Deep Learning-based Security and Usability Analysis Framework for Mobile Android Applications

Rathod Sai Vamshi Krishna[1], M. Sai Laxman Reddy[2], and N. Sanjay[3]
Department of AI & ML, Chaitanya Bharathi Institute of Technology, Gandipet, Hyderabad, India

***Abstract*—Fast mobile applications raise the importance for building effective scalable evaluation methods so as to answer issues pertaining to security issues to even those regarding usability issues. Typically, most conventional evaluation approaches do not support depth scalability thus becoming important enough to handle through advanced techniques of machine learning. The article makes a new proposition where there are large language models that also use the LSTM network providing an overarching framework used when assessing Android applications. Specif- ically, we make use of LLMs such as GPT (Generative Pre-trained Trans- former) that supports security assessment tasks, which include vulner- ability detection, code review, and identifying the patterns of security-related behaviors in application behavior. Meanwhile, we employ LSTM networks trained on user interaction data to generate usability scores in reflection of the general experience of the user. This is what synergistic integration of LLMs for security evaluation and LSTMs for usability as- sessment can offer: a dual focus approach to provide a more holistic view of an application's strengths and weaknesses. Our framework is designed to automate and scale the evaluation process, providing actionable in- sights to developers and stakeholders. In fact, our approach is validated by conducting extensive experiments on a diverse set of Android appli- cations, including security vulnerabilities and feedback from users. The results show that our approach is capable of giving accurate, timely, and cost-effective assessments of both security and usability, thus helping improve the overall evaluation process for mobile applications.**

***Index Terms*—Security Assessment · Usability Evaluation Deep Learning Models.**

I. INTRODUCTION

With mobile applications spreading rapidly into the aspects of modern life - in- cluding communication, entertainment, finance, and healthcare, just to name a few, the proliferation of such systems has been tremendous within recent years.

Android is the leading mobile operating system today and has an enormous set of applications that offer quite varied functionalities and services. Still, as the number and complexity of mobile apps are growing, so do related issues con- cerning their security and usability. Serious consequences such as unauthorized access to data, privacy breaches, and exploitation of sensitive user information may arise from the security vulnerabilities of mobile applications. Poor usabil- ity can seriously impact user engagement and satisfaction, thereby affecting the adoption of the app and its success in the long run.

Traditional approaches to mobile application evaluation rely on manual or rule-based methods for security analysis and user testing for usability assess- ments, which may not scale well. These methods are time consuming, labor intensive, and do not provide in-depth insights into the full range of potential vulnerabilities or usability issues. Moreover, as diversity in Android apps grows along with their complexity, available evaluation techniques cannot keep pace with the demands of the developers and stakeholders who are looking for fast, reliable, and automated solutions. In response to these challenges, this study proposes a synergistic deep learning-based approach that

combines LSTM-based security assessment with Google's Large Language Models (LLMs) for usability evaluation. LSTM networks are well suited to sequential data processing and can be used to both reveal security vulnerabilities and uncover patterns in code behavior. For example, this LLM, for instance GPT, is helpful in processing nat- ural language and will therefore be of use in the evaluation of the usability of an app-that is, to analyze what is being said about users from user feedback, an application description, and many more textual data associated with experience. We hope that through integration of these two techniques from deep learning to offer an automated framework that can serve all aspects of security usability related to Android applications.

The main objective of this research is to design an automated, scalable, and accurate evaluation system that can help developers and stakeholders identify security flaws and usability issues early in the development cycle, thus improv- ing app quality and user safety. This approach enhances the efficiency of the evaluation process and adds to the growing body of knowledge in applying deep learning techniques to real-world mobile app challenges.

II. LITERATURE SURVEY

In the present scenario, evaluation of security and usability in Android mo- bile applications consist of disparate methodologies with unique strengths and weaknesses. Usually, traditional tools and techniques for assessing these criti- cal aspects of mobile applications rely on static and dynamic analysis, manual testing, and user feedback, but they seem to fail when it comes to depth, scalibil- ity, or integration. With the complexity in the mobile applications, the need for more rigorous and comprehensive evaluation methods that can handle automa- tion has increased. Although there are existing methods in place, they cannot actually deliver the kind of overall and timely assessment in accordance with the increasing volumes of complexity in mobile applications presented by today's app ecosystem.

The DroidBox and AndroBugs in the family of dynamic analysis try to check an application at runtime as it appears to have some realworld usage scenario. The overall approach in the case of dynamic analysis is mostly followed for dis- covering the vulnerabilities as data leakage, excessive demand for permission, or for insecure communication with a network. In that way running an applica- tion within a controlled environment opens doors for the discovering of different behaviors of application in conditions and usage pattern. Thus, the dynamic analysis can detect certain runtime threats that may evade static analysis; it is also heavily reliant on the completeness of test cases, where it fails to find any other forms of vulnerabilities based on complex interrelations among various app components or based on user-dependent contexts.

This involves manual testing as well as security audits, during which a team of experts will analyze the code of the app along with permissions and total security measures taken. This would afford scope to have more accurate scrutiny on potential security weaknesses at their finest grain, if for instance the automated techniques can ignore even finer kinds of vulnerabilities. Manual checking, on the other hand is extremely time consuming and involves extensive resource allocation. Again, this greatly relies on individual test performers who have wide-ranging expertise; hence variations will also appear depending upon who does it. Manual audits do not scale for the amount of apps that need to be evaluated with the exponential growth of mobile applications.

Users' opinions, primarily as reviews on the Google Play Store, are equally important in understanding the user experience and performance of mobile appli- cations. This results in crashes, violation of privacy, or errors in the functionality of an application, but usually retrieves the most important information about a user's experience. The large-scale analysis of reviews is extremely difficult. Since the number of reviews is too high, manual processing and actionable extraction is practically impossible. Techniques applied in sentiment analysis are not very effective in capturing subtlety of opinions held by the users and thus result in possible misinterpretation of the feedback. This results in a lack of fully un- derstanding the impact of usability issues or security concerns, which may not clearly emerge

through technical analysis alone.

However, despite the strong advantages of these existing solutions, there are several gaps and shortcomings that remain. The greatest concern is that technical evaluations have a very narrow scope and do not consider most of the emerging security threats or vulnerabilities that come from app design and third-party dependencies. Most evaluation tools do not consider user experience, which is critical to the complete success of an application. Traditional security evalua- tions usually leave out usability factors such as interface designing, performance optimization, and ease of navigation since these do not focus on capturing the potential flaws within the app. Moreover, manual testing cannot be scaled be- cause serious security audits for large numbers of apps by human expertise would be highly time-consuming and expensive.

The current systems are ineffective in terms of integration because technical evaluations, user feedback analysis, and security features are tackled separately. Such fragmentation keeps the development of a single framework, which may be comprehensive enough for understanding security as well as usability. However, evaluation methods give solutions that are limited to prior vulnerabilities and known attack vectors and fail to deal with emergent threats and advanced tech- niques of attacks. Although the issue of scalability has yet to be addressed in any satisfactory way, the complexity presented by mobile applications requires so much more efficient and more automated methods of managing so many ap- plications being produced and constantly updated on a daily basis.

In this regard, surely, such limitations call for a more holistic and integra- tive approach to mobile application evaluation. Current techniques should be complemented by automatic techniques that scale efficiently with in-time data and from both the technical and user-centric point of view. Integration of ad- vanced machine learning models, such as LSTM networks and LLMs, is one promising solution. These techniques can automate the analysis of both security vulnerabilities and usability issues, providing actionable insights to developers and stakeholders alike and enabling more timely and cost-effective assessments.

III. CHALLENGES AND EVOLUTION

This is particularly relevant in the creation of innovative approaches concerning data systems and artificial intelligence, where several technical problems have come up that require constant progression in techniques and technologies. One significant challenge is computational complexity: working with large data sets and executing complex algorithms can be computationally intensive, thus call- ing for resource-constrained processing strategies. Moreover, concerns with data quality and preprocessing continue, as noise, or errors in data lead to similar problems with algorithms, underlining the importance of careful data manipula- tion. Real-time processing also carries significant challenges especially when the loop of processing is in real-time or involves online data, due to hardware restric- tions or in some instances the need to come up with efficient algorithms. The use of the new methodologies in conjunction with the existing or legacy systems also intensifies the problem because compatibility and integration commonly require specific approaches.

Such challenges have not left the field stagnant but rather developed con- siderably over time. The advancements in algorithms have contributed to the paradigm change that favors new methods more capable of executing larger and complex problems. Growing computational power due to hardware enhance- ments such as GPUs and emergence of cloud platforms facilitate significant anal- yses that were once prohibitively complex. Another significant shift is the appli- cation of artificial intelligence and machine learning within the lines of forecasts and automation, which has significantly improved modeling efficiency and flexi- bility. This evolution has also been supported by interdisciplinary collaboration, for instance between computer vision, signal processing, and machine learning to formulate more comprehensive and efficient solutions. These are the trends that are emerging as the next generation supports, they are laying the foundations for further evolutions, which are bound to present solutions to the current issues and will expand new horizons for the applications of big data.

IV. FUTURE DIRECTIONS

It is believed that future advances in the field of machine learning and other data systems will continue to raise the bar in terms of capabilities for prediction, sys- tems size, automation, and security as they apply to real-world circumstances. Deep learning and reinforcement learning are expected to be instrumental in improving the practicability and precision of forecasts and to make systems ca- pable to counteract changing conditions. Such models could spearhead perfectly acceptable microrationality responses in deterministic domains, producing prac- tically helpful neocognitives for science-based, more flexible decision making.

In parallel with these developments, developing methods for building large- scale, real-time systems that can evolve with data arrays and promptly update their architecture in response to changes in input sources will remain essential to apply such approaches in high-tempo operational environments. Another im- portant area will be data processing stream where efforts will be made to take out human effort and replace that kind of work with automated methods which will make the data flow faster and errors to a minimum. This is so because as data continues to become more central to most industries, protecting the data and ensuring privacy issues have to be solved to voice trust and to meet the set laws. However, if the IoT and edge computing could be used, data may be collected and processed asynchronously to be on time for response and may find practical applications in areas with limited resources depending on connectiv- ity and response time. Together, these signify more of a revolutionary mode for incorporating data and machine learning into the various future applications.

V. CONCLUSIONS

This research develops a more enhanced approach to evaluate the Android apps application's security and utility with the help of deep learning algorithms and user feedback information analysis. The system comprises keywords, which are obtained from comprehensive Google Play Store reviews for the analysis of pat- terns in users' language and an overall assessment of the app's quality and its features. Using deep learning, the framework identifies risks to the security of a system – including malware and privacy issues – and usability problems. The scalable and automated data pipelines make it possible for continuous, efficient evaluation and may be extended to host services such as Google Play Protect as an added layer of app security and thereby increase users' confidence when using Android applications. The first results show the effectiveness of the suggested framework in enhancing the security of the Android environment, including sand boxing and threats identification. In conclusion, this research brings focus and information in addressing the goals of enhancing the overall security and user interfaces in Android applications.

REFERENCES

[1] F. A. Almarshad, N. Almutairi, and S. Alajlan," Detection of An- droid Malware Using Machine Learning and Siamese Shot Learning Technique for Security," *IEEE Access*, vol. 11, pp. 15045-15054, 2023. https://doi.org/10.1109/ACCESS.2023.3245659

[2] M. Deypir and T. Zoughi," Risk Score Computation for Android Mobile Applications Using the Twin k-NN Approach," *Journal of Web Engineering*, vol. 23, no. 1, pp. 117-138, 2024. https://doi.org/10.13052/jwe2345-1242.2319

[3] S. Dong, L. Shu, and S. Nie," Android Malware Detection Method Based on CNN and DNN Hybrid Mechanism," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 3, pp. 672-684, 2024. https://doi.org/10.1109/TII.2024.3214567

[4] Y. Ban, J. Choi, and H. Kim," FAM: Featuring Android Malware for Deep Learning-Based Familial Analysis," *IEEE Access*, vol. 10, pp. 37895-37905, 2022. https://doi.org/10.1109/ACCESS.2022.3154321

[5] Z. Namrud, A. Abdulaziz, and S. Qaisar," Deep-Layer Clustering to Identify Per-

mission Usage Patterns of Android App Categories," *IEEE Access*, vol. 10, pp. 20491-20500, 2022. https://doi.org/10.1109/ACCESS.2022.3165775

[6] K. A. Dhanya, P. Shaji, and G. Resmi," Obfuscated Malware Detection in IoT Android Applications Using Markov Images and CNN," *IEEE Systems Journal*, vol. 17, no. 1, pp. 331-342, 2023. https://doi.org/10.1109/JSYST.2023.3267945

[7] B. Urooj, M. Ahmad, and A. Mansoor," Malware Detection: A Frame- work for Reverse Engineered Android Applications Through Machine Learning Algorithms," *IEEE Access*, vol. 10, pp. 40915-40925, 2022. https://doi.org/10.1109/ACCESS.2022.3225692

[8] M. İbrahim, K. Shaalan, and I. M. El-Henawy," A Method for Automatic Android

[9] Malware Detection Based on Static Analysis and Deep Learning," *IEEE Access*, vol. 10, pp. 29547-29555, 2022. https://doi.org/10.1109/ACCESS.2022.3192457

[10] H. Rodriguez-Bazan, G. Sidorov, and P. J. Escamilla-Ambrosio," An - droid Ransomware Analysis Using Convolutional Neural Network and Fuzzy Hashing Features," *IEEE Access*, vol. 11, pp. 21678-21688, 2023. https://doi.org/10.1109/ACCESS.2023.3305678

[11] Y. Wang, W. Gao, X. Hei, and Y. Du," Method and Practice of Trusted Embedded Computing and Data Transmission Protection Architecture Based on Android," *Chinese Journal of Electronics*, vol. 33, no. 2, pp. 157-167, 2024. https://doi.org/10.1049/cje.2024.0049

Table 1. Comparison outlining each study's authors, techniques, advantages, and dis- advantages in Android malware detection.

Authors	Methodology Used	Strengths	Limitations
Fahdah A. Almarshad et al. (2023)	Siamese One-shot learning with Drebin dataset for malware classification	Achieved 98.9% accuracy; robust with limited training data	High dependency on feature extraction quality
Mahmood Deypir, Toktam Zoughi (2024)	Twin k-NN using Hamming distance for risk score computation	Realistic security risk estimation; works on various datasets	Complex to implement in real-time; high computational cost
Shi Dong, Longhui Shu, Shan Nie (2024)	Hybrid CNN and DNN integrating permissions and API call graphs with ACO for dimensionality reduction	96.8% accuracy; improved detection performance	Overfitting risk; high false alarm rates with novel malware
Younghoon Ban et al. (2022)	Deep learning familial analysis using CNNs	98% accuracy in familial classification; tested on real-world data	Limited in detecting complex malware behaviors; dataset representativeness
Zakeya Namrud et al. (2022)	SOM and K-means clustering on Android permissions	Consistent permission patterns; improved classifier performance with SVM	Focus on permission patterns may miss sophisticated malware behaviors

Dhanya K. A. et al. (2023)	CNN trained on Markov images generated from app bytecodes to detect obfus- cated malware	High detection accuracy (99.81%) for obfuscated malware; cost-effective compared to feature engineering	Dependent on the quality of image generation from bytecodes
Beenish Urooj et al. (2022)	Reverse engineered Android apps; ensemble learning (AdaBoost, SVM) on static features like permissions, intents, API calls	High accuracy of 96.24%; large dataset used for training	High false positive rate (0.3%)
Mülhem İbrahim et al. (2022)	Static analysis with API-based deep learning model on features like file size, permissions, services	High malware detection performance (99.5% F1-score); uses recent dataset	Limited to static analysis; can't detect sophisticated malware behavior
Horacio Rodriguez-Bazan et al. (2023)	CNN for malware classification using fuzzy hashing and NLP preprocessing of APK code to grayscale images	High accuracy in ransomware detection; innovative image transformation method	Limited dataset for evaluation; dependence on quality of image transformation
Yichuan Wang et al. (2024)	Trusted embedded static measurement and ECDH key exchange for data protection in Android	Strong encryption and secure key handling using TrustZone	Moderate performance overhead; risk of exposure to side-channel attacks