Bowie State University: An Efficient Approach to Closing the Breach Detection Gap (BDG)

By: **Sydney Raymond, David Tan, and Khalil Davis**
Faculty Mentor: **Dr. David Anyiwo**
Graduate Mentor: **Jerry Godwin Diabor**

Presentation Overview

Introduction

Project Goals

Methods

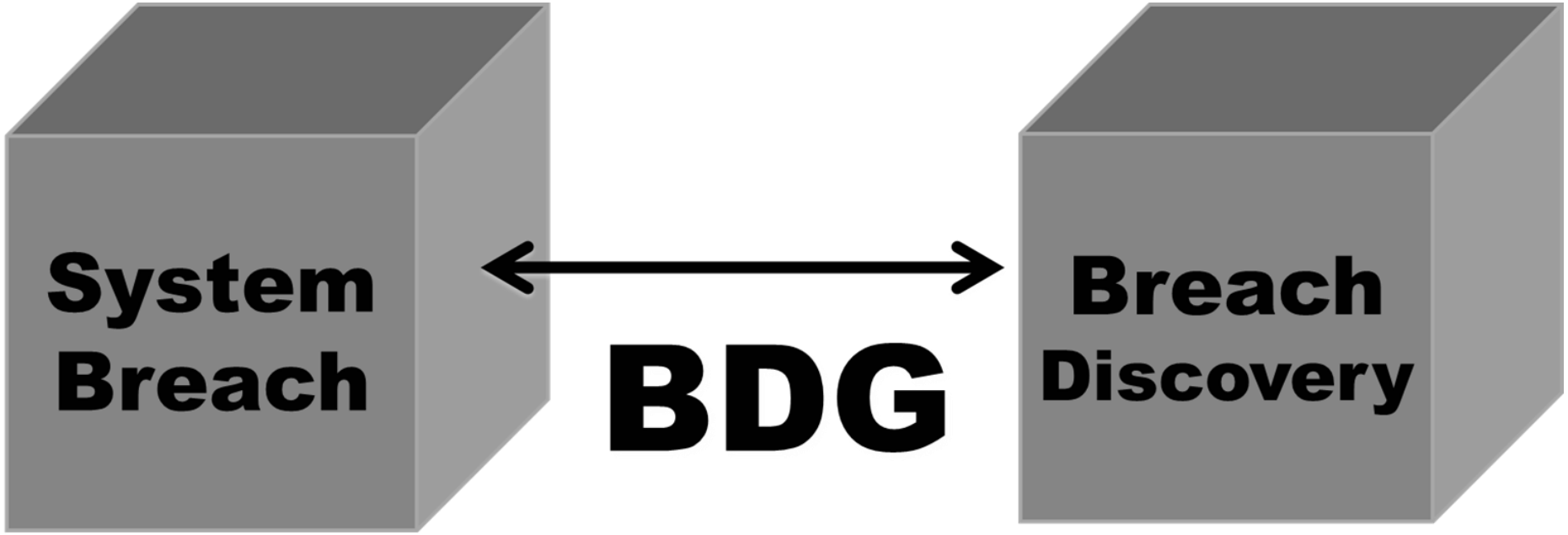
Results

Discussion

Conclusion/Future Works

Introduction: Research Idea

As the use of systems, applications, and software is on the rise globally, user privacy and data protection are of the utmost concern to individuals, firms, and businesses operating online. The time period between a data breach occurrence and its discovery, known as the BDG, can span months, leaving users' data vulnerable to exploitation. We intend to identify the technological solutions that target the most significant causes of the BDG.



Introduction: Literature Review

There has been significant prior research conducted establishing risk factors for data breaches. We examined the ways in which these risk factors could contribute to a wider BDG.

1. F. Kamoun & M. Nicho- “Human and Organizational Factors of Healthcare Data Breaches: The Swiss Cheese Model of Data Breach Causation And Prevention”
2. D. Dolezel & McLeod- “Managing Security Risk: Modeling the Root Causes of Data Breaches”
3. V. Jaganathan, P. Cherurveettil, & P. Muthu Sivashanmugam- “Using a Prediction Model to Manage Cyber Security Threats”
4. H. Tabrizchi & M. Kuchaki Rafsanjani- “A survey on security challenges in cloud computing: issues, threats, and solutions”
5. A. Steimers & M. Schneider- “Sources of risk of AI Systems”
6. D. Rios Insua, V.A. Couce, J.A. Rubio, W. Pieters, K. Labunets, and D. G. Rasines- “An Adversarial Risk Analysis Framework for Cybersecurity”

Introduction: Literature Review

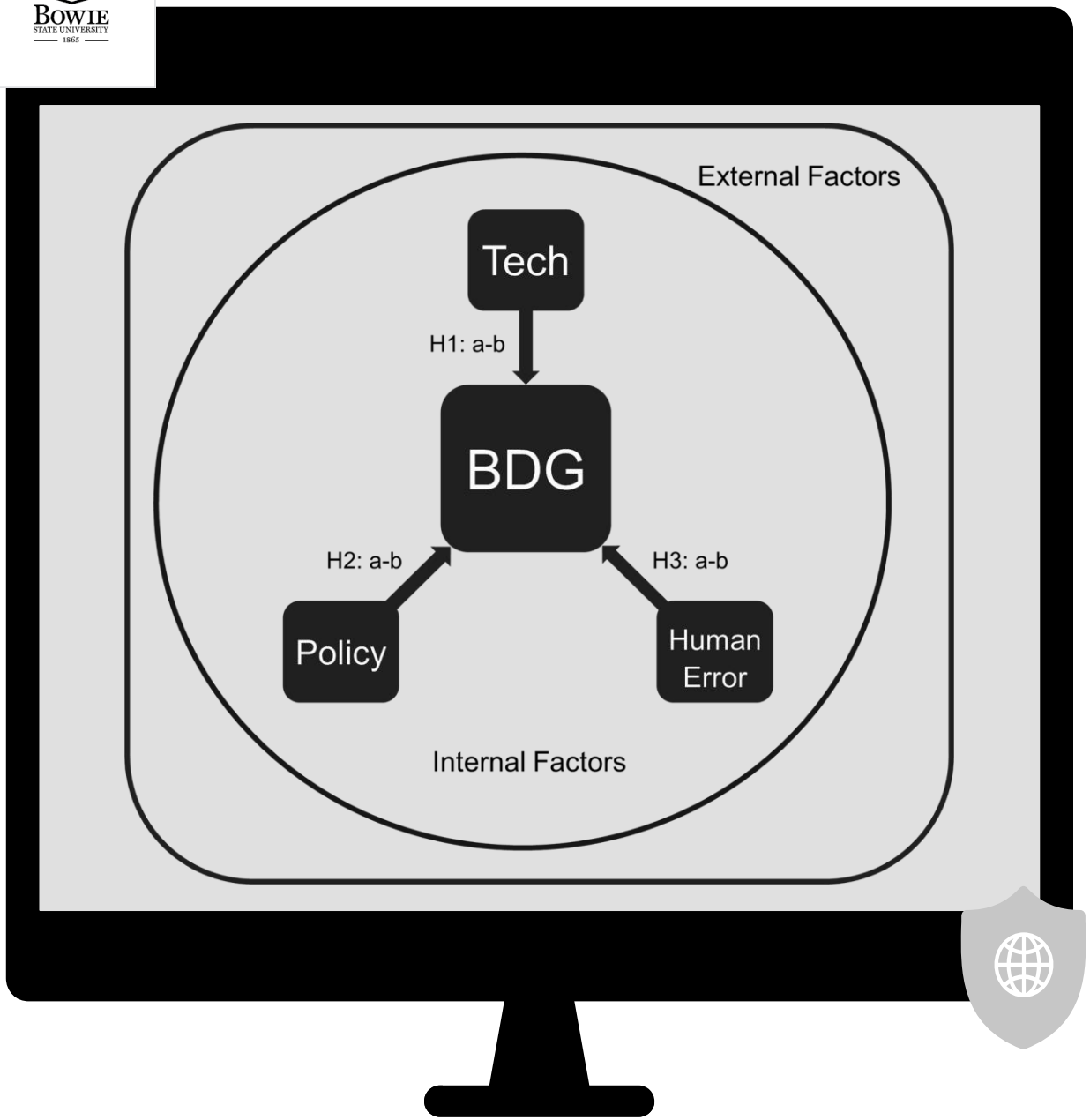
We additionally looked at case studies of major data breaches, including the Yahoo! and Equifax breaches.

1. Equifax- *Equifax Releases Details on Cybersecurity Incident, Announces Personnel Changes*. Atlanta: Equifax.
2. P. Wang & C. Johnson- “Cybersecurity Incident Handling: A Case Study of the Equifax Data Breach”
3. P. Wang & S.-A. Park- “Communication in Cybersecurity: A Public Communication Model for Business Data Breach Incident Handling”
4. Wang L, He R, Wang H, Xia P, Li Y, Wu L, Zhou Y, Luo X, Sui Y, Guo Y, Xu G. - “Beyond the virus: a first look at coronavirus-themed Android malware”
5. H. Zafar- “Security Risk Management at a Fortune 500 Firm: A Case Study”
6. M. Jain- “The Aadhaar Card: Cybersecurity Issues with India's Biometric Experiment”

Introduction: Literature Review

We reviewed current breach detection technologies, policies, and user behavior to determine the ways in which they contribute to the BDG and reviewed several pre-existing technological options for gap reduction.

1. B.M.C Silvia, P.R.M Inácio, and J.A. Santos- “Towards the Use of Blockchain in Mobile Health Services and Applications”
2. Q. Chen & U. Aickelin- “Anomaly Detection Using the Dempster-Shafer Method”
3. J. Botha et al.- “Pro-Active Data Breach Detection: Examining Accuracy and Applicability on Personal Information Detected
4. Cymulate- *The 3 Approaches of Breach & Attack Simulation Technologies.* Dallas: Cymulate”
5. X. LI- “A Blockchain-Based Verifiable User Data Access Control Policy for Secured Cloud Data Storage”
6. O.O. Malomo el al.- “Next-generation cybersecurity through a blockchain-enabled federated cloud framework”



Introduction: Hypotheses

1. H1a: An association exists between technology and the breach detection gap.
   H1b: There is a negative correlation between correct technology application and breach detection time.
2. H2a: An association exists between corporate and government policies and the breach detection gap.
   H2b: There is a negative correlation between effective policies and breach detection time.
3. H3a: An association exists between human error and the breach detection gap.
   H3b: There is a positive correlation between human error and breach detection time.
4. H4a: The external environment influences breach detection time.
   H4b: The internal influence always acts to delay breach detection.

Project Goals: Overall Goals

1. Produce a research paper concerning the most efficient and effective approach to closing the BDG.
2. Postulate a model that can serve as a template for organizations interested in breach detection.

Methods: Overview

RESEARCH DESIGN

Obtained data from literature review case studies and from press releases, news articles, and academic articles published on specific data breach incidents that we selected for study.

ANALYZE DATA

Graphically analyzed relationship between selected factors and breach detection time in the selected case studies and determined the most significant factor.

IDENTIFY AND CREATE

Identified the most effective solution based on the most significant factor and developed a technological solution that addressed the issue of breach detection time.

Methods: Sampling

Population	Sample	Reasoning
• The population we studied was organizations that have experienced data breaches	• Nine companies that have experienced data breach incidents • Three organizations discussed in our literature review plus six selected companies	• Seven companies experienced data breaches within the past three years • Organizations were additionally selected to provide a variety of breach detection times

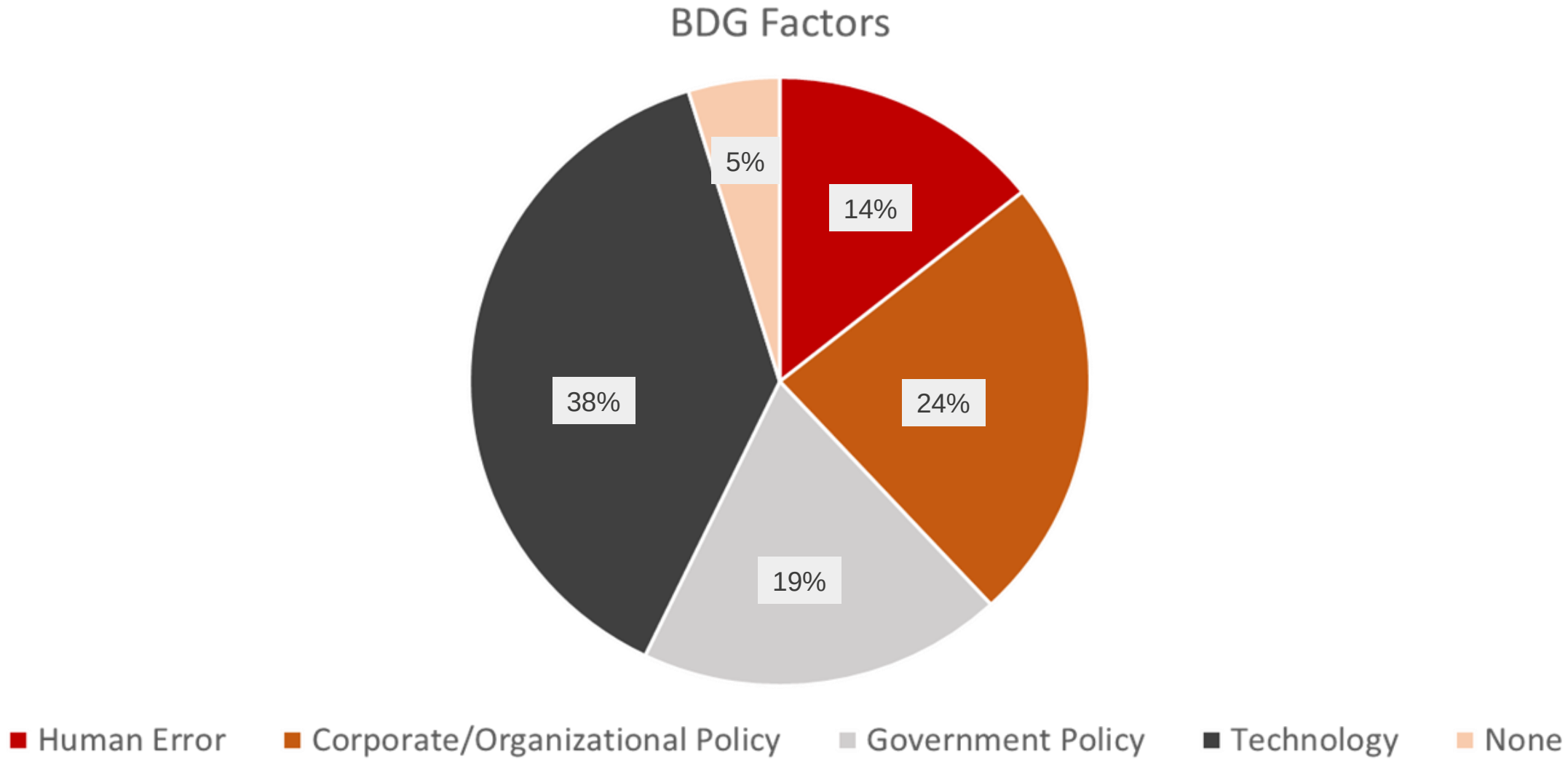
Methods: Analysis

Breach Incident	Present Factors	BDG
Logan Health	None	4 days
Service Employees International Union, Local 32BJ	Tech	11 days
Marriott International	Tech & Policy	~45 days
Equifax breach	Human error, Tech, & Policy	~75 days
Ethos	Tech	182 days
Facebook breach	Tech & Policy	~365 days
SolarWinds	Tech & Policy	~700 days
Syniverse	Human error, Tech & Policy	~1,825 days
Aadhaar breach	Human error, Tech, & Policy	~1825 days

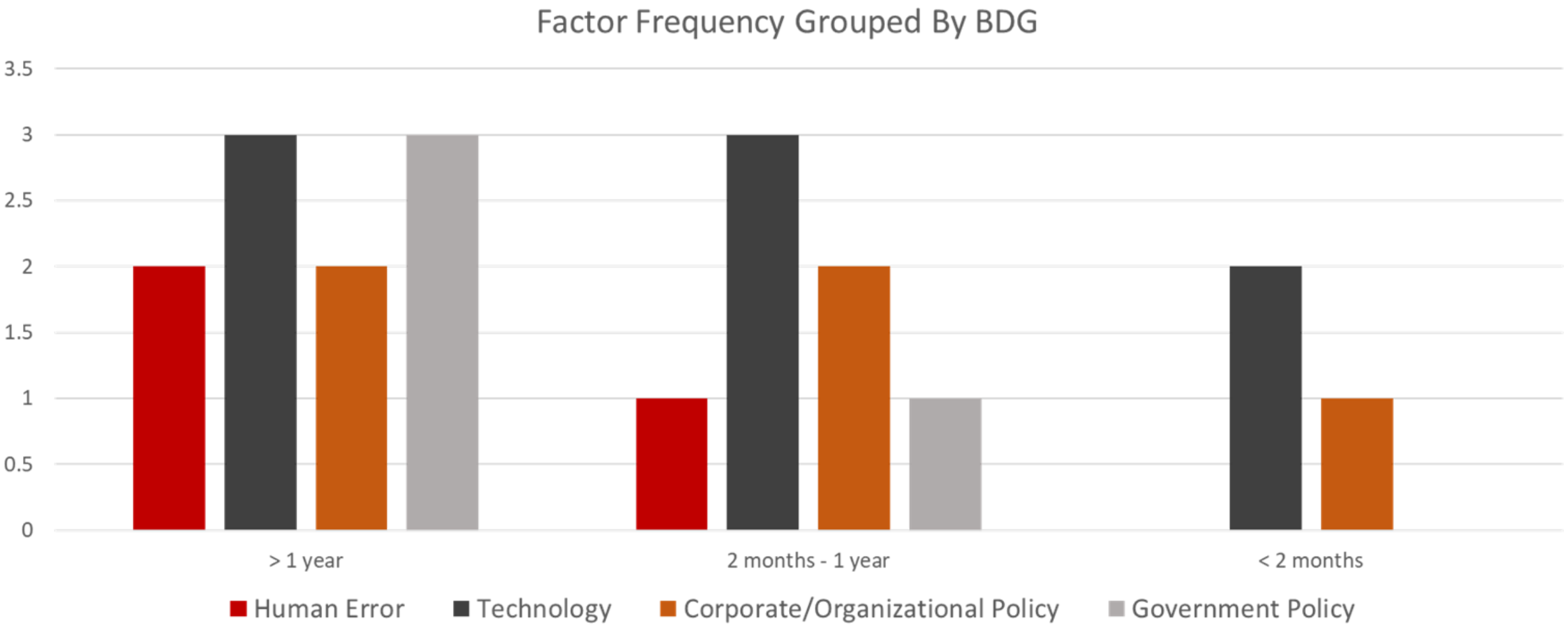
Results: Hypotheses

Hypothesis #1	H1a: Verified via content analysis
	H1b: Requires further analysis
Hypothesis #2	H2a: Verified via content analysis
	H2b: Requires further analysis
Hypothesis #3	H3a: Verified via content analysis
	H3b: Requires further analysis
Hypothesis #4	H4a: Requires further analysis
	H4b: Requires further analysis

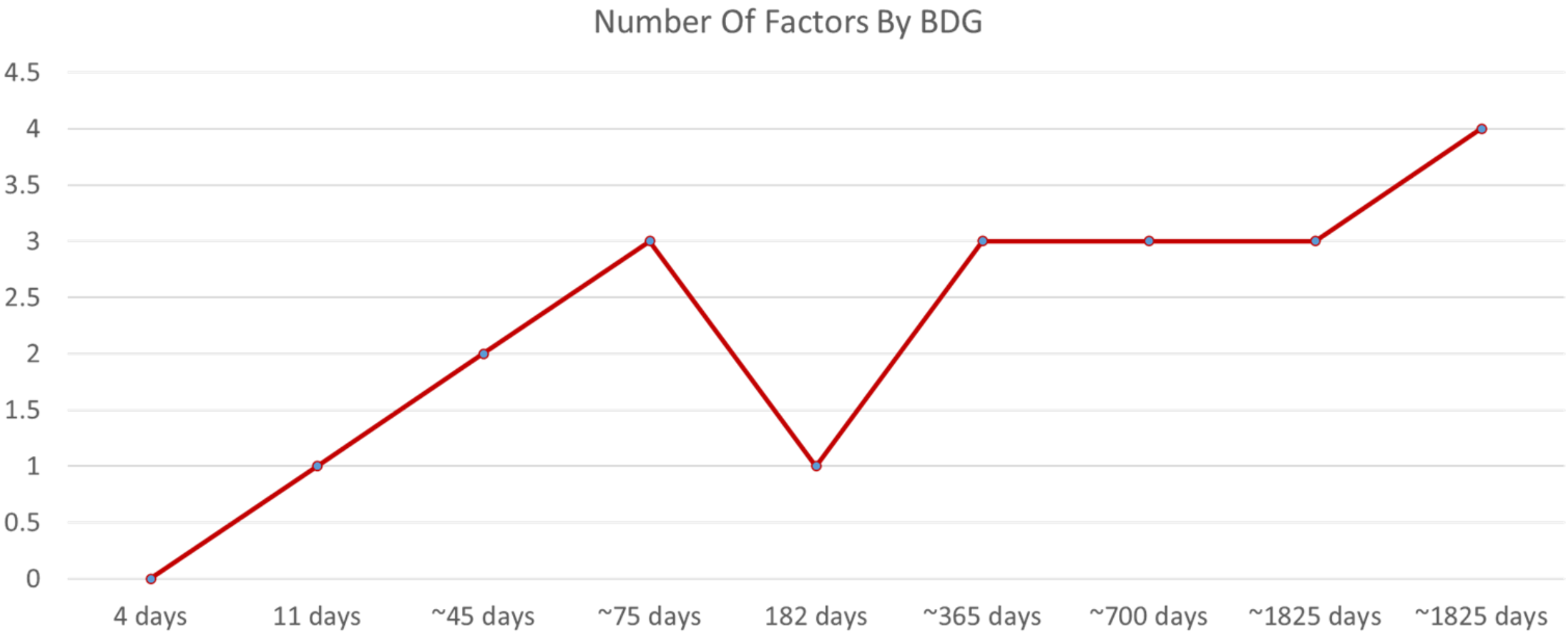
Results: Graphs



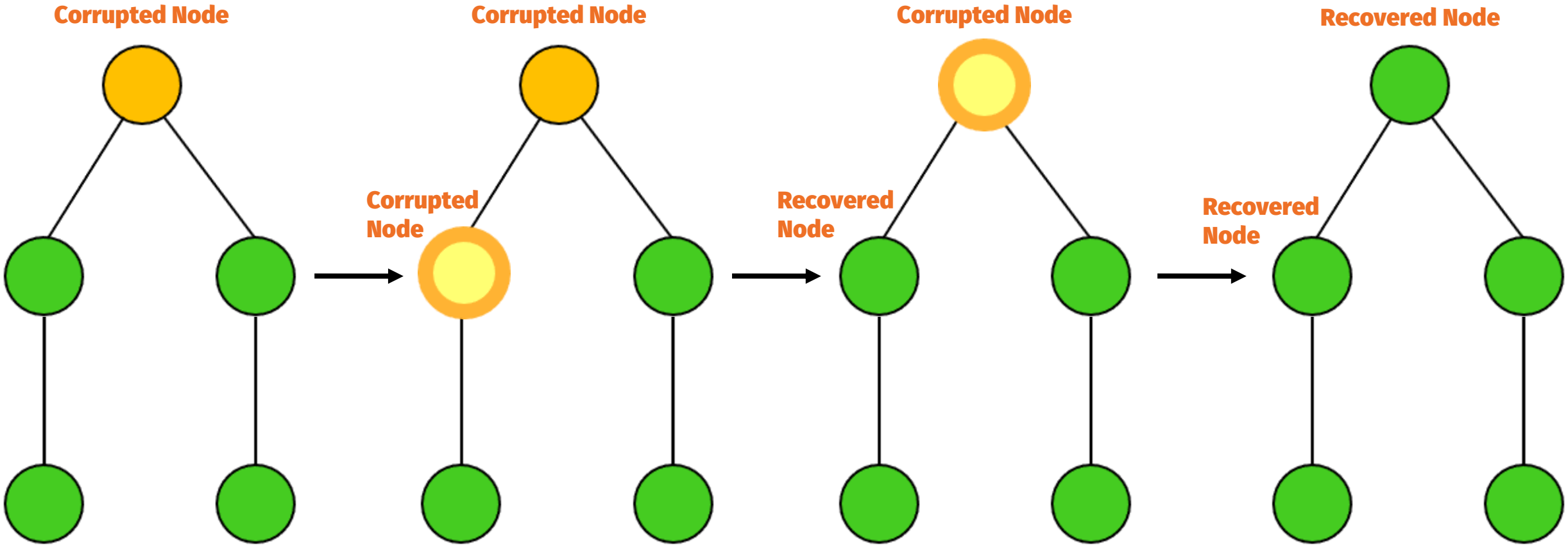
Results: Graphs



Results: Graphs



Results: Code Diagram



Results:
Code Snippet

```java
public class Graph {

    public static void main(String[] args) {
        ArrayList<Integer> x_values = new ArrayList<>();
        int size = 21;
        for (int i = 0; i < size; i++) {
            x_values.add(i);
        }
        Collections.shuffle(x_values);
        ArrayList<Node> nodes = new ArrayList<>();
        // generate 20 nodes
        for (int i = 0; i < size; i++) {
            nodes.add(new Node(x_values.get(i)));
        }
        // link them randomly
        linking(nodes);

        long start = System.currentTimeMillis();
        String firstTime = java.time.LocalTime.now().toString();

        Node finalNode = processing(nodes);

        System.out.println("Time start: " + firstTime);
        System.out.println("Time end: " + java.time.LocalTime.now());
        System.out.println("Total time taken for execution: "
                + (System.currentTimeMillis() - start) + " milli second");

        System.out.println("node " + finalNode.getNumber() + " was selected.");
        System.out.println(finalNode.connection());

    }

```

```
Node:  ->  { 14, 1 } -> 9 { true }
X-value: 248, 238 234
Node:  ->  { 2 } -> 10 { true }
X-value: 244 227
Node:  ->  { 5 } -> 11 { true }
X-value: 214 213
Node:  ->  { 3, 13, 4, 16, 6 } -> 12 { true }
X-value: 233, 241, 212, 236, 227 237
Node:  ->  { 11, 17, 1 } -> 13 { true }
X-value: 213, 221, 238 241
Node:  ->  { 8, 15, 2 } -> 14 { true }
X-value: 242, 240, 244 248
Node:  ->  { 7, 16, 9 } -> 15 { true }
X-value: 247, 236, 234 240
Node:  ->  { 9 } -> 16 { true }
X-value: 234 236
Node:  ->  { 12 } -> 17 { true }
X-value: 237 221
Node:  ->  { 4, 0 } -> 18 { true }
X-value: 212, 227 230
Node:  ->  { 5 } -> 19 { true }
X-value: 214 218
Node:  ->  { 4, 14 } -> 20 { true }
X-value: 212, 248 237

Time start: 02:27:43.568833
Time end: 02:27:46.475394
Total time taken for execution: 2916 milli second
node 8 was selected.
node 8 { 18, 14, 4, 17 } --> true
the new x value for node 8: 230, 248, 212, 221 252
```

Algorithm for BDG

Attack structure classification algorithm
Input: Detection and reporting attack structures
Output: Malware Attack category

1. if FGd3mmd=FGHmmd=FGlocalgen=MJd3mmd=ABSmmd=MJlocalgen=no then
2. malware ← Node-1
3. else
4. if delShdCpy=ovrFile=no then
malware ← Node-2
6. else
7. if FGd3mmd = FGHmmd = MJlocalgen=no then
8. malware ← Node-5
9. else
10. if FGd3sym = FGHmmdsym = MJlocalgensym = yes then
11. malware ← Node-3
12. else
13. malware ← Node-4
14. end if
15. end if
16. end if
17. end if=0

Results: Average Stabilization Time

Node Number	20	40	60	80	100	120	140	160	180	200
	26579	55572	72356	110259	132021	153226	210065	221569	243962	277231
	4207	41235	55465	63569	124436	110363	196362	206654	223651	256213
	3477	22578	52365	72569	105569	123554	153265	215565	241623	263316
	8791	32152	12258	82246	113659	151235	206543	213326	235663	271621
	9274	38426	36542	100698	98625	136854	192236	196354	243316	243165
	80	42315	39569	95562	105378	110369	206321	205698	233656	269563
	853	21589	71125	73256	121369	146559	194563	220364	237128	275363
	819	37856	46589	83465	115639	142396	186336	199856	231236	265436
	17971	51582	44569	103356	126963	152646	201136	200656	229633	256312
	29	26539	25349	756233	110639	134886	204563	211566	240361	273165
Total	72080	369844	456187	1541213	1154298	1362088	1951390	2091608	2360229	2651385
Average	7208	36984.4	45618.7	154121.3	115429.8	136208.8	195139	209160.8	236022.9	265138.5

Results: Code (Cont.)

```
public class PortScanner {
   public static void main(String []args) {
       long startTime = System.currentTimeMillis();
       ArrayList<Integer> openPorts = new ArrayList<>();
       for (int port = 1; port <= 65535; port++) {
         try {
            Socket socket = new Socket();
            socket.connect(new InetSocketAddress("localhost", port), 1000);
            openPorts.add(port);
            socket.close();
            System.out.println("Port " + port + " is open");
            }
         catch (Exception ex) {
        }
       }
       long endTime = System.currentTimeMillis();
       System.out.println(System.lineSeparator() + "Time to scan all possible locations: "
                          + (endTime - startTime) + " milliseconds");
       long startTime2 = System.currentTimeMillis();
       for(int i = 0; i < openPorts.size(); i++) {
           Socket s = new Socket();
           if(s.isClosed() == false) {
               try {
                 s.close();
                 openPorts.remove(i);
                 }
               catch (IOException e) {
                 e.printStackTrace();
               }
           }
           else {
               openPorts.remove(i);
           }
       }
       long endTime2 = System.currentTimeMillis();
       System.out.println("Time to close all previous open ports: " +
                          (endTime2 - startTime2) + " milliseconds");
   }
}
```

Discussion

CHALLENGES

- The identified firms have not yet responded to our questionnaires.
- We have not been able to collate the responses from these firms.

STRATEGIES

- Consult with our faculty mentor and our graduate mentor to decide best choice of action.

Conclusions

In our literature review, we identified three potential breach detection gap factors. Using content analysis of several case studies, we verified our hypotheses involving the association between each factor (excluding external factors) and the breach detection gap. We then developed a multi-tool framework that addresses these factors and therefore aims to reduce the breach detection gap. The program developed by our group reduces breach detection time and is highly accurate, eliminating errors resulting from false positives. We have also developed a survey that companies can use to determine the presence of specific BDG risk elements (BDG factor subgroups) in their organizational structure. The program addresses the link between technology and the BDG, and the survey allows companies to reformulate their policies and obviate some human error risks.

Recommendations

We highly recommend that organizations:

- Adopt the coding standards established in this project
- Train and coach employees/staff in proper breach detection practices
- Improve on their IT policies related to breach detection
- Properly secure their networks to avoid malicious attacks and ransomware

Future Work

Future research should focus on collecting data to determine the quantitative correlation between factors and breach detection time. Organizations could contribute their individual survey data to a joint effort to understand this relationship. Additionally, further studies must be taken to more clearly analyze the relationship between external factors (i.e. those that do not originate from within an organization) and breach detection time. Work should also be undertaken to measure the success of organizations that adopt our proposed strategies.

References

Santos, J. A., Inácio, P. R. M., & Silva, B. M. C. (2021). Towards the Use of Blockchain in Mobile Health Services and Applications. *Journal of Medical Systems*, *45*(2). https://doi.org/10.1007/s10916-020-01680-w

Chen, Q., & Aickelin, U. (2006). Anomaly Detection Using the Dempster-Shafer Method. *Proceedings of the 2006 International Conference on Data Mining*. Retrieved from https://arxiv.org/ftp/arxiv/papers/0803/0803.1568.pdf

Kamoun, F., & Nicho, M. (2014). Human and Organizational Factors of Healthcare Data Breaches: The Swiss Cheese Model of Data Breach Causation And Prevention. *International Journal of Healthcare Information Systems and Informatics*, 42-60.

Dolezel, D., & McLeod, A. (2019). Managing Security Risk: Modeling the Root Causes of Data Breaches. *The Health Care Manager, 38*(4), 322-330. Retrieved June 9, 2022

Jaganathan, V., Cherurveettil, P., & Muthu Sivashanmugam, P. (2015). Using a Prediction Model to Manage Cyber Security Threats. *Scientific World Journal*

Botha, J., Eloff, M., & Swart, I. (2016). Pro-Active Data Breach Detection: Examining Accuracy and Applicability on Personal Information Detected. *ICCWS 11th International Conference on Cyber Warfare and Security*, 47-55.

Equifax. (2017). *Equifax Releases Details on Cybersecurity Incident, Announces Personnel Changes.* Atlanta: Equifax.

Wang, P., & Johnson, C. (2018). Cybersecurity Incident Handling: A Case Study of the Equifax Data Breach. *Issues in Information Systems, 19*(3), 150-159.

Wang, P., & Park, S.-A. (2017). Communication in Cybersecurity: A Public Communication Model for Business Data Breach Incident Handling. *Issues in Information Systems*, 136-147.

Cymulate. (2022). *The 3 Approaches of Breach & Attack Simulation Technologies.* Dallas: Cymulate.

Questions?