Section IV: Discussion

By training neural networks on data from individual rounds of the cryptographic hash SHA-1, each network theoretically needed to model less complexity. Then, linking together these neural networks would have allowed reconstruction of the internal state history and consequently resulted in a preimage attack on the hash. However, training various neural network architectures on each round led consistently to accuracies of 0, which is clearly unhelpful. Even through strategic modifications such as keyspace restriction and fuzzing, accuracy remained at 0 and loss was hardly decreased.

This result is due to a fundamental theoretical issue with the proposed framework. Since the hash is internally built from a string of compression functions, there is not enough information to deduce the previous round from the next one. In other words, all candidate previous rounds are equally possible, so it is impossible to choose the “correct” one.

In addition, this project had several limitations. One major limitation to this project was computational power in training the neural networks. Google Colab had a file size limit which prevented the upload of larger datasets, severely limiting the training efficacy of models across all architectures. Training locally was not possible either due to lack of access to a powerful GPU. The attack could potentially be effective if a larger dataset was used. Another limitation was, of course, the fundamental issue described earlier.

These results expand upon previous work by Alani (2012), Greydanus (2017), Goncharov (2019), So (2020), and Liu et al. (2021), all of which trained neural networks directly on input-output pairs of their attack models. This project explored a different approach, round-based neural network chaining. Although this approach was not effective, it serves in expanding the general knowledge of the field. As research in neuro-cryptanalysis has mainly centered around neural distinguishers in recent years, this

research was helpful in at least clarifying what does not work in direct neural preimage attacks (Gohr, 2019; Benamira et al., 2021).

Given the lack of strong success from previous neural preimage attacks, as well as the inefficacy of the neural network chain approach, as demonstrated in this experiment, focusing on differential-neural or classical cryptanalytic approaches may be more productive for future work.

Future Research

Although the described approach was ineffective within the limitations of this paper, future research could potentially involve the combination of neural network chains with meet-in-the-middle attacks, due to having similar strategic paradigms. The technique may also be used in conjunction with SAT solvers. Alternative architectures or, as mentioned before, greater computational power, may also enable the neural network chain model to be effective. Any successes in this future work would allow preimage attack research to take advantage of the rapid advances being made in machine learning.

Section V: Conclusion

In this rapidly-advancing technological age, digital information security is becoming ever more important. Simultaneously, rapid advances are being made in the field of machine learning. To leverage this progress in furthering the field of cryptanalysis and improving information security, this project proposed a preimage attack on a modern hash, SHA-1, using chained neural networks. Various neural models were trained to reverse singular rounds of SHA-1, including on keyspace-restricted data. However, this was completely ineffective for all architectures tested, including standard feed-forward, recurrent, keyspace-restricted, and fuzzy models—accuracy was 0 in all cases. These models would have been chained to partially reconstruct internal state history of reduced SHA-1, from which the preimage can be deduced, but this was unable to be done due to the poor results on single layers. Ultimately, the

attack was not effective, serving to expand the general knowledge in an understudied area of cryptanalysis. In the future, it is therefore more valuable to focus on other cryptanalytic techniques for preimage attacks and neural attacks.