PHÁT HIỆN TẤN CÔNG DDOS ĐỐI KHÁNG BẰNG MÔ HÌNH GAN VỚI BỘ PHÂN BIỆT KÉP (GAN-DD)

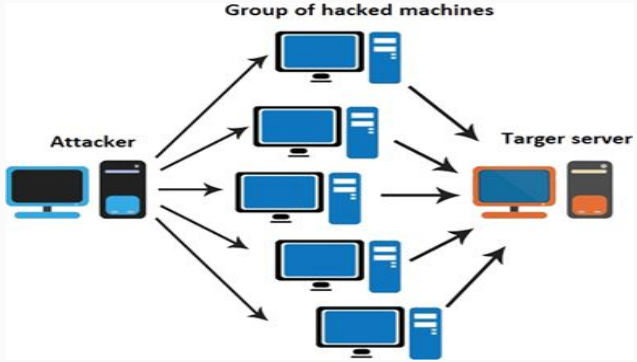
Kiều Hồng Khang - 240201042

Tóm tắt

- Lớp: CS2205.FEB2025
- Link Github của nhóm: https://github.com/khangkh19/CS2205.FEB2025
- Link YouTube video: https://www.youtube.com/watch?v=mziTuEJTaX0

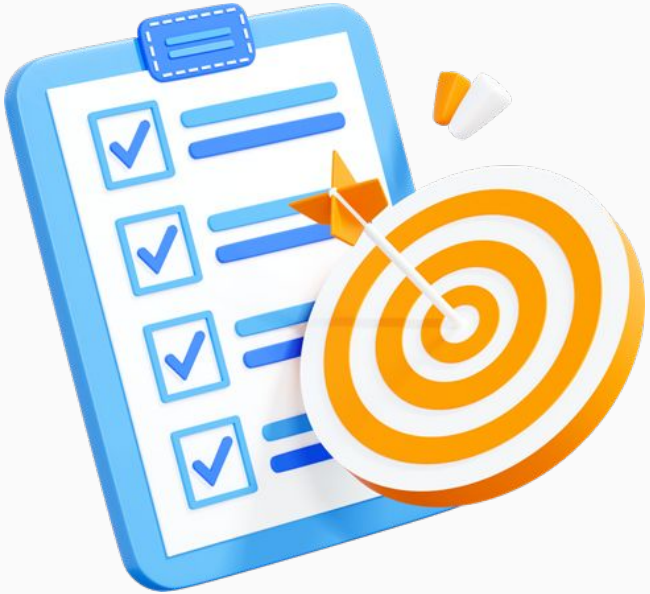
Kiều Hồng Khang

Giới thiệu



- Các cuộc tấn công DDoS ngày càng tinh vi hơn khi kết hợp kỹ thuật đối kháng (adversarial) để qua mặt hệ thống phòng thủ.
- Các mô hình học sâu truyền thống như CNN, LSTM... thường bị giảm hiệu quả khi gặp dữ liệu bị nhiễu này.
- Mô hình GAN với 2 bộ phân biệt (GAN-DD) có thể phân biệt rõ lưu lượng thật, giả, bị nhiễu đối kháng, tăng độ chính xác và tính ổn định khi phát hiện DDoS.
- Nâng cao khả năng phòng thủ mạng trong môi trường tấn công hiện đại.

Mục tiêu

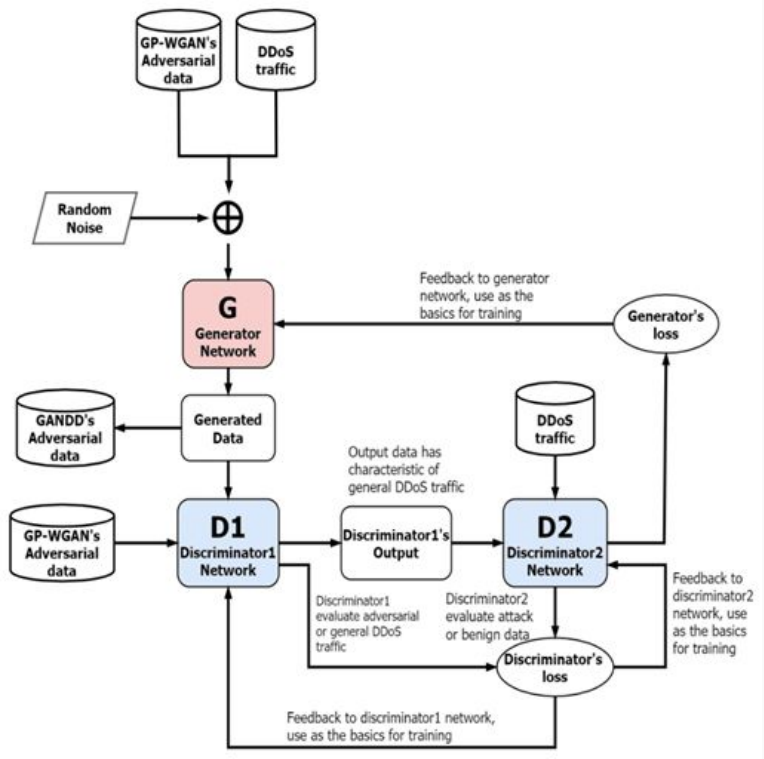


- Phát triển mô hình GAN-DD phát hiện DDoS đối kháng với độ chính xác cao.
- Đánh giá hiệu quả mô hình trên dữ liệu thực và dữ liệu bị nhiễu.
- Đề xuất hướng triển khai GAN-DD trong hệ thống mạng thực tế.

Nội dung và Phương pháp

Nội dung

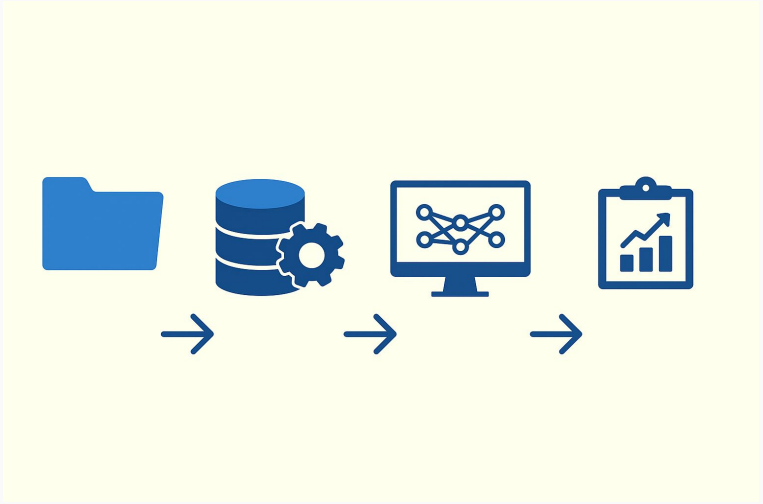
- Khảo sát các phương pháp phát hiện tấn công DDoS truyền thống.
- Tìm hiểu mô hình GAN-DD.
- Ứng dụng GAN-DD vào phát hiện tấn công DDoS.
- Đánh giá hiệu quả của GAN-DD.
- Phân tích thách thức và giải pháp.



Nội dung và Phương pháp

Phương pháp

- Tiền xử lý dữ liệu
- Xây dựng mô hình GAN-DD
- Huấn luyện mô hình
- Đánh giá hiệu quả
- Phân tích kết quả



Kết quả dự kiến

Việc ứng dụng GAN-DD (Generative Adversarial Networks with Dual Discriminators) vào phát hiện tấn công DDoS mang lại nhiều kết quả mong đợi bao gồm:

- Khả năng phát hiện hiệu quả cả tấn công thường và tấn công đối kháng với độ chính xác cao.
- Độ chính xác ≥ 90%, F1-score ≥ 85% trên dữ liệu thực nghiệm có nhiễu.
- So sánh tốt hơn so với các mô hình truyền thống như SVM, DNN và các biến thể GAN cơ bản.
- Đề xuất được hướng triển khai thực tế, đảm bảo mô hình hoạt động hiệu quả trong môi trường mạng hiện đại.

Tài liệu tham khảo

[1]. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. In Proceedings of the 27th International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; Volume 2, 2672–2680.

[2]. Gulrajani, I.; Ahmed, F.; Arjovsky, M.; Dumoulin, V.; Courville, A.C. Improved training of Wasserstein GANs. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; 5769–5779

[3]. Nguyen, T.D.; Le, T.; Vu, H.; Phung, D. Dual Discriminator Generative Adversarial Nets. In Proceedings of the Advances in Neural Information Processing Systems 30, Long Beach, CA, USA, 4–9 December 2017; pp. 2667–2677.

[4]. Zhang, X.; Zhao, Y.; Zhang, H. Dual-discriminator GAN: A GAN way of profile face recognition. In Proceedings of the 2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA), Dalian, China, 27–29 June 2020; 162–166

[5] Chin-Shiuh Shieh, Thanh-Tuan Nguyen, Wan-Wei Lin, Yong-Lin Huang, Mong-Fong Horng, Tsair-Fwu Lee, Denis Miu: Detection of Adversarial DDoS Attacks Using Generative Adversarial Networks with Dual Discriminators. Symmetry 14(1): 66 (2022). DOI: 10.3390/sym14010066