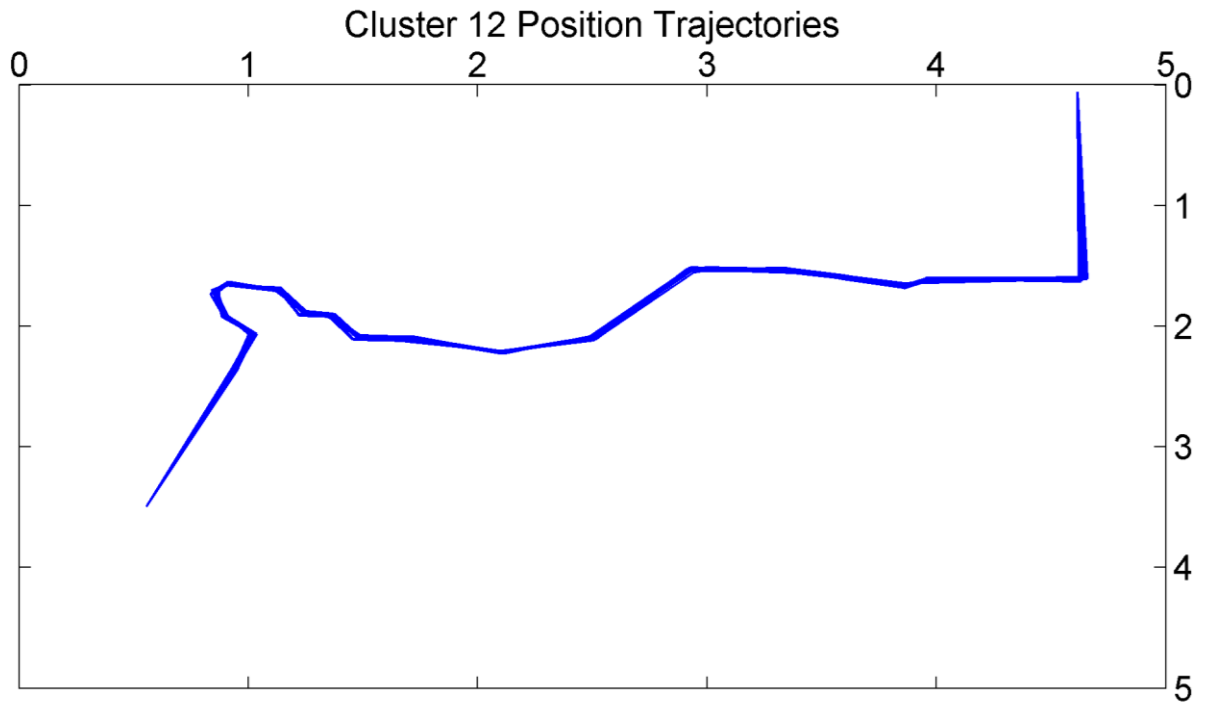
Figure 115. Cluster 12 Position trajectories illustrating a direct route to the roving patrol, then headquarters and finally roving patrol attack from the north unconcealed.

The final cluster to be analyzed is Cluster 11 shown in Figure 116 and Figure 117. This path, exceeding 50% probability of success, is unique among the paths shown because it mixes concealment and easy going by attacking the road guard from the east then traveling north on Suisun Valley Rd to attack the headquarters and finally engaging the roving patrol. It appears that the success of this cluster lies in the attack from the east. The benefit of an attack from this direction must be due to the orientation and line of sight of the road guard, leaving them susceptible from an eastern attack.
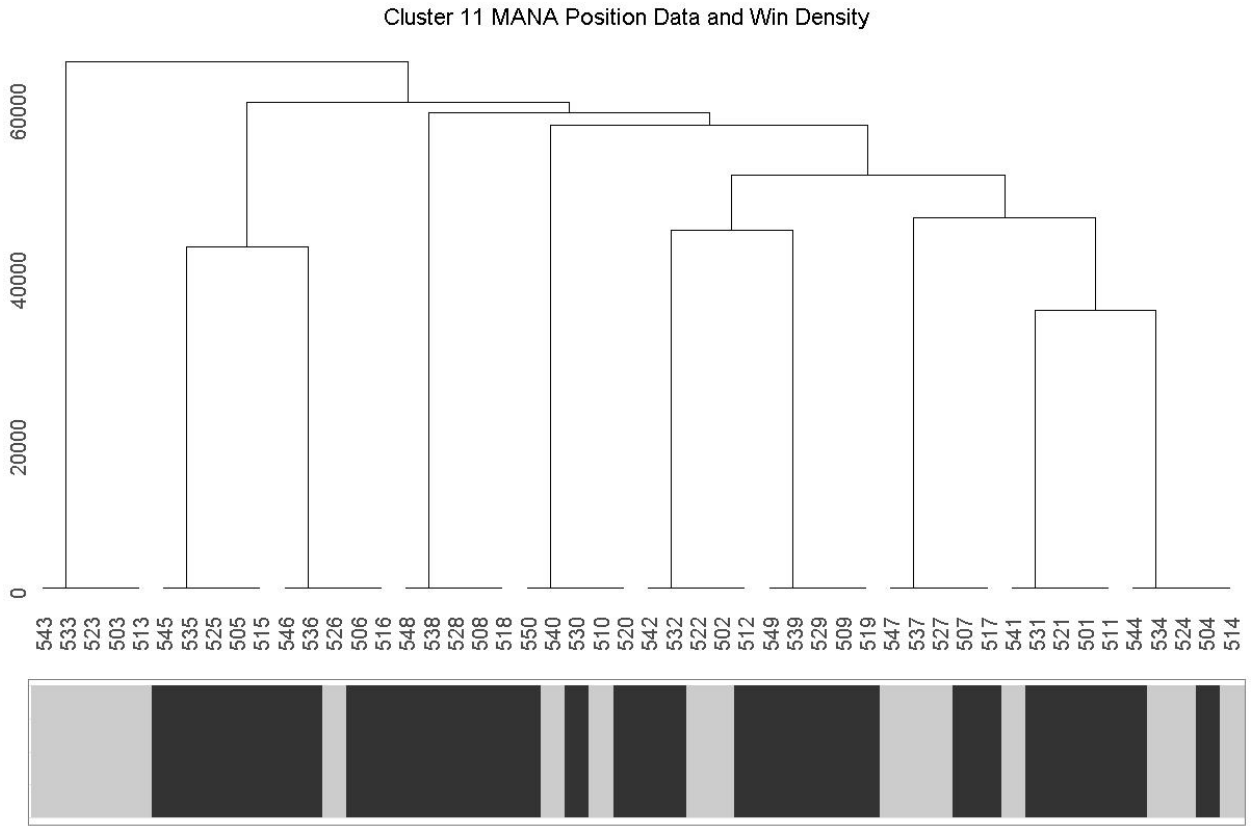
Figure 116. Cluster 11 Position dendrogram and win density.
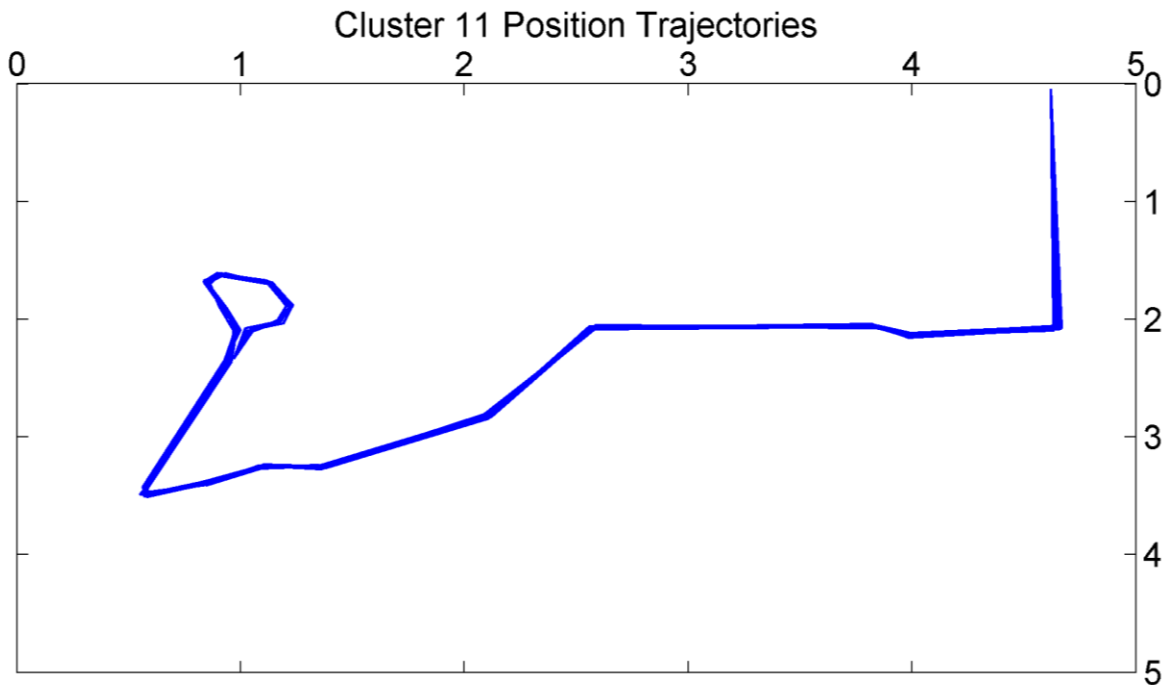
Figure 117. Cluster 11 Position trajectories showing a road going trajectory, mixes concealment and open terrain by attacking the road guard from the east then traveling north on Suisun Valley Rd to attack the headquarters and then the roving patrol.

As a result of the analysis performed, several significant clusters were expanded upon. Due to the size and p-value, breaking up the clusters further through HSig Optimization was not possible because if more similar clusters could be formed by breaking up a 'supercluster', it would be done at the cost of increasing the p-value. Additionally, in determining a most profitable COA, there have been several that exceeded the average probability of success for all cases run. These include Clusters 5, 6 and 11. However, in determining a CDP, only Clusters 6 and 12 are of interest. Both Clusters 6

and 12 are Northern concealed routes with identical attack sequences and can be seen below in Figure 118. However, their attack axis is not identical. The net result is increase or decrease in success probability depending on which COA is chosen. Although both COAs start from the origination point differently, their trajectories intersect. The navy blue trajectory, Cluster 12, opts for a road going trajectory, while the green trajectory, Cluster 6, takes a mountain pass then to surface roads. There is no engagement in this area, therefore the paths preceding the engagement axis do not affect the outcomes. The paths diverge prior to engaging the roving patrol, leading one COA to exhibit increased success probability and a reduction in win probability for the other. From this analysis and by viewing Figure 118, a strong inference can be made. By attacking the roving patrol from the South as opposed to the Eastern attack, a substantial change in the probability of success can be attained. The capability to determine which attack axis for a given attack sequence is advantageous. However, the lack of this presence in the data is telling as well.
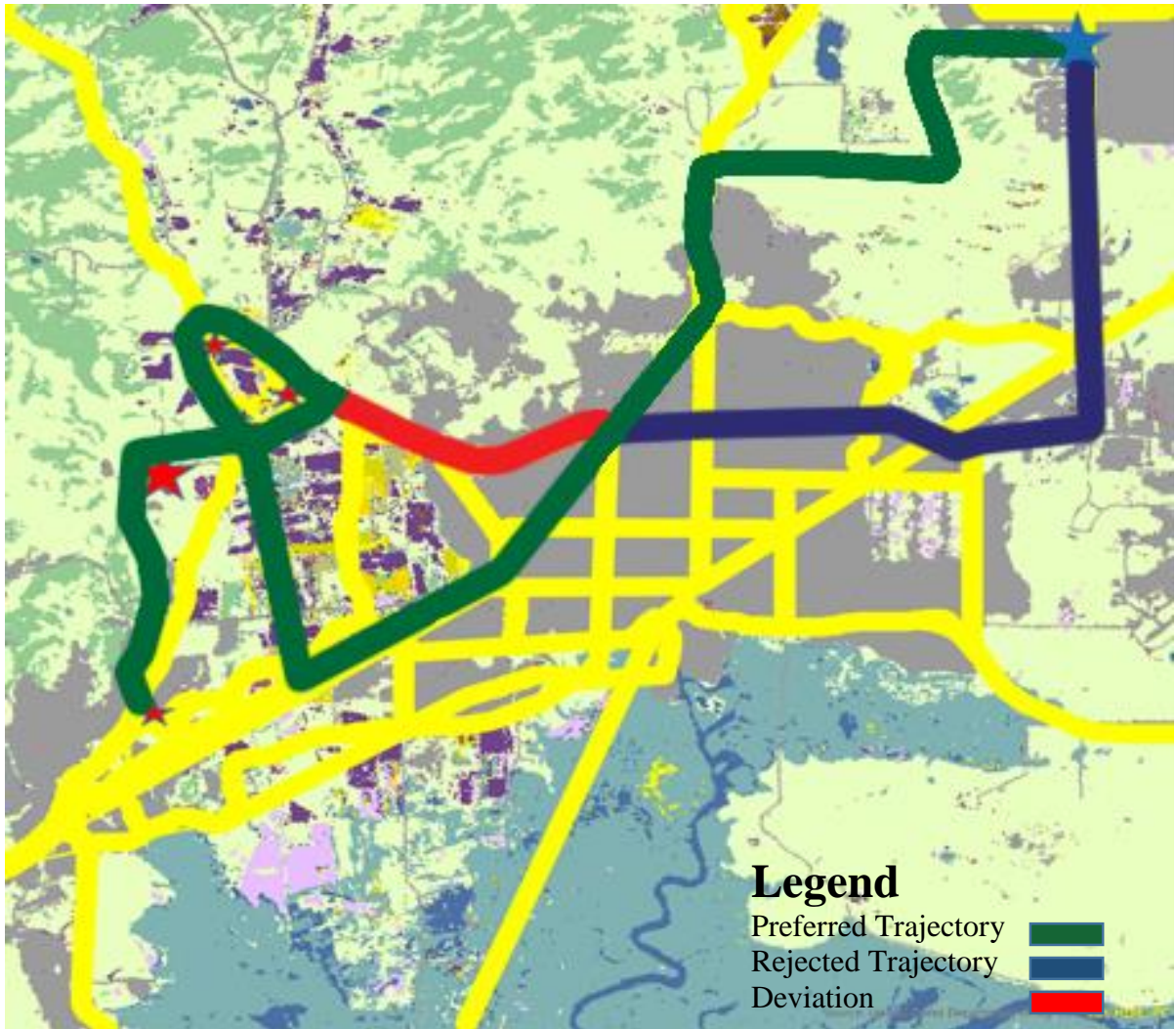
Figure 118. The similar COAs are contrasted showing the point where the COAs diverged yielding a higher probability of success for one and a lower success probability for the other.

Given the two types of data metrics used in this application problem, Position and Capability, the results were stark given the actionability of the COA clusters generated. For example, using Capability, which is a traditional performance measure used in the

planning process, did not show actionable COAs to pursue continued analysis. The utility

ended with showing a few COAs that were not present in the bulk of the Clusters. Analysis

using Capability identified a sort of outlier case that was shed when creating clusters.

However, analysis of the Position metric data using the CADRE Methodology, yielded

several 'good' and 'bad' COAs as well as a CDP. As Position is a non-traditional metric,

the tremendous insight offered by it is especially noteworthy. This illustrates an added

benefit of the CADRE Methodology. Shown here, there was disparity in the effectiveness

of each metric. This was due to the sensitivity of which metrics are of most importance

when analyzing COAs. This is expected to be situation dependent, so the results may

change from problem to problem, where a single or set of metrics are of more importance.

Therefore, in addition to selecting a proper Linkage Algorithm and deriving equitable

COAs and CDPs, this methodology can be used to determine the sensitivity of the metrics

to the simulation and further, the data set. This is also a powerful piece of knowledge to

have when attempting to classify the battlefield. If a commander is able to establish that a

battlefield is highly sensitive to Position and not Capability, less resources surgically

applied will yield much stronger outcomes.

5.4.2 Baseline Application

Although the information obtained from utilizing the methodology is actionable,

the relative performance against traditional analysis must be assessed. To do this, the

performance of all COAs generated will be assessed using two MOPs, Blue Casualties and Red Detections. These MOPs will be utilized to determine if there are any class of COA that perform well on either or both measures. This will be done by analyzing the behaviors and trends seen in the scattering of points representing each COA. Additionally, Pareto analysis on the plot will allow the identification of non-dominated COAs, whose performance cannot be improved without degrading performance in another dimension. By doing this type of analysis, it is expected that there will be a reduction in the number of COAs to analyses, as this is representative of a Multi-Attribute Decision Making paradigm. The best case would be that the best performing COA will be revealed by performing this process.

The first MOP, Blue Casualties, is a relatively simple and obvious choice in measure. It represents the tabulation of the number of blue agents killed during the simulation. If the accomplishment of an objective incurs heavy losses, it can be perceived as risky and also can subject the force to future risk, as the force will not be at full force for the next engagement. Furthermore, the removal of a unit from a battlespace due to incurring heavy losses will be logistically difficult and reduces the overall forces readiness. Therefore, identifying cases that perform well in this regard is of high importance as it is necessary for success.

The next MOP, Red Detections is slightly less obvious. This measure is the tabulation of how many red agents are detected throughout the simulation. The selection of axis and targets will enable or deter the ability of the blue force to sense enemy agents. Therefore, this is essentially a measure of how much information the blue force has about the red force. While it is expected that a high value of this MOP will be advantageous, it can also be a hindrance. This hindrance is primarily if there is a mutual detection and the red agents have some other advantage, forcing the blue agents to incur losses. Therefore, identifying COAs that have high values of Red Detections will be favorable if balanced suitably against the number of Blue Casualties.

The representation of the balance of Blue Casualties and Red Detections is illustrated below in Figure 119. This plot shows the performance of each COA with respect to the MOPs discussed previously. The ideal point in this case, is in the lower right corner of the plot which represents no Blue Casualties, and detecting all red agents. It is ideal because this scenario is not realized. However, there are non-dominated points that exist that compromise between both MOPs.
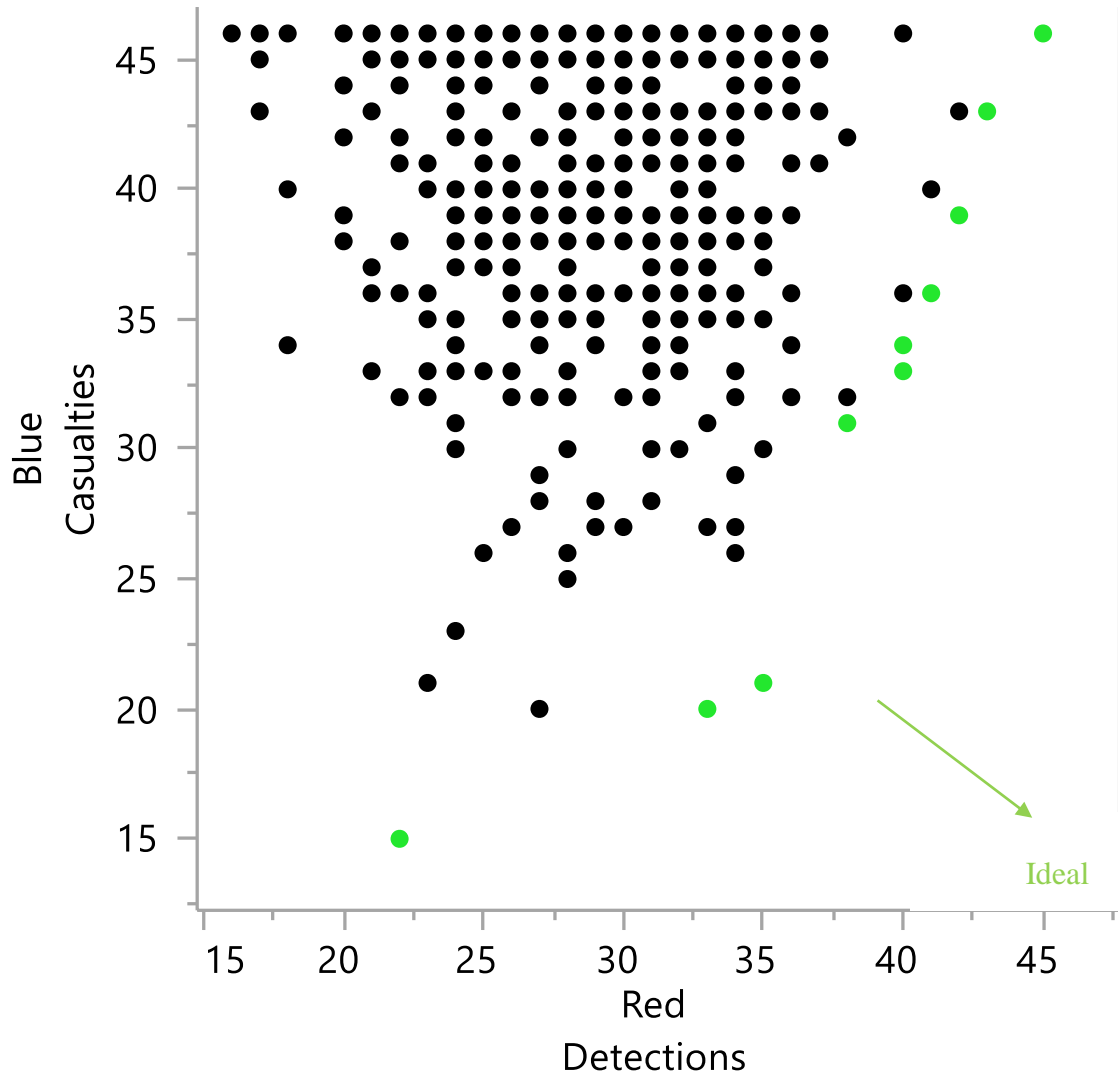
Figure 119. Plot of Blue Casualties vs. Red Detections with the ideal region existing in the bottom right of the plot and the non-dominated points colored green.

Utilizing the cases identified on the Pareto Frontier, the trajectories are plotted and

are shown below in

Figure 120. According to this plot, there is no easily identifiable or obvious COA to pursue. Additionally, the win probability is not known either, therefore Figure 119 is modified showing both wins and losses for each COA representation. Using this plot, only one trend can be seen; COAs that incur heavy losses regardless of enemy detections will be seen as generally losing. One thing to take not of is the appearance of the unequal probability of success in this plot. However, this is not the case as there are identical losses with the identical number of casualties and detections. The other COAs that do not show maximum losses appear to be more disbursed and unique. This appears to show a kind of non-linearity when observing this data set. While all of the COAs are different, COAs that incur heavy losses are collapsed into a tighter space when viewing from the casualty and detection space. COA alternatives that do not incur heavy losses are more disbursed in the casualty and detection space. In any case, the identification of advantageous or disadvantageous COAs is not easily done using this baseline analysis as opposed to the CADRE methodology.
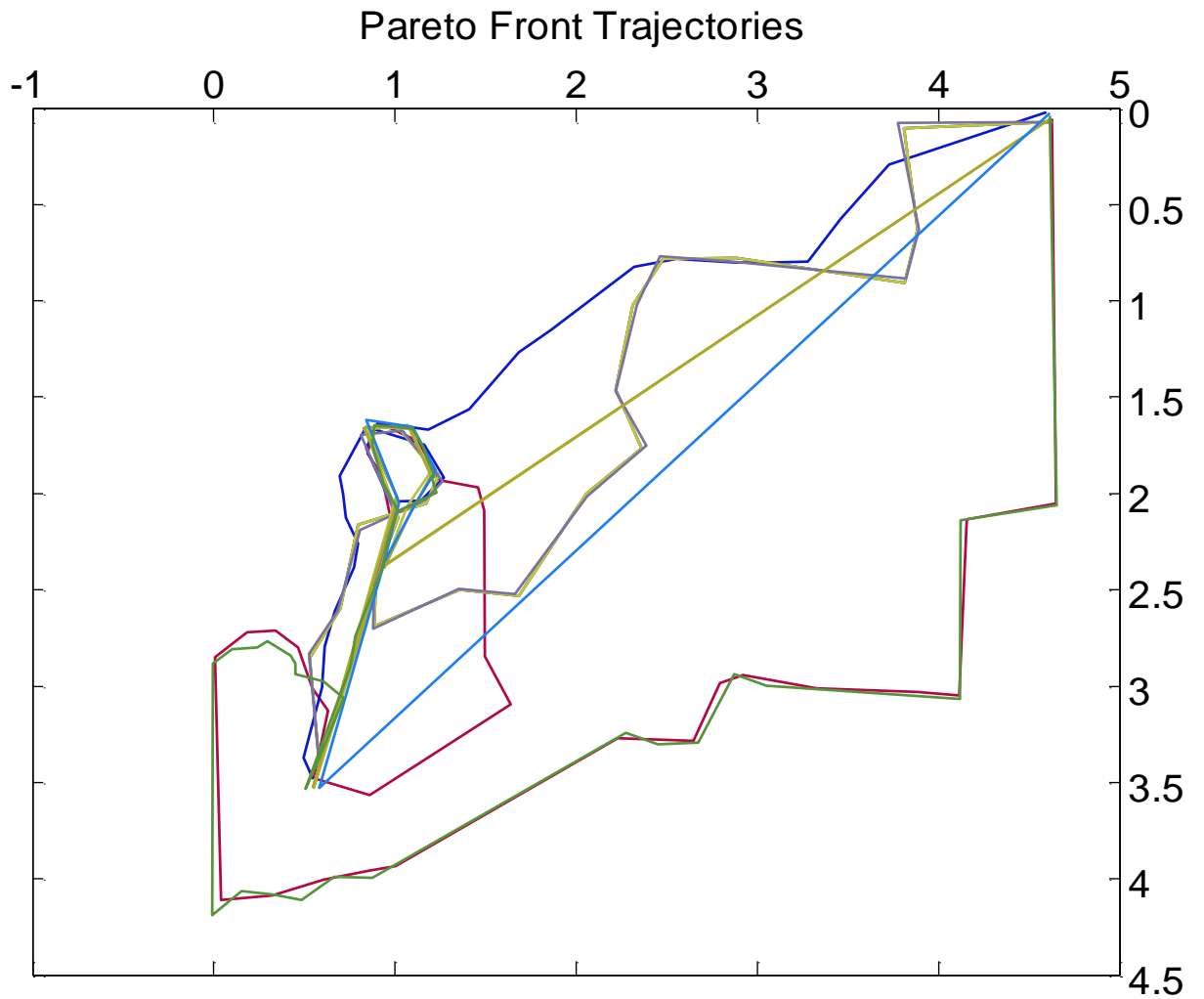
Figure 120. Pareto Frontier trajectories showing a lack of coherence on a specific COA to carry forward in utilizing Pareto Analysis.
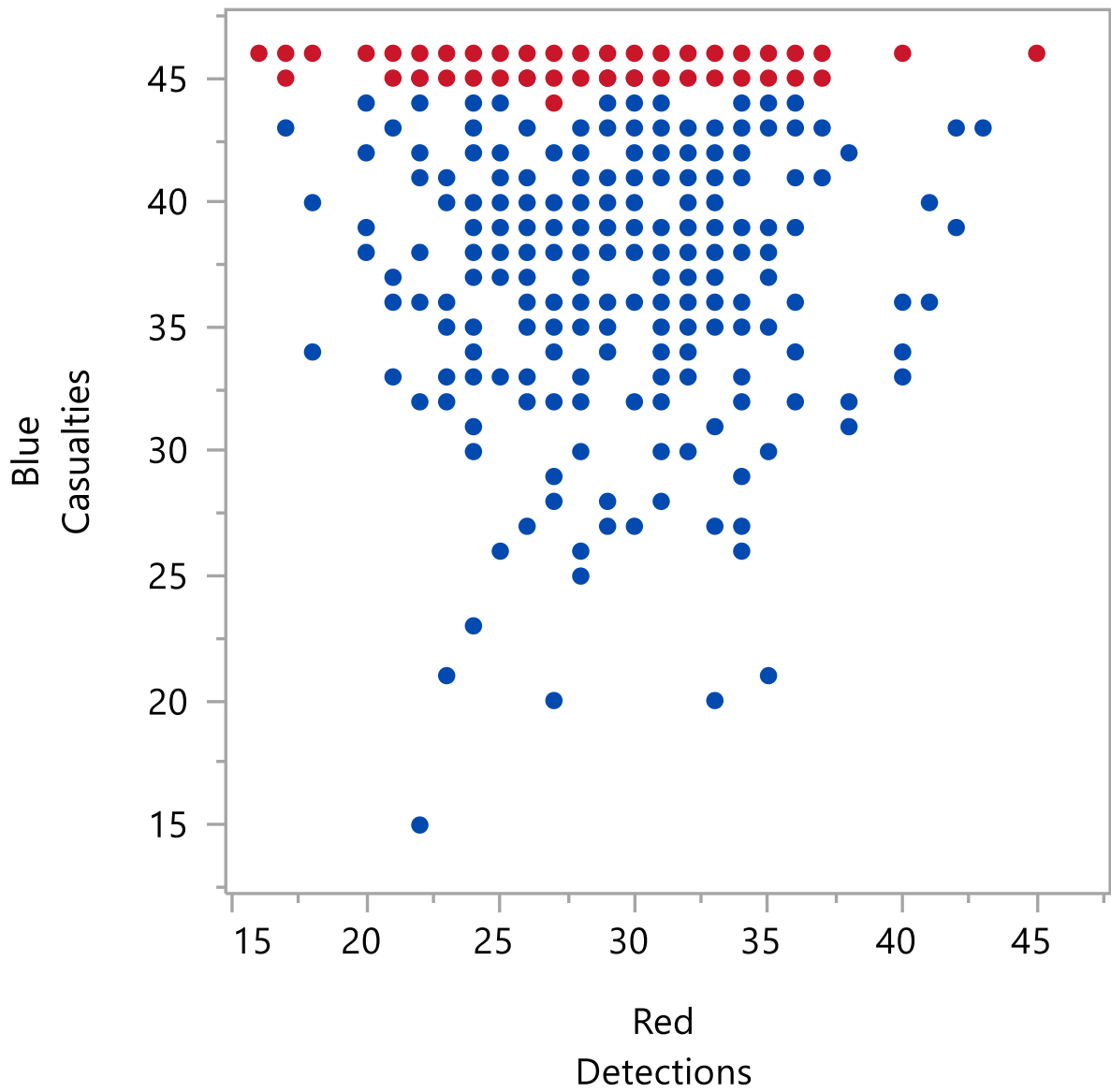
Figure 121. Plot showing wins and losses when considering the performance of each COA with respect to friendly casualties and enemy detections.

By using the same plots in the baseline analysis and adapting the information gained from the CADRE methodology, additional insights can be seen. Below in Figure 122 by

259

utilizing the clustering assignments of COAs, we can identify the CDP illustrated in the previous section. Recall that Cluster 6 was identified as advantageous and Cluster 12 was not advantageous. This can be seen below in Figure 122, where a majority of high casualty COAs are Cluster 12 and a majority of Pareto frontier points are Cluster 6. While this is an interesting observation and reinforces the information found in the methodology application, it is not possible with traditional Pareto analysis. However, this perspective of the methodology can be utilized in future representations. The ability to use cluster specific information on traditional scatter plots of MOPs enable the CADRE information to be viewed from multiple dimensions.
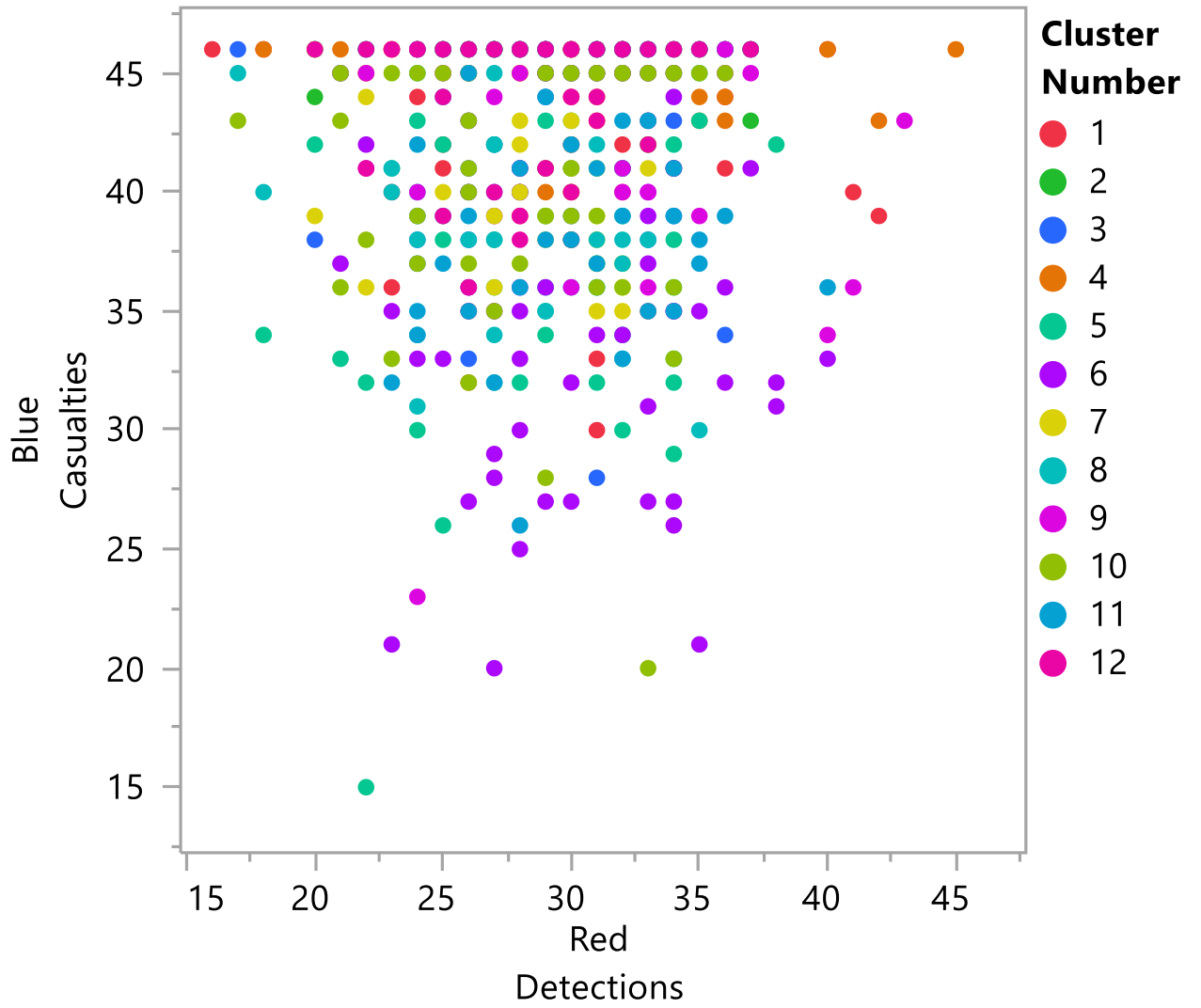
Figure 122. Scatter plot of all COAs using the cluster information from CADRE showing another dimension to the selection of the CDP identified in the methodology.

CHAPTER VI

CONCLUSIONS

COA wargaming and analysis limit the ability to rapidly alter plans in the face of today's uncertain and volatile battlespace. Presented here is a method to enable timely action through enhanced COA analyses and wargaming. The result was a process yields a drastically reduced decision space for analysis. As was shown in the application problem, there was on order of magnitude reduction between the initial sampled decision space and the final selection of clusters set for further analysis. However, the process yielded additional information on the generation and analysis of the alternative COAs presented here.

The necessary generation of alternative COAs was considered in the methodology. In using computer chess as a wargaming environment, generating alternative COAs is done by playing out many games by randomly choosing between feasible moves. The result of which is a COA consisting of the sequence of moves that make up the playout. Scaling to a full size board does not affect this step in the methodology. However, when scaling to an agent based wargame environment, additional complexity in the behavior of the agents is expected, however, the same process of choosing component elements of a COA randomly from a list of feasible moves is beneficial.

262

The next aspect of the methodology aims to facilitate the analysis of the wargame data for COA analysis. Due to the constraints on the form of the data, Hierarchical Agglomerative Clustering using Dynamic Time Warping was chosen as the method of mining attributes from the data. Additionally, it was determined that there is not a way to effectively select an aligning linkage algorithm without some experimentation. Therefore, as a contribution of this thesis, a set of algorithms were carried forward and analyzed by creating and using a novel validation index using the similarity of elements of a cluster matched with the significance of the outcome of the elements of a cluster. It was found that the HSig index exhibits a quadratic relationship with the selection of the clustering parameter, the number of identified clusters. Therefore, experimentation was proposed to establish the severity of this quadratic relationship which will inform the need to optimize these parameters to facilitate linkage algorithm selection. Furthermore, resulting from the experimentation, the algorithms were observed to exhibit varying levels of statistical significance by analyzing a spectrum of static cuts on the Dendrogram. Furthermore, these varying levels did not exhibit stark similarities between linkage algorithms, reinforcing the adequacy of the performance of the HSig Index and the overall optimum selection of the UPGMA linkage algorithm for the chess dataset and the Median and Single linkage algorithms selected for the combat simulation Capability and Position metrics respectively. Upon final selection of the linkage algorithm, decision point analysis is proposed as a final step to the methodology.

The chess games of ten clusters and the clusters themselves were analyzed in Experiment two. Analysis of random chess games proved to be quite difficult due to the vast decision space and the branching factor for each decision of the game. However, in most of the clusters, a distinct preference for the center of the chess board was seen reinforcing a known chess heuristic. Additionally, investigations in the timing and tempo of attacks were investigated but did not show any salient features as did the center board preference. Most importantly, Experiment 2 gave an indication of behavior and the dynamics of the HSig index and Optimization process.

In this experiment, it was determined that while it is possible to get clusters from HSig Optimization that favor similarity or significance, the analysis will mostly favor a balanced preference for both. In one case, if a cluster is carried forward that is significant but not similar, there can be no actionable knowledge extracted from the cluster. On the other hand, if a cluster is similar but not significant, there should be no actionable knowledge extracted because there lacks enough evidence to do so. The resolution in either case is to obtain more data. This phenomenon motivated the next experiment.

It is postulated that the complexity of interactions and the complexity of agent types present in a wargame environment will affect the ability to recognize decision points due to combinatorial nature of the problems selected. This, to an extent was shown in Experiment 2. Experiment 3 was meant to assess the direct impact of this complexity by analyzing reduced data sets. The complexity of behaviors and actions on a chess board

264

and a battlefield map to modeling and simulation through the vast number of alternatives present. Therefore, fractions of the chess data set were carried forward for analysis to simulate the increase in complexity. The outcome was a substantial decrease in quality clusters obtained. In the 400 case example, there were no usable clusters that could inform decision-making. In the 800 case example, four clusters were selected but after analysis, only one cluster was determined to be sufficient. As a result, it is determined that this methodology is data-philic, an attribute that scales with the decision and simulation space.

With the relevant experimentation compete in Experiments 1, 2, and 3, the methodology was utilized in an agent based model of a ground combat scenario between blue and red agents from the blue agent perspective. Data obtained from simulating the battlefield was utilized to influence the action of the agents as they attempt to meet objectives. Additionally, COA elements were developed as variations of routes taken and sequencing of attacks on red targets. There were a total of 600 COAs simulated, collecting Capability and Position data for each simulation. The CADRE methodology was performed on each metric and identification of COAs and CDPs were attempted.

The performance of the Capability metric was not particularly useful for COA determination. There appeared to be an identified region with increased casualty rates, but there were no indications of a preferred COA. This leads to the identification of an additional contribution of this work, which is to determine the relative sensitivity of one metric over another in the face of a particular data set and simulation. Given the results

form analyzing the Capability metric, it is determined that the application problem is not sensitive to the changes in Capability. The Position metric, however, was fruitful. To a high degree of significance, advantageous COAs were identified, as were blunderous COAs. A critical sequencing step in two similar clusters was identified that impacted the performance greatly. Therefore, this sequencing step, where the two COAs bifurcated represents the CDP for this simulation.

CONTRIBUTIONS

The primary contribution of this thesis is offering of enhanced COA analyses of wargame data allowing the identification of decision points and heuristics. Sequencing of decisions and actions on the battlefield were shown to have a dramatic effect on the outcomes of combat simulations. Providing the capability to recognize patterns in successful COA simulations to identify successful COAs, and more importantly, providing the capability to identify the time dependent set of actions that allowed them to be successful is a substantial contribution in the field of military operations research modeling, simulation, and analysis. The method described in this work that allows these capabilities is the CADRE methodology.

The methodology presented process highly dimensional data as metric time series, which are converted highly dimensional similarity space. This prevented the use of typical clustering algorithm parameters and validation indices based on 2-dimensional data representations. Additionally, a characteristic specific to military operations is a sharp focus on the outcomes of military engagements. Therefore, highly dimensional data was appended to yet another dimension further intensifying the intricacy. As a result, another contribution of this thesis was the creation of a validation index leveraging the data specific to this problem, and most importantly, the outcome associated with a set of decisions. Through formulation and experimentation, this work has resulted in the creation of a cluster validation index that leverages inner cluster similarity and the statistical significance of the

267

outcome of scenarios. This index is called the Height Significance Index (HSig Index) and can be utilized to select the clustering parameters needed for analysis: Linkage algorithm and number of clusters, resolution, for which to view the data.

With the HSig Index, algorithms can be applied to data yielding the most statistically significant groupings that yield a preferable outcome. However, due to the non-linear nature of the decision space, certain regions of the decision space may be more densely populated than other regions. Correspondingly, these regions of differing density must be analyzed at different resolution levels. The next contribution of this thesis is a process to analyze the data at differing resolutions, while still enabling similar COAs yielding desirable outcomes to be identified. This process is called the HSig Optimization process

An additional contribution in line with general concerns regarding the form and complexity of the data was motivated by the experimentation presented in the previous section. It was found that the quality of the output of this process, COAs and CDPs, was dependent on the factors surrounding the data retrieval process. Specifically, the complexity of the underlying process and the model used can affect the number of interactions on the battlefield, expanding the decision space. In addition, the number of actors and actions required to achieve a goal aids in the combinatorial nature of the problem, further expanding the decision space. Given these phenomena, the methodology and specifically, the HSig Index and HSig Optimization process presented a way to

quantitatively assess the effect of mission complexity and the corresponding investment into data retrieval, on the quality of decisions extracted using the CADRE methodology. Furthermore, this process can be utilized within other decision making methodologies utilizing similar data types to assess the effects of scaling.

The output of the CADRE methodology is a set of COAs and CDPs to subject to further analysis. The quality of these COAs and CDPs allow insights into the process as well. In the application problem it was seen that there was a disparity in the quality of the decisions perpetuated by the process using two different metrics: Position and Capability. Capability is a traditional performance measure, and is typically used in military planning. However, this measure did not yield notable results but did allow some reduction of the decision space. Performance, conversely, allowed a prominent set of COAs and CDPs to carry forward, exemplifying this disparity associated with the metrics selected. Therefore, as an additional contribution, this methodology enables the analysis of the sensitivity of performance metrics to the battlespace and associated decision space. Using this provision, planners may classify the engagement scenario based on the most sensitive metrics and plan accordingly.

An unintended contribution of enhancement to Pareto analysis were also identified through analysis of a baseline case. In the baseline case, the data was analyzed according to two MOPs: Blue Casualties, and Red Detections. In this type of analysis, it was shown that it was not possible to yield the insight resulting from the CADRE methodology.

269

Additionally, when appending the outcomes of the clustering process in CADRE, to the scatter plot of the MOPs, additional insights were identified in the data set that matched the CDP found in the example problem. Furthermore, this insight would not have been identified through the baseline case alone, and there would be no other way to obtain the cluster information from the baseline case as well. The ability to use cluster specific information on traditional scatter plots of MOPs enable the CADRE information to be viewed from multiple dimensions. In the case of using several MOPs in a scatterplot matrix, CADRE information can be viewed in each MOP comparison scatter and the relative behaviors for each cluster can be identified for each combination of MOPs. Thus, the CADRE methodology can be utilized to further enhance traditional decision-making paradigms.

FUTURE WORK

The primary output of the CADRE Methodology is a COA or set of COA clusters that can be subjected to further analysis, alleviating some of the burden of COA Analysis and Wargaming from military planners. This was successful. Additionally, these COAs set for further analysis were selected on the probability that they increase the probability of success. In one case, the success probability is increased by pursuing advantageous COAs while it is still increased by avoiding disadvantageous COAs. Given this capability, future work is required to add a means to accurately assess the actual within cluster probability of success.

This future work is needed to enhance the process with respect to the confidence in which outcomes can be predicted. While there the size of actionable COA clusters may be large enough to enable a frequentist evaluation of the probability of success, caution should be taken. A truly rigorous evaluation of success probability should be done by identifying the salient features in a COA cluster and resampling in that region of the space to verify the actual probability of success. This ensures that the decision space surrounding the cluster elements are present in the sampling, and offers more evidence in support of rejecting the null hypothesis. Modifying the simulation will affect this process.

271

If there is significant randomness in the simulation where the outcomes using the same inputs are sufficiently similar, the requirement to test COA in the vicinity of the identified cluster and not just repeating the simulating of alternatives is even more important. In this case, differing results can be had by re-running simulation, which is why Monte Carlo methods are required. However, in order to determine the actual probability of success, sampling in the vicinity of the cluster must be performed in addition to performing a Monte Carlo simulation on the alternatives.

Additionally, in performing the baseline application, a parallel was found using an adaptation of the traditional analysis reinforced with cluster information. Future work in this direction would enable the analysis of clusters in different dimensions of MOEs and MOPs without having to repeat the analysis using many evaluated metrics. If a subset of metrics can be used to get trustworthy cluster data, these cluster assignments can help to enhance traditional analyses, giving another dimension to the analysis. Therefore, in one plot, an analyst can identify performance trends in two typical dimensions of MOP vs. MOP as well as in a third dimension of what cluster each COA belongs to. This will enable more trends to be identified and allow the performance of each cluster as applied to many clusters to be assessed.

APPENDIX A: MANA SIMULAITON SCENARIO

This appendix shows the HTML file used to generate the MANA scenario. In order to utilize this, the user must have installed MANA. Additionally, in the folder where MANA is installed, the user must also have the terrain and elevation maps. If the maps are unavailable, the simulation can still be performed with the agent's behavior and locations. However, there will be no influence on the agents by the environment.

If decision analysis is desired, the coordinates for the waypoints below must be replaced with the intended routes and target. This is needed to show the influence sequencing and timing have on the outcomes.

```
<?xml version="1.0" encoding="UTF-8"?>

<specification>
<type>Mana Scenario File</type>
<version> 5.01.09 </version>

<description>
.. The country of Napa has invaded the neighboring country of Suisun and has taken the capitol of Suisun, the Suisun
Valley Wine Coop.  Napa is currently occupying the region of Suisun with a substantial presence in three locations
preventing unencumbered access to the wineries of Suisun.  The first location is located in the lower left portion of the
battle field and represents a road guard, protecting the enemy positon from southern attacks, and by preventing access
to the Coop via easy going terrain.  The next is the red team headquarters represented by a set of concentrated red
waypoints where a red force loops.  The final red location is a roving patrol looping around dispersed waypoints to the
north, which prevents access to the Coop from the East and North via easy going terrain.  A neighboring ally of Suisun,
Sacramento has staged a forward operating base North West of Suisun, represented by a single blue waypoint.
.
. </description>

<multirun>
  <seed> 0 </seed>
  <side> 1 </side>
  <BlueLimit> 48 </BlueLimit>
  <RedLimit> 5 </RedLimit>
```

273

```
    <NeutralLimit> 5 </NeutralLimit>
    <SquadReachedGoal> No </SquadReachedGoal>
    <SquadNumToReachGoal> 0 </SquadNumToReachGoal>
    <StopSquadDead> No </StopSquadDead>
    <SquadLimit> 1 </SquadLimit>
    <SquadNumToSufferLoss> 0 </SquadNumToSufferLoss>
    <MultiResultsName> multiruntest.csv </MultiResultsName>
    <maxruns> 2 </maxruns>
</multirun>

<Battlefield>
    <x_range> 5000 </x_range>
    <y_range> 5000 </y_range>
    <LowRealY> 0.0 </LowRealY>
    <HighRealY> 5000.0 </HighRealY>
    <LowRealX> 0.0 </LowRealX>
    <HighRealX> 5000.0 </HighRealX>
    <GlobalX_Min> 0.0 </GlobalX_Min>
    <GlobalX_Max> 5000.0 </GlobalX_Max>
    <GlobalY_Min> 0.0 </GlobalY_Min>
    <GlobalY_Max> 5000.0 </GlobalY_Max>
    <LowRealElev> 0 </LowRealElev>
    <HighRealElev> 255 </HighRealElev>
    <TimeScale> 1.0 </TimeScale>
    <LOSresolution> 1 </LOSresolution>
    <DistUnits> distMetres </DistUnits>
    <CombineLOSCalc> No </CombineLOSCalc>
    <DisableMsgLog> No </DisableMsgLog>
    <SAMode> Map Location </SAMode>
    <SAMgNewCt> Contact ID </SAMgNewCt>
    <LOSMode> Advanced </LOSMode>
    <UseNonStdTerrains> Yes </UseNonStdTerrains>
    <ContactRad> 6.01 </ContactRad>
    <TerrainRange> 5 </TerrainRange>
    <NumSquads> 4 </NumSquads>
    <image type="map" name="Suisun3Cropland+Roads.bmp" />
    <image type="background" name="Suisun3Background.bmp" />
    <image type="elevations" name="Suisun3Elevation.bmp" />
    <TimeStep> 0 </TimeStep>
    <MaxSteps> 8000 </MaxSteps>
    <settings>
        <ShowPath> Yes </ShowPath>
        <ShowFlags> Yes </ShowFlags>
        <CalcLOS> Yes </CalcLOS>
        <Networking> Yes </Networking>
        <IntraNet> Yes </IntraNet>
        <RecPositions> No </RecPositions>
        <RecFirstEnDetections> No </RecFirstEnDetections>
        <RecRedDetections> Yes </RecRedDetections>
        <RecComms> No </RecComms>
        <RecMDetects> No </RecMDetects>
        <RecSteps> No </RecSteps>
        <RecCasLocs> Yes </RecCasLocs>
        <SoundFx> Yes </SoundFx>
        <LockSeed> No </LockSeed>
```

```
  <StopBlueDead> Yes </StopBlueDead>
  <StopRedDead> No </StopRedDead>
  <ShowBackground> No </ShowBackground>
  <ShowAnimation> Yes </ShowAnimation>
  <ShowInteractions> Yes </ShowInteractions>
  <ShowFiring> Yes </ShowFiring>
  <ShowAgtOrientation> No </ShowAgtOrientation>
  <DoPathFinding> No </DoPathFinding>
  <RecAgentEndState> No </RecAgentEndState>
  <ShowDeadIcons> Yes </ShowDeadIcons>
  <RecDetects> No </RecDetects>
</settings>
<TerrainType>
  <TerrainName> BilliardTable </TerrainName>
  <Going> 1.00 </Going>
  <Cover> 0.00 </Cover>
  <Concealment> 0.00 </Concealment>
  <Red> 0 </Red>
  <Green> 0 </Green>
  <Blue> 0 </Blue>
</TerrainType>
<TerrainType>
  <TerrainName> Wall </TerrainName>
  <Going> 0.00 </Going>
  <Cover> 1.00 </Cover>
  <Concealment> 1.00 </Concealment>
  <Red> 192 </Red>
  <Green> 192 </Green>
  <Blue> 192 </Blue>
</TerrainType>
<TerrainType>
  <TerrainName> Hilltop </TerrainName>
  <Going> 0.90 </Going>
  <Cover> 0.10 </Cover>
  <Concealment> 0.95 </Concealment>
  <Red> 64 </Red>
  <Green> 64 </Green>
  <Blue> 64 </Blue>
</TerrainType>
<TerrainType>
  <TerrainName> Road </TerrainName>
  <Going> 1.00 </Going>
  <Cover> 0.00 </Cover>
  <Concealment> 0.00 </Concealment>
  <Red> 255 </Red>
  <Green> 255 </Green>
  <Blue> 0 </Blue>
</TerrainType>
<TerrainType>
  <TerrainName> LightBush </TerrainName>
  <Going> 0.75 </Going>
  <Cover> 0.10 </Cover>
  <Concealment> 0.30 </Concealment>
  <Red> 232 </Red>
  <Green> 255 </Green>
```

275

```
      <Blue> 191 </Blue>
    </TerrainType>
    <TerrainType>
      <TerrainName> DenseBush </TerrainName>
      <Going> 0.20 </Going>
      <Cover> 0.30 </Cover>
      <Concealment> 0.90 </Concealment>
      <Red> 147 </Red>
      <Green> 204 </Green>
      <Blue> 147 </Blue>
    </TerrainType>
    <TerrainType>
      <TerrainName> Vineyard </TerrainName>
      <Going> 0.10 </Going>
      <Cover> 0.00 </Cover>
      <Concealment> 0.00 </Concealment>
      <Red> 112 </Red>
      <Green> 68 </Green>
      <Blue> 137 </Blue>
    </TerrainType>
    <TerrainType>
      <TerrainName> Developed Area </TerrainName>
      <Going> 0.40 </Going>
      <Cover> 0.00 </Cover>
      <Concealment> 0.00 </Concealment>
      <Red> 155 </Red>
      <Green> 155 </Green>
      <Blue> 155 </Blue>
    </TerrainType>
  </Battlefield>

<Squad>
  <SquadType> Trooper </SquadType>
  <SquadActive> Yes </SquadActive>
  <SquadCommsEnable> Yes </SquadCommsEnable>
  <index> 1 </index>
  <ParentSqd> 0 </ParentSqd>
  <SquadName> Blue </SquadName>
  <NumAgents> 46 </NumAgents>
  <SquadOnly> Yes </SquadOnly>
  <Use_Momentum> No </Use_Momentum>
  <Use_Diag> Yes </Use_Diag>
  <MultiOccupy> No </MultiOccupy>
  <AutoRflAgt> No </AutoRflAgt>
  <Use_Going> Yes </Use_Going>
  <RooftopAllowed> No </RooftopAllowed>
  <Move_Together> No </Move_Together>
  <MoveSelectType> Precision </MoveSelectType>
  <Homes>
    <HomeNumberOf> 1 </HomeNumberOf>
    <HomeDistribution> Turns </HomeDistribution>
    <HomePos>
      <Home_x_coord> 4624 </Home_x_coord>
      <Home_y_coord> 24 </Home_y_coord>
      <HomeWid> 0 </HomeWid>
```

```
<HomeHgt> 0 </HomeHgt>
<HomeExclDist> 0 </HomeExclDist>
<HomeShape> HomeShapeRect </HomeShape>
<Waypoints>
  <numberOf> 24 </numberOf>
  <loop> No </loop>
  <waypoint>
    <x_coord> 921 </x_coord>
    <y_coord> 2378 </y_coord>
  </waypoint>
  <waypoint>
    <x_coord> 993 </x_coord>
    <y_coord> 2132 </y_coord>
  </waypoint>
  <waypoint>
    <x_coord> 921 </x_coord>
    <y_coord> 1904 </y_coord>
  </waypoint>
  <waypoint>
    <x_coord> 849 </x_coord>
    <y_coord> 1808 </y_coord>
  </waypoint>
  <waypoint>
    <x_coord> 867 </x_coord>
    <y_coord> 1664 </y_coord>
  </waypoint>
  <waypoint>
    <x_coord> 939 </x_coord>
    <y_coord> 1640 </y_coord>
  </waypoint>
  <waypoint>
    <x_coord> 1137 </x_coord>
    <y_coord> 1718 </y_coord>
  </waypoint>
  <waypoint>
    <x_coord> 1239 </x_coord>
    <y_coord> 1910 </y_coord>
  </waypoint>
  <waypoint>
    <x_coord> 1448 </x_coord>
    <y_coord> 1970 </y_coord>
  </waypoint>
  <waypoint>
    <x_coord> 1496 </x_coord>
    <y_coord> 2090 </y_coord>
  </waypoint>
  <waypoint>
    <x_coord> 1514 </x_coord>
    <y_coord> 2863 </y_coord>
  </waypoint>
  <waypoint>
    <x_coord> 1628 </x_coord>
    <y_coord> 3097 </y_coord>
  </waypoint>
  <waypoint>
```

277

```
            <x_coord> 843 </x_coord>
            <y_coord> 3583 </y_coord>
          </waypoint>
          <waypoint>
            <x_coord> 537 </x_coord>
            <y_coord> 3475 </y_coord>
          </waypoint>
          <waypoint>
            <x_coord> 651 </x_coord>
            <y_coord> 3109 </y_coord>
          </waypoint>
          <waypoint>
            <x_coord> 537 </x_coord>
            <y_coord> 3013 </y_coord>
          </waypoint>
          <waypoint>
            <x_coord> 453 </x_coord>
            <y_coord> 2809 </y_coord>
          </waypoint>
          <waypoint>
            <x_coord> 333 </x_coord>
            <y_coord> 2719 </y_coord>
          </waypoint>
          <waypoint>
            <x_coord> 189 </x_coord>
            <y_coord> 2731 </y_coord>
          </waypoint>
          <waypoint>
            <x_coord> 21 </x_coord>
            <y_coord> 2857 </y_coord>
          </waypoint>
          <waypoint>
            <x_coord> 51 </x_coord>
            <y_coord> 4098 </y_coord>
          </waypoint>
          <waypoint>
            <x_coord> 327 </x_coord>
            <y_coord> 4086 </y_coord>
          </waypoint>
          <waypoint>
            <x_coord> 609 </x_coord>
            <y_coord> 4020 </y_coord>
          </waypoint>
          <waypoint>
            <x_coord> 4626 </x_coord>
            <y_coord> 51 </y_coord>
          </waypoint>
          <HaveAltGoal> No </HaveAltGoal>
        </Waypoints>
      </HomePos>
   </Homes>
<FlagIcon> 1 </FlagIcon>
<Comms>
</Comms>
<SALockoutTime> 0 </SALockoutTime>
```

```
<OrgThreatRate> 30 </OrgThreatRate>
<InorgThreatRate> 30 </InorgThreatRate>
<OrgFuseRad> 0 </OrgFuseRad>
<OrgFuseTime> 0 </OrgFuseTime>
<InorgFuseRad> 0 </InorgFuseRad>
<InorgFuseTime> 0 </InorgFuseTime>
<ForceOrgUpdateByCtID> No </ForceOrgUpdateByCtID>
<ForceInorgUpdateByCtID> No </ForceInorgUpdateByCtID>
<ResOrgUnknown> No </ResOrgUnknown>
<ResInorgUnknown> No </ResInorgUnknown>
<PathWidth> 0 </PathWidth>
<TurretBearing> 0 </TurretBearing>
<CommsDelay> 0 </CommsDelay>
<Sound> 6 </Sound>
<Precision> 200 </Precision>
<FuelTank> 1000 </FuelTank>
<CommsQuality> Low </CommsQuality>
<Platform>
   <OrientToX> 0 </OrientToX>
   <OrientToY> 0 </OrientToY>
   <OrientAwayX> 0 </OrientAwayX>
   <OrientAwayY> 0 </OrientAwayY>
   <OrientToBearing> 0 </OrientToBearing>
   <OrientMode> Bearing </OrientMode>
</Platform>
<Patrol_Xmin> 0 </Patrol_Xmin>
<Patrol_Xmax> 2000 </Patrol_Xmax>
<Patrol_Ymin> 0 </Patrol_Ymin>
<Patrol_Ymax> 2000 </Patrol_Ymax>
<SelectZone> No </SelectZone>
<Weapon>
   <WpnSAType> AgentKE </WpnSAType>
   <WpnName> Untitled </WpnName>
   <WpnType> Simple </WpnType>
   <WpnAgtClass> Primary </WpnAgtClass>
   <WpnLockDefState> No </WpnLockDefState>
   <WpnLOF> Yes </WpnLOF>
   <WpnReloadAmmoTime> -1 </WpnReloadAmmoTime>
   <WpnAmmo> 1000 </WpnAmmo>
</Weapon>
<Sensor>
   <SensName> Untitled </SensName>
   <SensType> Simple </SensType>
   <SensClass> A </SensClass>
   <SensLockDefState> No</SensLockDefState>
</Sensor>

<state>
   <StateName> Default State </StateName>
   <SquadAlleg> 1 </SquadAlleg>
   <DefIcon> 1 </DefIcon>
   <Trigger>
     <fallback_to> Default State </fallback_to>
     <duration> 0 </duration>
     <CanInterrupt> Yes </CanInterrupt>
```

279

```
    <Interrupter> No </Interrupter>
    <OnceOnly> No </OnceOnly>
    <RandomDuration> No </RandomDuration>
    <HighestPriority> No </HighestPriority>
  </Trigger>
  <Threat> 3 </Threat>
  <AgtClass> 0 </AgtClass>
  <ReleaseFromParent> Yes </ReleaseFromParent>
  <ReleaseChildren> Yes </ReleaseChildren>
  <EmbussChildren> No </EmbussChildren>
  <ActivateConCombat> No </ActivateConCombat>
  <ActivateConCluster> No </ActivateConCluster>
  <ActivateConAdvance> No </ActivateConAdvance>
  <FormationShape> Wedge </FormationShape>
  <OnlyRespToMovAgt> No </OnlyRespToMovAgt>
  <Stealth> 0 </Stealth>
  <FuelRate> 0 </FuelRate>
  <Speed> 25 </Speed>
  <SpeedUnits> spKm_per_hour </SpeedUnits>
  <Mass> 0.1 </Mass>
  <TrackTarget> No </TrackTarget>
  <CrowdSeparation> 1 </CrowdSeparation>
  <RandomDebussAng> No </RandomDebussAng>
  <PlatformState>
    <SlewRate> 180 </SlewRate>
    <ScanSpan> 180 </ScanSpan>
    <Height> 2 </Height>
    <ScanMode> No Change </ScanMode>
    <ScanDir> Clockwise </ScanDir>
    <OrientUsingAgentSA> Yes </OrientUsingAgentSA>
    <OrientUsingSquadSA> No </OrientUsingSquadSA>
    <OrientUsingInorgSA> No </OrientUsingInorgSA>
    <OrientUsingInorgUnk> No </OrientUsingInorgUnk>
    <OrientTowardsShooter> No </OrientTowardsShooter>
    <ShotAtTime> 5 </ShotAtTime>
  </PlatformState>
  <WeaponState>
    <WpnEnabled> Yes </WpnEnabled>
    <WpnInterpSSKP> No </WpnInterpSSKP>
    <WpnTargetClosest> No </WpnTargetClosest>
    <WpnNumTargets> 100 </WpnNumTargets>
    <WpnShotRadius> 0 </WpnShotRadius>
    <WpnCollateral> No </WpnCollateral>
    <WpnArcOfConcern> 2 </WpnArcOfConcern>
    <WpnMinRange> 0 </WpnMinRange>
    <WpnMaxRange> 1000 </WpnMaxRange>
    <WpnPenetration> 0 </WpnPenetration>
    <WpnArmourPenSTD> 0 </WpnArmourPenSTD>
    <WpnMinThreat> 0 </WpnMinThreat>
    <WpnMaxThreat> 3 </WpnMaxThreat>
    <WpnMaxAge> 100 </WpnMaxAge>
    <WpnMinAp> 360 </WpnMinAp>
    <WpnMaxAp> 0 </WpnMaxAp>
    <WpnTargetUnknownsWait> -1 </WpnTargetUnknownsWait>
    <WpnProtectSelf> No </WpnProtectSelf>
```

280

```
<WpnProtectSqFr> No </WpnProtectSqFr>
<WpnProtectOthFr> No </WpnProtectOthFr>
<WpnProtectNeutrals> No </WpnProtectNeutrals>
<WpnProtectUnknowns> No </WpnProtectUnknowns>
<WpnProtectUsingOrgSA> No </WpnProtectUsingOrgSA>
<WpnProtectUsingInorgSA> No </WpnProtectUsingInorgSA>
<sskpTable>
 <sskpTablePoint>
   <sskpTableRange> 200  </sskpTableRange>
   <sskpTableProb> 10000  </sskpTableProb>
 </sskpTablePoint>
</sskpTable>
<AgtClassTable>
</AgtClassTable>
<nonTargetTable>
</nonTargetTable>
</WeaponState>
<SensorState>
  <SensStEnabled> Yes </SensStEnabled>
  <SensStClassify> 200 </SensStClassify>
  <SensStDetect> 200 </SensStDetect>
  <SensStMinAp> 360 </SensStMinAp>
  <SensStMaxAp> 0 </SensStMaxAp>
  <SensStSync> No </SensStSync>
  <SensSpecClasses> Yes </SensSpecClasses>
  <classTable>
   <classTablePoint>
     <classTableRange> 200  </classTableRange>
     <classTableProb> 10000  </classTableProb>
   </classTablePoint>
  </classTable>
  <detectTable>
   <detectTablePoint>
     <detectTableRange> 200  </detectTableRange>
     <detectTableProb> 0  </detectTableProb>
   </detectTablePoint>
  </detectTable>
  <SensClassTable>
  </SensClassTable>
</SensorState>
<weight><WgtName> Alive Enemy  </WgtName><WgtVal>  0   </WgtVal></weight>
<weight><WgtName> Alive Friends  </WgtName><WgtVal>  54   </WgtVal></weight>
<weight><WgtName> Injured Friends  </WgtName><WgtVal>  33  </WgtVal></weight>
<weight><WgtName> Alive Neutrals  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Next Waypoint  </WgtName><WgtVal>  90  </WgtVal></weight>
<weight><WgtName> Alternate Waypoint  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Easy Terrain  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Cover  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Concealment  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Centre  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Threat 1 (Low)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Threat 2 (Medium)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Threat 3 (High)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Squad Friends  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Other Friends  </WgtName><WgtVal>  0  </WgtVal></weight>
```

281

```
<weight><WgtName> Org Neutrals  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> Org Unknowns  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> Inorg Threat 1 (Low)  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> Inorg Threat 2 (Medium)  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> Inorg Threat 3 (High)  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> Inorg Friends  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> Inorg Neutrals  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> Inorg Unknowns  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> En Threat 1 (Low)  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> En Threat 1 (Medium)  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> En Threat 1 (High)  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> En Ideal  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> Formation  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> AgtEnOrient  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> OrgEnOrient  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> InorgEnOrient  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> Squad Awareness  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> Debuss Behaviour  </WgtName><WgtVal> 0  </WgtVal></weight>
<weight><WgtName> Dynamic Waypoint  </WgtName><WgtVal> 0  </WgtVal></weight>
<range><RangeName> Sensor Range  </RangeName><RangeVal> 1  </RangeVal></range>
<range><RangeName> Detection Range  </RangeName><RangeVal> 1  </RangeVal></range>
<range><RangeName> Movement Speed  </RangeName><RangeVal> 100  </RangeVal></range>
<range><RangeName> No. of Hits Required to Kill  </RangeName><RangeVal> 1  </RangeVal></range>
<range><RangeName> Min Distance to Enemies  </RangeName><RangeVal> 0  </RangeVal></range>
<range><RangeName> Min Distance to Friends  </RangeName><RangeVal> 0  </RangeVal></range>
<range><RangeName> Min Distance to Neutrals  </RangeName><RangeVal> 0  </RangeVal></range>
<range><RangeName> Min Distance to Next Waypoint  </RangeName><RangeVal> 0  </RangeVal></range>
<range><RangeName> Min Distance to Alternate Waypoint  </RangeName><RangeVal> 0
</RangeVal></range>
<range><RangeName> Cluster Constraint  </RangeName><RangeVal> 0  </RangeVal></range>
<range><RangeName> Combat Constraint  </RangeName><RangeVal> 0  </RangeVal></range>
<range><RangeName> Advance Constraint  </RangeName><RangeVal> 0  </RangeVal></range>
<range><RangeName> Refuel Range  </RangeName><RangeVal> 0  </RangeVal></range>
<range><RangeName> Probability of Refuel Enemy  </RangeName><RangeVal> 0  </RangeVal></range>
<range><RangeName> Probability of Refuel Friend  </RangeName><RangeVal> 0  </RangeVal></range>
<range><RangeName> Probability of Refuel Neutral  </RangeName><RangeVal> 0  </RangeVal></range>
<range><RangeName> Armour Thickness  </RangeName><RangeVal> 0  </RangeVal></range>
<range><RangeName> Waypoint Radius  </RangeName><RangeVal> 2  </RangeVal></range>
<range><RangeName> Max Distance to Enemies  </RangeName><RangeVal> 1000000  </RangeVal></range>
<range><RangeName> Max Distance to Friends  </RangeName><RangeVal> 1000000  </RangeVal></range>
<range><RangeName> Max Distance to Neutrals  </RangeName><RangeVal> 1000000  </RangeVal></range>
<range><RangeName> Max Distance to Next Waypoint  </RangeName><RangeVal> 1000000
</RangeVal></range>
<range><RangeName> Max Distance to Alternate Waypoint  </RangeName><RangeVal> 1000000
</RangeVal></range>
<range><RangeName> Min Distance to Organic Threat 1  </RangeName><RangeVal> 0  </RangeVal></range>
<range><RangeName> Min Distance to Organic Threat 2  </RangeName><RangeVal> 0  </RangeVal></range>
<range><RangeName> Min Distance to Organic Threat 3  </RangeName><RangeVal> 0  </RangeVal></range>
<range><RangeName> Min Distance to Organic Squad Friends  </RangeName><RangeVal> 0
</RangeVal></range>
<range><RangeName> Min Distance to Organic Other Friends  </RangeName><RangeVal> 0
</RangeVal></range>
<range><RangeName> Min Distance to Organic Neutrals  </RangeName><RangeVal> 0  </RangeVal></range>
<range><RangeName> Min Distance to Organic Unknowns  </RangeName><RangeVal> 0
</RangeVal></range>
```

```
    <range><RangeName> Max Distance to Organic Threat 1  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Organic Threat 2  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Organic Threat 3  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Organic Squad Friends  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Organic Other Friends  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Organic Neutrals  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Organic Unknowns  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Min Distance to Inorganic Threat 1  </RangeName><RangeVal>  0
</RangeVal></range>
    <range><RangeName> Min Distance to Inorganic Threat 2  </RangeName><RangeVal>  0
</RangeVal></range>
    <range><RangeName> Min Distance to Inorganic Threat 3  </RangeName><RangeVal>  0
</RangeVal></range>
    <range><RangeName> Min Distance to Inorganic Friends  </RangeName><RangeVal>  0
</RangeVal></range>
    <range><RangeName> Min Distance to Inorganic Neutrals  </RangeName><RangeVal>  0
</RangeVal></range>
    <range><RangeName> Min Distance to Inorganic Unknowns  </RangeName><RangeVal>  0
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Threat 1  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Threat 2  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Threat 3  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Friends  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Neutrals  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Unknowns  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Min Distance to Enemy Threat 1  </RangeName><RangeVal>  0  </RangeVal></range>
    <range><RangeName> Min Distance to Enemy Threat 2  </RangeName><RangeVal>  0  </RangeVal></range>
    <range><RangeName> Min Distance to Enemy Threat 3  </RangeName><RangeVal>  0  </RangeVal></range>
    <range><RangeName> Min Distance to Ideal Enemy  </RangeName><RangeVal>  0  </RangeVal></range>
    <range><RangeName> Max Distance to Enemy Threat 1  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Enemy Threat 2  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Enemy Threat 3  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Ideal Enemy  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Ideal Enemy Class  </RangeName><RangeVal>  0  </RangeVal></range>
    <range><RangeName> Max Distance for Combat Constraint  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance for Cluster Constraint  </RangeName><RangeVal>  1000000
</RangeVal></range>
```

283

    <range><RangeName> Max Distance for Advance Constraint </RangeName><RangeVal> 1000000 </RangeVal></range>
    <range><RangeName> Orientation To Enemy Angle (Agt SA) </RangeName><RangeVal> 0 </RangeVal></range>
    <range><RangeName> Orientation To Enemy Angle (Sqd SA) </RangeName><RangeVal> 0 </RangeVal></range>
    <range><RangeName> Orientation To Enemy Angle (Inorg SA) </RangeName><RangeVal> 0 </RangeVal></range>
    <range><RangeName> Min Distance En. Angle Orient (Agt SA) </RangeName><RangeVal> 0 </RangeVal></range>
    <range><RangeName> Min Distance En. Angle Orient (Sqd SA) </RangeName><RangeVal> 0 </RangeVal></range>
    <range><RangeName> Min Distance En. Angle Orient (Inorg SA) </RangeName><RangeVal> 0 </RangeVal></range>
    <range><RangeName> Max Distance En. Angle Orient (Agt SA) </RangeName><RangeVal> 1000000 </RangeVal></range>
    <range><RangeName> Max Distance En. Angle Orient (Sqd SA) </RangeName><RangeVal> 1000000 </RangeVal></range>
    <range><RangeName> Max Distance En. Angle Orient (Inorg SA) </RangeName><RangeVal> 1000000 </RangeVal></range>
    <range><RangeName> Formation Square Width </RangeName><RangeVal> 2 </RangeVal></range>
    <range><RangeName> Formation Circle Diameter </RangeName><RangeVal> 2 </RangeVal></range>
    <range><RangeName> Min Distance to Going </RangeName><RangeVal> 0 </RangeVal></range>
    <range><RangeName> Min Distance to Cover </RangeName><RangeVal> 0 </RangeVal></range>
    <range><RangeName> Min Distance to Concealment </RangeName><RangeVal> 0 </RangeVal></range>
    <range><RangeName> Max Distance to Going </RangeName><RangeVal> 1000000 </RangeVal></range>
    <range><RangeName> Max Distance to Cover </RangeName><RangeVal> 1000000 </RangeVal></range>
    <range><RangeName> Max Distance to Concealment </RangeName><RangeVal> 1000000 </RangeVal></range>
    <range><RangeName> Min Distance to Squad Awareness </RangeName><RangeVal> 0 </RangeVal></range>
    <range><RangeName> Max Distance to Squad Awareness </RangeName><RangeVal> 1000000 </RangeVal></range>
    <range><RangeName> Debuss Angle </RangeName><RangeVal> 0 </RangeVal></range>
    <range><RangeName> Dynamic Wpt Characteristic Dist </RangeName><RangeVal> 50 </RangeVal></range>
    <range><RangeName> Dynamic Wpt Tolerance </RangeName><RangeVal> 1 </RangeVal></range>
    <range><RangeName> Debuss Dist </RangeName><RangeVal> 1 </RangeVal></range>
    <range><RangeName> Embuss Wpt Tolerance </RangeName><RangeVal> 1 </RangeVal></range>
  </state>
</Squad>

<Squad>
  <SquadType> Trooper </SquadType>
  <SquadActive> Yes </SquadActive>
  <SquadCommsEnable> Yes </SquadCommsEnable>
  <index> 2 </index>
  <ParentSqd> 0 </ParentSqd>
  <SquadName> RedPatrol </SquadName>
  <NumAgents> 3 </NumAgents>
  <SquadOnly> Yes </SquadOnly>
  <Use_Momentum> No </Use_Momentum>
  <Use_Diag> Yes </Use_Diag>
  <MultiOccupy> No </MultiOccupy>
  <AutoRflAgt> No </AutoRflAgt>

```xml
<Use_Going> Yes </Use_Going>
<RooftopAllowed> No </RooftopAllowed>
<Move_Together> No </Move_Together>
<MoveSelectType> Precision </MoveSelectType>
<Homes>
  <HomeNumberOf> 1 </HomeNumberOf>
  <HomeDistribution> Turns </HomeDistribution>
  <HomePos>
    <Home_x_coord> 813 </Home_x_coord>
    <Home_y_coord> 1632 </Home_y_coord>
    <HomeWid> 0 </HomeWid>
    <HomeHgt> 0 </HomeHgt>
    <HomeExclDist> 0 </HomeExclDist>
    <HomeShape> HomeShapeRect </HomeShape>
    <Waypoints>
      <numberOf> 5 </numberOf>
      <loop> Yes </loop>
      <waypoint>
        <x_coord> 968 </x_coord>
        <y_coord> 2044 </y_coord>
      </waypoint>
      <waypoint>
        <x_coord> 1196 </x_coord>
        <y_coord> 2029 </y_coord>
      </waypoint>
      <waypoint>
        <x_coord> 1230 </x_coord>
        <y_coord> 1894 </y_coord>
      </waypoint>
      <waypoint>
        <x_coord> 1122 </x_coord>
        <y_coord> 1656 </y_coord>
      </waypoint>
      <waypoint>
        <x_coord> 840 </x_coord>
        <y_coord> 1624 </y_coord>
      </waypoint>
      <HaveAltGoal> No </HaveAltGoal>
    </Waypoints>
  </HomePos>
</Homes>
<FlagIcon> 2 </FlagIcon>
<Comms>
</Comms>
<SALockoutTime> 0 </SALockoutTime>
<OrgThreatRate> 30 </OrgThreatRate>
<InorgThreatRate> 30 </InorgThreatRate>
<OrgFuseRad> 0 </OrgFuseRad>
<OrgFuseTime> 0 </OrgFuseTime>
<InorgFuseRad> 0 </InorgFuseRad>
<InorgFuseTime> 0 </InorgFuseTime>
<ForceOrgUpdateByCtID> No </ForceOrgUpdateByCtID>
<ForceInorgUpdateByCtID> No </ForceInorgUpdateByCtID>
<ResOrgUnknown> No </ResOrgUnknown>
<ResInorgUnknown> No </ResInorgUnknown>
```

285

```
<PathWidth> 0 </PathWidth>
<TurretBearing> 0 </TurretBearing>
<CommsDelay> 0 </CommsDelay>
<Sound> 6 </Sound>
<Precision> 200 </Precision>
<FuelTank> 1000 </FuelTank>
<CommsQuality> Low </CommsQuality>
<Platform>
   <OrientToX> 0 </OrientToX>
   <OrientToY> 0 </OrientToY>
   <OrientAwayX> 0 </OrientAwayX>
   <OrientAwayY> 0 </OrientAwayY>
   <OrientToBearing> 0 </OrientToBearing>
   <OrientMode> Bearing </OrientMode>
</Platform>
<Patrol_Xmin> 0 </Patrol_Xmin>
<Patrol_Xmax> 2000 </Patrol_Xmax>
<Patrol_Ymin> 0 </Patrol_Ymin>
<Patrol_Ymax> 2000 </Patrol_Ymax>
<SelectZone> No </SelectZone>
<Weapon>
   <WpnSAType> AgentKE </WpnSAType>
   <WpnName> Untitled </WpnName>
   <WpnType> Simple </WpnType>
   <WpnAgtClass> Primary </WpnAgtClass>
   <WpnLockDefState> No </WpnLockDefState>
   <WpnLOF> Yes </WpnLOF>
   <WpnReloadAmmoTime> -1 </WpnReloadAmmoTime>
   <WpnAmmo> 1000 </WpnAmmo>
</Weapon>
<Sensor>
   <SensName> Untitled </SensName>
   <SensType> Simple </SensType>
   <SensClass> A </SensClass>
   <SensLockDefState> No</SensLockDefState>
</Sensor>

<state>
   <StateName> Default State </StateName>
   <SquadAlleg> 2 </SquadAlleg>
   <DefIcon> 26 </DefIcon>
   <Trigger>
     <fallback_to> Default State </fallback_to>
     <duration> 0 </duration>
     <CanInterrupt> Yes </CanInterrupt>
     <Interrupter> No </Interrupter>
     <OnceOnly> No </OnceOnly>
     <RandomDuration> No </RandomDuration>
     <HighestPriority> No </HighestPriority>
   </Trigger>
   <Threat> 3 </Threat>
   <AgtClass> 0 </AgtClass>
   <ReleaseFromParent> Yes </ReleaseFromParent>
   <ReleaseChildren> Yes </ReleaseChildren>
   <EmbussChildren> No </EmbussChildren>
```

```
<ActivateConCombat> No </ActivateConCombat>
<ActivateConCluster> No </ActivateConCluster>
<ActivateConAdvance> No </ActivateConAdvance>
<FormationShape> None </FormationShape>
<OnlyRespToMovAgt> No </OnlyRespToMovAgt>
<Stealth> 0 </Stealth>
<FuelRate> 0 </FuelRate>
<Speed> 25 </Speed>
<SpeedUnits> spKm_per_hour </SpeedUnits>
<Mass> 0.1 </Mass>
<TrackTarget> No </TrackTarget>
<CrowdSeparation> 1 </CrowdSeparation>
<RandomDebussAng> No </RandomDebussAng>
<PlatformState>
  <SlewRate> 180 </SlewRate>
  <ScanSpan> 180 </ScanSpan>
  <Height> 2 </Height>
  <ScanMode> No Change </ScanMode>
  <ScanDir> Clockwise </ScanDir>
  <OrientUsingAgentSA> Yes </OrientUsingAgentSA>
  <OrientUsingSquadSA> No </OrientUsingSquadSA>
  <OrientUsingInorgSA> No </OrientUsingInorgSA>
  <OrientUsingInorgUnk> No </OrientUsingInorgUnk>
  <OrientTowardsShooter> No </OrientTowardsShooter>
  <ShotAtTime> 5 </ShotAtTime>
</PlatformState>
<WeaponState>
  <WpnEnabled> Yes </WpnEnabled>
  <WpnInterpSSKP> No </WpnInterpSSKP>
  <WpnTargetClosest> No </WpnTargetClosest>
  <WpnNumTargets> 100 </WpnNumTargets>
  <WpnShotRadius> 0 </WpnShotRadius>
  <WpnCollateral> No </WpnCollateral>
  <WpnArcOfConcern> 2 </WpnArcOfConcern>
  <WpnMinRange> 0 </WpnMinRange>
  <WpnMaxRange> 0 </WpnMaxRange>
  <WpnPenetration> 0 </WpnPenetration>
  <WpnArmourPenSTD> 0 </WpnArmourPenSTD>
  <WpnMinThreat> 0 </WpnMinThreat>
  <WpnMaxThreat> 3 </WpnMaxThreat>
  <WpnMaxAge> 100 </WpnMaxAge>
  <WpnMinAp> 360 </WpnMinAp>
  <WpnMaxAp> 0 </WpnMaxAp>
  <WpnTargetUnknownsWait> -1 </WpnTargetUnknownsWait>
  <WpnProtectSelf> No </WpnProtectSelf>
  <WpnProtectSqFr> No </WpnProtectSqFr>
  <WpnProtectOthFr> No </WpnProtectOthFr>
  <WpnProtectNeutrals> No </WpnProtectNeutrals>
  <WpnProtectUnknowns> No </WpnProtectUnknowns>
  <WpnProtectUsingOrgSA> No </WpnProtectUsingOrgSA>
  <WpnProtectUsingInorgSA> No </WpnProtectUsingInorgSA>
  <sskpTable>
   <sskpTablePoint>
     <sskpTableRange> 200  </sskpTableRange>
     <sskpTableProb> 10000  </sskpTableProb>
```

```
        </sskpTablePoint>
      </sskpTable>
      <AgtClassTable>
      </AgtClassTable>
      <nonTargetTable>
      </nonTargetTable>
</WeaponState>
<SensorState>
      <SensStEnabled> Yes </SensStEnabled>
      <SensStClassify> 200 </SensStClassify>
      <SensStDetect> 200 </SensStDetect>
      <SensStMinAp> 360 </SensStMinAp>
      <SensStMaxAp> 0 </SensStMaxAp>
      <SensStSync> No </SensStSync>
      <SensSpecClasses> Yes </SensSpecClasses>
      <classTable>
       <classTablePoint>
        <classTableRange> 200 </classTableRange>
        <classTableProb> 10000 </classTableProb>
       </classTablePoint>
      </classTable>
      <detectTable>
       <detectTablePoint>
        <detectTableRange> 200 </detectTableRange>
        <detectTableProb> 0 </detectTableProb>
       </detectTablePoint>
      </detectTable>
      <SensClassTable>
      </SensClassTable>
</SensorState>
<weight><WgtName> Alive Enemy  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Alive Friends  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Injured Friends  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Alive Neutrals  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Next Waypoint  </WgtName><WgtVal>  50  </WgtVal></weight>
<weight><WgtName> Alternate Waypoint  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Easy Terrain  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Cover  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Concealment  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Centre  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Threat 1 (Low)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Threat 2 (Medium)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Threat 3 (High)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Squad Friends  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Other Friends  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Neutrals  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Unknowns  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Threat 1 (Low)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Threat 2 (Medium)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Threat 3 (High)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Friends  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Neutrals  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Unknowns  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> En Threat 1 (Low)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> En Threat 1 (Medium)  </WgtName><WgtVal>  0  </WgtVal></weight>
```

```
<weight><WgtName> En Threat 1 (High)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> En Ideal   </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Formation   </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> AgtEnOrient   </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> OrgEnOrient   </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> InorgEnOrient   </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Squad Awareness   </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Debuss Behaviour   </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Dynamic Waypoint   </WgtName><WgtVal>  0  </WgtVal></weight>
<range><RangeName> Sensor Range  </RangeName><RangeVal>  1  </RangeVal></range>
<range><RangeName> Detection Range  </RangeName><RangeVal>  1  </RangeVal></range>
<range><RangeName> Movement Speed  </RangeName><RangeVal>  100  </RangeVal></range>
<range><RangeName> No. of Hits Required to Kill  </RangeName><RangeVal>  1  </RangeVal></range>
<range><RangeName> Min Distance to Enemies  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Friends  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Neutrals  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Next Waypoint  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Alternate Waypoint  </RangeName><RangeVal>  0
</RangeVal></range>
<range><RangeName> Cluster Constraint  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Combat Constraint  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Advance Constraint  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Refuel Range  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Probability of Refuel Enemy  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Probability of Refuel Friend  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Probability of Refuel Neutral  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Armour Thickness  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Waypoint Radius  </RangeName><RangeVal>  2  </RangeVal></range>
<range><RangeName> Max Distance to Enemies  </RangeName><RangeVal>  1000000  </RangeVal></range>
<range><RangeName> Max Distance to Friends  </RangeName><RangeVal>  1000000  </RangeVal></range>
<range><RangeName> Max Distance to Neutrals  </RangeName><RangeVal>  1000000  </RangeVal></range>
<range><RangeName> Max Distance to Next Waypoint  </RangeName><RangeVal>  1000000
</RangeVal></range>
<range><RangeName> Max Distance to Alternate Waypoint  </RangeName><RangeVal>  1000000
</RangeVal></range>
<range><RangeName> Min Distance to Organic Threat 1  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Organic Threat 2  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Organic Threat 3  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Organic Squad Friends  </RangeName><RangeVal>  0
</RangeVal></range>
<range><RangeName> Min Distance to Organic Other Friends  </RangeName><RangeVal>  0
</RangeVal></range>
<range><RangeName> Min Distance to Organic Neutrals  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Organic Unknowns  </RangeName><RangeVal>  0
</RangeVal></range>
<range><RangeName> Max Distance to Organic Threat 1  </RangeName><RangeVal>  1000000
</RangeVal></range>
<range><RangeName> Max Distance to Organic Threat 2  </RangeName><RangeVal>  1000000
</RangeVal></range>
<range><RangeName> Max Distance to Organic Threat 3  </RangeName><RangeVal>  1000000
</RangeVal></range>
<range><RangeName> Max Distance to Organic Squad Friends  </RangeName><RangeVal>  1000000
</RangeVal></range>
<range><RangeName> Max Distance to Organic Other Friends  </RangeName><RangeVal>  1000000
</RangeVal></range>
```

```xml
<range><RangeName> Max Distance to Organic Neutrals   </RangeName><RangeVal>  1000000
</RangeVal></range>
   <range><RangeName> Max Distance to Organic Unknowns   </RangeName><RangeVal>  1000000
</RangeVal></range>
   <range><RangeName> Min Distance to Inorganic Threat 1   </RangeName><RangeVal>  0
</RangeVal></range>
   <range><RangeName> Min Distance to Inorganic Threat 2   </RangeName><RangeVal>  0
</RangeVal></range>
   <range><RangeName> Min Distance to Inorganic Threat 3   </RangeName><RangeVal>  0
</RangeVal></range>
   <range><RangeName> Min Distance to Inorganic Friends   </RangeName><RangeVal>  0
</RangeVal></range>
   <range><RangeName> Min Distance to Inorganic Neutrals   </RangeName><RangeVal>  0
</RangeVal></range>
   <range><RangeName> Min Distance to Inorganic Unknowns   </RangeName><RangeVal>  0
</RangeVal></range>
   <range><RangeName> Max Distance to Inorganic Threat 1   </RangeName><RangeVal>  1000000
</RangeVal></range>
   <range><RangeName> Max Distance to Inorganic Threat 2   </RangeName><RangeVal>  1000000
</RangeVal></range>
   <range><RangeName> Max Distance to Inorganic Threat 3   </RangeName><RangeVal>  1000000
</RangeVal></range>
   <range><RangeName> Max Distance to Inorganic Friends   </RangeName><RangeVal>  1000000
</RangeVal></range>
   <range><RangeName> Max Distance to Inorganic Neutrals   </RangeName><RangeVal>  1000000
</RangeVal></range>
   <range><RangeName> Max Distance to Inorganic Unknowns   </RangeName><RangeVal>  1000000
</RangeVal></range>
   <range><RangeName> Min Distance to Enemy Threat 1   </RangeName><RangeVal>  0  </RangeVal></range>
   <range><RangeName> Min Distance to Enemy Threat 2   </RangeName><RangeVal>  0  </RangeVal></range>
   <range><RangeName> Min Distance to Enemy Threat 3   </RangeName><RangeVal>  0  </RangeVal></range>
   <range><RangeName> Min Distance to Ideal Enemy   </RangeName><RangeVal>  0  </RangeVal></range>
   <range><RangeName> Max Distance to Enemy Threat 1   </RangeName><RangeVal>  1000000
</RangeVal></range>
   <range><RangeName> Max Distance to Enemy Threat 2   </RangeName><RangeVal>  1000000
</RangeVal></range>
   <range><RangeName> Max Distance to Enemy Threat 3   </RangeName><RangeVal>  1000000
</RangeVal></range>
   <range><RangeName> Max Distance to Ideal Enemy   </RangeName><RangeVal>  1000000
</RangeVal></range>
   <range><RangeName> Ideal Enemy Class   </RangeName><RangeVal>  0  </RangeVal></range>
   <range><RangeName> Max Distance for Combat Constraint   </RangeName><RangeVal>  1000000
</RangeVal></range>
   <range><RangeName> Max Distance for Cluster Constraint   </RangeName><RangeVal>  1000000
</RangeVal></range>
   <range><RangeName> Max Distance for Advance Constraint   </RangeName><RangeVal>  1000000
</RangeVal></range>
   <range><RangeName> Orientation To Enemy Angle (Agt SA)   </RangeName><RangeVal>  0
</RangeVal></range>
   <range><RangeName> Orientation To Enemy Angle (Sqd SA)   </RangeName><RangeVal>  0
</RangeVal></range>
   <range><RangeName> Orientation To Enemy Angle (Inorg SA)   </RangeName><RangeVal>  0
</RangeVal></range>
   <range><RangeName> Min Distance En. Angle Orient (Agt SA)   </RangeName><RangeVal>  0
</RangeVal></range>
```

290

```
    <range><RangeName> Min Distance En. Angle Orient (Sqd SA)  </RangeName><RangeVal>  0
</RangeVal></range>
    <range><RangeName> Min Distance En. Angle Orient (Inorg SA)  </RangeName><RangeVal>  0
</RangeVal></range>
    <range><RangeName> Max Distance En. Angle Orient (Agt SA)  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance En. Angle Orient (Sqd SA)  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance En. Angle Orient (Inorg SA)  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Formation Square Width  </RangeName><RangeVal>  2  </RangeVal></range>
    <range><RangeName> Formation Circle Diameter  </RangeName><RangeVal>  2  </RangeVal></range>
    <range><RangeName> Min Distance to Going  </RangeName><RangeVal>  0  </RangeVal></range>
    <range><RangeName> Min Distance to Cover  </RangeName><RangeVal>  0  </RangeVal></range>
    <range><RangeName> Min Distance to Concealment  </RangeName><RangeVal>  0  </RangeVal></range>
    <range><RangeName> Max Distance to Going  </RangeName><RangeVal>  1000000  </RangeVal></range>
    <range><RangeName> Max Distance to Cover  </RangeName><RangeVal>  1000000  </RangeVal></range>
    <range><RangeName> Max Distance to Concealment  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Min Distance to Squad Awareness  </RangeName><RangeVal>  0
</RangeVal></range>
    <range><RangeName> Max Distance to Squad Awareness  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Debuss Angle  </RangeName><RangeVal>  0  </RangeVal></range>
    <range><RangeName> Dynamic Wpt Characteristic Dist  </RangeName><RangeVal>  50
</RangeVal></range>
    <range><RangeName> Dynamic Wpt Tolerance  </RangeName><RangeVal>  1  </RangeVal></range>
    <range><RangeName> Debuss Dist  </RangeName><RangeVal>  1  </RangeVal></range>
    <range><RangeName> Embuss Wpt Tolerance  </RangeName><RangeVal>  1  </RangeVal></range>
  </state>
</Squad>

<Squad>
  <SquadType> Trooper </SquadType>
  <SquadActive> Yes </SquadActive>
  <SquadCommsEnable> Yes </SquadCommsEnable>
  <index> 3 </index>
  <ParentSqd> 0 </ParentSqd>
  <SquadName> RedHQ </SquadName>
  <NumAgents> 8 </NumAgents>
  <SquadOnly> Yes </SquadOnly>
  <Use_Momentum> No </Use_Momentum>
  <Use_Diag> Yes </Use_Diag>
  <MultiOccupy> No </MultiOccupy>
  <AutoRflAgt> No </AutoRflAgt>
  <Use_Going> Yes </Use_Going>
  <RooftopAllowed> No </RooftopAllowed>
  <Move_Together> No </Move_Together>
  <MoveSelectType> Precision </MoveSelectType>
  <Homes>
    <HomeNumberOf> 1 </HomeNumberOf>
    <HomeDistribution> Turns </HomeDistribution>
    <HomePos>
      <Home_x_coord> 901 </Home_x_coord>
      <Home_y_coord> 2377 </Home_y_coord>
```

```xml
<HomeWid> 0 </HomeWid>
<HomeHgt> 0 </HomeHgt>
<HomeExclDist> 0 </HomeExclDist>
<HomeShape> HomeShapeRect </HomeShape>
<Waypoints>
  <numberOf> 4 </numberOf>
  <loop> Yes </loop>
  <waypoint>
    <x_coord> 860 </x_coord>
    <y_coord> 2314 </y_coord>
  </waypoint>
  <waypoint>
    <x_coord> 806 </x_coord>
    <y_coord> 2353 </y_coord>
  </waypoint>
  <waypoint>
    <x_coord> 800 </x_coord>
    <y_coord> 2401 </y_coord>
  </waypoint>
  <waypoint>
    <x_coord> 840 </x_coord>
    <y_coord> 2441 </y_coord>
  </waypoint>
  <HaveAltGoal> No </HaveAltGoal>
</Waypoints>
  </HomePos>
</Homes>
<FlagIcon> 2 </FlagIcon>
<Comms>
</Comms>
<SALockoutTime> 0 </SALockoutTime>
<OrgThreatRate> 30 </OrgThreatRate>
<InorgThreatRate> 30 </InorgThreatRate>
<OrgFuseRad> 0 </OrgFuseRad>
<OrgFuseTime> 0 </OrgFuseTime>
<InorgFuseRad> 0 </InorgFuseRad>
<InorgFuseTime> 0 </InorgFuseTime>
<ForceOrgUpdateByCtID> No </ForceOrgUpdateByCtID>
<ForceInorgUpdateByCtID> No </ForceInorgUpdateByCtID>
<ResOrgUnknown> No </ResOrgUnknown>
<ResInorgUnknown> No </ResInorgUnknown>
<PathWidth> 0 </PathWidth>
<TurretBearing> 0 </TurretBearing>
<CommsDelay> 0 </CommsDelay>
<Sound> 6 </Sound>
<Precision> 200 </Precision>
<FuelTank> 1000 </FuelTank>
<CommsQuality> Low </CommsQuality>
<Platform>
  <OrientToX> 0 </OrientToX>
  <OrientToY> 0 </OrientToY>
  <OrientAwayX> 0 </OrientAwayX>
  <OrientAwayY> 0 </OrientAwayY>
  <OrientToBearing> 0 </OrientToBearing>
  <OrientMode> Bearing </OrientMode>
```

```
</Platform>
<Patrol_Xmin> 0 </Patrol_Xmin>
<Patrol_Xmax> 2000 </Patrol_Xmax>
<Patrol_Ymin> 0 </Patrol_Ymin>
<Patrol_Ymax> 2000 </Patrol_Ymax>
<SelectZone> No </SelectZone>
<Weapon>
  <WpnSAType> AgentKE </WpnSAType>
  <WpnName> Untitled </WpnName>
  <WpnType> Simple </WpnType>
  <WpnAgtClass> Primary </WpnAgtClass>
  <WpnLockDefState> No </WpnLockDefState>
  <WpnLOF> Yes </WpnLOF>
  <WpnReloadAmmoTime> -1 </WpnReloadAmmoTime>
  <WpnAmmo> 1000 </WpnAmmo>
</Weapon>
<Sensor>
  <SensName> Untitled </SensName>
  <SensType> Simple </SensType>
  <SensClass> A </SensClass>
  <SensLockDefState> No</SensLockDefState>
</Sensor>

<state>
  <StateName> Default State </StateName>
  <SquadAlleg> 2 </SquadAlleg>
  <DefIcon> 26 </DefIcon>
  <Trigger>
    <fallback_to> Default State </fallback_to>
    <duration> 0 </duration>
    <CanInterrupt> Yes </CanInterrupt>
    <Interrupter> No </Interrupter>
    <OnceOnly> No </OnceOnly>
    <RandomDuration> No </RandomDuration>
    <HighestPriority> No </HighestPriority>
  </Trigger>
  <Threat> 3 </Threat>
  <AgtClass> 0 </AgtClass>
  <ReleaseFromParent> Yes </ReleaseFromParent>
  <ReleaseChildren> Yes </ReleaseChildren>
  <EmbussChildren> No </EmbussChildren>
  <ActivateConCombat> No </ActivateConCombat>
  <ActivateConCluster> No </ActivateConCluster>
  <ActivateConAdvance> No </ActivateConAdvance>
  <FormationShape> None </FormationShape>
  <OnlyRespToMovAgt> No </OnlyRespToMovAgt>
  <Stealth> 0 </Stealth>
  <FuelRate> 0 </FuelRate>
  <Speed> 25 </Speed>
  <SpeedUnits> spKm_per_hour </SpeedUnits>
  <Mass> 0.1 </Mass>
  <TrackTarget> No </TrackTarget>
  <CrowdSeparation> 1 </CrowdSeparation>
  <RandomDebussAng> No </RandomDebussAng>
  <PlatformState>
```

293

```
    <SlewRate> 180 </SlewRate>
    <ScanSpan> 180 </ScanSpan>
    <Height> 2 </Height>
    <ScanMode> No Change </ScanMode>
    <ScanDir> Clockwise </ScanDir>
    <OrientUsingAgentSA> Yes </OrientUsingAgentSA>
    <OrientUsingSquadSA> No </OrientUsingSquadSA>
    <OrientUsingInorgSA> No </OrientUsingInorgSA>
    <OrientUsingInorgUnk> No </OrientUsingInorgUnk>
    <OrientTowardsShooter> No </OrientTowardsShooter>
    <ShotAtTime> 5 </ShotAtTime>
</PlatformState>
<WeaponState>
    <WpnEnabled> Yes </WpnEnabled>
    <WpnInterpSSKP> No </WpnInterpSSKP>
    <WpnTargetClosest> No </WpnTargetClosest>
    <WpnNumTargets> 100 </WpnNumTargets>
    <WpnShotRadius> 0 </WpnShotRadius>
    <WpnCollateral> No </WpnCollateral>
    <WpnArcOfConcern> 2 </WpnArcOfConcern>
    <WpnMinRange> 0 </WpnMinRange>
    <WpnMaxRange> 0 </WpnMaxRange>
    <WpnPenetration> 0 </WpnPenetration>
    <WpnArmourPenSTD> 0 </WpnArmourPenSTD>
    <WpnMinThreat> 0 </WpnMinThreat>
    <WpnMaxThreat> 3 </WpnMaxThreat>
    <WpnMaxAge> 100 </WpnMaxAge>
    <WpnMinAp> 360 </WpnMinAp>
    <WpnMaxAp> 0 </WpnMaxAp>
    <WpnTargetUnknownsWait> -1 </WpnTargetUnknownsWait>
    <WpnProtectSelf> No </WpnProtectSelf>
    <WpnProtectSqFr> No </WpnProtectSqFr>
    <WpnProtectOthFr> No </WpnProtectOthFr>
    <WpnProtectNeutrals> No </WpnProtectNeutrals>
    <WpnProtectUnknowns> No </WpnProtectUnknowns>
    <WpnProtectUsingOrgSA> No </WpnProtectUsingOrgSA>
    <WpnProtectUsingInorgSA> No </WpnProtectUsingInorgSA>
    <sskpTable>
     <sskpTablePoint>
      <sskpTableRange> 200  </sskpTableRange>
      <sskpTableProb> 10000  </sskpTableProb>
     </sskpTablePoint>
    </sskpTable>
    <AgtClassTable>
    </AgtClassTable>
    <nonTargetTable>
    </nonTargetTable>
</WeaponState>
<SensorState>
    <SensStEnabled> Yes </SensStEnabled>
    <SensStClassify> 200 </SensStClassify>
    <SensStDetect> 200 </SensStDetect>
    <SensStMinAp> 360 </SensStMinAp>
    <SensStMaxAp> 0 </SensStMaxAp>
    <SensStSync> No </SensStSync>
```

```
      <SensSpecClasses> Yes </SensSpecClasses>
      <classTable>
       <classTablePoint>
         <classTableRange> 200  </classTableRange>
         <classTableProb> 10000  </classTableProb>
       </classTablePoint>
      </classTable>
      <detectTable>
       <detectTablePoint>
         <detectTableRange> 200  </detectTableRange>
         <detectTableProb> 0  </detectTableProb>
       </detectTablePoint>
      </detectTable>
      <SensClassTable>
      </SensClassTable>
</SensorState>
<weight><WgtName> Alive Enemy  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Alive Friends  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Injured Friends  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Alive Neutrals  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Next Waypoint  </WgtName><WgtVal>  50  </WgtVal></weight>
<weight><WgtName> Alternate Waypoint  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Easy Terrain  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Cover  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Concealment  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Centre  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Threat 1 (Low)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Threat 2 (Medium)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Threat 3 (High)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Squad Friends  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Other Friends  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Neutrals  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Unknowns  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Threat 1 (Low)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Threat 2 (Medium)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Threat 3 (High)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Friends  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Neutrals  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Unknowns  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> En Threat 1 (Low)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> En Threat 1 (Medium)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> En Threat 1 (High)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> En Ideal  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Formation  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> AgtEnOrient  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> OrgEnOrient  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> InorgEnOrient  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Squad Awareness  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Debuss Behaviour  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Dynamic Waypoint  </WgtName><WgtVal>  0  </WgtVal></weight>
<range><RangeName> Sensor Range  </RangeName><RangeVal>  1  </RangeVal></range>
<range><RangeName> Detection Range  </RangeName><RangeVal>  1  </RangeVal></range>
<range><RangeName> Movement Speed  </RangeName><RangeVal>  100  </RangeVal></range>
<range><RangeName> No. of Hits Required to Kill  </RangeName><RangeVal>  3  </RangeVal></range>
<range><RangeName> Min Distance to Enemies  </RangeName><RangeVal>  0  </RangeVal></range>
```

```
<range><RangeName> Min Distance to Friends  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Neutrals  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Next Waypoint  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Alternate Waypoint  </RangeName><RangeVal>  0
</RangeVal></range>
<range><RangeName> Cluster Constraint  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Combat Constraint  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Advance Constraint  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Refuel Range  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Probability of Refuel Enemy  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Probability of Refuel Friend  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Probability of Refuel Neutral  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Armour Thickness  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Waypoint Radius  </RangeName><RangeVal>  2  </RangeVal></range>
<range><RangeName> Max Distance to Enemies  </RangeName><RangeVal>  1000000  </RangeVal></range>
<range><RangeName> Max Distance to Friends  </RangeName><RangeVal>  1000000  </RangeVal></range>
<range><RangeName> Max Distance to Neutrals  </RangeName><RangeVal>  1000000  </RangeVal></range>
<range><RangeName> Max Distance to Next Waypoint  </RangeName><RangeVal>  1000000
</RangeVal></range>
<range><RangeName> Max Distance to Alternate Waypoint  </RangeName><RangeVal>  1000000
</RangeVal></range>
<range><RangeName> Min Distance to Organic Threat 1  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Organic Threat 2  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Organic Threat 3  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Organic Squad Friends  </RangeName><RangeVal>  0
</RangeVal></range>
<range><RangeName> Min Distance to Organic Other Friends  </RangeName><RangeVal>  0
</RangeVal></range>
<range><RangeName> Min Distance to Organic Neutrals  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Organic Unknowns  </RangeName><RangeVal>  0
</RangeVal></range>
<range><RangeName> Max Distance to Organic Threat 1  </RangeName><RangeVal>  1000000
</RangeVal></range>
<range><RangeName> Max Distance to Organic Threat 2  </RangeName><RangeVal>  1000000
</RangeVal></range>
<range><RangeName> Max Distance to Organic Threat 3  </RangeName><RangeVal>  1000000
</RangeVal></range>
<range><RangeName> Max Distance to Organic Squad Friends  </RangeName><RangeVal>  1000000
</RangeVal></range>
<range><RangeName> Max Distance to Organic Other Friends  </RangeName><RangeVal>  1000000
</RangeVal></range>
<range><RangeName> Max Distance to Organic Neutrals  </RangeName><RangeVal>  1000000
</RangeVal></range>
<range><RangeName> Max Distance to Organic Unknowns  </RangeName><RangeVal>  1000000
</RangeVal></range>
<range><RangeName> Min Distance to Inorganic Threat 1  </RangeName><RangeVal>  0
</RangeVal></range>
<range><RangeName> Min Distance to Inorganic Threat 2  </RangeName><RangeVal>  0
</RangeVal></range>
<range><RangeName> Min Distance to Inorganic Threat 3  </RangeName><RangeVal>  0
</RangeVal></range>
<range><RangeName> Min Distance to Inorganic Friends  </RangeName><RangeVal>  0
</RangeVal></range>
<range><RangeName> Min Distance to Inorganic Neutrals  </RangeName><RangeVal>  0
</RangeVal></range>
```

    <range><RangeName> Min Distance to Inorganic Unknowns   </RangeName><RangeVal> 0
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Threat 1   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Threat 2   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Threat 3   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Friends   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Neutrals   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Unknowns   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Min Distance to Enemy Threat 1   </RangeName><RangeVal> 0 </RangeVal></range>
    <range><RangeName> Min Distance to Enemy Threat 2   </RangeName><RangeVal> 0 </RangeVal></range>
    <range><RangeName> Min Distance to Enemy Threat 3   </RangeName><RangeVal> 0 </RangeVal></range>
    <range><RangeName> Min Distance to Ideal Enemy   </RangeName><RangeVal> 0 </RangeVal></range>
    <range><RangeName> Max Distance to Enemy Threat 1   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Enemy Threat 2   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Enemy Threat 3   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Ideal Enemy   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Ideal Enemy Class   </RangeName><RangeVal> 0 </RangeVal></range>
    <range><RangeName> Max Distance for Combat Constraint   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance for Cluster Constraint   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance for Advance Constraint   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Orientation To Enemy Angle (Agt SA)   </RangeName><RangeVal> 0
</RangeVal></range>
    <range><RangeName> Orientation To Enemy Angle (Sqd SA)   </RangeName><RangeVal> 0
</RangeVal></range>
    <range><RangeName> Orientation To Enemy Angle (Inorg SA)   </RangeName><RangeVal> 0
</RangeVal></range>
    <range><RangeName> Min Distance En. Angle Orient (Agt SA)   </RangeName><RangeVal> 0
</RangeVal></range>
    <range><RangeName> Min Distance En. Angle Orient (Sqd SA)   </RangeName><RangeVal> 0
</RangeVal></range>
    <range><RangeName> Min Distance En. Angle Orient (Inorg SA)   </RangeName><RangeVal> 0
</RangeVal></range>
    <range><RangeName> Max Distance En. Angle Orient (Agt SA)   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance En. Angle Orient (Sqd SA)   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance En. Angle Orient (Inorg SA)   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Formation Square Width   </RangeName><RangeVal> 2   </RangeVal></range>
    <range><RangeName> Formation Circle Diameter   </RangeName><RangeVal> 2   </RangeVal></range>
    <range><RangeName> Min Distance to Going   </RangeName><RangeVal> 0   </RangeVal></range>
    <range><RangeName> Min Distance to Cover   </RangeName><RangeVal> 0   </RangeVal></range>

```
    <range><RangeName> Min Distance to Concealment  </RangeName><RangeVal> 0  </RangeVal></range>
    <range><RangeName> Max Distance to Going  </RangeName><RangeVal> 1000000  </RangeVal></range>
    <range><RangeName> Max Distance to Cover  </RangeName><RangeVal> 1000000  </RangeVal></range>
    <range><RangeName> Max Distance to Concealment  </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Min Distance to Squad Awareness  </RangeName><RangeVal> 0
</RangeVal></range>
    <range><RangeName> Max Distance to Squad Awareness  </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Debuss Angle  </RangeName><RangeVal> 0  </RangeVal></range>
    <range><RangeName> Dynamic Wpt Characteristic Dist  </RangeName><RangeVal> 50
</RangeVal></range>
    <range><RangeName> Dynamic Wpt Tolerance  </RangeName><RangeVal> 1  </RangeVal></range>
    <range><RangeName> Debuss Dist  </RangeName><RangeVal> 1  </RangeVal></range>
    <range><RangeName> Embuss Wpt Tolerance  </RangeName><RangeVal> 1  </RangeVal></range>
  </state>
</Squad>

<Squad>
  <SquadType> Trooper </SquadType>
  <SquadActive> Yes </SquadActive>
  <SquadCommsEnable> Yes </SquadCommsEnable>
  <index> 4 </index>
  <ParentSqd> 0 </ParentSqd>
  <SquadName> RedRoadGuard </SquadName>
  <NumAgents> 4 </NumAgents>
  <SquadOnly> Yes </SquadOnly>
  <Use_Momentum> No </Use_Momentum>
  <Use_Diag> Yes </Use_Diag>
  <MultiOccupy> No </MultiOccupy>
  <AutoRflAgt> No </AutoRflAgt>
  <Use_Going> Yes </Use_Going>
  <RooftopAllowed> No </RooftopAllowed>
  <Move_Together> No </Move_Together>
  <MoveSelectType> Precision </MoveSelectType>
  <Homes>
    <HomeNumberOf> 1 </HomeNumberOf>
    <HomeDistribution> Turns </HomeDistribution>
    <HomePos>
      <Home_x_coord> 632 </Home_x_coord>
      <Home_y_coord> 3439 </Home_y_coord>
      <HomeWid> 0 </HomeWid>
      <HomeHgt> 0 </HomeHgt>
      <HomeExclDist> 0 </HomeExclDist>
      <HomeShape> HomeShapeRect </HomeShape>
      <Waypoints>
        <numberOf> 2 </numberOf>
        <loop> Yes </loop>
        <waypoint>
          <x_coord> 524 </x_coord>
          <y_coord> 3415 </y_coord>
        </waypoint>
        <waypoint>
          <x_coord> 645 </x_coord>
          <y_coord> 3471 </y_coord>
```

298

```
      </waypoint>
      <HaveAltGoal> No </HaveAltGoal>
    </Waypoints>
  </HomePos>
</Homes>
<FlagIcon> 2 </FlagIcon>
<Comms>
</Comms>
<SALockoutTime> 0 </SALockoutTime>
<OrgThreatRate> 30 </OrgThreatRate>
<InorgThreatRate> 30 </InorgThreatRate>
<OrgFuseRad> 0 </OrgFuseRad>
<OrgFuseTime> 0 </OrgFuseTime>
<InorgFuseRad> 0 </InorgFuseRad>
<InorgFuseTime> 0 </InorgFuseTime>
<ForceOrgUpdateByCtID> No </ForceOrgUpdateByCtID>
<ForceInorgUpdateByCtID> No </ForceInorgUpdateByCtID>
<ResOrgUnknown> No </ResOrgUnknown>
<ResInorgUnknown> No </ResInorgUnknown>
<PathWidth> 0 </PathWidth>
<TurretBearing> 0 </TurretBearing>
<CommsDelay> 0 </CommsDelay>
<Sound> 6 </Sound>
<Precision> 200 </Precision>
<FuelTank> 1000 </FuelTank>
<CommsQuality> Low </CommsQuality>
<Platform>
   <OrientToX> 0 </OrientToX>
   <OrientToY> 0 </OrientToY>
   <OrientAwayX> 0 </OrientAwayX>
   <OrientAwayY> 0 </OrientAwayY>
   <OrientToBearing> 120 </OrientToBearing>
   <OrientMode> Bearing </OrientMode>
</Platform>
<Patrol_Xmin> 0 </Patrol_Xmin>
<Patrol_Xmax> 2000 </Patrol_Xmax>
<Patrol_Ymin> 0 </Patrol_Ymin>
<Patrol_Ymax> 2000 </Patrol_Ymax>
<SelectZone> No </SelectZone>
<Weapon>
   <WpnSAType> AgentKE </WpnSAType>
   <WpnName> Untitled </WpnName>
   <WpnType> Simple </WpnType>
   <WpnAgtClass> Primary </WpnAgtClass>
   <WpnLockDefState> No </WpnLockDefState>
   <WpnLOF> Yes </WpnLOF>
   <WpnReloadAmmoTime> -1 </WpnReloadAmmoTime>
   <WpnAmmo> 1000 </WpnAmmo>
</Weapon>
<Sensor>
   <SensName> Untitled </SensName>
   <SensType> Simple </SensType>
   <SensClass> A </SensClass>
   <SensLockDefState> No</SensLockDefState>
</Sensor>
```

299

```
<state>
  <StateName> Default State </StateName>
  <SquadAlleg> 2 </SquadAlleg>
  <DefIcon> 26 </DefIcon>
  <Trigger>
    <fallback_to> Default State </fallback_to>
    <duration> 0 </duration>
    <CanInterrupt> Yes </CanInterrupt>
    <Interrupter> No </Interrupter>
    <OnceOnly> No </OnceOnly>
    <RandomDuration> No </RandomDuration>
    <HighestPriority> No </HighestPriority>
  </Trigger>
  <Threat> 3 </Threat>
  <AgtClass> 0 </AgtClass>
  <ReleaseFromParent> Yes </ReleaseFromParent>
  <ReleaseChildren> Yes </ReleaseChildren>
  <EmbussChildren> No </EmbussChildren>
  <ActivateConCombat> No </ActivateConCombat>
  <ActivateConCluster> No </ActivateConCluster>
  <ActivateConAdvance> No </ActivateConAdvance>
  <FormationShape> None </FormationShape>
  <OnlyRespToMovAgt> No </OnlyRespToMovAgt>
  <Stealth> 0 </Stealth>
  <FuelRate> 0 </FuelRate>
  <Speed> 25 </Speed>
  <SpeedUnits> spKm_per_hour </SpeedUnits>
  <Mass> 0.1 </Mass>
  <TrackTarget> No </TrackTarget>
  <CrowdSeparation> 1 </CrowdSeparation>
  <RandomDebussAng> No </RandomDebussAng>
  <PlatformState>
    <SlewRate> 180 </SlewRate>
    <ScanSpan> 180 </ScanSpan>
    <Height> 2 </Height>
    <ScanMode> No Change </ScanMode>
    <ScanDir> Clockwise </ScanDir>
    <OrientUsingAgentSA> Yes </OrientUsingAgentSA>
    <OrientUsingSquadSA> No </OrientUsingSquadSA>
    <OrientUsingInorgSA> No </OrientUsingInorgSA>
    <OrientUsingInorgUnk> No </OrientUsingInorgUnk>
    <OrientTowardsShooter> No </OrientTowardsShooter>
    <ShotAtTime> 5 </ShotAtTime>
  </PlatformState>
  <WeaponState>
    <WpnEnabled> Yes </WpnEnabled>
    <WpnInterpSSKP> No </WpnInterpSSKP>
    <WpnTargetClosest> No </WpnTargetClosest>
    <WpnNumTargets> 100 </WpnNumTargets>
    <WpnShotRadius> 0 </WpnShotRadius>
    <WpnCollateral> No </WpnCollateral>
    <WpnArcOfConcern> 2 </WpnArcOfConcern>
    <WpnMinRange> 0 </WpnMinRange>
    <WpnMaxRange> 0 </WpnMaxRange>
```

```
<WpnPenetration> 0 </WpnPenetration>
<WpnArmourPenSTD> 0 </WpnArmourPenSTD>
<WpnMinThreat> 0 </WpnMinThreat>
<WpnMaxThreat> 3 </WpnMaxThreat>
<WpnMaxAge> 100 </WpnMaxAge>
<WpnMinAp> 360 </WpnMinAp>
<WpnMaxAp> 0 </WpnMaxAp>
<WpnTargetUnknownsWait> -1 </WpnTargetUnknownsWait>
<WpnProtectSelf> No </WpnProtectSelf>
<WpnProtectSqFr> No </WpnProtectSqFr>
<WpnProtectOthFr> No </WpnProtectOthFr>
<WpnProtectNeutrals> No </WpnProtectNeutrals>
<WpnProtectUnknowns> No </WpnProtectUnknowns>
<WpnProtectUsingOrgSA> No </WpnProtectUsingOrgSA>
<WpnProtectUsingInorgSA> No </WpnProtectUsingInorgSA>
<sskpTable>
  <sskpTablePoint>
    <sskpTableRange> 200  </sskpTableRange>
    <sskpTableProb> 10000  </sskpTableProb>
  </sskpTablePoint>
</sskpTable>
<AgtClassTable>
</AgtClassTable>
<nonTargetTable>
</nonTargetTable>
</WeaponState>
<SensorState>
  <SensStEnabled> Yes </SensStEnabled>
  <SensStClassify> 200 </SensStClassify>
  <SensStDetect> 200 </SensStDetect>
  <SensStMinAp> 360 </SensStMinAp>
  <SensStMaxAp> 0 </SensStMaxAp>
  <SensStSync> No </SensStSync>
  <SensSpecClasses> Yes </SensSpecClasses>
  <classTable>
    <classTablePoint>
      <classTableRange> 200  </classTableRange>
      <classTableProb> 10000  </classTableProb>
    </classTablePoint>
  </classTable>
  <detectTable>
    <detectTablePoint>
      <detectTableRange> 200  </detectTableRange>
      <detectTableProb> 0 </detectTableProb>
    </detectTablePoint>
  </detectTable>
  <SensClassTable>
  </SensClassTable>
</SensorState>
<weight><WgtName> Alive Enemy   </WgtName><WgtVal>  0   </WgtVal></weight>
<weight><WgtName> Alive Friends   </WgtName><WgtVal>  0   </WgtVal></weight>
<weight><WgtName> Injured Friends   </WgtName><WgtVal>  0   </WgtVal></weight>
<weight><WgtName> Alive Neutrals   </WgtName><WgtVal>  0   </WgtVal></weight>
<weight><WgtName> Next Waypoint   </WgtName><WgtVal>  50   </WgtVal></weight>
<weight><WgtName> Alternate Waypoint   </WgtName><WgtVal>  0   </WgtVal></weight>
```

301

```
<weight><WgtName> Easy Terrain  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Cover  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Concealment  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Centre  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Threat 1 (Low)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Threat 2 (Medium)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Threat 3 (High)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Squad Friends  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Other Friends  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Neutrals  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Org Unknowns  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Threat 1 (Low)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Threat 2 (Medium)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Threat 3 (High)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Friends  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Neutrals  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Inorg Unknowns  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> En Threat 1 (Low)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> En Threat 1 (Medium)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> En Threat 1 (High)  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> En Ideal  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Formation  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> AgtEnOrient  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> OrgEnOrient  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> InorgEnOrient  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Squad Awareness  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Debuss Behaviour  </WgtName><WgtVal>  0  </WgtVal></weight>
<weight><WgtName> Dynamic Waypoint  </WgtName><WgtVal>  0  </WgtVal></weight>
<range><RangeName> Sensor Range  </RangeName><RangeVal>  1  </RangeVal></range>
<range><RangeName> Detection Range  </RangeName><RangeVal>  1  </RangeVal></range>
<range><RangeName> Movement Speed  </RangeName><RangeVal>  100  </RangeVal></range>
<range><RangeName> No. of Hits Required to Kill  </RangeName><RangeVal>  2  </RangeVal></range>
<range><RangeName> Min Distance to Enemies  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Friends  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Neutrals  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Next Waypoint  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Min Distance to Alternate Waypoint  </RangeName><RangeVal>  0
</RangeVal></range>
<range><RangeName> Cluster Constraint  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Combat Constraint  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Advance Constraint  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Refuel Range  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Probability of Refuel Enemy  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Probability of Refuel Friend  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Probability of Refuel Neutral  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Armour Thickness  </RangeName><RangeVal>  0  </RangeVal></range>
<range><RangeName> Waypoint Radius  </RangeName><RangeVal>  2  </RangeVal></range>
<range><RangeName> Max Distance to Enemies  </RangeName><RangeVal>  1000000  </RangeVal></range>
<range><RangeName> Max Distance to Friends  </RangeName><RangeVal>  1000000  </RangeVal></range>
<range><RangeName> Max Distance to Neutrals  </RangeName><RangeVal>  1000000  </RangeVal></range>
<range><RangeName> Max Distance to Next Waypoint  </RangeName><RangeVal>  1000000
</RangeVal></range>
<range><RangeName> Max Distance to Alternate Waypoint  </RangeName><RangeVal>  1000000
</RangeVal></range>
<range><RangeName> Min Distance to Organic Threat 1  </RangeName><RangeVal>  0  </RangeVal></range>
```

    <range><RangeName> Min Distance to Organic Threat 2   </RangeName><RangeVal> 0   </RangeVal></range>
    <range><RangeName> Min Distance to Organic Threat 3   </RangeName><RangeVal> 0   </RangeVal></range>
    <range><RangeName> Min Distance to Organic Squad Friends   </RangeName><RangeVal> 0
</RangeVal></range>
    <range><RangeName> Min Distance to Organic Other Friends   </RangeName><RangeVal> 0
</RangeVal></range>
    <range><RangeName> Min Distance to Organic Neutrals   </RangeName><RangeVal> 0   </RangeVal></range>
    <range><RangeName> Min Distance to Organic Unknowns   </RangeName><RangeVal> 0
</RangeVal></range>
    <range><RangeName> Max Distance to Organic Threat 1   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Organic Threat 2   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Organic Threat 3   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Organic Squad Friends   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Organic Other Friends   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Organic Neutrals   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Organic Unknowns   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Min Distance to Inorganic Threat 1   </RangeName><RangeVal> 0
</RangeVal></range>
    <range><RangeName> Min Distance to Inorganic Threat 2   </RangeName><RangeVal> 0
</RangeVal></range>
    <range><RangeName> Min Distance to Inorganic Threat 3   </RangeName><RangeVal> 0
</RangeVal></range>
    <range><RangeName> Min Distance to Inorganic Friends   </RangeName><RangeVal> 0
</RangeVal></range>
    <range><RangeName> Min Distance to Inorganic Neutrals   </RangeName><RangeVal> 0
</RangeVal></range>
    <range><RangeName> Min Distance to Inorganic Unknowns   </RangeName><RangeVal> 0
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Threat 1   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Threat 2   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Threat 3   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Friends   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Neutrals   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Inorganic Unknowns   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Min Distance to Enemy Threat 1   </RangeName><RangeVal> 0   </RangeVal></range>
    <range><RangeName> Min Distance to Enemy Threat 2   </RangeName><RangeVal> 0   </RangeVal></range>
    <range><RangeName> Min Distance to Enemy Threat 3   </RangeName><RangeVal> 0   </RangeVal></range>
    <range><RangeName> Min Distance to Ideal Enemy   </RangeName><RangeVal> 0   </RangeVal></range>
    <range><RangeName> Max Distance to Enemy Threat 1   </RangeName><RangeVal> 1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Enemy Threat 2   </RangeName><RangeVal> 1000000
</RangeVal></range>

303

```
    <range><RangeName> Max Distance to Enemy Threat 3  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance to Ideal Enemy  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Ideal Enemy Class  </RangeName><RangeVal>  0  </RangeVal></range>
    <range><RangeName> Max Distance for Combat Constraint  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance for Cluster Constraint  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance for Advance Constraint  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Orientation To Enemy Angle (Agt SA)  </RangeName><RangeVal>  0
</RangeVal></range>
    <range><RangeName> Orientation To Enemy Angle (Sqd SA)  </RangeName><RangeVal>  0
</RangeVal></range>
    <range><RangeName> Orientation To Enemy Angle (Inorg SA)  </RangeName><RangeVal>  0
</RangeVal></range>
    <range><RangeName> Min Distance En. Angle Orient (Agt SA)  </RangeName><RangeVal>  0
</RangeVal></range>
    <range><RangeName> Min Distance En. Angle Orient (Sqd SA)  </RangeName><RangeVal>  0
</RangeVal></range>
    <range><RangeName> Min Distance En. Angle Orient (Inorg SA)  </RangeName><RangeVal>  0
</RangeVal></range>
    <range><RangeName> Max Distance En. Angle Orient (Agt SA)  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance En. Angle Orient (Sqd SA)  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Max Distance En. Angle Orient (Inorg SA)  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Formation Square Width  </RangeName><RangeVal>  2  </RangeVal></range>
    <range><RangeName> Formation Circle Diameter  </RangeName><RangeVal>  2  </RangeVal></range>
    <range><RangeName> Min Distance to Going  </RangeName><RangeVal>  0  </RangeVal></range>
    <range><RangeName> Min Distance to Cover  </RangeName><RangeVal>  0  </RangeVal></range>
    <range><RangeName> Min Distance to Concealment  </RangeName><RangeVal>  0  </RangeVal></range>
    <range><RangeName> Max Distance to Going  </RangeName><RangeVal>  1000000  </RangeVal></range>
    <range><RangeName> Max Distance to Cover  </RangeName><RangeVal>  1000000  </RangeVal></range>
    <range><RangeName> Max Distance to Concealment  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Min Distance to Squad Awareness  </RangeName><RangeVal>  0
</RangeVal></range>
    <range><RangeName> Max Distance to Squad Awareness  </RangeName><RangeVal>  1000000
</RangeVal></range>
    <range><RangeName> Debuss Angle  </RangeName><RangeVal>  0  </RangeVal></range>
    <range><RangeName> Dynamic Wpt Characteristic Dist  </RangeName><RangeVal>  50
</RangeVal></range>
    <range><RangeName> Dynamic Wpt Tolerance  </RangeName><RangeVal>  1  </RangeVal></range>
    <range><RangeName> Debuss Dist  </RangeName><RangeVal>  1  </RangeVal></range>
    <range><RangeName> Embuss Wpt Tolerance  </RangeName><RangeVal>  1  </RangeVal></range>
  </state>
</Squad>

</specification>
```

APPENDIX B: SIGNIFICANT CLUSTERS

Position Data Cluster Statistics for MANA Simulation

Cluster	V6	V2	V4	V5	V12	V11	V10	V9	V3	V8	V7	V1
p-value	1.13E-12	2.10E-09	1.19E-05	1.19E-05	0.001301	0.01642	0.101319	0.161118	0.239944	0.239944	0.335906	0.443862
number o	50	50	50	50	50	50	50	50	50	50	50	50
number o	48	45	40	40	36	33	30	29	28	28	27	26
polarity	0.96	0.9	0.8	0.8	0.72	0.66	0.6	0.58	0.56	0.56	0.54	0.52

Capability Data Cluster Statistics for MANA Simulation

Cluster	V15	V6	V20	V3	V22	V4	V25	V2	V18	V33
p-value	0.000919	0.006642	0.044766	0.050568	0.071653	0.125	0.125	0.166153	0.22569	0.25
number of trials	135	48	35	84	38	3	3	17	44	2
number of successes	86	33	23	50	24	3	3	11	25	2
polarity	0.637037	0.6875	0.657143	0.595238	0.631579	1	1	0.647059	0.568182	1
Cluster	V19	V16	V9	V12	V5	V13	V1	V8	V10	V14
p-value	0.290527	0.3125	0.321984	0.33882	0.363281	0.376953	0.427768	0.5	0.5	0.5
number of trials	13	4	42	23	8	10	30	3	5	17
number of successes	8	3	23	13	5	6	16	2	3	9
polarity	0.615385	0.75	0.547619	0.565217	0.625	0.6	0.533333	0.666667	0.6	0.529412
Cluster	V23	V21	V24	V28	V7	V11	V17	V26	V27	V29
p-value	0.5	0.59819	0.75	0.75	1	1	1	1	1	1
number of trials	7	16	2	2	1	1	1	1	1	1
number of successes	4	8	1	1	1	1	1	1	1	1
polarity	0.571429	0.5	0.5	0.5	1	1	1	1	1	1
Cluster	V30	V31	V32							
p-value	1	1	1							
number of trials	1	1	1							
number of successes	1	1	1							
polarity	1	1	1							

REFERENCES

1.   United States. Joint Chiefs of Staff., *Joint security operations in theater*, in *Joint publication 3-10*. 2010, Joint Chiefs of Staff: Washington, D.C.

2.   Macias, A., Bender, Jeremy, Gould, Skye *The 35 Most Powerful Militaries In The World*. Business Insider, 2014.

3.   Moffat, J., *Adapting modeling & simulation for Network enabled operations*. 2011, DTIC Document.

4.   Gertler, J. *Operation Odyssey Dawn (Libya): Background and Issues for Congress*. 2011. Congressional Research Service, Library of Congress.

5.   Hundley, R.O., *Past Revolutions, Future Transformations. What Can the History of Revolutions in Military Affairs Tell us about Transforming the US Military?* 1999, DTIC Document.

6.   Tilford Jr, E.H., *The revolution in military affairs: Prospects and cautions*. 1995, DTIC Document.

7.   Layne, C., *America as European hegemon.* The National Interest, 2003: p. 17-29.

8.   Pfaff, C.A., *Chaos, complexity and the battlefield.* Military Review, 2000. **80**(4): p. 82.

9.   Mitchell, P.T., *Network Centric Warfare: Coalition operations in the age of US military primacy*. 2013: Routledge.

10.   Neufeld, J., *Technology and the Air Force: A Retrospective Assessment*. 2009: DIANE Publishing.

11.   Alderson, D.L. and J.C. Doyle, *Contrasting views of complexity and their implications for network-centric infrastructures.* Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 2010. **40**(4): p. 839-852.

12.   Alberts, D.S., J.J. Garstka, and F.P. Stein, *Network Centric Warfare: Developing and Leveraging Information Superiority*. 2000, DTIC Document.

13.   United States. Joint Chiefs of Staff., *Joint operation planning*, in *Joint publication v 5-0*. 2011, Joint Chiefs of Staff: Washington, DC.

14.   Boukhtouta, A., et al. *A survey of military planning systems*. in *The 9th ICCRTS Int. Command and Control Research and Technology Symposium*. 2004.

15.   Kaste, R., et al., *Research in modeling and simulation for command and control*. 2000, DTIC Document.

16.   Clausewitz, C.v., *On war*, ed. M. Howard and P. Paret. 1976, Princeton University Press: Princeton, N.J.

17.   Kuylenstierna, J., J. Rydmark, and T. Fahraeus. *The Value of Information in War: Some Experimental Findings*. in *Proceedings of 5th International Command and Control Research and Technology Symposium*. 2000. Citeseer.

18.   Corps, U.M., *MCDP 6: Command and Control*. Washington, DC: Department of the Navy: Headquarters of the Marine Corps, 1996.

19.   Yao, W., et al., *Review of uncertainty-based multidisciplinary design optimization methods for aerospace vehicles*. Progress in Aerospace Sciences, 2011. **47**(6): p. 450-479.

20.   Thunnissen, D.P. *Uncertainty classification for the design and development of complex systems*. in *3rd annual predictive methods conference*. 2003.

21.   Perry, W.L. and J. Moffat, *Information Sharing Among Military Headquarters: The Effects of Decisionmaking*. 2004, DTIC Document.

22.   Kruger, K., U. Schade, and J. Ziegler. *Uncertainty in the fusion of information from multiple diverse sources for situation awareness*. in *Information Fusion, 2008 11th International Conference on*. 2008. IEEE.

23.   Vego, M.N., *Systems versus Classical Approach to Warfare*. 2009, DTIC Document.

24.   Neate, M.C., *The Battle of the Narrative*. 2010, DTIC Document.

25.   Flynn, G.J., *Decade of war, volume 1: Enduring lessons from the past decade of operations*. 2012, Joint Staff J-7, Joint and Coalition Operational Analysis (JCOA), Suffolk, VA.

26.   Harland, D., *United Nations Peacekeeping Today: Current Challenges and Required Responses*. United Nations as Peacekeeper and Nation-Builder Continuity and Change–What Lies Ahead, 2007: p. 169-183.

27.   Montgomery, E.B., *Defense Planning for the Long Haul: Scenarios, Operational Concepts, and the Future Security Environment*. 2009, DTIC Document.

28.   Wittnam, D.J., *Fratricide Prevention: An Opportunity to Develop a Joint Solution*. 2009, DTIC Document.

29.   Webb, C.M. and K.J. Hewett, *An analysis of US Army fratricide incidents during the Global War on Terror (11 September 2001 to 31 March 2008)*. 2010, DTIC Document.

30.   Nordland, R., *U.S. and Afghan Troops Start Combat for Kandahar*, in *New York Times*. 2010: New York, N.Y.

31.   Tom Bownman, Q.L., *U.S. Thrust Has Unintended Effect On Kandahar City*, in *NPR*. 2010, NPR.

32.   Whittaker, G., *Asymmetric Wargaming: Toward A Game Theoretic Perspective*. 2000, DTIC Document.

33.   Firby, R.J. *An investigation into reactive planning in complex domains*. in *AAAI*. 1987.

34.   Lempert, R., et al., *Characterizing climate-change uncertainties for decision-makers. An editorial essay.* Climatic Change, 2004. **65**(1): p. 1-9.

35.   Holloway, C.A., *Decision making under uncertainty: models and choices*. 1979: Prentice-Hall Englewood Cliffs, NJ.

36.   Morgan, M.G., M. Henrion, and M. Small, *Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis*. 1992: Cambridge University Press.

37.   Haasnoot, M., et al., *Dynamic adaptive policy pathways: A method for crafting robust decisions for a deeply uncertain world.* Global Environmental Change, 2013. **23**(2): p. 485-498.

38.   van der Pas, J., et al., *Operationalizing adaptive policymaking.* Futures, 2013. **52**: p. 12-26.

39.   Dessai, S. and M. Hulme, *Assessing the robustness of adaptation decisions to climate change uncertainties: A case study on water resources management in the East of England.* Global Environmental Change, 2007. **17**(1): p. 59-72.

40.   Dessai, S. and J.P. Sluijs, *Uncertainty and climate change adaptation: A scoping study*. Vol. 2007. 2007: Copernicus Institute for Sustainable Development and Innovation, Department of Science Technology and Society Utrecht.

41.   Walker, W.E., S.A. Rahman, and J. Cave, *Adaptive policies, policy analysis, and policy-making.* European journal of operational Research, 2001. **128**(2): p. 282-289.

42.   Walker, W.E., et al., *Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support.* Integrated assessment, 2003. **4**(1): p. 5-17.

43.   Walker, W.E., V.A. Marchau, and D. Swanson, *Addressing deep uncertainty using adaptive policies: Introduction to section 2.* Technological Forecasting and Social Change, 2010. **77**(6): p. 917-923.

44.   McCray, L.E., K.A. Oye, and A.C. Petersen, *Planned adaptation in risk regulation: An initial survey of US environmental, health, and safety regulation.* Technological Forecasting and Social Change, 2010. **77**(6): p. 951-959.

45.   Wardekker, J.A., et al., *Operationalising a resilience approach to adapting an urban delta to uncertain climate changes.* Technological Forecasting and Social Change, 2010. **77**(6): p. 987-998.

46.   Lempert, R.J. and D.G. Groves, *Identifying and evaluating robust adaptive policy responses to climate change for water management agencies in the American west.* Technological Forecasting and Social Change, 2010. **77**(6): p. 960-974.

47.   Moench, M., *Responding to climate and other change processes in complex contexts:: Challenges facing development of adaptive policy frameworks in the Ganga Basin.* Technological Forecasting and Social Change, 2010. **77**(6): p. 975-986.

48.   Eriksson, E.A. and K.M. Weber, *Adaptive foresight: navigating the complex landscape of policy strategies.* Technological Forecasting and Social Change, 2008. **75**(4): p. 462-482.

49.   Marchau, V.A., W.E. Walker, and G. Van Wee, *Dynamic adaptive transport policies for handling deep uncertainty.* Technological forecasting and social change, 2010. **77**(6): p. 940-950.

50.   *The Operations Process.* 2010, 2010-03-26.

51.   Smith, R., *The long history of gaming in military training.* Simulation & Gaming, 2009.

52.   Smith, J.A., et al., *An army-centric system of systems analysis (SoSA) definition*, in *Arl-Tr 5446.* 2011, Army Research Laboratory: White Sands Missile Range, NM.

53.   Roessler, R.R., *SSC SAN DIEGO.* 2003.

54.   Joint Staff, J. and V. Suffolk, *Commander's Handbook for Assessment Planning and Execution.* 2011.

55.   Domercant, J.C. and D. Mavris, *Understanding and Evaluating Command & Control Effectiveness by Measuring Battlespace Awareness.* 2014, DTIC Document.

56.   Alberts, D.S. and R.E. Hayes, *Understanding command and control.* 2006, DTIC Document.

57.   Perry, W., et al., *Measurements of Effectiveness for the Information-Age Navy: The Effects of Network-Centric Operations on Combat Outcomes.* 2002, DTIC Document.

58.   Davis, J.K., et al., *Air/missile defense, counterproliferation and security policy planning : implications for collaboration between the United States and the Gulf Co-operation Council countries.* 1999, Abu Dahbi, United Arab Emirates: Emirates Center for Strategic Studies and Research. xvi, 144 p.

59.   Cares, J.R., *An information age combat model.* Produced for the United States Office of the Secretary of Defense, USA, 2004.

60.   Corps, U.M., *Marine Corps Planning Process.* 2010, Marine Corps Warfighting Publication (MCWP).

61.   Atwell, R.J. and D.G. McBryde, *Theater-Level Ground Combat Analyses and the TACWAR Submodels*. 1991, DTIC Document.

62.   Hernandez, R.M., *Defining a simulation capability hierarchy for the modeling of a seabase enabler (SBE)*. 2010, DTIC Document.

63.   Blais, C. and J. Ruck, *Connecting Simkit, CombatXXI, and Naval Simulation System (NSS) via XMSF Web Services*. 2004.

64.   Mavris, D.N., A.P. Baker, and D.P. Schrage, *IPPD through robust design simulation for an affordable short haul civil tiltrotor*. 1997.

65.   Weill, P., M. Subramani, and M. Broadbent, *Building IT infrastructure for strategic agility*. Sloan management review, 2002. **44**.

66.   Dekker, A.H., *Measuring the agility of networked military forces*. 2006.

67.   Klein, G.A., R. Calderwood, and A. Clinton-Cirocco. *Rapid decision making on the fire ground*. in *Proceedings of the Human Factors and Ergonomics Society annual meeting*. 1986. SAGE Publications.

68.   O'Donnell, D., et al., *The tyranny of the Balanced Scorecard in the innovation economy*. Journal of Intellectual Capital, 2006. **7**(1): p. 43-60.

69.   Davis, P.K., R.D. Shaver, and J. Beck, *Portfolio-analysis methods for assessing capability options*. Vol. 655. 2008: Rand Corporation.

70.   Agoglia, J., M. Dziedzic, and B. Sotirin, *Measuring Progress in Conflict Environments*. 2010.

71.   STAFF, J., *UNIVERSAL JOINT TASK MANUAL*. 2011.

72.   Schlabach, J.L., C.C. Hayes, and D.E. Goldberg, *FOX-GA: a genetic algorithm for generating and analyzing battlefield courses of action*. Evolutionary Computation, 1999. **7**(1): p. 45-68.

73.   Gonzales, D.R., *Evolution of the Air Campaign Planning Process and the Contingency Theater Automated Planning System (CTAPS)*. 1996, DTIC Document.

74.   Browne, C.B., et al., *A survey of monte carlo tree search methods*. IEEE Transactions on Computational Intelligence and AI in Games, 2012. **4**(1): p. 1-43.

75.   Aldridge, I., *High-Frequency Trading: A Practical Guide to Algorithmic Strategies and Trading Systems*. 2013: John Wiley & Sons.

76.   Romanycia, M.H. and F.J. Pelletier, *What is a heuristic?* Computational Intelligence, 1985. **1**(1): p. 47-58.

77.   Jaiswal, N.K., *Military operations research: Quantitative decision making*. Vol. 5. 2012: Springer Science & Business Media.

78.   *Strategy and force planning*. 4th ed. ed, ed. R.M. Lloyd and T.N. Castle. 2004, Naval War College Press: Newport, RI.

79.   Borger, J., *Wakeup call.* Guardian Weekly. Sydney, 2002. **20**.

80.   Gerecke, W., D. Enas, and S. Gottschlich, *A Distributed Collaborative Workflow Based Approach To Data Collection and Analysis*. 2004, DTIC Document.

81.   Brutzman, D., A. Buss, and C. Blais, *Connecting Simkit Discrete Event Simulation (DES) and the Naval Simulation System (NSS) via Web Services for Extensible Modeling & Simulation (XMSF)-Capable Analysis.* Naval Postgraduate School, 2004. **27**.

82.   Bourassa, N.R., *Modeling and simulation of fleet air defense systems using EADSIM*. 1993, Monterey, California. Naval Postgraduate School.

83.   Stone III, G.F. and G.A. McIntyre. *JWARS: the joint warfare system (JWARS): a modeling and analysis tool for the defense department*. in *Proceedings of the 33nd conference on Winter simulation*. 2001. IEEE Computer Society.

84.   Cil, I. and M. Mala, *A multi-agent architecture for modelling and simulation of small military unit combat in asymmetric warfare.* Expert Systems with Applications, 2010. **37**(2): p. 1331-1343.

85.   Moffat, J., *Complexity theory and network centric warfare*. 2010: DIANE Publishing.

86.   Dreyer, P. and P.K. Davis, *A Portfolio-analysis Tool for Missile Defense (PAT-MD): Methodology and User's Manual*. 2005, DTIC Document.

87.   Antunes, C.M. and A.L. Oliveira. *Temporal data mining: An overview*. in *KDD workshop on temporal data mining*. 2001.

88.   Drachen, A., et al., *Game data mining*, in *Game Analytics*. 2013, Springer. p. 205-253.

89.   Alpaydin, E., *Introduction to machine learning*. 2014: MIT press.

90.   Fu, T.-c., *A review on time series data mining.* Engineering Applications of Artificial Intelligence, 2011. **24**(1): p. 164-181.

91.   Rushing, Ś., et al., *Augmenting Wargame AI with Data Mining Technology*. 2004.

92.   Yannakakis, G.N. and J. Togelius, *A panorama of artificial and computational intelligence in games.* 2014.

311

93.   Madeira, C.A., V. Corruble, and G. Ramalho. *Designing a Reinforcement Learning-based Adaptive AI for Large-Scale Strategy Games*. in *AIIDE*. 2006.

94.   Esling, P. and C. Agon, *Time-series data mining*. ACM Computing Surveys (CSUR), 2012. **45**(1): p. 12.

95.   Drachen, A., et al. *Guns, swords and data: Clustering of player behavior in computer games in the wild*. in *Computational Intelligence and Games (CIG), 2012 IEEE Conference on*. 2012. IEEE.

96.   Drachen, A., et al., *A comparison of methods for player clustering via behavioral telemetry*. arXiv preprint arXiv:1407.3950, 2014.

97.   Missura, O. and T. Gärtner. *Player modeling for intelligent difficulty adjustment*. in *Discovery Science*. 2009. Springer.

98.   Weber, B.G. and M. Mateas. *A data mining approach to strategy prediction*. in *Computational Intelligence and Games, 2009. CIG 2009. IEEE Symposium on*. 2009. IEEE.

99.   Yannakakis, G.N. *Game AI revisited*. in *Proceedings of the 9th conference on Computing Frontiers*. 2012. ACM.

100.   Bauckhage, C., A. Drachen, and R. Sifa, *Clustering Game Behavior Data.* 2014.

101.   Liao, T.W., *Clustering of time series data—a survey*. Pattern recognition, 2005. **38**(11): p. 1857-1874.

102.   Berkhin, P., *A survey of clustering data mining techniques*, in *Grouping multidimensional data*. 2006, Springer. p. 25-71.

103.   Bouveyron, C. and C. Brunet-Saumard, *Model-based clustering of high-dimensional data: A review*. Computational Statistics & Data Analysis, 2014. **71**: p. 52-78.

104.   Liao, T., et al. *Understanding and projecting the battle state*. in *23rd Army Science Conference, Orlando, FL*. 2002.

105.   Aghabozorgi, S., A.S. Shirkhorshidi, and T.Y. Wah, *Time-series clustering–A decade review*. Information Systems, 2015.

106.   Xu, R. and D. Wunsch, *Clustering*. Vol. 10. 2008: John Wiley & Sons.

107.   Oates, T., M.D. Schmill, and P.R. Cohen. *A method for clustering the experiences of a mobile robot that accords with human judgments*. in *AAAI/IAAI*. 2000.

108.   Vlachos, M., G. Kollios, and D. Gunopulos. *Discovering similar multidimensional trajectories*. in *Data Engineering, 2002. Proceedings. 18th International Conference on*. 2002. IEEE.

312

109. Senin, P., *Dynamic time warping algorithm review*. University of Hawaii, 2008.

110. Müller, M., *Dynamic time warping*. Information retrieval for music and motion, 2007: p. 69-84.

111. Murtagh, F. and P. Contreras, *Methods of hierarchical clustering*. arXiv preprint arXiv:1105.0121, 2011.

112. Jain, A.K., M.N. Murty, and P.J. Flynn, *Data clustering: a review*. ACM computing surveys (CSUR), 1999. **31**(3): p. 264-323.

113. Carteron, A., et al., *Assessing the efficiency of clustering algorithms and goodness-of-fit measures using phytoplankton field data*. Ecological Informatics, 2012. **9**: p. 64-68.

114. Shannon, W., R. Culverhouse, and J. Duncan, *Analyzing microarray data using cluster analysis*. Pharmacogenomics, 2003. **4**(1): p. 41-52.

115. El-Hamdouchi, A. and P. Willett, *Comparison of hierarchic agglomerative clustering methods for document retrieval*. The Computer Journal, 1989. **32**(3): p. 220-227.

116. Punj, G. and D.W. Stewart, *Cluster analysis in marketing research: review and suggestions for application*. Journal of marketing research, 1983: p. 134-148.

117. Milligan, G.W., *Ultrametric hierarchical clustering algorithms*. Psychometrika, 1979. **44**(3): p. 343-346.

118. Fernández, A. and S. Gómez, *Solving non-uniqueness in agglomerative hierarchical clustering using multidendrograms*. Journal of Classification, 2008. **25**(1): p. 43-65.

119. Gan, G., *Data clustering : theory, algorithms, and applications*. ASA-SIAM series on statistics and applied probability ;, ed. C. Ma and J. Wu. 2007, SIAM Society for Industrial and Applied Mathematics ;: Philadelphia, Pa.

120. Aggarwal, C.C., *Data Clustering Algorithms and Applications*. Chapman & Hall/CRC data mining and knowledge discovery series. 2013, CRC Press: Hoboken, NJ.

121. Liu, Y., et al., *Understanding and enhancement of internal clustering validation measures*. Cybernetics, IEEE Transactions on, 2013. **43**(3): p. 982-994.

122. Liu, Y., et al. *Understanding of internal clustering validation measures*. in *Data Mining (ICDM), 2010 IEEE 10th International Conference on*. 2010. IEEE.

123. Arbelaitz, O., et al., *An extensive comparative study of cluster validity indices*. Pattern Recognition, 2013. **46**(1): p. 243-256.

124. Hartigan, J.A. and M.A. Wong, *Algorithm AS 136: A k-means clustering algorithm.* Journal of the Royal Statistical Society. Series C (Applied Statistics), 1979. **28**(1): p. 100-108.

125. Dan Sears, C.S., Tim Mann, *WinBoard*, T. Mann, Editor. 2014, GNU: GNU Operating System.

126. Heusden, F.v., *POS - An experimental chess program*, F.v. Heusden, Editor. 2008: https://www.vanheusden.com/.

127. Dembélé, D. and P. Kastner, *Fuzzy C-means method for clustering microarray data.* bioinformatics, 2003. **19**(8): p. 973-980.

128. Schaeffer, J., M. Müller, and A. Kishimoto, *Go-bot, go.* Spectrum, IEEE, 2014. **51**(7): p. 48-53.

129. Levy, D., *Computer chess compendium*. 2013: Springer Science & Business Media.

130. Lauren, M. and R. Stephen, *Map-aware non-uniform automata (MANA)-a new zealand approach to scenario modelling.*

131. Anderson, M.A., et al., *MANA (Map Aware Non-Uniform Automata), Version 4 User Manual*. 2007, New Zealand: Operations Analysis Section, Defense Technology Agency.

132. Giorgino, T., P. Tormene, and M.T. Giorgino, *The dtw Package.* 2007.

133. Kadish, S. *Utilization of Squares by Chess Masters*. 2014 [cited 2016 06/02/2016]; Available from: http://vizual-statistix.tumblr.com/post/78821780083/.

134. Lasker, E. and J. Du Mont, *Chess strategy*. 1921, London, New York: G. Bell and sons ltd.; Brentano; [etc., etc.]. xiii, 282 p.

135. Desktop, E.A., *Release 10.* Redlands, CA: Environmental Systems Research Institute, 2011.