Simplifying Cyber Security Maturity Models through National Culture: A Fuzzy Logic Approach

Jongkil Jay Jeong
Deakin University & Cyber Security CRC
jay.jeong@deakin.edu.au

Marthie Grobler
CSIRO's Data61
marthie.grobler@data61.csiro.au

M.A.P. Chamikara
CSIRO's Data61 & Cyber Security CRC
chamikara.arachchige@data61.csiro.au

Abstract

Different assessment models exist to measure a country's cyber security maturity levels. These levels serve as a benchmark for indicating how well prepared a nation is against a cyber security attack and how resilient it would be in recovering from such an attack. However, results from these maturity assessments are either too general, overly complex, or resource intensive to apply and guide important national cyber security strategies and frameworks. To address this we propose a model to link national culture with a country's cyber security maturity through fuzzy logic mapping to ensure that a more uniform reflection of the cyber security maturity level within a country can be measured. In this paper, we present additional research towards optimising our model. The extended model incorporates input from two cyber security assessment models, and validates the refined output models on 11 countries to compare the maturity levels from the traditional assessment model with our optimised fuzzy model. Our results show that it is viable to reduce the resources required to conduct a national cyber security maturity assessment.

1. Introduction

As a mechanism for countries to measure, analyse and improve their cyber security practices, different Cyber security Maturity Assessment (CMA) models have been developed. These models are used to assess a country's cyber security readiness and resilience across multiple facets of a nation's digital capacity, and provide nations the ability to establish a benchmark across these facets to enable better facilitation of targeted interventions to the most vulnerable aspects of a nation's cyber presence [1]. As such, a number of different CMAs have been developed over the past decade, each with its own assessment criteria, scoping boundaries, and ways to present cyber security maturity ranking.

Two of the most widely adopted CMAs are the Global Cybersecurity Index (GCI) and the Cybersecurity Maturity Model (CMM), with 134 countries participating in the GCI[1] and 70 countries carrying out the CMM assessment[2] to date. Although these two CMAs measure similar aspects of a country's overall cyber security maturity level, it is unclear whether the recommendations provided from these models are comparable, since the process, methodology, and data behind each method differ. Considering that the results from CMAs have a direct impact on a nation's cyber security strategy, [2], it is important to investigate if the recommendations derived from these CMAs are similar or provide mixed results.

Therefore, this study conducts a cross-comparison of results between the GCA and CMM on similar criteria to ensure that the recommendations provided to governments can be considered solid, accurate, and optimised in terms of available data and assessment methodology. We do this by drawing upon a prior study in which we determined an indicative level for GCI based on certain dimensions of national culture using fuzzy logic [3]. Fuzzy logic enables the mapping of the impreciseness of certain concepts with precise logic; hence, fuzzy logic is used to solve many complex real-world problems that involve complex human-specific dynamics (e.g. voice recognition) [4]. This property of fuzzy logic allowed the linkage between national culture with a country's cyber security maturity through fuzzy relationships. By comparing the outcomes from the GCI and CMM, based on our tests incorporating national culture elements with fuzzy logic, the aim of this study is to address the following research questions:

- **RQ1:** How can we refine and improve the existing fuzzy logic model by applying data from different CMA models?
- **RQ2:** What differences exist when comparing the results between CMA models?

The background literature section provides oversight into the dimensions of national culture and its impact

[1] ITU Publications. 2019. https://www.itu.int/en/ITU-D/Cybersecurity/Documents/draft-18-00706_Global-Cybersecurity-Index-EV5_print_2.pdf

[2] University of Oxford. 2019. https://www.sbs.ox.ac.uk/cybersecurity-capacity/content/gcscc-global-impact-publication

URI: https://hdl.handle.net/10125/80000
978-0-9981331-5-7

on cyber security, and the details of CMAs selected to refine the fuzzy logic model. We present the research model and discuss the data sources that were used in the fuzzy logic model optimisation. We then provide a detailed case study country overview, with data across all relevant data sources. We further detail the data analysis and provide our experimentation results and conclude by highlighting the key implications from our study.

2. Background Literature

This section details the existing links between national culture and cyber security, as well as background information on selected CMA models.

2.1. National Culture and Maturity

Many studies have explored how national culture has an effect on cyber security related activities [5]. National culture nests a subculture around cyber on the intentional and unintentional manner in which entities make use of cyber space, based on assumptions, beliefs, values, and knowledge of users in a digital environment [6]. As our focus is on CMAs at country level, we consider these macro-cultures to form the foundation of the fuzzy model optimisation that we perform in this research study.

On this foundational level, national culture mediates the level of cyber security capacity that a country has by directly influencing a dominant 'clan mentality' that integrates various economic and social aspects, which in turn determines the perceived norms, attitudes, and behaviours associated with cyber security [7]. Examples of how this impacts a country's overall cyber security maturity level can be observed through the average strength of passwords established by nationals, their general attitude, and perception towards security and privacy, and the uptake, and efficacy of cyber security training and education programs [8, 9]. Further studies suggest that national culture has a direct impact on the investment, development, and efficacy of systems utilised by a country to develop its cyber security capacity and maturity levels [10]. Moreover, national culture plays an important role in setting the landscape for cyber security related business and policy, which also influences the capacity for private and public entities to build, develop and invest in cyber security capacity building [8].

Significant work has been done in terms of the influence of a person's environment and exposure to the Internet on their online security behaviour [11]. For example, studies clearly show the link between poor security behaviours learned by students and carried with them through to the workforce [12]. A poor behavioural approach to password use in younger people would therefore form the foundation for adult password use and cultural adoption as people are found to retain fragments of previous habits, such as using a root word as basis for all passwords, that result in long-term and extended password reuse, i.e. development of national cyber culture [13].

In the context of our study, we consider social identity and cultural appropriation as an individual user's self-classification as identifying with or belonging to an online social group associated with a specific geographical region or country. Social identity as a driving force of human behavior research is a fairly new research domain, still to be fully explored [14].

2.2. Cyber Security Capacity Building

Although processes and technologies can be created to be theoretically secure, true security depends on the people involved in the implementation, application, and usage [15]. Therefore, research strongly suggests that the availability and uptake of cyber security related educational programs and awareness campaigns will have a significant impact on the overall cyber capacity of a country[2]. Cyber security related training has been known to mitigate the risk from potential threats, and lead to higher levels of compliance and less risky behaviour within the general population [16, 17]. This implies that the level of education and training received significantly boosts situational awareness and cyber security abilities.

Awareness campaigns and programs that are designed to improve compliance towards cyber security policies or ensure that the risks surrounding cyber crimes and attacks are fully apprehended can influence a country's maturity [18] because users have certain perceptions that can either positively or negatively impact the security process. It is especially cyber security misbehaviour that has a negative impact on cyber security culture, causing resistance to cyber security measures that can compromise its effectiveness and have an impact on a national level [19]. Finally, different groups of people and organisations would require specific cyber security training and interventions based on their own individual processes and needs [20, 21]. This implies that customised, systematic, and tailor-made training and education is an integral part of cyber security maturity.

2.3. CMA Models for Countries

A number of different CMAs exist to identify and assess the level of cyber security maturity and capacity through the structured collection of practices and processes across different areas [22], providing a benchmark for the current level of cyber security maturity within that specific area. We consider two CMAs developed explicitly to measure the maturity level for countries.

The *Global Cybersecurity Index (GCI)* is a composite index compiled by the International Telecommunication Union (ITU) as a means of measuring countries' commitment towards cyber security. It consists of five pillars – legal, technical, organisational, capacity building, and cooperation – to monitor and compare a country's cyber security commitment. Each pillar is divided into a number of indicators that collectively present a overview of the country's commitment to the pillar. The GCI is specifically aimed at helping countries identify areas of improvement by providing a global benchmark ranking[1].

The *Cyber Security Maturity Model (CMM)* is developed by the Global Cyber Security Capacity Centre to provide a comprehensive and nuanced understanding of the cyber security capacity landscape for countries[2]. It covers a broad expanse of areas that need to be considered when seeking to enhance cyber security capacity across five dimensions – cyber security policy and strategy; cyber culture and society; cyber security education, training and skills; legal and regulatory frameworks; and standards, organisations and technologies. Each dimension is divided into factors that describe what it means to possess cyber security capacity.

Although differences exist between the two CMAs, the structure of their models is largely similar. Both the GCI and CMM identify and measure the maturity and capacity pertaining to cyber security through a systematic collection of practices [1]. **Pillars / Dimensions** represent the key concepts and nuances of cyber security capacity across multiple facets of society. **Indicators / Factors** are the objectives that have to be fulfilled in each of the areas of the model and visualize the progress towards the objectives. **Score / Aspect** is the assessment based on the level of fulfilment for each indicator/factor within the pillars/dimensions of countries. These results determine the maturity level (ranging from 'start-up' to 'dynamic'), and the ability to rapidly adapt to changes in the cyber security landscape.

3. Research Methodology

We developed a fuzzy logic model to help scope the boundaries of and impact between cyber security maturity, national culture and login credential strength. Specifically, it enables the modelling of concrete logic onto the traditionally intangible concepts of cyber security maturity and national culture based on their fuzzy relationships. In developing the fuzzy logic CMA model, we used a custom login credentials dataset and the Hofstede Cultural Dimensions (HCD) model (both shown on the left of Fig. 1 as input to the process). This was to model the scope of national culture onto login strength, and login strength onto the Cyber Maturity Level (CML) surface to optimise the assessment of CML based only on a country's national culture values as input. To validate the model functionality, we make use of the GCI and CMM figures (shown at the top of Fig. 1) as inputs into the fuzzy logic model.

3.1. Data Sources

In an earlier exploratory study, we established a preliminary connection between national culture and a country's overall cyber security maturity level (CML). This connection was established through testing the hypothesis of a definitive link between national culture and real world login strength data within a country, and affirming the link between a country's real world login strength data and CML. We have now extended this mapping to include the GCI and CMM model datasets.

3.1.1. Meta Information and Data Preparation With the focus of our research on the national culture aspect of CMA, the Hofstede's Cultural Dimensions (HCD) dataset[3] provides the base culture data from 111 countries and sub cultural groupings. The original Hofstede model was developed in 1973 from a survey with more than 117,000 employees across 50 global offices [23]. Each entry of the cultural model normalised to 100, across its six cultural dimensions. The original login strength dataset contains 256 country profiles, and one profile for all emails with no distinctive country code (such as *gmail.com*). We obtained the GCI data directly from the ITU, creating a GCI dataset of 194 country profiles. To date, the CMM review has been carried out across 85 countries.

We used these country profiles in two ways, as shown in Fig. 1. In both instances, we used the ISO 3166 country code as the common denominator[4] to identify a new subset of country data. In ①, we created a subset of data with only the cultural model and login strength datasets. After the data cleaning, 65 country profiles were identified that had complete information from both datasets. These country profiles are considered in developing the fuzzy model to determine the existing linkage between national culture dimensions and login credential strength, In ②, we created a subset of data from all four datasets, excluding country code mappings that did not present an exact match and custom country subsets within the cultural model dataset based on cultural groupings (rather than geographical boundaries).

Of the initial country profiles across the four datasets, only 11 countries satisfied the criteria for non-null values (data sparseness) for all six HCD dimensions, non-null values in the CD dataset, and quantitative assessment

[3] https://geerthofstede.com/research-and-vsm/dimension-data-matrix/

[4] https://www.iso.org/iso-3166-country-codes.html

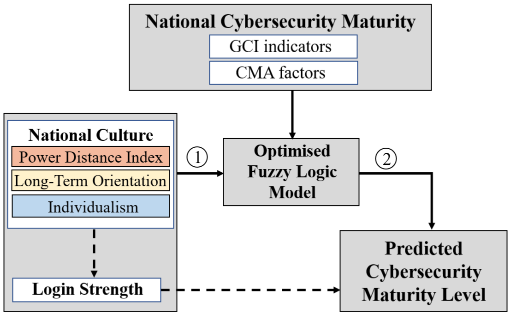


Figure 1: Influence of national culture on cyber security maturity. *Solid lines represents our approach to optimise CML assessment by using the fuzzy logic model; dashed lines considers CML by connecting national culture and login strength.*

values in both the GCI and CMM datasets. These 11 countries form the basis for our case studies. Our validation is dependant on complete sets of data values for the countries used in the comparison. Although future research may consider the finetuning of the fuzzy logic CMA model when not all values are available, the scope of this specific experiment is to validate the use of the model with CMA data obtained from other maturity models; in this case GCI and CMM.

3.1.2. Custom Login Credentials Dataset The custom dataset (CD) comprises statistical analysis of username and password combinations that are available in the publicly available Anti Public (AP) Combo list and Exploit.in (EX) datasets which both contain real breached user login-details collated and compiled over a number of years [5] [6]. In total, 257 unique country code top-level domains were identified in the AP and EX datasets.

The CD dataset presents statistical analysis on the original datasets according to the country code top-level domain[7] in the associated email addresses. The CD dataset further presents an analysis on 14 login credential parameters (shown in Tab. 1), populated for a series of countries. The parameters summarise the manipulation of the original datasets to provide insights from analysing the global authentication trends in username and password pairs, and insight into these login credential parameters on a national scale. We use the knowledge gained from the CD dataset to devise a fuzzy model that derives the output for the GCI index of a particular country. We refer to this as our research study base model.

3.1.3. Hofstede Cultural Dimensions Model The HCD model[8] dataset provides the base culture data from 111 countries and sub cultural groupings. Although the model does not consider individual behaviours within the group setting, it has been applied extensively across numerous cross-national, longitudinal, and validation studies have since its inception in 1973 to provide valuable insights into the dynamics of cultural relations [24]. The model is designed to represent six independent dimensions that distinguish a country's cultural values from each other [25]. Based on the quantitative score for each cultural dimension, countries are placed into one of three categories: Low (0-33), Mid (34-66), or High (67-100). Fig. 2 provides an overview of the six dimensions, followed by a summary of what each of these dimension scores represents in the context of national culture [8]. Our experimental analysis in Section 3.2 shows that only three of the HCD dimensions have a connection with a country's CML, and as such, only the three relevant dimensions are shown in Fig. 1.

Keeping public concerns about the rigidity of the HCD model in mind, we consider the categorisation for each of the dimensions as only a starting point for understanding national culture. We accept that the HCD data for a country may not present a fluid representation of all individuals within that country, but accept it as a snapshot of a representative portion of the individuals within that country, at a specific point in time, and not as a rigid and representative view of the national culture.

3.1.4. GCI and CMM Figures The latest datasets included in this study are the index values and figures from the GCI and CMM reports since they jointly shape the building blocks of a national cyber security culture[1]. In particular, we focus on **GCI Indicator 4.1** (IND4.1 *Public awareness campaigns*) and **GCI Indicator 4.4** (IND4.4 *National educational courses and programs*) from the GCI, and **CMM Factor 1.1** (FCT1.1 *National cyber security strategy*), and **GCI Indicator 4.4** (IND4.4 *National educational courses and programs*) from the CMM. Based on prior analysis, these dimensions were identified to best represent the affirmation between national culture and login strength, and also the transitive dependency between login strength and CML [3]. For instance, the weighting of IND4.1 and IND4.4 are respectively 0.036 and 0.032, jointly contributing 34% of the overall GCI *Capacity building* pillar[1]. This contribution is significant in that these two indicators are amongst the highest contributors within the GCI Pillar 4's seven indicators. Therefore, we argue that these indicators are strongly aligned and connected with login strength and cyber security maturity.

Next, we mapped IND4.1 to FCT1.1, and IND4.4 to FCT3.2 and FCT3.3 respectively. It must be noted that for a more aligned comparison, we average the values for FCT3.2 and FCT3.3 by weighting each at 0.5 to enable

[5] https://databases.today/

[6] https://www.hackread.com/anti-public-combo-list-with-billions-of -accounts-leaked/

[7] https://www.iana.org/domains/root/db

[8] https://geerthofstede.com/research-and-vsm/dimension-data-matrix/

a one-to-one comparison with IND4.4. The mappings were made based on both IND4.1 and FCT1.1 focuses on measuring the government's effort to promote certain agendas pertaining to cybersecurity across a variety of different stakeholders [26, 27]. In a similar fashion, **GCI Indicator 4.4** was mapped to **CMM Factor 3.2** and **CMM Factor 3.3**. This is because both measure the level of education courses and programs to train cybersecurity related skills on a national level.

3.2. Culture Dimensions Evaluation

In addressing RQ1, we evaluate the cultural dimensions to further refine the existing fuzzy logic model, see Fig. 1 ①. We considered the HCD and CD datasets to conduct a full Spearman correlation evaluation [28] to determine which of the national culture dimensions have a strong link with login credential strength. Spearman correlation is often used to evaluate relationships using ordinal values; in this case, each HCD dimension is cross-correlated against the country profiles in the CD dataset, and all entries with incomplete data are excluded to ensure a fully representative statistical correlation.

We conducted Spearman statistical correlation tests between the six HCD dimensions and the 14 login credential parameters from the CD dataset[9]. The rows labeled 'Rho' present the correlation values, and the subsequent rows present the significance for each dimension. The *two-tailed* value represents the p-values for the two-tailed test. A p-value less than 0.05 indicates that the corresponding Rho values differ significantly from zero. A positive correlation means that an increase in one parameter increases the other parameter. We investigate the correlations between particular HCD dimensions and any of the login credential parameters, using both left and right one-tailed tests with the two-tailed tests to identify any possible single-sided correlations. For example, the two-tailed test between PDI and the parameter 'total' shows no correlation with the two-tailed test (p-value=0.0845); the left-tailed test returns a p-value of 0.0422, confirming a negative correlation between PDI and 'total'.

In our analysis, all the p-values that are less than 0.05 are regarded as significant. Of the six cultural dimensions, PDI, IDV, and LTO have the highest number of yellow cells compared with MAS, UAI, and IND, indicating the highest impact of the six HCD dimensions. Majority of the login credential parameters show a negative correlation with PDI. Hence, *the higher the PDI, the lower the password strength (and vice versa)*. However, majority of the login credential parameters show a positive correlation with IDV, indicating that *the higher the IDV, the higher the password strength (and vice versa)*. LTO shows a similar trend as IDV. Assuming that all the login credential attributes have a similar impact on password strength, we can order the dimensions in terms of significance based on the number of yellow columns in the table; i.e. LTO (13 columns) > IDV (11 columns) > PDI (10 columns) > MAS (6 columns) > IND (1 column) > UAI (0 columns). As the three remaining dimensions are less significant, we consider only the first three dimensions to derive the fuzzy model and conduct further analysis and discussions.

3.3. Fuzzy Model Optimisation

Our fuzzy logic model is based on experimental analysis that showed that national culture (specifically, the PDI, LTO, and IDV dimensions) has a direct connection with the login strength demonstrated by individuals of that country. The positive causal relationship between login strength and CML identified earlier is affirmed not only by the GCI, but is supported by a number of other Cyber Security Maturity Models, such as C2M2 [2] and CCSMM [29]. These documents affirm that education and awareness of the general public on the importance of secure login details is a key indicator in measuring CML. Our research methodology aims to optimise the fuzzy logic model by extending the experimental design for our earlier fuzzy logic CMA model to enable the optimisation of cyber security maturity assessments.

Specifically, we aim to validate the application of our model to different datasets and will refine and finetune the fuzzy logic model based on further experiments. As such, the initial rule matrix used to generate the rule surface was adjusted to incorporate the additional parameters from the GCI and CMM data sources. Fig. 3 shows the rule matrix for the optimised fuzzy models, incorporating membership functions for the base (original) fuzzy model and the parameters of all four data sources to be used in the creation of the model surfaces. As each input has three membership functions, it resulted in 27 ($= 3 \times 3 \times 3$) rules to be defined among the inputs and outputs. These fuzzy rules reflect the outcomes of the case studies.

In declaring a fuzzy model, we model input parameters (in this case, PDI, IDV, and LTO) with the output parameters (in this case, GCI or CMM indices), using a rule base (refer to Fig. 3) and a collection of membership functions (refer to Fig. 4). In a previous study [3], we identified that the PDI, IDV, and LTO dimensions of HCD have the most substantial influence on the selected GCI and CMM indices. We further identified correlations between higher PDI and lower password strength values, as well as higher LTO values and higher password strength values. Considering this knowledge, we declare the base model to reflect these patterns to model the three inputs to each of the respective output parameters. We used the triangular membership function (considering

[9] please refer to [3] for the corresponding results

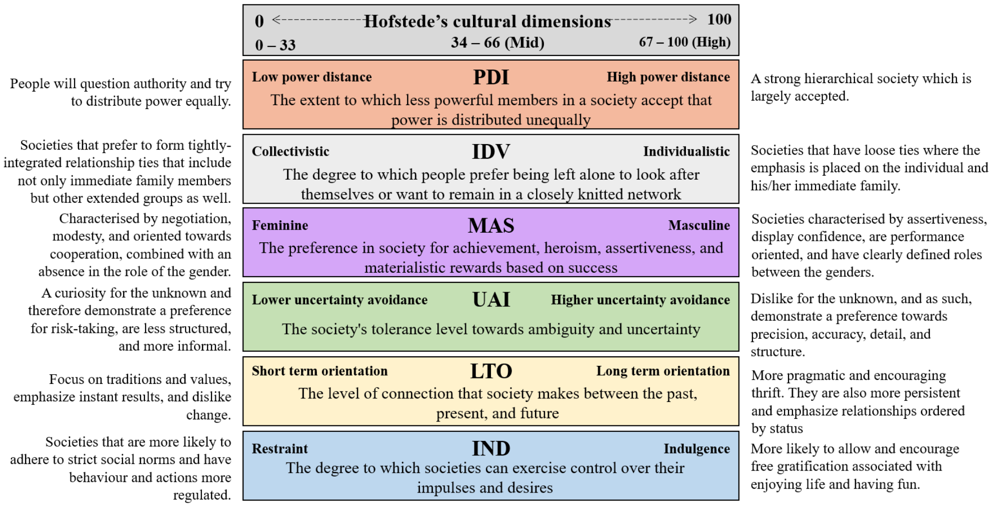

0 Hofstede's cultural dimensions 100
0 – 33
34 – 66 (Mid)
67 – 100 (High)
People will question authority and try to distribute power equally.
Low power distance
PDI
High power distance
The extent to which less powerful members in a society accept that power is distributed unequally
A strong hierarchical society which is largely accepted.
Societies that prefer to form tightly-integrated relationship ties that include not only immediate family members but other extended groups as well.
Collectivistic
IDV
Individualistic
The degree to which people prefer being left alone to look after themselves or want to remain in a closely knitted network
Societies that have loose ties where the emphasis is placed on the individual and his/her immediate family.
Characterised by negotiation, modesty, and oriented towards cooperation, combined with an absence in the role of the gender.
Feminine
MAS
Masculine
The preference in society for achievement, heroism, assertiveness, and materialistic rewards based on success
Societies characterised by assertiveness, display confidence, are performance oriented, and have clearly defined roles between the genders.
A curiosity for the unknown and therefore demonstrate a preference for risk-taking, are less structured, and more informal.
Lower uncertainty avoidance
UAI
Higher uncertainty avoidance
The society's tolerance level towards ambiguity and uncertainty
Dislike for the unknown, and as such, demonstrate a preference towards precision, accuracy, detail, and structure.
Focus on traditions and values, emphasize instant results, and dislike change.
Short term orientation
LTO
Long term orientation
The level of connection that society makes between the past, present, and future
More pragmatic and encouraging thrift. They are also more persistent and emphasize relationships ordered by status
Societies that are more likely to adhere to strict social norms and have behaviour and actions more regulated.
Restraint
IND
Indulgence
The degree to which societies can exercise control over their impulses and desires
More likely to allow and encourage free gratification associated with enjoying life and having fun.

Figure 2: CDM score scale

PDI	L									M									H								
IDV	L			M			H			L			M			H			L			H			M		
LTO	L	M	H	L	M	H	L	M	H	L	M	H	L	M	H	L	M	H	L	M	H	L	M	H	L	M	H
Base Model	L	M	H	L	H	H	M	H	H	L	L	M	L	M	M	M	H	H	L	L	L	L	M	M	L	H	H
GCI 4.1	L	M	H	L	H	H	M	H	H	H	L	M	M	M	H	M	H	H	L	H	M	H	M	M	L	H	H
GCI 4.4	L	L	L	L	L	L	H	H	L	M	L	L	L	L	M	M	L	L	L	L	L	L	L	L	L	L	L
CMM 1.1	L	M	H	L	H	H	M	H	H	L	L	L	L	M	M	M	H	H	L	L	L	M	M	M	M	H	H
CMM 3.2_3.3	L	M	H	L	H	H	M	H	H	L	L	M	H	M	M	M	M	H	L	L	L	L	L	M	L	M	H

Figure 3: Optimised fuzzy model rule matrix

three levels: L:LOW, M:MODERATE, and H:HIGH) for the inputs and the gaussian membership functions (considering three levels: L:LOW, M:MEDIUM, and H:HIGH) for the output (refer to Fig. 4 following the patterns in the original values and the statistical analysis. We optimised the rule bases and the membership functions of the base model to obtain four optimised fuzzy models that generate outputs for the aforementioned GCI and CMM indices. As shown in these figures, the output membership function shapes were kept the same to represent the original variations of the outputs. However, we changed the input membership functions to reflect the dynamics of the values represented in the case studies.

We applied the semantics identified during the statistical analysis: negative correlation means an increase in PDI leads to a decrease in GCI, and positive correlation means an increase in IDV and LTO leads to an increase in GCI. After defining the rules for the fuzzy modeling, we obtain the rule-surface, as based on the fuzzy membership functions shown in Fig. 4. This figure represents the fuzzy membership functions used to generate the optimised GCI index on the three inputs (PDI, IDV, and LTO). The membership functions were optimised to obtain four different fuzzy models that generate optimised outputs for IND4.1 and IND4.4, as well as FCT1.1 and FCT3.2/FCT3.3 (combined) for the 11 case study countries (refer to Section 3.1.4). These figures show the modifications conducted on the original fuzzy model (base model in the figures) on the GCI to obtain four separate models that generate values for IND4.1 and IND4.4 of the GCI, and FCT1.1 and FCT3.2/FCT3.3 (combined) of the CMM, respectively. As shown in Fig. 4, the ranges and the choice of shapes of the membership functions for the output variables remain the same while the input variables' membership functions are adjusted to reflect the outcome of the case studies. This fuzzy model can be used to define the level of GCI based on any combination of values for PDI, IDV, and LTO.

4. Case Study Country Overview

We focus on the 11 countries with sufficient data in all four datasets (HCD, CD, GCI, and CMM) to test the two research questions. This case study data are represented in Tab. 1, with the countries referred to by their ISO 3166 country codes[10]. The values are displayed in their original format and range from the data sources. All ranked values (HCD, GCI and CMM) are normalised from 0 to 1 in the fuzzy model to make the model generalisable to any value. We consider low, medium/mid, and high values for the fuzzy logic rules.

The *HCD data*[11] are represented in the first section of the table, with rows marked as PDI, IDV, and LTO. These values are shown in a range of 1 to 100, with val-

[10] https://www.iso.org/standard/63545.html

[11] The country specific explanations are summarised from Hofstede Insights at https://www.hofstede-insights.com/country-comparison/.

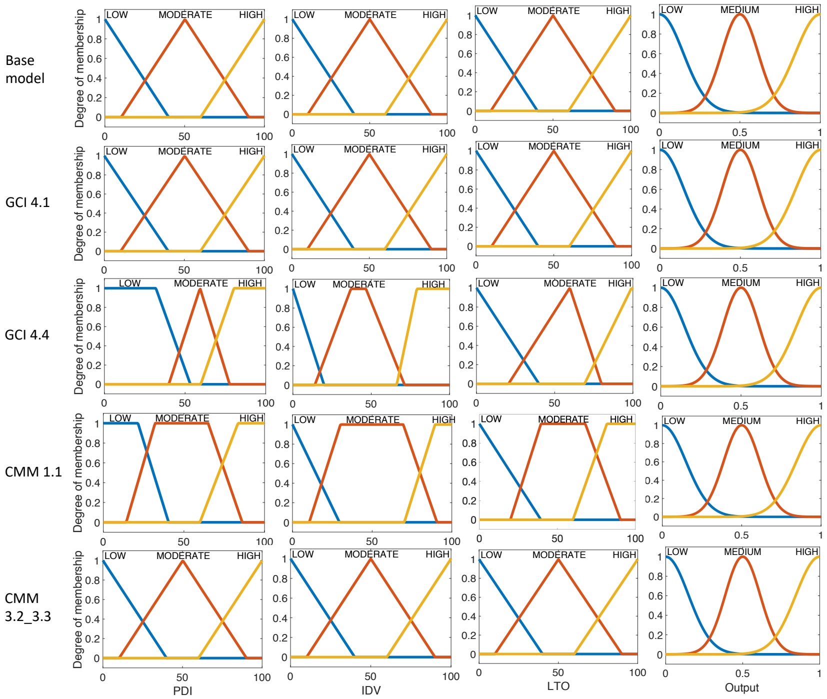


Figure 4: Fuzzy membership functions of the optimised models. *Each fuzzy variable (input/output) has three membership functions (namely, LOW: in blue, MEDIUM/MODERATE: in red, and HIGH: in orange). The first row shows the fuzzy variables and their membership functions used in the base model. The subsequent rows show the optimized membership functions for each index.*

ues between 0 and 33 representing a low classification (coloured red), values between 34 and 66 representing a mid classification (coloured yellow), and values between 67 and 100 representing a high classification (coloured green). The *CD data* are represented in the second section of the table, the rows marked TP to MXE representing the values for the different username/password strength parameters (refer to ** below the table). These values are not normalised but reflect actual data values and statistical measures. The *GCI data* are represented next, with the rows representing the values for IND4.1 and IND4.4, respectively. Tab. 1 displays these values on a range of 1 to 5, with (1) being the lowest value and considered the Start-up stage (coloured red), (2) considered the Formative stage (coloured orange), (3) considered the Established stage (coloured yellow), (4) considered the Strategic stage (coloured light green) and (5) considered the Dynamic stage and the highest maturity (coloured darker green). None of the countries in our selection has reached full cyber security maturity on this scale.

5. Results and Discussion

To determine whether the fuzzy logic model can be applied using data from different CMA models, we compared the results from the case study countries. These countries were selected as the only ones with reliable data (with least sparseness) for all parameters across the datasets. We used an extensive input dataset to refine and improve the existing fuzzy logic model based on the relationship derived between national culture and CML.

We created an optimised fuzzy model rule matrix across all the input parameters and developed separate membership functions for the base model and the CML inputs. The four new fuzzy models developed based on the rule matrix are applied to the datasets for the 11 countries. Fig. 5 shows the values generated by our optimised models coloured in gray, compared to the original GCI/CMM values coloured in blue and the values returned by the base model coloured in orange. In the proposed model, the fuzzy output gives an overall impression of cyber security maturity. The model infers the possibility of deriving a value relative to the GCI value based on only three significant parameters (PDI, IDV, LTO), as opposed to a large number of difficult to measure indicators traditionally used by CMA models. The fuzzy output provides a relative notion to the original GCI values of the countries. We apply this to the case study countries to evaluate and test the model by statistically cross-correlating the output from the model with country specific data on cyber security.

To address the first research question, the optimisations conducted on the fuzzy model have allowed the four individual models to determine CML values that are closer to the original GCI/CMM values than the base model values. We demonstrated this in Fig. 1 by conducting the final evaluation of the four models based on the

Table 1: Case study summary of data

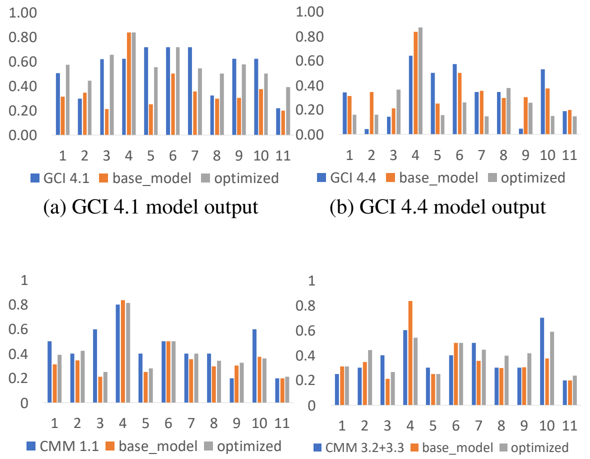
Data / Country*		BGD	CHL	COL	GBR	IDN	LTU	MEX	PER	SLV	URY	VEN
Hofstede's Cultural Dimensions	PDI	80	63	67	35	78	42	81	64	66	61	81
	IDV	20	23	13	89	14	60	30	16	19	36	12
	LTO	47	31	13	51	62	82	24	25	20	26	16
Custom Login Credentials**	TP	7,999	396,112	580,510	25,403,240	2,205,976	444,111	2,057,816	88,170	7,658	49,869	43,264
	UE	4,746	248,333	430,531	15,608,501	1,334,755	232,560	1,290,972	50,102	5,055	30,566	28,449
	UP	3,565	217,711	336,625	9,162,835	804,070	195,099	961,835	43,606	4,263	29,293	22,945
	UU	4,307	205,537	384,660	15,608,501	1,245,267	204,303	1,187,974	43,737	4,806	28,266	26,079
	UD	1,948	35,637	77,052	929,663	1,334,755	19,279	46,944	7,981	7,658	4,408	9,659
	UNP	355	6,164	8,092	197,466	13,734	11,497	961,835	1,350	150	1,328	629
	PIU	2,033	7,751	13,615	338,427	33,502	10,832	52,071	2,354	984	1,288	5,256
	PM	109	3,750	4,298	123,476	8,958	6,300	17,608	748	90	576	266
	MEL	6	8	8	8	8	8	8	8	8	8	8
	MNL	0	0	0	0	0	0	0	0	0	0	0
	MXL	73	79	109	212	119	131	86	71	65	87	87
	MEE	14	18	18	18	17	18	18	18	17	18	16
	MNE	0	0	0	0	0	0	0	0	0	0	0
	MXE	82	96	102	81	144	144	96	88	88	166	272
Global Cyber security Index	4.1	2.53	1.48	3.1	3.1	3.58	3.58	3.58	1.61	3.1	3.1	1.09
	4.4	1.71	0.22	0.72	3.21	2.5	2.86	1.73	1.73	0.23	2.65	0.94
Cyber security Maturity Model	1.1	2.5	2	3	4	2	2.5	2	2	1	3	1
	3.2/3.3	1.75	1.5	2	3	1.5	2	2.5	1.5	1.5	3.5	1

* BGD = Bangladesh; CHL = Chile; COL = Colombia; GBR = United Kingdom; IDN = Indonesia; LTU = Lithuania; MEX = Mexico; PER = Peru; SLV = El Salvador; = Urugauy; VEN = Venezuela

** TP = Total Passwords; UE = Unique Emails; UP = Unique Passwords; UU = Unique Usernames; UD = Unique Domains; UNP = Usernames in Passwords; PIU = Passwords in Usernames; PM = Pure Match; MEL = Mean Length; MNL = Min Length; MXL = Max Length; MEE = Mean Entropy; MNE = Min Entropy; MXE = Max Length

values from 11 case study countries. Considering the four sub figures we find that the optimised model is significantly more accurate in assessing the CML than the base model. In sub figure (a), seven of the 11 model evaluations showed an improved CML assessment, whilst one assessment remained the same as the base model. In sub figures (c) and (d), respectively, seven and four model evaluations show improvements, whilst respectively one and three model evaluations remain the same. Sub figure (b) shows the least improvement with only three model evaluations resulting in a more accurate assessment. Based on these values, we accept that refinement and improvement in our existing fuzzy logic model are possible by applying data from different CMA models.

To address the second research question, we compared the results obtained between the two CMA models. We note that both our selected models showed varying improvement, with the CMM showing slightly more improvement. Overall, the CMM showed 11 improved assessments, seven worse off assessments, and four assessments remaining the same. Of these, both assessments for Colombia, United Kingdom, and Mexico improved, both assessments for country El Salvador and Venezuela were worse off, and both assessments for Lithuania remained consistent with the base model assessments. The GCI showed overall 10 improved assessments, 11 worse off assessments, and one assessment remaining the same. Of these, both assessments for El Salvador improved, and both assessments for Venezuela were worse off. We observe that both the GCI IND4.1 and CMM FCT1.1 model outputs showed a consistently improved assessment across seven countries, with one country assessment worse than the base model and three countries' assessments remaining the same. This is evidence that the optimisation of the fuzzy model can help improve the accuracy with which the CMLs can be assessed.



(a) GCI 4.1 model output

(b) GCI 4.4 model output

(c) CMM 1.1 model output

(d) CMM 3.2 / 3.3 (Combined) model output

Figure 5: Model output comparison. *Comparing the optimised model outputs of each index with the base model and the original index values.*

Although there are noticeable overlaps in the countries that show the same outcomes (Bangladesh, Colombia, Indonesia, and Mexico show improvement whilst Venezuela's maturity deteriorates in both models), these model evaluations cannot be regarded as a rule based on the current statistical analysis and model evaluations. Although these outputs show promise in terms of the preliminary output of our fuzzy logic model in refining the model based on various different CMA models, further analysis on the correlation between national culture and its causal effect on awareness and education factors is required to enhance the accuracy of the model. Furthermore, this will also provide additional insights into how cultural indicators impact digital societies within a particular cultural subclass.

6. Limitations and Future Research

Three main aspects are suggested for possible future rework and investigation to further improve the accuracy with which the model can assess a country's CML. Firstly, we have mapped GCI IND4.4 and CMM FCT 3.2/3.3 (combined) to enable a one-to-one comparison between the GCI indicator and the CMM factor. However, these model outputs did not show similarities in our analysis as GCI IND4.4 showed three improvements and eight worse off assessments, whilst CMM FCT 3.2/3.3 (combined) showed four improvements, four worse off and three assessments remaining the same. The only similar assessments in both output models are for Venezuela, which was a worse off assessment. Although we initially opted to map CMM FCT 3.2/3.3 (combined) to GCI IND4.4 by averaging out the values, we consider this as the most probable reason for why the outputs across the two models do not show similar outputs, particularly since the outputs between GCI IND4.1 and CMM FCT1.1 are closely aligned. In future work, we will consider an **alternative weighting of qualitative data to estimate the maturity level**, particularly for the combination of related factors, to ensure a more consistent model output.

Secondly, further refinement of the model is encouraged through future studies. With our current analysis, only Venezuela showed a consistent output across all four output models; that is, the assessment was worse off than the base model assessments. All the other countries showed values with different combinations of improvements, worse off, and consistent assessments. Improved accuracy in the CML assessments can be achieved if different variations of the model based on a variety of **case study characterisation** can be further explored.

Thirdly, further investigation may be required in terms of the CMA models selected for comparison. We selected the GCI and CMM as input models, but these models use different scales for representing CMLs and different methodologies to collect and analyse data. The methodology utilised by these models also differs, with the GCI using a standard questionnaire across decision makers, whereas the CMM adopts a focus group approach with pre-identified stakeholders of the country to gain an overview of the country CML. To enable comparison of these models, we applied normalisation to the model outputs. We concur with Schlienger and Teufel [30] that a cyber security culture should reflect on the social, cultural, and ethical aspects of the users in order to alter the users' overall cyber security behaviour. We argue that culture is evident in the behaviour of the users and that the cultural dimension values and real user behaviour data in the form of the login credential parameters are essential in the accurate assessment of CML.

7. Conclusion

As countries progress on their digital transformation journey, there is a need for effective cyber security capacity building strategies and clear merits for countries to establish their baseline cyber security capacity through maturity assessment models. We argue that there is a need to simplify the assessment process to reduce the time and effort spent compared to existing methods, but also to enable governments to make informed decisions based on how their citizens perceive privacy and security in relation to cyber security. In addressing this challenge, our research proves that it is viable to reduce the number of inputs and resources required to determine a country's cyber security maturity level. We demonstrate this through our fuzzy logic cyber security assessment model which can simplify cyber security maturity models for a quick indicative assessment based on national culture.

Our theory shows that modelling the HCD data and the real world login credential data from countries with full datasets can lead to an improved and optimised rule matrix that enables a quick assessment of CML values. Our results open new possibilities for reducing the time required to provide governments with indicative figures pertaining to their country's relative cyber security maturity, without the current extensive and time consuming data collection through workshops, focus groups, or considerable effort to complete questionnaires that accurately represent national scale. These models can be used to assess a country's cyber security capacity that will, in turn, enable prompt facilitation of targeted interventions by governments to where it is most critical. The optimisations applied to our fuzzy logic CML assessment model allow us to generate comparative CML values on a national scale, with less effort, and in a shorter timespan.

The inherent contribution of our proposed model is its application to validate and prove that the transitive dependency identified between national culture and cyber security maturity is a simplification of quickly determining a country's relative CML. Although our model does not replace the need for traditional and more comprehensive CMAs, our model is useful for quick analysis to enable governments to develop more strategies that can assist with targeted interventions, particularly those relevant to cyber security awareness, education, and training, as well as cyber security strategy (GCI IND 4.1 and IND4.4, CMM FCT1.1, v3.2, and FCT3.3). We emphasise the consideration of human behavioural aspects as prominent drivers in measuring maturity for the continuous improvement of cyber security awareness. Our research affirms that national culture dimensions have a direct impact on human behaviour in terms of login strength and can be used to establish a benchmark across multiple facets of cyber security capacity and be applied to assess a country's relative CML.

Acknowledgment

The work has been supported by the Cyber Security Cooperative Research Centre Limited (CSCRC) whose activities are partially funded by the Australian Government's Cooperative Research Centres Program.

References

[1] Angel Marcelo Rea-Guaman et al. 'Comparative study of cybersecurity capability maturity models'. In: *International Conference on Software Process Improvement and Capability Determination*. Springer. 2017, pp. 100–113.

[2] Jason D Christopher et al. 'Cybersecurity capability maturity model (C2M2)'. In: *Department of Homeland Security* (2014), pp. 1–76.

[3] Jongkil Jay Jeong et al. 'Fuzzy logic application to link national culture and cybersecurity maturity'. In: *5th International Conference on Collaboration and Internet Computing (CIC)* (2019).

[4] Mahawaga Arachchige Pathum Chamikara et al. 'PPaaS: Privacy Preservation as a Service'. In: *Computer Communications* 173 (2021), pp. 192–205.

[5] Waldo Rocha Flores and Mathias Ekstedt. 'A Model for Investigating Organizational Impact on Information Security Behavior'. In: *Proceedings of the 7th Pre-ICIS Workshop on Information Security and Privacy, Orlando, December 15, 2012*. 2012.

[6] Balbir S Barn, Ravinder Barn and Jo-Pei Tan. 'Young people and smart phones: An empirical study on information security'. In: *47th Hawaii International Conference on System Sciences*. IEEE. 2014, pp. 4504–4514.

[7] Chun Wei Choo. 'Information culture and organizational effectiveness'. In: *International Journal of Information Management* 33.5 (2013), pp. 775–779.

[8] Diane Henshel et al. 'Integrating cultural factors into human factors framework and ontology for cyber attackers'. In: *Advances in Human Factors in Cybersecurity*. Springer, 2016, pp. 123–137.

[9] Gianfranco Walsh et al. 'Cross-cultural fear of online identity theft: A comparison study and scale refinement'. In: *39th International Conference on Information Systems (ICIS)*. Association for Information Systems. 2018.

[10] Sunthoshan Govender, Elmarie Kritzinger and Marianne Loock. 'The influence of national culture on information security culture'. In: *IST-Africa Week Conference*. IEEE. 2016, pp. 1–9.

[11] Paul Van Schaik et al. 'Risk perceptions of cybersecurity and precautionary behaviour'. In: *Computers in Human Behaviour* 75.2017 (2017), pp. 547–559.

[12] Margaret Tan and Kathrine Sagala Aguilar. 'Risk perceptions of cyber-security and precautionary behaviour'. In: *Information Management & Computer Security* 20.5 (2012), pp. 364–381.

[13] Elizabeth Stobert and Robert Biddle. 'The password life cycle: user behaviour in managing passwords'. In: *Proc. SOUPS*. 2014.

[14] Q Yan, L Wu and L Yi. 'Influence of social identity on information release in microblog'. In: *2012 Second International Conference on Intelligent System Design and Engineering Application* (Jan. 2012).

[15] R Reid and J Van Niekerk. 'From information security to cyber security cultures'. In: *Information Security for South Africa (ISSA) Conference*. 2014.

[16] Gizem Öğütçü, Özlem Müge Testik and Oumout Chouseinoglou. 'Analysis of personal information security behavior and awareness'. In: *Computers & Security* 56 (2016), pp. 83–93.

[17] Steve Sheng et al. 'Who falls for phish? A demographic analysis of phishing susceptibility and effectiveness of interventions'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2010, pp. 373–382.

[18] Daniel Pienta, Wenxi Pu and Russell Purvis. 'The Impact of Culture on Information Security: Exploring the Tension of Flexibility and Control'. In: *International Conference on Information Systems*. Association for Information Systems. 2017.

[19] N Gcaza et al. 'A general morphological Analysis: delineating a cyber-security culture'. In: *Information Computer Security*. Vol. 25. 2017.

[20] Jongkil Jeong et al. 'Towards an Improved Understanding of Human Factors in Cybersecurity'. In: *2019 IEEE 5th International Conference on Collaboration and Internet Computing (CIC)*. IEEE. 2019, pp. 338–345.

[21] Robert W Proctor and Jing Chen. 'The role of human factors/ergonomics in the science of security: Decision making and action selection in cyberspace'. In: *Human Factors* 57.5 (2015), pp. 721–727.

[22] Ngoc T Le and Doan B Hoang. 'Can maturity models support cyber security?' In: *IEEE 35th International Performance Computing and Communications Conference (IPCCC)*. IEEE. 2016, pp. 1–7.

[23] Geert Hofstede. 'The cultural relativity of organizational practices and theories'. In: *Journal of International Business Studies* 14.2 (1983), pp. 75–89.

[24] Michael Jones. 'Hofstede - Culturally Questionable?' In: *Oxford Business Economics Conference* (2007).

[25] Geert Hofstede. 'Dimensionalizing cultures: The Hofstede model in context'. In: *Online Readings in Psychology and Culture* 2.1 (2011), p. 8.

[26] ITU. *Global Cybersecurity Index (GCI, International Telecommunication Union) 2018*. Tech. rep. 2019. URL: shorturl.at/uzA59.

[27] GCSCC University of Oxford. 'Global impact: Knowledge and policy contributions from the first five years.' In: (2019). URL: shorturl.at/lEGP1.

[28] Leann Myers and Maria J Sirois. 'Spearman correlation coefficients, differences between'. In: *Encyclopedia of Statistical Sciences* 12 (2004).

[29] Gregory B White. 'The community cyber security maturity model'. In: *IEEE International Conference on Technologies for Homeland Security*. IEEE. 2011, pp. 173–178.

[30] T Schlienger and S Teufel. 'Information security culture – From analysis to change'. In: *Security in the Information Society*. Springer. 2002, pp. 191–201.