HARNESSING ARTIFICIAL INTELLIGENCE FOR PENETRATION TESTING

Structured Approaches, Automation, and AI Advancements

Kiera Conway

INTRODUCTION TO AUTOMATED PENETRATION TESTING

Penetration Testing (PT)

- Proactive cybersecurity approach
- Simulates cyber attacks to identify vulnerabilities

Significance in Cybersecurity

- Essential for protecting digital assets
- Helps organizations anticipate and mitigate potential attacks

Artificial Intelligence (AI) in PT

- Automates complex and time-consuming processes
- Enhances efficiency and effectiveness

CURRENT CHALLENGES

Lack of Research

- Limited foundational research in AI-driven PT
- Unexplored areas ripe for investigation

Reliance on Human Intervention

- Current solutions require significant human oversight
- Limits the full potential of automation in PT

Complex Technical Hurdles

- Ensuring data accuracy
- Maintaining model robustness
- Adapting to environmental changes

PENETRATION TESTING: PRACTICAL INTRODUCTION & TUTORIALS

Stephen Watts

Structured PT Approach

- Information Gathering
- Scanning
- Exploiting
- Maintaining Access
- Covering Tracks
- Reporting

Explores PT methodology

- Significance
  - Uncover Weak Points,
  - Understand Potential Attack Scenarios
  - Analyze The Vulnerability Severity
- Continuous Training
- Practical Demonstrations

Importance

- Bridge Theory and Practice
- Continued Training
- Automation as a Solution

REINFORCEMENT LEARNING FOR INTELLIGENT PENETRATION TESTING

Mohamed C. Ghanem, Thomas M. Chen

Phase 1: Gathering Information

Objective

- Create Intelligent agent that Mimics Experts
- Intelligent Automated Penetration Testing System (IAPTS)

Advanced Techniques

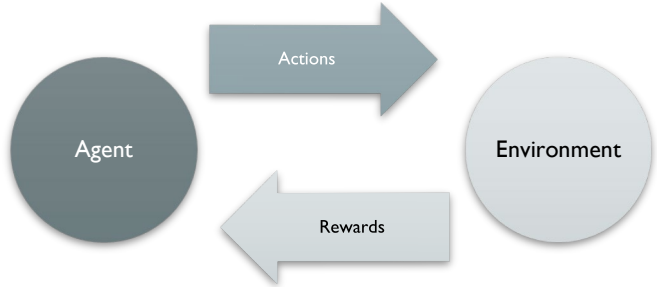
- Reinforcement Learning (RL)
- Partially Observable Markov Decision Processes (POMDPs)

Training

- Reward System
- Save High Reward Strategies

Outcome

- Learned From Past Experiences
- Efficient Reuse Saves Time
- Accuracy In Secure Networks



PENETRATION TESTING PROCEDURE USING MACHINE LEARNING

Reevan Seelen Jagamogan, Saiful Adli Ismail, Noor Hafizah Hassan, Hafiza Abas

Phase 2: Scanning

Hypothesis

- ML Tools Outperform Non-ML Tools

GyoiThon

- Default vs ML Mode
- Compare Data from Port 80 and Port 443
  - Port 80 – HTTP
  - Port 443 - HTTPS

Outcome

- ML Found More Vulnerabilities
- Potential for ML in Scanning Phase

Challenges

- Reliance on National Vulnerability Database (NVD)
- Requires further Refinement

Resources

- https://cve.mitre.org/
- https://nvd.nist.gov/vuln/search
- https://nvd.nist.gov/vuln/full-listing

VULNERABILITY EXPLOITATION USING REINFORCEMENT LEARNING

Anas AlMajali, et al

Phase 3: Exploiting

Objective

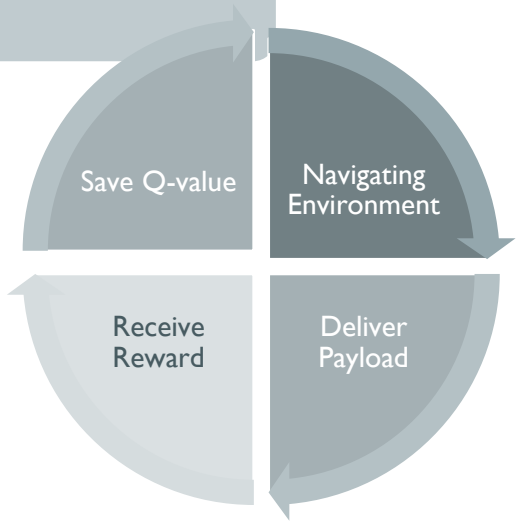
- Create Adaptable AI to Obtain Reverse Shell

Reinforcement Learning (RL)

- Guess-and-Reward System
- Q-Learning Algorithm (Q-Values)

Outcome

- Performs Better than Traditional Methods
- Effective and Efficient
  - average success rate: 83.64%
  - average exploit time: 8.26s
- Learned from Past Experiences



AUTONOMOUS SECURITY ANALYSIS AND PENETRATION TESTING

Ankur Chowdhary, DIJIANG Huang, Jayasurya Sevalur Mahendran, Daniel Romo, Yuli Deng, Abdulhakim Sabur

Autonomous Security Analysis and Penetration Testing (ASAP)

Advanced Techniques

- Reinforcement Learning (RL)
- Deep-Q Networks (DQN)

Attack Plans

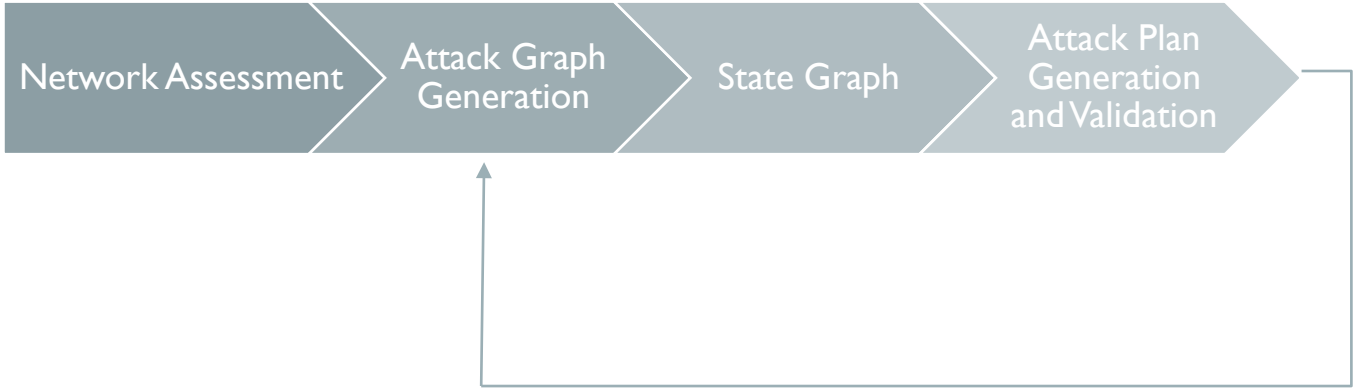
- Highly Detailed Series of Steps
- Provide Domain-Specific Rewards

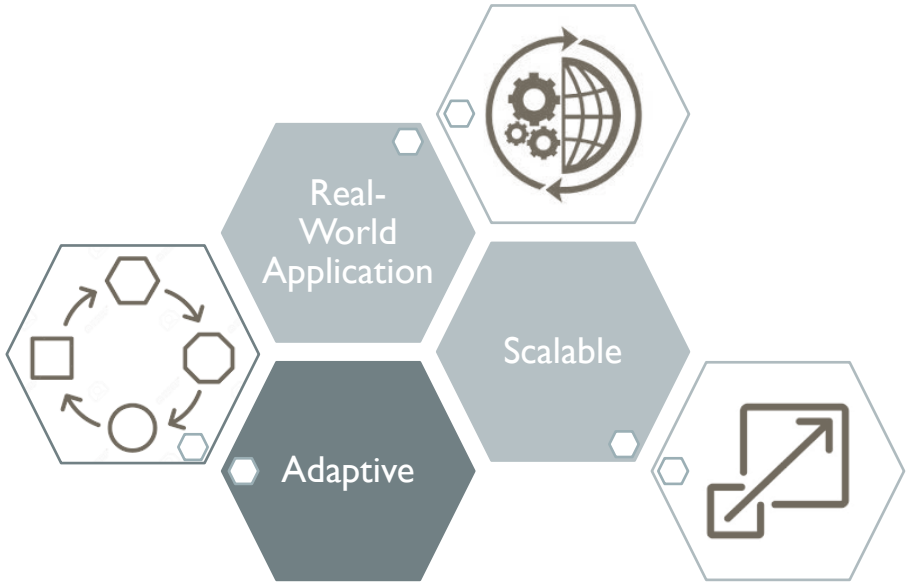
Limitations of Traditional Automation

- Lack Flexability
- Network Size Costraints
- Real-world Complexity Challenge

AUTONOMOUS SECURITY ANALYSIS AND PENETRATION TESTING

Ankur Chowdhary, Dijiang Huang, Jayasurya Sevalur Mahendran, Daniel Romo, Yuli Deng, Abdulhakim Sabur





AUTOMATED DECISION-MAKING SYSTEMS

Automated Decision-Making Systems (ADMS)

- Decision Support Systems (DSS)
- Attack Graph Modeling
- Vulnerability Databases (CVE, NVD)

Advanced Techniques

- Deep Reinforcement Learning (DRL)
- Transfer Learning
- Natural Language Processing (NLP)
- Text-Mining

Benefits

- Reduction in Human Reliance
- Addresses Data Accuracy Issues, Model Robustness, Environmental Adaptability

Examples

- Vulnerability Identification
- Automated Exploit Execution

Consideration

- Data Accuracy

SELF-IMPROVING SYSTEMS

Self-Improving Systems

- Feedback Loops
- Real-Time Data Processing
- Task Graphs
- Fault Tolerance

Benefits

- Adaptive to New Threats
- High Throughput, Low Latency
- Minimized Human Oversight
- Addresses Data Accuracy Issues, Model Robustness, Environmental Adaptability

Examples of Applications

- Real-Time Threat Detection
- Continuous Micro-Simulations

HUMAN-IN-THE-LOOP

Human-in-the-loop (HITL)

- Human Insight
- AI Capabilities

Benefits

- Minimize human involvement
- Addresses Data Accuracy Issues, Model Robustness, Environmental Adaptability

Examples

- Interactive Parameter Adjustment
- Real-Time Interaction Monitoring

CONCLUSION

Systems

- Automated Decision-Making Systems (ADMS)
- Self-Improving Systems
- Human-in-the-loop (HITL)

Impact on Challenges

- Reduction in Human Reliance
- Data Accuracy
- Model Robustness
- Environmental Adaptability

REFERENCES

1. Big Data: What it is and why it matters.
2. S. Watts , Penetration Testing: Practical Introduction & Tutorials, 2022.
3. M. C. Ghanem and T. M. Chen, "Reinforcement Learning for Intelligent Penetration Testing," in Second World Conference on Smart Trends in Systems, Security and Sustainability, London, 2018.
4. R. S. Jagamogan, S. A. Ismail, N. H. Hassan and H. Aba, "Penetration Testing Procedure using Machine Learning," in International Conference on Smart Sensors and Application (ICSSA), Kuala Lumpur, 2022.
5. A. AlMajali, L. Al-Abed, R. Mutleq, Z. Samamah, A. A. Shhadeh, B. J. Mohd and K. M. Ahmad Yousef, "Vulnerability Exploitation Using Reinforcement Learning," in Jordan International Joint Conference on Electrical Engineering and Information Technology, Amman, 2023.
6. A. Chowdhary, D. Huang, J. S. Mahendran, D. Romo, Y. Deng and A. Sabur, "Autonomous Security Analysis and Penetration Testing," in 16th International Conference on Mobility, Sensing and Networking (MSN), Tokyo, 2020.
7. R. Kaur, D. Gabrijelčič and T. Klobučar, "Artificial intelligence for cybersecurity: Literature review and future research directions," in Information Fusion, 2023.
8. T.-M. Georgescu, B. Iancu and M. Zurini, "Named-Entity-Recognition-Based Automated System for Diagnosing Cybersecurity Situations in IoT Networks," Sensors, vol. 19, no. 15, 2019.
9. R. Nishihara, P. Moritz, S. Wang, A. Tumanov, W. Paul, J. Schleier-Smith, R. Liaw, M. I. Jordan and I. Stoica, "Real-Time Machine Learning: The Missing Pieces," March 2017.
10. X. Wu, L. Xiao, Y. Sun, J. Zhang, T. Maa and L. He, "A survey of human-in-the-loop for machine learning," Future Generation Computer Systems, vol. 135, pp. 364-381, 2022.