BEND Battle: An Agent Based Simulation of Social-Cyber Maneuvers*

Matthew Hicks[0000−0002−0614−8845] and Kathleen M. Carley[0000−0002−6356−0238]

CASOS Center, Software and Societal Systems, Carnegie Mellon University, Pittsburgh, USA
{mhicks3,kathleen.carley}@cs.cmu.edu

Abstract. BEND Battle is an agent-based model that simulates social media users conducting BEND social-cyber maneuvers in order to influence each other on a single topic. It provides a visualization of two sides conducting maneuvers against each other and the effects of those maneuvers. Previous BEND maneuver simulations are grounded in medium-specific implementations. This simulation attempts to normalize confounding factors - such as varying network structures - and focuses on how BEND maneuvers interact on a level-playing field represented by an attention-limited broadcast domain. Results suggest that explain and negate maneuvers provide a broad counter to a wide range of opposing BEND maneuvers.

Keywords: BEND · simulation · social-cyber · agent-based.

1 Introduction

Looking at cyberspace through the lens of warfare is not new. Interactions between adversaries within cyberspace have often been referred to in military terms of attack and defense [9]. Cyberspace simulations have been used to model these conflicts, often closely emulating current physical military doctrine [10]. However, these simulations focus primarily on the cyber-terrain itself - accurately deducing that terrain has a large impact on the outcome of conflict [10]. However, just as the physical domain of warfare stretches into the digital space, so too does the information domain. This is social cyber-security [6] [7].

BEND provides a framework for discussing social-cyber interactions through the lens of maneuver warfare [2]. BEND is shorthand for the social-cyber maneuvers: back, build, bridge, boost, engage, explain, excite, enhance, neutralize,

* The research for this paper was supported in part by the Office of Naval Research (ONR) under grant N00014182106, the United States Army under grant W911NF20D0002, and the center for Informed Democracy and Social-cybersecurity (IDeaS). The views and conclusions are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the ONR, the United States Army, or the US Government.

nuke, narrow, neglect, dismiss, distort, dismay, and distract. These maneuvers and their definitions are taken from Beskow and Carley's 2019 work Social cybersecurity: an emerging national security requirement [2] as refined and validated by Blane et al. in 2022 [4].

BEND maneuver comparisons and interactions suffer from many of the same problems as tasks in other military domains. Attempts to compare the interaction of military tactical tasks fails due to the overwhelming influence of all the other factors involved. In the US Army, this is called METT-TC. From Army Field Manual 3-21.8, METT-TC stands for Mission, Enemy, Terrain/Weather, Troops available, Time available, and Civilian considerations [11]. When planning an attack, all these factors influence the success or failure of the planned task - often more than the form of the task selected.

For an even simpler example, while an attacker may generally consider a 3:1 advantage in combat power sufficient to conduct an attack, if the defender occupies especially favorable terrain, then they might require a 5:1 or even a 10:1 advantage. Other factors almost always matter more than the maneuver.

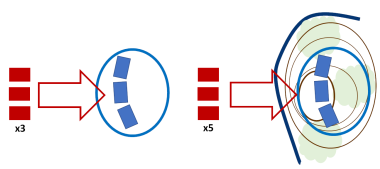

Fig. 1. Generally an attacker requires a 3:1 advantage to succeed (left), but given additional factors - high ground, a water obstacle, dense forests, etc. - the ratio may increase (right).

However, this is not to suggest that any attempt at comparison is fruitless. It is important to know that, generally speaking, attackers require a 3:1 advantage – just as it would be important to know that Dismay is more effective than Engage (or vice versa). In order to make such comparisons the confounding factors need to be equalized.

This is where a simulations can be helpful. BEND Battle is an agent-based model that simulates social media users conducting BEND social-cyber maneuvers in order to influence each other on a single topic. The focus of BEND Battle is on providing a visualization of two sides simultaneously conducting maneuvers against each other and the effects of those maneuvers. The winner is determined by evaluating which side used BEND maneuvers to effectively retain agents of their own and pull opposing agents over to their side.

2 Related Works

The twitter_sim2.0 model – as outlined by Blane, Moffit and Carley in 2021 [5] – is the closest model to BEND Battle. However, it is distinctively more focused on Twitter-specific interactions. The twitter_sim2.0 model also has additional capabilities that make it a more comprehensive model than BEND Battle. The simulation accounts for both emotion and logic – ensuring tweets that emotionally correspond with a recipient have magnified effects.

Table 1. Docking Lite with twitter_sim2.0

Feature	BEND Battle	twitter_sim2.0
Media Agents		✓
Opinion Leaders	✓	✓
Information Access		✓
General Memory	✓	✓
Homophily	✓	✓
Limited Attention	✓	✓
Dynamic Network		✓
Emotional Response		✓
All BEND Maneuvers	✓	
Live Visualization	✓	

3 Model Description

BEND Battle is built on the NetLogo simulation platform [15]. The game space within NetLogo for BEND Battle is comprised of two opposite stances on a single topic – denoted by red and blue squares at the far left and right of the model visualization. Between the two stance squares, users - depicted by circles - compete to push and pull each other towards one stance or the other through the application of BEND maneuvers. The environment is not Twitter, Reddit, or YouTube; rather, the environment is a broadcast information transmission medium through which users enact BEND effects upon other users.

3.1 User Attributes

The entities being modeled are individual users of a social media. In order to facilitate the visual understanding of BEND maneuver effects, BEND Battle uses four primary attributes for all users (shown in Fig. 2.): strength, topicality, affiliation and leadership.

Strength Strength is a measure of a user's stance on the topic. It is visualized on the x-axis of BEND Battle, ranging from -100 (fully anti-stance) to 100 (fully pro-stance). A user with a strength greater than 0 is on the blue, pro-stance side and a user with a strength less than 0 is on the red, anti-stance side. Users initialised with a strength of exactly 0 are randomly assigned, while users that have their strength changed to 0 retain their previous side. Within BEND Battle a user's strength is only changed by BEND maneuvers - it does not decay or grow based upon that user's actions.

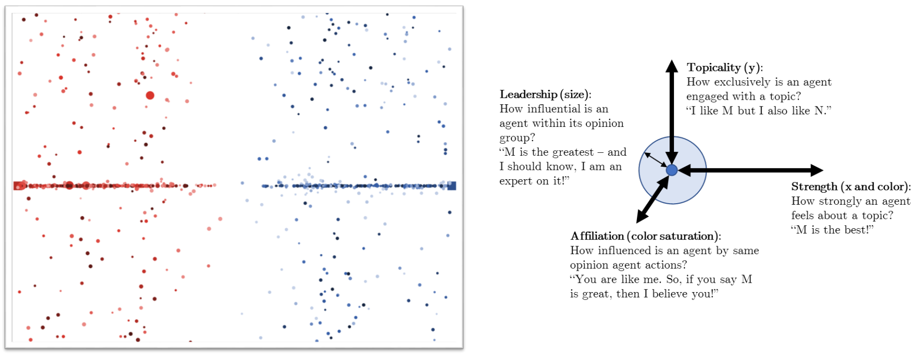

Fig. 2. BEND Battle visualization of agents is based on four attributes.

Topicality Topicality is a measure of the exclusivity of a user's engagement with the topic. Because BEND Battle is meant to be a simplified, normalized encounter on a single topic, all engagement with other topics is amalgamated into a single attribute per user - topicality. This can be thought of as the overall focus of a user on this particular topic vice other topics.

Topicality is visualized on the y-axis of BEND Battle, ranging from 0 (solely focused on this topic) to |100| (fully preoccupied with other topics). At initialization, each user is randomly determined to vary either from 0 to 100 or 0 to -100.

Together strength and topicality - variation of stance engagement and variation of topic engagement - determine the overall engagement of a user. Mathematically, the engagement of a user is determined by calculating the distance between themselves and their current stance box - read at -100,0 and blue at 100,0. Overall engagement is important because it determines how often a user conducts a maneuver. The stronger they feel about their stance and the more focused they are on the topic - the more likely they are to act each tick. This is determined by an exponential function, where D is the absolute value of the distance between a user and their stance box:

$$P[act] = (e^{-.069D})$$

Leadership Leadership is a measure of the influence a user has over the BEND battlefield. It is visualized by the overall size of the user ranging from 0 to 100.

BEND Battle adheres to a form of limited attention [12] such that not all users are able to process all received information. There are many factors that influence the information to which users pay attention. While the underlying structure of a network plays a dominant role [12], the influence of the originating individuals is also important [12]. Because the network structure of BEND Battle is normalized as a single broadcast domain, the influence of the originating user becomes the primary factor for attention consideration.

Instead of every user receiving every BEND maneuver and then calculating which to pay attention to based upon the influence of the originating user, the originating user's leadership is used to determine a limited number of users upon which to apply the BEND maneuver's effects.

Affiliation Affiliation is measure of how influenceable an agent is (visualized by color saturation). Affiliation is used here for disambiguation and to address a number of related concepts. Generally, BEND maneuvers that would typically target network structure - Bridge, Narrow, etc. - and influence homophily, cohesion, or related measures, instead affect the affiliation attribute of users within BEND Battle. This models these users becoming further separated from their group and reducing the in-group influence and increasing the out-group influence of BEND maneuvers.

Affiliation ranges from 0 to 100 but is signed by the color of the user. This allows for each user's affiliation to range from -100 to 100 based upon the colors of interacting users. Affiliation is used within a logistic function to determine the influence BEND maneuvers have on a user. A blue BEND maneuver will influence a blue user with high affiliation (80) a lot more than it will a red user with high affiliation (-80). When determining the the value of a new attribute based upon a BEND maneuver's affect, the affiliation logistic function is combined with that attribute's exponential function. Where X is the current value of the attribute, X' is the future value of the attribute, I is the magnitude of the change, and A is the affiliation value with respect to the color of the originating user:

$$X' = X + I\left(\frac{1}{1+e^{-.05A}}\right)\left(e^{-.069X}\right)$$

3.2 BEND in the Model

Each BEND maneuver has an effect on one or more of the user attributes - strength, topicality, affiliation, and leadership (STAL). Because one of the primary motivations behind BEND Battle is to normalize the effects of maneuvers with respect to one another, each of the maneuvers is given three points to allot towards changing these attributes. The English language definition of each BEND maneuver has been translated into effects on STAL with a gross magnitude of three.

For instance, Engage as a maneuver provides "more arguments for better associations with the narrative" [1]. Thus, within STAL, Engage increases strength and topicality while leaving affiliation and leadership alone. Given a total of three points, Engage increases strength by two and topicality by one.

Distract is used to "misdirect the audience offering a new distracting topic or adding noise and confusion" [1]. This translates to three points of change in topicality away from the baseline.

There are several complications involved with translating BEND maneuver effects in this way. First, BEND Battle applies the effects of an executed BEND maneuver across a random set of users within the battlespace. This means that the maneuver effect translations need to account for how maneuvers affect both friendly and enemy users.

Second, because BEND battle does not consider the sentiment of the actual maneuver, the number of distinct maneuvers is effectively reduced. The Bridge maneuver adds "linkages between various groups and communities" [1] and Narrow is meant to "polarize communities, isolate groups and break connections between them" [1]. These maneuvers might look very different in reality but in an effects-only simulation where the effects are randomly applied across both friendly and enemy users, they appear identical. Due to this simplification, Build, Bridge, Boost, Nuke, Narrow, and Neglect all have identical effects within BEND Battle.

A full layout of all maneuver effects is provided in Table 2.

Table 2. BEND Maneuver Effects

	Friendly				Enemy			
	Strength	Topicality	Affiliation	Leadership	Strength	Topicality	Affiliation	Leadership
Engage	++	+			- -	-		
Explain	+	+		+*	-	-		
Excite	+	+	+		-	-	-	
Enhance	+++				- - -			
Dismiss	+++				- - -			
Distort	+	+	+		-	-	-	
Dismay	++		+		- -		-	
Distract		+++				- - -		
Back				+++				
Build	+		++		-		- -	
Bridge	+		++		-		- -	
Boost	+		++		-		- -	
Negate								- - -
Null	+		++		-		- -	
Narrow	+		++		-		- -	
Neglect	+		++		-		- -	

3.3 Model Overview

BEND Battle uses a simple three step algorithm for resolving the interplay of user initiated BEND maneuvers. These three steps are check-action, take-action, and resolve-action. All three steps are run each tick (unit of time/iteration within NetLogo).

check-action: All users determine if they are going to act this tick based upon the distance from the user to their stance box
take-action:
A) Acting users determine which BEND maneuver they will execute this tick - only one maneuver is allowed per tick - and the maneuver is chosen based upon the probabilities given for each of the BEND maneuvers
B) Acting users determine which other users will be the target of their BEND

maneuver
C) All acting users conduct their BEND maneuvers in a random order
resolve-action: All users that were affected by BEND maneuvers appropriately change their position, color, and size - they also recalculate their chances for acting next turn.

4 Experiment

For the virtual experiment, we will manipulate the probability of each of the 16 BEND maneuvers. Every maneuver will be pitted against every other maneuver for a total of 256 test cases. The default probabilities are taken from Alieva et al., 2022, which they gathered from tweets about Alexei Navalny from August 2020 until August 2021.[1] Because this dataset allowed for multiple maneuvers within each tweet, these real world probabilities are used as ratios in order to determine probability of each maneuver within the simulation. All maneuvers are given their Navalny dataset ratio by default and then this ratio is increased for each by .2 when it is that maneuver's turn to be pitted against another.

We initialize each run by setting up 500 agents per team with random strength (1 to 100/ -1 to -100), random topicality (1-100), random affiliation (1-100), and random leadership (1-10).

From the dependent variables, we extracted either a winner or a tie from each combination of the BEND probabilities. A winner is any team with at least 525 users out of 1000 after 500 time-ticks. We replicated each of the 256 combinations of the BEND maneuvers 100 times.

Table 3. 16x16 Virtual Experiment

Independent Variables	# Test Cases	Values Used
Build	2	.7/.9
Bridge	2	.2/.4
Boost	2	.35/.55
Back	2	.25/.45
Null	2	.16/.36
Negate	2	.15/.35
Neglect	2	.05/.25
Narrow	2	.1/.3
Distract	2	.2/.4
Distort	2	.15/.35
Dismiss	2	.26/.46
Dismay	2	.25/.45
Explain	2	.36/.56
Excite	2	.3/.5
Enhance	2	.3/.5
Engage	2	.4/.6
Control Variables	# Test Cases	Values Used
Time Periods	1	500 ticks
Red/Blue Starting Entities	1	500
In/Out Percent / Time Period	1	0.01
Dependent Variables		Values Expected
# of agents on the red team		0-500
# of agents on the blue team		0-500
16x16	256 cells	
Replications per cell	100	
25,600 total runs (12,800,000 data points)		

5 Results and Discussion

In order to visually represent the variations across the runs, for each set of variables we subtracted the number of runs that red won from the number of runs that blue won. Thus, a 100 represents that blue won every iteration for that combination of variables, a -100 means that red won every iteration, and a 0 indicates that they either tied every iteration or the number of red wins equaled the number of blue wins. In general, positive values indicated that blue is winning more often, whereas negative values indicated that red is winning more often. These results are mapped in Fig. 3.

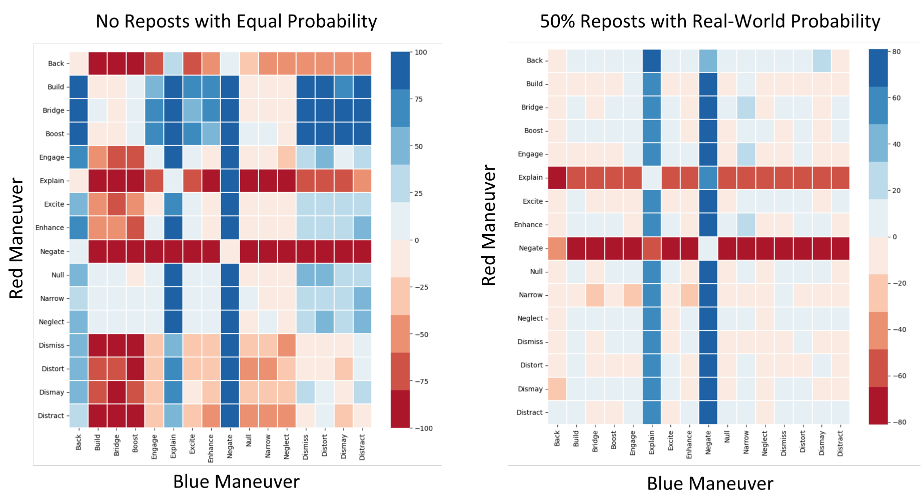

Fig. 3. Results of all of BEND Maneuver head-to-head match-ups. On the left are the results with no reposts and a default equal probability for all maneuvers. On the right are the results from the full virtual experiment.

The results indicate that increasing the likelihood of Explain and Negate maneuvers is the most effective way to predict a winning outcome. This can be seen by observing that these maneuvers produce darker red results - more wins - when used by Red (anti-stance) and darker blue - more wins - when used by Blue (pro-stance). This makes sense as the both Explain and Negate maneuvers affect leadership, which has an amplifying or subduing affect on all other maneuvers conducted by the targeted user. While Back also affects leadership, any Back likelihood increase is dampened because Back is already being conducted as a part of the 50% of maneuvers that are reposts. You can see this dampening clearly in Fig. 3, where Back performs better without reposts as compared to with reposts.

6 Validation

Any model that simulates the personal beliefs of individual users and the influence other users might have on those beliefs is going to be difficult to validate. However, face validity - though difficult since the simulation is built to avoid directly representing any specific social media platform - is accomplished through the help of domain experts. Additionally, BEND Battle takes advantage of several stylized facts outlined in Table 4.

Table 4. Stylized Facts

Summary	Effect	Source
Users trend toward their own beliefs	Sigmoid and logistics functions	Sunstein, 2022.[14] Cinelli et al. 2021. [8]
Attentions spans limit how many users are affected	Reach is determined by leadership	Kang and Lerman, 2013. [12] Weng et al., 2012. [12] Lu et al., 2014. [13]
Commonality of reposts	50% of messages are derivative	Beskow and Carley's "Agent Based Simulation of Bot Disinformation Maneuvers in Twitter" from 2019.[3]
Which messages are being reposted	Power law distribution	Lu et al., 2014. [13]
Real world likelihood of BEND maneuvers	BEND maneuver probabilities from tweets about Alexei Navalny from August 2020 until August 2021	Alieva et al., 2022.[1]

7 Future Work

BEND Battle should handle diverse bots and bot actions that reflect historical bot use. This should include bots that conduct multiple maneuvers as well as bots that are sometimes users and users that are sometimes bots.

BEND Battle currently only allows for one maneuver at a time from an agent. Future iterations should allow for multi-maneuver actions since most social-cyber maneuvers in the real world are comprised of multiple individual maneuvers. [1]

Finally, the simulation results tell us more about how BEND maneuvers interact than what actions any group should take. The simulation results are not prescriptive – you cannot necessarily control what users on your side will do. A more interactive model would include randomized user actions on a user-manipulable side but then allow control over some portion of friendly users in order to counteract enemy bot and user actions.

8 Conclusion

BEND Battle provides a simulation for comparing how BEND maneuvers interact with each other and visually showcasing that interaction. The BEND Battle simulation results suggest that stance sides that prioritize Explain and Negate maneuvers are more effective than any other combination of BEND maneuvers.

References

1. Alieva, I., Moffitt, J.D. & Carley, K.M.: How disinformation operations against Russian opposition leader Alexei Navalny influence the international audience on Twitter. Soc. Netw. Anal. Min. **12**, 80 (2022). https://doi.org/10.1007/s13278-022-00908-6
2. Beskow, D.M. and Carley, K.M.. Social Cybersecurity: An Emerging National Security Requirement. Military Review **99**(2):117–127. (2019)
3. Beskow, D.M. and Carley, K.M.: Agent Based Simulation of Bot Disinformation Maneuvers in Twitter. 2019 Winter Simulation Conference (WSC). National Harbor, MD, USA, 2019, pp. 750-761. (2019) https://doi.org/10.1109/WSC40007.2019.9004942.
4. Blane, J.T.,, Bellutta, D., Carley, K.M.: Social-Cyber maneuvers analysis during the COVID-19 vaccine initial rollout. Journal of Medical Internet Research. 08/01/2022 24(3):e34040 PMID: 35044302, DOI: https://doi.org/10.2196/34040 (2022)
5. Blane, J.T., Moffitt, J.D., Carley, K.M.: Simulating Social-Cyber Maneuvers to Deter Disinformation Campaigns. Social, Cultural, and Behavioral Modeling: 14th International Conference, SBP-BRiMS 2021, Virtual Event, July 6–9, 2021. Proceedings, 2021, Pages 153–163. (2021) https://doi.org/10.1007/978-3-030-80387-2_15
6. Carley, K.M., Cervone, G., Agarwal, N., Liu, H.: Social Cyber-Security. In: Thomson, R., Dancy, C., Hyder, A., Bisgin, H. (eds) Social, Cultural, and Behavioral Modeling. SBP-BRiMS 2018. Lecture Notes in Computer Science(), vol 10899. Springer, Cham. https://doi.org/10.1007/978-3-319-93372-6_42 (2018)
7. Carley, K.M.: Social Cybersecurity: An Emerging Science. Computational and Mathematical Organization Theory, 26(4): 365-381. (2020)
8. Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M.: The echo chamber effect on social media. Proceedings of the National Academy of Sciences, 118(9). (2021) https://doi.org/10.1073/pnas.2023301118
9. F. Cohen: Simulating Cyber Attacks, Defenses, and Consequences, IEEE Symposium on Security and Privacy Special 20th Anniversary Program, Berkeley, CA, May, 1999. https://doi.org/10.1016/S0167-4048(99)80115-1 (1999)
10. Dobson, G.B., Carley, K.M.: Cyber-FIT: An Agent-Based Modelling Approach to Simulating Cyber Warfare. In: Lee, D., Lin, YR., Osgood, N., Thomson, R. (eds) Social, Cultural, and Behavioral Modeling. SBP-BRiMS 2017. Lecture Notes in Computer Science(), vol 10354. Springer, Cham. https://doi.org/10.1007/978-3-319-60240-0_18 (2017)
11. Headquarters, Department of the Army: The Infantry Rifle Platoon and Squad (FM 3-21.8) (2016) https://armypubs.army.mil/epubs/DR_pubs/DR_a/pdf/web/ATP 3-21x8 FINAL WEB INCL C1.pdf
12. Kang, J., Lerman, K.: LA-CTR: A Limited Attention Collaborative Topic Regression for Social Media. AAAI Conference on Artificial Intelligence, North America, jun. 2013. https://doi.org/10.1609/aaai.v27i1.8451 (2013)
13. Lu, Y., Zhang, P., Cao, Y., Hu, Y., & Guo, L.: On the Frequency Distribution of Retweets. Procedia Computer Science, 31, 747–753. (2014) https://doi.org/10.1016/j.procs.2014.05.323
14. Sunstein, C.R.: The Law of Group Polarization. Journal of Political Philosophy, 10: 175-195. 2002. https://doi.org/10.1111/1467-9760.00148
15. Wilensky, U.: NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL. (1999)