A Spatial Multiagent Model of Border Security for the Arizona–Sonora Borderland

Maciej M. Latek¹, Seyed M. Mussavi Rizi¹, Andrew Crooks¹, and Mark Fraser²

¹ Department of Computational Social Science
George Mason University
Fairfax, VA 22030
mlatek, smussavi, acrooks2@gmu.edu
² EADS North America
Herndon, VA 20171
mark.fraser@eads-na.com

Abstract. We report the results of the first sprint of a project to build decision support tools for border security that incorporate interactions among border security forces, smugglers and the population and represent integrated technology architectures made up of fixed and mobile sensor and surveillance networks. To demonstrate the feasibility of social simulation for the security of the Southwestern U.S. border, we first describe open-source data on Customs and Border Patrol (CBP) and smuggling organizations, replicating for 2009 the landscape of gateway organizations and cartels in Sonora along with the border security architecture for the Tuscon sector of the CBP. We then recount the architecture of the model, connecting a disaggregated view of these organizations to a high-fidelity representation of the physical environment and sensor networks. Finally, we discuss model dynamics and validity, and the generalizability of our approach.

1 Introduction

1.1 Research objectives

Legitimate commerce, organized crime, migration and demographic change meet in borderlands. Providing sufficient risk mitigation guarantees faces considerable cost and technological challenges, especially when addressing rapidly adapting and increasingly sophisticated opponents like the Mexican drug cartels and their U.S. counterparts [1–3]. Responding to this challenge, we have set out to build the next generation of tools to support border security analytics. These tools incorporate interactions among border security forces, smugglers and the population into an integrated technology architecture made up of fixed and mobile sensor and surveillance networks and can create a unified representation of the border security architectures with organizational, social and policy contexts; address budgetary challenges by emphasizing cost–benefit and economic spillovers.
at all levels, and present anticipated interactions and reactions of adversaries and stakeholders.³

1.2 Challenges to the security of the Southwestern U.S. borders

Designing our tool requires separating challenges common to most border security problems from those specific to the U.S.–Mexico border. Common challenges include representing regulatory and budgetary realities of securing borders that shape the performance of border security programs to a large extent; local, state and federal law enforcement agencies and other government bodies, businesses and local populations that differ in how they perceive the current situation of the borderland; what it should be, and whether it is converging to what they desire, and lastly, border security organizations in the U.S. and across the border that do not fall under a unified command structure and pursue independent goals.

Issues specific to the U.S.–Mexico border go beyond the enormous geographic, organizational and environmental diversity of the border. Narcoinsurgency in Mexico has nurtured an ecosystem of adaptive and increasingly sophisticated cartels that engage in influence and psychological operations [4], seeking silence and impunity, and managing perceptions by various means [5, 6] designed to shock and intimidate: Beheadings, kidnappings, assassinations and bombings, or to subvert and undermine the legitimacy of the Mexican state: Staging blockades, checkpoints and demonstrations; writing songs and news stories extolling cartel virtues, and corrupting journalists on both sides of the border. Drug cartels are also increasingly more technologically savvy and use intelligence tradecraft to learn about the behaviors of their competitors and law enforcement agencies.

1.3 Motivation for social simulation

Probabilistic risk assessment (PRA) is the first tool for high-level decision support in border security [7]. Effective deployment of PRA is limited, because it requires elicitation of a vast number of highly subjective probabilities and cannot directly express available geospatial, technological and organizational aspects of border security. Other methods of decision support for border security can interact with the data more directly, in particular, geospatial analysis [8, 9], operations research methods [10–12] and statistical analysis [13]. However, these methods ignore the adaptive nature of the adversary and the organizational context of border security. In order to pass the litmus test of (a)–(e), we turned to multiagent simulations as a potential solution.

³ The materials contained herein represent the opinions and views of the authors and should not be construed to be the opinions or views of George Mason University or EADS North America. ECO# 2012-ENA-0508-02-DG. Uncontrolled Unclassified Information (UUI): This documentation does not contain ITAR-controlled data within the definition of ITAR and has been reviewed and approved for release to non-US persons.
Multiagent simulation is a relatively new modeling paradigm that explores how interactions among individual entities form aggregate phenomena [14]. A multiagent simulation is essentially composed of agents and the world they inhabit. What distinguishes these agents from other modeling approaches is that they are not centrally governed, and do not have to maximize any single common objective function; instead they have rules or reasoning mechanisms that dictate agents’ behavior. By simulating agents’ learning and interactions with each other and their environment, multiagent models can generate outcomes from the bottom up rather than the top down.

Multiagent simulation has already been applied to border security issues. For example, [15] explored the effects of using mini UAVs along with other surveillance assets on the Southwestern U.S. border; [16,17] studied dynamics of sensor networks and wireless communication networks during patrols; [18] developed a simulation to support a tabletop war gaming application, and [19] applied computational game theory for point of entry security; [20] to patrol planning by Federal Air Marshals aboard commercial flights. [21] optimized intelligence analysis processes by developing a multiagent model of the interactions of terrorist and antiterrorist organizations; [22] built cognitive architectures and simulation methodologies for robust courses of action under strategic uncertainty, and [23] applied simulation to study counternarcotics in Afghanistan. To the best of our knowledge, our effort is the first to use social simulation to move beyond supporting tactical, day-to-day operations.

Using multiagent simulations as the paradigm to view borderlands as complex adaptive systems, we organized our research into concurrent modeling, data collection and analysis tracks. We used simulation to represent cognitive and organizational bases of decision making that give rise to emergent outcomes for Blue and Red interactions. These outcomes provide futures that can be used to assess the performance of courses of action or to evaluate investments into and modifications of security programs, procedures, plans and policies. Moreover, we can determine the robustness of proposed courses of action when facing adaptive and anticipatory adversaries and stakeholders who may resort to deception. Finally, the model provides guidance on how to structure operational data. This functionality helps analysts to structure intelligence, especially by inferring and encoding the structure of the adversary organization and expressing the range of operations it can wage. Secondarily, it highlights intelligence inconsistent with history or organizational objectives of adversaries.

2 Outline of the solution

2.1 Agents, environments and behaviors

Our current model incorporates the following components:

1. Detailed geography of the border area.
2. The adversarial ecosystem including Blue and Red organizations and their information sharing processes, zones of control, preferred market niches, modes of penetration and alliances.

3. Autonomous cognitive agents that form organizations mentioned above, each satisfying individual and organizational goals with heterogeneous decision making and learning.

4. Selected licit and illicit markets in order to calculate revenues for cartels and their contractors and to set baseline opportunity costs for recruitment.

For example, in our simulation Red is divided into cartels, gateway organizations and individual smugglers. A cartel supplies cargo only to the gateway organizations willing to take it across the border with the lowest expected cost. A gateway organization is more complex: It is composed of individuals driven by greed, fear and personal loyalties; pursues several goals concurrently; has some idea of who its opponents, competitors, suppliers and buyers are; can share information with other gateway organizations, and is responsible for recruiting individual smugglers.

Gateway organizations give smugglers a cargo at a certain location and provide them with the location of a pick-up point. Smugglers are free to choose any route that matches their endurance and time constraints, minimizing their perceived risk of being detected and intercepted. This risk is a mix of historical knowledge of areas in which other smugglers have been caught, the prevalence of Blue patrols observed by smugglers from the same or allied organizations and knowledge of fixed sensors and their viewsheds.

2.2 Data requirements and collection

The key enabler in our model is the data collection process we designed to support modeling. We collected and collated past and present patterns of illegal border activities and border security architectures, and information on the environment and human terrain and sought to understand individual and group decision making for the players involved from various narrative sources.

Environmental data layers came first. Figure 1 outlines the key input layers for the Arizona–Sonora scenario, covering both sides of the border: Landcover information, roads and tracts and terrain roughness derived from terrain relief. We used these layers to calculate possible routes and time, costs of traversing each route on foot and on vehicle, and viewsheds for various sensors. We used information on population density, night-time lights and land ownership to provide additional inputs for the initial phases of the routing problem by smugglers, and to limit and delay access by Blue patrols. Information on water sources, intermediary resupply points and potential shelters is used by smugglers and human traffickers traveling on foot.

All geospatial layers are subject to significant preprocessing. A “gluing” process links layers from various sources. For example, road network and population data layers can be merged from national geospatial sources such as the
Fig. 1. Environmental data layers for Arizona-Sonora scenario.
Census Bureau, the U.S. Geological Survey, the Mexican Geography and Census Bureau and the International Boundary and Water Commission, or taken from a unified data source like OpenStreetMap. Landcover data can be taken from low-resolution sources with global coverage such as Moderate Resolution Imaging Spectroradiometer and Advanced Very High Resolution Radiometer [24] or stitched together from higher-resolution, but differently encoded national sources. We gave special treatment to landcover and relief data that make modeling perceptions of individuals and mobile sensor platforms computationally tractable.

Another dataset describes Blue concept of operations. Figure 2 presents a sample initial border security architecture. The simulation requires locations, force allocations and rotation schedules of CBP stations and forward operating bases along with locations of mobile platforms, integrated fixed towers and unmanned ground sensors. For the Arizona–Sonora border, we used openly available information by the CBP Tuscon sector [25] and environmental impact studies [26]. Collaboration between Border Patrol and other federal land owners like the U.S. National Park Service is reportedly far from perfect. For example, roads and vehicles are not allowed on certain federal lands alongside the Arizona–Sonora border, keeping the Border Patrol out while illegal border crossers can use vehicles, horses and bikes in areas restricted to licit and governmental traffic [27, 28]. That is why users of our tool can outline boundaries of responsibility for Blue organizations and rules for crossing these boundaries on the land ownership layer.

![Fig. 2. Sample Arizona–Sonora border security architecture.](image)

The Red side is described by outlining the initial population of cartels, gateway organizations and their human and technical resources. Figure 3 outlines the competitive landscape of the Arizona–Sonora borderland around 2009. South and Central American cartels have long used analytical and surveillance technologies for intelligence and counterintelligence. As reported in [29], in 1996 the Cali cartel used link analysis on a database of phone records of Cali residents to cross-reference phone calls among its members and American and Columbian
counternarcotics officials. It managed to detect, capture and kill at least 12 informants. More recently, [30] reported that the Mexican cartels have deployed an encrypted distributed radio network across almost all Mexican states. At the tactical level, radio traffic monitoring, coordination of movement and opportunistic use of small UAVs is reportedly widespread among smugglers [31]. We also included data on the operational and tactical sophistication of gateway organizations, along with their locations, preference for smuggling drugs or trafficking humans and border penetration mode\(^4\).

\(\text{Fig. 3.} \) Competitive landscape of cartels and gateway organization of Arizona-Sonora borderland around year 2009. We used this data to initialize cartels and gateway organizations and to set information sharing and contracting limits among them.

\(^4\) We currently only cover border penetration via terrain. We have not represented the use of ultralights, sewage, storm water systems and point of entry with corrupted officials or forged documents.
2.3 System architecture

The architecture of our system, outlined in Figure 4, is organized around a data warehouse for storing physical, geographic and social data, and the Mason simulation engine. We provide users with a work process using open source platforms for inputting and editing organizational and network data [32] and for managing geospatial data [33].

![Figure 4. Current system architecture.](image)

The Mason simulation core [34] is a Java framework that can be used for a single run on a laptop with a graphical user interface or on a desktop for small-scale experiments. It decouples visualization from simulation execution, making it possible to deploy the simulation to a cluster or cloud for large-scale parameter sweeps. We can run the current version of the Arizona-Sonora simulation for a two-year period at high spatial, temporal and organizational resolutions once in 30 minutes on a single core with 4 Gb of memory. During model development, we routinely conducted overnight unit checking and parameter sweeps of 5000 to 10000 runs on a small cluster.

The Mason graphical user interface helps the user to monitor and inspect model dynamics at runtime by displaying the following data:

1. Trails and current positions of Blue patrols and Red smuggling groups.
2. Current vision ranges and sensor coverage for Blue and Red, including Blue common operating picture and current perceptions by Red of the locations of Blue patrols.
3. Perception of detection and interception risks for each Red organization.
4. Time series on successful penetrations, drug load deliveries and seizures, smuggling groups spotted and intercepted.
5. Information on the history of detections for each sensor and Blue organization.
6. Readiness statistics for Blue and Red, including information on captured and recaptured smugglers and new recruits.
7. Information on financial resources of Red.

The user interface can be used to perturb the model during runtime. Various additional statistics at a high level of disaggregation are saved to the database for offline analysis. This allows the user to perform face-validity checks, statistically compare various scenarios and course of actions and visualize the results.

3 Validity of the social simulation approach

We are currently performing initial validation of model outputs, investigating qualitative properties of the model and preparing for quantitative validation. The validation plan follows the pattern set forth by [35] and includes both verification or face validation by subject matter expert to determine conceptual validity, comparisons with reimplementation of operations research and statistical models and the external validation against real-world observations at micro and macro levels. For example, macro-level pattern matching may include comparing simulated smuggling corridors to ones reported in the real life. Such a comparison is presented in Figure 5. The outlines of the real-life corridors have been created by collating newspaper articles and narratives of routes taken by smugglers and illegal immigrants that mention geographic features. The simulated view on Figure 5(b) presents an “ant trail” of recent smuggling traces after 6 months of simulated interactions.

![Approximation of real-life smuggling corridors](image1.png)  ![Snapshots of the traces of simulated smugglers](image2.png)

Fig. 5. Real-life smuggling corridors and a sample of simulated movement trails used for face validity and qualitative model checking.
Sparse data exists in open sources on exact routes, their spatiotemporal variability and results of Red interactions with Blue. In Figure 6, we present two samples of high-resolution event data that could constitute desired targets for quantitative validation: Locations in which bodies of deceased immigrants were found [36] and locations of seizures of marijuana and cocaine loads [37]. Unfortunately, neither dataset includes additional data such as date, amount of drug seized and route taken. Nevertheless, this and other data we already collected is sufficient to start the validation process.

![Fig. 6. Sample data layers for quantitative validation. (a) Deaths (b) Drug seizures](image)

## 4 Future outlook

We are working on code hardening and documentation, empirical validation and mechanisms for scenario analysis. We have expanded data farming and large-scale computational experimentation to test assumptions on border security architectures, and Blue organizational behavior, force levels and patrol posture. Large-scale computational experimentation in our effort is key to capture the uncertainty of the system; design corresponding robust courses of actions and portfolios of policies, plans, procedures and programs that perform well in this uncertain environment, and to estimate “known unknowns”\(^5\) [38]. We are also working on exposing the simulation to subject matter experts to obtain better face validity and follow the validation steps outlined in the previous section.

In the current version of the model, we do not have a disaggregated view of the population of the borderland, representation of points of entry or of the logistics and economic layers. We will therefore work on representing and synthesizing the underlying rural and urban population of the borderland at the individual resolution. We can then develop a module where the population can perceive the economic and security effects of interactions between the CBP border architecture and the activities of the transnational criminal organizations and local gangs.

\(^5\) For example, unobservable or hard-to-collect quantities such as the tonnage of drugs and the number of immigrants that go through the border without being intercepted or observed at any point.
Finally, we are working on extending the common ontology of subsystems and actors, and their objectives, capabilities and constraints by collecting data on the organizational features and technological properties of joint border security arrangements. This includes adding new environments like the Carrizo Cane forests alongside the Rio Grande in Texas, allowing for riverine operations by Blue and Red. These steps continue the move our effort toward a comprehensive computational framework that connects security, economics, and politics of border security with current and future border security technologies.

References

28. USA Today: Federal Agents Say Environmental Laws Hamper Work. USA Today (October 2011)
30. Ackerman, S.: Radio Zeta: How Mexico’s Drug Cartels Stay Networked (December 2011)
Discussion of data flows in the model

Figure 8 presents how information moves through the simulation: From input data fusion, scenario authoring and experiment design to post–processing simulation results. The process is engineered to enable iterative testing and benchmarking as we proceed with model implementation and data collection. Determining proper metrics for spatial multiagent models of strategic interactions is an active research question, so our key model benchmarks are evolving. Available tests and model views also coevolve with data collection. The empirical data we have collected so far can be used to match model outputs with the spatial distribution and volume of interdiction and apprehension events, relative prices of labor and drugs on both sides of the border, and quarterly and annual data with numbers of various interdiction events for each CBP station. However, matching model outputs with spatial time series data requires developing models of additional sectors.

In this appendix, we demonstrate how we performed parameter sweeps for model calibration, sensitivity analysis, and empirical validation of drug interdiction locations. Table 1 lists swept parameters along with their ranges and definitions, and 8 model outputs collected from each run to classify the outcomes of interactions between the CBP and smugglers. We quickly built networks of individual runs based on output similarity and used visual cues to data mine relationships between model parameters and outputs. Figure 8 shows an example of such a network.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Range</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>blue force multiplier</td>
<td>(0, 2)</td>
<td>Multiplier of the baseline number of CBP patrols</td>
</tr>
<tr>
<td>blue sensor fraction</td>
<td>(0, 1)</td>
<td>Fraction of functioning integrated fixed towers</td>
</tr>
<tr>
<td>blue bias and red bias</td>
<td>(0, 5)</td>
<td>Exploration bias for both sides</td>
</tr>
<tr>
<td>blue lambda and red lambda</td>
<td>(0, 5)</td>
<td>Exploitation weight for both sides</td>
</tr>
<tr>
<td>(\Phi)</td>
<td>(0, 1)</td>
<td>Shared rate of belief adaptation</td>
</tr>
<tr>
<td>maximum observation equivalents</td>
<td>(0, 10)</td>
<td>Upper bound on information certainty</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Output</th>
<th>Mean, standard deviation</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>interceptions with load</td>
<td>3175, 207</td>
<td>Accumulated number of interceptions of inbound smugglers</td>
</tr>
<tr>
<td>interceptions without load</td>
<td>1116, 177</td>
<td>Accumulated number of interceptions of outbound smugglers</td>
</tr>
<tr>
<td>blue--red detections</td>
<td>26994, 10595</td>
<td>Accumulated number of detections of smugglers by CBP</td>
</tr>
<tr>
<td>red--blue detections</td>
<td>11071, 3969</td>
<td>Accumulated number of detections of CBP patrols by smugglers</td>
</tr>
<tr>
<td>volume of successful shipments</td>
<td>2946, 724</td>
<td>Accumulated number of successful shipments</td>
</tr>
<tr>
<td>total gateway money</td>
<td>1000, 513</td>
<td>Accumulated profits of all gateway organization (10K USD)</td>
</tr>
<tr>
<td>total active smugglers</td>
<td>191, 67</td>
<td>Final number of active smugglers</td>
</tr>
<tr>
<td>total employed smugglers</td>
<td>3946, 332</td>
<td>Total accumulated number of smugglers employed</td>
</tr>
</tbody>
</table>

Table 1. Description of parameters for one sweep, together with key model outputs used for low-dimensionality mapping. Averages and standard deviations of the output variables are computed from a sweep with 8500 runs. See Figure 8 for a possible visualization of results of that parameter sweep.
Fig. 8. A 10–nearest neighbors network of 8500 runs, with the Euclidean distance between 8 outputs, each rescaled to mean 1 as the underlying metric, shown in Table 1. Each node corresponds to an individual run and is colored by the value of the parameter blue patrol multiplier (the number of CBP patrols). Node size is proportional to the value the output variable total gateway money (total accumulated profits of all gateway organization).
Let us move to empirical validation. On panel A of Figure 9 we present a heatmap derived by a kernel density estimate of data points presented on Figure 6(b). Note that the points on Figure 6(b) are derived from a multiagency fused dataset and include interdictions by law enforcement agencies not represented in the model such as tribal and local police departments, Arizona State Police and various federal agencies like DEA. In our validation framework, the heatmap can be made smoother or more focused by controlling the kernel half-width. Heatmaps can be computed from event data collected from a single run or multiple runs. Lower resolution heatmaps derived from multiple runs can be used to extract salient features of tactical coadaptation.

On panel B of Figure 9 we present a heatmap derived from data generated by a single run of a development version of the model. The number of events in the real-life dataset was 3750, this particular instance of the simulation resulted in 4100 events. Panel C is a lower-resolution heatmap showing an average difference between simulated and real-life event densities. By inspecting panels A and B we can quickly observe the following:

1. Real-life patterns are in principal more “diffused”.
2. In real life, interdictions are focused on the border, whereas in that particular version of the model, the bulk of interdictions happen farther away from the border.
3. In Tohono O’Odham Indian Territory, Organ Pipe Cactus National Monument Park, Cabeza Prieta Refuge Park and Eastern part of Coronado National Forest, no interdictions are observed.
4. The number of interdictions around points of entry such as Nogales or Douglas is overestimated, clearly visible on panel C.

These observations inform our subsequent model development. For example, one of the primary real-life missions of CBP is to push drug smugglers and associated violence away from the population centers. In the tested version of the model, the security of the population was not part of CBP’s fitness: Our agents focused on maximizing the number of interdictions, which was achieved more easily if some of the patrols were pushed back from the border to shorten patrol lengths. Similarly, we hardcoded limitations in access to some of the federal land, given open-source narratives on coordination problems between CBP and National Park Service and Department of Agriculture. The fused real-life data seem to suggest that even in 2009 such limitations did not open obvious holes in security coverage, as areas such as Tohono O’Odham Indian Territory do feature interdiction events, perhaps by local police department. A great number of interdictions took place on or before highways. This tells us that it is paramount that we work on representing other law enforcement agencies, interagency lines of responsibility and intelligence collection and sharing arrangements to better understand not only what happens on the Red side, but also how complex are all the moving elements on the Blue side.
Fig. 9. Heatmap of real-life drug seizure events for marijuana (block A), identical heatmap produced off of a single model run (block B) and an difference between real-life density and average of 8500 simulated densities (block C). The characteristic resolution for blocks A and B is 2 km. For block C we used density smoothing with characteristic resolution of 10 km.