

What Makes an Industrial Control System Security Testbed Credible and Acceptable? Towards a Design Consideration Framework

Uchenna D. Ani^a and Jeremy M. Watson

Department of Science Technology Engineering and Public Policy, University College London, U.K.

Keywords: ICS Testbeds, Security Modelling & Simulations, Security Modelling, Cyber Security Simulations, Model Credibility, Model Acceptance.


Abstract: The convergence of Industrial Control System (ICS) with Information Technologies (IT) coupled with the resulting and widely publicized cyber security incidents have made ICS security and resilience issues of critical concern to operators and governments. The inability to apply traditional IT security practice to ICSs further complicates the challenges of effectively securing critical industrial systems. To investigate these challenges without impacting upon live system operations, testbeds are being widely used as viable options to explore, develop and assess security risks and controls. However, how an ICS testbed is designed, and its attributes, can directly impact not only on its viability but also its credibility and acceptance for use as a whole. Through a systematic review and analysis of ICS security testbed design factors, a novel outline conceptual mapping of design factors for building credibility and acceptance is proposed. These design considerations include: design objectives, implementation approach, architectural component coverage, core operational characteristics, and evaluation approach.

1 INTRODUCTION

Industrial Control Systems (ICSs) play crucial roles in operating and controlling critical infrastructure (CI) systems that support essential societal services such as: transport, energy/power, and water treatment. Although technological advances have improved the functionalities of ICSs through their design, setup, and operational scope, other more challenging issues exist. ICS deployment life cycles can span decades – leading to outdated and insecure legacy systems, connecting with modern, more secure deployments (Sadeghi, Wachsmann and Waidner, 2015). These introduce new security risks in ICSs, which have become issues of growing concern due to the challenges posed by cyber-attacks (Knowles *et al.*, 2015). ICSs that are part of CIs must exhibit high levels of system safety, security, and resilience. These requirements relate to factors where the impact of ICS failures on environmental and human safety, economic and national security viewpoints would be highly damaging (Candell, Zimmerman and Stouffer,

2015). As a consequence it is often not feasible to test these requirements on live operational ICS environments as severe functional disruptions can occur (Ani, He and Tiwari, 2017).

To overcome the concerns about conducting research and testing on live ICSs, modelling and simulation (M&S) provides a viable alternative using testbeds to replicate ICS environments for exploring and addressing cyber security challenges (Davis and Magrath, 2013). This can be approached in various ways, providing exploratory platforms where training, experimentation, data aggregation, and evaluation can be safely performed, while avoiding undesirable interference, performance degradation, and/or damage to system operations (Pahi, Leitner and Skopik, 2017). This process is referred to as ‘*simulation*’, and is widely acknowledged as effective in experimenting, studying, analysing, and developing ICS security practices (Frank, Leitner and Pahi, 2018). While ‘*simulation*’ is common term in the ICS community; it is often referred to as *testbeds* – creating an experimental platform for executing

^a <https://orcid.org/0000-0001-6064-480X>

activities and processes as if in real life. The motive is to avoid compromising system performance and reliability on live ICS (Holm, Ekstedt and Andersson, 2012). Hence for the remainder of this paper we will use 'testbeds' as an all-inclusive term.

ICS testbeds are being used to evaluate or address security-related challenges (Holm *et al.*, 2015), and follow different development approaches (Frank, Leitner and Pahi, 2018). Mostly, the testbeds are unavailable for public use, or focus on specific sectors/applications, hence they lack detailed information about their use and results. Also, they are characterised by dissimilar approaches to M&S security research (Craggs *et al.*, 2018). Thus, the trustworthiness and reliability of such works are unclear. Also, there are no available design benchmarks, or a structured and clearly defined set of design considerations to provide guidance on ICS security testbeds. This presents a challenge when developing capabilities to support research objectives, and when evaluating the quality of a testbed and related research. Often, it is difficult to state or demonstrate how these works strongly support or improve confidence in reproducing real ICS scenarios. Thus, a benchmarking structure is required to guide the proof of relevance and effectiveness of ICS security testbeds and associated works (Gardiner *et al.*, 2019).

To address this challenge, this study draws from a systematic study of existing ICS security testbed works to identify the relevant design factors that can provide guidance on ICS security testbed development, and can support confidence on the trustworthiness and use of such testbeds and associated research outcomes. An outline and mapping of design factors and attributes is then proposed that can support the above need. This can assist security developers and decision-makers in determining suitable design approaches and/or factors peculiar to their requirements. It can also support establishing and/or enhancing acceptability requirements for security-related ICS testbed work.

1.1 Acceptance of Simulations Models and Testbeds

M&S acceptance is a common topic in various fields including computing, engineering, communications and psychology. It has been linked to; 'believability', 'reliability', 'trustworthiness', 'accuracy', 'objectivity' and 'any combinations thereof' (Hilligoss and Rieh, 2008). Acceptance means agreeing with a claimed assertion that a simulation model is a reasonable representation of reality. This requires a higher level of belief in a

M&S with evidence supporting such claim (Patterson and Whelan, 2017). This builds M&S credibility, which depends on perceived quality from simultaneous multiple dimensions of evidence, enabling the assessments of trustworthiness and expertise (Fogg and Tseng, 1999). Credibility comes from confidence in the lack of biased knowledge, skills and experience of the source, and centres on attributes that make a context likely or worthy of being believed, (Young Rieh and Hilligoss, 2008). It can be based on presumption, reputation, surface appearance, or first-hand experience (Fogg and Tseng, 1999).

From an M&S perspective, credibility and acceptance are often expressed by the disposition to agree, and base decisions on the attributes (information) acquired from a model (Schruben, 1980). Hence, judging simulation credibility to drive acceptance relies on clear data collection and good documentation by developers. In terms of testbeds, this describes how well a security testbed (setup, system, process, and/or output) expresses and promotes confidence, belief and acceptance as a reasonable representation of a real system (Law, 2009). It includes a testbed's suitability for exploring realistic scenarios. Thus, the absence of certain design attributes that commonly contribute to the reasonable representativeness of an actual system can degrade credibility and inhibit acceptance. In ICS security simulation testbeds, this can make or mar an accurate understanding, analysis and the resolution of security and safety issues.

2 RELATED WORK

Security testbed works were reviewed considering their close alignment with variants of ICSs, and relevant cyber security measures. From reviewing prior testbeds, a conceptual design of a cyber-physical production system testbed is presented by Salunkhe *et al* (2018). Testbeds were analysed based on their application with cyber security being a dominant area of focus.

A three scenario-based ICS security testbed demonstration is described (Candell Jr., Zimmerman and Stouffer, 2015). Details of processes, components, architecture, protocols, modelling approach, and security context were muddled. The lack of a clear outline for structural design attributes makes it difficult to easily recognise relevant design requirements and attributes. The work also fails to consider some credibility-supporting factors that can improve acceptance such as 'evaluation modes and

outputs' (McLean *et al.*, 2011; Government Office for Science, 2018). Lacking corroboration from external parties on the testbed quality, such works can be presumed to lack sufficient evidence to support claimed representativeness and can be considered weak. Holm *et al.* (2015) surveyed thirty ICS testbeds proposed for scientific research. Most of them were designed for: vulnerability analysis, testing defence mechanisms, and educational purposes. Simulation fidelity was more heavily emphasised, and factors like repeatability and safe execution were not well-addressed by the surveyed testbed articles.

Design factors for security testbeds is a topic of interest across relevant communities and stakeholders. However, what is not seen in community discussions and literature is the context of relevant requirements that can help to support design credibility and acceptance, or how this may be achieved. While design and development considerations have appeared directly or implicitly in some works, they are fragmented. This limits the ability to identify a broader set of critical requirements that support confidence and acceptance of ICS security testbed design and associated research. Clearly, there are design factors and guidelines which can be considered and followed to build ICS testbeds that can appear credible. However, there are diverse views on how this may be achieved. Considering such diversities, having a uniform and structured outline of design considerations becomes necessary. This can help in organising existing ideas. A mapping outline can provide a focal guide for developing future ICS security testbeds with simplified uniform evaluation. This can also shape the direction for benchmarking and standardising testbed design.

3 METHODOLOGY

A systematic review (Grant and Booth, 2009) was used in this work starting with unstructured searches on Google on 'ICS security', and applying related titles for a period covering 2008 to 2019. Relevant keywords were identified and applied to a structured search in SCOPUS and Web of Science (WoS) databases to identify relevant articles. This enabled the benefits of access to a wider resource coverage and concurrency (Salisbury, 2009). Keywords used involved Boolean combinations of 'ICS' OR 'SCADA' AND 'Security Testbeds' OR 'Testbeds' to identify relevant literature on ICS security-related testbed works.

77 articles were identified based on their match to applied search parameters. Each article was considered for relevance from reading the title and abstract. Duplications were discarded, and 41 articles were found useful.

3.1 Design Factors

ICS security testbeds that can support evidence-based decision-making on security policies and controls typically rely on an assured correctness representation of system and operations. This increases confidence in the testbed's reliability to satisfy specific purposes (Government Office for Science, 2018). Typically, a testbed's composition would depend on the knowledge being drawn. The testbed's quality and the confidence it can evoke depends on the quality of the underpinning theory. Hence relevant existing literature on ICS security testbeds was reviewed based on its ability to capture design attributes.

NIST SP 800-82 Rev 2 (Stouffer *et al.*, 2015) presents a broad guide to ICS security. The 'Basic ICS Architecture' is shown to consist of; Physical Process (PP), Field Devices (FD), Communications Gateway (CG), and Control Centre (CC) component functionality groups. In M&S, well-defined 'design/usage objectives' – the purpose(s) intended for a testbed prior to its development – often drive good design considerations (McDonald and Richardson, 2009). Prior work (Holm *et al.*, 2015) highlights that representativeness and cost trade-offs amongst other factors explain why design considerations and decisions need to be guided by a testbed's intended use. *Usage objectives* need to align with design configurations (Holm *et al.*, 2015), and defined early to direct and scope the development process, eliminate ambiguity, and support functional and operational validity.

Other factors that appear to significantly contribute to the reliability and trustworthiness of ICS security testbed projects include: 'component architecture' and 'operational requirements' (Zhao, Peng and Xie, 2013; Vaughn and Morris, 2016). *Component architecture* refers to the combination of component functionality groups that are part of ICS setup as outlined in NIST SP 800-82 Rev 2. Again, this are; PP, FD, CG, and CC, including aspects of communications protocols; IP routable and/or IP non-routable protocols (Candell, Zimmerman and Stouffer, 2015). Including these types of information at more granular levels can help to achieve better clarity of security issues around specific components. This can enable a sense of the broader security implications of certain component issues for the

entire ICS network, and better representation of an ICS from both architectural and operational viewpoints. All of these promote acceptance of the resulting testbed and the research activities that use it.

Operational Requirement refer to the features that underpin the structure and operations of a testbed, which can be functional or non-functional (McLaughlin *et al.*, 2016). *Functional features* refer to behavioural attributes of testbed operations including the ability to; mirror real system form (*fidelity*), add or remove components or test scenarios (*modularity*), and log status of test scenarios (*monitoring and logging*). *Non-functional features* refer to performance attributes of testbeds, and include the ability to; easily use the testbed for an intended application (*usability*), adapt to new applications or scenarios (*adaptability*), and be open to improvements and modifications (*scalability*), normally drawn from functional features (Siboni *et al.*, 2019). The relevance of these operational requirements in supporting the credibility and acceptance of a testbed has been acknowledged (Holm *et al.*, 2015; McLaughlin *et al.*, 2016). Demonstrating functional and non-functional features in an M&S testbed contributes to assurance, hence promotes trustworthiness, credibility, and acceptability.

The *design approaches* for simulation testbeds also contribute to testbed reliability (Vaughn and Morris, 2016). This refers to the structural formation of components that make up a simulation testbed. Typically, this can take one of three approaches: (i) *Physical Simulation (PS)*; involving real components (hardware and/or software), (ii) *Semi-Physical Simulations (SPS)*; involving a combination of real and emulation and/or virtualisation of ICS components, (iii) *Software-based Simulations (SBS)*; involving a purely software simulation of ICS (Zhao, Peng and Xie, 2013). These three testbed forms are often referred to as; *real system, computer emulations or virtualisation (including hardware-in-the-loop)*, and *pure software-based simulations* (Holm *et al.*, 2015) or *live, virtual, and constructive simulations* (Kavak *et al.*, 2016). Selecting an approach is often influenced by the desired degree of representativeness, cost and time for development (Zhao, Peng and Xie, 2013), and the experience or expertise of those involved (GSE Systems, 2017).

Experience/expertise is particularly crucial as it can affect the level of detail in a simulation testbed. Physical, real, or live simulations typically enable the nearest representation of the real system, hence these are most likely to lend better credibility and acceptance than the other two. The *evaluation process* for ICS security testbeds can also influence

design quality (McLean *et al.*, 2011). This refers to the procedure(s) for performing assessments to understand how well the testbed design or related outputs are correct, and(or) acceptable. To drive simulation testbed reliability, scenarios, data, and outcomes all rely on suitable evaluations, and demonstration of proof that relevant reliability factors and operational characteristics are achieved, especially for the intended use (McLean *et al.*, 2011). This proof is good when produced by authors and implementers (*a verification*). It can be rather better if produced by sources other than the authors (*a validation*). The most valuable endorsements are from public institutions, standards or certification agencies (*an accreditation*) (DoD, 2010).

Thus, in security testbed M&S, an evaluation process can transition from *verification* to *validation* and to *accreditation* over time; demonstrating an incremental appraisal process, building stronger evidence and grounds for incremental trust and acceptance. Such multi-level evaluation processes for simulation models and testbeds can help provide a secure basis for critical decision-making.

This study draws insights from reviewing 41 relevant articles to identify and evaluate the significance of the following key criteria for building reliable and acceptable ICS security testbeds. These include: (i) *Design Objectives*, (ii) *Implementation Approach*, (iii) *Architectural Coverage*, (iv) *Operational Requirements*, and (v) *Evaluation Process*.

4 ANALYTICAL RESULTS

4.1 Design Objectives

In determining the design objectives from both explicit and implicit descriptions, eight broad themes were found. These include *threat analysis, vulnerability analysis, attack analysis, impact analysis, defence mechanism test/analysis, education and training, creation of policies and(or) standards, and performance/quality of service analysis* (See Table 1).

These results overlap with the outcomes of similar evaluations (Holm *et al.*, 2015), and suggest that most ICS security-related works seem to address a common range of security areas including the investigation of; cyber-attack feasibilities, effectiveness of security controls and defence mechanisms, and the consequences and impacts of successful cyber-attacks and failed security measures. Interest in understanding security vulnerabilities is also found. Note that it is good practice to define at the beginning, the key

objectives in scope and application, which can help clarify understanding of the requirements for testbed design and implementation.

4.2 Implementation Approach

The ICS testbed simulation approaches found can be broadly categorised into three groups. These include; (i) *Physical Simulation (PS)*, which involves purely real infrastructure components; (ii) *Semi-Physical Simulations (SPS)*, which involves combining real, emulated and/or virtualised abstractions of ICS components; and (iii) *Software-based Simulations (SBS)*, which involves simulating ICS components on purely software-based platform. These approaches may also be combined to achieve a desired hybrid system setup (Lu *et al.*, 2014). Results in Table 1 show the different implementation approaches. Combining multiple simulation approaches seems to be more favoured than using them singly. This might be motivated by the possibility of gaining greater representativeness, and/or having one approach compensate for the limitations of another. However, the expertise/experience of developers, development time, and costs are also influencing factors in deciding on an appropriate implementation approach.

4.3 Architectural Component Coverage

An ICS architecture can be categorised into four functional areas: (i) *Physical Process (PP)*, (ii) *Field Devices (FD)*, (iii) *Communications Gateway (CG)*, and (iv) *Control Centre (CC)*. These can be implemented at different levels of abstraction in a system; they constitute the basic elements of a typical ICS needing to be considered in any design project.

From Table 3, nearly half of the reviewed works reflect design/component structures that covered all four functional ICS areas. Between 3 and 11 works covered three functional areas, and on average 2 works covered two functional areas. *Communications gateway (CG)* appears to be the most widely included functional area, with a frequency of 95.12% of all the works reviewed. Covering all the functional areas in a testbed architecture can support representing wider contexts which lends to a better architectural and operational specification of a real ICS.

4.4 Core Operational Characteristics

Fourteen categories of re-occurring operational characteristics emerged as shown in Table 2. Analysis shows that 70.7% of the reviewed works contained statements associated with one or more of the

fourteen categories of operational characteristics. ICS security testbed *'fidelity'* (41.46%) appeared to be most highly acknowledged. This is followed by *'scalability'* also referred to as *'extensibility'* with 26.83%. For credibility that can drive acceptance and use, it is important for ICS security testbeds and research to demonstrate some (if not most) of the identified operational characteristics (Holm *et al.*, 2015). This can increase the confidence of adopters, decision-makers, and other stakeholders to trust the approach and its related outputs.

4.5 Evaluation Process

From the earlier identified mechanisms for achieving an incremental build-up of credible evidence involving verification, validation, and accreditation (DoD, 2010), results show that a significant proportion of works reviewed lacked any form of evaluation (See Table 2). This is despite the acknowledged significance of testbed evaluations. While a few works covered *verification* and *validation*, no reviewed work demonstrated any form of evaluation to the level of accreditation by any formal body to support its use. This supports earlier arguments on the lack of sufficient research emphasis and efforts to establishing trustworthiness to promote acceptance and the use of ICS security testbeds and their associated outputs.

Table 1: Results from ICS Security Design Objectives and Implementation Approaches.

Design/Simulation Objectives (Security-Centric)	Percent of Total (%)
Attack Analysis	63.41
Defence Mechanism Tests/Analysis	56.10
Impact Analysis	41.46
Vulnerability Analysis	36.59
Education and Training	24.39
Threat Analysis	9.76
Performance/QoS Analysis	2.44
Creation of Policies and(or) Standards	2.44
Design/Simulation Approach:	Percent of Total (%)
SBS + SPS	21.95
SBS + SPS + PS	19.51
PS	17.07
SBS	17.07
SPS + PS	12.20
SPS	12.20
<i>Key to notations:</i>	
- <i>Software-Based Simulation</i> = SBS,	
- <i>Semi-Physical Simulation (Emulation or Virtualisation / HIL)</i> = SPS, <i>Physical Simulation</i> = PS,	
- '+' used to reflect combination of approaches.	

Table 2: Analysis of Testbed Evaluation Process.

Process Category	Evaluation Method	Percent of Total (%)
-	Not Mentioned	56.10
<i>Verification</i>	User-defined Requirements	14.63
<i>Validation</i>	Standards and Reference Model	19.51
	Prior Works	4.88
	Real ICS	2.44
-	Unreferenced Architecture	2.44

Table 3: Analysis of Architectural Component Simulation Coverage.

Design/Simulation Coverage:	Percent of Total (%)
CC + CG + PP + FD	46.34
CC + CG + PP	26.83
PP + CG + FD	9.76
CC + CG + FD	7.32
CC + FD	4.88
CC + CG	4.88
Characteristics	Percent of Total (%)
Fidelity	41.46
Scalability or Extensibility	26.83
Flexibility or Adaptability	19.51
Reproducibility or Repeatability	19.51
Modularity	17.07
Cost-Effectiveness	9.76
Measurability & Measurement Accuracy	9.76
Isolation or Safe Execution	7.32
Usability	4.88
Diversity	4.88
Interoperability	2.44
Monitoring & Logging	2.44
Openness	2.44
Complexity	2.44
<i>Key to notations:</i>	
- Communications Gateway = CG	
- Physical Process = PP, Control Centre = CC	
- Field Device/Components = FD	
- '+' used to reflect combination of component classes covered.	

5 MAPPING CREDIBILITY-SUPPORTING DESIGN FACTORS FOR ICS SECURITY TESTBEDS

From the results, the occurrence and frequency of certain factors and themes suggest their reasonable relevance in ICS security testbed works. We find that; (i) clearly defined security-related design/usage

objectives and scenarios, (ii) the type(s) of modelling and simulation approach(es) and the degree of abstractions applied, (iii) the architectural design components covered, (iv) the indication and clear mappings to core operational system characteristics, and (v) the testbed evaluation modes covered, can all contribute to building the credibility, acceptance, and use of ICS security testbeds and(or) associated research.

Clearly defined security-related objectives defined from the outset can direct an ICS testbed development process and can support the evaluation of reliability. Clearly indicating the testbed simulation approach clarifies understanding of the tools and techniques being used and their simulation capabilities. Defining simulation approach(s) also provides crucial detail that can support reproducing and evaluating security testbed M&S and associated research. Defining the various architectural components, combined with defined simulation testbed approach(s), enhances the potential for achieving scientific research rigour that is repeatable with reproducible results. Having an evaluation step helps to demonstrate endorsements from authors or external parties on the reliability of an ICS security testbed's claim. All of these come together to create or add confidence and credibility to any claimed quality state of an ICS security testbed. A mapping structure is provided in Fig 1 showing how the above factors can be considered during design process to advance belief and acceptance based on the core operational characteristics found. *Fidelity* in M&S emphasises the degree of structural, operational, and process correlation between a testbed or test predictions, and real-world observations (Alves, Das and Morris, 2016) and can be demonstrated by defining a testbed's *simulation design approach*. Commonly, Physical simulation (PS) tends to provide the most fidelity (Kavak *et al.*, 2016). *Scalability or Extensibility* refers to a testbed's ability to be easily expanded in functionality and size (Hahn *et al.*, 2013). This can be defined in *design objective(s)* and architectural components covered and demonstrates a capacity to add or migrate components from testbeds without any significant need for entire system re-design. *Flexibility or Adaptability* refers to a testbed's dynamic attribute of supporting re-definition and repurposing for alternative use case(s) (Kavak *et al.*, 2016). For example, showing in *design objectives* that a testbed initially meant to analyse vulnerabilities, can be easily re-purposed for analysing security impacts. *Repeatability or Reproducibility* describes a testbed's ability to produce consistent results when replicated (Koutsandria *et al.*, 2015). This can be

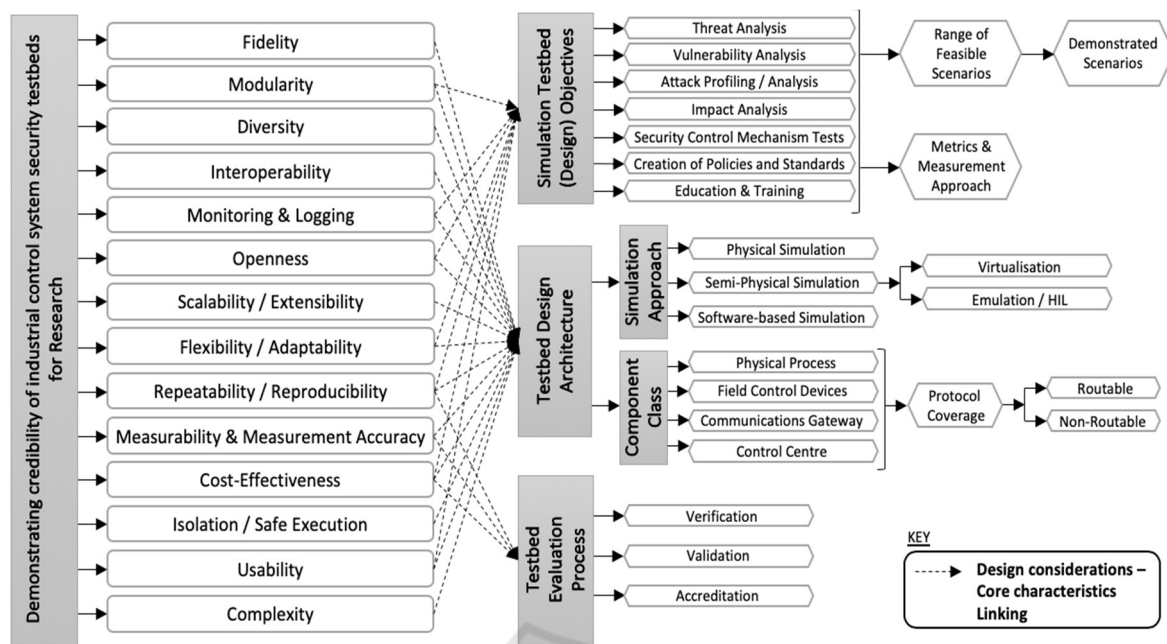


Figure 1: Outline and conceptual Mapping of design consideration for credible ICS testbeds.

demonstrated through appropriate definitions of *architectural components*, *process configurations*, and *design objectives*; enabling others to easily reproduce experimental parameters and obtain similar results. *Modularity* allows a testbed to be efficiently changed to accommodate new design or functional requirements (Ahmed *et al.*, 2016). This enables ICS testbeds to accept continuous improvements; indicated in *design objective(s)*. Modularity can help structurally, to realise incremental validation, credibility, and acceptance.

Measurability and Measurement accuracy are two linked attributes. *Measurability* describes the capacity of a testbed to enable the quantification of tests. Measurement accuracy quantifies the degree to which measurement processes interfere (or not) with corresponding results (Alves, Das and Morris, 2016). Both attributes can be achieved by including tools and defining relevant metrics in the *architectural component* definition to support authentication of performance. *Interoperability* testbed attributes describe the capacity to support any combination of *simulation approaches*; software-based, semi-physical or purely physical simulation, to interface, communicate, and use information for desired objectives. *Cost-effectiveness* describes the attributes of a testbed to be within defined budgets (Holm *et al.*, 2015). Often, there are trade-offs between testbed development cost and its fidelity (Gao *et al.*, 2014). The *Safe execution or Isolation* attribute pertains to real world test cases linked to testbeds. It ensures that

security M&S activities are bounded and isolated to avoid any impacts on real system functions (Bergman *et al.*, 2009). Network segmentation (Fovino *et al.*, 2010) and access control (Bergman *et al.*, 2009) can be applied at the *architectural component* level to help show that the outputs from tests are unaltered. This is a proof of testbed simulation integrity.

Usability describes the ability for a testbed to be used for purposes defined. This reduces the likelihood of testbed misuse (Holm *et al.*, 2015), and is necessary to assist users with varying competencies (contextual knowledge and skills). It can be promoted at the *architectural component* level by applying design structures and components with friendly user interfaces (Almalawi *et al.*, 2013). This can support credibility and acceptance if seen to support a wider range of users with diverse skillsets. *Diversity* describes the capacity of a testbed to include different vendor components without undermining scalability or extensibility as discussed earlier. Device heterogeneity is a common feature where IoT converges into ICS, and needs to be replicated in testbeds design. Diversity can be demonstrated at the *architectural component* and *simulation approach* levels. *Monitoring and Logging* describe testbed attributes of observing process executions and optimising data logging for security purposes (Green *et al.*, 2017; Gardiner *et al.*, 2019). These can be shown from the results of *evaluation processes*. Credibility and acceptance can be improved by the ability to record and review recorded outputs for

future reference. *Complexity and Openness* are two related attributes. *Openness* defines a simulation testbed's capacity to support data openness or remote access (Gardiner *et al.*, 2019). This can be demonstrated at *architectural component configuration* and *evaluation process* levels. *Complexity* assures that *architectural components* are represented transparently so that a single point of data access or extraction can be supported.

6 CONCLUSIONS

This paper presented a critical review and analysis to identify relevant factors that can provide guidance on ICS security testbed development and use, upon which credibility can be based. A novel conceptual mapping of design considerations for credible ICS security testbeds is presented. Building or enhancing testbed credibility typically comes mostly from architectural coverage, augmented by the adopted implementation approach, selected components, and a demonstration of some reasonable degree of evaluation. ICS security researchers and developers must strive to achieve fundamental architectures that are representative of real-world systems and can allow appropriate, yet realistic testing.

It may not be necessary or feasible to capture all the core characteristics outlined within a testbed setup. However, choosing those compliance characteristics viewed as important within a particular project may well depend on the project's core objectives and scope. Considering the available resources and capabilities, some characteristics may be incorporated or maximised at the expense of others. New attributes can also be considered based on evolving dynamics or system context.

The proposed mapping structure can promote effective and well-organized procurement of systems and sub-system components guided by clearly defined design requirements – in response to system and functional dynamics, and the endorsement of the relevant community of stakeholders. This can thereby streamline the task of setting requirements and reduce the costs of both infrastructure development and sub-system integration. Also, it can lead to greater consistency and efficiency in growing research related to ICS security testbeds. Most conveniently, combining this with the growing trend and capability for federating ICS security testbeds as keenly advocated and explored in recent initiatives, the potential to make testbeds more available and interoperable. Furthermore, it can minimise the diversity in design structures amongst different and

physically dispersed testbeds in a federated system. For future work, we will explore further acceptance-supporting design considerations for ICS testbeds. We will also explore how to build credibility-supporting ICS security testbeds, and how such models can be better evaluated to drive and simplify acceptance and use.

ACKNOWLEDGEMENTS

The research is supported by the EPSRC-funded PETRAS National Centre of Excellence for IoT Systems Cyber Security.

REFERENCES

- Ahmed, I. *et al.* (2016) 'A SCADA System Testbed for Cybersecurity and Forensic Research and Pedagogy', in *2nd Annual Industrial Control System Security Workshop on - ICSS '16*. Los Angeles, CA, USA: ACM, pp. 1–9.
- Almalawi, A. *et al.* (2013) 'SCADA-VT-A framework for SCADA security testbed based on virtualization technology', in *Conference on Local Computer Networks, LCN*. Sydney, NSW, Australia: IEEE, pp. 639–646.
- Alves, T., Das, R. and Morris, T. (2016) 'Virtualization of Industrial Control System Testbeds for Cybersecurity', in *2nd Annual Industrial Control System Security Workshop on - ICSS '16*. Los Angeles, CA, USA: ACM, pp. 10–14.
- Ani, U. P. D., He, H. (Mary) and Tiwari, A. (2017) 'Review of cybersecurity issues in industrial critical infrastructure: manufacturing in perspective', *Journal of Cyber Security Technology*. Taylor & Francis, 1(1), pp. 32–74.
- Bergman, D. C. *et al.* (2009) 'The virtual power system testbed and inter-testbed integration', in *2nd conference on Cyber Security Experimentation and Test*. Montreal, Canada: USENIX Association Berkeley, CA, USA, p. 5.
- Candell Jr., R., Zimmerman, T. A. and Stouffer, K. A. (2015) 'NISTIR 8089 - An Industrial Control System Cybersecurity Performance Testbed'. Gaithersburg, Maryland, United States: National Institute of Standards and Technology (NIST), pp. 1–47.
- Candell, R., Zimmerman, T. and Stouffer, K. (2015) 'An industrial control system cybersecurity performance testbed (NISTIR 8089)'. National Institute of Standards and Technology.
- Craggs, B. *et al.* (2018) 'A Reference Architecture for IIoT and Industrial Control Systems Testbeds.', in *2nd Conference on Living in the Internet of Things*. London, UK: Institution of Engineering and Technology (IET).
- Davis, J. and Magrath, S. (2013) *A Survey of Cyber Ranges and Testbeds*. Edinburgh South Australia.

- DoD (2010) 'Department of Defense Modeling and Simulation Best Practices Guide'. US Department of Defense.
- Fogg, B. J. and Tseng, H. (1999) 'The elements of computer credibility', in *CHI 99 Human Factors in Computing Systems Conference*. Pittsburgh, PA: ACM Press, pp. 80–87.
- Fovino, I. N. *et al.* (2010) 'An experimental platform for assessing SCADA vulnerabilities and countermeasures in power plants', in *3rd International Conference on Human System Interaction, HSI'2010 - Conference Proceedings*. Rzeszow, Poland: IEEE, pp. 679–686.
- Frank, M., Leitner, M. and Pahi, T. (2018) 'Design Considerations for Cyber Security Testbeds: A Case Study on a Cyber Security Testbed for Education', in *15th International Conference on Dependable, Autonomic and Secure Computing, IEEE 15th International Conference on Pervasive Intelligence and Computing, IEEE 3rd International Conference on Big Data Intelligence and Compu.* Orlando, FL, USA: IEEE, pp. 38–46.
- Gao, H. *et al.* (2014) 'An Industrial Control System Testbed Based on Emulation, Physical Devices and Simulation', in Butts, J. and Sheno, S. (eds) *IFIP Advances in Information and Communication Technology*. Critical I. Arlington, Virginia, USA: Springer Berlin Heidelberg, pp. 79–91.
- Gardiner, J. *et al.* (2019) 'Oops I Did it Again: Further Adventures in the Land of ICS Security Testbeds', in *Proceedings of the 2019 Workshop on Cyber-Physical Systems Security and Privacy*. London, UK: ACM.
- Government Office for Science (2018) *Computational Modelling: Technical Futures*. London.
- Grant, M. J. and Booth, A. (2009) 'A typology of reviews: An analysis of 14 review types and associated methodologies', *Health Information and Libraries Journal*, 26, pp. 91–108.
- Green, B. *et al.* (2017) 'Pains, Gains and PLCs: Ten Lessons from Building an Industrial Control Systems Testbed for Security Research', in *10th USENIX Workshop on Cyber Security Experimentation and Test (CSET '17)*. Vancouver, Canada: USENIX Association Berkeley, CA, USA, pp. 1–8.
- GSE Systems (2017) *Fidelity Matters: What 'High-fidelity' Really Means, Simulation & Training Blog*.
- Hahn, A. *et al.* (2013) 'Cyber-physical security testbeds: Architecture, Application, and Evaluation for Smart Grid', *IEEE Transactions on Smart Grid*, 4(2), pp. 847–855.
- Hillgoss, B. and Rieh, S. Y. (2008) 'Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context', *Information Processing and Management*. Elsevier (ScienceDirect), 44(4), pp. 1467–1484.
- Holm, H. *et al.* (2015) 'A Survey of Industrial Control System Testbeds', in *In: Buchegger S., Dam M. (eds) Secure IT Systems. Lecture Notes in Computer Science*. Switzerland: Springer, Cham, pp. 11–26.
- Holm, H., Ekstedt, M. and Andersson, D. (2012) 'Empirical analysis of system-level vulnerability metrics through actual attacks', *IEEE Transactions on Dependable and Secure Computing*, 9(6), pp. 825–837.
- Kavak, H. *et al.* (2016) 'A Characterization of Cybersecurity Simulation Scenarios', in *19th Communications and Networking Simulation Symposium (CNS'16)*. CA, USA: Society for Modelling & Simulation International (SCS) & ACM, pp. 1–8.
- Knowles, W. *et al.* (2015) 'A survey of cyber security management in industrial control systems', *International Journal of Critical Infrastructure Protection*. Elsevier, 9, pp. 52–80.
- Koutsandria, G. *et al.* (2015) 'A Real-Time Testbed Environment for Cyber-Physical Security on the Power Grid', in *First ACM Workshop on Cyber-Physical Systems-Security and/or PrivaCy - CPS-SPC '15*. Denver, CO, USA: ACM, pp. 67–78.
- Law, A. M. (2009) 'How to build valid and credible simulation models', in Rossetti, M. D. *et al.* (eds) *2009 Winter Simulation Conference (WSC)*. Austin, TX, USA: IEEE, pp. 24–33.
- Lu, T. *et al.* (2014) 'Cyber-Physical Security for Industrial Control Systems Based on Wireless Sensor Networks', *Downloads.Hindawi.Com*, 2014.
- McDonald, M. J. and Richardson, B. T. (2009) *Position Paper: Modeling and Simulation for Process Control System Cyber Security Research, Development and Applications, Center for Information Management, Integration and Connectivity - Position Papers*.
- McLaughlin, S. *et al.* (2016) 'The Cybersecurity Landscape in Industrial Control Systems', *Proceedings of the IEEE*, 104(5), pp. 1039–1057.
- McLean, C. *et al.* (2011) 'Modeling and Simulation of Critical Infrastructure Systems for Homeland Security Applications', *NIST Special Publications*. National Institute of Standards and Technology, p. 86.
- Pahi, T., Leitner, M. and Skopik, F. (2017) 'Data Exploitation at Large: Your Way to Adequate Cyber Common Operating Pictures', in *16th European Conference on Cyber Warfare and Security*. Reading, Berkshire, UK: Academic Conferences and Publishing International Ltd, pp. 307–315.
- Patterson, E. A. and Whelan, M. P. (2017) 'A framework to establish credibility of computational models in biology', *Progress in Biophysics and Molecular Biology*. Elsevier Ltd, 129, pp. 13–19.
- Sadeghi, A.-R., Wachsmann, C. and Waidner, M. (2015) 'Security and Privacy Challenges in Industrial Internet of Things', in *52nd Annual Design Automation Conference on - DAC '15*. San Francisco, CA, USA: ACM, pp. 1–6.
- Salisbury, L. (2009) 'Web of Science and Scopus: A comparative review of content and searching capabilities', *The Charleston Advisor*, July, pp. 5–19.
- Salunkhe, O. *et al.* (2018) 'Cyber-Physical Production Testbed: Literature Review and Concept Development', *Procedia Manufacturing*. Elsevier B.V., 25, pp. 2–9.
- Schruben, L. W. (1980) 'Establishing the credibility of simulations', *Simulation*. Simulation Councils Inc., 34(3), pp. 101–105.

- Siboni, S. *et al.* (2019) 'Security Testbed for Interne-of-Things Devices', *IEEE Transactions on Reliability*, 68(1), pp. 23–44.
- Stouffer, K. *et al.* (2015) 'Guide to Industrial Control Systems (ICS) Security - NIST.SP.800-82r2'. Gaithersburg, Maryland: NIST, US Department of Commerce, pp. 1–247.
- Vaughn, R. B. and Morris, T. (2016) 'Addressing Critical Industrial Control System Cyber Security Concerns via High Fidelity Simulation', in *11th Annual Cyber and Information Security Research Conference on - CISRC '16*. Oak Ridge, Tennessee, USA: ACM, pp. 1–4.
- Young Rieh, S. and Hilligoss, B. (2008) 'College Students' Credibility Judgments in the Information-Seeking Process', in Metzger, M. J. and Flanagin, A. J. (eds) *Foundation Series on Digital Media and Learning*. The John D. Cambridge, MA: The MIT Press, pp. 49–72.
- Zhao, W., Peng, Y. and Xie, F. (2013) 'Testbed Techniques of Industrial Control System', in *3rd International Conference on Computer Science and Network Technology, ICCSNT 2013*. Dalian, China: IEEE, pp. 61–65

