Exploring how Component Factors and their Uncertainty Affect Judgements of Risk in Cyber-Security*

Zack Ellerby, Josie McCulloch, Melanie Wilson, and Christian Wagner

Computer Science, University of Nottingham, UK,
{zack.ellerby,josie.mcculloch,
melanie.wilson,christian.wagner}@nottingham.ac.uk

Abstract. Subjective judgements from experts provide essential information when assessing and modelling threats in respect to cyber-physical systems. For example, the vulnerability of individual system components can be described using multiple factors, such as complexity, technological maturity, and the availability of tools to aid an attack. Such information is useful for determining attack risk, but much of it is challenging to acquire automatically and instead must be collected through expert assessments. However, most experts inherently carry some degree of uncertainty in their assessments. For example, it is impossible to be certain precisely how many tools are available to aid an attack. Traditional methods of capturing subjective judgements through choices such as *high*, *medium* or *low* do not enable experts to quantify their uncertainty. However, it is important to measure the range of uncertainty surrounding responses in order to appropriately inform system vulnerability analysis. We use a recently introduced interval-valued response-format to capture uncertainty in experts' judgements and employ inferential statistical approaches to analyse the data. We identify key attributes that contribute to hop vulnerability in cyber-systems and demonstrate the value of capturing the uncertainty around these attributes. We find that this uncertainty is not only predictive of uncertainty in the overall vulnerability of a given system component, but also significantly informs ratings of overall component vulnerability itself. We propose that these methods and associated insights can be employed in real world situations, including vulnerability assessments of cyber-physical systems, which are becoming increasingly complex and integrated into society, making them particularly susceptible to uncertainty in assessment.

Keywords: cyber-security · uncertainty · interval-values · intervals

1 Introduction

Cyber-security professionals play a vital role in assessing and predicting vulnerabilities within cyber-physical systems, which often form part of an organisa-

* Supported by EPSRCs EP/P011918/1 grant and by the UK National Cyber Security Centre (NCSC).

tion's or state's critical digital infrastructure. As outsider threats become more prevalent and sophisticated, there is increasing pressure on experts to provide timely and comprehensive assessments within the context of the rapidly changing cyber-physical ecosystem. As cyber-systems increase in both ubiquity and complexity, methods to quantify and handle error in subjective measurements from experts need to be developed [13]. It has been demonstrated across many industry sectors that as complexity increases accurate risk assessment decreases [10]. Enabling the effective reconstruction of overall assessments from component and attribute ratings would streamline the process of updating overall system vulnerability assessments, in line with shifts in this ecosystem.

Both objective and subjective measures of risk provide useful information to aid decision making in vulnerability assessment [4, 5]. Several different methods can be used to assess vulnerability and risk in a cyber-security system, such as vulnerability scanning tools [18] or the Common Vulnerability Scoring System (CVSS) [15], which gives qualitative severity ratings of *low*, *medium*, and *high*, and CVSS Version 3, which extends the ratings to include *none* and *critical* [8]. However, when using CVSS, the necessary information to complete the calculation may be missing. Hubbard and Seiersen [11] argue that assessing risk in terms of *low*, *medium*, and *high* ratings is highly subjective and open to error. It is therefore suggested that cyber-security risk should be described quantitatively. This would help quantify what areas of risk are perceived as important to cyber-security professionals and, from that, move towards how those risks might correspond to the actually enacted attacks and their success or failure.

Assessment of risk or the likelihood of an attack are inherently uncertain [1, 2]. Objective measures may carry uncertainty because the measures themselves are imprecisely defined [2]. Subjective assessments (collected from experts) carry uncertainty because, for example, the experts are not familiar with the particular technology, there is inherent uncertainty caused by insufficient detail in the scenario, or due to individual personality [12, 16, 20].

Between-expert uncertainty is often modelled implicitly, for example, through probability distributions [7] or uncertainty measures [5]. These methods model the between-expert uncertainty, but they do not capture within-expert uncertainty. Choi et al. [4] capture within-expert uncertainty by enabling experts to express knowledge and uncertainty through terms such as *very small*, *small* and *large* and using fuzzy sets to represent the uncertainty of these words. However, this assumes that the experts share the same degree of uncertainty regarding the meanings of these terms. After capturing uncertainty, it may be modelled and handled through methods such as Dempster-Shafer evidence theory [6, 9] or fuzzy logic [4, 14, 19].

We propose explicitly capturing uncertainty in experts' individual judgements using an interval-valued response format, as previously introduced in [17]. Experts provide ratings along a continuous scale to quantify, for example, the perceived difficulty of an attack on a component. Using an interval captures both the experts' rating (position on the axis) and the degree of uncertainty associated with the response (width of the interval). Fig. 1 shows an example of a narrow

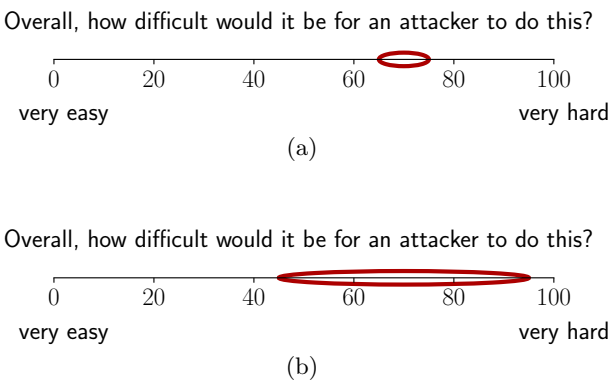

Fig. 1. Illustration of narrow (a) and wide (b) interval-valued responses capturing different degrees of uncertainty.

rating (slightly uncertain) and a wide rating (highly uncertain). In this manner, uncertainty is captured as an integral aspect of the judgement itself, through a coherent and intuitive response-format. The novel output of this paper is to show that an interval-valued response scale can be used to effectively capture uncertainty in expert judgements, and that this can be used to better predict both the magnitude and the uncertainty of risk in vulnerability assessments.

In this paper, we assess the importance of a variety of attributes in determining the overall vulnerability of components being attacked or evaded within a cyber-system; these include maturity of the technology and the frequency that a given attack is reported (the full set of attributes are listed in Tables 1 and 2). We also capture the overall difficulty of attacking or evading each component. Our core aim is to understand how the component attributes contribute to the overall difficulty of an attack and, equally importantly, how uncertainty in component attributes affects not only the uncertainty in the overall vulnerability of a component, but also the overall difficulty itself. For example, experts might perceive an attack as more difficult if there are *fewer* tools available to aid in this effort. However, experts might also perceive an attack as more difficult if they are *uncertain* about the availability of tools. From this, we can learn what additional insight is gained by capturing uncertainty through interval-valued responses. Specifically, we wish to answer

- how is overall vulnerability of a component affected
  - by attribute ratings?
  - by uncertainty around attribute ratings?
- how is uncertainty around overall vulnerability of a component affected
  - by attribute ratings?
  - by uncertainty around attribute ratings?

We find that although cyber-security experts only assessed attributes that were previously deemed likely to be important to component vulnerability, only some

of these attributes have a significant effect. Similarly, we discover that although uncertainty around some attributes affects the uncertainty around the component as a whole, this is not consistent for all attributes. We also find that the uncertainty around some attribute ratings makes a significant contribution to overall vulnerability ratings themselves.

2 Methods

2.1 Data Collection

The data was collected from experts in CESG (Communications-Electronics Security Group), which was the information security arm of GCHQ (Government Communication Head Quarters) in the United Kingdom[1]. A total of 38 cyber-security experts at CESG assessed a range of components that are commonly encountered during a cyber-attack. They rated these on both overall difficulty to either attack or evade, as appropriate, and on several attributes that might affect this difficulty. Two types of components (also referred to as hops) were assessed, those that require an attacker to attack the component (referred to as *attack*) and those that require the attacker to bypass the component (referred to as *evade*). Examples of the hops assessed include *bypass gateway content checker* and *overcome client lockdown*, but any hop may be assessed using this method.

Tables 1 and 2 list the attack and evade attributes, respectively (variable notations are also provided, which are used in the next section). Attack hops are defined by seven attributes and evade hops by three attributes. In addition, both are described by their *overall difficulty*. Our aim is to understand how the hop attributes and the uncertainty around these attributes relate to the perceived overall difficulty of attacking/evading the hop, and to the uncertainty around this difficulty. Experts were asked to provide interval-valued ratings to enable them to coherently express the uncertainty associated with their responses; these ratings were provided on a scale from 0 to 100.

The cyber-security components chosen for this study are designed to be representative of a mainstream government system, which would include assets at Business Impact Level 3 (BIL3) – an intermediate category of impact [3]. CESG describes such a system as typically including remote site and mobile working, backed by core services and back-end office integrated systems, such as telemetry devices and associated systems used by the emergency services. Compromising assets at BIL3 might have large-scale negative effects, including but not limited to: disruption to regional power supply, key local transport systems, emergency or other important local services for up to 24 hours, local loss of telecoms, risk to an individual's safety, damage to intelligence operations, hindrance to low level crime detection and prosecution, or financial loss to the UK government or a leading financial company in the order of millions (GBP) [3].

[1] CESG has since been replaced by the NCSC (National Cyber Security Centre).

Table 1. Attributes used to describe attack hops, the question used in the study, and the variable name used in the analysis.

var	attribute	description
c	complexity	How complex is the target component (e.g. in terms of size of code, number of sub-components)?
t	interaction	How much does the target component process/interact with any data input?
f	frequency	How often would you say this type of attack is reported in the public domain?
a	availability of tool	How likely is it that there will be a publicly available tool that could help with this attack?
d	inherent difficulty	How inherently difficult is this type of attack? (i.e. how technically demanding would it be to do from scratch, with no tools to help.)
r	maturity	How mature is this type of technology?
g	going unnoticed	How easy is it to carry this attack out without being noticed?
o	overall difficulty	Overall, how difficult would it be for an attacker to do this?

2.2 Analysis

We use linear mixed effects modelling (an extension of linear regression) to determine the contribution of each of the hop attributes, as rated by experts, to overall hop difficulty. We also assess the contribution of the associated uncertainty in these ratings, captured through the interval-valued response-format. The midpoint (m) of the interval-valued response is used as a single-valued numeric rating of the attribute, and the width (w) of the response is used to represent the uncertainty around this rating. Note, of course, that higher widths are only possible towards the centre of the scale. That is, as the midpoint approaches the edge of the scale, a wide interval cannot exist. Also, note that while experts provided ratings in the range $[0, 100]$, these data were standardised, through z-transformation, before entry into the model.

This approach estimates the contribution of each attribute's midpoint and width together upon the same outcome variable, in the form of β weights. These variables are entered as fixed effects. The inclusion of random intercepts also allows the model to account for potential between expert and between hop differences in baseline ratings. In addition, this technique allows us to examine the combined effects of attribute rating and uncertainty (e.g. high certainty may have an opposite effect on overall difficulty when relating to a high or low attribute

Table 2. Attributes used to describe evade hops, the question used in the study, and the variable name used in the analysis.

var	attribute	description
c	complexity	How complex is the job of providing this kind of defence?
a	availability of information	How likely is that there will be publicly available information that could help with evading defence?
r	maturity	How mature is this type of technology?
o	overall difficulty	Overall, how difficult would it be for an attacker to do this?

rating). We model this through the inclusion of two-way interaction terms ($m \cdot w$) pertaining to the midpoint and width of each attribute.

Four separate analyses are reported. These were conducted for the dependent variables of

- attack hop overall difficulty rating (interval midpoint)
- attack hop overall uncertainty (interval width)
- evade hop overall difficulty rating (interval midpoint)
- evade hop overall uncertainty (interval width)

For each of these analyses, an initial model was created. These included fixed effects of all hop attribute ratings, all hop attribute widths and all two-way interactions, along with random intercepts for both expert and hop. Following this, a stepwise variable reduction process was applied to each model, in order to remove variables that were found not to significantly contribute to the respective outcome variable. β weights of the variables retained into the final models were then interpreted as estimates of the (significant) contribution of each of these factors to the respective outcome variable.

Table 1 lists the variables used to denote the attributes of attack hops. The sum of all simple effects for the attack hop attribute midpoints (ratings) is

$$A_m^z = \beta_1^z x_{i,j}^{cm} + \beta_2^z x_{i,j}^{tm} + \beta_3^z x_{i,j}^{fm} + \beta_4^z x_{i,j}^{am} + \beta_5^z x_{i,j}^{dm} + \beta_6^z x_{i,j}^{rm} + \beta_7^z x_{i,j}^{gm} \quad (1)$$

where β is the coefficient, $x_{i,j}^{cm}$ is the value m (midpoint) of attribute c (complexity) for i (a given expert) and j (a given hop), and z reflects the model's outcome variable, which may be either m (midpoints) or w (widths) of the overall difficulty.

The sum of all simple effects for the attack hop attribute widths (uncertainty) is

$$A_w^z = \beta_8^z x_{i,j}^{cw} + \beta_9^z x_{i,j}^{tw} + \beta_{10}^z x_{i,j}^{fw} + \beta_{11}^z x_{i,j}^{aw} + \beta_{12}^z x_{i,j}^{dw} + \beta_{13}^z x_{i,j}^{rw} + \beta_{14}^z x_{i,j}^{gw} \quad (2)$$

where $x_{i,j}^{cw}$ is the width w of attribute c (complexity) for i (a given expert) and j (a given hop).

The sum of the interactions between the midpoints and widths of the attack hop attributes is

$$A_{mw}^{z} = \beta_{15}^{z}(x_{i,j}^{cm} \cdot x_{i,j}^{cw}) + \beta_{16}^{z}(x_{i,j}^{tm} \cdot x_{i,j}^{tw}) + \beta_{17}^{z}(x_{i,j}^{fm} \cdot x_{i,j}^{fw}) + \beta_{18}^{z}(x_{i,j}^{am} \cdot x_{i,j}^{aw}) + \beta_{19}^{z}(x_{i,j}^{dm} \cdot x_{i,j}^{dw}) + \beta_{20}^{z}(x_{i,j}^{rm} \cdot x_{i,j}^{rw}) + \beta_{21}^{z}(x_{i,j}^{gm} \cdot x_{i,j}^{gw}) \quad (3)$$

Our initial model formula to explain the overall difficulty rating midpoints ($\gamma_{i,j}^{Aom}$) and widths ($\gamma_{i,j}^{Aow}$) of attack hops is then

$$\gamma_{i,j}^{Aoz} = \beta_0^z + A_m^z + A_w^z + A_{mw}^z + \mu_i^z + \mu_j^z + \epsilon_{i,j}^z \quad (4)$$

where z reflects the model's outcome variable, which may be either m (midpoints) or w (widths), for expert i on hop j; β_0 denotes the fixed intercept; μ_i and μ_j denote respective random intercepts for expert and hop; and ϵ represents the error. The remaining β terms (within A_m, A_w and A_{mw}) denote the coefficients of the fixed effects of the hop attributes.

We perform likewise calculations for the evade hops (variables listed in Table 2). The sum of effects for the midpoints of the evade hops is

$$E_m^z = \beta_1^z x_{i,j}^{cm} + \beta_2^z x_{i,j}^{am} + \beta_3^z x_{i,j}^{rm}, \quad (5)$$

for the widths is

$$E_w^z = \beta_4^z x_{i,j}^{cw} + \beta_5^z x_{i,j}^{aw} + \beta_6^z x_{i,j}^{rw}, \quad (6)$$

and for the interactions is

$$E_{mw}^z = \beta_7^z(x_{i,j}^{cm} \cdot x_{i,j}^{cw}) + \beta_8^z(x_{i,j}^{am} \cdot x_{i,j}^{aw}) + \beta_9^z(x_{i,j}^{rm} \cdot x_{i,j}^{rw}) \quad (7)$$

Our initial model formula to explain the overall difficulty ratings for midpoints ($\gamma_{i,j}^{Eom}$) and widths ($\gamma_{i,j}^{Eow}$) of evade hops is then

$$\gamma_{i,j}^{Eoz} = \beta_0^z + E_m^z + E_w^z + E_{mw}^z + \mu_i^z + \mu_j^z + \epsilon_{i,j}^z. \quad (8)$$

Each of the initial models, as presented above, was then subjected to a backwards stepwise variable elimination procedure. During this, fixed effects were iteratively assessed and those that did not significantly contribute to the overall model were removed. Specifically, this process began by selection, from the pool of all non-significant fixed effects, of the effect with the t-statistic closest to zero. This variable was then removed, and the resulting model directly compared with the preceding one, using the Theoretical Likelihood Ratio test. This was implemented through the MATLAB $fitlme$ and *compare* functions. If the benefit of retaining the variable in question was calculated to be non-significant, then the model with the lower Bayesian Information Criterion (BIC) was retained into the next iteration. This procedure continued until a final model was determined, within which all fixed effects were statistically significant.

3 Results

3.1 Attack Hops

Table 3 shows all effects retained in the final model with the outcome variable of overall attack hop difficulty (midpoints), following the stepwise variable reduction process. These results indicate that a number of factors make a substantial contribution. Attacks were rated less difficult if they are frequently reported or have a large availability of tools. By contrast, attacks were rated as more difficult if they have a greater inherent difficulty or relate to more mature technologies. Attacks were also rated as more difficult when technological maturity was uncertain and, perhaps surprisingly, when easier to go unnoticed. The latter might relate to some underlying factor - for instance, some attacks may be difficult to conduct, but also difficult to detect. The significant interaction term $(m \cdot w)$ indicates a combined effect of reported tool availability and uncertainty around this. This likely reflects that a hop is rated as being more difficult to attack when experts are certain about availability being low, but less difficult when experts are certain about availability being high. Unsurprisingly, the inherent difficulty rating was found to have the most robust effect.

Table 3. Results showing significant effects of hop attribute midpoints (m), widths (w) and two-way interactions ($m \cdot w$) on midpoints of overall attack hop difficulty ratings.

Fixed Effects Estimates	β	SE	t	p
Intercept : $(_0)$	.012	.066	.175	.861
Frequency m : $(x_{i,j}^{fm})$	-.223	.044	-5.065	$<$ .001
Availability Tool m : $(x_{i,j}^{am})$	-.201	.044	-4.574	$<$ .001
Inherent Difficulty m : $(x_{i,j}^{dm})$	.357	.030	11.890	$<$ .001
Maturity m : $(x_{i,j}^{rm})$	.126	.030	4.159	$<$ .001
Going Unnoticed. m : $(x_{i,j}^{gm})$	.142	.027	5.194	$<$ .001
Maturity w : $(x_{i,j}^{rw})$	.071	.027	2.612	.009
Availability Tool $m \cdot w$: $(x_{i,j}^{am} \cdot x_{i,j}^{aw})$	.077	.036	2.168	.031
Random Effects Estimates	μ			
Expert intercept $(_i)$	.183			
Hop intercept $(_j)$	.204			
Residual $\epsilon_{i,j}$	.502			

N = 532, DF = 524, AIC = 896.7, BIC=943.6

Table 4 shows all effects retained in the final model with the outcome variable of uncertainty surrounding overall attack hop difficulty (widths). Even more factors were retained in this final model. Results indicated that experts were more certain about the vulnerability of hops on which attacks were reported more frequently - likely due to familiarity. They were also more certain regarding hops that relate to mature technologies or when tool availability is low. By contrast,

Table 4. Results showing significant effects of hop attribute midpoints (m), widths (w) and two-way interactions ($m \cdot w$) on widths of overall attack hop difficulty ratings.

Fixed Effects Estimates	β	SE	t	p
Intercept : $(_0)$	-.031	.036	-.857	.392
Frequency m : $(x_{i,j}^{fm})$	-.116	.045	-2.614	.009
Availability Tool m : $(x_{i,j}^{am})$	.131	.045	2.934	.003
Maturity m : $(x_{i,j}^{rm})$	-.093	.031	-3.013	.003
Frequency w : $(x_{i,j}^{fw})$	.141	.035	4.034	$<$ .001
Availability Tool w : $(x_{i,j}^{aw})$	.095	.039	2.420	.016
Inherent Difficulty w : $(x_{i,j}^{dw})$	.406	.037	10.959	$<$ .001
Going Unnoticed w : $(x_{i,j}^{gw})$	.268	.036	7.399	$<$ .001
Maturity $m \cdot w$: $(x_{i,j}^{rm} \cdot x_{i,j}^{rw})$	-.122	.035	-3.484	$<$ .001
Going Unnoticed $m \cdot w$: $(x_{i,j}^{gm} \cdot x_{i,j}^{gw})$	-.080	.035	-2.270	.024
Random Effects Estimates	μ			
Expert intercept $(_i)$	.127			
Hop intercept $(_j)$	.000			
Residual $\epsilon_{i,j}$	.609			

N = 532, DF = 522, AIC = 1066.3, BIC=1121.7

overall uncertainty significantly increased in line with attribute uncertainty for reported attack frequency, tool availability, inherent difficulty, and ease of going unnoticed. Two interaction terms were also retained, indicating that the effects of uncertainty around these attributes on overall uncertainty were significantly modulated by attribute rating, or vice versa. These can be interpreted together with main effects, depending upon direction. For example, in the case of ease of going unnoticed, there is a relatively large positive main effect of attribute uncertainty and a smaller, but significant, negative interaction term. This indicates that overall ratings were most uncertain when going unnoticed was considered difficult but uncertain. However, overall ratings were most certain when going unnoticed was considered difficult with certainty. The effect of attribute uncertainty was reduced, though still substantial, around hops rated as easier to go unnoticed when attacking. For maturity, however, there is a negative main effect of attribute rating and a negative interaction term, of comparable size. This indicates that overall uncertainty was greatest when maturity was rated low but uncertain. Also, while an increase in maturity rating tended to increase the certainty of overall ratings, this effect was driven by cases in which maturity was itself uncertain. Of all effects in this analysis, uncertainty surrounding the inherent difficulty rating was found to be the most robust.

3.2 Evade Hops

Table 5 shows all effects retained in the final model with the outcome variable of overall evade hop difficulty (midpoints). Four fixed effects were retained. Experts

Table 5. Results showing significant effects of hop attribute midpoints (m), widths (w) and two-way interactions ($m \cdot w$) on midpoints of overall evade hop difficulty ratings.

Fixed Effects Estimates	β	SE	t	p
Intercept : $(_0)$	-.023	.133	-.173	.863
Availability Information m : $(x^{am}_{i,j})$	-.240	.049	-4.895	$<$ .001
Maturity m : $(x^{rm}_{i,j})$	.177	.051	3.459	$<$ .001
Availability Information w : $(x^{aw}_{i,j})$	.142	.049	2.878	.004
Complexity $m \cdot w$: $(x^{cm}_{i,j} \cdot x^{cw}_{i,j})$	-.105	.053	-1.993	.047
Random Effects Estimates	μ			
Expert intercept $(_i)$	.457			
Hop intercept $(_j)$	.340			
Residual $\epsilon_{i,j}$	.772			

N = 418, DF = 413, AIC = 1081.8, BIC=1114.0

Table 6. Results showing significant effects of hop attribute midpoints (m), widths (w) and two-way interactions ($m \cdot w$) on widths of overall evade hop difficulty ratings.

Fixed Effects Estimates	β	SE	t	p
Intercept : $(_0)$	-.000	.058	-.000	$>$ .999
Complexity w : $(x^{cw}_{i,j})$	.241	.046	5.200	$<$ .001
Availability Information w : $(x^{aw}_{i,j})$	.440	.045	9.683	$<$ .001
Maturity w : $(x^{rw}_{i,j})$	.134	.045	2.982	.003
Random Effects Estimates	μ			
Expert intercept $(_i)$	.070			
Hop intercept $(_j)$	.159			
Residual $\epsilon_{i,j}$	.643			

N = 418, DF = 414, AIC = 863.0, BIC=891.2

rated a hop as less difficult to evade when more information is available to aid with this, but more difficult to evade when they were uncertain about the availability of such information. Overall evasion difficulty was also higher for hops relating to more mature technologies. A negative interaction term was evident for ratings of hop complexity, with certainty around a more complex hop being associated with it being more difficult to evade, but certainty around a hop being less complex associated with it being easier to evade. The availability of information was found to be have the most robust effect.

Table 6 shows all effects retained in the final model with the outcome variable of uncertainty surrounding overall evade hop difficulty (widths). These results show that experts were more uncertain in their overall hop rating when they were more uncertain about each of the three attributes of a given hop: complexity, information availability, or maturity. However, no significant main effects

of attribute rating position, nor any interaction terms were found. Uncertainty surrounding information availability was found to have the most robust effect.

4 Conclusions

We analyse ratings provided by cyber-security experts that pertain to a range of potentially important component (hop) attributes, previously identified as commonly occurring within attack vectors of mainstream government cyber-systems. Importantly, these ratings were obtained through interval-valued responses, which enable experts to indicate both their rating and the uncertainty associated with this rating in a single, integrated response.

Our analyses provide a 'proof of concept' for interval-valued data capture applied to the field of cyber-security. We identify key factors that contribute to both component vulnerability and uncertainty, depending on whether a hop requires compromising or only bypassing. For example, the availability of information has the largest impact on the overall difficulty of evading a component, while uncertainty around the inherent difficulty of an attack has the largest impact on its overall uncertainty.

Uncertainty in experts' attribute ratings is found to be valuable in determining not only overall uncertainty, but also overall hop vulnerability. In a number of specific cases, this information explained variance over and above the discrete midpoints of attribute ratings. For instance, when predicting the overall difficulty of attack hops, uncertainty around the maturity of technology of a given hop was associated with a significant increase in difficulty rating for that hop. In other cases, we found that in order to best explain overall difficulty ratings it was necessary to consider an interaction effect, between the position and width of responses. Sometimes, the combination of both factors provided a better predictor than either did alone. It was the novel use of an interval-valued response format that made it possible to coherently capture this uncertainty, alongside traditional ratings.

This study provides initial empirical evidence for the potential added-value offered by capturing interval-valued responses to model expert uncertainty. This is demonstrated in the case of modelling vulnerabilities within cyber-systems comprising multiple components, each with varying attributes, as is characteristic of cyber-physical systems. The benefit of using interval-valued responses was found using a comparatively low-complexity linear modelling approach. While such an approach is unsuited to capturing varying effects of responses along the response scale (for instance, it cannot account for a tendency for responses to saturate towards the extremities), its simplicity facilitates interpretation of the results. In future work, we will investigate the use of generalised additive models, within which the relationships between independent and dependent variables may be non-linear. Additionally, we are pursuing an ongoing programme of research to demonstrate the efficacy of interval-valued responses, both in terms of capturing uncertainty and improving predictive power, with reference to real-world ground-truth.

References

1. Aven, T., Renn, O.: On risk defined as an event where the outcome is uncertain. Journal of risk research **12**(1), 1–11 (2009)
2. Black, P.E., Scarfone, K., Souppaya, M.: Cyber security metrics and measures. Wiley Handbook of Science and Technology for Homeland Security pp. 1–15 (2008)
3. CESG: Extract from HMG IA Standard No.1 Business Impact Level Tables. CESG (2009)
4. Choi, H.H., Cho, H.N., Seo, J.W.: Risk assessment methodology for underground construction projects. Journal of Construction Engineering and Management **130**(2), 258–272 (2004)
5. Duan, Y., Cai, Y., Wang, Z., Deng, X.: A novel network security risk assessment approach by combining subjective and objective weights under uncertainty. Applied Sciences **8**(3) (2018). https://doi.org/10.3390/app8030428, http://www.mdpi.com/2076-3417/8/3/428
6. Feng, N., Li, M.: An information systems security risk assessment model under uncertain environment. Applied Soft Computing **11**(7), 4332–4340 (2011)
7. Fielder, A., Konig, S., Panaousis, E., Schauer, S., Rass, S.: Uncertainty in cyber security investments. arXiv preprint arXiv:1712.05893 (2017)
8. FIRST: Cvss v3.0 specification document, https://www.first.org/cvss/specification-document
9. Gao, H., Zhu, J., Li, C.: The analysis of uncertainty of network security risk assessment using dempster-shafer theory. In: 2008 12th International Conference on Computer Supported Cooperative Work in Design. pp. 754–759. IEEE (2008)
10. Gardner, D.: Risk: The science and politics of fear. Random House (2009)
11. Hubbard, D.W., Seiersen, R.: How to measure anything in cybersecurity risk. John Wiley & Sons (2016)
12. Kahneman, D., Slovic, S.P., Slovic, P., Tversky, A.: Judgment under uncertainty: Heuristics and biases. Cambridge university press (1982)
13. Koubatis, A., Schonberger, J.Y.: Risk management of complex critical systems. International journal of critical infrastructures **1**(2-3), 195–215 (2005)
14. Linda, O., Manic, M., Vollmer, T., Wright, J.: Fuzzy logic based anomaly detection for embedded network security cyber sensor. In: 2011 IEEE Symposium on Computational Intelligence in Cyber Security (CICS). pp. 202–209. IEEE (2011)
15. Mell, P., Scarfone, K., Romanosky, S.: A complete guide to the common vulnerability scoring system version 2.0. In: Published by FIRST-Forum of Incident Response and Security Teams. vol. 1, p. 23 (2007)
16. Miller, S., Appleby, S., Garibaldi, J.M., Aickelin, U.: Towards a more systematic approach to secure systems design and analysis. International Journal of Secure Software Engineering (IJSSE) **4**(1), 11–30 (2013)
17. Miller, S., Wagner, C., Aickelin, U., Garibaldi, J.M.: Modelling cyber-security experts' decision making processes using aggregation operators. computers & security **62**, 229–245 (2016)
18. Munir, R., Disso, J.P., Awan, I., Mufti, M.R.: A quantitative measure of the security risk level of enterprise networks. In: 2013 Eighth International Conference on Broadband and Wireless Computing, Communication and Applications. pp. 437–442. IEEE (2013)
19. Sikos, L.F.: Handling uncertainty and vagueness in network knowledge representation for cyberthreat intelligence. In: 2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). pp. 1–6. IEEE (2018)
20. Slovic, P.: The perception of risk. Routledge (2016)