: Latex Research using Google Scholars

Kelvin Ndungu Wainaina

13 September 2021.

1 Machine Learning

1.1 Automotive safety and machine learning:

Abstract

Machine learning (ML) applications generate a continuous stream of success stories from various domains. ML enables many novel applications, also in safety-critical contexts. However, the functional safety standards such as ISO 26262 did not evolve to cover ML. We conduct an exploratory study on which parts of ISO 26262 represent the most critical gaps between safety engineering and ML development. While this paper only reports the first steps toward a larger research endeavor, we report three adaptations that are critically needed to allow ISO 26262 compliant engineering, and related suggestions on how to evolve the standard.Initial results from a study on how to adapt the ISO 26262 safety standard Share on.

2 Cryptography

2.1 cryptosystem based on DNA cryptography and randomly generated mealy machine:

Abstract

Nowadays, the amount of data produced and stored in computing devices is increasing at an alarming rate. Tremendous amounts of critical and sensitive data are transmitted between all these devices. Thus, it is very imperative to guarantee the security of all these indispensable data. Cryptography is a commonly used technique to ensure data security. The fundamental objective of cryptography is to transmit data from the sender to the receiver in the most secure way, so that an attacker is unable to extract the original data content. This paper proposes a novel cryptosystem based on Deoxyribonucleic Acid (DNA) cryptography and finite automata theory. The system is made of three entities, namely a key pair generator, a sender and a receiver. The sender generates a 256-bit DNA based secret key based on the attributes of the receiver, and this key is used for data encryption. Then, a randomly generated Mealy machine is used for coding the DNA sequence, which makes the ciphertext more secure. The proposed scheme can protect the system against numerous security attacks, such as brute force attack, known plaintext attack, differential cryptanalysis attack, cipher text only attack, man-in-the-middle attack

and phishing attack. The results and discussions show that the proposed scheme is efficient and secure than the existing schemes.

3 Articicial Inteligence

3.1 AI, you can drive my car: How we evaluate human drivers vs. self-driving cars:

Abstract

AI, you can drive my car: How we evaluate human drivers vs. self-driving cars

This study tests how individuals attribute responsibility to an artificial intelligent (AI) agent or a human agent based on their involvement in a negative or positive event. In an online, vignette experimental between-subjects design, participants (n = 230) responded to a questionnaire measuring their opinions about the level of responsibility and involvement attributed to an AI agent or human agent across rescue (i.e., positive) or accident (i.e., negative) driving scenarios. Results show that individuals are more likely to attribute responsibility to an AI agent during rescues, or positive events. Also, we find that individuals perceive the actions of AI agents similarly to human agents, which supports CASA framework's claims that technologies can have agentic qualities. In order to explain why individuals do not always attribute full responsibility for an outcome to an AI agent, we use Expectancy Violation Theory to understand why people credit or blame artificial intelligence during unexpected events. Implications of findings for practical applications and theory are discussed.

4 Cyber Security:

4.1 Visualization evaluation for cyber security:

Abstract

The Visualization for Cyber Security research community (VizSec) addresses longstanding challenges in cyber security by adapting and evaluating information visualization techniques with application to the cyber security domain. This research effort has created many tools and techniques that could be applied to improve cyber security, yet the community has not yet established unified standards for evaluating these approaches to predict their operational validity. In this paper, we survey and categorize the evaluation metrics, components, and techniques that have been utilized in the past decade of VizSec research literature. We also discuss existing methodological gaps in evaluating visualization in cyber security, and suggest potential avenues for future research in order to help establish an agenda for advancing the state-of-the-art in evaluating cyber security visualizations