*As per UGC guidelines an electronic bar code is provided to seure your paper*

*International Journal for Modern Trends in Science and Technology*, 9(05): 901-911, 2023

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0905154
Available online at: http://www.ijmtst.com/vol9issue05.html

# An Examination of the Role of Cognitive Models on Cyber Security andtheir Effectiveness

**Karan Chawla**

Ashoka Unviersity

**To Cite this Article**
Karan Chawla. An Examination of the Role of Cognitive Models on Cyber Security andtheir Effectiveness. International Journal for Modern Trends in Science and Technology 2023, 9(05), pp. 901-911. https://doi.org/10.46501/IJMTST0905154

**Article Info**
Received: 28 April 2023; Accepted: 24 May 2023; Published: 25 May 2023.

## ABSTRACT

*One of the challenges for game-theoretic/ML algorithms is that they usually rely on massive amounts of data to match the model parameters. In sectors like cybersecurity, it often involves human involvement and the gathering of a sizable quantity of data on human decision-making to properly grasp how various defense algorithms could function in actual conditions. Unfortunately, these approaches are rather challenging. Cognitive models are typically starting to play a direct part in applications where predictive models of human decision-making take the function of humans in the activity. Due to a lack of such decision data, it is challenging to comprehend the attacker's decision-making in the cybersecurity field. Defense algorithms frequently work under the presumption that attackers make logical choices and follow the optimal line of action. Researchers found that for the Markov Security Games, the calibrated IBL model performed better than the ACT-R model in accounting for human assessments under both patching scenarios. One potential cause is that the model appears to be unable to gather human data when using the default ACT-R parameters. Recalibrating these variables, however, greatly improved the model's own performance. Studies show that hackers' decisions were influenced by their perception of vulnerability, and there are several ways to shape this perception. To trick computer networks, for instance, one may utilize honeypots, which are simple to attack systems. Utilizing intrusion-detection systems (IDSs) and alerting hackers to their presence and accuracy is the second strategy. This review paper addresses the role of cognitive models in three domains namely Instance based models for masking, ACT-R model for Markov Based Games and Cybonto for increasing threat protection levels in cyber security.*

*KEYWORDS:Machine Learning, Cognitive Models, ACT-R, Intrusion Detection Systems,Markove Based Games, Cyber-Security, Algorithms.*

## 1. INTRODUCTION

The practice of protecting networks, computers, servers, mobile devices, electronic systems, and data from hostile assaults is known as cyber security. The expression is used in a broad variety of contexts, including business and mobile computing, and may be divided into a few main categories. As the global cyber threat evolves swiftly, there are more data breaches every year. When compared to the same period in 2018, this amount is more than twice (112%) the number of data disclosed. Because they gather financial and medical data, some of these industries are more interesting to cybercriminals than others, but any firms that utilize networks might be the target of customer data theft, corporate espionage, or consumer assaults. As the severity of the cyber threat is anticipated to keep increasing, there is an inevitable increase in global spending on cybersecurity solutions. According to Gartner, worldwide investment on cybersecurity will top $260 billion by 2026 and reach $188.3 billion in 2023. In response to the growing cyber danger, governments all over the world have issued recommendations to support businesses in putting good cyber-security practices into place. A framework for cyber security has been created by the National Institute of Standards and Technology (NIST) in the United States. According to the design, all electronic resources should be continuously monitored in real-time to stop the propagation of malicious code and aid in its early detection. Network security is the process of defending a computer network against intruders, such as malicious software that exploits weaknesses or deliberate assaults. The goal of application security is to protect software and hardware from damage. If an application is compromised, the data that it is designed to protect may become accessible. Effective security starts at the design phase, long before a software or gadget is put to use. Information security is used to protect data integrity and privacy during storage and transfer. The methods and decisions used to manage and protect digital assets fall under the category of operational security. This comprises the rules governing where and how data may be stored and shared, as well as the privileges users have when using a network. Defense-related game theory/ML algorithms are frequently data-driven and frequently do not take into account insights about human behavior. In these circumstances, cognitive models might be applied in a variety of ways, such as offering an interpretation of human behavior or serving as a data source for ML algorithms by making precise predictions about human data

## 2. INSTANCE BASED LEARNING (IBL)

Interest in creating efficient cyberdefense methods utilizing game theory and machine learning (ML) techniques has grown as cybercrime has become more prevalent. Deception (i.e., intentional acts made to induce attackers to perform, or not take, specific actions Cohen, 1998) is one method of cyberdefense. Masking is a cyber deception tactic used to mask network characteristics and hide information that an attacker may seize during the reconnaissance phase. The majority of masking strategy research to now has either been theoretical or has only been tested in simulations. Therefore, it is uncertain if such defense techniques would work in real-world situations against human attackers. In a recent investigation, we actually discovered that a masking technique that in principle seemed to work well against human assailants was no more effective than a random camouflage technique. The assumption that these algorithms make regarding the "rationality" of human attackers may be one explanation for the present outcomes of masking tactics. Humans can only be boundedly reasonable since they generally have several cognitive limitations. Humans have limited memory capacity and absorb information sequentially, which frequently leads to biases in judgment. Attackers may be susceptible to these biases and err in ways that affect cybersecurity. As an example, they utilized opposing human variables to take advantage of prejudices and shortcomings connected to poor attention to stop cyberattacks. biases of many types, including the illusion of control, the sunk cost fallacy, illogical escalation, and attentional tunneling, have been identified. Similar decision-making biases including anchoring bias, confirmation bias, and take-the-best heuristic bias have been seen among cybersecurity specialists. Unfortunately, existing defense algorithms don't take into account the biases produced by human memory and instead disregard them. Cyberdefense algorithms also don't take into account defender biases, which might hinder their ability to protect themselves. The impact of gain and loss framing biases on defenders' actions was shown in a network defense scenario. Defenders who started out using gain framing—that is, with a network already in quarantine—used a quarantine approach more akin to those who started out using loss framing. How to lessen these biases in defenders has not received much attention up to this point. On the attacker's side, it has been shown how defense algorithms can exploit biases (such as confirmation bias) in human attackers by using cognitive models that computationally mimic the attacker's decision-making process. With the aid of a straightforward job, it has been demonstrated that it is feasible to inform the defense algorithms on the attacker's behavior and enhance the game theory/ML algorithms by making them more responsive to the specific attacker's activities. Researchers have proposed a framework for the creation and use of game-theory defense algorithms on experimental testbeds. The effectiveness (i.e., utility of the defender) of defense algorithms is assessed using an experimental testbed with human participants (e.g., attackers). Importantly, game-theoretic algorithms for adaptive defense are informed by cognitive models that simulate human decision-making. This broad concept of adaptive cyberdefense is built on cognitive models to show how personalized and adaptive signals are generated in a straightforward insider assault scenario. The fact that game-theoretic/ML algorithms frequently rely on vast volumes of data to match the model parameters presents one of their obstacles. It typically takes human interaction and the collection of a significant amount of human decision-making data in fields like cybersecurity to fully understand how various defense algorithms might operate in real-world circumstances. These interventions are unfortunately quite difficult. In applications where predictive models of human decision-making assume the function of humans in the work, cognitive models are generally beginning to play a direct role. For many years, IBL models have been used in a variety of fields, such as repeated binary choice decisions, multi-choice sequential decisions, predictions of human reliance on automation, predictions of human Theory of Mind in gridworlds, and predictions of cognitive biases in human decision making (such as confirmation bias, anchoring and adjustment, probability matching, and base rate neglect). IBL models have been frequently utilized in the field of cybersecurity to simulate how people make decisions in a range of activities requiring deception in insider attack games, intrusion detection systems, and susceptibility to phishing emails. Nevertheless, despite their success, current IBL models of human attackers frequently have them perform relatively easy tasks that abstract the complexity of cyber-related circumstances. Additionally, such challenges need frequent attacker-defender interactions since doing so enables IBL models to collect experiential learning and produce more precise predictions. When conducting network reconnaissance, Thakoor suggested using a Risk-Based Cyber Camouflage Game (also known as a masking algorithm) to change the answers given to attackers' questions. Their algorithm is based on a general sum Stackelberg game model, in which the attacker scans the network and selects a system to attack based on the system's responses, and the defender configures the network with a deception strategy (i.e., how the system should respond to scan queries from an attacker). In this case, the gains for the attackers and the losses for the defenders could differ. In order to determine the deception approach that maximizes utility, the masking algorithm evaluates the worst-case scenario for a risk-averse attacker (i.e., computes the minimal utility that a given deception method would generate). The authors demonstrate the NP hardness of this problem and offer a mixed-integer linear programme to calculate the best answer. Each computer in a network has a unique True Configuration (TC), which reflects its unique characteristics and vulnerabilities. To make the Observed Configuration (OC) from the attacker's point of view considerably different from the Technical Configuration (TC) of a computer, the defender works to obscure the properties. Any machine with TC has associated values that, if successfully attacked, benefit the attacker and harm the defense, respectively. Due to the sequential structure of the choices, the interaction between the attacker and the defense is modeled as a Stackelberg Security Game (SSG). The leader who is aware of the real status of the network, or the number of TC machines, is the defender. With this knowledge, the defender masks the TCs with OCs. This assignment approach is expressed as an integer matrix, with each entry indicating the number of machines with TC that are covered up by OC. Using these techniques is subject to the following domain constraints: There are two limitations on masking any TC with an OC: 1) a feasibility restriction (i.e., some OCs cannot realistically mask with particular TCs); and 2) a budget for the defender. A defense strategy is created within these limitations. A logical attacker engages in an attack on a pair that maximizes predicted utility. Weak Stackelberg Equilibria result from the constraint to a pure strategy, which forces the defender to take into account the worst-case tie-breaking for the attacker in the event of indifference. Therefore, the defense tries to select a tactic to maximize utility, supposing a rational attacker who maximizes utility. This tactic, which presupposes rational attackers, is known as the WSE Model. According to prospect theory, risk-averse attackers make judgements based on a value transformation function that is monotonic rising and concave. Any reward (namely, the value of the targeted computer) is seen as. With capturing the attacker's risk aversion, and, an appropriate constant, is a typical parametric form suggested in literature. It is difficult to learn the parameters. This may be achieved by gathering attacker answers to methods that were produced randomly, then generating a maximum probability estimate of the supplied instances. Once calculated, the defender determines the risk-averse attacker's best course of action by altering the WSE method and substituting the modified values for the valuations. This tactic is known as the Prospect Theory (PT) Model. The resemblance between the current circumstance and the precedent cases stored in memory is computed whenever a new judgment has to be taken. The model calculates an estimated utility for each choice option by blending the average of previous results weighted by how likely they are to be remembered in the future. Calculating the likelihood of memory retrieval involves comparing an instance's memory activation to that of all other instances stored in memory. The ACT-R cognitive architecture has a formal definition for the idea of an instance being activated. The activation is influenced by the contextual resemblance to prior instances, the frequency of encountering comparable cases, and the recentness of a prior occurrence. This may be achieved by gathering attacker answers to methods that were produced randomly, then generating a maximum probability estimate of the supplied instances. Once calculated, the defender determines the risk-averse attacker's best course of action by altering the WSE method and substituting the modified values for the valuations. This tactic is known as the Prospect Theory (PT) Model. The model is initialized with instances from the practice round that correspond to either successful attacks or unsuccessful attacks to start the job. These initial examples depict the payment expectations that human participants are probably going to pick up during the practice round and employ in the actual task rounds. The model initially processes each of the task's 10 rounds' worth of choice alternatives from the matrices. The blending method is used by the attacker's model to determine the predicted utility for each of the choice possibilities. Each of these examples of the alternatives considered, together with their blended values, are stored in the model. By analysing the information provided in the form of the matrix, this technique simulates how people would scan various devices during the exploration phase and generate expectations. The option with the greatest blended value is the one the model chooses to attack after calculating the expected utility of each alternative. The chosen choice and the actual result are then retained in memory. For each of the task's ten rounds, the exploration and exploitation procedure is repeated. By simulating each individual attacker using this IBL model, we were able to compare the model's performance to the actual outcomes of the two experimental situations, WSE and PT, as previously mentioned. For the CyberVAN experiment, the IBL model mentioned above was run 1500 times in each condition to get reliable estimations of the participants' performance. Although the 1500 agents running in this simulation contribute to produce "stable" predictions for the model, the stochastic nature of the model causes the data they create to vary much like human data does. The model (i.e., agent) underwent the identical process as each human participant throughout each run. The results of the model's simulation versus actual data under the WSE and PT conditions are then shown. Overall and for each of the matrices, the model forecasts more defender losses in the WSE than the PT defense plan. Note that we normalize the defender's loss between 0 and 1 before computing the RMSEs for the defender's loss. For the WSE and PT tactics, the corresponding RMSE values for the defender's losses are 0.076 and 0.069, respectively. The IBL model can forecast the defender's losses properly in the majority of matrices, according to the RMSE values for defender losses in both the WSE and PT algorithms. Aggarwal et al. (2020b) used human tests to show that human attackers exhibit a risk-aversion bias while deciding whether to launch a cyberattack. Thakoor et al. (2020) used Prospect Theory (PT) to create a masking approach in order to take advantage of the risk-aversion bias of human attackers. In particular, Thakoor et al. (2020) created two masking algorithms: the first approach (WSE) assumes full rationality while the second method (PT) takes advantage of restricted rationality in the form of risk aversion. In this study, we conduct an experiment to investigate the effectiveness of PT and WSE masking tactics against human attackers.

## 3. MARKOV SECURITY GAMES

The usage of the Internet has lately seen a boom, and it is now prevalent across all socioeconomic strata. As the Internet has expanded, it has become more challenging to prevent unauthorized access to online data. Hackers, or those who attack computer networks, are always developing new techniques for taking advantage of flaws in computer systems. Security analysts, who work to secure computer systems, may apply software updates to close holes and defend against cyber-attacks. These software updates may be successful in removing vulnerabilities that are present in computer systems. These software updates might, however, be less successful as they could only partially fix an existing vulnerability or introduce a brand-new one. Consequently, a less effective software patch may only address vulnerabilities in a discrete area of the computer system. Previous studies have suggested that game theory is a useful tool for studying. human judgment in cyberattack circumstances. According to earlier studies, the impact of the efficiency of the patching operations on Markov security games may be used to study cyber decision-making. Human players take on the roles of hackers and analysts in the Markov security game, where hackers can take attack or not-attack actions and analysts can take defend (patch) or not-defend (not-patch) activities. Both participants may get payoffs (outcomes) as a result of their activities, and the interaction between hackers and analysts occurs repeatedly during the course of a game. According to the Markov assumption, an analyst's decision in the previous round affects how vulnerable the computer system is in the current round to an attack. Most of the time, patching vulnerabilities can increase a computer system's security (i.e., make it less susceptible to cyberattacks); however, in some circumstances, patching vulnerabilities can also result in unresolved vulnerabilities (i.e., patching may be ineffective, leaving the computer system open to attacks). In a preliminary study, Markov security game decision-making was examined. In the absence of fixes, cyberattacks may result in damages that grow problematic as the attacks spread throughout computer systems, according to the cited source. Conversely, when analysts are able to promptly fix existing vulnerabilities in computer systems, damages to those systems are reduced. These results are consistent with the dynamics of Markov security games. Researchers have used mathematical simulation approaches to make predictions about the Nash equilibria using Markov security games; however, these authors did not conduct an empirical analysis of human activities in relation to Nash predictions. A recurring 2 × 2 zero-sum game is the Markov security game. In this game, a hacker and an analyst compete against one another. The goal for both opponents is to maximize individual payoffs by continually selecting choices over a number of rounds (both opponents are unaware of the final destination). The analyst can take a defend (d) and a not-defend (nd) action, but the hacker can only take an attack (a) and a not-attack (na) action. Defend activities include correcting computer system vulnerabilities, whereas attack actions involve assaulting a computer system. When playing the game against another human, one human player is chosen at random to play the hacker, and the other human player is chosen to play the analyst. Between two succeeding rounds, the transition from state v to state nv or from state nv to state v depends on how well the patching procedure worked. If the patching procedure is successful, the likelihood of transitioning from state nv to state v is low (= 0.1) while the probability of transitioning from state v to state nv is high (= 0.8). If the patching is less successful, the likelihood of transitioning from state nv to state v and from state v to state nv is equal (= 0.5). People often want to maximise their perceived reward across acts, according to IBLT. In IBLT, the blended values calculated for various activities are what define the apparent payoffs. These players would likely have distinct perceived payoffs in both scenarios since their opponents acting as hackers and analysts would have different payoffs under effective and ineffective patching situations. We also anticipate variations in cognitive characteristics relating to dependence upon recency and frequency of outcomes, attention to opponent's activities, and cognitive noise across different effective and less-effective patching settings based on IBLT. Furthermore, IBLT predicts that human judgements will dramatically diverge from their Nash proportions under various patching settings. Most of the time, patching vulnerabilities can increase a computer system's security (i.e., make it less susceptible to cyberattacks); however, in some circumstances, patching vulnerabilities can also result in unresolved vulnerabilities (i.e., patching may be ineffective, leaving the computer system open to attacks). In a preliminary study, Markov security game decision-making was examined. In the absence of fixes, cyberattacks may result in damages that grow problematic as the attacks spread throughout computer systems, according to the cited source. Conversely, when analysts are able to promptly fix existing vulnerabilities in computer systems, damages to those systems are reduced. These results are consistent with the dynamics of Markov security games. To interpret human behaviour in the Markov security game, Researchers created a model based on IBLT. A circumstance in a task (a collection of qualities that characterise the decision situation), a decision in a task, and an outcome as a result of that decision in that situation make up an instance, or the smallest unit of experience, in the IBL model. Through a generic decision-making process, several components of an instance are built: a scenario is created from task attributes, a choice and anticipation of an outcome are made while making a judgment, and the outcome is updated in the feedback stage when the actual outcome is known. Instances are utilized frequently to make choices in the IBL model after they accumulate over time in memory and are retrieved from memory. A statistical method known as activation, which was first used in the ACT-R cognitive architecture, is used to gauge this availability. By only enabling two single-person models to continually engage with each other in the game, we may construct our model for two-player security games. Each instance in the IBL model is made up of a label that designates the choice each participant was given—for the hacker, it was to attack or not to attack, and for the analyst, it was to defend or not to defend—as well as the result that was reached. The structure of an instance for both players is just (alternative, result), as the circumstance is constant for each binary choice. Calculating the blended value of the choices is the first step in the process of choosing decision alternatives in the model for each round t in the game. The choice that has the highest blended value is then made. The likelihood of retrieving examples from memory that correspond to those outcomes and the outcomes happening in the option determine an option's blended value. Furthermore, the likelihood of retrieving instances from memory depends on how recently and frequently those instances have been retrieved from memory, which determines how likely it is that those instances will be retrieved. Researchers implemented two models, (1) Calibrated model, in which the values of the free parameters were determined by calibration using a genetic algorithm; (2) ACT-R model, in which the model's free parameters were set to their default values according to ACT-R. In the Markov security game, two identical IBL model agents played participants for 50 rounds under two distinct settings, just as two human participants would have. A trial's outcomes were decided by the choices made by both agents, who employed independent blending and activation methods with different sets of parameters, in each of the model settings. The three unrestricted parameters for each agent were: w attentiveness to opponent's activities, s noise, and d decay. Instances are utilised frequently to make choices in the IBL model after they accumulate over time in memory and are retrieved from memory. A statistical method known as activation, which was first used in the ACTR cognitive architecture, is used to gauge this availability. By only enabling two single-person models to continually engage with each other in the game, we may construct our model for two-player security games. Each instance in the IBL model is made up of a label that designates the choice each participant was given—for the hacker, it was to attack or not to attack, and for the analyst, it was to defend or not to defend—as well as the result that was reached. The structure of an instance for both players is just (alternative, result), as the circumstance is constant for each binary choice. Calculating the blended value of the choices is the first step in the process of choosing decision alternatives in the model for each round t in the game. The choice that has the highest blended value is then made. The likelihood of retrieving examples from memory that correspond to those outcomes and the outcomes happening in the option determine an option's blended value. Furthermore, the likelihood of retrieving instances from memory depends on how recently and frequently those instances have been retrieved from memory, which determines how likely it is that those instances will be retrieved. There is an urgent need to fix systems since cyberattacks are increasing quickly. Computer systems include vulnerabilities. The vulnerability patching procedure might not be perfect, though. However, in other situations, patching may be less successful and leave computer systems open to assaults. In other scenarios, patching may be effective and make the computer systems less vulnerable to cyberattacks. Results showed that whether patching operations were effective and less-effective, the proportion of attack and defense activities was comparable. Additionally, in most situations and states, both players greatly diverged from their ideal Nash proportions. First, it was discovered that both patching situations had a comparable proportion of attack and defense activities. The closeness in payment magnitudes and valances between the two patching circumstances may be a contributing factor to this conclusion. IBLT claims that people maximize their perceived reward across acts, as was already noted. Participants acting as hackers and analysts saw comparable payoffs under various patching situations, hence they probably had comparable perceived payoffs in both circumstances. The proportion of attack and defend activities, we discovered, greatly differed from their Nash proportions. Once more, the IBLT may be used to illustrate this expectation. Human participants, according to IBLT, have cognitive constraints on memory and recall processes, and people frequently base their judgements on the recentness and frequency of results. It seems that our experiment's dependence on recency and frequency processes prevented participants from developing the best Nash expectations for their behavior, leading them to dramatically diverge from Nash proportions in a number of circumstances and states.

## 4. WORKPLACE

The unique Cybonto conceptual framework seeks to offer broad guidelines for responding to the concerns raised about the vision of DTs and HDTs for cybersecurity. The framework uses an HDT within a DT system to simulate the malicious actor's thought process. The Cybonto ontology's behavioral/cognitive component serves as the definition of cognitive space. The HDT's repertoire of stored actions, its capacity for creating original movements, and the interaction interfaces of the other DTs all serve to constrain the action space. From the body of research in behavioural and cognitive psychology, fifty ideas were first chosen. The research potential, applicability to criminology and cybersecurity, and coherence of each theory were taken into account while ranking them. Each hypothesis was then formalised as tuples consisting of an entity, a "influence" connection, and an entity. Analysing the Cybonto ontology led to the creation of the Cybonto conceptual framework. Each DT's internal environment (INE) is personal. It consists of both cognitive and non-cognitive elements. The social environment (SOE), in contrast to the interior environment, is a public space. The in-group environment (IGE) bridges the gap between INE and SOE. Bronfenbrenner's Ecological System Theory, which defines influences as progressive, variable, and reciprocal forces among people and surroundings, governs all ecosystems. For instance, an apparently distant public event could yet have an impact on specific internal brain functions. The intended HDT is related to the IEG and the SOE. The micro- and mesosystems of Bronfenbrenner are analogous to the IEG. The microsystem, which includes members like family, close friends, school, lovers, and mentors, is the most influencing external environment. The Exo-, Macro-, and Chrono-systems of Bronfenbrenner are the equivalent of SOE. Four individuals from four DT groups are required to participate in the Cybonto conceptual framework. Both an attacker and a defender HDT are required. An attacker HDT must gather the data on its own, unlike older models, to which data and feature specifications were explicitly given. If the essential requirements of the group are not satisfied, group-related facts cannot be derived. Thus, to present IEG and SOE IDs, we then require at least two additional DTs. An HDT is capable of two different sorts of behaviours: those that create or modify artefacts and those that do not. A complicated noncognitive digital twin can be considered an artefact, as can a single line of code. Viewing a malware's source code is a non-artifact behaviour, however running the code, if it modifies other artefacts, may be an artifactalteringbehaviour. The internal and exterior environments (IEG and SOE) and the perceptual layer are in close proximity to one another. The identical data streams will be seen differently by various perceptual layers working together with various cognitive systems. Only a small portion of a digital cognitive system is made up of refined perceptions. Numerous cognitive pathways for handling initial impressions are described in great detail by the Cybonto ontology. Either a non-artifact behaviour or an artefact creating/altering behaviour is the outcome of a cognitive processing chain. The behaviours (data streams) are monitored by other HDTs, and a new cycle of feedback loops starts. It is important to remember that an in-group setting might allow for the concealment of a behaviour. Out of more than thirty options, Cybonto selected the Basic Formal Ontology (BFO) as its top-level ontology. Out of more than thirty options, Cybonto selected the Basic Formal Ontology (BFO) as its top-level ontology because it is the only top-level ontology that rejects materialism and commits to real-world possibilities. The only top-level ontology that embraces materialism, affirms real-world possibilities, and possesses an intensional identity requirement is BFO. By using the Mental Functioning (MF) as its mid-level ontology, the Cybonto Core (the behavioral/cognitive component) is deeper grounded. MF adheres to the OBO Foundry's recommended practices and coordinates with other initiatives in the Cognitive Atlas, a cutting-edge collaborative information source for cognitive science. A fundamental tenet of DT tactics, materialism sees the world as a collection of materialized objects existing in space and time. Cybonto's representation of mental structures is fundamentally different when one commits to materialism through BFO. Cognitive processes were long thought to be abstract details that could only be expressed through language. The majority of behavioral components in cybersecurity ontologies are language-based due to this heritage. Measurements of brain processes that correlate to certain cognitive conceptions are now possible because of recent advancements in the brain-machine interface, such as those made by Neuralink. As a result, behavioral and cognitive ontologies may now be grounded in materialism. Conceptual objects, various language descriptions of the same actual things, process-based objects, and qualitative object labels that are impossible to assess in the real world are all rejected by Cybonto. A greater link value is shown by a link color that is darker. Automatically, nodes were placed in a multi-circle structure with nodes closer to the center having a greater betweenness centrality. Top Authority Central (AC) constructions are influenced by those that have the most sway over other constructs. The constructions with the shortest pathways among other constructs are known as top Betweenness Central (BC) constructs. BC constructions have the potential to act as both gateways and bridges to other constructs and processes. Top Eigenvector Central (EC) structures act as the clique's ringleaders. A collection of constructs known as a clique consists of individuals who are connected to one another. A clique may stand for a potent cognitive/behavioral pattern in the context of the cognitive digital twin. The top EC constructions have connections with members of other cliques in addition to their own clique members. When the neighbours at one end of a link are the most dissimilar from the neighbours at the other end, the link has the highest contribution weight. Top Incoming Central (IC) constructions are impacted by the most significant in-coming neighbours, whilst top Out-degree Central (OC) constructs have the greatest out-links (influencing) to others. The most influential neighbours are connected to the top PageRank constructions in some way, either through inbound or outbound interactions. Behaviour, Arousal, Goals, Perception, Self-efficacy, Circumstances, Evaluating, Behavior-Controlabiity, Knowledge, and Intentional Modality are the top 10 components across 20 network centrality measurements. Only the items on this list—although some are implicitly implemented—are components of current digital cognitive architectures: Behaviours, Goals, Perception, Evaluating, and Knowledge. Before this study, significant cognitive structures may have been examined separately for different usecases and hence were unable to be jointly brought to the notice of traditional cognitive system designers. Now that we have a broad perspective of 20 behavioural theories, we may give these top 10 components more attention. We may think about implementing Goals, Knowledge, Perception, and Evaluating explicitly and at a finer granularity within cognitive architectures. For instance, experience extends beyond fleeting sensory perception. For instance, whether Bob is with Alice or not, Alice continues to think of Bob as a decent person. We should also think about include Arousal and Intentional Modality. Despite being a non-cognitive construct, arousal is placed second and has an impact on several of the top 10 cognitive constructs, including evaluating and intentional modality. Sadly, there is currently very little study on arousal as a component of the digital cognitive process. There are a few studies looking at the consequences of general emotions, according to SOAR-related research results. There are just four studies in the ACT-R research library that examine how arousal affects memory management. Another non-cognitive concept that significantly affects behavioural outcomes is the Circumstance. The research recommends incorporating non-physical environment factors including urgency, group dynamics, and social attitudes into the current Mental Image module in existing cognitive architectures.

## 5. ANALYSIS

One of their challenges comes from the fact that game-theoretic/ML algorithms typically rely on enormous amounts of data to match the model parameters. To completely comprehend how different defence algorithms could function in practical situations, it often involves human involvement and the gathering of a sizable quantity of data on human decision-making in sectors like cybersecurity. These therapies, regrettably, are exceedingly challenging. Cognitive models are often starting to play a more direct part in applications where human decision-making prediction models take over the job of people in the task. Given the tendency for people to choose the safe choice, it is crucial to note that the PT method would attempt to avoid creating matrices in which a true configuration would perfectly match an observable configuration. The only instance where the PT method yielded a greater defense loss and more attacker success than the WSE approach, according to our observations, was when the matrix included a single sure choice. This again implies that humans are unable to avoid such bias towards certainty and that a revision to the PT algorithm would be necessary to ensure that such instances are avoided. Our human experiment is underpowered due to the small number of participants, which may restrict the conclusions we can draw from it. It is often difficult to find participants with specialized knowledge, such as cybersecurity experience, and gathering 45 individuals required a large effort in data collection. We can simulate many more participants in each condition thanks to the IBL model's precise duplication of the choices. The masking techniques are capable of adding network limitations that are relevant in other realistic circumstances, even if the algorithm and the tests in this research have been carried out for a small number of nodes and straightforward network architectures. Similar to this, IBL models could also be able to adapt to various cybersecurity circumstances, albeit there may be a run-time restriction, which might be a bottleneck for big systems. We created a knowledge of how human attackers make judgements through trials and simulations. Attackers don't make rational decisions; instead, they follow biases like certainty and risk aversion. Human attackers make less-than-optimal judgements, diverging from the anticipated ideal activities that certain defense systems presume. Defense algorithms may lessen overall losses from cyberattacks if they are created to take advantage of such biases in attacker decision-making. For the Markov Security Games, Researchers discovered that the calibrated IBL model outperformed the ACT-R model in accounting for human judgements under both patching situations. The fact that the model appears to be unable to collect human data when using the default ACT-R settings is one plausible explanation. However, recalibrating these parameters significantly enhanced the model's own performance. First, given our findings, we predict that analysts will continue to overpatch computer systems in the real world, regardless of how effective and optimum these patching choices are. Second, it appears that hackers do not consider whether computer systems have been adequately patched while targeting networks. Hackers are concerned about how susceptible computer systems are to their attacks, though. Therefore, this sense of vulnerability is likely to have an impact on hackers' choices about cyberattacks. It can be crucial to present computer networks as less susceptible to cyberattacks in the real world. There are several ways to achieve this, including through social media, publications, reports, and multimedia. Additionally, IBLT-based models might be applied to account for cyber choices. The hacker models, for instance, might be used to mimic hacker choices about updates and vulnerabilities. Analyzer models may also be used to defend against various cyberattacks by automating the patching procedures. The cognitive aspects of the patching procedures may also be evaluated with the use of this research. Data was first gathered from people with degrees in computer science and backgrounds in the STEM fields. These individuals could still vary from hackers in the real world, though. Second, based on the hacker's view of the system's possible vulnerability, this study made the basic assumption that the hacker would continually choose to either attack or refrain from attacking the system. The analyst had a choice as to whether to fix the system or not in the meantime. As a result, the experiment in this study had a simple architecture and might not have been totally compatible with other automated hacker operating systems. For instance, in opportunistic/light touch attacks, a hacker may try to take advantage of a weakness against a large number of systems. These assaults could be widespread and they might not entail individual system-level decisions made by humans. Similar to this, there may be several weaknesses in focused assaults. These assaults could be widespread, and they might not entail individual system-by-system decisions made by humans. Similar to targeted assaults, a single system may be subjected to several vulnerability tests. For instance, in targeted assaults, the hacker may progressively test a large library of known exploits for a certain system of interest. Once more, targeted assaults could be launched using automated procedures rather than ones that need human judgment at every stage. As a result, they might not follow the hacker's modus operandi as intended. In this situation, the suggested analyst's model may aid in understanding their cognitive and decision-making processes against such automated attacks.

## 6. CONCLUSION

Overall and for each of the matrices, the model forecasts more defender losses in the WSE than the PT defense plan. Note that we normalize the defender's loss between 0 and 1 before computing the RMSEs for the defender's loss. For the WSE and PT tactics, the corresponding RMSE values for the defender's losses are 0.076 and 0.069, respectively. The IBL model can forecast the defender's losses properly in the majority of matrices, according to the RMSE values for defender losses in both the WSE and PT algorithms. The IBL model's results showed that it can, on average, forecast defenders' success and loss rates under both PT and WSE scenarios. A model that performs well on average may, however, be unable to account for the heterogeneity in individual decisions. In this part, we compare the human data to the distribution of the individual machine selection predicted by the IBL model. In both the WSE and PT techniques, the IBL model can typically describe the distributions of human preferences. The IBL model captures both the selection behavior under both situations as well as human behavior at the aggregate level. The noise in the IBL model might be the cause of some of the variations that are seen in the figure.

## 7. FUTURE DIRECTIONS

According to studies, the sense of vulnerability affected hackers' choices, and there are a variety of techniques to mould this view. For instance, one may use honeypots, which are easily attackable systems, to deceive computer networks. The second option is to use intrusion-detection systems (IDSs) and inform hackers of their presence and precision. For instance, if hackers are informed that IDSs are absent or that they are there but are less effective, this knowledge is likely to affect how vulnerable the network is to their assaults. Again, in this instance, the IDSs can be useful in forcing hackers to target specific systems (like honeypots) over others and in making them fall victim to such assaults. Hackers and analysts might not have access to knowledge regarding an opponent's behavior in the real world. Therefore, it would be fascinating to explore how hackers' and analysts' judgements are impacted by the availability and lack of knowledge on adversaries' actions. Along with these concepts, we also want to look at how other cognitive processes, such as similarity, spreading activation, and cognitive inertia, play a part in our models. Additionally, this research made a practice-unlike assumption that the analyst would have complete freedom in trying to optimize the application of fixes. For instance, in the real world, the majority of patching may be controlled by an operational policy (such as the Common Vulnerability Scoring System) that allows analysts to deploy patches addressing vulnerabilities of a specific severity within a specified time frame. Therefore, vital patches can be applied quickly, but non-critical fixes can take a while to apply. Therefore, scenarios where analysts choose between essential and less-critical fixes may be tested in future research, and the success of patching may apply to both types of patches. We will start working on some of these concepts right away as part of our current research programme in game theory and cyber-security.

### Conflict of interest statement

Authors declare that they do not have any conflict of interest.

### REFERENCES

[1] P. D. Bruza, Z. Wang and J. R. Busemeyer, "Quantum cognition: A new theoretical approach to psychology", Trends Cogn. Sci., vol. 19, no. 7, pp. 383-393, 2015.

[2] J. R. Busemeyer and P. D. Bruza, Quantum Models of Cognition and Decision, Cambridge, UK:Cambridge Univ. Press, 2014.

[3] J. R. Busemeyer and Z. Wang, "Hilbert space multidimensional theory", Psycholo. Rev., vol. 125, no. 4, pp. 572-591, 2018.

[4] R. I. G. Hughes, The Structure and Interpretation of Quantum Mechanics, Cambridge, MA, USA:Harvard Univ. Press, 1989

[5] A. N. Kolmogorov, Foundations of the Theory of Probability, Vermont, United States:Chelsea Publishing, 1956

[6] J. von Neumann, Mathematical Foundations of Quantum Mechanics, Princeton, New Jersey, USA:Princeton Univ. Press, 1955.

[7] J. von Neumann and O. Morgenstern, Theory of Games and Economic Behavior, Princeton, New Jersey, USA:Princeton Univ. Press, 1953

[8] A. Tversky and D. Kahneman, "Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment", Psychol. Rev., vol. 90, no. 4, pp. 293-315, 1983.

[9] H. J. Wilson, "The Cognitive Usefulness of the Internet of Things", Harvard Business Review, Nov. 2014.

[10] H. Abie and I. Balasingham, "Risk-Based Adaptive Security for Smart IoT in eHealth", BODYNETS, vol. 2012, pp. 269-275, 2012

[11] .D. Miessler, "HP Study Reveals 70 Percent of Internet of Things Devices Vulnerable to Attack", HP Fortify, July 2014.

[12] A. Kott, "Towards fundamental science of cyber security" in Network Sci. Cybersecur., Berlin, Germany:Springer, pp. 1-13, 2014.

[13] "Cybersecurity in healthcare: A narrative review of trends threats and ways forward", Maturitas, pp. 113, 2018.

[14] .Karen Zita Haigh and Craig Partridge, Can artificial intelligence meet the cognitive networking challenge?, 2011, [online] Available: http://www.cs.cmu.edu/~khaigh/2011-haigh-EURASIP-JWCN.pdf.

[15] W.J. Chisum and B.E. Turvey, "Evidence dynamics: Locard's exchange principle & crime reconstruction", Journal of Behavioral Profiling, vol. 1, no. 1, 2000.

[16] V. Lenders, A. Tanner and A. Blarer, "Gaining an Edge in Cyberspace with Advanced Situational Awareness", IEEE Security & Privacy, vol. 13, no. 2, pp. 65-74, 2015.

[17] A. Yelizarov and D. Gamayunov, "Adaptive Visualization Interface That Manages User's Cognitive Load Based on Interaction Characteristics", Proc. of VINCI, pp. 1-8, 2014.

[18] J. Cho et al., "Stram: Measuring the trustworthiness of computer-based systems", ACM Computing Survey, 2019

[19] F. Ullah and M.A Babar, "Architectural Tactics for Big Data Cybersecurity Analytic Systems: A Review", Arxiv.Org., pp. 1-48, 2018.

[20] Sandeep Narayanan et al., "Cognitive Techniques for Early Detection of Cybersecurity Events", Aug 2018

[21] Arild B. Torjusen, Habtamu Abie, Ebenezer Paintsil, Denis Trcek and ÅsmundSkomedal, "Towards Run-Time Verification of Adaptive Security for IoT in eHealth", Proceedings of ECSAW '14, pp. 8, 2014.

[22] J. Cho et al., "Stram: Measuring the trustworthiness of computer-based systems", ACM Computing Survey, 2019.

[23] Abbasi, Y. D., Short, M., Sinha, A., Sintov, N., Zhang, C., and Tambe, M. (2015). "Human adversaries in opportunistic crime security games: evaluating competing bounded rationality models," in Third Annual Conference on Advances in Cognitive Systems ACS-2015. Atlanta, GA.

[24] Aggarwal, P., Gonzalez, C., and Dutt, V. (2016). "Cyber-security: role of deception in cyber-attack detection," in Advances in Intelligent Systems and Computing, vol. 501 (Orlando, FL: Springer), 85–96

[25] Alpcan, T., and Başar, T. (2010). Network Security: A Decision and Game-Theoretic Approach. Cambridge, UK: Cambridge University Press.