Opponent Modeling of the Taliban in Afghanistan's Emerging Peace Talks

Leonard Klenner
Johns Hopkins SAIS
Washington, DC
lklenne1@jhu.edu

Abstract

Terrorist groups can be modeled as rational actors that adapt the structure of their networks to perform reward maximizing actions in a given environment. In the current talk-and-fight environment of Afghanistan, the Taliban need to adapt their network to two domains, negotiations and combat. Opponent modeling can be used to understand how the Taliban concurrently optimize the structure of their network across both domains. Conceptually, it is first shown that negotiations have a different reward distribution – which yields a different optimal structure of the Taliban network – than combat. Empirically, it is validated that relevant changes in the structure of the network of the Taliban can be inferred from 2017-2018 data on Taliban attacks and Coalition airstrikes. Two alternative cases of a best response to these structural changes are provided that enable focused diplomatic engagement of the Taliban in the emerging Afghan peace talks.

1 Introduction

To counter a terrorist group, we first need to understand what actions the group is likely to perform and what broader strategy it follows in performing these actions. This is a non-trivial problem as the structure of terrorist groups is generally fuzzy – we may not know who the groups' members are and how they are organized. Even if these parameters are known, the search space of possible actions for the group remains vast – should we prepare to counter a shooting, an ambush, a hostage taking, a suicide attack, an insider attack, a combination of these; when and where?

In most cases, no direct answer to these questions can be found. A feasible means to reduce the number of options is to identify what the group would optimally choose to do, given the constraints of the environment that it operates in. Modeling the optimal choice of an opponent can be summarized under the term *opponent modeling* (see 1.2 for details on relevant research).

Opponent modeling comes in multiple conceptual implementations all of which assume that terrorists are rational actors, defined as choosing the action that maximizes their rewards. In an opponent model, a terrorist group optimizes available parameters – its command-and-control structure, its resource allocation, its recruiting etc. – to perform an action that at time $\mathbf{t_1}$ yields the highest known reward.

Up until late 2017, an opponent model of the Taliban would have been complete by focusing on how the group optimizes its parameters to maximize rewards in combat.

Final course paper for *International Bargaining and Negotiation*, taught by Professor Sinisa Vukovic Dep. of Conflict Management, Fall 2019.

With the emergence of backchannels between the Taliban's Qatar Office and U.S. diplomats, an opponent model of the groups needs to optimize for both domains, combat and negotiations.

The opponent model proposed here aims to identify (1) the optimal structure of the Taliban network in each domain and (2), by approximating observed behavior of the Taliban to these optima, which domain appears to yield the relatively higher rewards for the Taliban. While it seems plausible that the Taliban optimize rewards across both domains, we can assume that at some points rewards in one domain will exceed those in the other domain. A prediction that rewards in negotiations will exceed those in combat can be assumed to have started the current backchannel talks. Hence, it is important to understand when the distribution of rewards changes across domains and – following the above prediction – the Taliban become fully committed to negotiations or resort to continuing the conflict militarily.

Based on data on Taliban attacks and Coalition airstrikes against Taliban leadership for 2017-2018, it is possible to validate changes in the reward maximization of the Taliban. However, the data cannot be reduced to either a full domain-shift to combat or negotiations. This aligns with our intuition about the conflict, given the early and ambiguous stage of the peace talks. While Afghanistan is still covered in fog of war, the fog of peacemaking is settling in, removing established coordinates from the landscape and demanding that parties reroute their problem solving from the trodden paths of the 'longest war'. To provide diplomats with a heuristic in conceptualizing their best response to the Taliban's best response in the changing environment of Afghanistan, two hypothetical alternative cases of full domain shifts of the Taliban's reward distribution are proposed.

The next sub-section introduces the logic of the approach taken for the opponent model.

1.1 Background

For the purpose of the opponent model, negotiations and combat are treated as problems of decisionmaking under uncertainty. The goal of the Taliban is to reduce this uncertainty. It is assumed that low uncertainty facilitates performing a reward-maximizing action, while high uncertainty leads to inferior decisions.

As a result, the focus of the model is not on what constitutes a reward maximizing action for the Taliban – should a specific attack be carried out or demand be made in negotiations? – but on the optimal means to reduce uncertainty. The sole parameter that the model optimizes is the structure of the Taliban network. Structure is defined as the configuration of individual nodes inside the network with respect to a common end, i.e. launching an attack or starting peace talks.

In the model, the Taliban can achieve the common end only if they have the right information management in place to reduce the uncertainty of actions that are directly related to achieving the common end. To give an example, *common end*: ambush a convoy → *action*: send a group of fighters → *information management*: give fighters autonomy over when and where to attack or give them specific and defining directions.

To simplify the information management of the Taliban network, we divide the network into one global decisionmaker – the information demand – and a large set of local members – the information supply. To optimally reduce uncertainty in an environment, information supply has to meet information demand. The optimal equilibrium of supply and demand changes based on different ends. As will be shown later, the uncertainty in negotiations is different from the uncertainty in combat. Therefore, each of the domains requires a different 'right' information management for its respective ends, such as ambushing a convoy or opening a backchannel. From what has been established, we see that changes in equilibrium demand and supply to provide optimal information management represent changes in the configuration of nodes inside the Taliban network. Therefore, these changes represent changes in the structure of the network itself.

Different structures of terrorist groups yield different patterns of actions. To give an example, a network that is structured into many cells with low overall connectivity often produces simple attack patterns for each cell. The network structure then acts as a force multiplier by scaling the cells into a swarm of independent attack points (Arquilla and Ronfeldt 1999). Given this connection between structure and attacks, we can use data on Taliban attacks to approximate the structure of the Taliban network.

Recall that network structure is assumed to be a product of information demand and supply. Demand and supply adjust based on the optimal equilibrium for reducing the uncertainty of a given reward maximizing action. Importantly, the type of uncertainty changes based on the domain in which rewards are maximized. Hence, a specific network structure as it becomes manifest in an attack is indicative of the domain that is reward maximizing – even if an attack is carried out, we can infer whether or not a domain shift to negotiations would yield higher rewards for the network structure that underlies the attack. Therefore, through the steps outlined above, we are able to derive an approximate reward distribution of the Taliban across domains based on publicly available attack data. With this information, diplomats can develop an optimal counter-strategy and, importantly, adapt it to changes in the Taliban's reward maximization.

We can re-state the logic of the opponent model as follows: observed actions → network structure → information equilibrium → type of uncertainty managed → domain which maximizes the Taliban's rewards and defines their optimal behavior.

Section 2 provides a toy example of the opponent model, section 3 formally introduces the model, section 4 validates the model based on data on Taliban attacks and Coalition counter-attacks for 2017-2018 and section 5 discusses two alternative cases of domain-shifts of the Taliban's reward distribution.

1.2 Related Work

Opponent modeling. Opponent modeling is formally introduced in Carmel and Markovitch (1996) as method to classify the optimal adaptation to opponents' strategies in multi-agent systems. Schadd et al. (2007) discuss reduction of uncertainty in opponent modeling through building hierarchies of models. Enders and Su (2007), Braynov (2009), and Rios et al. (2015) apply game-theoretical opponent models to terrorist networks. The present opponent model is most similar to Enders and Su (2007) in that it focuses on optimizing network structure as a means to perform actions that maximize rewards.

Inverse optimization. Problems in which behavior is observed but the reward distribution underlying this behavior is unknown are studied in inverse optimization, which aims to infer the reward distribution of a given optimal behavioral pattern. Prominent implementations of inverse optimization are Abeel and Ng (2001), Kuleshov and Schrjivers (2014) and Molloy et al. (2017).

Inverse optimization is highly formal and generally studied on constrained problems. To ensure that the present model generalizes to the complex Afghan conflict, inverse optimization is translated into a less formal approach for deriving the reward distribution of the Taliban attack behavior.

Nonlethal targeting of insurgents. The problem of optimally influencing local populations in counter-insurgency operations is known as the nonlethal target assignment problem – whom is it best to reach out to in a village for maximizing influence over its members? Hung (2010), and Hung et al. (2013) propose generative optimization models that yield an optimal allocation of counter-insurgency assets to nonlethal targets among the local population. The present model inverts this approach and assumes that the parallel problem of assigning information supply to demand exists inside the Taliban network, which adapts its structure to ensure that its assets are optimally allocated to match information demand with supply. Once it is identified who in the Taliban network is optimally talking to whom, diplomats can tailor their outreach to maximize influence over the Taliban network.

2 Toy Example

Consider the current leader of the Taliban, Mawlawi Hibatullah Akhundzada (MHA)'s task to make informed decision about combat operations and negotiations. For both domains, MHA needs to establish channels to acquire information from the wider Taliban network. We assume that the knowledge about the world that MHA can provide by himself does not sufficiently reduce the uncertainty of the environments that his decisions will be carried out in. Hence, MHA needs to establish paths within the Taliban network that optimally supply his information demand and reduce uncertainty. Paths refer to the configuration of nodes inside the Taliban network that MHA includes in his decisionmaking loops and which shape the optimal structure of the Taliban network.

The graph in Figure 1 illustrates a simple set up of two decision loops of MHA for combat and negotiations. MHA has explored the network up unto nodes **X** and **Y**, which are each included in one loop respectively and which provide information to MHA. Black arrows represent exploration, blue arrows represent information provision. Dotted lines represent nodes accessible via **X** and **Y**, which can be explored in subsequent cycles.

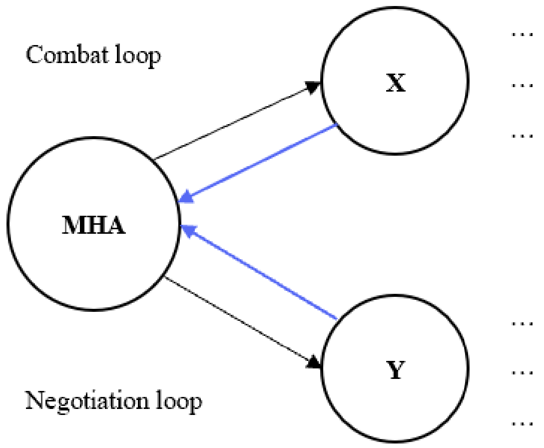


Figure 1: One-cycle decision loops for combat and negotiations.

Assuming multiple cycles, the operations available to MHA are adding additional nodes to each loop, receiving and evaluating their information supply and then either deciding to end a loop and carry out the associated decision or deciding to either add more nodes to the loop to increase knowledge or remove nodes that are degrading knowledge. Importantly, more nodes do not necessarily result in better knowledge about the world and can add redundancy that slows down information transfer on the network's active paths. As will be discussed in the next section, redundancy in the Taliban network can be valuable in combat but does not provide benefits for negotiations.

As we assume that uncertainty in combat is different from uncertainty in negotiations, the optimal paths for each loop will be different and will lead to different types of information management to reduce uncertainty. We will differentiate between *coordination* and *planning* as means to reduce uncertainty in carrying out decisions. Combat decisionmaking requires coordination, while negotiation decisionmaking requires planning.

The toy example shows the basic setup of the opponent model.

We assume that information demand – in the example, MHA – explores the Taliban network – in the example, nodes X and Y – to identify the optimal supply of information that reduces the uncertainty of a reward maximizing action.

Each domain establishes different equilibria of information demand and supply, and therefore yields different information foraging paths by MHA. These paths shape the optimal structure of the Taliban network, relative to the action that MHA aims to perform and whose uncertainty he aims to reduce.

Based on the distinction between coordination and planning, the validation part of the model will approximate which network structures underly current attack patterns of the Taliban. This in turn allows us to approximate the reward distribution of the Taliban across domains and whether rewards in one domain exceed those in the other domain.

3 Method

The central assumption in our opponent model concerns different types of uncertainty for combat and negotiations. We define uncertainty as uncertainty of the decisionmaker about the rewards of her optimal action. Thus, uncertainty can be expressed in the following question: How to know what rewards an action in either domain will yield?

Consequently, we derive different forms of uncertainty for each domain based on the domains' different *reward architecture*, which we define in the next sub-section.

3.1 Reward Architecture

We define as reward architecture of a domain the collection of the domain's reward values, win-set, reward distribution and reward signal. We discuss each of these components for the reward architectures of combat and negotiations in turn.

Reward values. Reward values refer to how valuable the reward of action is on average. We use reward values as a measure to compare rewards across domains. We assume that an average action in combat yields relatively low rewards compared to an average action in negotiations. This is intuitive as actions in negotiations have a higher level of generality – negotiating the removal of American troops from Afghanistan represents a larger shift that cannot be achieved by military means, say, by launching attacks on army bases.

Win-set. The notion of a win-set features prominently in two-level game theory (Putnam 1988) as a measure of the robustness of a party's rewards relative to a proposed action. We specify this notion of reward robustness to obtain the *margin of partial rewards* of an action. This margin represents the range across which parts of the reward of an optimal action can be obtained for different, sub-optimal performances of the action. To give an example, if an attack on an army base results only in the *partial* destruction of the base, does the attacker receive partial rewards and if yes, in what proportion to the rewards for a complete destruction of the base? We assume that an average action in combat has a higher win-set than an average action in negotiations. The partial destruction of an army base is likely going to result in proportionally higher partial rewards than partial agreement on the negotiation of a ceasefire.

Reward distribution. Reward distribution refers to the allocation of rewards in an environment. We distinguish between sparse and dense reward distributions. A sparse distribution means it is comparatively hard to explore an action that yields rewards. A dense distribution means it is comparatively easy to explore an action that yields rewards. We assume the distribution of rewards in combat is relatively dense and that the distribution of rewards in negotiations is relatively sparse. Consider that in negotiations, rewards often exclusively emerge vertically from packaging multiple actions (Vukovic 2015), whereas in combat rewards emerge horizontally from individual actions, such as attacking a single army base.

Reward signal. Reward signal refers to the likelihood that the reward of an action will become invalid after the action has been taken. The invalidation of a reward is assumed to be binary – the reward stays constant or diminishes to zero. We distinguish between instable and stable reward signals. Instable reward signals represent a high likelihood of a reward becoming invalid, stable reward signals represent a low likelihood of a reward becoming invalid. We assume that the reward signal in combat is instable and that the reward signal in negotiations is stable. Consider that actions in combat are conducted under counter-attacks, which potentially disrupt the performance of an action. For the Taliban, these counter-attacks are Coalition airstrikes. We can model the impact of counter-attacks by way of the instable reward signal, where the disruption of an action is modeled as the reward of the action becoming invalid. In negotiations, spoilers present a similar potential for disruption. However, given the scale of Coalition airstrikes against the Taliban, we assume that the reward signal in negotiations is – comparatively – stable.

Based on this overview, we see that combat and negotiations have different reward architectures. The rewards in combat are low in value, high in win-set, dense in distribution and instable in signal. The rewards in negotiations are high in value, low in win-set, sparse in distribution and stable in signal. Based on their different reward architectures, each domain gives rise to a different type of uncertainty about what rewards an action might yield. This results in different optimal approaches to reducing uncertainty to maximize rewards. In combat this optimal approach is coordination, in negotiations it is planning.

3.2 Coordination and Planning

Coordination and planning are different approaches to information management. Each approach results in different information foraging paths inside the Taliban network, based on the underlying type of uncertainty, which the decisionmaker aims to reduce. Coordination and planning can be distinguished as horizontal and hierarchical forms of organizing the network. Coordination is horizontal, planning is hierarchical. Going back to the earlier example of the Taliban's ambush on a convoy, whereas coordination entails giving the group of fighters autonomy over when and where to attack, planning entails providing the group with specific and defining directions.

We discuss each approach in turn as it relates to combat and negotiations. As in the toy example, we assume the perspective of MHA, who constitutes information demand exploring the Taliban network for optimal information supply. For each approach we want to understand why it satisfies the equilibrium of information demand and supply of the respective domain.

Coordination. To maximize rewards in an environment with low reward values, high density of rewards, a high win-set and instable reward signals, MHA establishes the following policy for his decisionmaking: *low reward values* → rewards are maximized through multiple, horizontal actions; *high density of rewards* → actions can be simple in nature; *high win-set* → actions can be delegated to low level operators as the associated rewards are robust with respect to varying degrees of success; *instable reward signal* → building up an accurate understanding of the environment before performing an action is of limited value as the environment is likely to change based on enemy counter-attacks.

This policy determines how MHA sets his information demand in combat and, in turn, explores the Taliban network to obtain optimal information supply. Based on this policy, MHA will build a wide foraging path to include a large number of nodes in his decisionmaking loop, even if this configuration does not yield the most accurate level of information supply. Given the high win-set and instable reward signal, the uncertainty of this environment rests not on *how* an action will be performed but *that* it will be performed.

We define this type of information management policy as coordination. It prioritizes adaptation and on the ground learning, which entails shifting agency away from the central decisionmaker, MHA. Importantly, learning about the environment occurs largely 'out of the loop', once MHA has terminated the decisionmaking loop and the associated decision is carried out. While coordination results in actions that are less predictable from the perspective of MHA, it ensures that the informational process of the decisionmaking is sufficiently robust to withstand disruptions in the reward signal. Figure 2 gives a graphical representation of the information foraging path for coordination inside the network. The decisionmaker node, **A**, has explored the network up unto the row of nodes from **I** to **O**. The parameter ***k*** represents the knowledge of the world of **A** at a specific node. The value of ***k*** for the last node on a path is always representative of **A**'s knowledge about the world given all explored nodes (in Figure 2, ***k***=0.2 represents **A**'s knowledge based on nodes (**A**, …, **O**)). Arrows are directed upwards, representing information supply from the network to the decisionmaker.

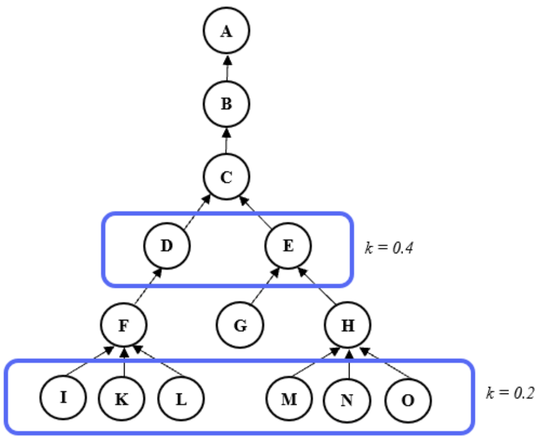


Figure 2. Foraging path for coordination. Values for knowledge parameter ***k*** are illustrative.

We assume that ***k***(**D**, **E**) > ***k***(**I**, …, **O**) as coordinated information management prioritizes the distribution of information supply in the network over the accuracy of the supplied information to decisionmaker, **A**. Although the decisionmaker's knowledge of the world at (**D**, **E**) is larger than at (**I**, …, **O**), she will not prune the path back to (**D**, **E**) as redundancy is needed to mitigate the instable reward signal and allows taking advantage of the dense reward distribution. Hence, we see that the underlying uncertainty managed through coordination cannot be reduced by building up a detailed and centralized control mechanism. Reducing the uncertainty requires that a substantial number of decentralized elements of the network are included in the decisionmaking loop of **A**. These elements perform their actions largely autonomously. From the perspective of the Taliban, the uncertainty managed through coordination represents the uncertainty of combat.

Planning: To maximize rewards in an environment with high reward values, low density of rewards, a low win-set and stable reward signals, MHA establishes the following policy for his decisionmaking: *high reward values* → reward maximization can be achieved with a smaller number of actions, *low density of rewards* → actions need to be sufficiently complex and are vertical, *low win-set* → actions need to be performed with strong strategic oversight as the rewards are not robust with respect to varying degrees of success, *stable reward signal* → building up an accurate understanding of the environment before performing an action is important and feasible as changes in the environment are likely to be limited.

This policy determines how MHA sets his information demand in negotiations and, in turn, explores the Taliban network to obtain optimal information supply. Based on this policy, MHA will build a narrow foraging path as the stable reward signal makes accurate information about the environment desirable and is required to perform vertical actions on sparse rewards. For a wide path, the high number of nodes in the decisionnaking loop could reduce the accuracy of information supply and would conflict with the low win-set, given that operations need to be carried out with strong strategic oversight of the decisionmaker as partial rewards are relatively small. Consequently, the uncertainty in this environment rests not on *that* an action will be performed but *how* it will be performed.

We define this type of information management policy as planning. It prioritizes the construction of vertical actions by way of centralized information foraging and thus shifts agency towards the decisionmaker, MHA. Importantly, learning about the environment occurs largely ‘in the loop’, before MHA has terminated the decisionmaking loop and the associated decision is carried out. Planning results in actions that are predictable from the perspective of MHA but the informational process of the decisionmaking is not robust enough to withstand disruptions in the reward signal or, based on the low win-set, autonomous decisionmaking by operators.

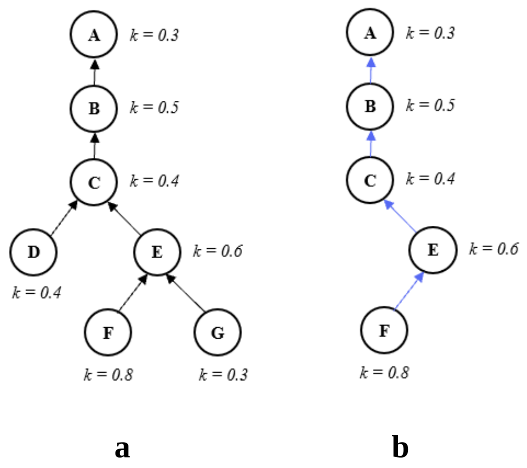


Figure 3. Foraging path for planning. Graph a represents the total network explored. Graph b represents the pruned network. Values for ***k*** are illustrative.

A graphical representation of the information foraging path for planning inside the network is given in Figure 3. Whereas coordination does not require pruning the network, planning entails pruning to remove redundancy and reduce the path to only those nodes which lead to the highest known ***k***.

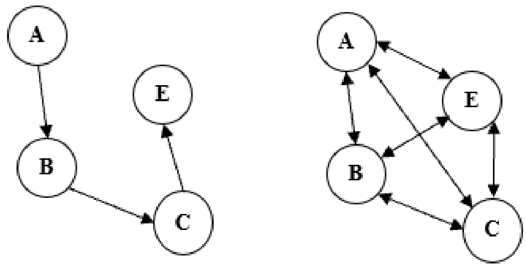
In Figure 3, we assume that the decisionmaker first explores the network (a) but before terminating the decisionmaking loop, the explored path is pruned (b). In (a), the decisionmaker explores the network on two paths. The deeper path extends unto nodes **F** and **G**. The shallower path unto extends unto node **D**. Since ***k***(**F**) > ***k***(**D**) > ***k***(**G**), **F** will be adopted whereas **D** and **G** will be pruned in (b) to obtain the narrow foraging path needed to perform vertical actions on sparse rewards with a low win-set. Hence, we see that the underlying uncertainty managed through planning cannot be reduced by distributing the decisionmaking across a decentralized network whose elements act largely autonomously. Reducing the uncertainty requires that a centralized and detailed control mechanism is built up in the decisionmaking loop of **A**. From the perspective of the Taliban, uncertainty managed through planning represents the uncertainty of negotiations.

In the next section, we address how the information foraging paths of coordination and planning shape the optimal network structure of the Taliban.

3.3 Network Structure

The information foraging paths for coordination and planning represent the optimal structure of the Taliban network with respect to the information demand of a decisionmaker. Therefore, the information foraging paths are optimal network structures only under the constraint of the information demand of the decisionmaker, an element of the network. By default, this constraint entails that the information foraging paths are not necessarily optimal for the information demand of the network itself, which would constitute the unconstrained optimal case.

It follows that if only one decisionmaker exists in the network, the constrained optimum is equal to the unconstrained optimum. This is the case for negotiations as planning entails that control rests with a single decisionmaker. As a result, the Taliban's optimal network structure for negotiations is identical to the information foraging path of planning, in which nodes are hierarchically connected.



a **b**

Figure 4. Optimal network structure of the Taliban Graph a represents the optimal structure for negotiations. Graph b represents the optimal structure for combat.

For coordination, the constrained optimum does not equal the unconstrained optimum as control is distributed across multiple decisionmakers. Following the established logic of reward maximization, each of these decisionmakers forages for information supply in the network. The resulting foraging paths present a series of local optima that are variations of the first foraging path, i.e. for the Taliban, the foraging path of their leader, MHA. The global optimum has to satisfy the information demand of all decisionmakers and therefore can be represented as the set of the decisionmakers' respective information foraging paths. Recall that these paths should overlap given their shared origin on the first information foraging path and deviate individually as each decisionmaker has slight variations in information demand leading to different patterns of exploration of the network. Therefore, as the set of the decisionmakers' information foraging paths, the optimal structure of the Taliban network for combat is decentralized, based on the deviations from the first foraging path, and fully connected, based on the overlap of multiple foraging paths.

4 Analysis

To understand whether the Taliban's reward maximization has changed in response to the emergence of negotiations in Afghanistan, we analyze data on Taliban attacks and Coalition airstrikes against Taliban leadership for 2017-2018 available through the ESRI Terrorist Attacks dataset (ESRI 2018) and The Bureau of Investigative Journalism Drone Warfare dataset (TBIJ 2018).

4.1 Categories

For the Taliban, we focus on *equilibrium attacks*. An equilibrium attack takes place when information demand and supply inside the network have been matched to enable the correspondence of multiple parts in a unified attack. Examples of equilibrium attacks are large-scale attacks on army bases or parallel attacks on multiple checkpoints or insider attacks on Coalition leadership. Importantly, equilibrium attacks encompass both equilibria established by coordination and planning.

For the Coalition airstrikes, we focus on *offset strikes*. Offset strikes take place when tactical or strategic leadership of the Taliban are killed, thereby offsetting the capabilities of the Taliban network to perform equilibrium attacks. An example of a tactical+ offset strike is a strike that kills a Taliban commander or higher, an example of a strategic offset strike is a strike that kills a Taliban senior commander or higher. Recall that the Coalition airstrikes are modeled in the Taliban's reward architecture as the reward signal, which becomes more unstable as offset strikes increase and becomes less unstable as offset strikes decrease.

4.2 Coding

The coding of the data was performed manually in one iteration. The labeled dataset is available online.[1]

In the ESRI dataset, Taliban attacks are identified equilibrium attacks if the provided description of the attacks is indicative of sufficient complexity, such as multiple parallel attacks or assassinations that suggest deep infiltration of Coalition forces and result in senior-level casualties. Based on 381 attacks in the dataset, 21 are identified as equilibrium attacks, or approximately 5.5% of the total.

In the TBIJ dataset, Coalition airstrikes are identified as tactical+ if regional leadership of the Taliban or higher are killed, and as strategic if supra-regional leadership of the Taliban are killed. For an overview of the Taliban leadership structure, see Afsar et al. (2008) and Giustozzi (2017). Importantly, the TBIJ dataset distinguishes between strikes confirmed by U.S. sources and strikes confirmed only by non-U.S. sources. Here, this distinction is suspended. Both types of strikes are included in the model. In addition, we obtain the number of total weapons released for Coalition airstrikes from the TBIJ dataset. This provides a measure of the overall intensity of the Coalition air campaign against the Taliban.

Note that there are several constraints on coding the data. First, the descriptions available on the Taliban attacks are often ambiguous with respect to the complexity

[1] Complete dataset at `https://git.io/fhLtZ`

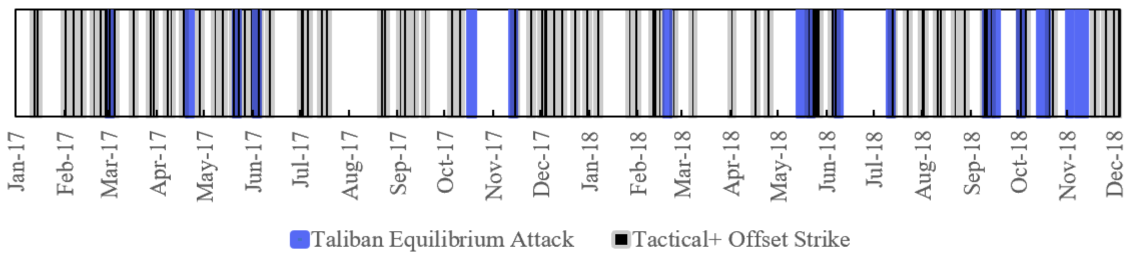


Figure 4. Distribution of Taliban Equilibrium Attacks and Tactial+ Offset Coalition Airstrikes

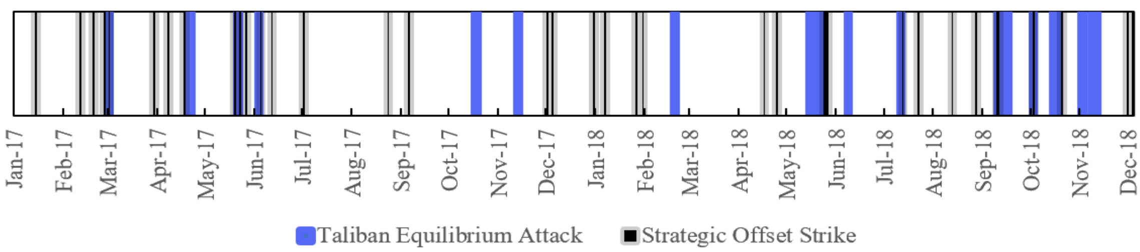


Figure 5. Distribution of Taliban Equilibrium Attacks and Strategic Offset Coalition Airstrikes

of the attack. Second, the Taliban organizational structure is not perfectly understood and hence the impact of a commander's death or the death of a related rank on mission capabilities remains ambiguous. In addition, the data on Taliban attacks and Coalition airstrikes that is publicly available is likely not fully complete. As a result of these constraints, the coding should be seen as approximative.

4.3 Results

The results of the analysis are given in Table 1. In line with news reporting on the Afghan conflict in late 2018 (Philips 2018), we observe an intensification of the Coalition air campaign against the Taliban taking place year-on-year for 2017 and 2018. Total weapons released (TWR) of all Coalition strikes increases by 61.08% from 4361 in 2017 to 7025 in 2018.

Concurrently, tactical+ offset strikes decrease by 22.92% from 48 in 2017 to 37 in 2018. The subset of strategic offset strikes remains approximately constant, decreasing by 0.05% from 19 in 2017 to 18 in 2018. Thus, the intensification of the air campaign does not show with respect to tactical or strategic leadership of the Taliban killed in strikes. Instead, the trend becomes inverted on the tactical level.

	2017	2018	Pct. Change
Coalition airstrikes			
TWR	4361	7025*	+61.08%
Tactical+	48	37**	-22.92%
Strategic	19	18**	-0.05%
Taliban attacks			
Total	140	241*	+72.14%
Equilibrium	7	14**	+100.00%

Table 1. Data on Coalition airstrikes against the Taliban and Taliban attacks for 2017 and 2018. TWR refers to total weapons released for all strikes in that year. (*) indicates that the data is extrapolated based on previous months' average. For TWR, we extrapolate for the last three months of 2018. For total Taliban attacks, we extrapolate for the last month of 2018. (**) indicates that the data is capped at 12/06/18.

For the Taliban, the total number of attacks conducted increases by 72.14% from 140 in

2017 to 241 in 2018. The subset of equilibrium attacks increases by 100% from 7 in 2017 to 14 in 2018. Thus, the Taliban reciprocate by intensifying their attacks against Coalition forces. Importantly, the increases in equilibrium attacks exceed those in total attacks by a significant margin.

4.4 Visualization

We visualize the data on Coalition airstrikes and Taliban equilibrium attacks in Figure 4 and 5 to gain insights into the monthly distribution of the data. We observe

The graphs show that the density of tactical+ strikes is substantially higher in the first half of 2017 than in the first half of 2018. The second halves of both years remain roughly equal in density of tactical+ strikes.

We observe that the increase in Taliban equilibrium attacks occurs predominantly in the last quarter of 2018. Similarly, tactical+ offset Coalition airstrikes enjoy the highest density for 2018 in the last months of the year. The distribution of strategic offset strikes remains roughly constant for both years, with the highest density in the first quarters of 2017 and 2018.

Overall, the graphs suggest that as the parties' involvement in negotiations has firmed up in late 2018, combat between them has intensified.

5 Discussion

While the coding remains approximative, two trends can be extracted from the data, which can be linked to changes in the reward architecture of the Taliban and hence to an optimal adaptation of the structure of the Taliban network. The first trend concerns the decrease in tactical+ offset Coalition airstrikes with near constant strategic offset airstrikes. The second trend concerns the over proportional increase in the Taliban's equilibrium attacks. Based on the logic of the opponent model, these trends can be translated into two different scenarios of developments in the Taliban network.

Scenario 1. Following reductions in strikes against their tactical leadership, the Taliban attain a higher equilibrium of information demand and supply in combat as connections between decisionmakers increase and improve coordination between them. In the opponent model, the 22.92% decrease in tactical+ offset strikes with near constant strategic offset strikes means that the reward signal of the Taliban becomes more stable for tactical leadership, facilitating their information foraging, which in turn leads to a more connected network at sub-strategic levels. A more deeply connected network structure makes coordination more effective, leading to a higher reward maximization in combat relative to negotiations. Given the near constant strategic strikes, planning, which requires concentration of agency at the strategic level, becomes relatively less attractive for reward maximization. In this scenario, the 100% increase in equilibrium attacks is accounted for exclusively by higher levels of coordination.

Scenario 2. Following the emergence of negotiations, the Taliban shift their reward maximization from combat to negotiations. As a result, the information management switches from coordination to planning, resulting in a more hierarchically organized network with greater capabilities for conducting equilibrium attacks. In this scenario, the 100% increase in equilibrium attacks is accounted for predominately by planning, with the 22.92% decrease in tactical+ offset strikes partially augmenting this trend through higher coordination among the parts of the network that are pruned from the foraging path of the decisionmaking loop of MHA.

Both of the scenarios align with the logic of the opponent model. Although, Scenario 1 appears to have deeper integration with the model's parameters, Scenario 2 cannot be discounted as the equilibrium attacks could be *planned* not *coordinated*.

The current reality of the talk-and-fight environment of the Afghan conflict may not yet have evolved into either Scenario 1 or 2 and indeed it seems reasonable to assume that a hybrid scenario has the best fit for the current stage of the conflict. However, hybrid scenarios are difficult to conceptualize in terms of a best response of Western diplomats conducting negotiations with elements of a hybrid optimal Taliban network structure. To facilitate the conceptualization of a best response to the Taliban's best response in the conflict, we can return to the two scenarios.

Scenario 1 and 2 constitute two different full domain shifts in the Taliban's reward architecture with clear consequences for the optimal structure of the network and hence clear options for Western diplomats to optimally engage the Taliban.

In Scenario 1, the Taliban fully shift their reward maximization to combat. This entails a weaker strategic level of leadership and a strengthened tactical level. In Scenario 1, Western diplomats need to expand their outreach away from MHA and his deputies towards commanders and other tactical leaders in the Taliban. Backchannels with the strategic level alone do not result in enough leverage over the network itself. At this stage, public diplomacy and scaled numbers of diplomatic personnel for local solution building become necessary.

In Scenario 2, the Taliban fully shift their reward maximization to negotiations. This entails a stronger strategic level of leadership and a weaker tactical level. In Scenario 2, Western diplomats can continue to focus on backchannel engagement with the strategic level of the Taliban. Importantly, the model shows that this shift to negotiations is accompanied by an increase in equilibrium attacks, which become structurally feasible through planning, although a hierarchical network structure is long-term suboptimal for combat. Diplomatic efforts engaging the Taliban must be robust enough to withstand this increase in equilibrium attacks and should recognize it as a collateral effect of the welcomed underlying domain shift from combat to negotiations.

With these two scenarios as options, Western diplomats have a range in which to conceptualize their best response to the Taliban's best response. As data on the progress of the Afghan conflict becomes available, this range can be narrowed down to deliver more precise points of optimal engagement.

6 Conclusion

We establish an opponent model of the Taliban that allows us to connect observations on attack patterns to the Taliban's domain-specific reward maximization. The focus of the opponent model constitutes the Taliban's optimal network structure for combat and negotiations. We validate empirically that changes in the optimal network structure occur either within a domain (combat) or across domains (combat → negotiations). Based on these changes in the network structure, we provide options for a best response of Western diplomats conducting negotiations with the Taliban. How well these options fit the Afghan conflict depends on the data we can extract from it. Across possible datasets, the architecture of the opponent model proposed here enables us to channel simple observations on conflict events into insights into the optimal structure underlying these events.

References

S. Afsar, C. Samples and T. Wood. 2008. The Taliban – An Organizational Analysis. Military Review (Mai-June 2008), 59–72.

J. Arquilla and D. Ronfeldt. 1999. Networks and Netwars – The Future of Terror, Crime and Militancy. RAND, Santa Monica, CA.

S. Braynov. 2009. Adversarial Planning in Networks, in S. Argamon and N. Howard (eds.) Computational Methods for Counterterrorism. Springer, Berlin, GER.

D. Carmel and S. Markovitch. 1996. Incorporating Opponent Models into Adversarial Search. AAAI-96 Proceedings, 120 – 125.

W. Enders and X. Su. 2007. Rational Terrorists and Optimal Network Structure. Journal of Conflict Resolution, Vol. 51(1), 33–57.

A Giustozzi. 2017. Afghanistan: Taliban's organization and structure. LandInfo, Oslo.

B. Hung. 2010. Optimization-Based Selection of Influential Agents in a Rural Afghan Social Network. Master's Thesis, MIT, Operations Research Center.

B. Hung, S. Kolitz, A. Ozdaglar. 2013. Optimization-Based Influencing of Village Social Networks in a Counterinsurgency. International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction 2013, 10–17.

R. Putnam. 1988. Diplomacy and Domestic Politics: The Logic of Two-Level Games. International Organization (Summer 1988), Vol. 42(3), 427–460.

J. Rios, D. Banks and D. Insua. 2015. Modeling Opponents in Adversarial Risk Analysis. Risk Analysis (April 2006), vol. 36(4), 742–755.

F. Schadd, S. Bakkes and P. Spronck. 2007. Opponent Modeling in Real Time Strategy Games. Gamecon 2007, 133–145.

S. Vukovic. 2015. Multiparty Mediation and Conflict Management. Routledge, NYC, New York.